High-Efficiency, Wideband RF Power Amplifiers for Cellular Infrastructure

A thesis submitted to Cardiff University

In candidature for the degree of

Doctor of Philosophy

by

Richard J. Wilson

Center for High Frequency Engineering

School of Engineering

Cardiff University

United Kingdom

April 2020

Abstract

Over the past decade, the exponential increase in demand for data content has led to many challenges for the cellular networks. More spectrally efficient modulation formats place tight linearity requirements on the system, and the high peak to average ratio of the signals requires the use of power amplifier architectures with high efficiency at backed-off power levels. Also, the push toward multi-band radios demands the use of wideband power amplifiers in the place of multiple single band amplifiers.

This work focuses on the following areas of research:-

RF Bandwidth

Much research focuses on achieving wideband solutions at low frequency (sub 1GHz) or low power (<20W) through the absorption of device parasitics into the matching structures. This work focuses on Doherty amplifier topologies with bandwidths (up to 40% fractional) and power levels (over 100 watts) appropriate for cellular infrastructure applications. The bandwidth of each element in the Doherty amplifier is analyzed across frequency when load modulated. Several novel wideband Doherty amplifier topologies are presented, and two demonstration amplifiers are designed, achieving the state of the art performance.

Linearity

In recent years, there has been considerable research focus to enhance the linearity of the RF power amplifier when linearized in a digital pre-distortion (DPD) system. Much of this research focuses on the output baseband impedance of the device and circuit. However, until very recently, little work focused on the impact of the device input. This research focuses on the effects of the input baseband impedance of the device and circuit, with a novel input matching topology

proposed to enable virtually ideal impedance characteristics. A novel integrated passive device is developed to enable the proposed topology, and the enhanced DPD correction, when compared to the current state of the art, is demonstrated using a 60-watt LDMOS device.

Performance Scaling with Power and Frequency

This research focuses on the minimization of performance impact due to device power scaling at high frequency. A proposed waveform engineering analysis and optimization method using 3D electromagnetic simulation with load-pull is proposed. In addition, a novel matching topology using multi-level, high Q integrated passive devices (IPD) is proposed. The analysis method is used to demonstrate a reduced performance degradation through enhanced voltage and current waveform uniformity across the power transistor. For the first time, the concept of active harmonic impedances due to the distributed effects of a high power RF device is presented.

Acknowledgments

I have had the fortune and pleasure of working with many outstanding engineers during my career and have appreciated the individual skills and insights each has provided. Although I would love to list many more, I would like to particularly thank the core team of smart guys I have worked alongside over the past decade. Haedong Jang, Bjorn Herrmann, Tim Canning, Zul Mokhti, Bayaner Arigong, Madhu Chidurala, Marvin Marbell, Saurabh Goel, and Simon Ward – thank you for the many open and honest discussions, for being a safe place to brainstorm potentially (and often) silly ideas, and your contributions to this work.

Specifically, I would like to thank Saurabh Goel for the late-night characterization of the constant impedance Doherty presented in 3.3.1.2 as we ran up to the IMS submission deadline. To my co-author, Dr. Haedong Jang, for the module layout and characterization of the zero phase peaking, constant impedance Doherty presented in section 3.3.4. I regret that we did not have the opportunity to progress this work to a transaction paper as we had intended. And to Dr. Bayaner Arigong, for the baseline device design shown in 4.16 (a), the benchmark to which the work in chapter four was compared.

Thank you to both of my parents. The sacrifices you made throughout my childhood led me to where I am today.

Also, a big thank you to my supervisors Dr. Jonny Lees and Professor Paul Tasker, for the discussions during this research and the significant time investment in the thesis review. Your inputs have been invaluable.

Finally, the biggest thank you must go to my family. For the many, many weekends and evenings they missed out on time with me for this work to be undertaken and for this thesis to be written. I dedicate this work to my wife, Lucy, and children Amethyst, Jack, Rose, Magenta, James, and Jude.

Table of Contents

Chapter One - Introduction

1.1 Research Motivation

Over the past decade, there has been a linear increase in mobile device subscribers worldwide. Once considered a luxury for the wealthy individual or business executive has grown to a one to one ratio between mobile subscription and world population. Between 2010 and 2020, mobile subscriptions increased from 5 billion to just over 8 billion, with a forecasted 1 billion new subscribers in the next five years [2]. Figure 1.0 shows the total worldwide subscriptions over time, broken down into technology type.

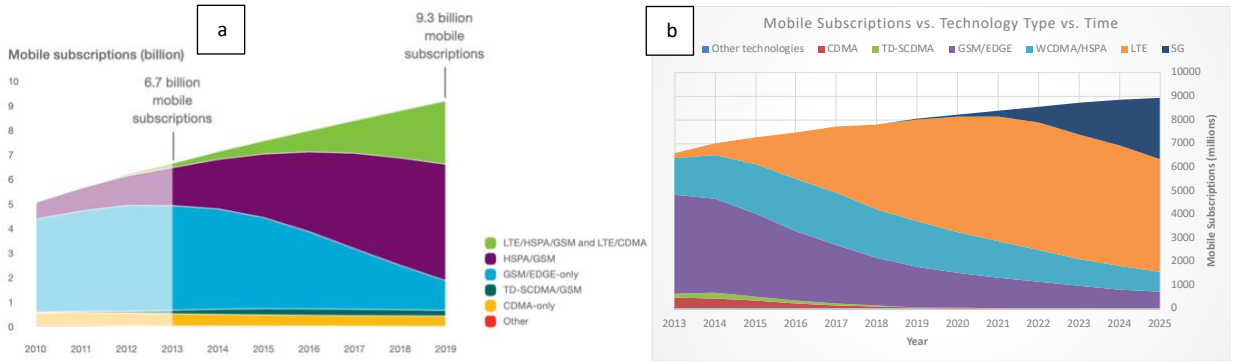

Figure 1.1 Mobile subscriptions by technology: (a) historical and predicted in 2013 [1], (b) historical and predicted in 2019 [2].

Figure 1.1 shows that the predicted number of subscribers in 2019 is slightly lower than the 2013 predictions. However, while the number of subscriptions is below the predictions six years prior, the amount of global monthly data traffic in 2019, seen in figure 1.2, is 40 exabytes per month [2], double the 19 exabytes per month predictions from 2013 [1]. Therefore, while predictions for the number of subscribers proved reasonably accurate, the exponential increase in data traffic grew at double the expected rate.

Data traffic growth driven by the following [2]: -

- Increase in the hardware screen sizes and resolutions.
- Changing user behavior – video anywhere, any time.
- Additional video components in most online content.
- Faster network speeds, enabling an improved user experience.
- Emerging immersive media formats, HD, UHD, 360-degree video, AR, VR.

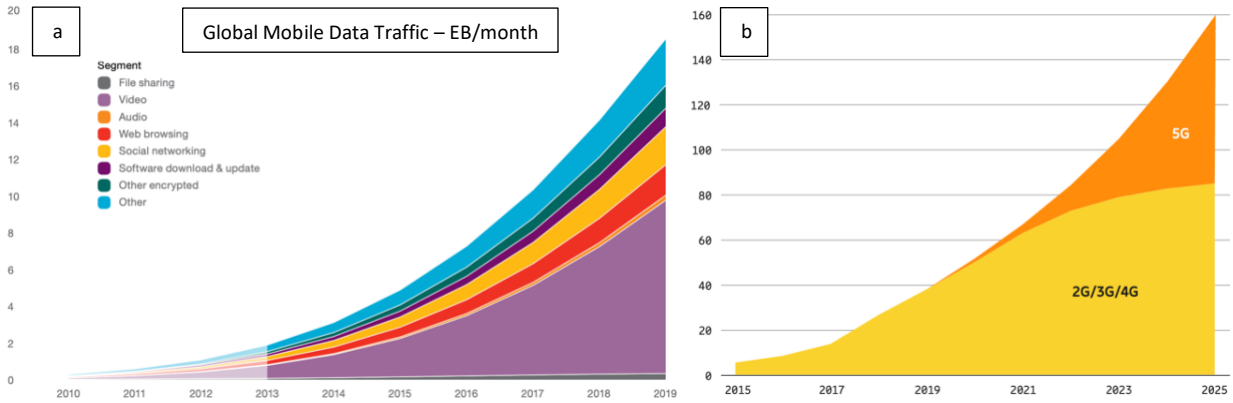

Figure 1.2 Global mobile data traffic: (a) historical and predicted in 2013 [1], (b) historical and predicted in 2019 [2].

This drive for higher and higher levels of content places an increased demand on the cellular network data capacity. In addition, the increase in traffic also raises environmental concerns for the global carbon footprint of the cellular network required to support the growth. Since 2007, the global CO_2 emissions from cellular infrastructure have increased from 86 $MtCO_2e$ to 235 $MtCO_2e$ in 2020 [4]. The highest power component within the radio is the RF power amplifier, typically consuming at least 60% of the total energy budget. Therefore, methods to improve the transmitted efficiency of the radio translate to a reduction in the network energy consumption and a corresponding reduction in CO_2 production.

Networks now employ more advanced modulation schemes such as 64, 128, or even 256QAM to enable higher data throughput per MHz of the frequency spectrum. The characteristics of these modulation schemes are high transmitted peak to average ratios and tight error vector magnitude

(EVM) requirements. These characteristics force the power amplifier to operate for the majority of time highly backed off from full power, while still capable of delivering peaks of up to 10dB over the average power level. The efficiency of traditional power amplifiers is low in this back off power region; therefore, techniques to enable high efficiency at high power back-off are necessary for the cellular infrastructure market. The tight EVM specifications necessitate highly linear power amplifiers, traditionally achieved using class-A bias and high back-off, the antitheses of the high-efficiency amplifier. Therefore, methods to enhance the high-efficiency amplifier linearity are popular research topics in this market space.

1.2 Research Objectives

1. Methods to enhance the RF bandwidth of the Doherty RF power amplifier.
2. Methods to enhance the linearized performance of the RF power amplifier within a digital pre-distortion system.
3. Methods to analyze and reduce the performance degradation incurred in scaling the power and frequency of the RF power transistor.

1.3 Novel Contribution

1.3.1 Contribution to Research Objective 1

1. A proposed Doherty topology providing a significant bandwidth increase at the back off average power level.
2. A proposed Doherty topology providing a significant bandwidth increase to the peaking amplifier off-state impedance.
3. An impedance compensation technique enabling extended bandwidth to shunt matched high-power RF devices.
4. A 100W LDMOS Doherty amplifier covering three cellular bands (1.8 – 2.2GHz) at >50% efficiency.

5. A 225W GaN Doherty amplifier demonstrating the proposed topologies and covering 1.8 – 2.7GHz at >45% efficiency.

1.3.2 Contribution to Research Objective 2

6. A proposed input matching topology for an RF power transistor with enhanced baseband impedance characteristics, and demonstrated improvements across multiple digital pre-distortion platforms
7. A proposed input matching topology for RF power transistors with enhanced RF bandwidth.

1.3.3 Contribution to Research Objective 3

8. The first known published work demonstrating the active load pulling effect on transistor cells from adjacent cells in the high-power transistor array. This causes non-uniform impedances to be presented to the cells at both the fundamental and the harmonic frequencies.
9. A proposed method to analyze the power and efficiency degradation of high-power RF power amplifiers due to physical distributed effects.
10. A novel multi-level integrated passive (IPD) device concept, enabling a reduction in performance degradation due to the internal matching network.
11. A high-power RF device design using novel IPD, demonstrating the following performance improvements: -
    - a reduction in the impact from distributed effects on from peak power of the transistor from 0.27dB to 0.01dB at 2GHz.
    - a reduction in the impact from distributed effects on from peak efficiency of the transistor from 2.85% to 0.34% at 2GHz.

1.4 Transactions, Conference Papers and Presentations

1. Wilson, R, “Increasing Bandwidth Trends Within the Wireless Infrastructure Sector”, WMG: Broadband PAs for Wireless Communications, IEEE IMS 2012.

2. Wilson, R., Goel, S., Singerl, P., “A Novel Input Matching Topology for Improved Digital Pre-Distortion of RF Power Amplifiers,” 2015 IEEE Topical Conference on Power Amplifiers for Wireless and Radio Applications (PAWR).

3. Wilson, R., Jang, H., Arigong, B., “Enhanced Instantaneous Bandwidth LDMOS RF Power Transistor Using Integrated Passive Devices,” 2016 IEEE MTT-S International Microwave Symposium (IMS).

4. Jang, H., Wilson, R., Canning, T., Seebacher, D., Shuberth, C., Arigong, B., “Self Out-phasing Chireix Power Amplifier Using Device Input Impedance Variation,” 2016 IEEE MTT-S International Microwave Symposium (IMS).

5. Jang, H., Wilson, R., Canning, T., Seebacher, D., Shuberth, C., Arigong, B., Trang, F., Ward, S., “RF-Input Self-Outphasing Doherty Chireix Combined Amplifier,” IEEE Transactions on Microwave Theory and Techniques, Volume 64, Issue 12, December 2016.

6. Canning, T., Herrmann, B., Jang, H., Mokhti, Z., Wilson, R., “A Novel Approach to Selecting Doherty Amplifier Asymmetry,” 2018 IEEE/MTT-S International Microwave Symposium (IMS).

7. Jang, H., Wilson, R., “1.8-2.7GHz Broadband Doherty Power Amplifier with Zero-Phase Shift Peaking Amplifier,” 2018 IEEE/MTT-S International Microwave Symposium (IMS).

8. Jang, H., Mokhti, Z., Herrmann, B., Wilson, R., “Nonlinear Embedding of FET Devices for High Efficiency Power Amplifier Design,” 2018 IEEE BiCMOS and Compound Semiconductor Integrated Circuits and Technology Symposium (BCICTS).

1.5 Patents

1.5.1 Patents pertaining to thesis chapter three

1. 8193857, June 2012, Wideband Doherty Amplifier Circuit
2. 20120319780, December 2012, Wideband Doherty Amplifier Circuit Having a Constant Impedance Combiner
3. 8717099, May 2014, Wideband Doherty Amplifier Circuit with Peaking Impedance Absorption
4. 20130076446, March 2013, RF Device with Complementary Resonator Matching Topology
5. 20150243649, August 2015, Power Transistor Die with Capacitively Coupled Bond Pad
6. 20190089317, March 2019, Broadband Harmonic Matching Network

1.5.2 Patents pertaining to thesis chapter four

7. 8736379, May 2014, Input Match Network for a Power Circuit
8. US8970308, March 2015, Input Match Network with RF Bypass Path

1.5.3 Patents pertaining to thesis chapter five

9. 20190088642, March 2019, RF Amplifier Package with Biasing Strip
10. Patent pending, Multi-level Integrated Passive Matching Structure

1.6 Thesis Organization

As described in section 1.1, the past decade has seen an exponential increase in cellular data demand. This demand has led to technical challenges in all base station radio areas, particularly the power amplifier. The demands fall into three performance areas for the power amplifier: -

1. Wide RF bandwidth – driven by the increase in cellular bands, and the flexibility of a multi band radio. Multiband designs lead to a reduction in manufacturer SKU's and enable smart allocation of power across the operation bands. For example, a single 80W band 1 + 3 radio enables a 40W + 40W power allocation between bands, and also enables a 60W + 20W allocation. A 2 x 40W single band design does not enable power allocation flexibility.
2. Wide signal bandwidth support and DPD correction. While the amplifier may provide wide RF bandwidth at high back-off efficiency, a concurrent multi band design requires both wide RF and signal bandwidth. The inability to correct the amplifier distortion to the specified levels disqualifies the radio from network deployment. Therefore, a wideband RF amplifier must also enable the linearization of wideband modulated signals.
3. High efficiency at high output power back off levels, driven by the high peak to average ratio of modulation formats used for cellular infrastructure. High efficiency topologies enable lower running costs and enable the installation of a higher power amplifier given the size and weight restrictions of the radio on the cell tower and health and safety weight limits for radio installers.

The research presented in this thesis seeks to enhance the state of the art in each of these areas as follows: -

In chapter three, the RF bandwidth capability amplifier is analyzed, and the learnings are used to develop two novel wideband Doherty amplifier topologies. Based on these topologies, two reference designs are developed with RF bandwidths capable of covering multiple cellular frequency bands.

Chapter four considers the amplifier signal bandwidth, focusing on the input matching topology of the RF device, an area that has received little research focus. A novel matching topology is introduced and an evaluation circuit is designed to demonstrate the matching topology's

performance benefits in a digital pre-distortion (DPD) system. Linearized results are presented and compared to a current state of the art input matched device. A 2-3dB improvement in linearized performance is demonstrated with multiple modulation formats in two DPD platforms.

Chapter five focuses on maintaining the power and efficiency of the RF device at higher frequencies and power levels. A novel analysis method is presented to quantify the device performance loss as the power, and therefore the device periphery, is increased. For the first time, the active load pulling effect on each die cell due to adjacent cells (or adjacent die) is shown. This effect also uses waveform engineering to demonstrate the resultant variation in the intrinsic current and voltage waveforms, verifying the previous observations during load-pull measurement [6] of cyclical hot and cold regions of the die under different load impedance conditions. The observation of an active harmonic impedance at the current generator, first presented in [5], is demonstrated using this analysis. For the first time, the concept of an active harmonic load due to distributed effects and adjacent die non-linearity is presented.

Conclusions from the research, and recommendations for further work are presented in chapter six.

1.7 References

[1] "Ericsson Mobility Report," November 2013.

[2] "Ericsson Mobility Report," November 2019.

[3] "Smart 2020: Enabling the low carbon economy in the information age," A report by The Climate Group on behalf of the Global eSustainability Initiative (GeSI), 2008.

[4] "The Global Carbon Footprint of Mobile Communications – The Ecological and Economic Perspective," IEEE Communication Magazine 49, August 2011.

[5] Mokhti, Z., Lees, J., Cessan, C., Alt, A., Tasker, P., "The Nonlinear Drain-Source Capacitance Effects on Continuous-Mode Class-B/J Power Amplifiers," IEEE Transactions MTT Volume 67, No.7, July 2019.

[6] Sevic, J.F., Albright, G., Schuerch, W., Simpson, G.M., "Simultaneous Load-Pull and Real Time Infrared Thermal Imaging of RF/Microwave Power Transistors," ARFTG Conference Digest, Spring 2004, pp.13-20.

Chapter Two – Overview of High Efficiency RF Power Amplifiers

In this chapter, an overview of methods to achieve high-efficiency amplification at backed-off power levels is presented. The architectures can broadly be divided into two categories – load modulating architectures, whereby the load impedance to the amplifiers in a variable with respect to output power, as seen in figure 2.1(a) – or supply modulating architectures, whereby the amplifier supply voltage is a variable with respect to output power, figure 2.1(b). Switch mode architectures, whilst demonstrating high-efficiency audio frequency amplification, have yet to show suitability for RF amplification. This is largely due to switching losses, the lack of complementary RF technologies that are commonplace at audio frequencies, and the coding efficiency of the modulators.

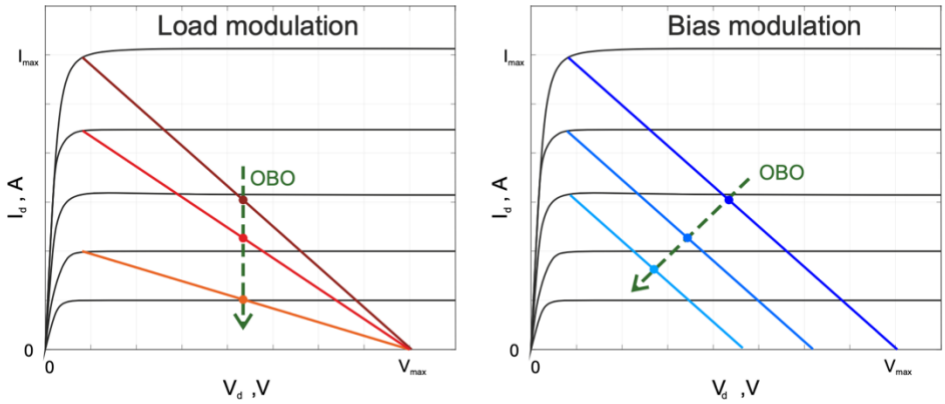

Figure 2.1 Load vs. supply modulation load line.

In general, supply modulation has found favor in applications with a narrow modulation bandwidth in combination with a wide RF bandwidth. This is because the design of wideband RF amplifiers is a well-proven field, and as the modulation bandwidth of the signal increases, the efficiency of the envelope modulator decreases.

Conversely, load modulating architectures have found favor in the cellular infrastructure market where the amplifier is targeted to a specific cellular band. This is because they are relatively agnostic to the signal bandwidth, but struggle to achieve wideband RF performance due to the band limiting nature of elements in the system, and the need for the system to operate into a range of load impedances. This topic is discussed in detail in chapter three.

2.1 Load Modulating Architectures

2.1.1 Doherty

The Doherty amplifier, named after the inventor William Doherty [1], dates back to 1936, conceived for high power tube amplifiers for broadcast applications. Since around 2004, the Doherty architecture began replacing quadrature combined or push-pull amplifiers and is now the de facto standard for the base station RF power amplifier. The popularity is because, unlike other techniques such as envelope tracking, the topology requires minimal implementation overhead. Also, the characteristics of narrow RF bandwidth, coupled with minimal signal bandwidth limitations, align well with the market demands for RF amplifiers in this sector, targeting specific cellular bands occupying under 10% fractional bandwidth. In recent years, much research has focused on increasing the RF bandwidth of the Doherty to achieve concurrent multi-band operation. The following market factors have driven this: -

- Multi operator cell site sharing
- More efficient RF power sharing between cell bands
- Lower hardware equipment variants
- Reduced radio form factor

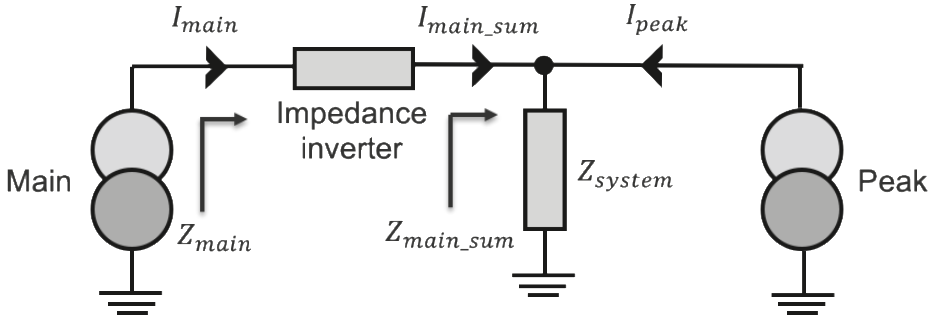

Figure 2.2 Doherty amplifier schematic.

As can be seen in figure 2.2, the Doherty amplifier operates using the concept of active load modulation. The main amplifier drives a system impedance defined in (2.1).

$$Z_{system} = \frac{Z_{opt}}{\left(1 + \frac{I_{peak\_max}}{I_{main\_max}}\right)} \tag{2.1}$$

Where I_{main_max} and I_{peak_max} is the peak current delivered from main and peak amplifiers, and Z_{opt} is the optimum load impedance presented to the main amplifier for full-power operation. It can be seen in (2.2) that through the action of the impedance inverter, the main amplifier sees a load modulated impedance higher than Z_{opt} in back-off.

$$Z_{main} = \frac{{Z_{inverter}}^2}{Z_{system}} \tag{2.2}$$

Therefore, at back-off power levels, the main amplifier operates into a higher load impedance, providing higher efficiency with reduced output power capability.

As the peaking amplifier conducts, the impedance seen by the main amplifier is modulated as described in (2.3).

$$Z_{main\_sum} = Z_{system} \times \left(1 + \frac{I_{peak}}{I_{main}}\right) \tag{2.3}$$

And at the main amplifier current generator from (2.4)

$$Z_{main} = \frac{{Z_{inverter}}^2}{Z_{system} \times \left(1 + \frac{I_{peak}}{I_{main}}\right)} \tag{2.4}$$

Figure 2.3 shows the resultant efficiency of the Doherty amplifier vs. output voltage.

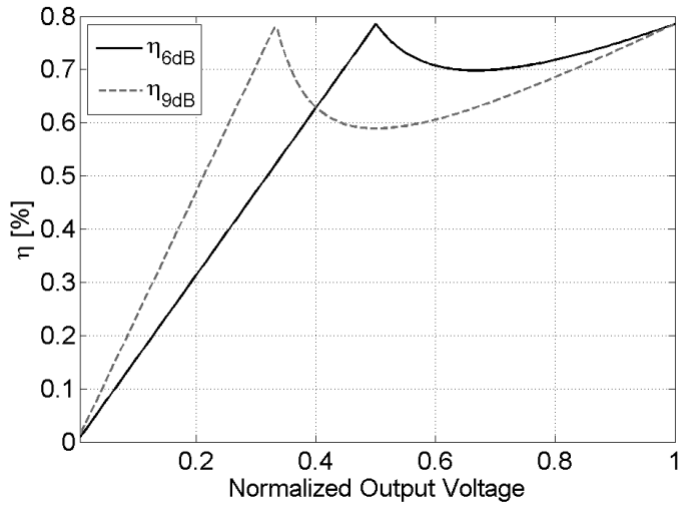

Figure 2.3 Doherty efficiency vs. output back off.

2.1.2 Chireix

The Chireix amplifier, named after the inventor Henri Chireix was first proposed in 1935 [2]. While similar to the Doherty in that both techniques employ load modulation to enhance the amplifier efficiency at the back off power levels, the method in achieving this differs significantly. Whereas the Doherty employs a peaking device to load-modulate the main amplifier between two load impedance states – the Chireix uses the two amplifiers simultaneously across all drive conditions. Figure 2.4 shows a conceptual schematic for the Chireix amplifier.

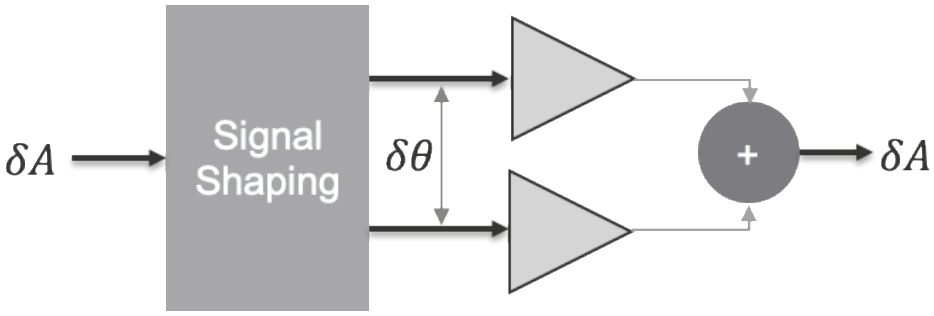

Figure 2.4 Chireix outphasing amplifier schematic.

The two amplifiers are connected using a non-isolating combiner, and the load modulation of both amplifiers is accomplished through applying a delta in phase between the two. The effect of the phase delta between the amplifiers is a reactance generated in the load impedance and can be calculated in (2.5) and (2.6). The corresponding load impedances are plotted in 2.5a.

$$Z_{PA\_1} = \frac{R_{load}}{2}(1 - \cot\theta) \tag{2.5}$$

$$Z_{PA\_2} = \frac{R_{load}}{2}(1 + \cot\theta) \tag{2.6}$$

The Chireix combiner uses compensation reactance on each branch, tailoring the load trajectory of the two amplifiers vs. out phasing angle. The load impedances trajectory of the Chireix after compensation can be seen in figure 2.5.

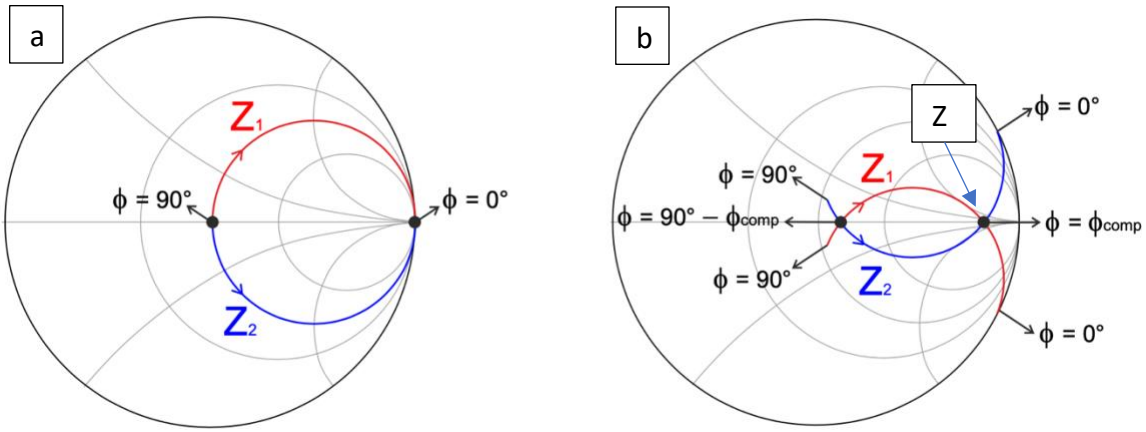

Figure 2.5 Chireix load modulation [3]: (a) without reactive compensation, (b) with reactive compensation.

The amount of reactive compensation applied to the Chireix combiner determines the output back off (OBO) at which peak efficiency occurs. High compensation provides a back off-peak at low power back off, and conversely, little compensation provides a back off-peak at higher OBO, sacrificing efficiency between the back off-peak (labeled Z_b in figure 2.5b) and full power as the presented load impedance deviates from real to real load modulation. The resultant efficiency curves of two different levels of compensation are shown in figure 2.6 in comparison to the Doherty amplifier tuned to the same OBO peak through the peak to the main asymmetric ratio.

Figure 2.6 shows both the advantages and disadvantages of the Chireix in comparison to the Doherty amplifier. For topologies requiring a back-off peak at higher OBO, the efficiency degradation due to the low efficiency of the peaking amplifier at low power levels causes an efficiency degradation between the OBO peak and full power levels in the Doherty amplifier.

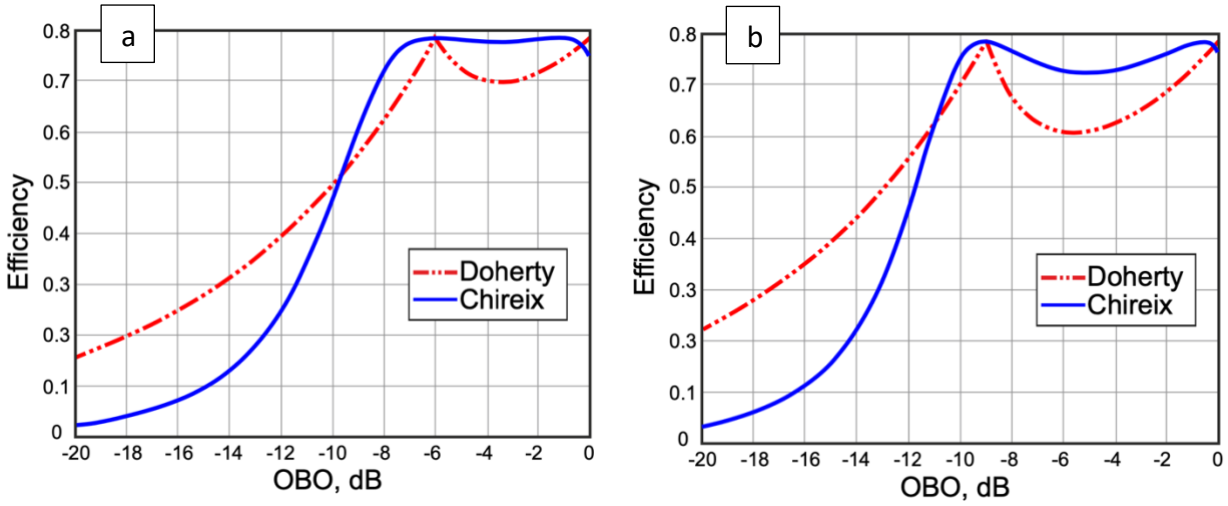

Figure 2.6 – Chireix vs. Doherty amplifier efficiency vs. OBO, tuned for: (a) 6dB OBO, (b) 9dB OBO.

Conversely, both the amplifiers in the Chireix operate at full swing into high load modulation in this region, providing greater efficiency that the Doherty. However, the disadvantage of the Chireix becomes apparent at power levels below the back off efficiency peak, where the PA load impedances deviate heavily and become highly reactive. This is seen in figure 2.5(b) and the impact seen in figure 2.6.

While the approach has merit, the adoption of the Chireix amplifier has found little traction within the cellular infrastructure market due to: -

- Faster efficiency degradation than Doherty at power levels below back off OBO peak
- Flatter efficiency between OBO peak and full power can also be achieved using n way Doherty (where n>2), whereby the low efficiency from the peaking amplifier is mitigated through the use of a Doherty peaking amplifier.
- Requires increased complexity of input signal shaping
- Practically achievable efficiency is not as high as theoretical due to device shunt losses and high amplifier load modulation ratio.

As with the Doherty amplifier, many of the issues of the Chireix may be alleviated through advances in digital signal processing. As described previously, the fast efficiency degradation below the OBO peak is due to the divergence in the amplifier load impedances. In this case, controlling the minimum phase angle between the amplifiers to prevent these load conditions can alleviate this undesirable characteristic. Also, higher-order Chireix implementations have been shown in [5] to provide considerably more ideal load modulation behavior digitally controlled drive signals.

2.1.3 Load Modulated Balanced Amplifier

Arguably one of the most interesting architectures to emerge in the last decade is the load modulated balanced amplifier, first proposed in 2016 by [4]. The concept uses active load-pull, applied using a control signal injected into the isolation port of a hybrid combiner, shown in figure 2.7.

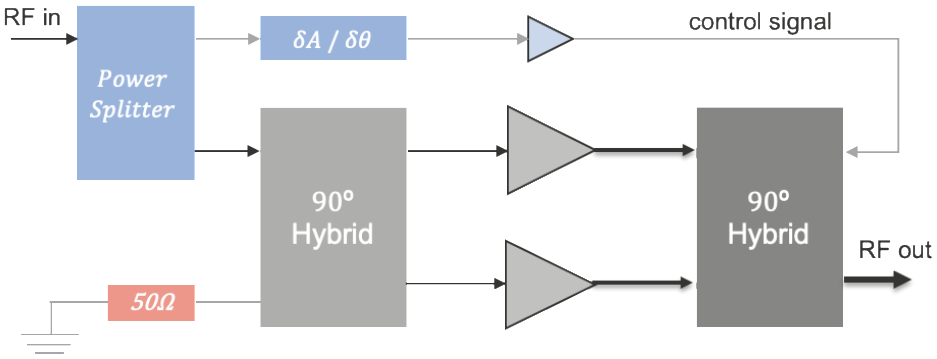

Figure 2.7 Load modulated balanced amplifier schematic.

The magnitude and phase of the control signal correspond to the load modulation magnitude and phase presented to the amplifiers, according to (2.7).

$$Z_{PA\_1,2} = Z_0\left(1 - \sqrt{2}\frac{I_c e^{j\emptyset}}{I_{pa}}\right) \tag{2.7}$$

Where Zo is the hybrid coupler impedance (typically 50Ω), and I_c and I_{pa} represent the control signal current and the PA current, respectively.

The load modulated balanced amplifier offers a number of advantages over other techniques: -

- A wideband high-efficiency amplifier can be achieved through the use of hybrid couplers and amplitude and phase control of the control signal amplifier
- The control signal is theoretically summed with the PA signal at the output; therefore, RF efficiency is not degraded through wasted RF generation.
- The balanced topology provides cancellation of the individual amplifier S11. Therefore, the amplifier presents a low VSWR load to the driver amplifier stage.

2.2 Supply Modulating Architectures

2.2.1 Envelope Tracking

The envelope tracking amplifier, shown in figure 2.8, employs a high-speed power supply modulator to provide the RF amplifier with the necessary DC bias to deliver the instantaneous RF envelope power.

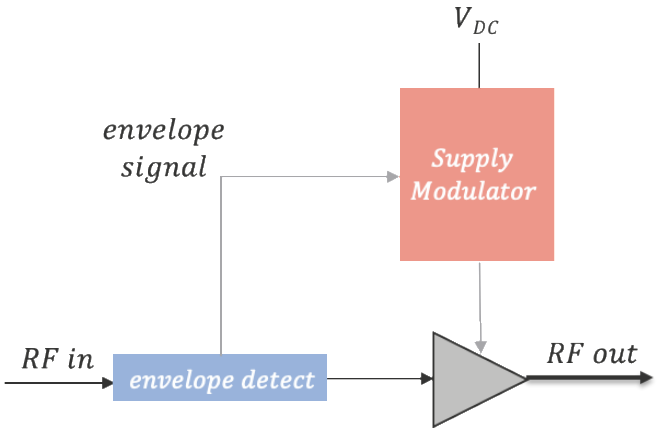

Figure 2.8 Envelope tracking schematic.

At lower power levels, the RF amplifier supply will operate at a significantly reduced supply voltage, therefore improving the efficiency of the device in this region, as shown in figure 2.9. The supply modulator may also act as a linearization method through the use of a supply voltage shaping function. The purpose of the shaping function is to determine the supply voltage corresponding to the signal envelope. Therefore, the modulator may be configured to provide a shaping function in order to optimize the characteristics of the system, for linear gain, or the highest efficiency.

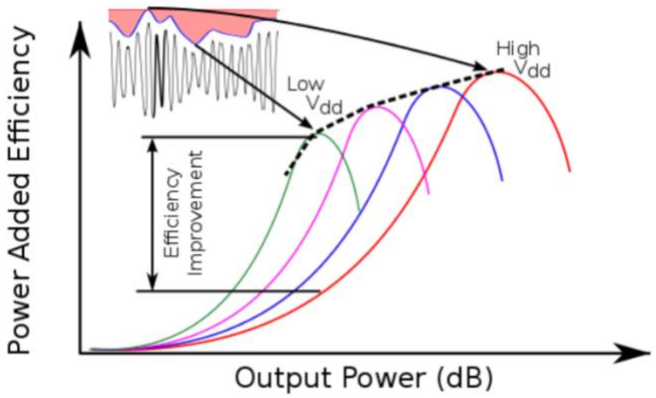

Figure 2.9 Envelope tracking efficiency enhancement.

While some impressive efficiency numbers have been achieved using envelope tracking amplifiers at cellular frequencies [7], state-of-the-art modulators have lagged the bandwidth requirements demanded from the market. Despite much research into bandwidth enhancement techniques [6], this has led to a relatively weak adoption in the base station radio transmitter.

2.2.2 Envelope Elimination and Restoration

The envelope elimination and restoration (EER) architecture, shown schematically in figure 2.10, was first proposed in 1952 by Leonard Kahn [8]. While some similarity exists to the envelope tracking amplifier, the two approaches fundamentally differ in the approach.

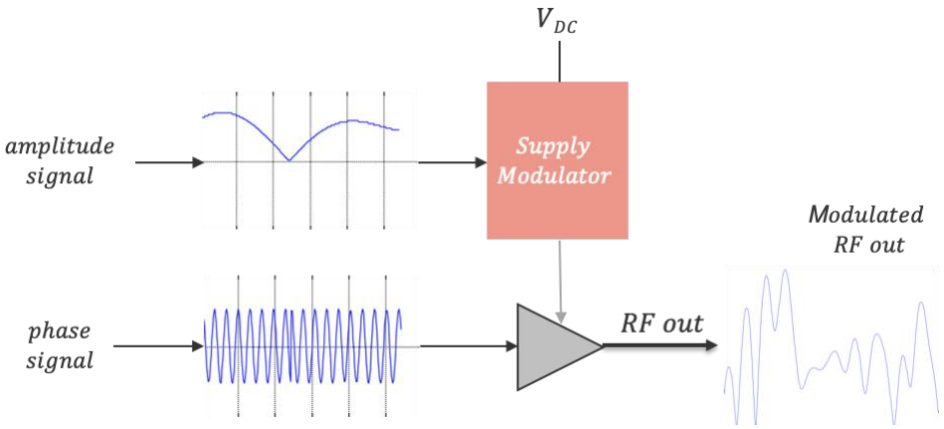

Figure 2.10 Envelope Elimination and Restoration Concept Schematic

In the EER topology, the amplitude and phase modulation are separated. The phase-modulated, constant amplitude signal is applied to the RF amplifier input; therefore, the RF amplifier operates at a constant, saturated output power level with high efficiency. The amplitude signal is amplified by a baseband amplifier, modulating the signal to the RF amplifier supply voltage. Therefore, the output from the RF amplifier is a 'restored' composite of the phase-modulated RF carrier and the amplitude modulated supply voltage.

The EER amplifier shares many of the same advantages and disadvantages of the envelope tracking amplifier. Conceptually, the technique enables the use of a high-efficiency RF amplifier operating mode such as class F or F-1. Also, the approach enables wide RF bandwidth since a broadband RF amplifier can be used independently of the applied modulator. However, the technique has gained little traction in the cellular base station market due to the signal bandwidth requirements exceeding the maximum practical modulation rate achievable from the AM amplifier. In addition, the increased complexity of input signal separation, and supply limitations with high PAR signals impact the adoption of the approach.

2.3 Hybrid

Many of the concepts presented in this chapter are complementary and therefore used together to enhance the amplifier performance. A non-exhaustive list is described in this section.

2.3.1 Supply Modulated Doherty

Conceptually, the two-way Doherty amplifier suffers from two main limitations: -

1. Efficiency drop at power levels between peak OBO and peak power due to low peaking amplifier efficiency at low power levels.
2. Efficiency drop below the OBO peak due to the limited main amplifier load modulation range.

Both of these limitations can be addressed using higher-order Doherty combinations, such as the 3-way Doherty alleviates efficiency drop off (1) by using a Doherty peaking amplifier to increase the efficiency at low power levels, thereby reducing the 'sag' in the 2-way efficiency vs. OBO curve.

However, both of these disadvantages can also be addressed using hybrid approaches. A two-way Doherty may use an envelope tracking modulator on the main amplifier branch to enhance the efficiency of the amplifier at power levels below the Doherty OBO peak. Conversely, the envelope tracking modulator may be applied to the peaking amplifier of a highly asymmetric 2-way amplifier to offset the efficiency degradation between the OBO peak and full power due to low peaking amplifier efficiency.

In practice, the use of envelope tracking on the main amplifier of a symmetric Doherty is the most popular topology. The. ET modulator extends the back-off efficiency peak past the Doherty 6dB OBO point, and the efficiency drop at peak turn on is lower compared to a highly asymmetric design tuned for a higher OBO back-off peak. Also, when using technologies with non-linear C_{ds} such as silicon LDMOS, the off-state impedance of the peaking amplifier becomes a function of supply voltage. Therefore the 'open circuit' phase angle cannot be maintained across the modulator supply range.

2.3.2 Doherty - Chireix

In Just as envelope tracking can be applied to the main or peaking amplifiers in the Doherty to improve device efficiency at lower individual power levels, an out-phasing amplifier may also be used to provide a stage with high efficiency in back-off. In this case, the advantage of the flat efficiency vs. OBO can be used to replace the individual main or peaking amplifier stage to alleviate the efficiency reduction at lower power levels (Chireix main), or between the OBO peak and full power (Chireix peaking).

While every possible topology combination has not described in this section, it can be seen that amplifiers using a combined load and supply modulation enable possible performance advantages over single approach topologies. The decision to develop a hybrid approach vs. a more complex derivative of the same topology (i.e., Doherty + ET vs. multi-way Doherty) needs to be evaluated against the system requirements in terms of complexity, signal bandwidth, RF bandwidth, and practically realizable efficiency in balancing device series and shunt losses.

2.4 Switch Mode

The switch-mode amplifier, shown in figure 2.11, has seen significant growth in the audio frequency range, enabling high fidelity, high power amplifiers without the need for the bulky heatsinks necessary for class A or high-power AB designs. However, the adoption of the switch-mode amplifier into the RF power domain has been less successful. The reasons for this are: -

- The Coding efficiency of the modulators is low and impacts the overall PA efficiency
- Lack of complementary RF devices, commonplace at audio frequencies
- Gate voltage swing to the upper device in the inverter is greater than output voltage swing with non-complementary devices
- High-speed switching final-stage gate capacitance takes significant current – necessitating high current driver
- Bandwidth is limited

- Output switching noise, shown in figure 2.12, requires filtering and is contrary to tight FCC emission requirements

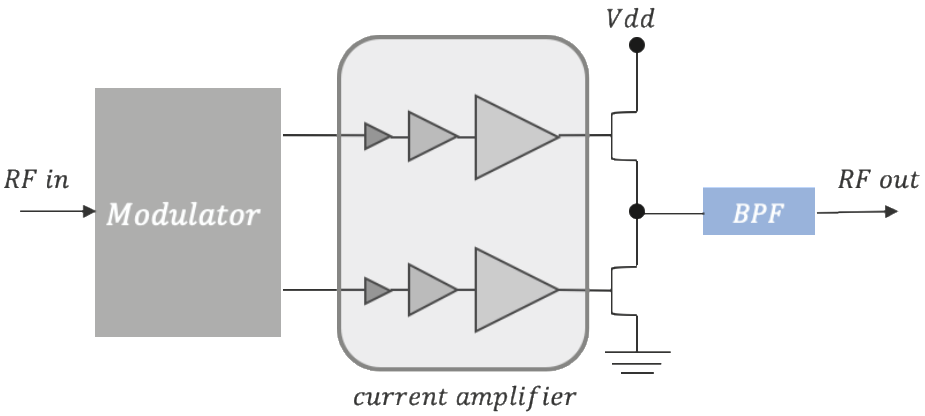

Figure 2.11 Typical switch mode RF amplifier concept schematic.

For these reasons, to the author's knowledge, no compelling demonstrations of switch-mode RF amplifiers have been made to suggest a viable alternative to the well-proven approaches described in this chapter.

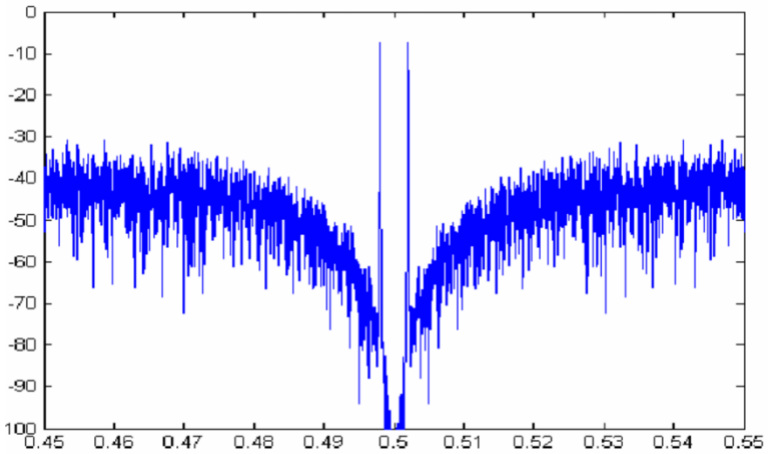

Figure 2.12 Delta Sigma Switch Mode PA Output Spectrum

2.5 Summary

In this section, an overview of high-efficiency amplifier architectures has been presented. While many techniques are known, this author has focused this research on the Doherty architecture for the following reasons: -

- The characteristics align with the cellular infrastructure requirements of narrow operating bands, but wide modulation bandwidth. The narrow RF bandwidth claims of the Doherty architecture enable this research to focus on wider band techniques for multi band cellular operation.
- The Doherty enables high efficiency operation without added implementation complexity, as is required for other techniques such as a supply modulator for an envelope tracking amplifier, or input signal shaping for a Chireix amplifier.
- Disruptive innovation aligns with future market needs. For the reasons described above, the Doherty has emerged as the de facto standard for the cellular base station power amplifier, and has proved to be fertile topology for research to reduce the RF bandwidth limitations.

2.6 References

[1] Doherty, W.H., "A New High Efficiency Power Amplifier for Modulated Waves," Proceedings of the Institute of Radio Engineers, Vol.24, Issue No.9, September 1936, pp.1163-1182.

[2] Chireix, H., "High Power Outphasing Modulation," Proceedings of the Institute of Radio Engineers, Vol.23, Issue No.11, November 1935, pp.1370-1392.

[3] Ramella, C. et.al., "High Efficiency Power Amplifiers for Modern Mobile Communications: The Load Modulation Approach," MDPI, November 2017.

[4] Shepphard, D., Powell, J., Cripps, S., "An Efficient Broadband Reconfigurable Power Amplifier Using Active Load Modulation," IEEE Microwave and Wireless Components Letters, Volume 26, Issue 6, June 2016, pp. 443-445.

[5] Barton, T., Perreault, D., "Four-Way Microstrip-Based Power Combining for Microwave Outphasing Power Amplifiers," IEEE Transactions on Circuits and Systems, Vol 61, No. 10, pp. 2987-2998, October 2014

[6] Godoy, P., Chung, S., Barton, T., Perreault, D., Dawson, J., "A 2.4-GHz, 27-dBm Asymmetric Multilevel Outphasing Power Amplifier in 65-nm CMOS", IEEE Journal of Solid State Circuits, Vol. 47, No.10, October 2012.

[7] Kimball, D.F., Jeong, J., Hsia, C., Draxler, P., Lanfranco, S., Nagy, W., Linthicum, K., Larson, L.E., Asbeck, P.M.,"High-Efficiency Envelope-Tracking W-CDMA Base Station Amplifier Using GaN HFET," IEEE Transactions on Microwave Theory and Techniques, Vol. 54, No.11, November 2006, pp. 3848-3856.

[8] Kahn, L., "Single Sideband Transmission by Envelope Elimination and Restoration," Proc. IRE, Vol. 40, Issue 7, July 1952, pp.803-806.

Chapter Three – Wideband Doherty Power Amplifiers

In this section, the bandwidth limitations inherent in the Doherty amplifier are investigated.

In Section 3.1, the Doherty amplifier structure is broken up into the band limiting elements shown in figure 3.1, and the bandwidth performance of each section is examined in terms of the impedance environment it is presented with within the Doherty itself.

In section 3.2, an impedance compensation technique is explored, whereby a series connected network of complementarity characteristics is added to the parallel shunt matched power transistor, and an increase in the matching bandwidth demonstrated.

Section 3.3 takes what is learned from section 3.1 and section 3.2, and implements these in a number of novel Doherty architectures. Measured results from two published wideband Doherty designs are presented.

3.1 Bandwidth Analysis of the Doherty Amplifier Elements

In this section, the bandwidth of each block of the Doherty amplifier is investigated in isolation. In later sections of the chapter, the impact of the blocks is characterized within the Doherty amplifier, and cascaded frequency compensation effects are considered. When considering the bandwidth of the structure, the location within the amplifier operation needs consideration, as certain blocks of the Doherty amplifier are exposed to load modulation, whereas other blocks operate in a fixed impedance environment. The bandwidth of each block therefore needs to be considered under all impedance conditions of operation.

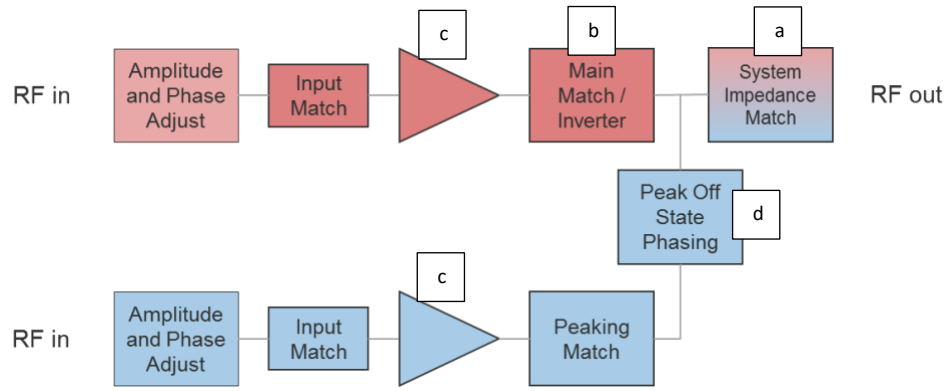

Figure 3.1 Doherty amplifier block diagram.

3.1.1 Summing Node Transformer

The summing node transformer (block (a), 'System Impedance Match' in figure 3.1) is responsible for defining the main and peak amplifier combining node impedance, and therefore the corresponding amplifier matching impedances. This is the only structure in the Doherty amplifier output network that operates outside of the amplifier load modulation. The term system impedance (Z_{sys}) defines the impedance at the summing node of the main and peak amplifiers, assuming no current flow from main or peaking branches. The term summing node impedance defines the instantaneous impedance presented to the amplifier given the current contribution from the other amplifier, as is calculated in 3.3. While the summing node transformer does not experience or play a direct role in the amplifier load modulation, it does provide the terminating impedance into which current flow from main and peak amplifiers are combined, therefore causing mutual load modulation. Given the fixed impedance transformation that occurs between the Doherty amplifier external load impedance and system impedance, wideband structures may be used in this location to transform the external impedance (typically 50Ω for transmitters) to the system impedance. In this section, the bandwidth of the six summing node transformer structures depicted in figure 3.2 (a-f) is analyzed, and the bandwidth and insertion loss of each is characterized.

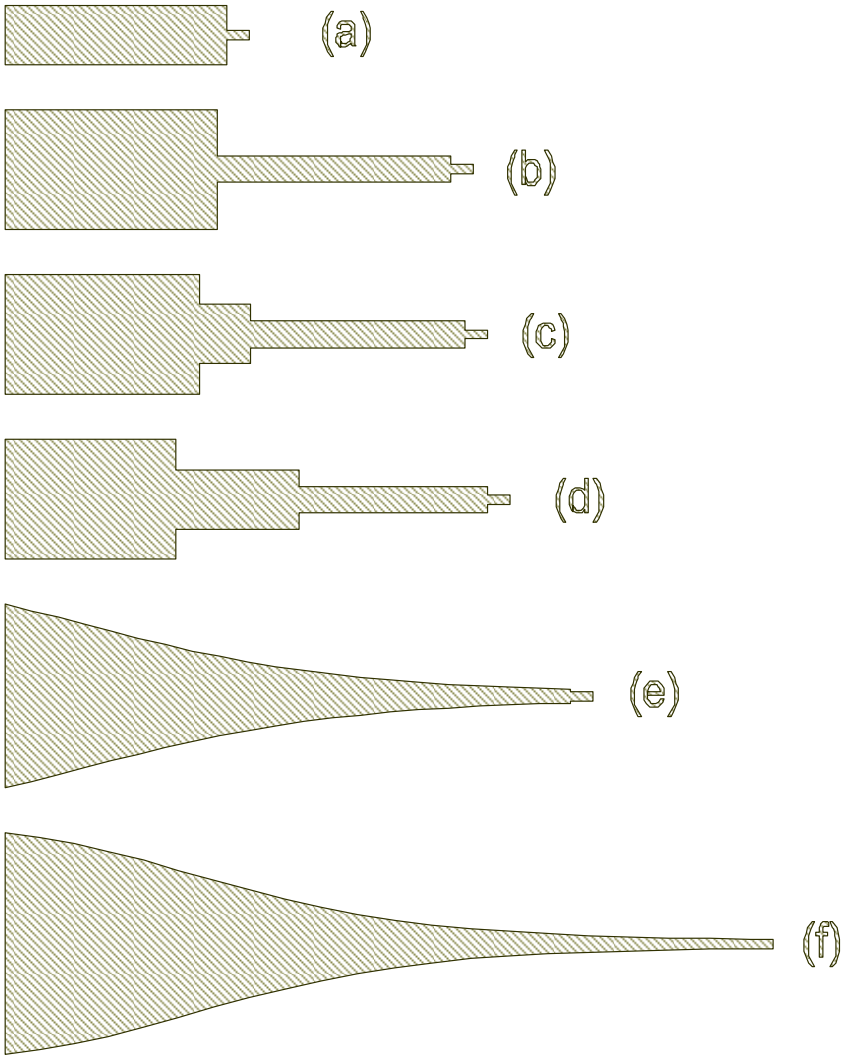

Figure 3.2 Summing node transformer physical size comparison, to scale: (a) single section transformer, (b) two-section transformer, (c, d) two-section transformer with geometric mean rotator, (e, f) Klopfenstein taper.

3.1.1.1 Single 90° Section Transformer

The single section 90° transformer, depicted in figure 3.2 (a), forms the simplest method for defining the Doherty system impedance, which can be calculated from 3.1.

$$Z_{sys} = \frac{Z_{transformer}{}^{2}}{Z_{termination}} \tag{3.1}$$

Since the electrical length of the single section transformer will only equal 90° at the target frequency f_c, real to real impedance transformation will only occur at this frequency. At frequencies above f_c, the inverter will 'over-rotate' the impedances past the target Z_{sys}, and between 90° and 180°, impedances will fall into the inductive hemisphere of the smith chart. Conversely, at frequencies below f_c, the inverter will 'under rotate' the impedances. As the electrical length of the inverter appears between 0° and 90°, the impedances will fall into the capacitive hemisphere of the smith chart. Figure 3.3 (a) shows the bandwidth of the single section transformer with a 10:1 transformation ratio, and (b) the corresponding insertion loss. The insertion loss is extracted from the maximum available gain based on 3D electromagnetic simulations. This characterization is needed in order to remove the mismatch loss from the measurement since this is considered when correlating target load impedance matching across frequency.

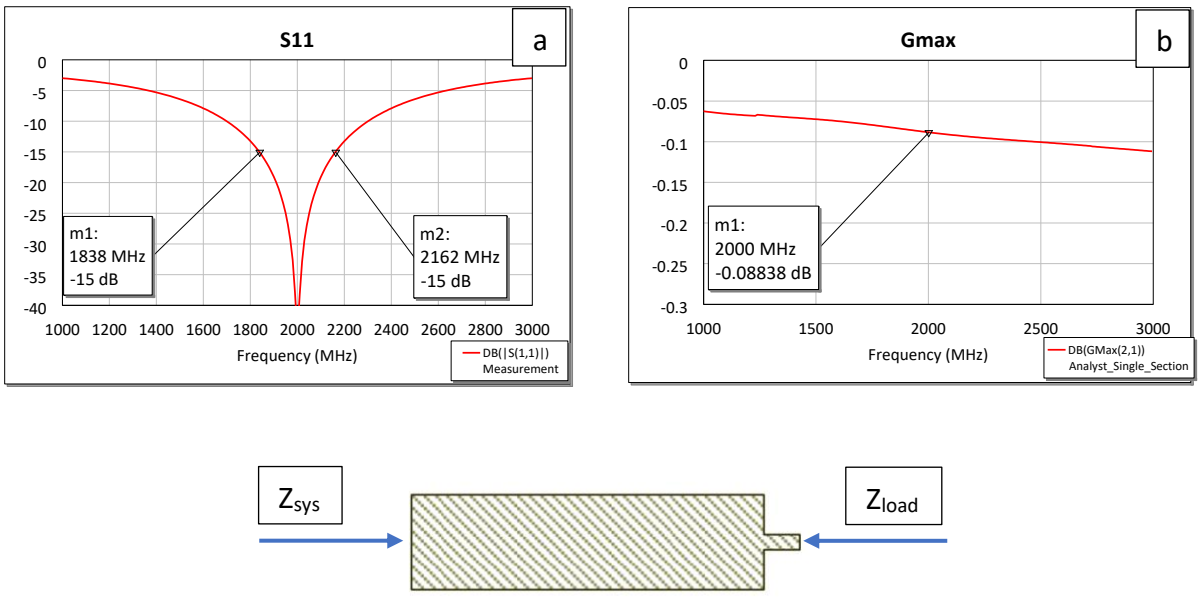

Figure 3.3 Single section transformer (a) bandwidth, (b) insertion loss.

It can be seen in figure 3.3 that the bandwidth of the single section transformer is limited to 16.2% when considering achieving a 15dB return loss from a 10:1 transformation ratio. However, due to the short 90° electrical length of the structure, a low insertion loss of 0.088 dB is achieved from the structure at 2GHz when implemented on a 25 mil Rogers 3010 PCB material.

3.1.1.2 Multi Section Transformer

The two-section transformer, depicted in figure 3.2 (b), forms a two section, 90° series connected transmission line transformer, 180° in electrical length. In this case, the bandwidth is increased in comparison to the single section presented in 3.1.1.1 by reducing the transformation ratio of each section from 10:1 (single section) to √10:1 (per section of the two-section match).

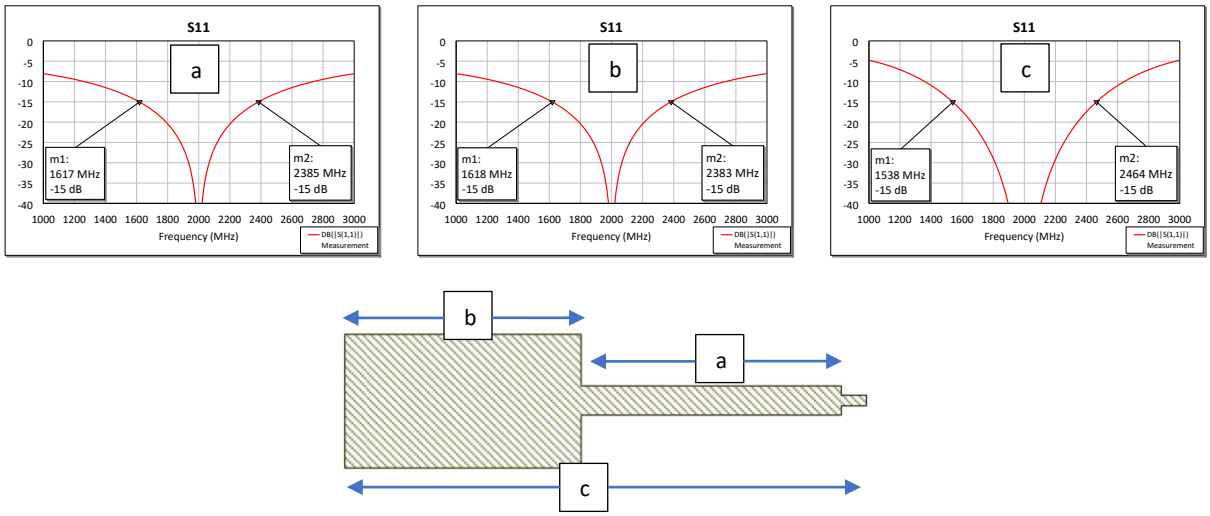

Figure 3.4 Transformation bandwidth: (a) load to geometric mean, (b) geometric mean to DUT, (c) combined load to DUT.

The two-section transformer is formed by the series connection of two band-limited networks, each performing a transformation to the geometric mean of the terminating source and load impedances. At first glance, it would be expected that the composite response of the full network is reduced in bandwidth when compared to each section in isolation. However, assuming system impedance specification of -15dB return loss or greater for the Doherty amplifier, the out of band

impedance response of the first matching section, when transformed through the second matching section remains within a -15dB return loss circle across a wider bandwidth than each of the matching sections in isolation. This can be seen in figure 3.4, where (a) and (b) depict the response of each individual section, achieving 38.3% fractional bandwidth, while the cascaded network, shown in (c) achieves a 46.3% fractional bandwidth. This is due to the compensation effect of the two transmission lines across frequency. At the intermediate impedance point, frequencies greater than the center frequency (f_c) are rotated >90°, thereby necessitating a second section transformation of >90° for transformation to the system impedance. Conversely, frequencies less than the center frequency (f_c) are rotated <90°, thereby necessitating a second section transformation of <90° for transformation to the system impedance. These conditions are satisfied by the second matching section. Therefore, the impedance dispersion of the first section provides the correct frequency compensation for the second section, which will achieve a tighter impedance grouping than the second section operating from Z_{int} to Z_{sys}.

3.1.1.3 Multi Section Transformer with Geometric Mean Rotator (GMR)

The structures in 3.2 (c, d) build on the concept described in section 3.1.1.2, whereby the bandwidth of the complete network is increased due to the out of band impedance trajectory of the preceding matching section. The two-section transformer with GMR utilizing the design methodology described below.

The first transformer section is designed to set the high side pole of the two-section matching network (shown at 2.45GHz in figure 3.6). Therefore, a quarter-wave transmission line at the target frequency of the high side pole is used for the first matching section. This provides a real to real impedance transformation at the upper frequency, and frequencies below this reside in the capacitive hemisphere of the smith chart, as shown in figure 3.5 (a)

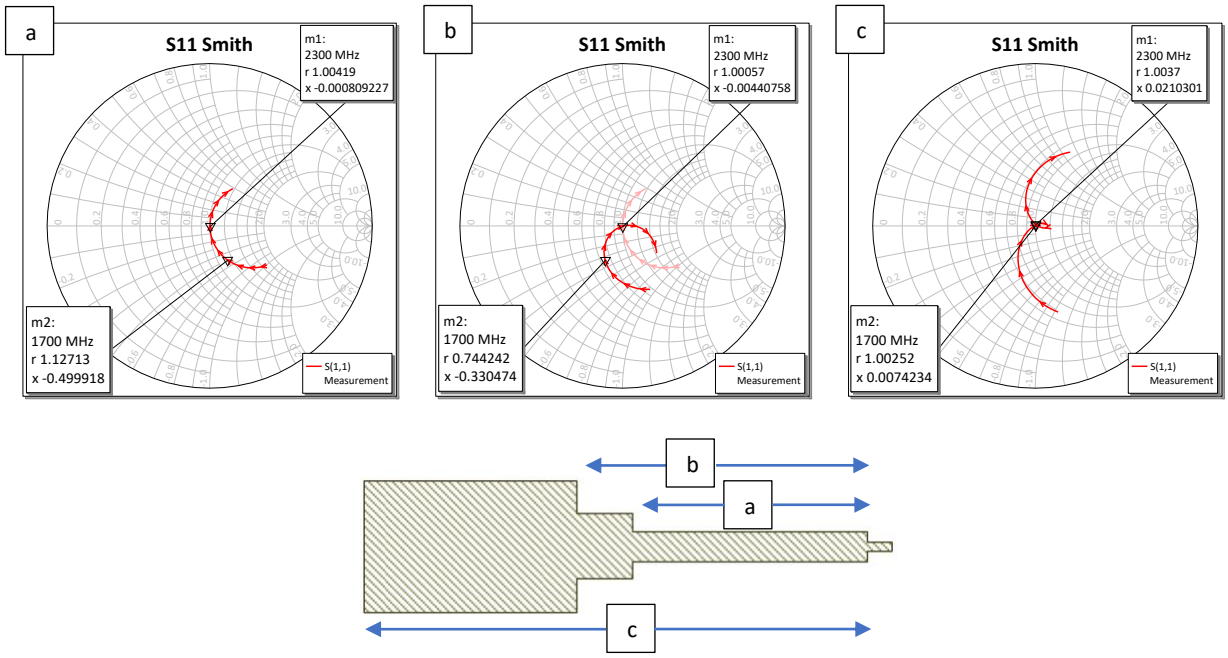

Figure 3.5 Impedance transformation: (a) first section, (b) first section with phase rotation from GMR, (c) first section + GMR + second section.

At this point, two characteristics of the transmission line transformer are utilized. Firstly, due to the greater electrical length per unit physical length at higher frequencies, impedances at higher frequencies rotate faster than those at lower frequencies through the same matching structure. Secondly, a transmission line transformer rotates impedances about the characteristic impedance of the line. Therefore, a load impedance equal to the characteristic impedance of the line will remain at a constant impedance through the transformer. By contrast, a load impedance unequal to the line impedance will exhibit phase rotation at a constant VSWR around the transformer characteristic impedance.

These two transmission line characteristics are utilized in the design. The GMR is used to rotate the phase of the lower frequency pole in order to present the correct phase angle for the final transformer to rotate both upper and lower impedances to the real axis. The impedance alignment occurs since the rotation of M1 at 2300MHz will be higher than M2 at 1700MHz, thus compensating for the increased transformation required. Figure 3.5 shows the impedance at the

nodes of the two section + GMR transformer, (b) at intermediate impedance and (c) the load impedance, or system impedance when used in the Doherty amplifier.

The bandwidths of the multi section transformer, and the multi section transformer with GMR are shown in figure 3.6. The fractional bandwidth of the multi section transformer is increased from 46.3% (two-section) to 68.3% with the addition of the GMR.

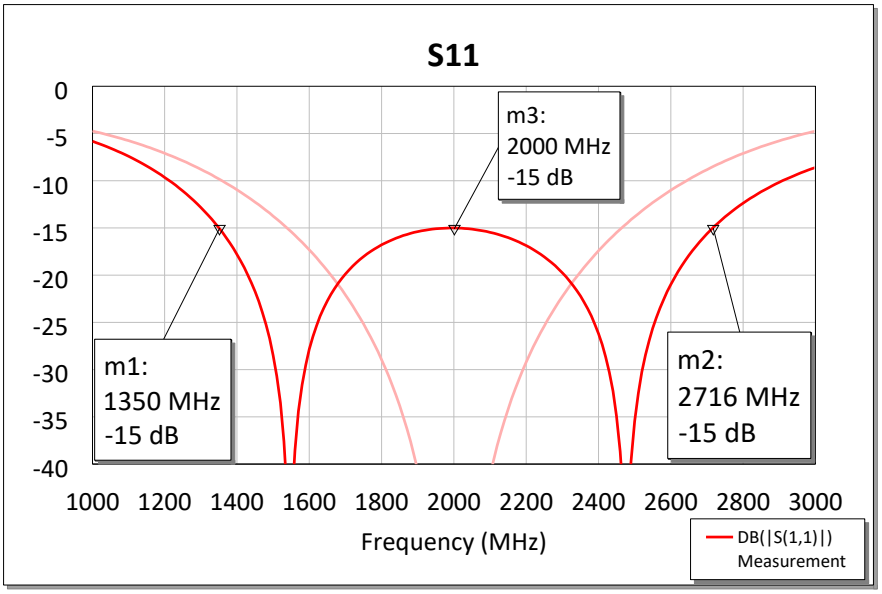

Figure 3.6 Bandwidth of two section match with GMR.

It can be seen in figure 3.6 that, unlike the two-section match, the two-section match plus GMR presents a bandpass response with finite ripple, the magnitude of the ripple being a function of bandwidth and number of matching sections. Therefore, the comparison metric of interest is achievable bandwidth vs. return loss specification across the amplifier operating bandwidth. This comparison is made in figure 3.7.

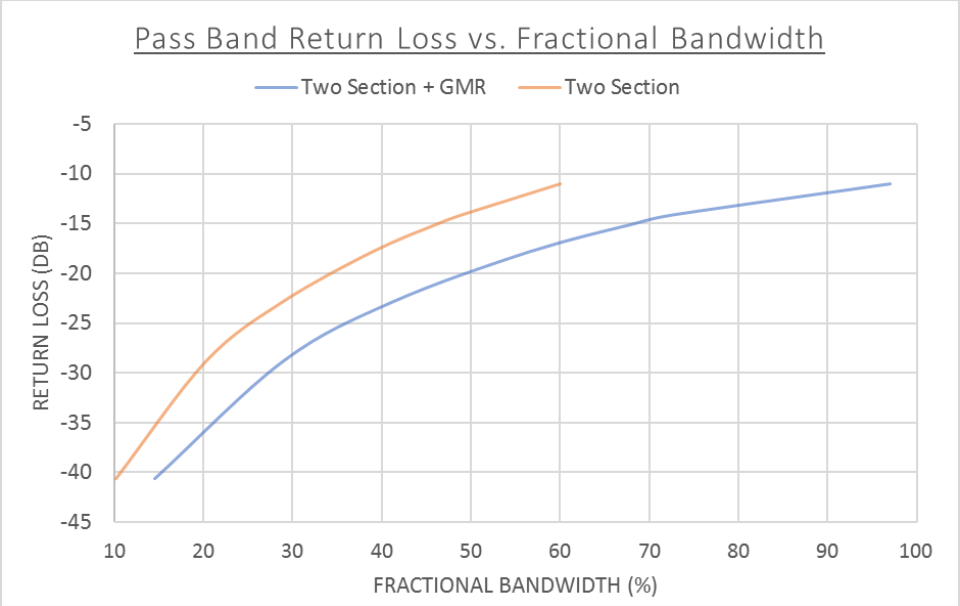

Figure 3.7 Two-section matching topologies, fractional bandwidth vs. pass band ripple.

It can be seen in figure 3.7 that at all target return loss specifications, the bandwidth of the two-section match with GMR is greater than the two-section match alone. Therefore, this matching technique would appear to provide a suitable wideband method for defining the Doherty system impedance. This assumption is validated by the insertion loss comparison in figure 3.8, where an insertion loss delta of 0.002dB is observed between the two structures, a negligible difference. It should be noted, however, that due to the greater physical and electrical size of the structures, both of the implementations in this section exhibit 0.06dB higher insertion loss than the narrower band, single section match shown in figure 3.2a and described in 3.1.1.1. Therefore, for narrow-band amplifier designs targeting the highest achievable efficiency, the simple quarter-wave transformer still provides merit.

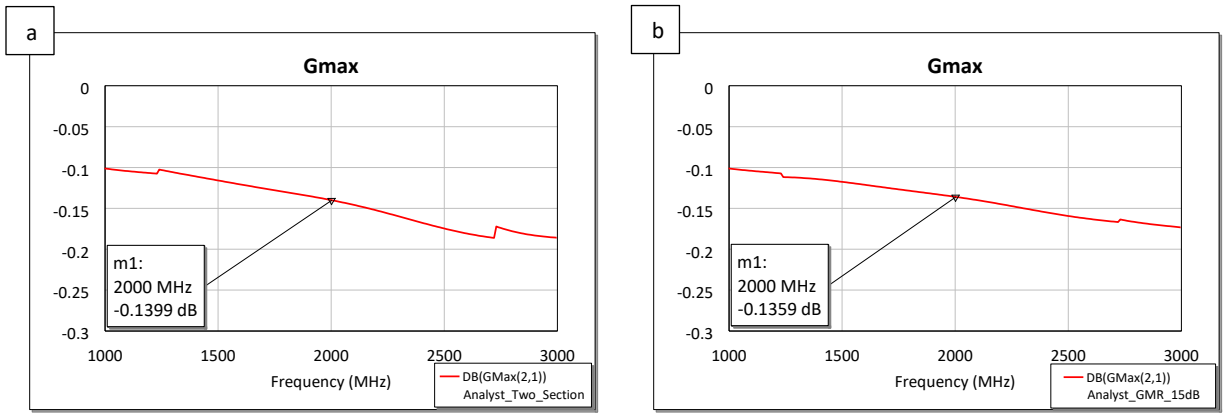

Figure 3.8 Insertion Loss Comparison: (a)two-section match, (b) two-section match plus GMR.

3.1.1.4 Klopfenstein Taper

As is described in detail in [1] - [4], the Klopfenstein Taper presents a Chebyshev high pass response, with a low-frequency cut off determined by both the electrical length of the taper, and the target passband return loss. The design tradeoffs between these two parameters are shown in figure 3.9, where the response of three tapers of equal physical length, designed for a passband return loss of 10, 20, and 40dB respectively, are shown.

The corresponding variation in impedance along the line for each structure in figure 3.9 is shown in figure 3.10, showing the change in the tapered shape for the three passband ripple conditions. The impedance uses the corrected formulation proposed in [4].

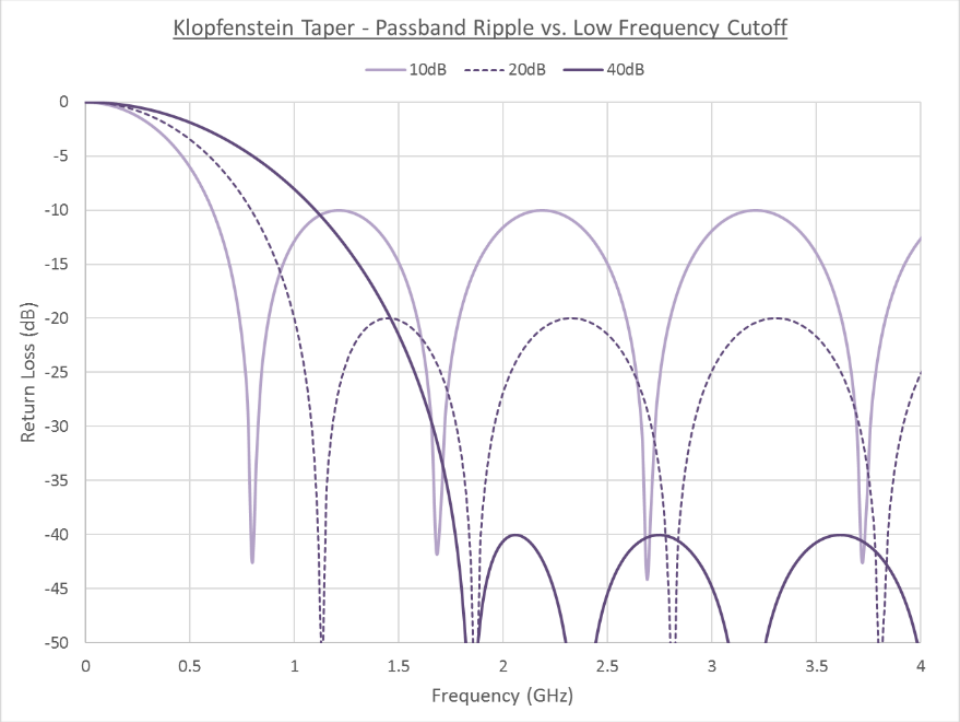

Figure 3.9 Klopfenstein taper return loss vs. electrical length.

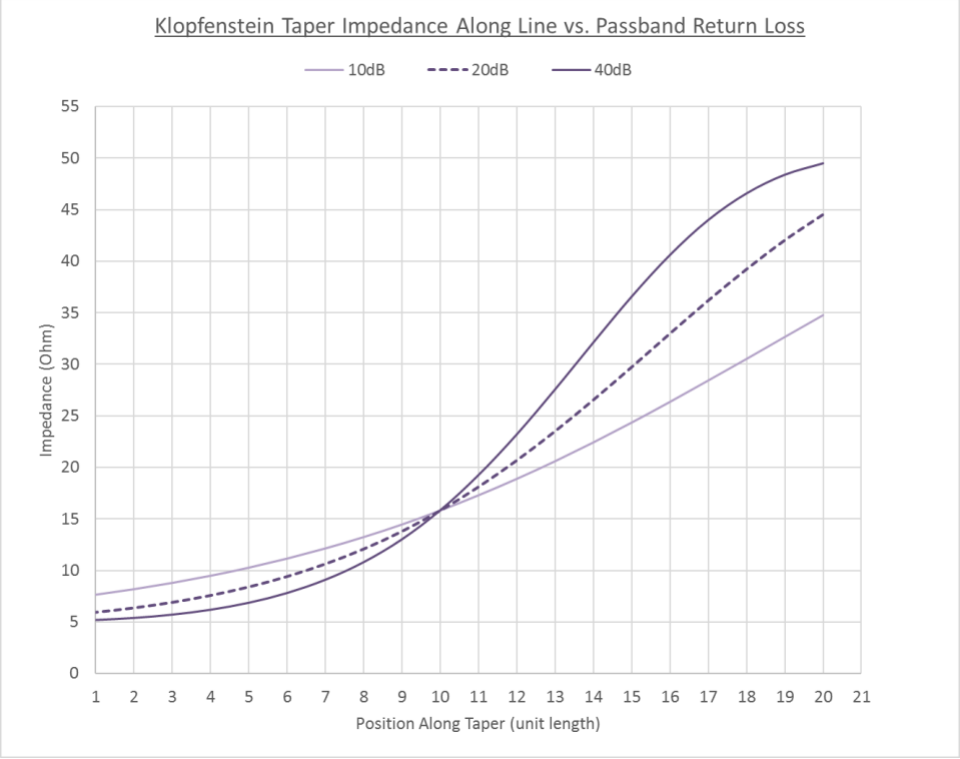

Figure 3.10 Klopfenstein taper impedance vs. n^{th} physical position along the taper.

A comparison is made between the two-section with GMR matching structure in 3.1.1.3 and the Klopfenstein taper. A Klopfenstein is designed with an identical low-frequency response to the two-section match described in 3.1.1.3 and the insertion loss, physical size, and passband response compared. Figure 3.11 shows the response of the taper designed with a low-frequency cutoff of 1350MHz and a passband ripple of -15dB return loss, with the two-section match with GMR from 3.1.1.3 shown for reference.

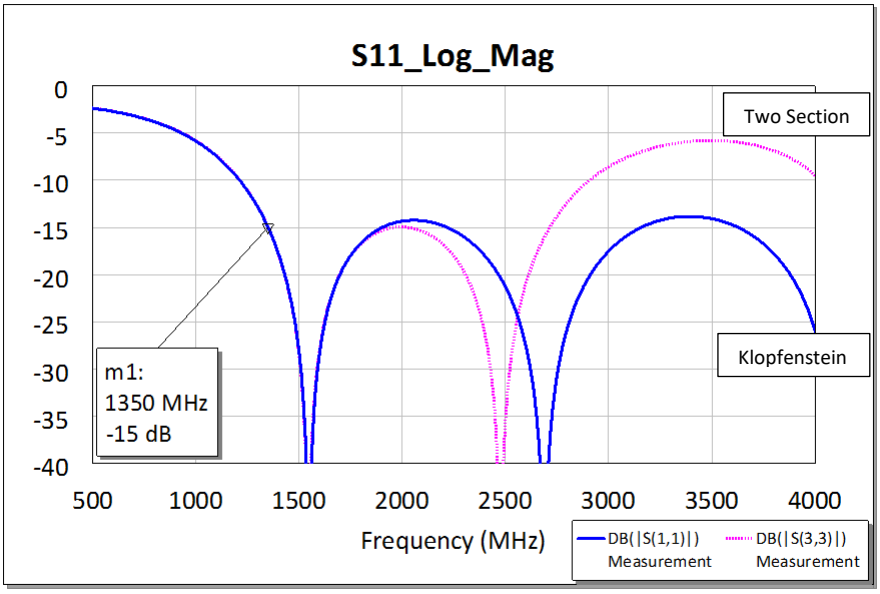

Figure 3.11 Klopfenstein taper bandwidth compared to two section with GMR.

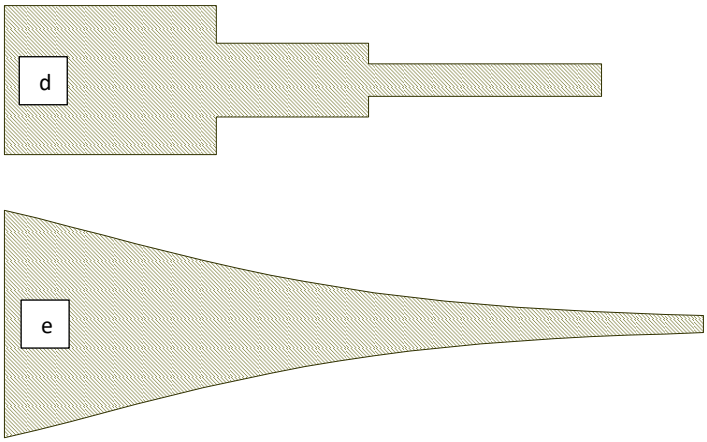

Figure 3.12 Physical size: (d) Two Section with GMR and (e) Klopfenstein Taper, 15dB Return Loss (labels correspond to those in figure 3.2).

When designed to an identical low-frequency cut-off and 15dB return loss specification, the response of the two-section + GMR and the Klopfenstein are very similar in physical size (shown in figure 3.12), insertion loss (figure 3.13).

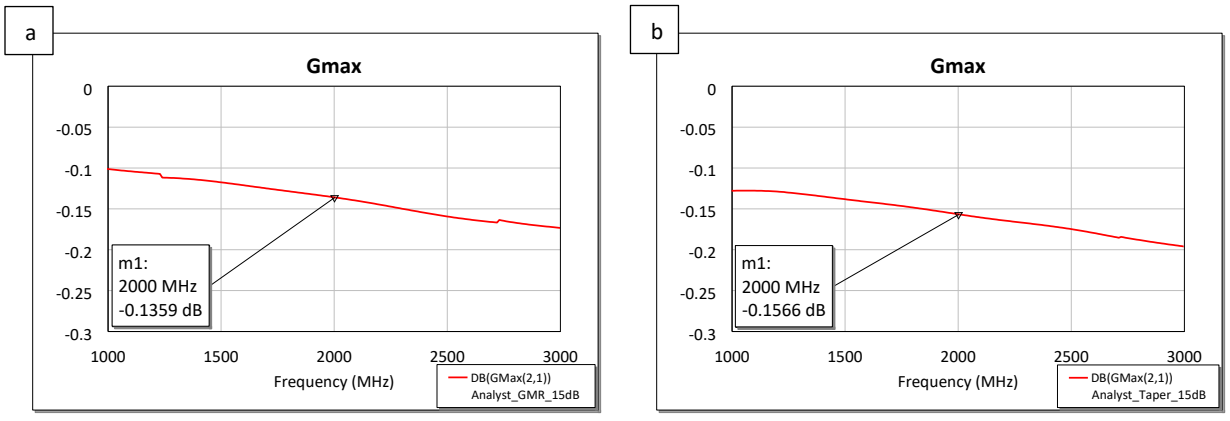

Figure 3.13 Insertion loss with 15dB passband ripple: (a) two-section transformer with GMR, (b) Klopfenstein taper.

The bandwidth of the structures is similar, apart from the inherent characteristics of the structures, whereby the two-section GMR match exhibits a passband response, and therefore the high-frequency response is limited. In contrast, the Klopfenstein taper exhibits a high pass response and is, therefore, only limited at the low-frequency region.

A similar comparison is made assuming a higher return loss specification of 30dB at the summing node. In this case, a two-section GMR match is designed to achieve a 30dB return loss, giving rise to a low-frequency cut off at 1722MHz. The Klopfenstein taper is designed to the same return loss and low-frequency cutoff specification, and the response of the two networks are compared in figure 3.14.

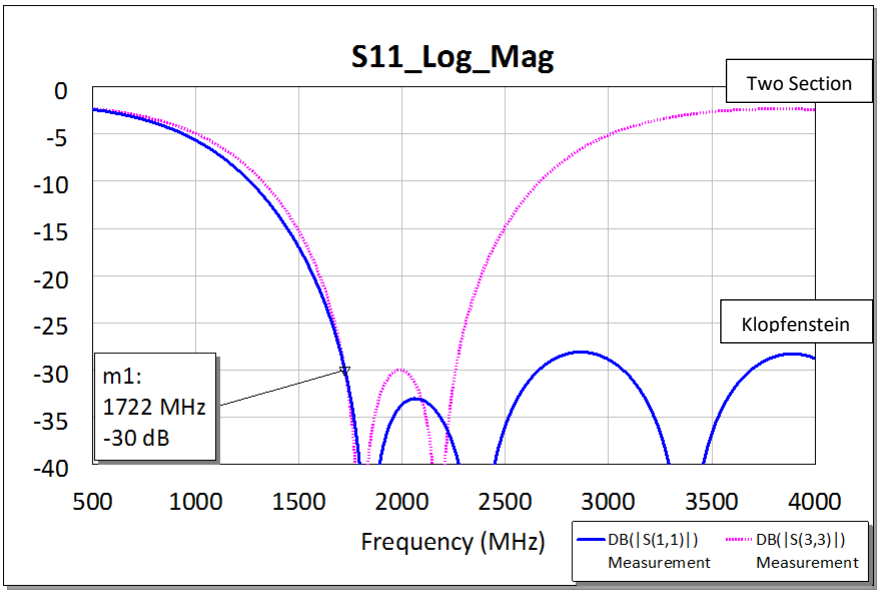

Figure 3.14 Klopfenstein taper bandwidth compared to two section with GMR.

As is seen in figure 3.14, the increased return loss specification reduces the bandwidth of the two-section match from 69 to 27% fractional bandwidth. By contrast, the Klopfenstein taper maintains high return loss at frequencies above the low-frequency cut-off. Therefore, the structure provides a wideband summing junction impedance at all frequencies above the low-frequency cut-off.

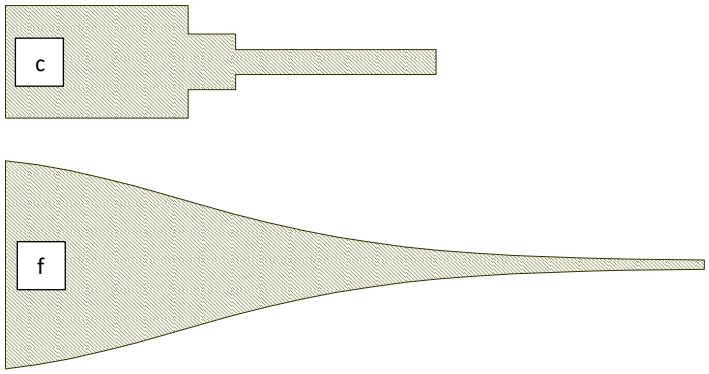

Figure 3.15 – Physical size: (c) two section with GMR and (f) Klopfenstein taper – 30dB return loss (labels correspond to those in figure 3.2)

However, as can be seen in figure 3.15, the cost penalty of the Klopfenstein is the physical size of the structure is increased from 1.1x (figure 3.12) to 1.7x (figure 3.15). Therefore, for practical applications where a compact form factor may be required, the additional real estate occupied by the Klopfenstein may prove to be non-ideal. In addition, given the uniform width of each section of the two-section match, a folded structure is relatively easy to optimize using electromagnetic simulations to achieve a more compact footprint. By contrast, a folded Klopfenstein taper is difficult to design since the structure is continually variable in width. The development of a design approach for the folded Klopfenstein is a proposed area of further work in section 3.4.

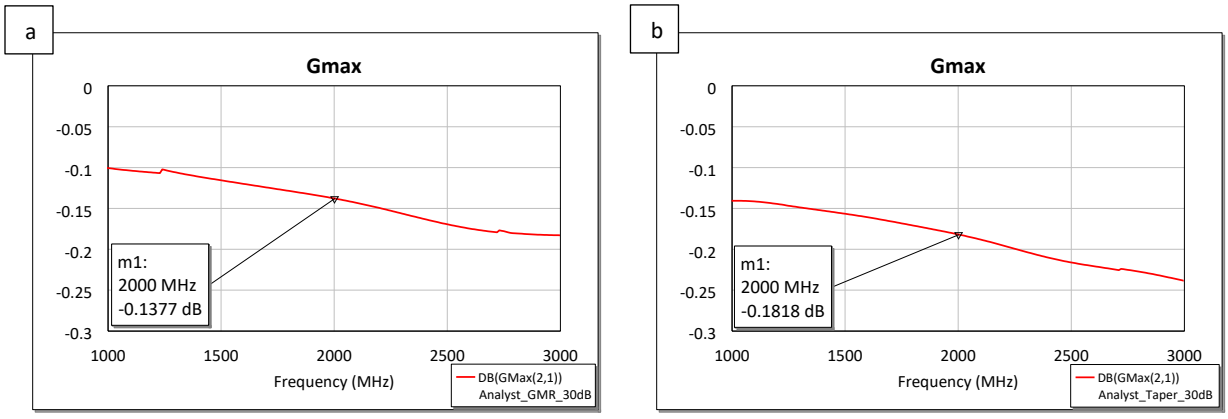

Figure 3.16 Insertion loss with 30dB passband ripple: (a) two-section transformer with GMR, (b) Klopfenstein taper.

Finally, the bi-product of the additional size of the structure is an increase in the insertion loss, as can be seen in figure 3.16. At the center frequency of 2GHz, it can be seen that the insertion loss is increased by 0.044dB over the two-section GMR transformer. The impact of the insertion loss on the amplifier efficiency is shown in figure 3.17 and correlates to a 1% efficiency reduction.

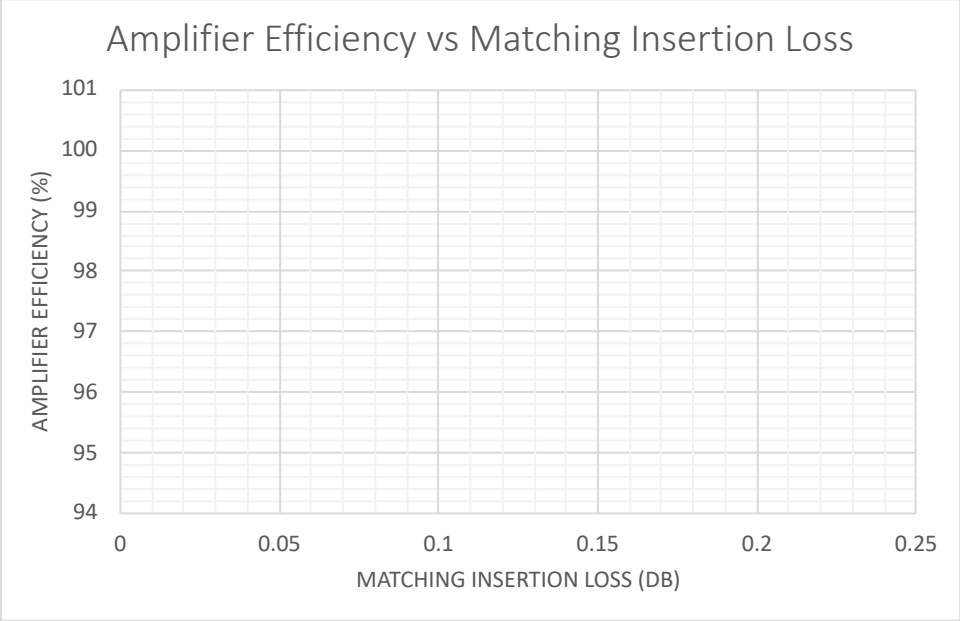

Figure 3.17 Impact of insertion loss on amplifier efficiency.

3.1.1.5 Summary

In this section, three methods of defining the Doherty amplifier system impedance are compared. The suitability of each method is broken down into three categories of fractional RF bandwidth.

For narrow bandwidth applications, up to around 15% fractional bandwidth, the most straightforward implementation comprising of a single, quarter-wave transformer provides the most compact and lowest insertion loss method of defining the Doherty amplifier summing node impedance. This design approach is commonly utilized on many single-band Doherty amplifier designs in the field today.

For moderate fractional bandwidths, from 15 - 65%, the use of a two-section matching network provides a low insertion loss and compact method of defining the Doherty system impedance. It should be noted, however, that in all cases, the two-section match with GMR achieves higher bandwidth than the traditional two-section quarter-wave transformer match with no detrimental effect on either physical size or insertion loss. Therefore, for moderate bandwidth Doherty designs, the two-section plus GMR is considered a superior choice when compared to the two-section match.

For wideband applications targeting 60% fractional bandwidth or higher, particularly when a low summing node VSWR is targeted, the high pass response of the Klopfenstein taper provides the optimal method for setting a wideband system impedance to the Doherty amplifier. The high pass response provides effectively infinite bandwidth, provided the taper low-frequency cutoff meets the lowest frequency required. In this analysis, the overlap region between Two Section + GMR and the Klopfenstein occurs at around 60% fractional bandwidth and 15dB target return loss, with both structures approximating similar size, insertion loss, and bandwidth. For higher insertion loss requirements, the bandwidth limitations of the two-section match are reduced, as shown in figure 3.7, and the gap in bandwidth between the two matching techniques increases, albeit at the expense of size and insertion loss in the case of the Klopfenstein.

While not analyzed here but discussed in section 3.4 – Conclusions and Further Work, two other approaches in setting the system impedance are the use of impedance transforming balun or hybrid couplers. These approaches necessitate the use of two Doherty amplifiers in parallel but can provide a wideband system impedance into which the Doherty amplifier can operate.

3.1.2 Main Amplifier Impedance Inverter

The main amplifier impedance inverter (block (b), 'Main Match / Inverter' in figure 3.1), performs two functions within the Doherty amplifier. As shown in equation 3.2, the impedance presented to the main amplifier (Z_{load}) is inversely related to the impedance at the summing node. Therefore, an increasing summing node impedance will cause a corresponding reduction in Z_{load}.

$$Z_{load} = \frac{Z_{inverter}{}^2}{Z_{sum}} \tag{3.2}$$

Injection of peaking amplifier current into the summing node occurs as the class C peaking amplifier conducts (at 6.02dB OBO for the classical, symmetric Doherty amplifier). From equation 3.3, it can be seen that as current from the peaking amplifier is increased, the apparent summing node impedance increases from Z_{system} to Z_{sum}.

$$Z_{sum} = Z_{system} \times \left(1 + \frac{I_{peak}}{I_{main}}\right) \tag{3.3}$$

Therefore, for the symmetric Doherty, between 6.02dB output back off and full power, the summing node impedance is increasing due to the increasing peaking amplifier current delivered. Figure 3.18b shows the dynamic impedance of the summing node, as seen by each amplifier, corresponding to the Doherty drive up curves in figure 3.18a.

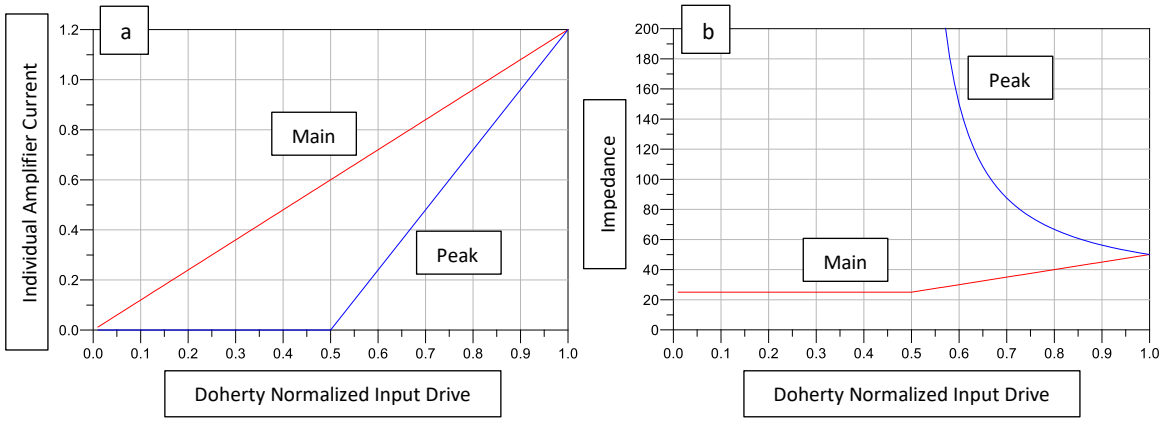

Figure 3.18: (a) Doherty drive up curves, (b) dynamic summing node impedances.

However, the requirement for correct main amplifier operation is a reducing main amplifier load impedance as the drive level, and therefore Doherty output power level increases. This condition enables the main amplifier to remain in a voltage saturated condition, with a progressively

increasing drain current due to the reducing load impedance. Ensuring this operation is the function of the Doherty impedance inverter, ensuring a decreasing main amplifier load impedance due to increasing peaking amplifier current contribution to the summing node. This is accomplished through the characteristics of a 90° transmission line, as can be seen from equation 3.4.

$$Z_{main} = \frac{{Z_{inverter}}^2}{Z_{sum\_eff}} \tag{3.4}$$

Figure 3.19 shows the dynamic load impedance for main and peaking amplifiers in a symmetric Doherty with an and assumed full power optimum load impedance (R_{opt}) of 5Ω. The impedance of the main amplifier would be: -

$$Z_{main\_full\_power} = \frac{{Z_{inverter}}^2}{Z_{sum\_eff}} = \frac{15.81^2}{50} = 5\ \Omega \tag{3.5}$$

and

$$Z_{main\_backoff} = \frac{{Z_{inverter}}^2}{Z_{sum\_eff}} = \frac{15.81^2}{25} = 10\ \Omega \tag{3.6}$$

Therefore, the load presented to the main amplifier is high at the back off power levels and reduced by the load modulation ratio at full power. This impedance trajectory provides correct operation, enabling the main amplifier to operate into a high-efficiency, high impedance load at a back-off power level, and a high-power, low impedance load at the full power. Figure 3.6 shows the plots for both main and peaking amplifiers at the summing node (b) and the DUT (a) assuming a 5Ω optimal amplifier load.

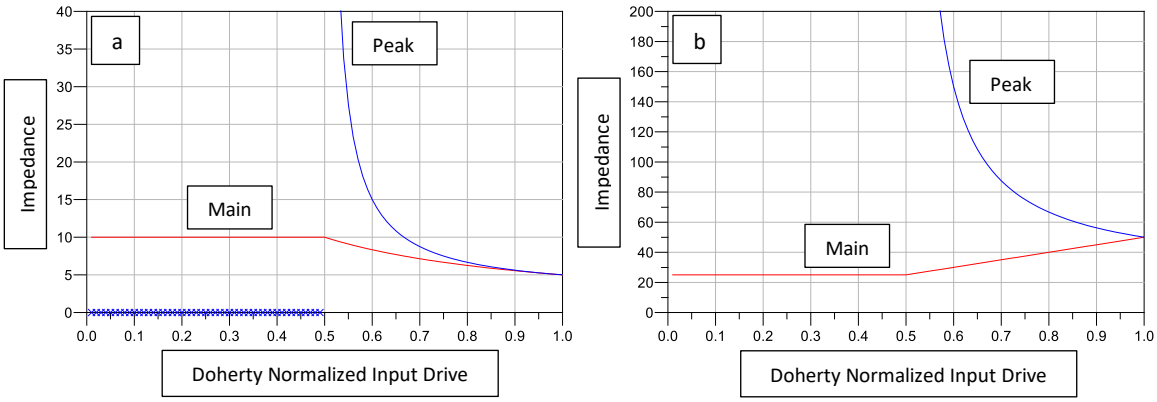

Figure 3.19 Dynamic impedance: (a) at amplifier reference plane, (b) at summing node reference plane

Section 3.2.1 presented the bandwidth comparison between three matching structures under fixed load and source impedance conditions, as would be the case between amplifier load and amplifier system impedance where no load modulation occurs. In this section, the bandwidth of these structures is analyzed in a load modulated environment, as would occur between both main and peaking amplifier loads and the effective system impedance. The impedance range from Figure 3.19 is used for bandwidth analysis, therefore assuming two 50Ω matched amplifier stages driving a 25Ω summing node, and each targeting a 5Ω Z_{opt} and 10Ω Z_{mod}.

3.1.2.1 Single Section Transformer

The simplest embodiment of the impedance inverter is implemented using a single section transmission line. The bandwidth of the inverter is proportional to the transformation ratio, and therefore under the conditions described in 3.1.2, increases in back-off as the transformation ratio decreases from 10:1 at full power (50Ω to 5Ω) to 2.5:1 at back off (25Ω to 10Ω). Figure 3.20 shows the change in inverter bandwidth over a 2:1 load modulation ratio.

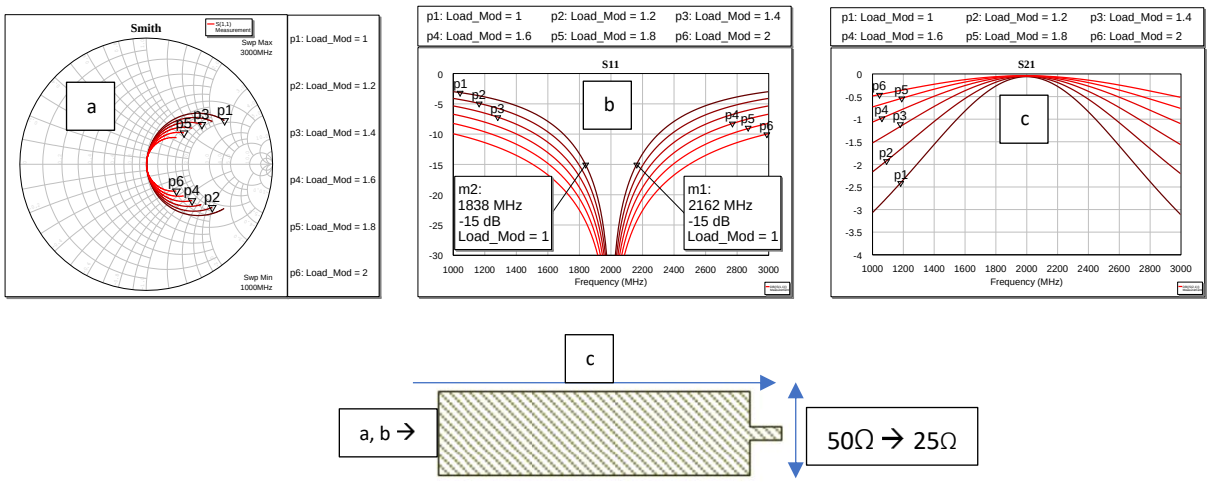

Figure 3.20 Behavior of single section matching transformer under load modulation: (a) normalized impedance, (b) return loss into load modulated impedance, (c) S21 into load modulated impedance.

It can be seen that a fractional bandwidth of 16.2% is achieved at full power, and 49.55% at 2:1dB load modulation with a -15dB return loss. Figure 3.21 shows the fractional bandwidth vs. load modulation.

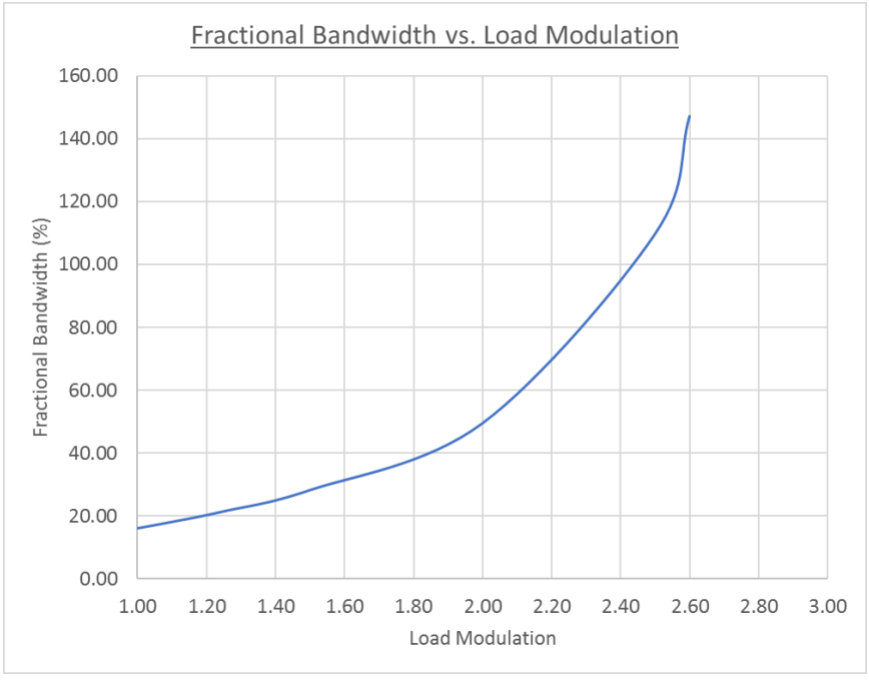

Figure 3.21 Fractional bandwidth vs. load modulation, single section impedance inverter.

Figure 3.21 shows that as the load modulation ratio increases, the fractional bandwidth of the main amplifier increases exponentially. The bandwidth increase is because the summing node impedance decreases as the load modulation ratio increases, thereby reducing the transformation ratio of the impedance inverter. Figure 3.22 plots the inverter transformation ratio vs. load modulation.

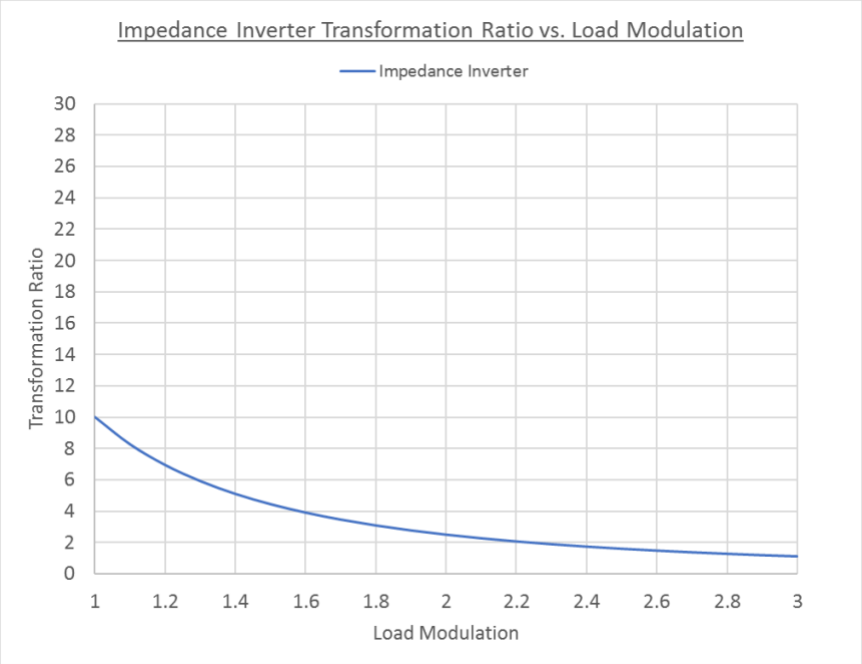

Figure 3.22 Transformation ratio, single section Doherty impedance inverter.

3.1.2.2 Two Section Transformer

The two-section transformer is typically implemented within an amplifier matching network to increase the bandwidth in a fixed impedance environment, as has been shown in 3.1.1. Figure 3.3 (a), and 3.4 (c) show the increase in bandwidth of the two-section transformer over the single section transformer. The single section transformer enables a bandwidth of 324MHz (16.2%) compared to the two-section transformer, which provides an increased bandwidth of 926MHz (46.3%).

When discussing the bandwidth of the two-section transformer structure in the output of a Doherty under load modulation, the concept of an 'inverting' or 'non-inverting' matching network requires consideration. These terms have historically been used loosely in describing the Doherty topology, with a minimal formal definition. Any Doherty amplifier output matching structure must be an inverting topology, since the current injection from the peaking amplifier into the Doherty output summing node causes an increase of the summing node impedance at high power levels, as shown in 3.19 (b). Therefore, an impedance inversion between the summing node and the main amplifier current generator is necessary in order for the main amplifier to deliver full current into its optimum load (Z_{opt}), and deliver high efficiency under load modulation. Figure 3.23 shows the mismatch introduced due to load modulated of the two-section transformer without the inclusion of the additional impedance inverter since the impedance increase at the current generator is proportional to that of the Doherty output summing node. (3.10) and (3.11) further illustrate the non-inverting characteristic where the load impedances developed by the matching network are derived in the example of transforming a 50Ω load impedance to a 5Ω transistor-level impedance.

$$Z_{intermediate\_full\_power} = \sqrt{Z_{DUT} \times Z_{Load}} = \sqrt{5 \times 50} = 15.811\ \Omega \tag{3.7}$$

$$Z_{first\_section} = \sqrt{Z_{intermediate\_full\_power} \times Z_{Load}} = \sqrt{15.811 \times 50} = 28.12\ \Omega \tag{3.8}$$

$$Z_{second\_section} = \sqrt{Z_{intermediate\_full\_power} \times Z_{DUT}} = \sqrt{15.811 \times 5} = 8.89\ \Omega \tag{3.9}$$

$$Z_{intermediate\_backoff} = \frac{{Z_{first\_section}}^2}{Z_{load\_backoff}} = \frac{28.117^2}{25} = 31.62\ \Omega \tag{3.10}$$

$$Z_{DUT\_backoff} = \frac{Z_{second\_section}{}^{2}}{Z_{intermediate\_backoff}} = \frac{8.89^2}{31.62} = 2.5\ \Omega \tag{3.11}$$

As is shown in (3.11), the two-section matching network presents a 'non-inverting' impedance relationship between summing node impedance and device-level impedance. This non-inverting relationship is seen by the reduction in the impedance at the back-off power levels derived in (3.11) from a 5Ω Z_{opt} at full power to a 2.5Ω Z_{mod} at back-off power levels. This impedance relationship is the inverse of the required Doherty amplifier load modulation characteristics, and is due to the 180° electrical length of the structure, compared to the 90° electrical length of the single section match. Figure 3.23 shows the non-inverting behavior, whereby a mismatch is introduced in all load modulated load conditions.

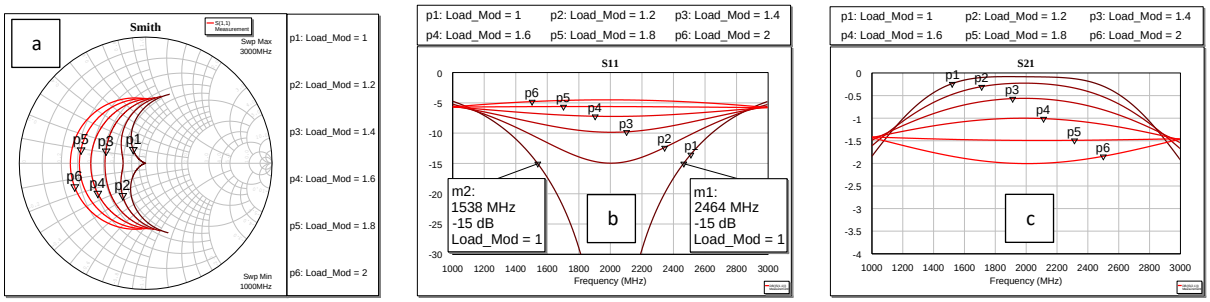

Figure 3.23 Behavior of two section matching transformer with load modulation: (a) normalized load impedance, (b) return loss into load modulated impedance, (c) S21 into load modulated impedance.

In order to provide correct load modulation, the addition of a further impedance inverting section is necessary in order to provide impedance inversion to a matching network with non-inverting properties. The impedance inverter may take the form of a 90° transmission line between the summing node and the main amplifier matching network. The resulting, modified impedance at the current generator reference plane is derived in 3.12 – 3.14, where the inverting relationship

is seen by the increase in the impedance at the back-off power levels derived from a 5Ω Z_{opt} at full power to a 10Ω Z_{mod} at back-off power levels. This impedance relationship is the required Doherty amplifier load modulation characteristics.

$$Z_{presented\_to\_match} = \frac{Z_{TLine\_sum\_to\_match}{}^{2}}{Z_{summing\_node\_backoff}} = \frac{50^2}{25} = 100\,\Omega \tag{3.12}$$

$$Z_{intermediate\_backoff} = \frac{Z_{first\_section}{}^{2}}{Z_{load\_backoff}} = \frac{28.117^2}{100} = 7.91\,\Omega \tag{3.13}$$

$$Z_{DUT\_backoff} = \frac{Z_{second\_section}{}^{2}}{Z_{intermediate\_backoff}} = \frac{8.89^2}{7.91} = 10\,\Omega \tag{3.14}$$

Figure 3.24 shows how the target current generator impedance matches the target device impedance across frequency and with load modulation of the Doherty amplifier summing node due to the peaking amplifier current. This impedance matching vs. Doherty summing node impedance illustrates that the load modulation at the current generator is correctly aligned with the addition of the impedance inverter. In this case, the device level impedance decreases as the current from the peaking amplifier increases, causes the effective summing node impedance to increase at high power levels.

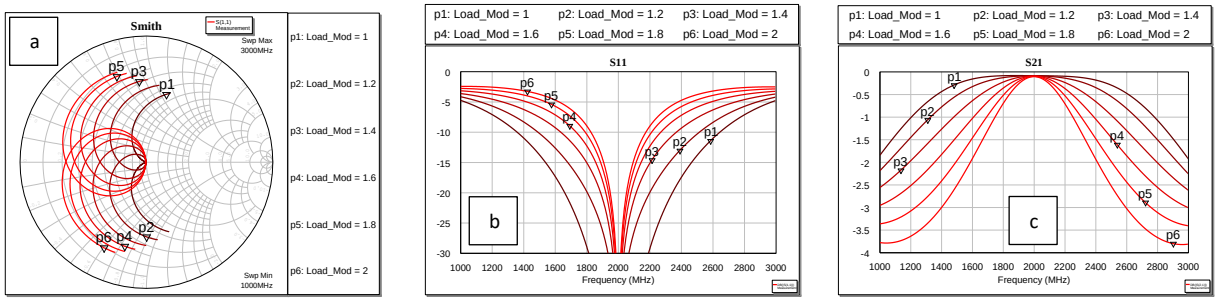

Figure 3.24 Behavior of two section matching transformer with addition of impedance inverter with load modulation: (a) normalized load impedance, (b) return loss into load modulated impedance, (c) S21 into load modulated impedance.

Figure 3.24 shows the results of the current generator matching vs. frequency and load modulation, demonstrating a wide RF bandwidth of 46.3% at full power, but a reduced bandwidth of 9.3% at 6dB OBO at 2:1 load modulation. Figure 3.25 plots the fractional bandwidth vs. load modulation.

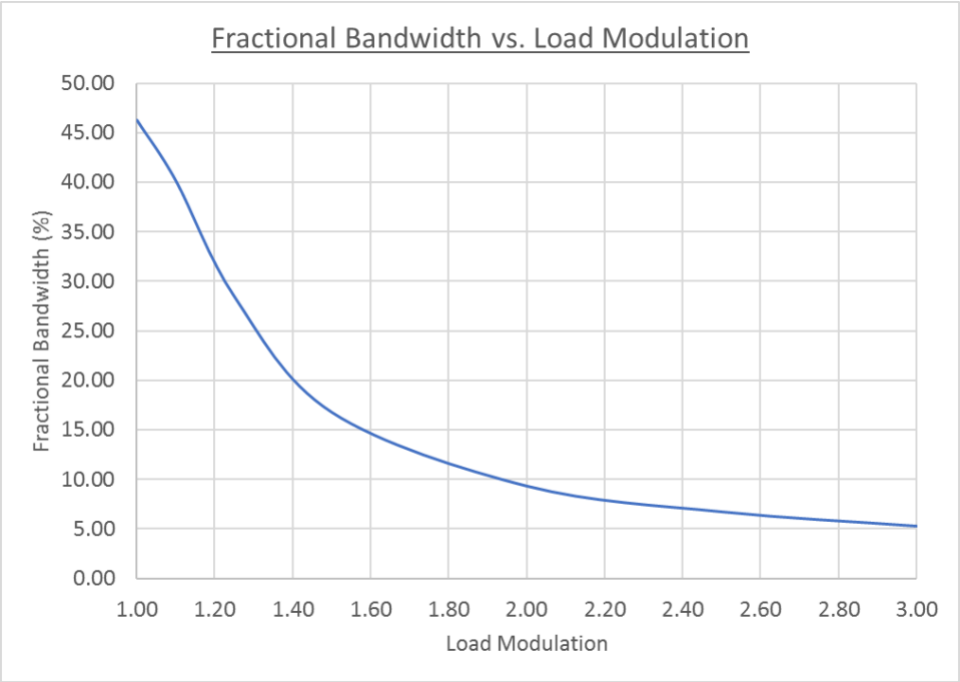

Figure 3.25 Fractional bandwidth vs. load modulation, two section matching network with impedance inverter.

Figure 3.25 shows that the two-section matching network provides a wide fractional bandwidth when operating into the fixed, intended circuit impedances, with a maximum transformation ratio of 3.16 per matching section. However, under load modulation, the increased load impedance presented to the amplifier matching circuit causes an increase in the transformation ratio of the first matching section by the square of the load modulation. Therefore, in a popular 2:1 asymmetric Doherty amplifier, with a 3:1 load modulation presented to the main amplifier, the transformation ratio of the first matching section is increased from 3.16 to 28.44. The transformation ratios through the matching circuit, from the summing node to the transistor, are shown in figure 3.26.

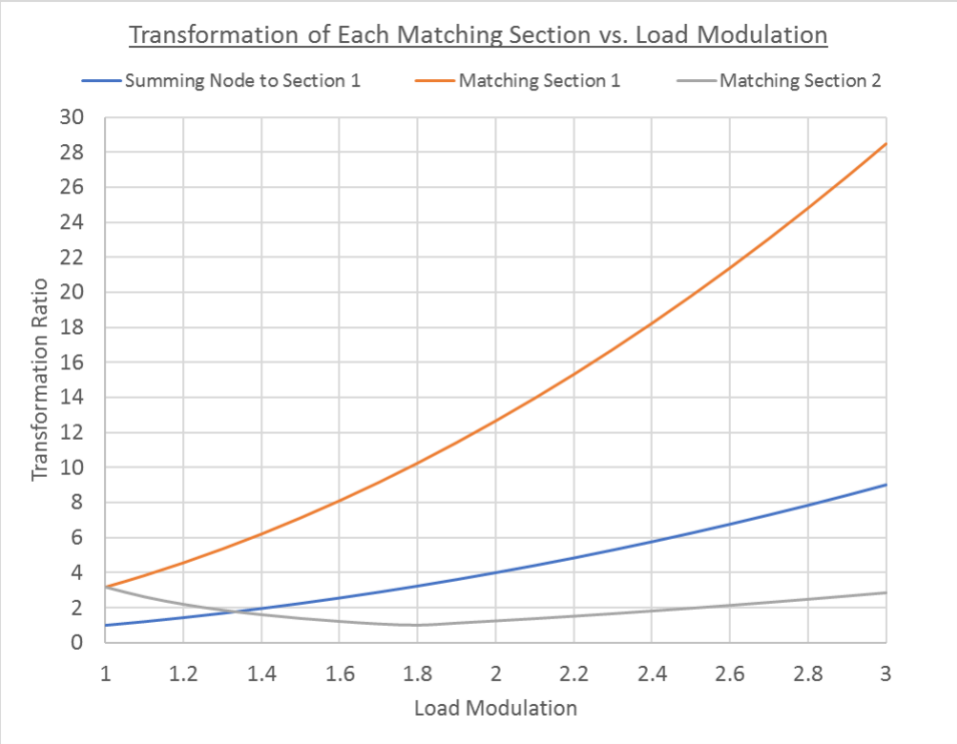

Figure 3.26 Transformation ratio per matching element, two section matching network plus impedance inverter.

3.1.2.3 Klopfenstein Taper

The characteristics of the Klopfenstein taper in a fixed load environment are described in detail in section 3.1.1. In this section, the suitability of the Klopfenstein taper for operation in a load modulated environment is analyzed. Unlike the single or multi-section matching networks, described in sections 3.1.2.1 and 3.1.2.2, comprising of a number of 90° matching sections, the Klopfenstein taper forms a progressive impedance transformation along the length of the structure, as shown in figure 3.10. The electrical length of the taper determines the lowest frequency of operation, and the characteristic impedance vs. length determines the passband ripple. Therefore, the taper can function as an effective impedance transformer at any frequency above the low-frequency cutoff, whereas the single and multi-section transformer provides correct transformation at multiples of 90°.

Therefore, in order to provide correct operation in a load modulated environment, the tapered design necessitates consideration of the phase of the load modulation, in addition to low-frequency cutoff and passband ripple. In the simplest implementation, this can be absorbed into the electrical length of the taper, thus avoiding the need for additional phasing lines at either the source or load termination. The response of the taper discussed in section 3.1.1, designed without consideration for the load modulated phase, is shown in figure 3.27.

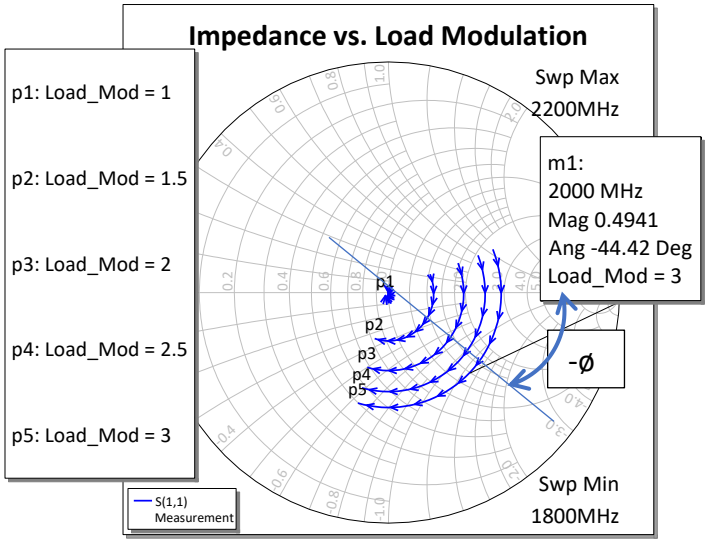

Figure 3.27 Klopfenstein taper load impedance vs. source impedance load modulation.

In this case, the design is not ideal for a load modulated amplifier since a non-realizable -22.2° phasing line is required to align the load modulated impedances correctly, as labelled -ø in figure 3.27. Therefore, the options to enable the structure to operate under load modulation correctly are: -

- Reduce the taper electrical length, therefore increasing the low frequency cut off frequency, shown in 3.28 (a)
- Increase the taper electrical length, therefore reducing the low-frequency cutoff, but increasing both size and loss of the structure, as shown in figure 3.28 (b)
- Add a phasing line at either load, figure 3.28 (c) or source, figure 3.28 (d) to ensure real - real load modulated impedances

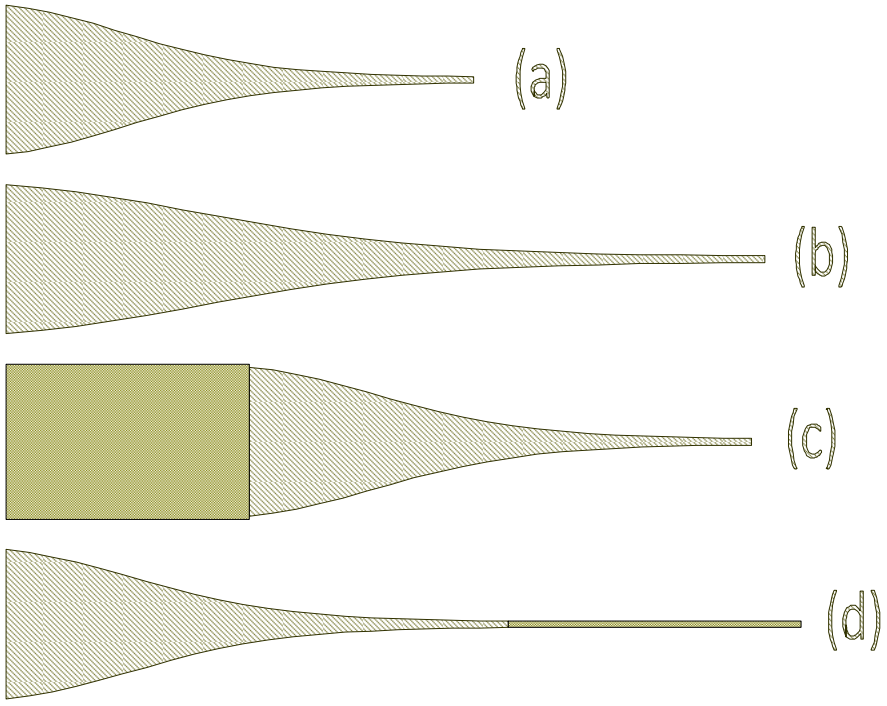

Figure 3.28 Physical implementations of Klopfenstein tapers with correct load modulation.

Figures 3.29 (a-d) show the resultant impedance of the tapers in figure 3.28 vs. source load modulation. In contrast to the results shown in figure 3.27, the load impedance modulation occurs along the real axis in response to the source load modulation of the summing junction.

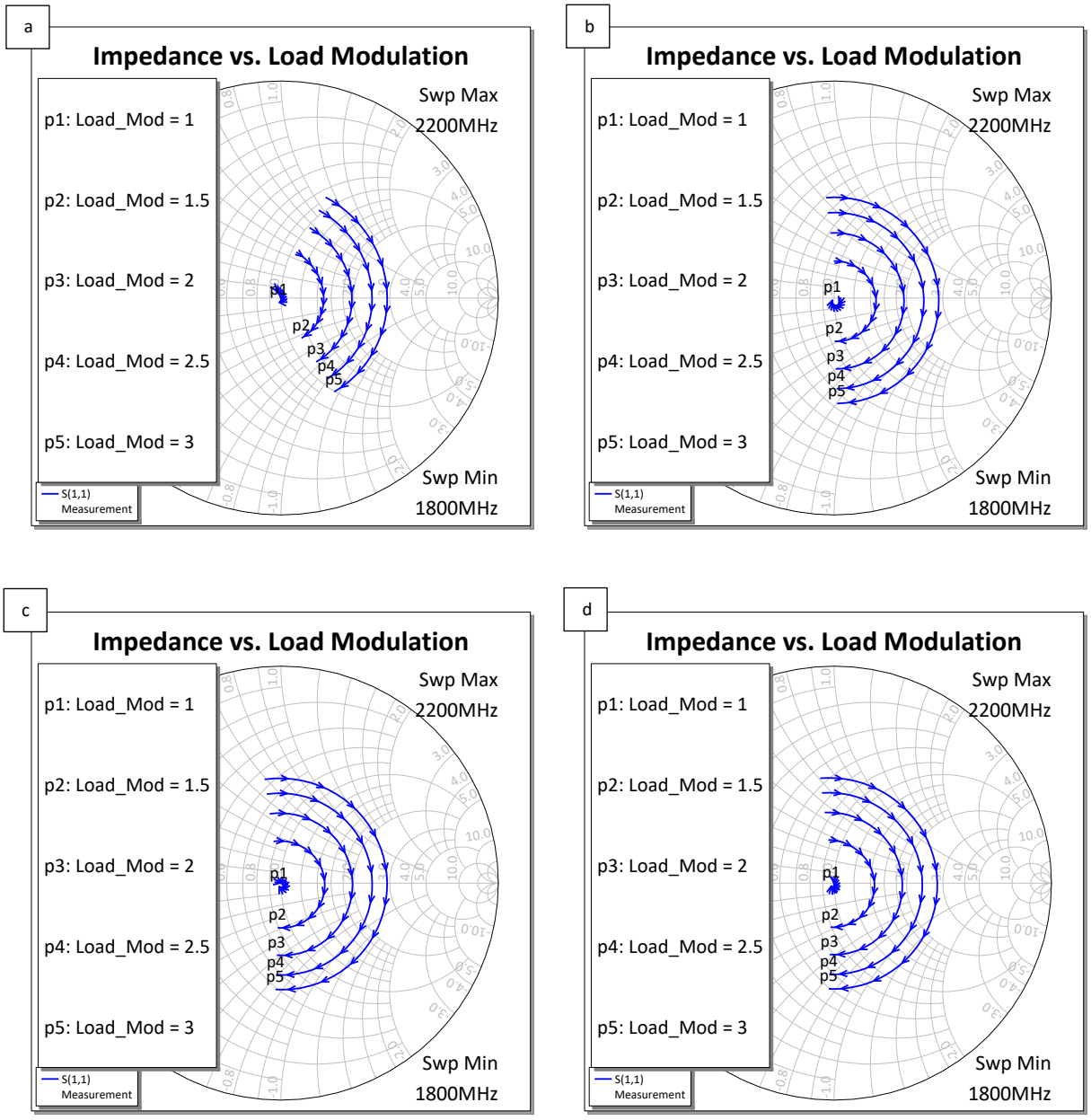

Figure 3.29 Impedance vs. load modulation behavior of Klopfenstein tapers from figure 3.28, after optimization for load modulation.

Figures 3.29 and 3.30 show the extra impedance dispersion, and therefore bandwidth reduction, incurred with load modulation for tapers (c) and (d) due to the addition of the electrical length required for correct load modulation phase characteristics. By contrast, taper (a), designed for the

correct load modulation phase at 2GHz, exhibits reduced dispersion and increased bandwidth, and a smaller physical size.

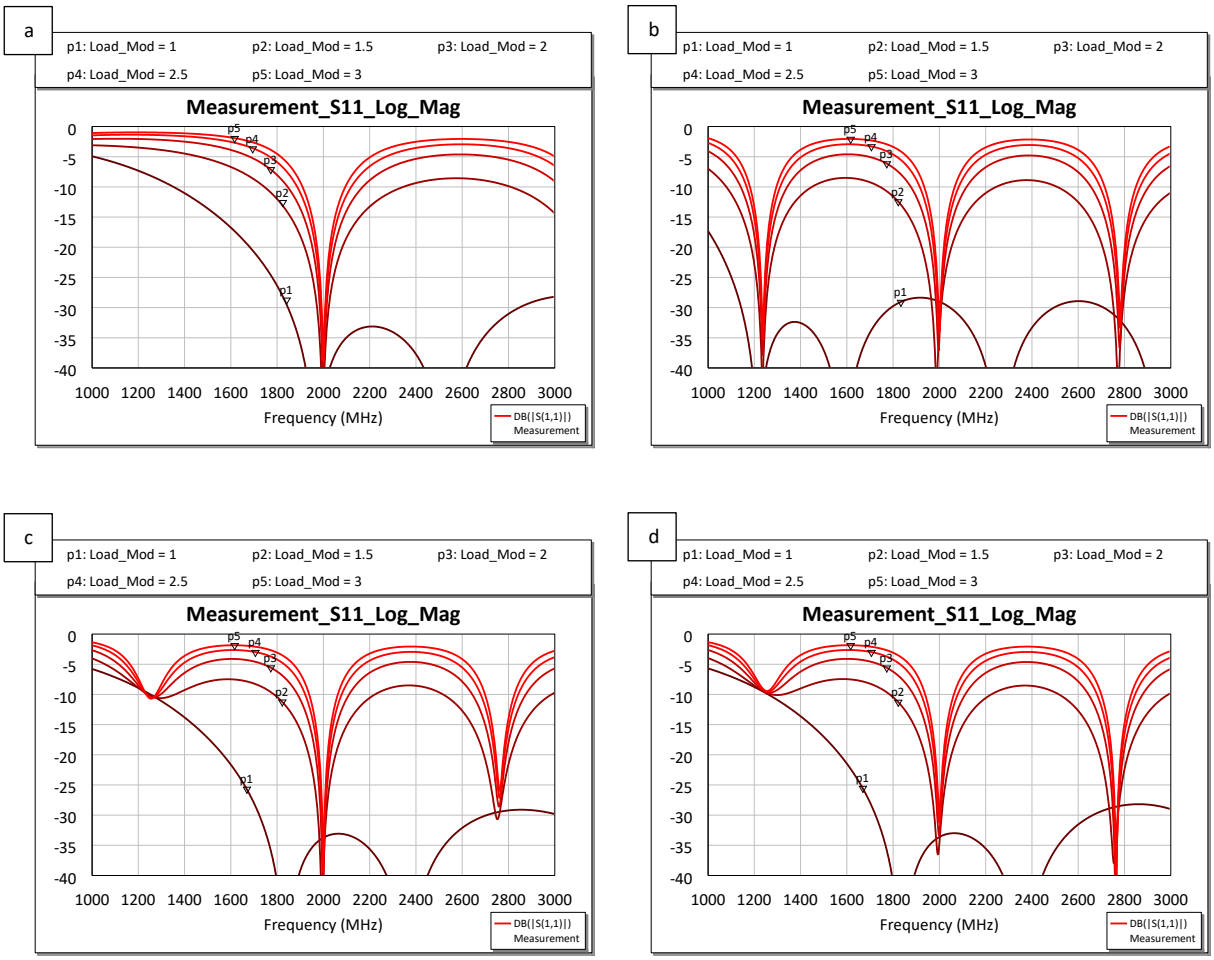

Figure 3.30 Bandwidth of Klopfenstein tapers from figure 3.28, after optimization for load modulation

Figure 3.31 shows a summary of the Klopfenstein taper performance after optimization for load modulation. From this, two conclusions can be drawn: -

- The taper with shortest electrical length, while still achieving the desired low-frequency cutoff, will provide optimal bandwidth under load modulation.

- The addition of a source or load phasing line to a non-optimal taper gives rise to sub-optimal performance in comparison to a taper optimized for real-to-real impedance transformation under load modulation.

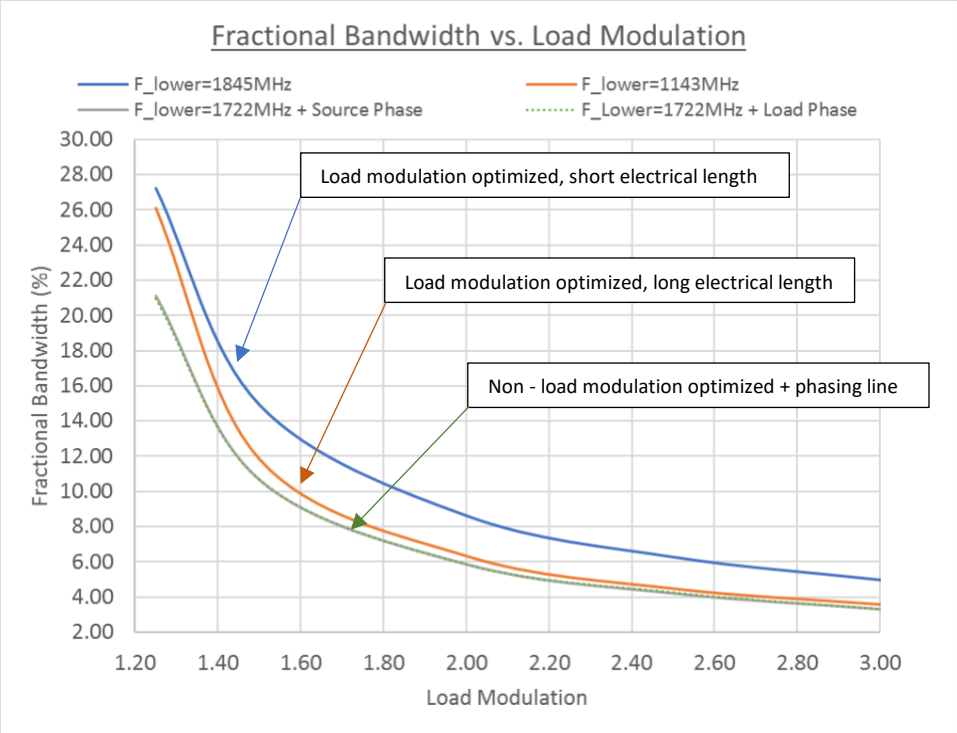

Figure 3.31 Fractional bandwidth vs. load modulation, Klopfenstein impedance inverter.

The load modulation axis in figure 1.2 does not extend down to 1.0 (no-load modulation) since, under these conditions, all the Klopfenstein structures exhibit a high pass response, limited only by the low frequency cut off frequency. Therefore, since there is no high-frequency limit – the bandwidth of all the structures tends to infinity.

3.1.2.4 Summary

In this section, the analysis of the matching networks presented in 3.1.1 is extended to consider the response of these networks under load modulation, as would be seen by the power amplifier matching networks. The results highlight the difference in the requirements between the

summing node transformer, discussed in 3.1.1, operating in a static impedance environment, and the amplifier matching network/impedance inverter, discussed in this section, while operating in a load modulated environment. The differences are highlighted in the suitability of the structures under different conditions. In 3.1.1, it was shown that the Klopfenstein taper enables the potential for a virtually infinite bandwidth summing node, limited only by the size of the taper and the associated insertion loss. Therefore, it is possible to realize an ultra-wideband summing node in this way.

However, when the additional complexity of load modulation is included, as would need consideration for the amplifier matching networks, the suitability of the networks is reversed, and the simple, single section transformer enables excellent flexibility to provide a wideband response through careful selection of the summing node impedance. As is shown in figure 3.21, the bandwidth of the single section inverter increases with load modulation. This bandwidth increase is because the reducing system impedance, and corresponding increasing device load impedance, both tend toward the characteristic impedance of the inverter, thereby reducing the transformation ratio, and increasing the bandwidth. With correct summing node impedance selection, this characteristic enables a combiner design with no transformation occurring in the back off region, and when used in conjunction with a Klopfenstein taper to set the summing node, an infinite bandwidth is achieved at the back off power [6].

By contrast, the multi-section structures achieve wideband performance operating into the target load impedances through minimizing the transformation ratio through the network. However, when used in a varying load environment, the idealized transformation ratios are no longer valid, and the network results in a cascade of high transformation ratio structures. This is highlighted in figure 3.26, showing the increase in the first transformation from 3.16 to 28.44, with the corresponding bandwidth collapse shown in figure 3.25. In addition, the 180-degree electrical length of the two-section match does not provide impedance inversion, and therefore an additional impedance inverter is required to ensure correct load modulation at the transistor reference plane.

Finally, while the Klopfenstein taper provides a highly suitable method for defining the fixed summing node impedance, in order to operate correctly in a load modulated environment, the

structure requires a third parameter to be considered in the design in addition to the low-frequency cutoff and the passband return loss. This parameter is the load phase trajectory under source impedance modulation (shown in 3.27 and 3.29). The performance of the taper is improved if the load modulation behavior is considered in the design of the taper, rather than resorting to the source or load phasing lines shown in figure 3.31. However, even after optimization, the performance under load modulation of the Klopfenstein taper does not achieve the same bandwidth as a single impedance inverter with optimal selection of system impedance, achieved through the use of a wideband impedance matching structure such as a Klopfenstein taper.

3.1.3 Device Drain to Source Capacitance

In this section, the impact of the transistor drain-to-source capacitance (C_{ds}) on the bandwidth of the load modulated amplifier is investigated. While it is true that the elimination of the C_{ds} would remove the bandwidth limitations at the transistor output, the achievable bandwidth with a finite C_{ds} is highly dependent on the loaded Q of the network presented to the device. The calculation for loaded Q is shown in 3.15, showing the increases in Q with the load impedance (R_L), a function of the transistor operating voltage, and the load modulation applied.

$$Q_{Loaded} = w_0 R_L C_{ds} \tag{3.15}$$

Where R_L is the load impedance presented to the device and C_{ds} is the transistor drain to source capacitance.

3.1.3.1 C_{ds} Impact vs. Transistor Operating Voltage

The output power of the transistor is determined by the available current swing (I_{max}) and the voltage swing (V_{swing}), and is calculated in 3.16.

$$P_{out} = \frac{I_{max}.2.(V_{dd} - V_{knee})}{8} \quad (3.16)$$

Where V_{dd} is the drain supply voltage to the transistor, and V_{knee} is the knee voltage for the device which reduces the assumed RF voltage swing. It can be seen from (3.16) that for a fixed Pout, the voltage and current are inversely proportional. Therefore, a fixed output power is achievable from:

-

- A low voltage technology, operating into a low impedance load at a high current.
- A high voltage technology, operating into a high impedance load at a low current.

Traditional thinking tells us that for high power devices, a higher voltage technology enables increased amplifier bandwidth since less transformation ratio is required to achieve a typically 50Ω load impedance. However, it has been shown in 3.1.1 that a wideband transformation to low impedance is achievable, and the Q of the device's C_{ds} should be considered into these load impedances. Figure 3.32 show plots of calculated bandwidth vs. drain supply voltage for a fixed RF output power of 100 watts, assuming a constant parasitic capacitance of 0.3pF/W, and operating into a 2:1 load modulation (symmetric Doherty).

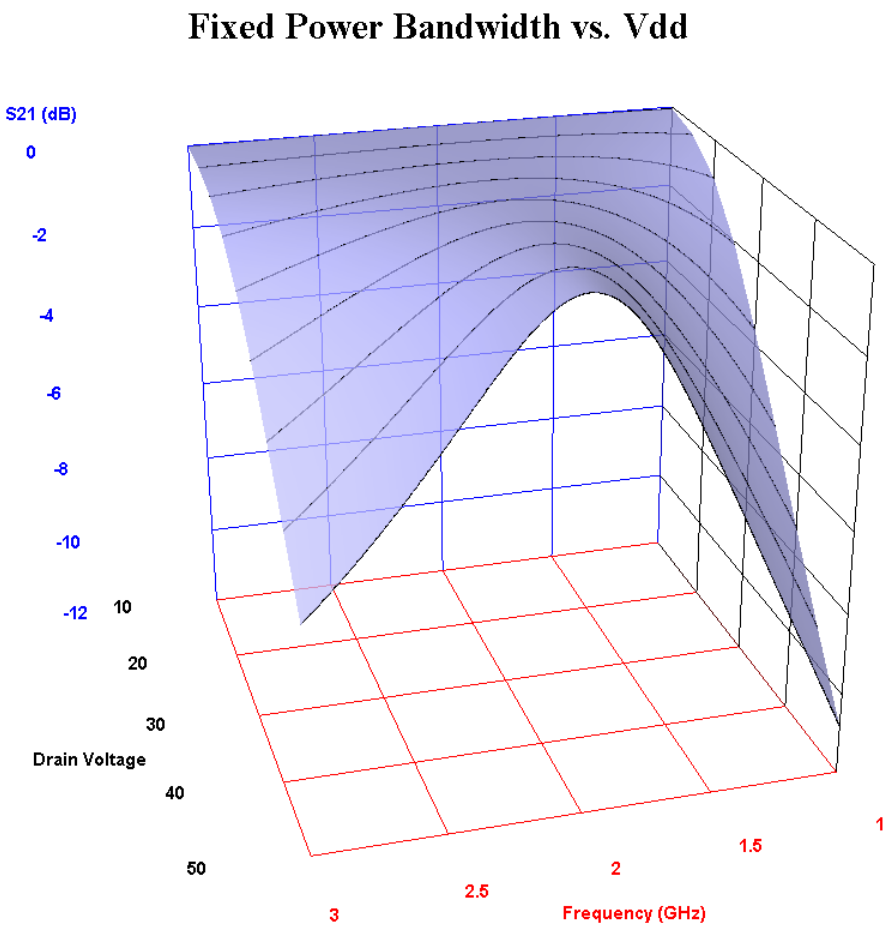

Figure 3.32 Device bandwidth for fixed C_{ds} to P_{out} ratio vs. drain voltage

It can be seen that as the drain voltage increases, the bandwidth of the output parallel network (C_{ds} // $L_{internal\_match}$) decreases. This is because the higher drain voltage requires a higher load impedance, which, as shown in 3.15, causes an increase in Q factor of the output network. Therefore, a lower voltage device technology will enable wider bandwidth, assuming the parasitic capacitance per watt is comparable, and the matching impedances are realistic to achieve. In order to compare two technology types, the above analysis should be run for the specific C_{ds}/W inherent in each technology, in order to determine an accurate bandwidth comparison.

3.1.3.2 - C_{ds} Impact vs. Transistor Load Modulation

In this analysis, the impact of the Doherty load modulation is demonstrated for a fixed technology operating at a drain voltage of 30v, and assuming 100w device with 0.3pF/W of parasitic C_{ds}. These figures are typical of a silicon LDMOS transistor at the time of writing.

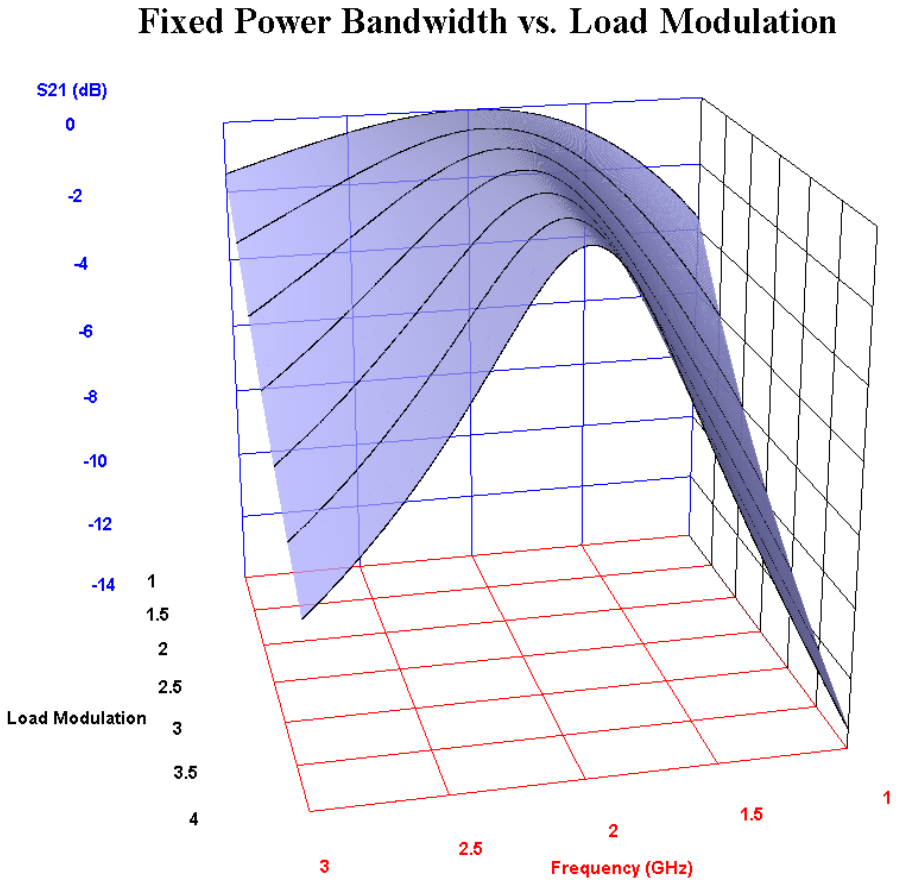

Figure 3.33 Device bandwidth with fixed C_{ds} vs. load modulation

(Peak P_{out}=100W, C_{ds} = 0.3pF/W, V_{dd} = 30v)

It can be seen that the transistor bandwidth and load modulation are inversely proportional since the loaded Q of the C_{ds} and the internal shunt match increase with the load modulation ratio. In this case, achieving wideband performance becomes more difficult as the Doherty asymmetry

ratio is increased, since the load modulation ratio increases with asymmetry, as shown in chapter five, figure 5.12.

3.1.3.3 - Impact of C_{ds} per Unit Power (C_{ds}/W)

Assuming all other parameters equal, a reduction in the device C_{ds} enables a proportional increase in the bandwidth of the device. This can be seen in figure 3.34, whereby the bandwidth is calculated based on a 100W device operating at 30v V_{dd} into a 2:1 Doherty load modulation. It can be seen that an increase in device C_{ds} on the y-axis gives rise to a progressive reduction in bandwidth from the device.

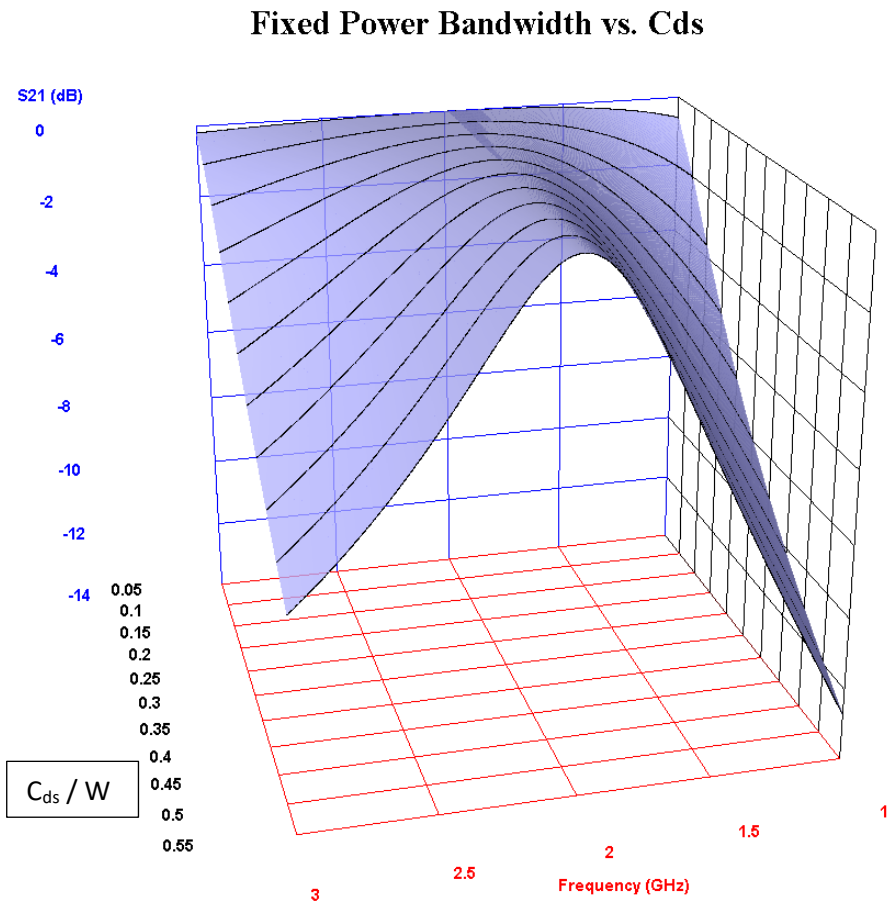

Figure 3.34 Device bandwidth vs. C_{ds} per watt.

(100W, 30v V_{dd}, 2:1 Load Modulation)

3.1.4 Peaking Amplifier Off State Impedance

At all power levels below the peaking amplifier turn-on threshold, the peaking amplifier and corresponding circuity form a parasitic load at the summing node impedance. At center frequency, the peaking amplifier output impedance will be phased in order to present a high impedance at the main and peak summing junction, as depicted by the 'peak off-state phasing' in figure 3.35.

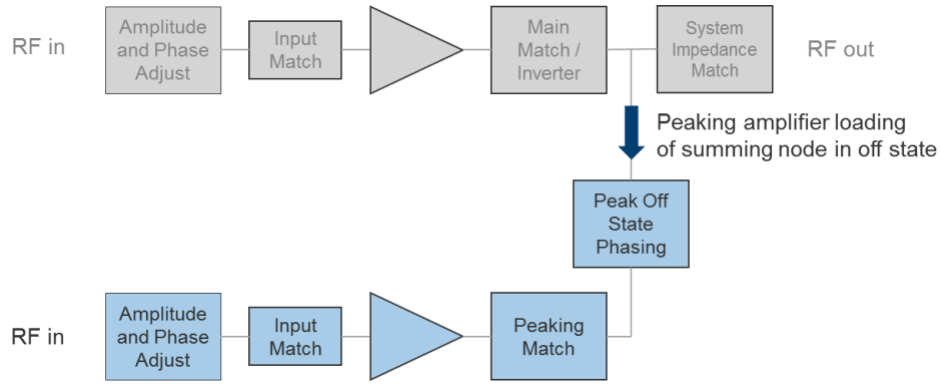

Figure 3.35 - Peaking amplifier off state.

The impact of the peaking device parasitic loading effect, in this case, is a shunt loss at the summing node, whereby power dissipation will occur in the parallel off-state resistance. Before considering the bandwidth impact of the peaking amplifier, an analysis is run to consider the ratio between the summing node impedance, and the off-state impedance presented by the peaking amplifier at the Doherty summing node. This ratio determines the peaking off-state losses in the Doherty since the peaking device loading appears as a parallel loss at the summing node.

Given that the Doherty system impedance and the amplifier matching impedance are variables, an analysis is run to examine the impact of these effects. For example, the symmetric Doherty may be designed using two amplifiers, each matched to 50Ω and driving a system impedance of 25Ω. Alternatively, the Doherty can be designed using two amplifiers matched to 10Ω, and driving a 5Ω system impedance. In the former case, the peaking amplifier output matching network transforms the high intrinsic off state of the transistor through the matching network, designed to

transform 50Ω to Z_{opt}, and presents this impedance in parallel with the 25Ω summing node. In the latter case, the peaking amplifier matching network transforms the same off-state impedance through the matching network designed to transform 10Ω to Z_{opt}, and presents this impedance in parallel with the 5Ω summing node. The parallel loss in the system is, therefore, a function of ratio between off-state impedance and summing node impedance, characterized in figure 3.36.

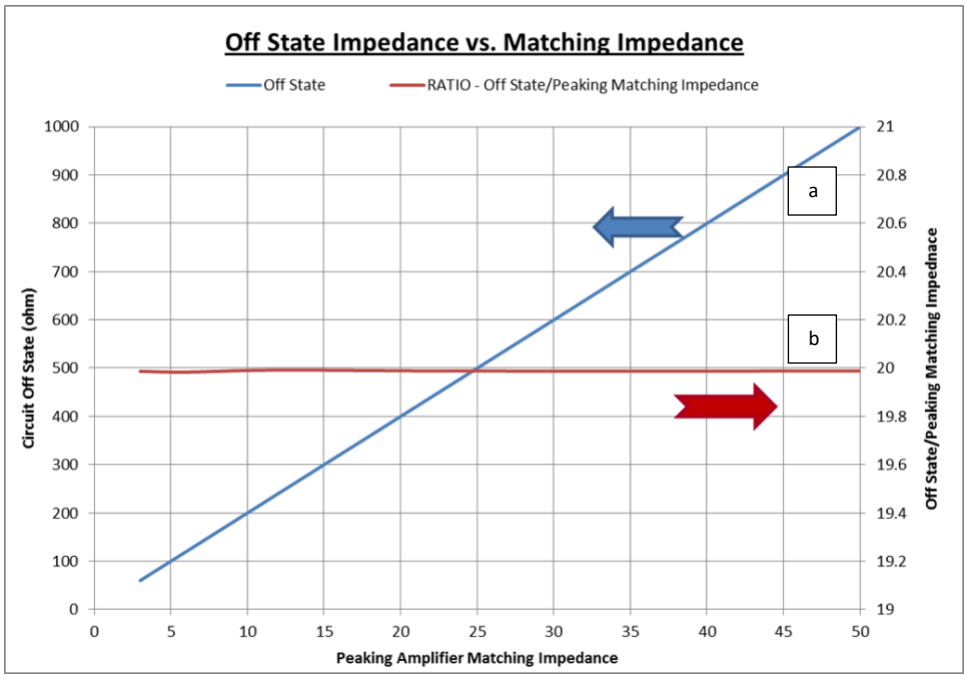

Figure 3.36 (a) off state impedance vs. device matching Impedance, (b) ratio of off-state impedance to Doherty summing node impedance

Figure 3.36 (a) shows that the off-state impedance presented by the peaking amplifier at the Doherty summing node is reduced as the Doherty summing node, and therefore amplifier matching impedance, is reduced. However, the ratio between the two, shown in figure 3.36 (b), remains constant, therefore no reduction in shunt loss is seen through the manipulation of the Doherty summing node impedance and amplifier matching impedance.

If the same analysis is now run vs. the Doherty asymmetry, it can be seen that while the summing node: off-state ratio is also independent of matching impedance, the peak off-state to summing node impedance is reduced in comparison to the symmetric case. Therefore, higher shunt losses

will be inherent in the asymmetric Doherty in comparison to the symmetric case. Figure 3.37 shows the off-state ratio normalized to the symmetric case.

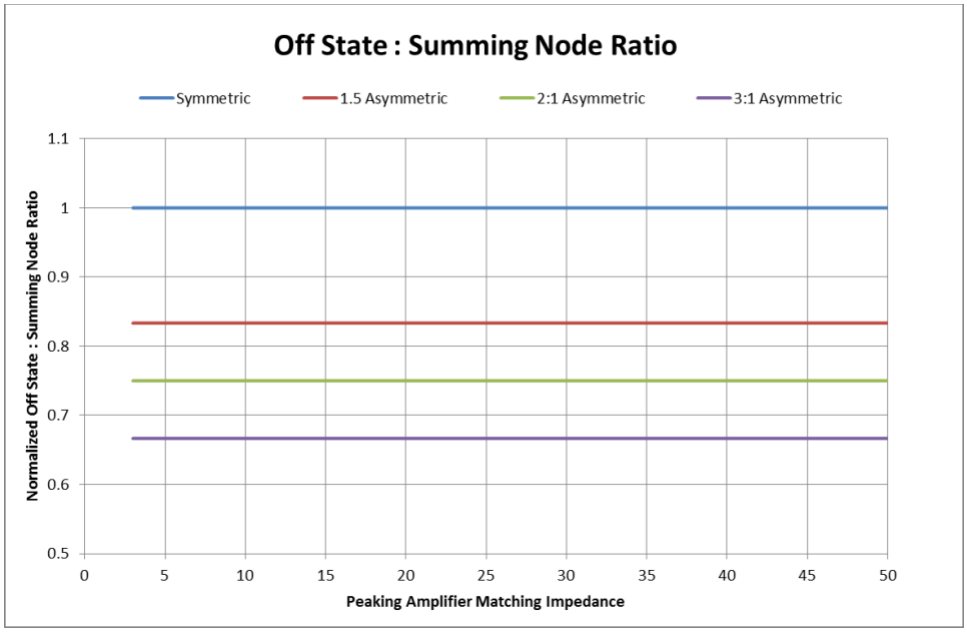

Figure 3.37 Off state to summing node impedance ratio vs. matching impedance and Doherty asymmetry.

As the frequency deviates from the target center frequency, the off-state impedance will appear as a reactive loading element at the summing node of the Doherty. Figure 3.38 plots the peaking off-state phase angle for a 100W transistor with a C_{ds} of 30pF, operating at a drain voltage of 30V into a 2:1 load modulation. These characteristics are typical of a silicon LDMOS RF power transistor at the time of writing.

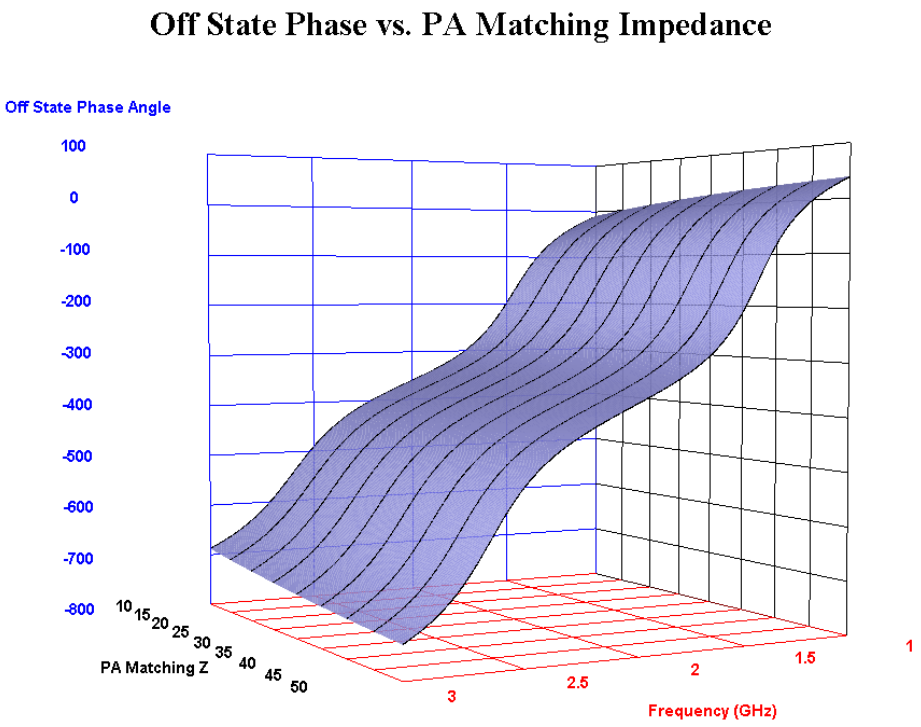

Figure 3.38 Off state phase angle vs. PA matching impedance.

Figure 3.38 shows that the selection of the Doherty amplifier system impedance, and therefore the impedance to which the amplifier is matched, have minimal impact on the off-state loading at the Doherty summing node due to the peaking amplifier.

In the above conditions, the peaking amplifier requires an impedance matching network to transform the Doherty summing node impedance to the optimal load impedance of the peaking transistor, and provide the correct high impedance phase when viewed from the summing node. In order to satisfy the latter case, the peaking matching network must be constructed of n*180° sections in order to rotate the phase of the high off-state impedance at the transistor reference to present a high impedance when viewed from the Doherty summing junction, indicated in figure 3.35, labelled 'Peaking amplifier loading of summing node in off state'.

The traditional Doherty amplifier is designed with a value of n=1. Therefore, a phase of 180° is required between the current generator and the Doherty summing node. However, the requirements for correct Doherty operation are met with any value of n, and therefore the

amplifier may be designed with a system impedance selected to enable n=0, a technique first proposed by this author in [1]. The system impedance (Z_{sys}) for n=0 is calculated in 3.17

$$Zsys_{n=0} = \frac{Z_{opt_{main}}}{\left(1 + \frac{I_{peak}}{I_{main}}\right)} \tag{3.17}$$

It can be seen from 3.17 that a system impedance less than the device Z_{opt} is required for the correct operation of the peaking n=0 Doherty amplifier. This is because the 'effective' summing node impedance seen by each amplifier is greater than the static system impedance due to the active load pulling to a higher impedance caused by the current from the other device.

Figure 3.39 shows the comparison of the high impedance phase angle presented at the Doherty summing node for the device characterized in figure 3.38 for both the n=0 and n=1 Doherty.

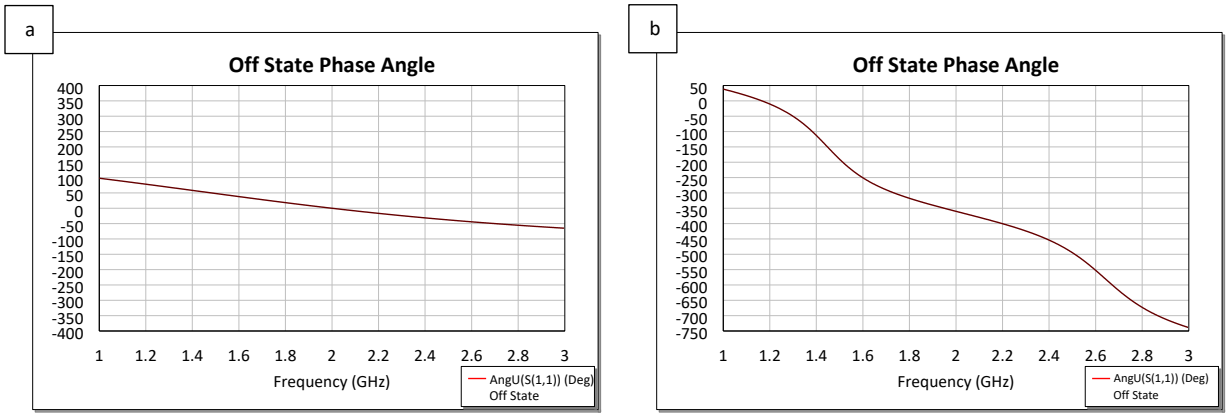

Figure 3.39 Off state phase angle at Doherty summing node, (a) n=0, (b) n=1.

It can be seen that the off-state phase angle is reduced from 777° (n=1) to 163° (n=0) across a 2GHz sweep. The corresponding impact on the load modulation presented to the main amplifier in back off is shown across a 1GHz bandwidth in figure 3.40.

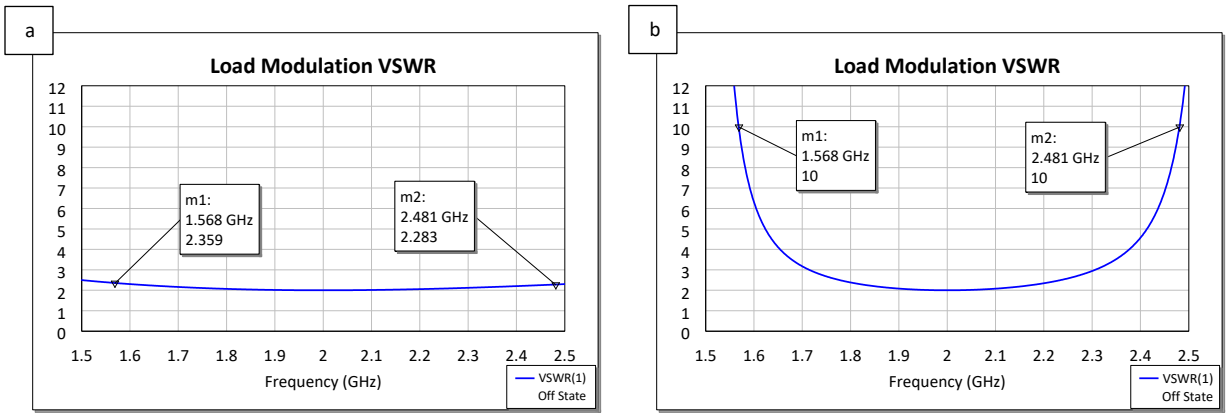

Figure 3.40 Load modulation at summing node, (a) n=0, (b) n=1

Figure 3.40 demonstrates the importance of this finding for the design of the wideband Doherty amplifier. In the case of the n=0 Doherty, the load pulling effect of the peaking amplifier on the main amplifier in back-off is relatively small – with the load modulation increasing from 2:1 at the center frequency to 2:5:1 at the extremes of a 1GHz bandwidth sweep. This effect would raise the impedance presented to the main amplifier in back-off, thus reducing the peak power available from the main amplifier in back-off. However, under the same conditions, the impact of the peak amplifier loading of the n=1 Doherty gives rise to an exponential increase in the main amplifier load modulation out of band, quickly forcing the main amplifier into a high impedance load, where it is unable to deliver the current required to achieve the target Pout. The reason for the exponential increase in load modulation observed in figure 3.40 (b) can be understood when looking at the phase angle plots in 3.39 (b), whereby the off-state phase is rotated though 360 - 180 degrees at 1.5GHz and 360 + 180 degrees at 2.6GHz. Therefore, in both cases, the open circuit presented at center frequency is rotated to present a short at these frequencies, hence the high VSWR seen at the Doherty summing node.

If the impact of the device C_{ds} is now considered, it can be seen in figure 3.41 that the use of device technology with low parasitic C_{ds}/W in conjunction with the n=0 Doherty topology negates a limitation in the Doherty architectures. This is the shorting of the Doherty summing node caused by the peaking amplifier off state impedance rotating to a low impedance. It can be seen in figure 3.41 (a) that a device with 0.05pF/W (marker P1), a typical value for GaN HEMT devices at the time

of writing, will not cause any significant peak off-state loading across a 50% fractional bandwidth. By contrast, the loading effect of the same technology used in the n=1 Doherty topology will increase the load modulation at the Doherty summing node from 2:1 to 3:1 across the same bandwidth. While this loading will not cause the non-functionality as would be the case with the higher C_{ds} technology shown in 3.40 (b), the main amplifier will suffer from a reduction in peak power in the back-off region, giving rise to significant AM-AM distortion around the back-off peak prior to peaking amplifier conduction.

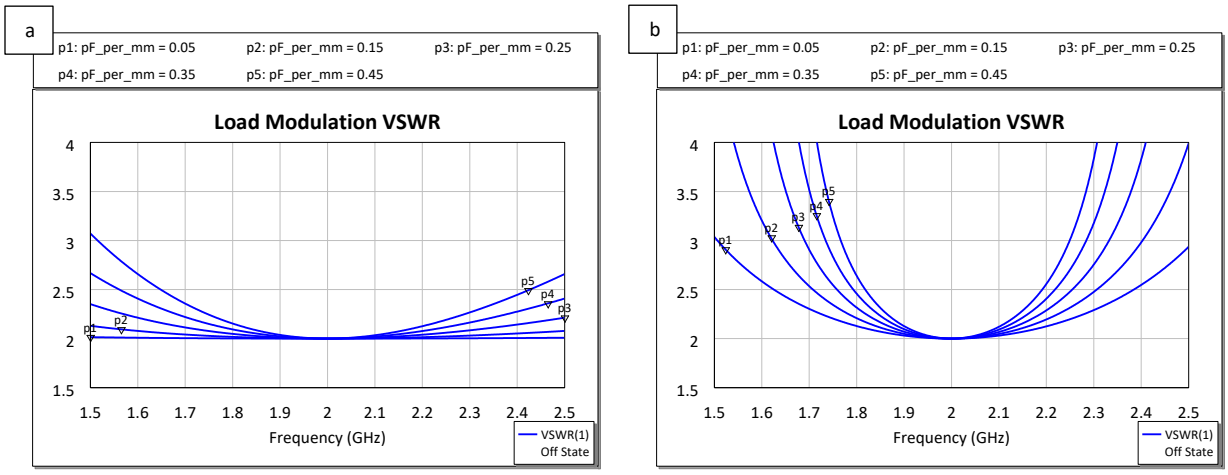

Figure 3.41 Load modulation at summing node vs. C_{ds}, (a) n=0, (b) n=1.

A summary of the fractional bandwidth achievable for the device analyzed in this section vs. C_{ds} is shown in figure 3.42, whereby a 2.5:1 specification is defined for the maximum summing node VSWR due to off-state impedance pulling. As can be seen, the n=0 Doherty achieves an inverse linear bandwidth scaling vs. parasitic capacitance, whereas the n=1 topology bandwidth does not scale linearly with C_{ds}.

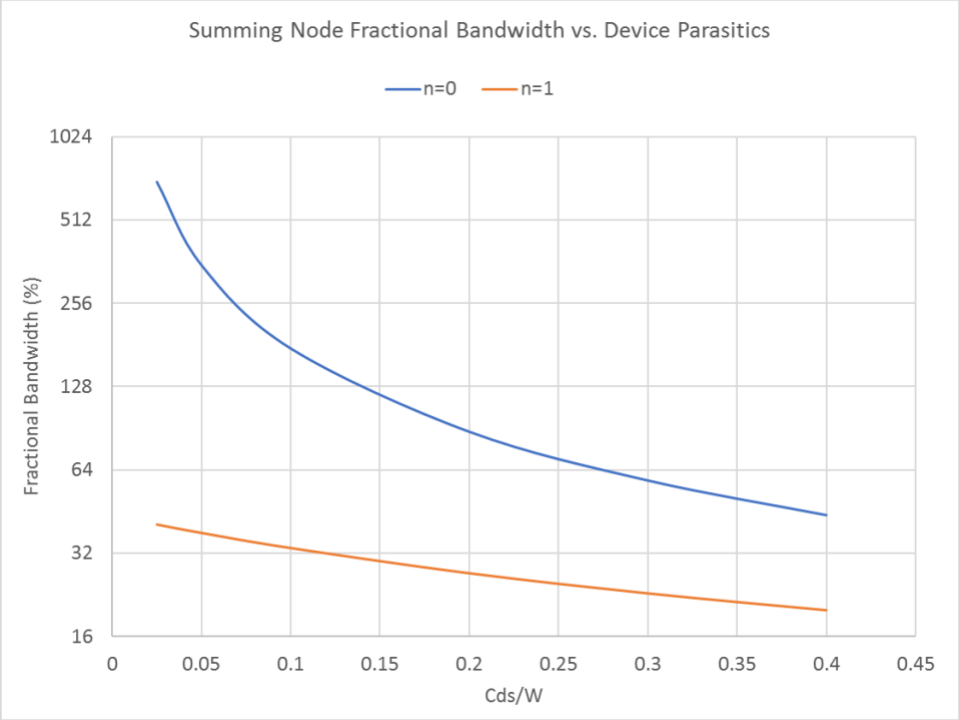

Figure 3.42 Summing node bandwidth with peak loading vs. device C_{ds} per watt.

Figure 3.43 shows the impact of the peaking amplifier operating voltage on the available bandwidth of the Doherty summing node. The bandwidth limitation occurs due to the parallel loading of the peaking amplifier off-state impedance at the summing node with the target system impedance. In this analysis, a constant C_{ds}/W is assumed, and a bandwidth reduction of the Doherty system impedance occurs as the transistor operating voltage increases. At higher voltages, the constant phase rotation of the peaking amplifier off-state impedance, returning to an open circuit, can be seen at the frequency extremes at higher drain voltages. However, the greater summing node bandwidth achieved from a device operating at a lower V_{dd} is clear from the plot.

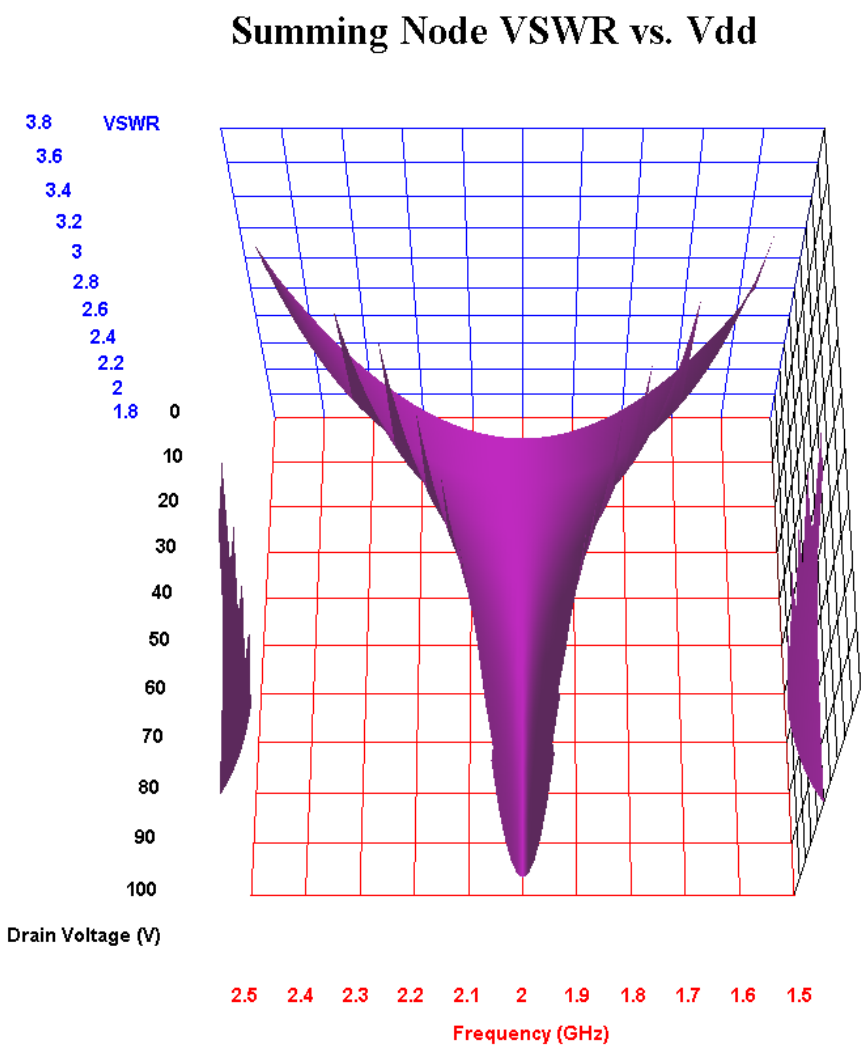

Figure 3.43 Summing node bandwidth with peak loading vs. device operating supply voltage.

The final comparison is considering the impact of output power for a fixed technology on the off-state loading of the Doherty amplifier. In this case the peaking amplifier output power is scaled, and the corresponding load impedance, parasitic capacitance and output matching network parameter values calculated. Figure 3.44 shows the resultant off-state impedance loading at the summing junction. It can be seen that the off-state impedance dispersion is independent of the Doherty power level, and therefore a low power Doherty will incur the same off state loading issues described in this section as a very high-power design.

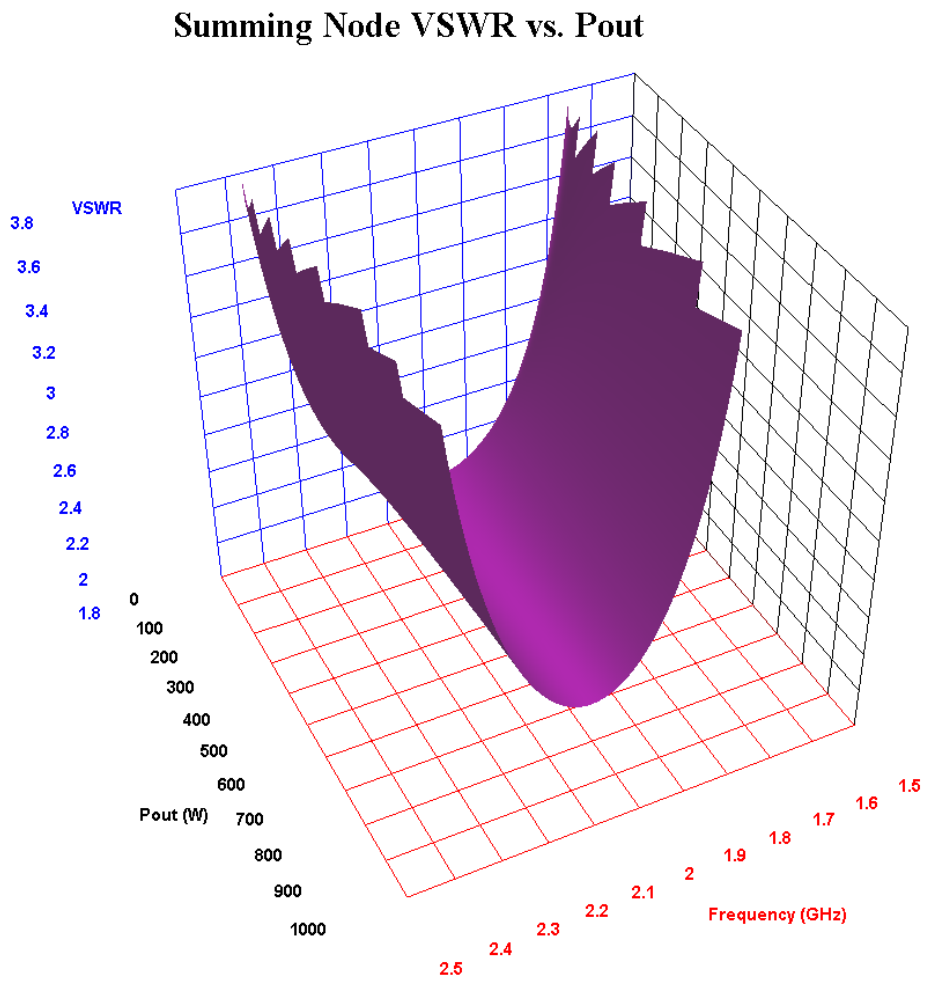

Figure 3.44 Summing node bandwidth vs. output power.

(30v Technology, 0.3pF/W, 10W-1kW)

Figure 3.45 shows the relationship between the peaking amplifier off-state impedance dispersion, the device C_{ds}, and the amplifier matching impedance. The deviation from the high impedance phase angle is plotted over a 200MHz RF bandwidth. This mirrors the data characterized in this section, showing the 'step' function at a circuit matching impedance of 3Ω as the Doherty transitions from the n=1 to the n=0 topology. It can also be seen that the percentage impact due to device C_{ds} is higher in the n=0 case (verifying figure 3.42) and the insensitivity to the PA matching impedance, verifying figure 3.38.

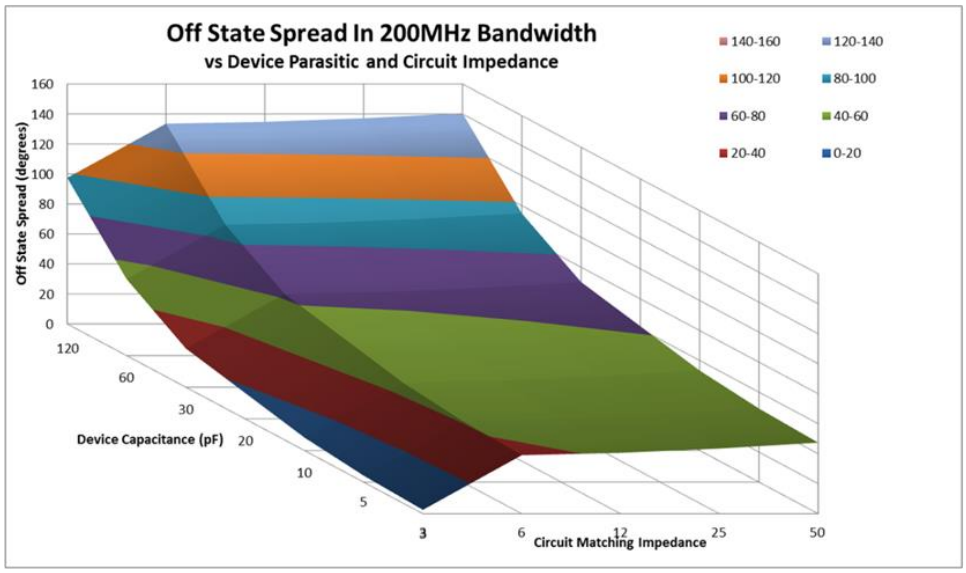

Figure 3.45 Peaking amplifier off-state impedance dispersion at Doherty summing node at fixed P_{out} in 200MHz RF bandwidth vs. circuit matching impedance and device capacitance.

3.1.6 Bandwidth Analysis Conclusions

In section 3.1, a significant amount of data is presented, showing techniques to increase the bandwidth of the Doherty amplifier. In this section, a summary of the key findings is presented, subdivided into the four categories (a) – (d), as shown in figure 3.1.

3.1.6.1 System Impedance Transformer

The transformation from external impedance to the Doherty system impedance is considered one of the critical areas in which wideband operation is achieved. Since no load modulation occurs in this section of the Doherty, wideband structures enable a transformation to the desired system impedance, thus enabling the use of low transformation ratios in the impedance inverter (an area in which wideband structures prove un-suitable). Of the matching structures analyzed in section 3.1.1, the use of either a Klopfenstein taper or a two-section matching network using a geometric mean rotator to enhance the bandwidth over a traditional two-section match prove to be the optimal networks. The choice between the two will be determined by the desired bandwidth,

return loss, and insertion loss specifications for the Doherty. In simple terms, for applications up to 60% fractional bandwidth, the two-section + GMR transformer offers lower insertion loss and a more compact footprint than the Klopfenstein taper. However, for wider bandwidths, or if a system impedance with <-20dB passband ripple is targeted, the Klopfenstein proves to be a more optimal matching technique.

While not analyzed in this thesis, the concept of a parallel Doherty using either a 90-degree hybrid coupler or 180-degree balun structure, such as the Marchand balun, to define the system impedance of a parallel Doherty pair could form an excellent basis for a wideband Doherty. This is discussed in more detail in further work in chapter six.

3.1.6.2 Main Amplifier Impedance Inverter

In section 3.1.2, an extensive analysis of the bandwidth of wideband structures was presented. In conclusion, a case of Occam's Razor is found, i.e., "the simplest solution is the right one." While the more complex structures provide significant bandwidth under static load conditions, none of these provide a satisfactory solution when used as an impedance inverter, since the wideband transformation assumes specific load and source termination impedances. When the termination impedances are changed, the low transformation ratio rapidly increases, leading to a narrow-band performance from the structures under load modulation. Therefore, a design approach using the system impedance transformer to perform a significant portion of the necessary impedance transformation, and enabling the impedance inverter no more than two times the load modulation is a recommended approach for wideband design.

As mentioned in 3.1.6.1, an opportunity for further research would be achieving load modulation from coupled structures and is discussed in chapter six, further work.

3.1.6.3 Device Characteristics

When considering the device characteristics, two cases emerge in enabling wideband Doherty operation. Firstly, the use of a high voltage technology permits a reduction in the impedance transformation required for the power amplifier matching network. Therefore, it could be concluded that this enables wideband RF performance since a lower matching transformation ratio enables wider band operation. However, this fails to consider several factors. Firstly, independent of the PA transformation ratio, an impedance inverter will be necessary in order to provide correct load modulation to the main amplifier. In this case, the minimum impedance transformation at both full power and back off will be the load modulation ratio of the Doherty. Therefore, the advantage of the higher voltage technology is the potential for a higher summing node impedance. However, it has been demonstrated in 3.1.2 that the bandwidth of the summing node is not a limitation since wideband structures enable the transformation across the bandwidth needed.

Perhaps less obvious is that the impact of the transistor C_{ds} is greater as the load impedance increases since the ratio between the C_{ds} shunt reactance to the shunt load impedance is reduced. Additionally, since an internal shunt match is ubiquitous in high power RF transistors, the loaded Q of this resonant network increases as the load increases, either though the Doherty load modulation or the use of higher voltage technology.

Therefore, assuming all other parameters being equal (specifically the C_{ds} per watt), a lower voltage technology should provide wider operating bandwidth in the Doherty amplifier, assuming the impedances are realizable at the PCB level.

3.1.6.3 Peaking off State

The peaking amplifier off-state impedance presents a significant challenge in achieving wideband Doherty performance. This is because, as the non-conducting peaking transistor output impedance is rotated through the peaking amplifier matching network, significant S22 dispersion occurs, which in turn appears in parallel with the system impedance. The impact is an increase in

the summing node VSWR, thus a load pulling effect to the main amplifier Z_{mod} load impedance. As the reflection of the peaking amplifier deviates from n*360 degrees (open) through an odd multiple of 180 degrees, the peaking amplifier and matching network will present a short circuit, and effectively short circuit the summing junction. This behavior can be seen in figure 3.40 (b), whereby the VSWR at the summing node is plotted with the peak amplifier loading. It is shown that achieving >50% fractional bandwidth from the Doherty topology with n > 0 becomes impractical due to the inherent shorting of the summing node by the peaking amplifier.

A significant enabler in reducing the peak-off state bandwidth limitation is the use of the n=0 topology, proposed in [5,8]. In this topology, the Doherty system impedance is selected to present the optimal load impedance to the peaking amplifier, thus removing the need for a 180-degree peaking matching network. This topology enables a significant increase in summing node bandwidth over the 50% fractional bandwidth limitation of the n>0 architecture, as shown in figure 3.42. Also, the bandwidth of this architecture scales 1:1 with reduced device parasitic, unlike the n>0 architecture where the bandwidth to parasitic capacitance ratio scales much less than 1:1, as shown in figure 3.42.

3.2 Frequency Compensation - The Complementary Resonator

Figure 3.46 shows the typical schematic for the internal matching of a high-power RF transistor. The topology is almost universally used for high power LDMOS devices, and is gaining popularity for GaN HEMT devices several reasons: -

- It facilitates realistic matching impedances at the package level
- It enables the implementation of a 90-degree matching network for wideband Doherty design
- It enables wide RF matching bandwidth
- It facilitates an RF short circuit at C_{block}. This enables a node for connection of circuitry such as a baseband termination with minimal loading at the RF frequency, as described in chapter four.

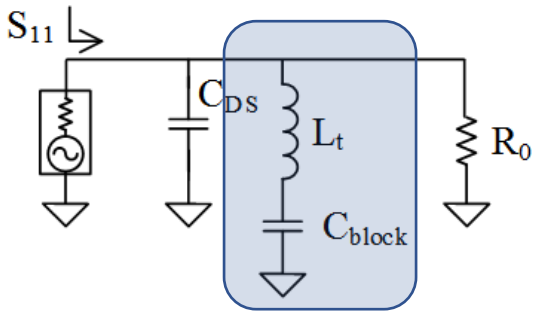

Figure 3.46 Typical internal matching network (shaded) of high-power RF transistor.

The bandwidth of the network can be analyzed when presented with R_o, the optimal load for the current generator given the I_{max} and the breakdown voltage for the device. The response of the network, as seen by the current generator, can be seen in figure 3.47 (a) impedance and (b) bandwidth.

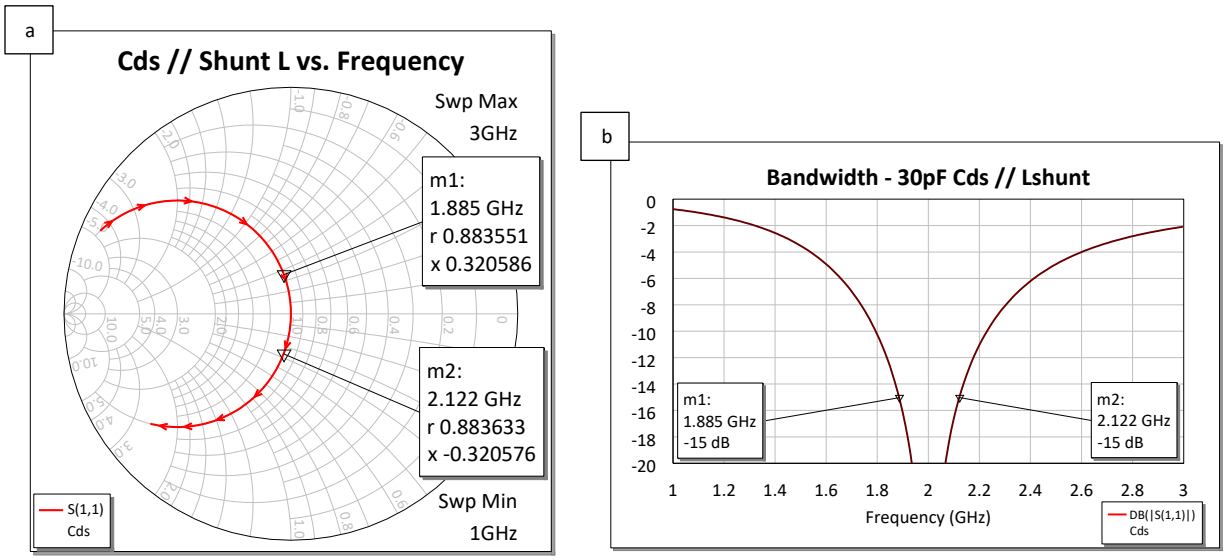

Figure 3.47 (a) Impedance of parallel network as seen by current generator, (b) corresponding bandwidth.

In order to compensate for the impedance dispersion over frequency, a compensation network is added to the transistor, whereby the network provides the inverse impedance trajectory over frequency to the parallel resonance of C_{ds} and the internal matching inductor L_t, first proposed by [7]. The network is implemented as a series resonant network, as shown in figure 3.48.

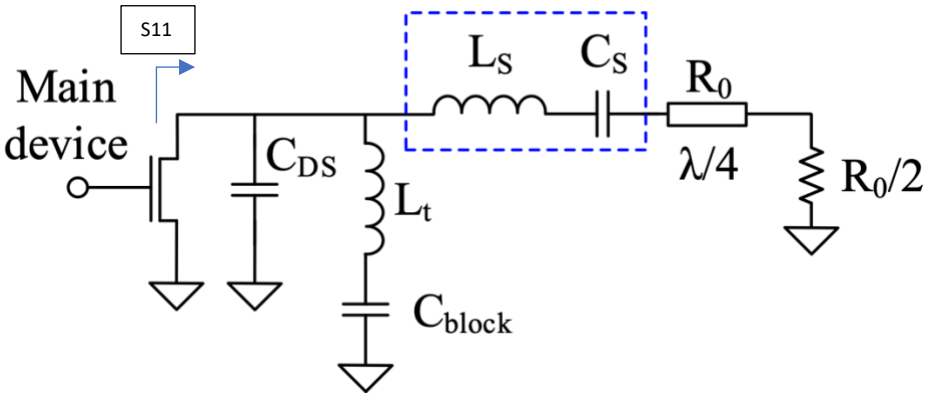

Figure 3.48 Internal matching network of high-power RF transistor with complementary resonator circuit

The response of the series resonator in isolation is shown in figure 3.49 and can be seen to mirror the impedance trajectory over frequency of the transistor C_{ds} and internal shunt match L_t.

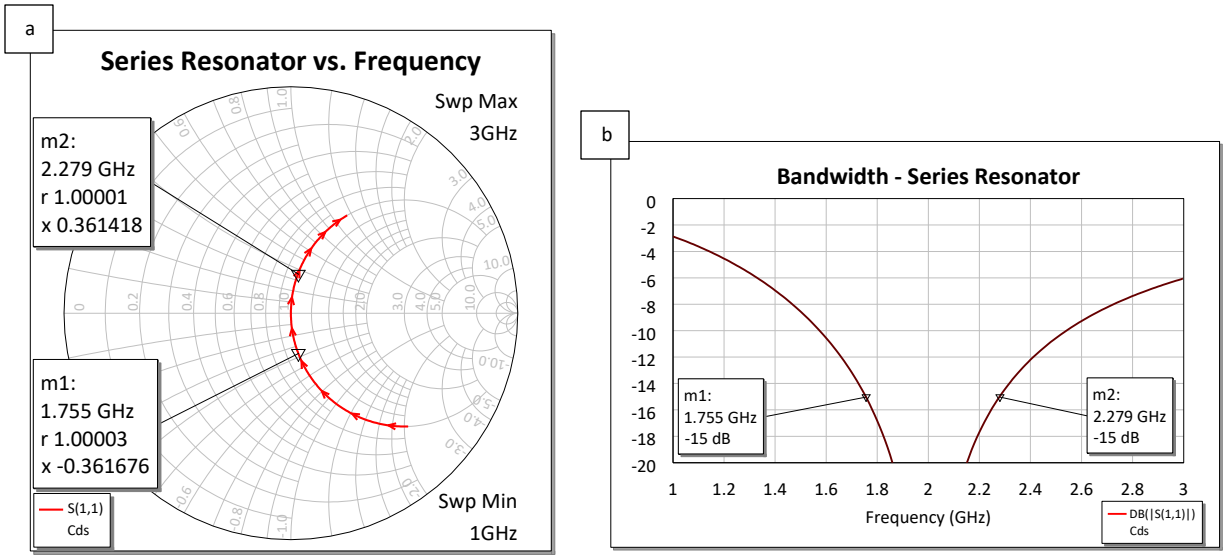

Figure 3.49 (a) Impedance of series network as seen by current generator, (b) corresponding bandwidth

The cumulative frequency response of the parallel resonator comprising of the transistor C_{ds} and internal shunt match L_t, and the compensation series resonator is shown in figure 3.50, where the fractional matching bandwidth can be seen to increase from 26.2% to 33.2%.

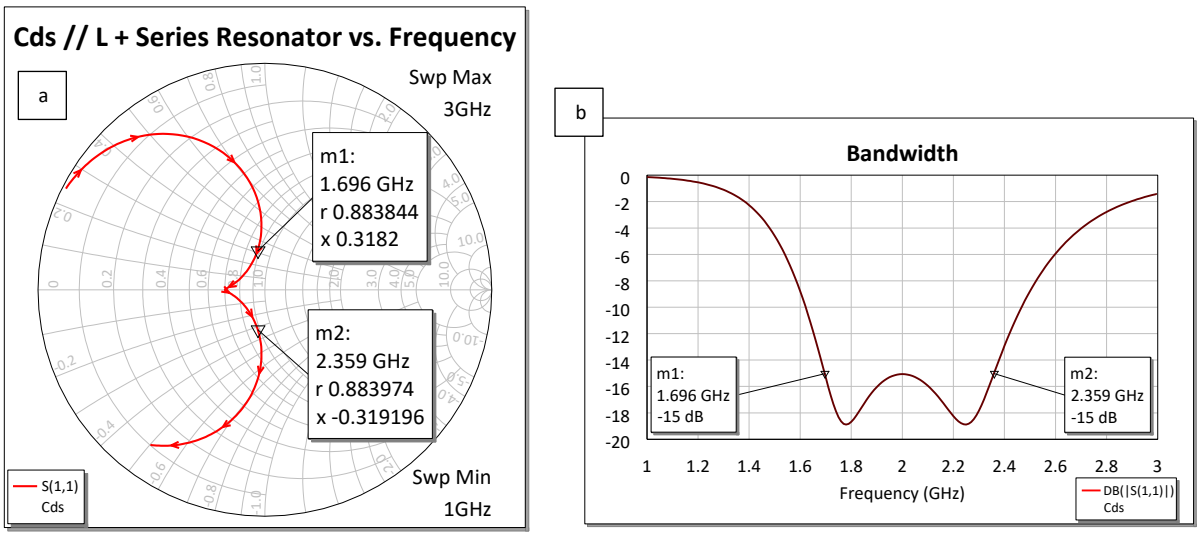

Figure 3.50 (a) Impedance of combined shunt and series networks as seen by current generator, (b) corresponding bandwidth.

In addition to the increase in matching bandwidth, the technique enables the absorption of the series bond wire inductance required to connect the die to the RF package. Therefore, zero-phase shift is achieved between the die and the package reference plane. This provides a package level transistor with load-pull contours aligned on a 'real to real' impedance plane, which will, therefore, match the impedance presented by the Doherty impedance inverter. Figure 3.51 shows the measured load-pull contours of a transistor adopting this topology.

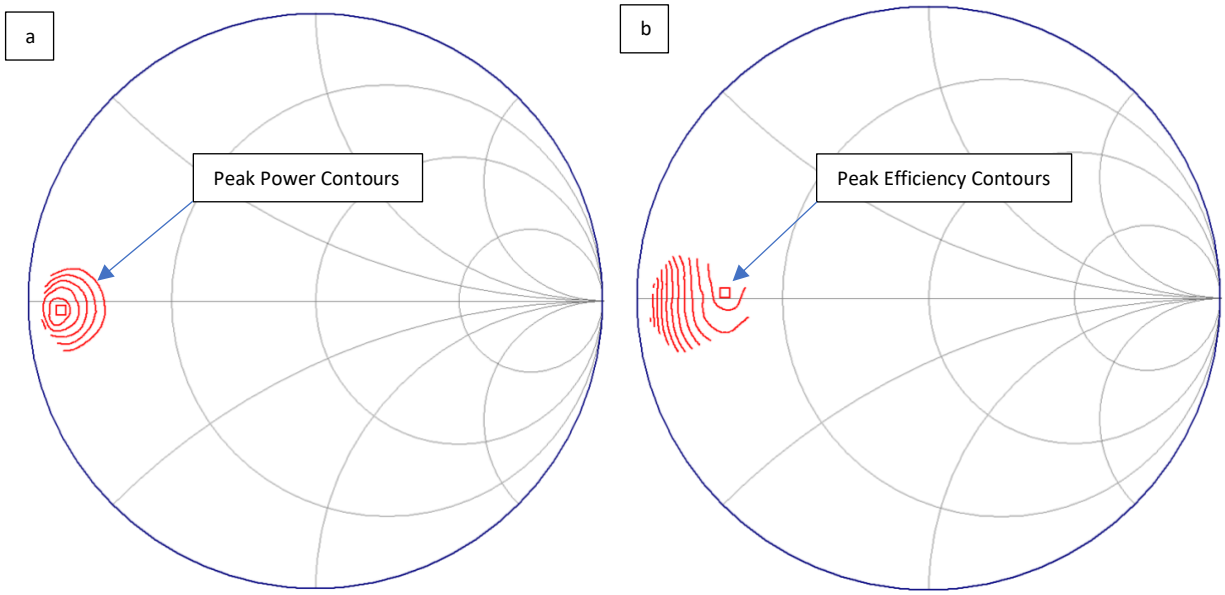

Figure 3.51 Measured load pull contours for 60 watt packaged LDMOS transistor with series compensation network: (a) peak power contours, (b) Peak Efficiency Contours.

It should also be noted that an additional application for this topology is in the n=0 Doherty described in 3.1.4, whereby a zero-phase connection to the summing node is required. In this case, the zero-phase connection to the summing node enables both the correct load modulation to the peaking device (rather than requiring a reactive load due to the connecting inductance of the bond-wires) and also a correctly centered off-state impedance at the summing node (which would be shifted by the series L of the bond wires).

3.3 Doherty Implementations

In this section, the learnings from analyzing the behavior of each component in the Doherty combiner in sections 3.1 and 3.2 are combined in proposing two wideband Doherty architectures. The bandwidth enhancements from these approaches are described, and two wideband Doherty reference circuits are designed to demonstrate the bandwidth achievable from the topologies.

3.3.1 Constant Impedance Doherty

The constant impedance Doherty derives its name from the constant impedance nature of the output combining network. The embodiment described is shown in figure 3.52, whereby a Klopfenstein taper defines a system impedance equal to the main amplifier Z_{mod} impedance. An impedance inverter with a characteristic impedance equal to Z_{mod} connects the summing node and the main amplifier, thus providing a 'constant impedance.' Since the Klopfenstein taper provides a wideband impedance down the high-pass cutoff frequency, and the characteristic impedance of the inverter is equal to the Klopfenstein taper load impedance, a wideband Z_{mod} impedance is presented to the main amplifier at back-off power levels.

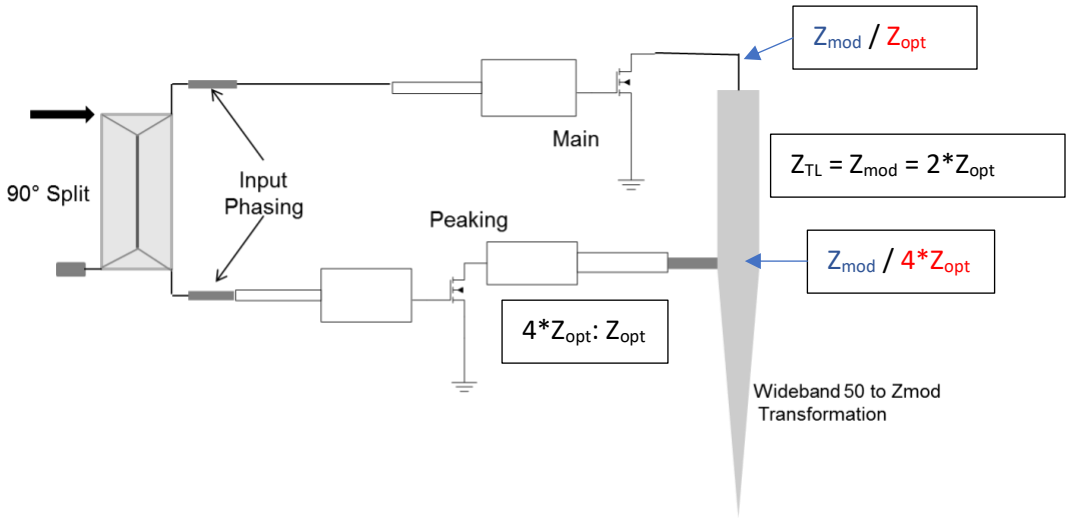

Figure 3.52 Constant impedance Doherty node impedances indicated at back-off (blue) and full power (red).

Figure 3.53 shows a simulated comparison between the constant impedance topology shown in figure 3.52, and a textbook 80-watt symmetric Doherty amplifier, comprising of two 50Ω matched amplifiers with two-section, 180-degree matching networks, combined at a 25Ω summing node.

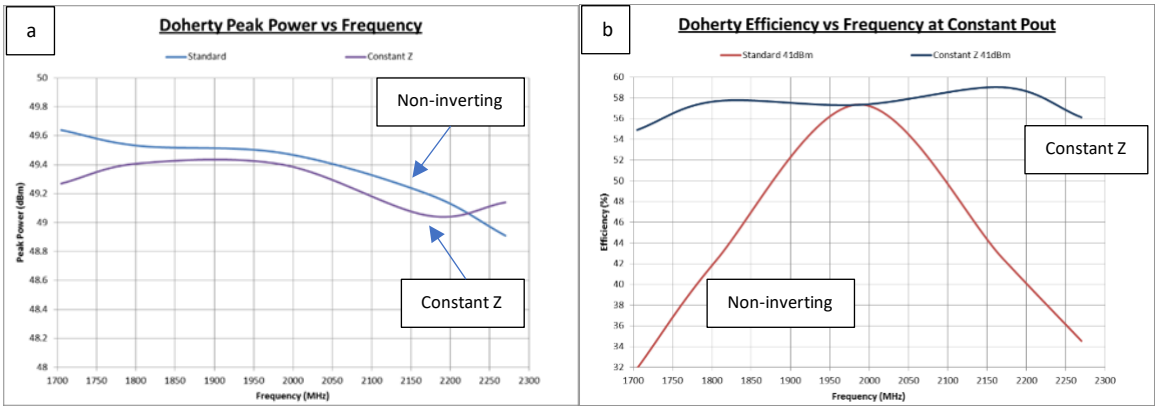

Figure 3.53 Constant impedance vs. standard non-inverting Doherty performance comparison, (a) peak power, (b) efficiency at back-off.

3.3.1.1 Constant Impedance Doherty with Complementary Resonator

A detailed analysis of the compensation network is presented in 3.2, whereby the out of band frequency response of the device C_{ds} and internal shunt match is compensated by a series resonant network with complementary characteristics at frequencies either side of the center frequency. In this section, the impact of the technique is analyzed in the constant impedance Doherty amplifier topology. Main amplifier characteristics, both with and without the compensation, are shown across frequency.

Figure 3.54 shows the impact of the compensation network on the efficiency of the main amplifier. In each case, four values of capacitance are evaluated, with the corresponding series inductance to resonate at the center frequency of 2GHz.

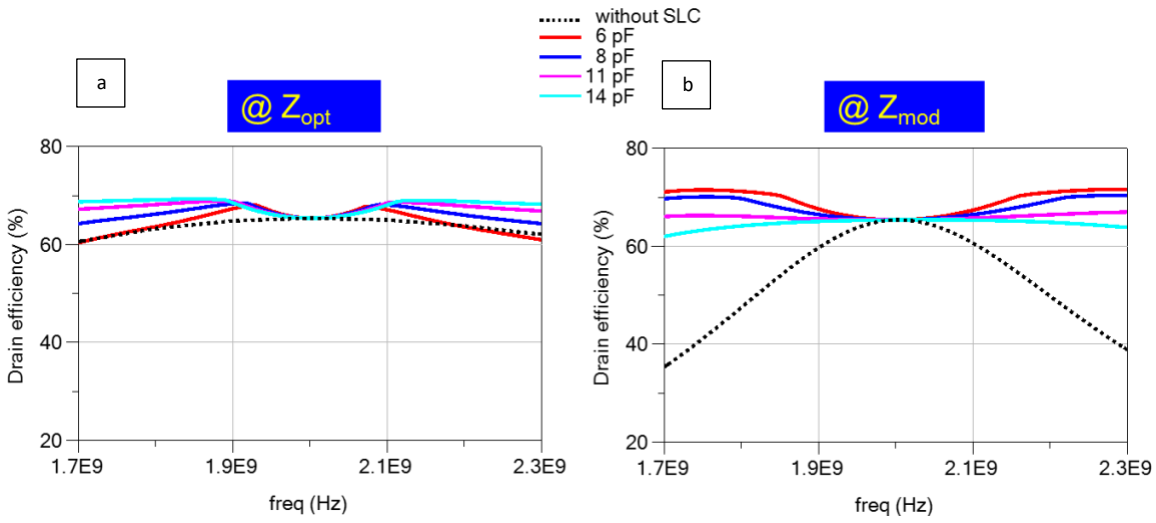

Figure 3.54 Main amplifier drain efficiency with and without Compensating Resonator, (a) Z_{opt}, (b) Z_{mod} load impedances.

It can be seen that the technique is highly effective in maximizing the efficiency of the main amplifier in back-off – where the increase in load impedance inherently increases the Q of the shunt resonator (C_{ds} // $L_{internal}$), thus reducing the bandwidth. This can be seen in figure 3.54 (b) as the efficiency drops off away as the frequency deviates from the center frequency, and also in figure 3.56 (b) as the impedance dispersion of the matching network is overlaid on the efficiency contours of the device.

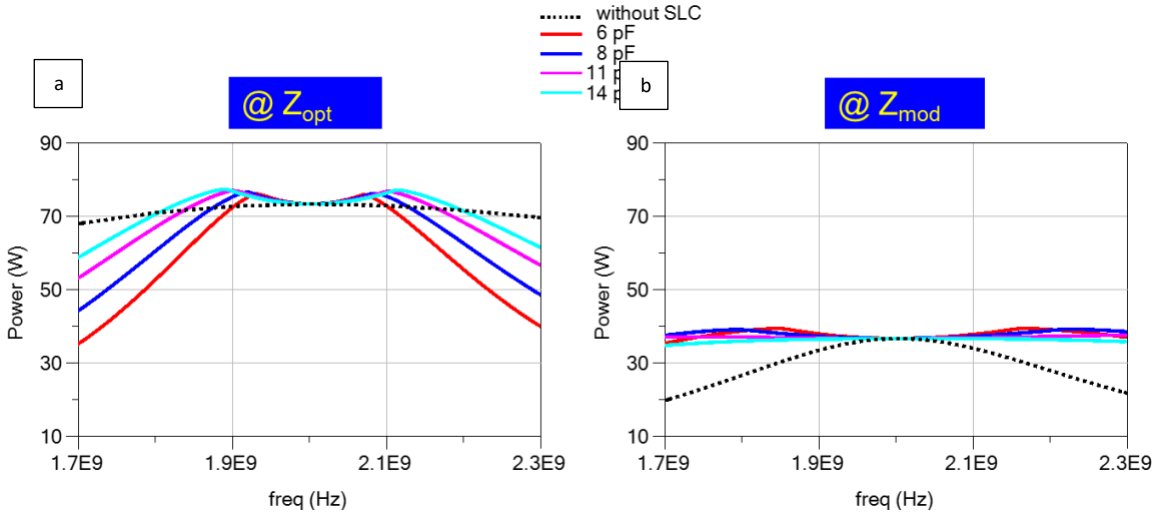

Figure 3.55 Main amplifier output power with and without compensating resonator, (a) Z_{opt}, (b) Z_{mod} load impedances.

Figure 3.55 shows the corresponding impact on the peak power of the main amplifier in both the Z_{opt}, figure 3.55 (a), and Z_{mod}, figure 3.55 (b). In this case, the Z_{mod} power can be seen to flatten as matching network load impedance aligns with the power contours of the device, shown in figure 3.56. However, the full power case shows where caution needs to be taken to ensure the resonator does not impact the peak power capability of the main amplifier, and therefore the 14pF capacitor would be optimal the selection in this case.

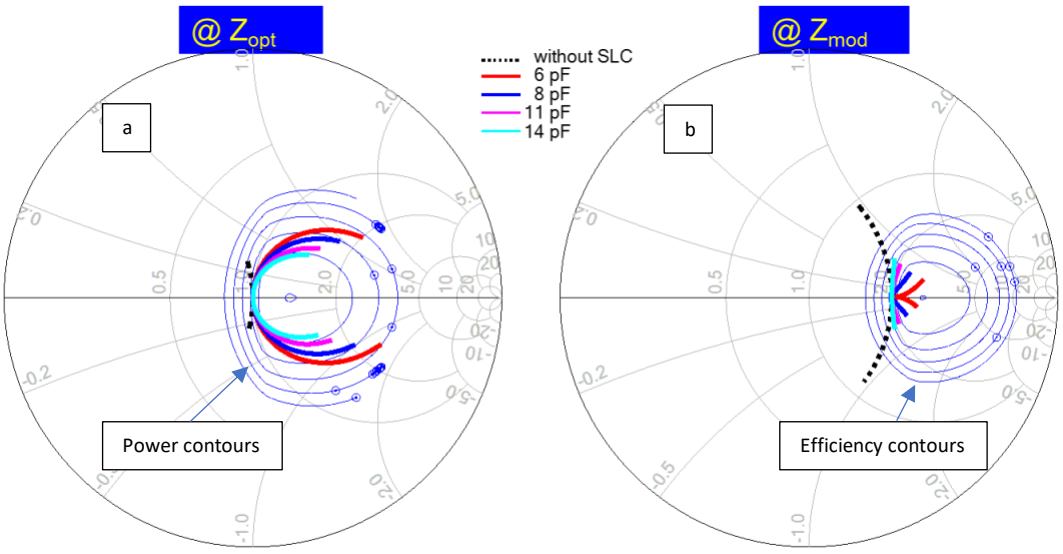

Figure 3.56 Main amplifier intrinsic load pull contours with matching network trajectory across frequency overlaid, with and without compensating resonator, (a) peak power, (b) peak efficiency.

3.3.1.2 Constant Impedance 120W, 1.8-2.2GHz, LDMOS Doherty with Complementary Resonator

In order to demonstrate the performance of the constant impedance topology, a 120W Doherty is designed using 2 x 60mm LDMOS transistors, mounted in a TO265 package. The main amplifier uses external drain bias pins, attached to the internal shunt matching MOSCAP to enable the

implementation of the (DC blocking) complementary resonator described in 3.2. The resonator uses a package lead mounted capacitor (indicated by 'resonator' in figure 3.54), to which the series drain wires are attached, thus forming a series L-C resonant network.

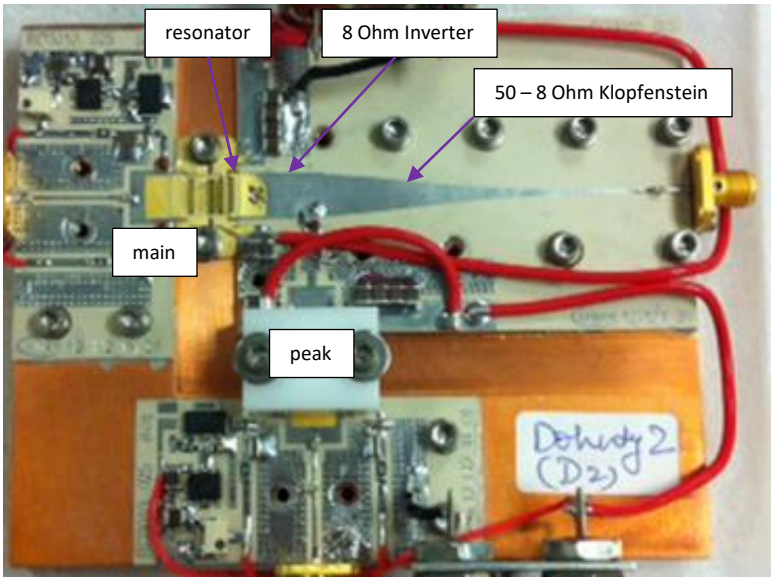

Figure 3.54 120W constant impedance 1.8 – 2.2GHz Doherty demonstration circuit.

In this design, the Z_{opt} impedance of the transistor is 4 Ohms, with a Z_{mod} of 8Ω from the symmetric 2:1 Load modulation. A Klopfenstein taper is used to define a wideband 8Ω system impedance at the summing junction, and an 8 Ohm impedance inverter is used to transform the $2*Z_{mod}$ (16Ω) summing node impedance to $Z_{mod}/2$ (4Ω) at full power to enable the main amplifier to deliver full current at high power levels. The peaking amplifier uses a 4:1 matching network, transforming from $2*Z_{mod}$ to Z_{opt} at full power.

The circuit is constructed using a 25 mil thick Rogers 3010 laminate, selected to enable a close match between the impedance inverter trace and the package lead width.

Measured P3dB results for the amplifier shown in figure 3.54 are shown in figure 3.55. It can be seen that the amplifier achieves a P3dB of 120 watts or higher from 1.8 – 2.2GHz. Efficiency, shown in figure 3.56, is greater than 50% at a fixed output power of 44dBm across RF the frequency range. To the authors knowledge this is the highest power LDMOS amplifier to

simultaneously achieve 20% fractional bandwidth and over 100 watts of peak power. A comparison to other published work is shown in table 3.1.

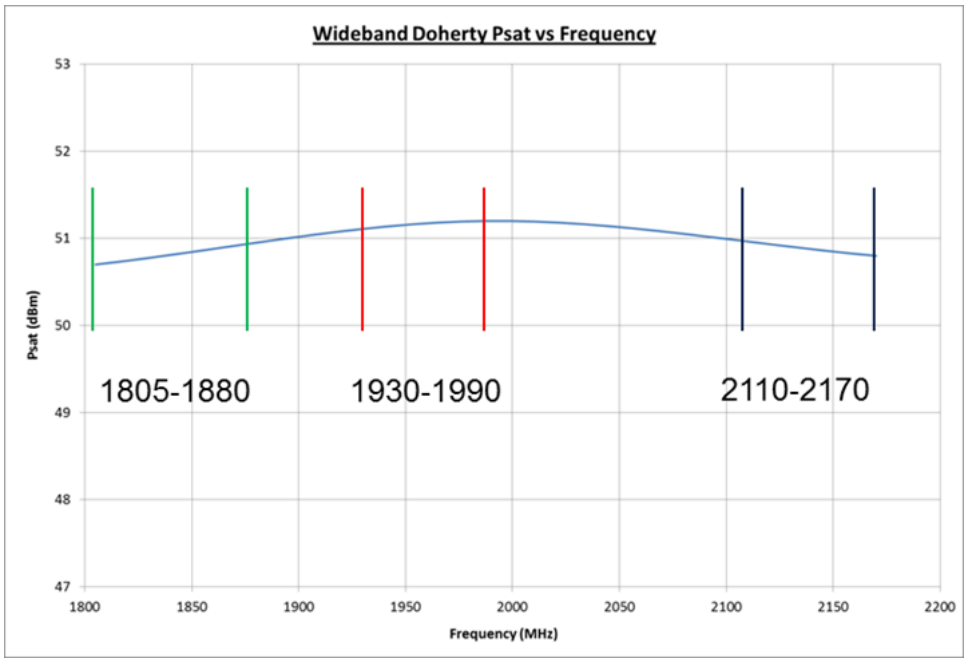

Figure 3.55 Constant impedance Doherty demonstration circuit, peak power.

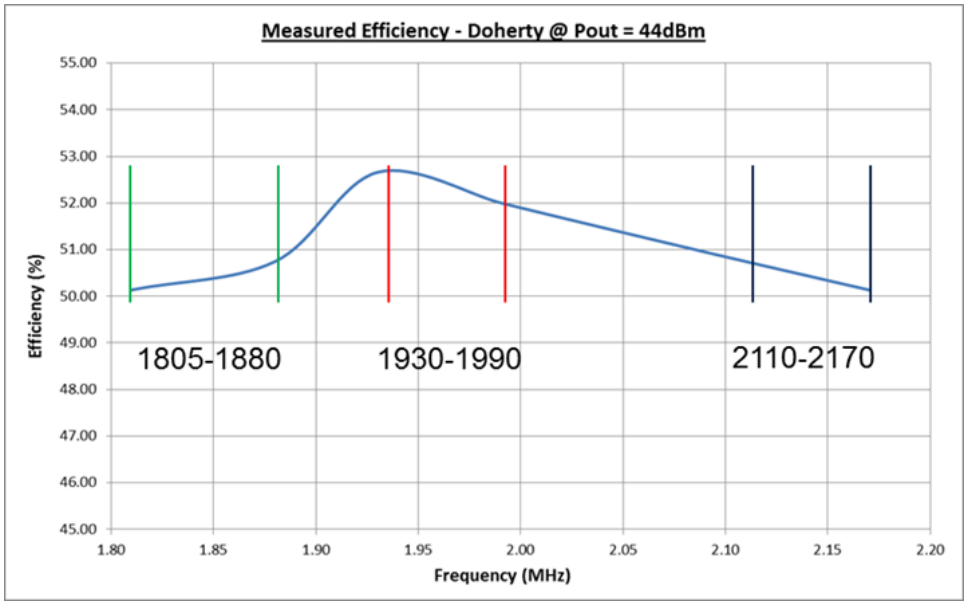

Figure 3.56 Constant impedance Doherty demonstration circuit, efficiency at 6dB output back-off

Reference	Technology	Frequency Range (GHz)	Peak Pout (dBm)	Efficiency (%) 6dB OBO (min – max)
10	LDMOS	1.7 – 2.3	42	34 – 54
11	GaN	1.9 – 2.4	40.5	35 - 40
This Work	LDMOS	1.8 – 2.2	50.7	50 – 52.8

Table 3.1 Constant Impedance Doherty, performance comparison vs. current state of the art.

3.3.2 Zero Phase n=0 Peaking Doherty

The zero-phase n=0 Doherty takes the findings from section 3.14 to maximize the Doherty bandwidth through minimization of the parasitic loading of the peaking amplifier off-state at the Doherty summing node. A summing node impedance below the individual Z_{opt} of the amplifiers is required, calculated in 3.17. In the case of the symmetric Doherty, the summing node = $R_{opt}/2$, since both the main and the peaking amplifiers deliver equal summing node current at full power.

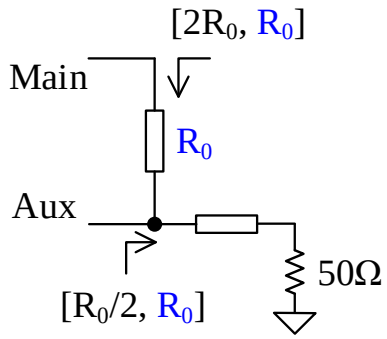

Figure 3.57 Zero Phase n=0 peaking symmetric Doherty simplified schematic.

The low summing node impedance enables the removal of the 180-degree peaking amplifier matching network since, at full power, the summing node presents the correct Z_{opt} to the peaking transistor. A comparison of the n=0 Doherty to a textbook, non-inverted symmetric Doherty, is shown in figure 3.58. This plot can be directly compared to the constant impedance implementation, shown in figure 3.53.

Figure 3.58 Zero phase n=0 Doherty vs. standard non-inverting Doherty performance comparison, (a) peak power, (b) efficiency at back-off.

3.3.3 Mixed Technology, Zero Phase n=0 Peaking Amplifier Used with Constant Impedance Combiner

Figure 3.59 shows the simplified schematics for (a) the n=0 zero peaking phase Doherty, and (b) the constant impedance Doherty. As described in sections 3.1.4 and 3.2.1, each topology provides key bandwidth enhancements. Specifically, configuration (a) minimizes the peaking amplifier off-state impedance dispersion at the summing node but incurs a 4:1 impedance inverter transformation ratio at back-off. Configuration (b) provides a combiner with no bandwidth limitations at back-off, but the necessary 4:1 matching network on the peaking branch causes a considerable bandwidth penalty to the summing node VSWR.

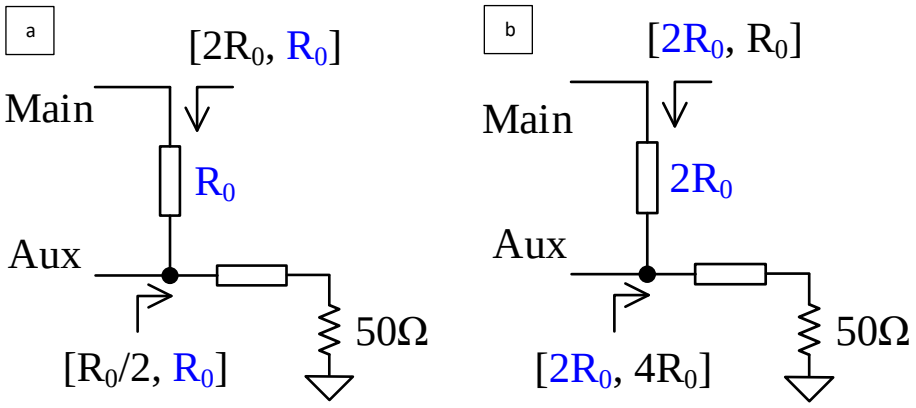

Figure 3.59 Simplified schematics, (a) Zero Phase n=0 Peaking Doherty, (b) Constant Impedance Doherty.

The optimum load (R_{opt}) for a 100W device operating at a 50v drain voltage, with an assumed 5v knee voltage is calculated in 3.18 – 3.20.

$$R_{opt} = \frac{2.(V_{dd} - V_{Knee})}{I_{max}} \tag{3.18}$$

Where: -

$$I_{max} = \frac{8.P_{out}}{2.(V_{dd} - V_{knee})} \tag{3.19}$$

Therefore: -

$$R_{opt} = \frac{90}{8.89} = 10.13\ \Omega \tag{3.20}$$

For the symmetric zero phase Doherty, the summing node impedance is calculated in 3.21.

$$Z_{system} = \frac{10.13}{2} = \ 5.06\ \Omega \tag{3.21}$$

Therefore, for a fixed technology, 200-watt symmetric n=0 Doherty, the impedances are: -

- System impedance 5.06 Ω
- Impedance Inverter 10.13 Ω
- Z_{mod} (main – back-off) 20.25 Ω
- Z_{opt} (main - full power) 10.13 Ω

In this case, the impedance inverter provides a 4:1 impedance transformation at back-off power levels, shown in 3.2.2 to impact the Doherty efficiency bandwidth in comparison to the constant impedance Doherty in 3.2.1. However, if a mixed technology Doherty is considered, whereby the main amplifier uses a lower voltage device technology to the peak amplifier, the main amplifier Z_{opt} impedance is reduced. Figure 3.60 plots the Z_{opt} impedance for 100W transistor vs. operating V_{dd}. and V_{knee}. The increase in Z_{opt} through the use of a device technology optimized for the main amplifier V_{dd} can be seen in comparing the plot 'Knee-5v' (50v device run at lower V_{dd}) and 'Knee=10% of Vdd' (device technology selected for operating V_{dd}).

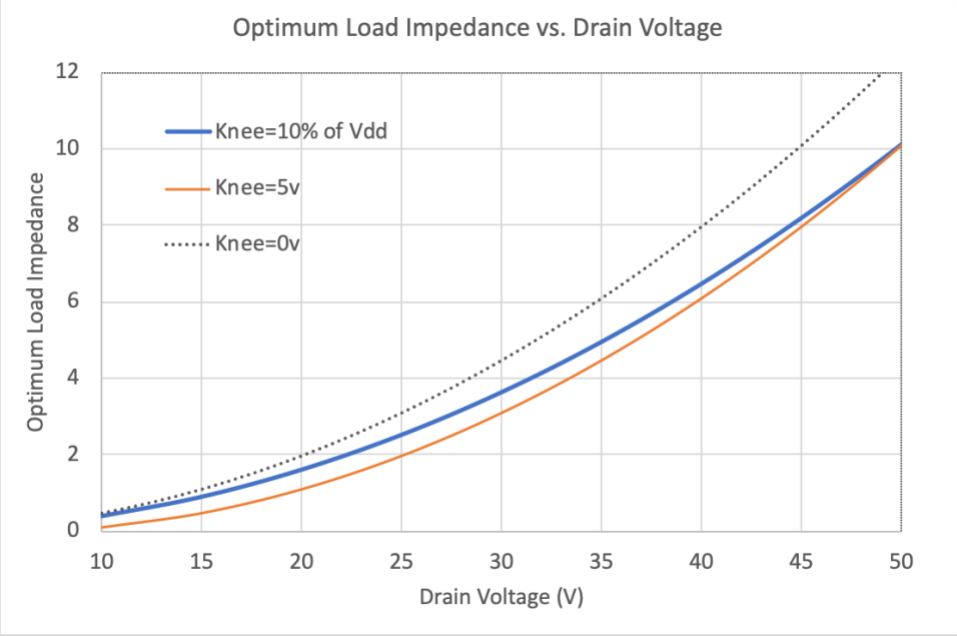

Figure 3.60 Optimum load impedance for 100W transistor vs. drain voltage.

The target impedances for a constant impedance Doherty are: -

- System impedance 5.06Ω
- Impedance Inverter 5.06Ω
- Z_{mod} (main – back-off) 5.06Ω
- Z_{opt} (main - full power) 2.53Ω

Figure 3.61 (b) shows the main amplifier drain voltage vs. Z_{opt} impedance. From here, it can be seen that for a symmetric two-way Doherty, a main amplifier drain voltage of 25v with Z_{opt} = 2.53Ω is targeted in order to provide constant impedance Doherty operation.

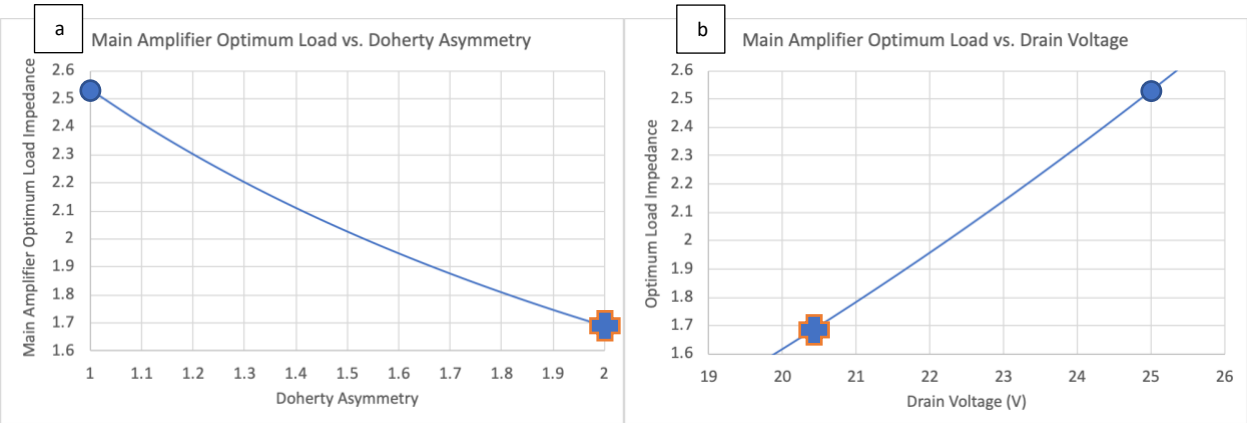

Figure 3.61 Main amplifier optimum load impedance in 200W Doherty vs. (a) Doherty asymmetry and (b) drain voltage.

Figure 3.61 (a) shows the target main amplifier load impedance vs. asymmetry is for a 200-watt Doherty amplifier, with the power delivered per branch in each case shown in figure 3.62. In this case, markers indicate the required load impedance and main amplifier for a symmetric and a 2:1 asymmetric Doherty amplifier.

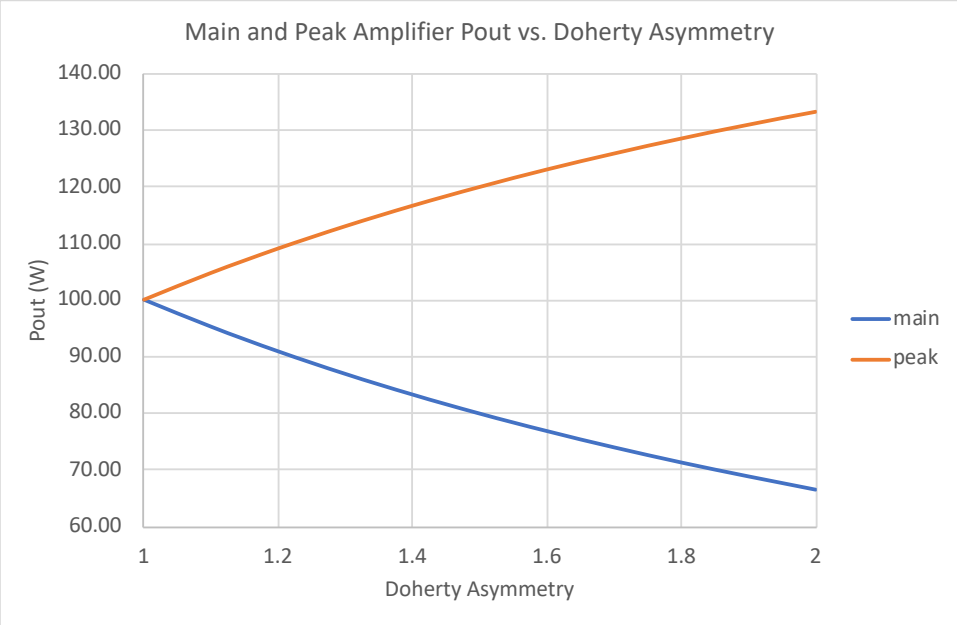

Figure 3.62 200W Doherty, main and peaking amplifier output power vs. asymmetry.

3.3.4 Zero Phase Peaking, Constant Impedance, 1.8-2.7GHz GaN Doherty Demonstrator

In order to demonstrate the concept of the zero-phase peaking, constant impedance Doherty, a 225-watt wideband GaN-based Doherty is developed to cover the popular cellular infrastructure bands between 1800MHz (lower end of cellular band 3) and 2700MHz (upper end of cellular band 41). This frequency grouping is popular because many transmission bands fall within this frequency range, and either side, a significant frequency jump occurs, down to the sub 1GHz bands, or up to bands 42 and 43 occupying the 3.4 – 3.8GHz range.

However, amplifiers used for base station transmitters do not cover this 40% fractional bandwidth for the following reasons: -

- GaN-based amplifiers have only recently started appearing in products, and the parasitic capacitance inherent in LDMOS based transistors precludes designs with these fractional bandwidths.
- Even with GaN transistors, there are many design challenges to achieve this bandwidth from the Doherty amplifier, while still hitting the efficiency targets of single or narrower

band designs. These design challenges will likely be a topic of much research over the next decade, as the market demand for wideband amplifier designs increases.

In this demonstrator, due to the availability of die for the prototype, the Cree 0.4um 50v die is used for both main and peaking amplifiers. Given the lower main amplifier supply voltage, a shorter gate length transistor would have been the preferable choice. However, the 0.4um die proved to have good performance at lower voltages, and after load pull evaluation vs. supply voltage, the following configuration was arrived at.

Main Amplifier	1 Die	20mm die	Vdd=28v
Peak Amplifier	2 Die	7.2mm die	Vdd = 50v
Theoretical Doherty Asymmetry	1.4 to 1		

Table 3.2 – Doherty Demonstrator Summary

A simplified schematic for the Doherty amplifier is shown in figure 3.63, with a picture of the assembled measured board shown in figure 3.64. The novel design approaches used for this design are described below.

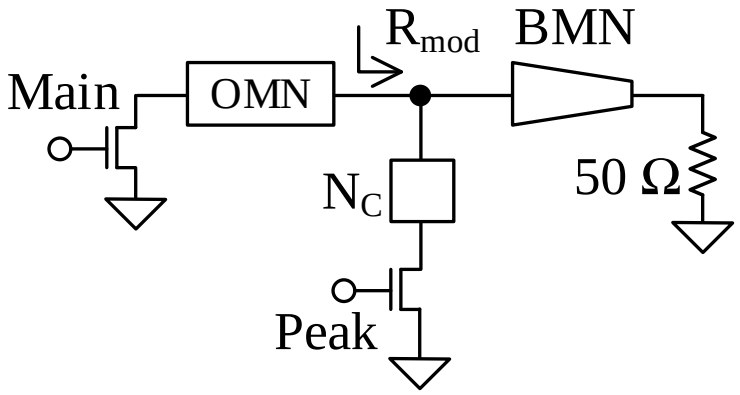

Figure 3.63 Simplified Schematic of Zero Phase Peaking Doherty Amplifier

Firstly, a wideband impedance transformer is used to transform the 50Ω load to the system impedance of 6.5Ω. Two DC blocking capacitors are absorbed into the transformer in order to block the peaking amplifier drain voltage. The peaking amplifier is split into two smaller die. This technique provides two advantages over a single peaking amplifier design: -

1. The peaking current is symmetrically fed into the summing junction, giving rise to a more uniform impedance across the main amplifier manifold at the end of the impedance inverter.
2. The split peak has significantly less electrical length per drain manifold than a single die of twice the size. This ensures the electrical length along the impedance inverter into which the peaking amplifier current injection occurs is reduced by a factor of 2.

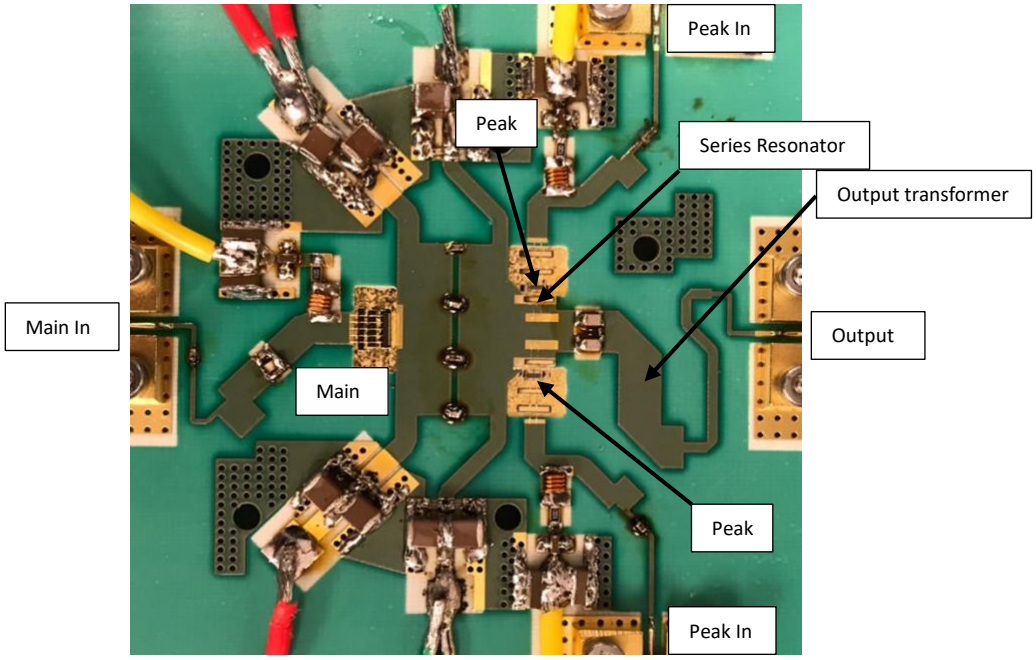

Figure 3.64 –Assembled Doherty Amplifier

Both peaking amplifiers use the compensation resonator introduced in section 3.1.5. In this case, the resonator enables the connection of both the peak dies to the summing node while mitigating the reactance of the series bond wires. This inductance has the effect of adding a phase offset to the open circuit off-state impedance, as well as adding series inductance to the load presented to the peaking die, pulling the fundamental impedance match. A novel approach to enable DC bias is introduced to the resonator circuit [9], whereby a second inductor is bonded from the MOSCAP

top plate to the PCB. The inductance has sufficient reactance to prevent the loading of the RF network. The implementation can be seen in figure 3.65, labeled 'DC Feed.'

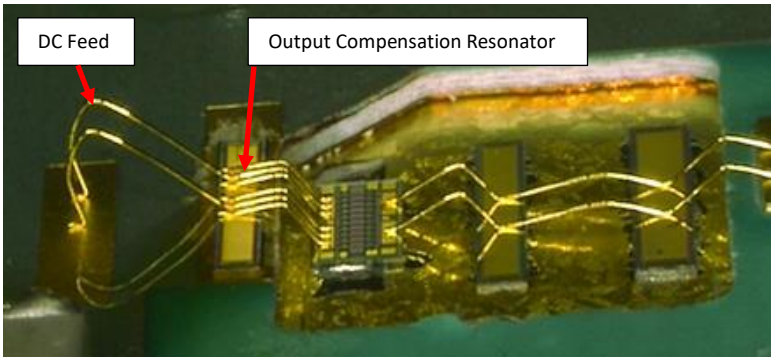

Figure 3.65 – Peaking amplifier output resonator, including novel DC feed.

Measured results on the Doherty are taken using a two-path DPD system, whereby the amplitude and phase of the input signal can be controlled. For the following measurements, the amplitude and phase delta between the main and peak sides is optimized at each frequency but fixed with amplitude drive up. Therefore, no idealized digital turn on characteristics vs. drive level is used. Figure 3.66 shows the amplitude and phase delta of the peaking path relative to the main path across frequency.

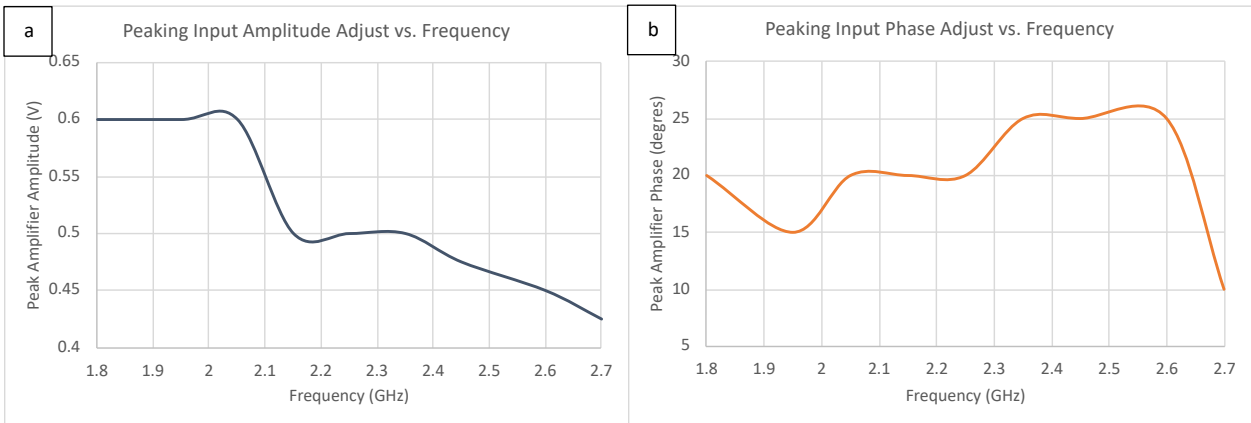

Figure 3.66 Peaking amplifier input (a) amplitude, (b) phase adjust relative to main path.

Figure 3.67 shows the simulated vs. measured S parameters, and a close correlation occurred between the two. The correlation is attributed to the significant use of 3D FEM analysis for the

board, wires, and capacitors in advance of PCB tape out, as well as accurate ASM based models of the GaN die.

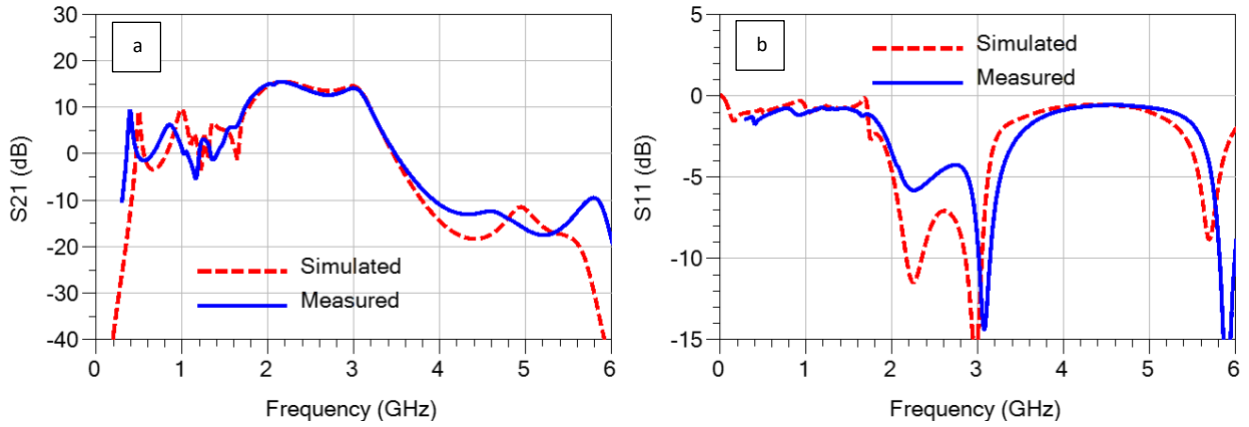

Figure 3.67 Simulated vs. measured small signal S parameters (a) S21 and (b) S11.

Those experienced in Doherty design may note an interesting observation in figure 3.67 (a), whereby the traditional 'nulls' either side of the center frequency in the S21 plot, typically seen in Doherty amplifiers, are not present in this measurement. Since these nulls are the result of the peak amplifier off-state loading, this measurement gives an indication of the increased off-state bandwidth achieved using this n=0 design methodology.

Figure 3.68 shows the measured drain efficiency and gain plots for the complete Doherty PA. A minimum P3dB of 53.5dBm was measured from the module, thus giving a theoretical OBO peak at 46dBm for the voltage and current scaling used in the design.

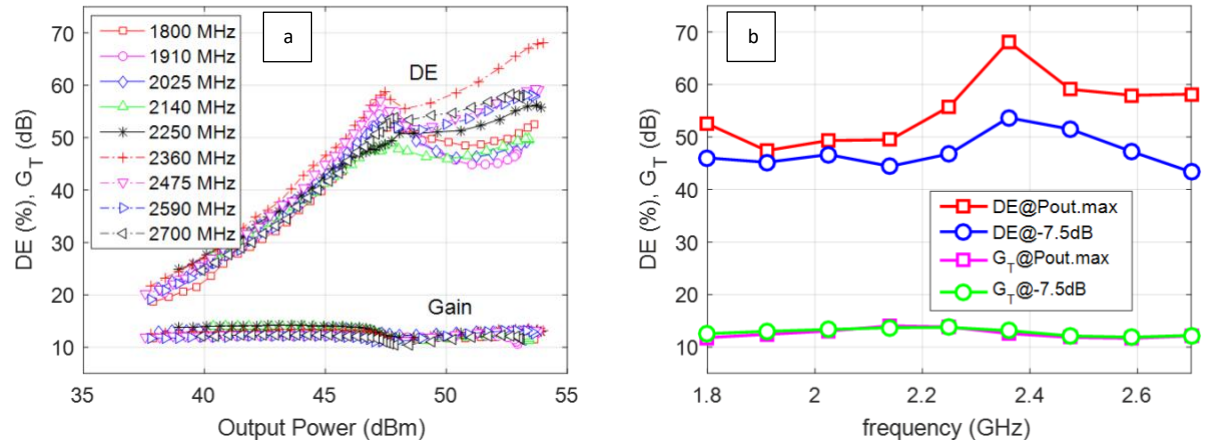

Figure 3.68 Measured drain efficiency and gain vs. (a) output power, (b) frequency.

An important metric for the power amplifier intended for use in the cellular infrastructure market is the linearizability since the transmit path needs to adhere to mandatory emissions specifications. The linearizability of the Doherty is characterized across frequency, and the ACPR is plotted at the average power level. The DPD system used is a Texas Instruments reference board, with 250MHz RX and TX bandwidth. The algorithm used for DPD is a General Memory Spline model, with 16 segments, a memory depth of 2, and an envelope delay of 1.

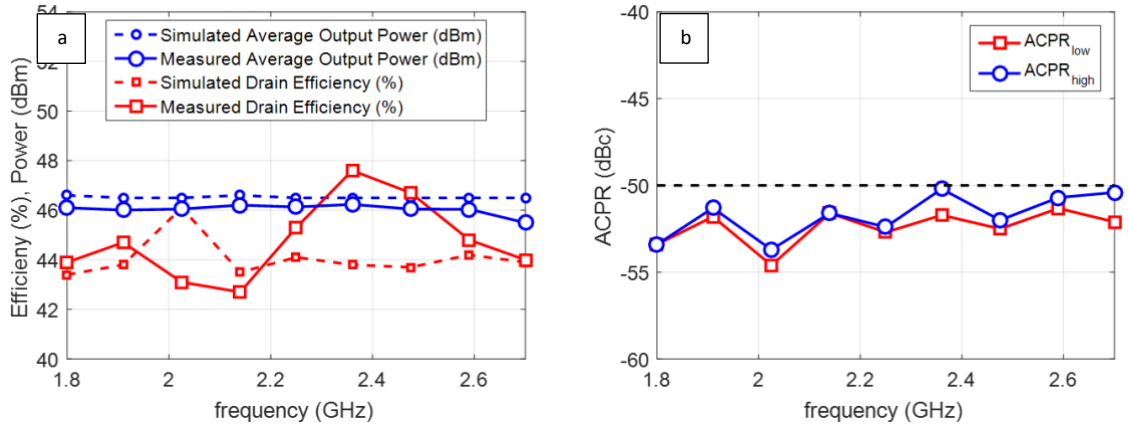

Figure 3.69 (a) Efficiency at average output power, (b) ACPR after DPD correction at P_{avg}.

Figure 3.70 shows the pre and post-DPD down-converted envelope of the (a) amplitude and (b) phase response of the amplifier. The inflection during peak turn on can be seen in the pre corrected response; this, however, did not pose a problem for linearization, as can be seen by the corrected amplitude and phase plots.

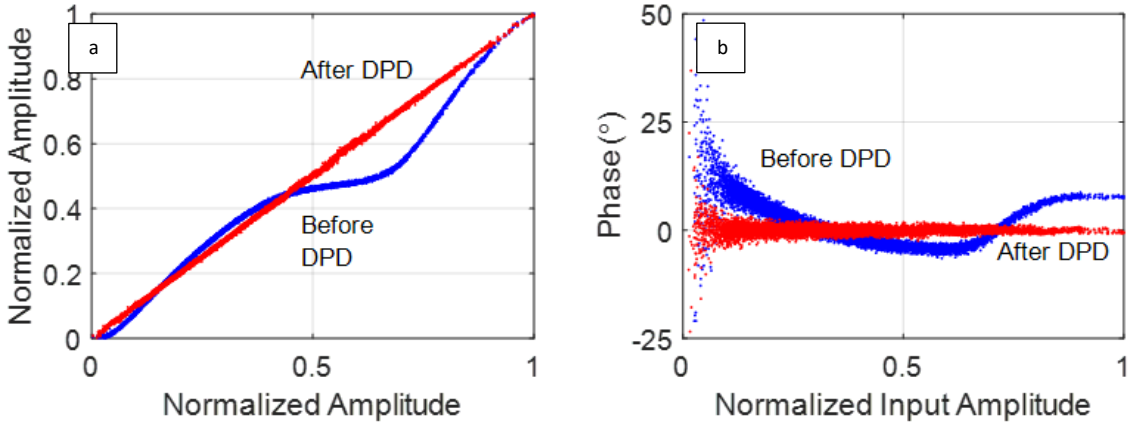

Figure 3.70 Pre and post DPD: (a) Pin vs. Pout, (b) AM-PM.

A performance comparison of this work compared to current state of the art is shown in table 3.3. To the authors knowledge this is the first Doherty amplifier with over 200 watts of peak power covering 1.8 – 2.7GHz at 50% efficiency.

Reference	Technology	Frequency Range (GHz)	Peak Pout (dBm)	Efficiency (%) (min – max)
12	GaN	1.7 – 2.7	52.7	39.0 – 47.0 (7.5dB OBO)
13	GaN	0.6 – 0.9	47.0	51.7 – 52.3 (11.75 OBO)
14	GaN	1.55 – 2.35	43.7	50.4 – 56.2 (8dB OBO)
This Work	**GaN**	**1.8 – 2.7**	**53.5**	**50.0 – 52.8 (7.5dB OBO)**

Table 3.3 Zero phase peaking constant impedance Doherty, performance comparison vs. current state of the art.

3.4 Summary

This chapter presents an extensive bandwidth analysis of each section of the Doherty amplifier. The results from the analysis led to the development of several novel Doherty implementations. These implementations were prototyped and demonstrated the intended bandwidth enhancements, providing state of the art bandwidth and power results summarized in tables 3.1 and 3.2.

In chapter four addresses the requirement for the wideband high-efficiency amplifier to support concurrent wideband signal transmission. A novel internal device matching topology is presented, and the enhancements to both RF and signal bandwidth are demonstrated.

3.5 References

[1] Klopfenstein, R.W., "A Transmission Line Taper of Improved Design", Proc. IRE, Vol.44, pp.31-35, January 1956.

[2] Shirasaki, H., "The Design Charts by Waveguide Model and Mode-Matching Techniques of Microstrip Line Taper Shapes for Klopfenstein and Hecken Type Microstripline Tapers", 2000 Asia Pacific Microwave Conference Proceedings.

[3] Pozar, D.M., Microwave Engineering 4th Edition, John Wiley & Sons, 2011.

[4] Kajfez, D., Prewitt, J., "Correction to "A Transmission Line Taper of Improved Design"", IEEE Transactions on Microwave Theory and Techniques Letters, May 1973.

[5] Wilson, R.J., "Wideband Doherty Amplifier Circuit", US Patent 8193857B1.

[6] Wilson, R.J., "Wideband Doherty Amplifier Circuit Having a Constant Impedance Combiner", US Patent 20120319780, December 2012.

[7] Wilson, R.J., "RF Device with Complementary Resonator Matching Topology", US Patent 20130076446.

[8] Wilson, R.J., "Impedance spreading wideband Doherty amplifier Circuit with Peaking Impedance Absorption", US Patent 20130241639A1, September 2013.

[9] Jang, H., Canning, T., Herrmann, B., Mokhti, Z., Trang, F., Wilson, R., "RF Power Amplifier with Frequency Selective Impedance Matching Network", US Patent 20190229686A1, September 2019.

[10] Qureshi, J.H., Li, N., Neo, W.C.E., van Rijs, F., Blednov, I., de Vreede, L.C.N., "A Wide-Band 20W LMOS Doherty Power Amplifier", IEEE MTT-S International Microwave Symposium, May 2010.

[11] Akbarpour, M., Helaoui, M., Ghannouchi, F., "A Transformer-Less Load-Modulated (TLLM) Architecture for Efficient Wideband Power Amplifiers", IEEE Transactions on Microwave Theory and Techniques, Vol. 60, No. 9, September 2012.

[12] Chen, X., Chen, W., Zhang, Q., Ghannouchi, F., Feng, Z., "A 200 Watt Broadband Continuous-Mode Doherty Power Amplifier for Base-Station Applications", IEEE MTT-S International Microwave Symposium, June 2017.

[13] Barthwal, A., Rawat, K., Koul, S., "A Design Strategy for Bandwidth Enhancement in Three-Stage Doherty Power Amplifier with Extended Dynamic Range", IEEE Transactions on Microwave Theory and Techniques, Volume 66, Issue 2, February 2018.

[14] Pang, J., He, S., Dai, Z., Huang, C., Peng, J., You, F., "Design of a Post-Matching Asymmetric Doherty Power Amplifier for Broadband Applications", IEEE Microwave and Wireless Components Letters, Volume 26, No. 1, January 2016.

Chapter Four – RF Amplifier Signal Bandwidth and DPD Correction

The topic of extending RF amplifier signal bandwidth has proved to be a growing area of research as the trend toward multi-band solutions for wireless infrastructure increases. Much of this research [1] has focused on the output baseband impedance presented to the output current generator of a power transistor, influenced by internal matching components, and the interaction with external PCB bias and RF matching networks. However, until very recently [2], little research has targeted the input networks, which also contribute to the signal bandwidth limitations and linearized performance of the power amplifier in a digital pre-distortion system.

In this chapter, the topic of the influence of RF amplifier input networks is explored in detail. A novel input matching topology is proposed, with significantly enhanced baseband characteristics compared to the current state of the art. The concept is evaluated using both a simulated DPD test bench and a measured comparison of two 60-watt LDMOS devices with only input topology changes. The two devices are evaluated in identical circuits, other than the input RF matching, and using two commercially available DPD platforms. Comparison results of both the device characteristics and linearized performance are presented.

4.1 Signal Bandwidth Limitations in RF Power Amplifiers

The definition of 'RF amplifier signal bandwidth' can be a somewhat ambiguous term to define, since a transistor and circuit with a wide RF bandwidth will functionally transmit a signal with significant carrier separation. Therefore, in defining the signal bandwidth, a proposed starting place is to consider modulation frequency dependencies inherent in both the device in isolation and the device and circuit interaction. Figure 4.1 shows a summary of the relative time constants of the signal frequency dependencies in solid-state RF power devices. These dependencies are referred to as ‘memory effects’ since the instantaneous device characteristics are dependent on the historical signal conditions.

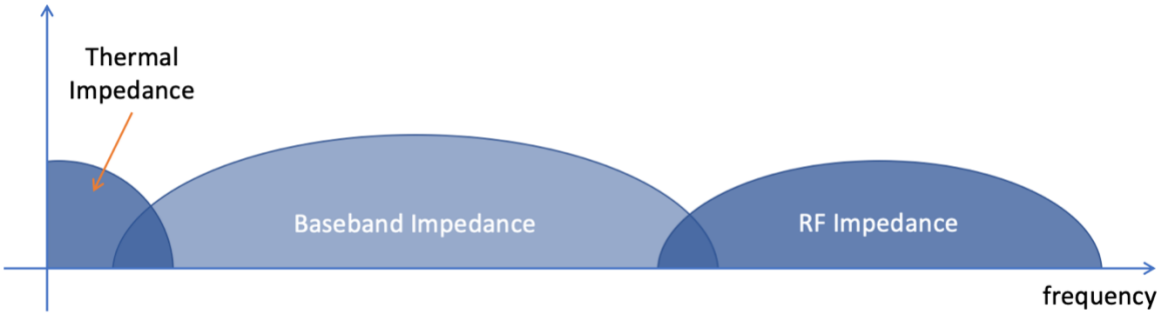

Figure 4.1 Memory effects, mechanisms vs. envelope frequency.

$$f_{baseband} = f_{rf2} - f_{rf1} \tag{4.1}$$

At frequencies less than 1MHz, the memory effects are dominated by the thermal time constant of the transistor [3]. At these low frequencies, the transistor junction undergoes temperature modulation by the slowly varying signal envelope, causing a variation to the instantaneous RF characteristics as a function of time.

At frequencies above 1MHz, the thermal time constant of the transistor diminishes the impact of thermal memory contribution. Above this frequency, the electrical memory effects start to dominate, as the impedance from inductance in the bias circuits causes an increase in the baseband impedance presented to the current generator. As the frequency increases further, a parallel resonance can form between the external bias inductance and the internal capacitance of the device and internal matching capacitors. This resonant frequency is called the video bandwidth (VBW) resonance. Since the parallel resonance forms a high impedance at the VBW resonant frequency, the transistor exhibits high gain at this frequency; therefore, as the signal envelope approaches the VBW resonance, the magnitude of the second-order distortion products, calculated in equation 4.1, increases. These distortion terms sum with the amplifier output at an arbitrary phase angle, giving rise to the commonly observed IMD asymmetry at the amplifier output, whereby the arbitrary phase angle of the vector causes unequal magnitude intermodulation products [4].

At very high modulation rates, the bandwidth of the RF matching networks can become a contributor to the signal bandwidth limitation of the device. At the bandwidth extremes, the RF matching networks have rapidly changing amplitude and phase responses. Therefore, as the signal bandwidth tends to the network RF bandwidth, asymmetry of the signal and IMD products is introduced.

4.2 Input Signal Bandwidth

Before considering the input signal bandwidth of the amplifier, the impact of distortion generating mechanisms though the amplifier needs to be considered. Figure 4.2 shows a simplified schematic for a FET power amplifier, and the core transistor with the finite parasitic capacitance between the drain and the source (C_{ds}), the drain to the gate (C_{dg} or 'miller' capacitance) and the gate to the source (C_{gs}) nodes.

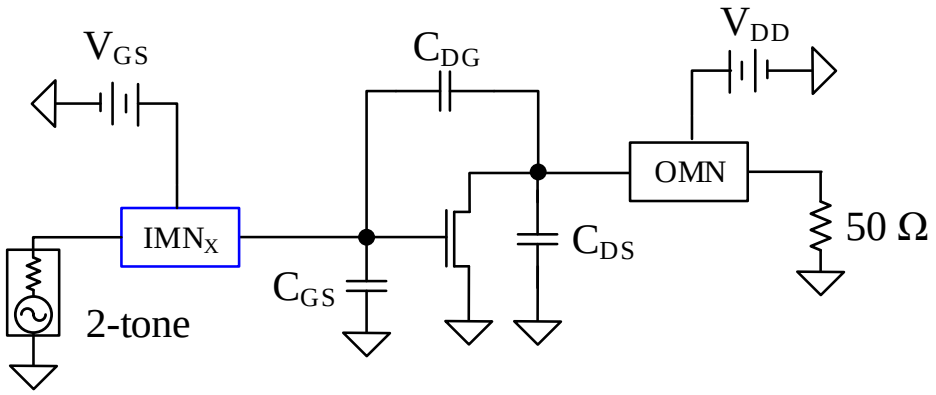

Figure 4.2 Typical RF power amplifier schematic with linear patristic capacitances.

Assuming the transistor in figure 4.2 exhibits linear transconductance, and the parasitic elements, while finite, exhibit linear capacitance vs. signal level characteristics, the resultant network response at baseband frequencies under two-tone RF excitation with a 20MHz carrier separation is shown in figure 4.3.

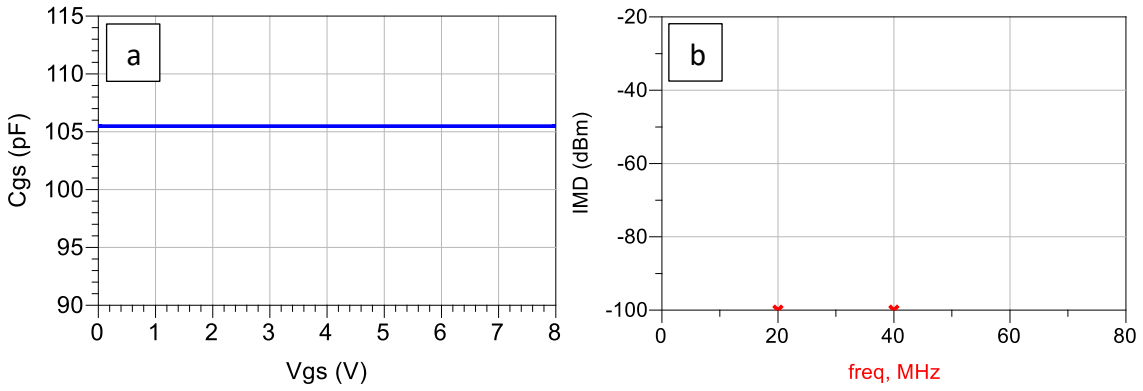

Figure 4.3 (a) C_{gs} characteristics vs. V_{gs} (signal level), (b) baseband gate signal generated with 20MHz carrier separation.

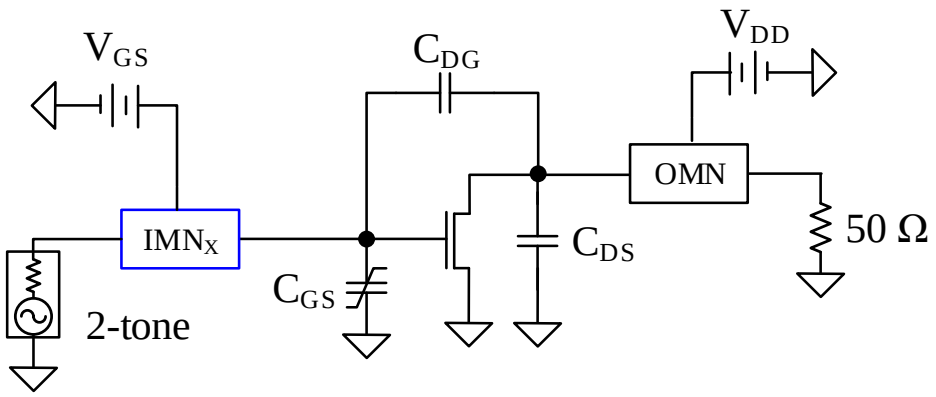

Figure 4.4 Typical RF power amplifier schematic with non-linear C_{gs}.

If the linear capacitance characteristic from figure 4.3 is now replaced with the non-linear characteristic shown in figure 4.4, second-order baseband distortion products are generated at the RF carrier difference frequency (f2-f1). The simulated characteristics are shown in figure 4.5.

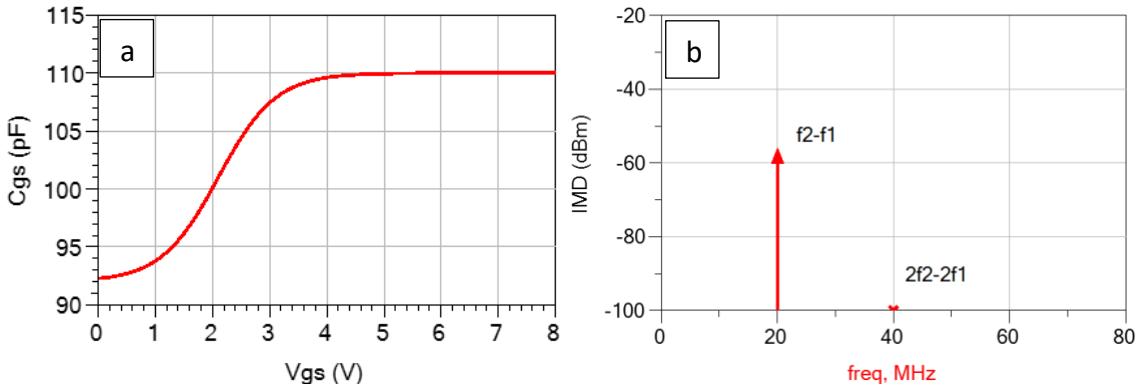

Figure 4.5 (a) C_{gs} characteristics vs. V_{gs} (signal level), (b) baseband gate signal generated with 20MHz carrier separation.

It should be noted that the magnitude of the second-order distortion terms is a function of the impedance presented at this frequency.

4.2.1 Signal Bandwidth Performance of Traditional RF Device

High power RF power devices typically use an internal chip and wire-based matching network in order to transform the FET gate input impedance to a more easily matched intermediate impedance, using PCB based matching structures at the package reference plane. A typical input schematic for a high-power device is shown in figure 4.6, whereby an internal MOS capacitor (C1) is used in conjunction with the bond wires L1 and L2 to form a low pass matching topology.

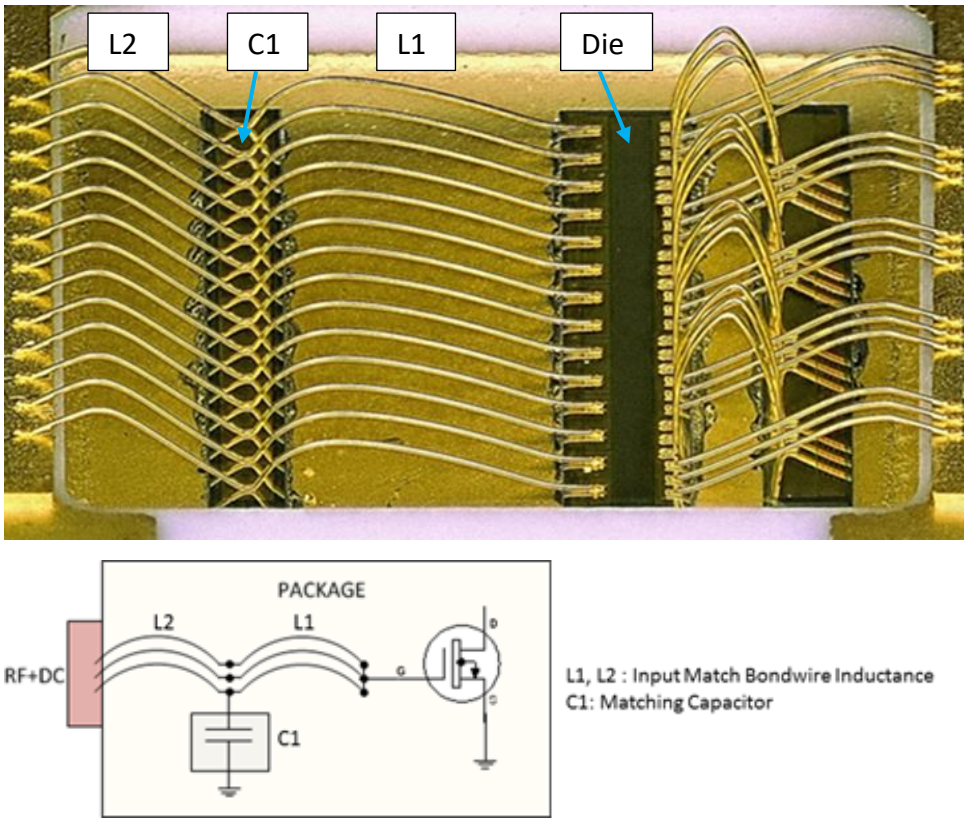

Figure 4.6 Typical high power LDMOS Internal Matching Network.

The typical values of C_{gs} and C1 for a 100W LDMOS transistor are shown in table 1

C_{gs}	90-110pF
C1	25-45pF

Table 1 Typical input capacitance for 100W LDMOS transistor.

The internal input capacitance of the transistor appears in parallel with the external bias inductance used for DC gate bias. The parallel resonant frequency of the network can be derived from equation 4.2.

$$f_{resonance} = \frac{1}{2.\Pi.\sqrt{L_{bias}.C_{internal}}} \quad (4.2)$$

Taking the mean values from table 1, and assuming an external bias network inductance of 18nH, a baseband resonant frequency of 102MHz is calculated. The simulated baseband characteristics, as seen from the current generator for the topology shown in figure 4.6, are shown in figure 4.7.

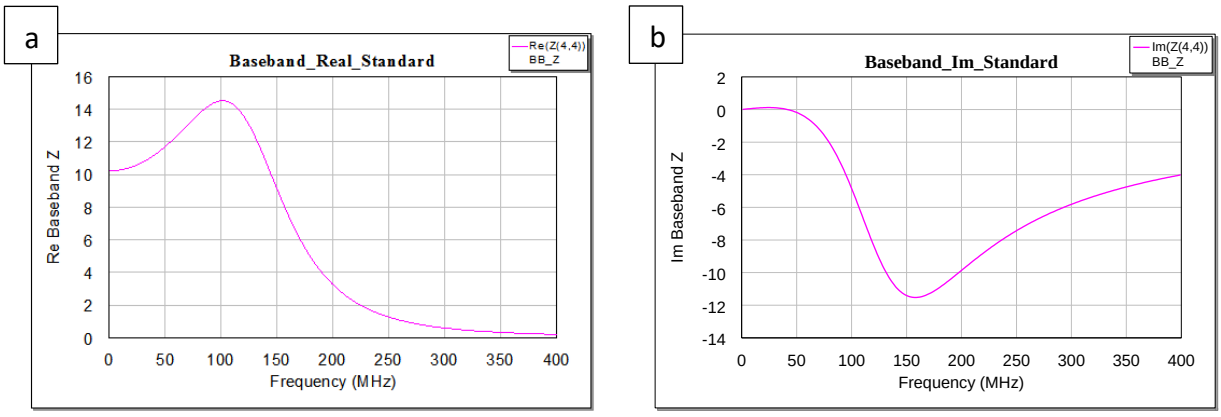

Figure 4.7 Baseband impedance: (a) real and (b) imaginary as seen from current generator.

Figure 4.7 shows that the matching topology shown in figure 4.6 presents a significant impedance variation across the baseband frequency range, where significant second and fourth-order intermodulation products would be generated.

4.2.2 A Novel Input Matching Topology with Enhanced Input Signal Bandwidth

The properties required from the internal input matching network are twofold: -

- To transform the very low gate impedance at RF frequencies (due to the high device C_{in}) to an impedance to which the device can be matched at the circuit board level.
- To provide a very low baseband impedance, in order to minimize the magnitude of the even order non-linear terms generated within the baseband frequency range due to device

non linearity, as demonstrated in figure 4.5. The impact of the baseband impedance on the magnitude of the even order terms generated can be seen in figure 4.11.

A novel input matching topology is proposed, whereby impedance transformation is performed at the RF frequency range, while simultaneously presenting a wideband, low impedance termination impedance across the baseband frequency range. Figure 4.8 shows the schematic for the proposed topology.

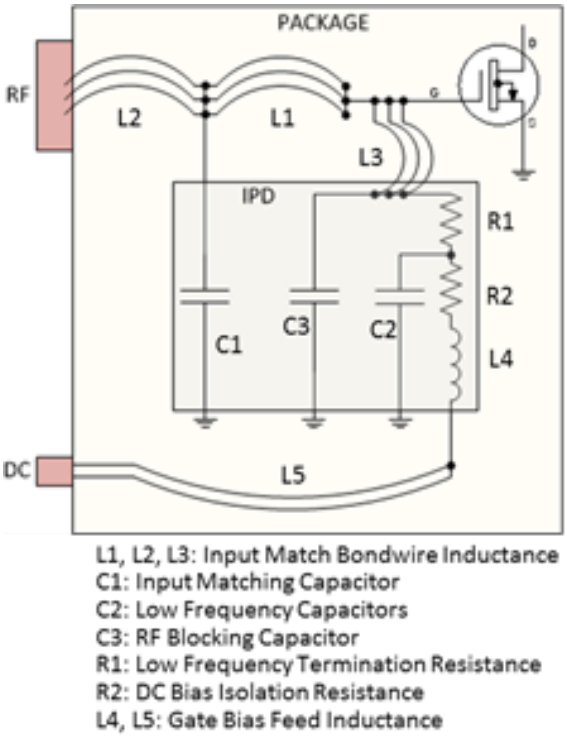

Figure 4.8 Proposed input matching topology.

The low pass matching topology in figure 4.6 is replaced with a high pass network. RF matching is dominated by L3, with C3 presenting a low impedance at RF frequencies (typically 100-300pF). In isolation, this network will only decrease the resonant frequency of the topology in figure 4.6, with the increase in capacitor value, reducing the baseband resonance from 102MHz to 68MHz. In order to mitigate this resonance, components R1, R2, C2, and L4 are added to the integrated

passive device (IPD), with L5 formed by a bond wire connection from IPD to package lead. Components R1 and C2 form a low-frequency bypass network, connecting to the RF short node, provided by C3, and provide a low impedance termination to the transistor gate. Components R2, L4, and L5 form a high impedance DC connection, from a bias feed pin on the package to the low-frequency decoupling capacitor within the IPD. These components negate the need for PCB mounted DC bias components. Figure 4.9 shows the internal image of the new transistor with the IPD. For the implementation shown, C1 and L2 were not included since PCB matching impedances were achievable without the need for additional internal matching sections.

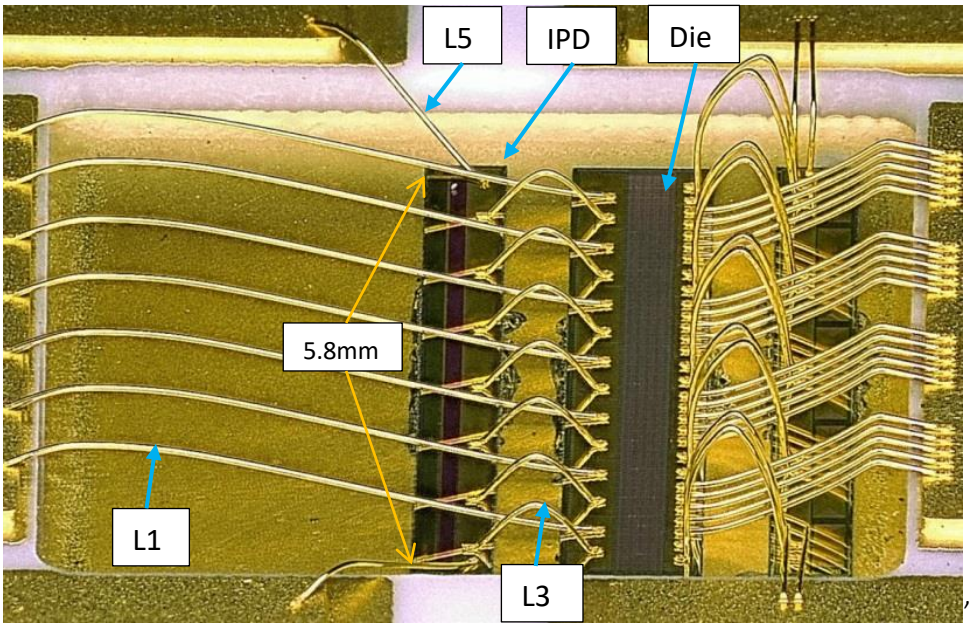

Figure 4.9 LDMOS transistor with proposed internal matching network.

Figure 4.10 shows the comparison of the baseband characteristics of the new device shown in figure 4.9, from the perspective of the current generator, alongside the response of the traditional input matching network shown in figure 4.6. It can be clearly seen that a very wideband, consistent low baseband impedance, is now presented at the gate of the transistor using the proposed matching topology.

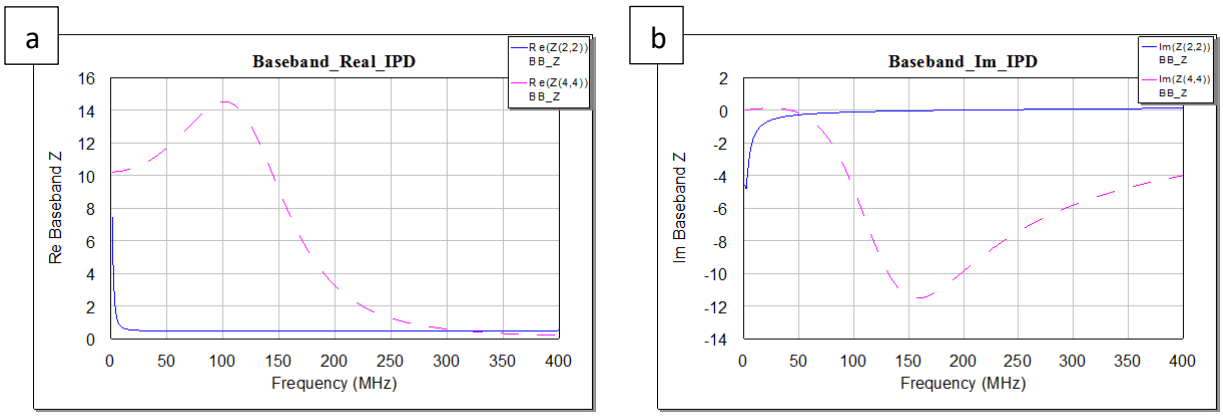

Figure 4.10 Baseband impedance: (a) real and (b) imaginary as seen from current generator.

Figure 4.11 shows the impact of the difference in baseband impedance between the standard and the new device in terms of IMD generation. The new device presents a reduced impedance termination to the 2nd order terms generated by the nonlinear input capacitance under two-tone RF excitation (20MHz tone spacing). This impedance reduction results in a corresponding reduction in the magnitude of the distortion products at the gate node. Figure 4.11 shows the magnitude of the IMD at the gate node for both (a) the standard device, and (b) the new device.

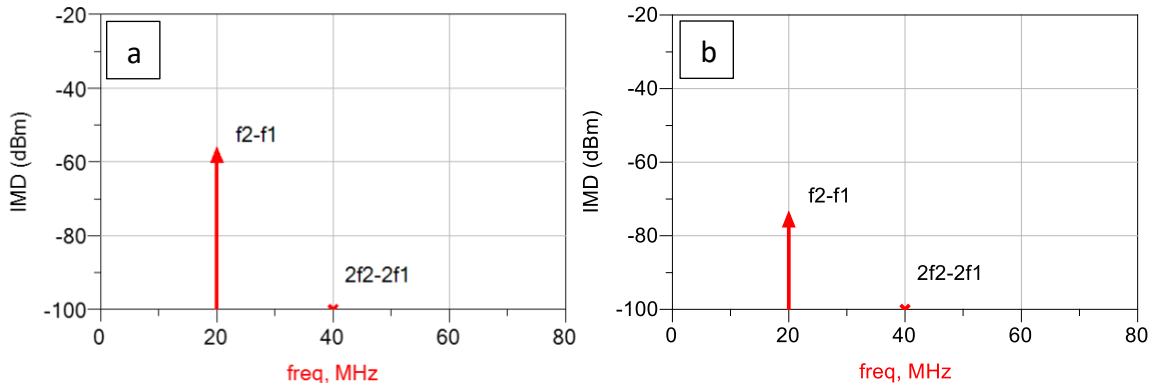

Figure 4.11 Input baseband distortion terms at gate node: (a) standard device, (b) proposed device.

The core LDMOS transistor has significant available gain in the baseband frequency range. Consequently, a significant amplification of baseband distortion terms appearing at the gate node of the FET gate occurs. The amplified input distortion terms appear as a low-frequency modulation

at the drain of the transistor, with a magnitude dependence proportional to the output baseband impedance. Figure 4.12 shows the resultant output terms for both devices, each presented with an equal output baseband impedance.

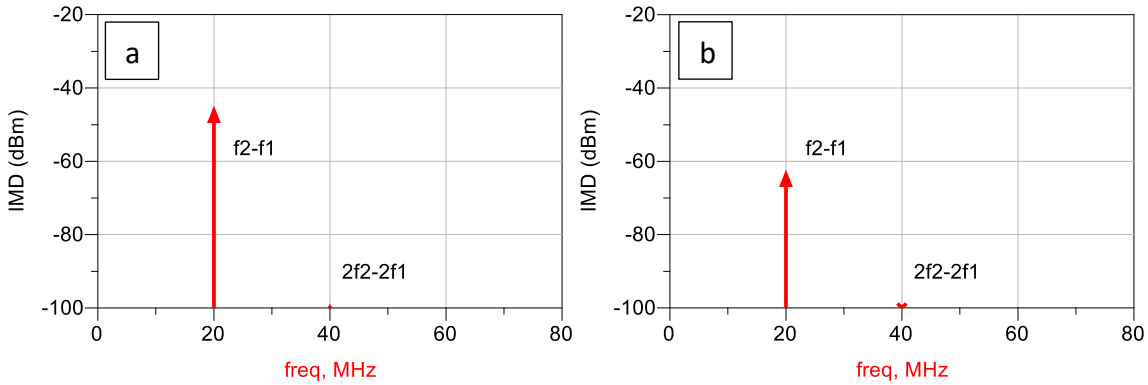

Figure 4.12 – Output baseband distortion terms at drain node: (a) standard device, (b) proposed device.

Figure 4.12 shows that the generation of intermodulation products through input nonlinearity manifests as an additive signal to any nonlinear terms generated through other mechanisms, such as nonlinear transistor transconductance, or nonlinear C_{ds}.

4.2.3 Device Simulated DPD Comparison

Before running hardware evaluations, the proposed input matching network is compared with a traditional matching network using a simulated DPD test bench in Keysight ADS. The test bench represents the hardware blocks used for the measured DPD results in section 4.2.2.3. Figure 4.13 shows a simplified block diagram for the system.

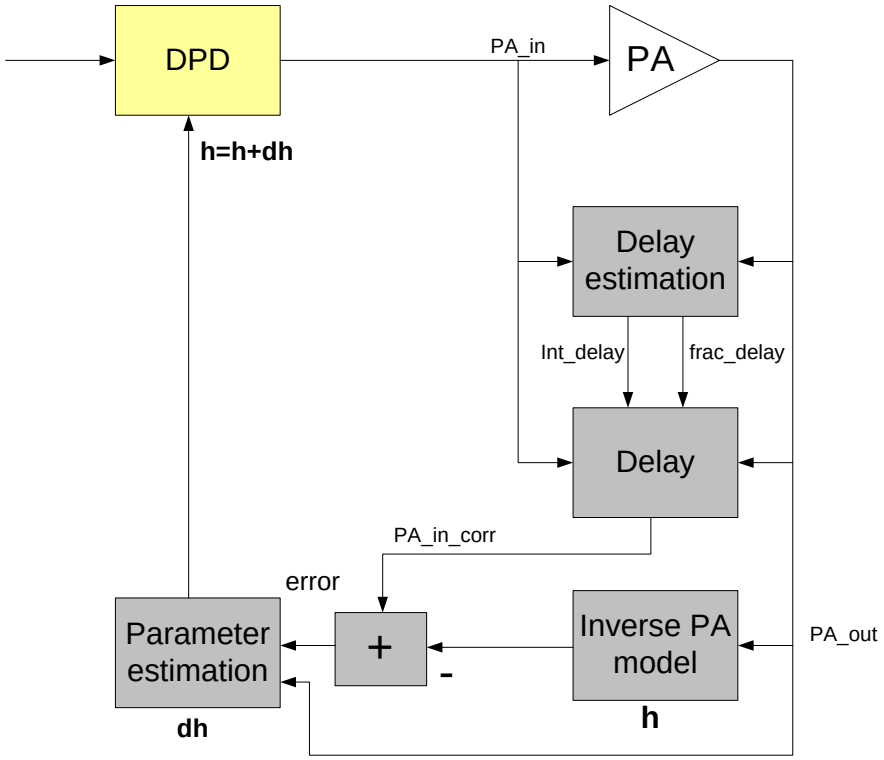

Figure 4.13 Simulated DPD block diagram.

The DPD uses an indirect learning architecture, whereby the PA output is fed into an inverse PA model, and a comparison made between the output and reference input signal. The model is iterated based on the resultant delta term (labeled 'error' in 4.13) until the model converges to a minimum delta between the PA Out through the inverse model and the reference input signal. Figure 4.14 shows the resultant error term for both devices vs. time. As is seen in the plot, the proposed device converges to a delta 5dB lower than the standard device.

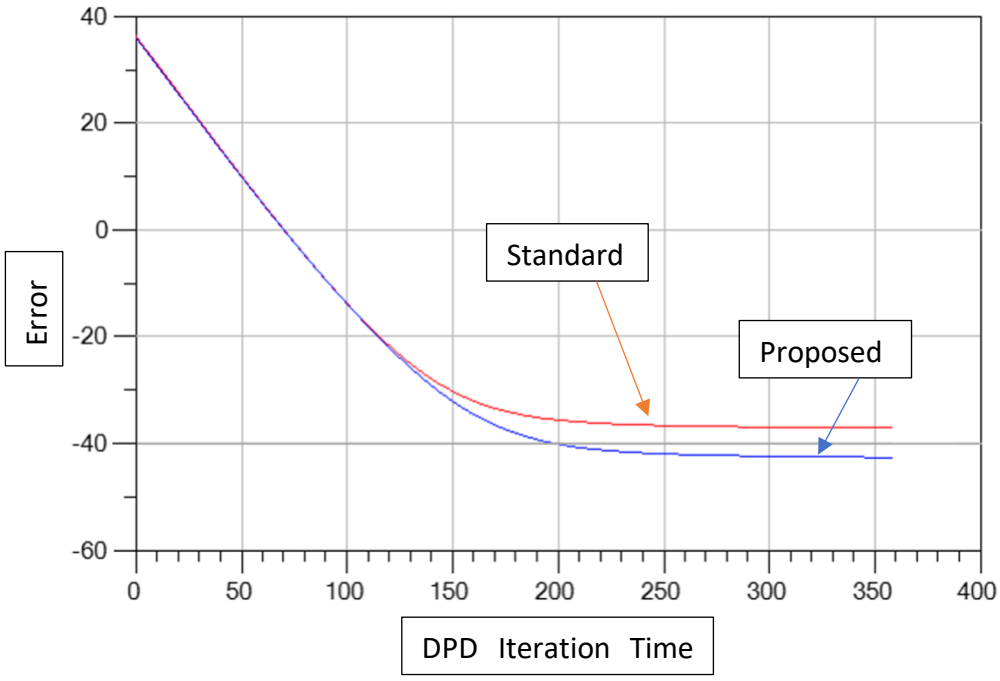

Figure 4.14 – Delta between inverse PA model output and reference signal vs. iteration time.

Figure 4.15 shows the simulated DPD results using a 20MHz bandwidth, 7dB PAR signal, with the linearity results summarized in table 2. The reduction in the input baseband impedance and, therefore, input even-order distortion products shows corresponding improvements in the simulated DPD results, with an average improvement of 4.75dB in ACLR.

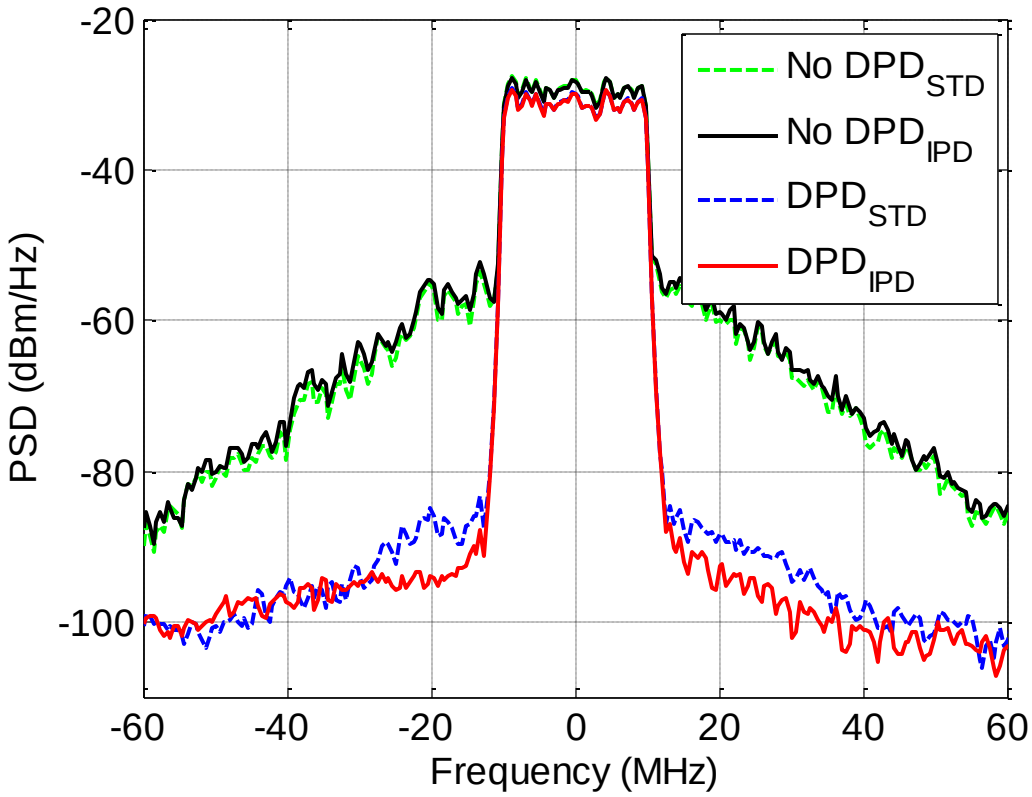

Figure 4.15 Simulated DPD, 20MHz signal bandwidth, 7dB PAR.

	ACP lower No DPD	ACP upper No DPD	ACP lower with DPD	ACP upper with DPD
Standard Device	-43.5	-43.4	-56.4	-57.2
This Work	-41.8	-41.7	-62.0	-61.1

Table 2 – Simulated DPD ACP Summary, 20MHz signal bandwidth, 7dB PAR.

4.2.4 Device Measurement Comparison

In order to assess the practical impact of the improved baseband impedance characteristics of the proposed device topology, two devices and circuits are designed and compared. The output matching network for both devices (both internal and external) is consistent, the only change being to the internal input match, and the corresponding PCB match designed to match the DUT impedances. Both circuits use the same baseband decoupling and output matching networks, and the circuits are measured to have <0.1dB delta in P3dB. Figure 4.16 shows photographs of both internal matching and reference circuits.

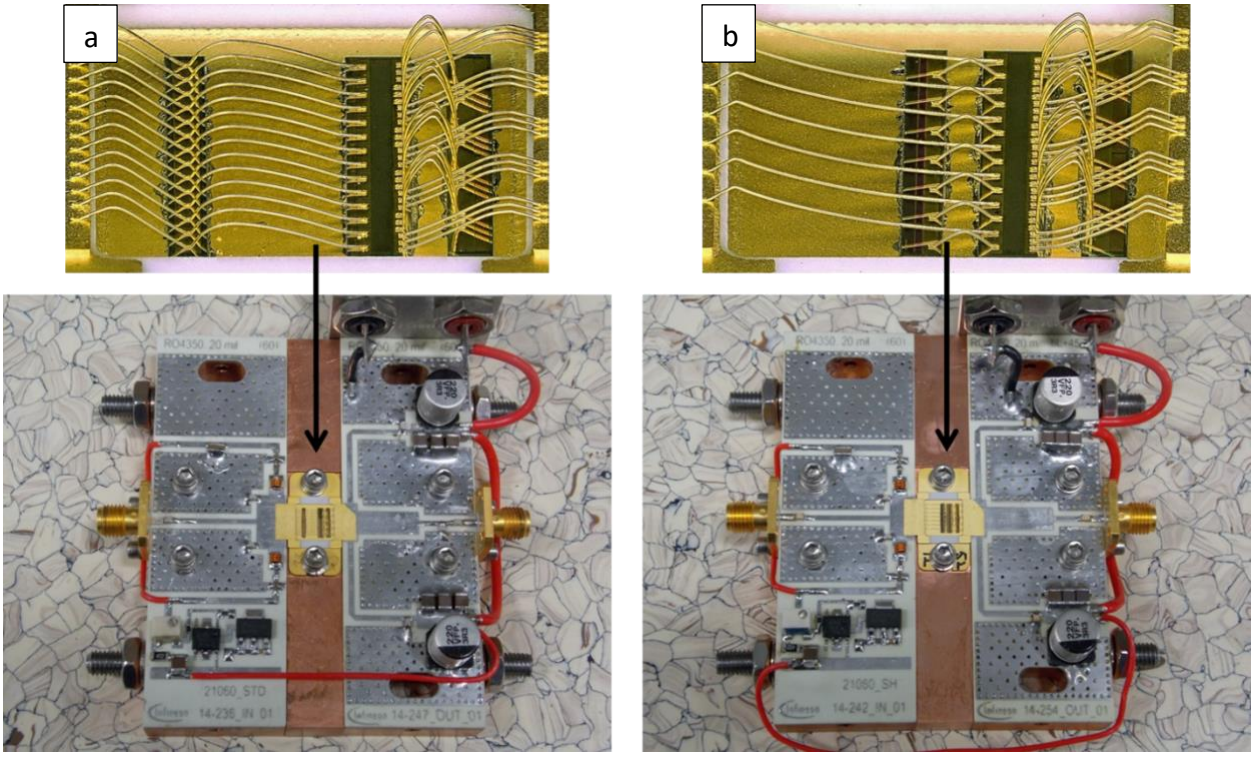

Figure 4.16 Device and circuit designs for DPD comparison: (a) standard device, (b) proposed device

As mentioned already, the load impedance tuning for the devices was identical, and both boards achieved a drain efficiency of between 59.5 – 62.5% efficiency from 1805-2170MHz (cellular bands 1,2, and 3). P3dB for the boards was flat at 47.4dBm (55 watts) across the same frequency range.

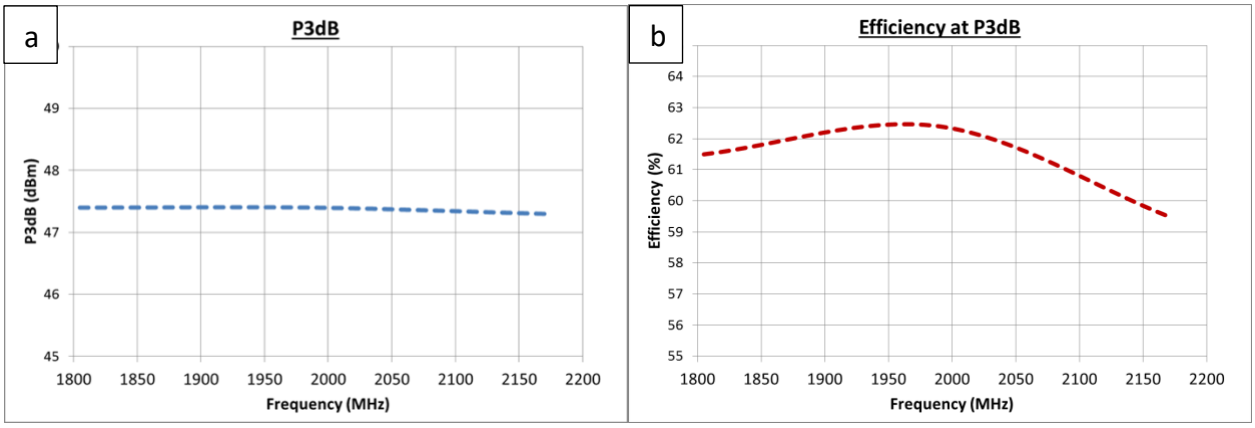

Figure 4.17 Circuit level performance: (a) P3dB, (b) efficiency at P3dB.

4.2.4.1 Device Measurement Comparison Differences

The input match of the proposed device provides out of band impedance characteristics well aligned with a single section external matching network. In the case of the standard device, the impedance location of the high frequency is closer to 50 ohms than the low frequency. Consequently, when matched using a transmission line section, impedance dispersion occurs since the high frequencies rotate faster than low frequencies. In the case of the proposed device, the internal matching network presents the reverse input impedance locations. Therefore, when rotating through a fixed electrical length matching network, the proposed device impedances converge as the high frequencies rotate toward the low frequencies.

Figure 4.18 shows the resultant input S11 for (a) the standard input match, and (b) the proposed input match.

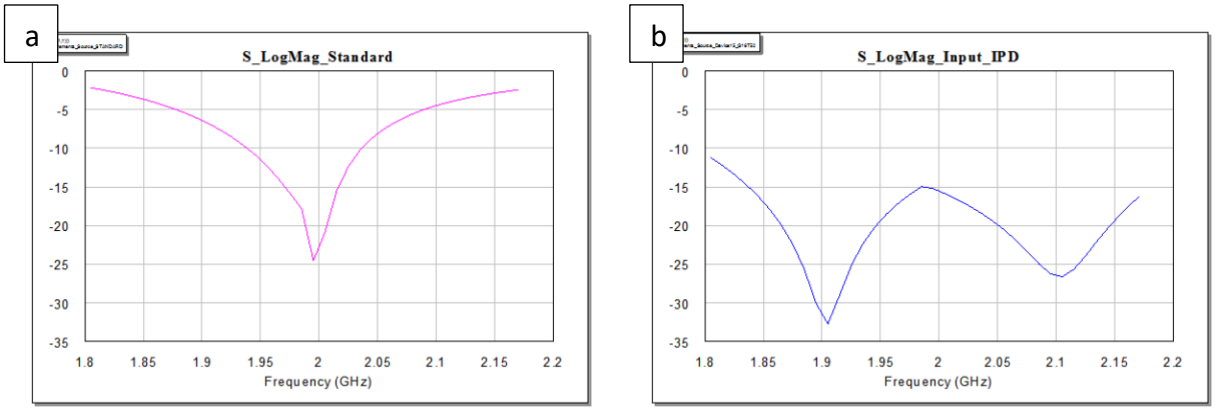

Figure 4.18 Measured input S11 for reference circuits at SMA Connector: (a) standard device, (b) proposed device.

The measured baseband characteristics, characterized by the S21 gain response of the device, showed a significant attenuation in the sub 500MHz baseband region. Figure 4.19 shows the peak baseband gain of the device is reduced by 39dB.

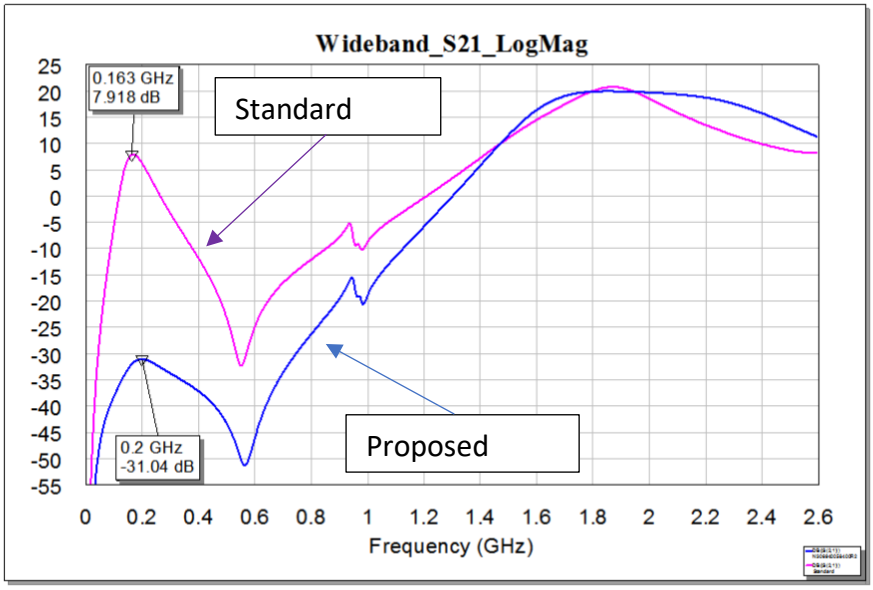

Figure 4.19 Wideband S21 for reference circuits at SMA Connector: (a) standard device and (b) proposed device.

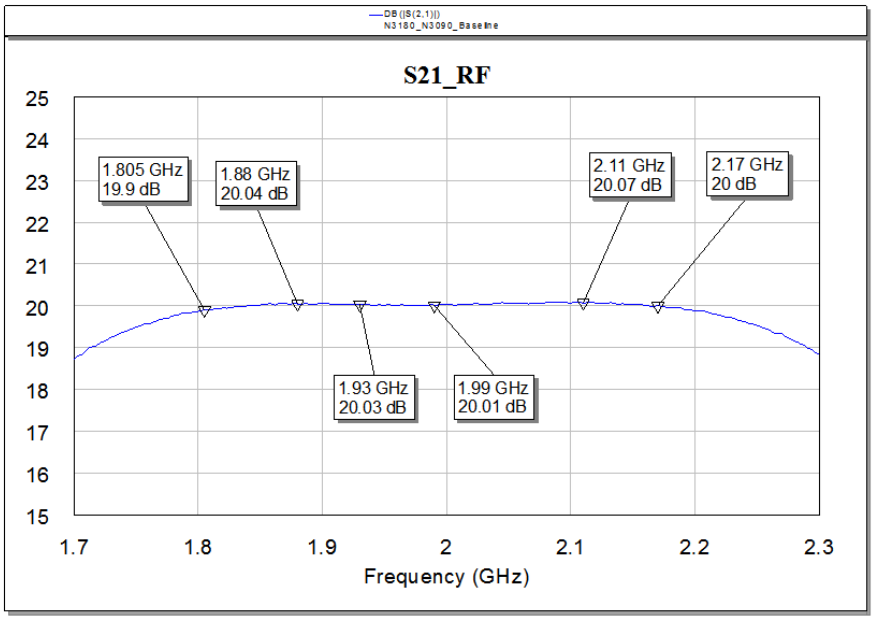

Figure 4.20 Measured gain response of proposed device: 20dB gain with 0.17dB flatness across 365MHz RF bandwidth.

4.2.2.3 Measured DPD Comparison

A DPD comparison of the two circuits was run on two systems in order to compare the linearized performance. The systems used were: -

- Texas Instruments reference circuit with a simple memory polynomial algorithm and 250MHz transceiver bandwidth.
- Optichron (now Broadcom) OP6100 reference circuit and 307MHz transceiver bandwidth

The TI system utilized a highly flexible algorithm with variable polynomial order and memory depth. The Optichron system enables enhanced linearization performance, but the DPD algorithm is a locked solution. Therefore, no comparisons of DPD coefficients can be made, so less insight is available from the measurements.

Initial results were taken using the TI system with a 40MHz LTE carrier. Figure 4.21 shows results for the standard and the proposed device at the same output power back-off from P3dB

compression. In this case, the DPD memory taps were disabled to enable observation of the residual memory effects in the RF circuits.

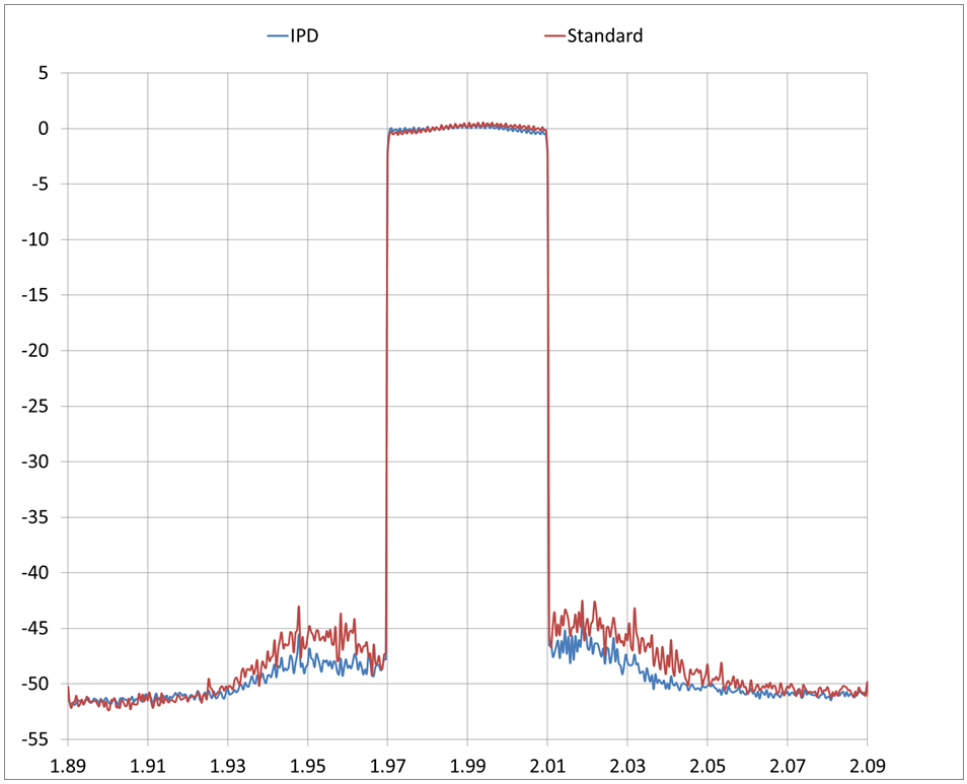

Figure 4.21 40MHz signal bandwidth, Texas Instruments DPD system, standard and proposed device after DPD correction.

A more extensive analysis vs. output back off from P3dB is run on the Optichron DPD system using a six carrier GSM signal. This test case was of high market interest at the time of evaluation, as the multi-carrier GSM linearity specifications proved extremely challenging to achieve with any DPD system at the time (-70dBc for a class 1 radio). Figure 4.22 shows that the proposed device settles at an IMD floor of -64.5dBc, compared to -62dBc from the standard device. This delta reflected the typical improvement of 3dB observed from systems during the evaluation of the matching topology.

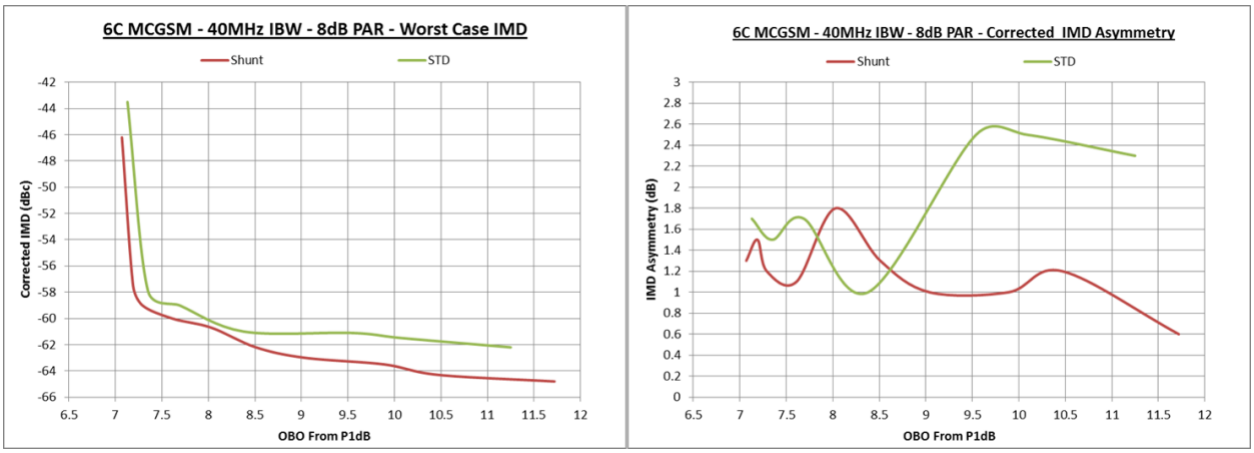

Figure 4.22 40MHz signal bandwidth, multi carrier GSM modulation, Optichron DPD system: (a) standard and proposed device linearity after DPD correction, (b) upper and lower IMD asymmetry after correction.

4.3 Summary

This chapter presents a novel input matching topology for a high-power RF transistor. The performance enhancements are quantified by comparing the device to a current state of the art LDMOS transistor.

The RF bandwidth enhancements were quantified by comparing reference circuits designed for each transistor. The reference device provided approximately 100MHz of RF bandwidth, whereas the novel topology extended the RF bandwidth to 400MHz, as shown in figures 4.18 and 4.20.

The novel topology also significantly reduced the input baseband impedance, shown in figures 4.10 and 4.19. This characteristic enabled improved linearized performance when evaluated in multiple digital pre-distortion platforms. The corrected performance is characterized in figure 4.21 (Texas Instruments DPD platform, 250MHz b/w) and 4.22 (Optichron OP6100, 307MHz b/w), with improved post linearization performance in each case.

Chapter five explores the efficiency degradation mechanisms within the RF power device. A detailed method to analyze the impact of device scaling, both in terms of power and frequency, is

presented, and a novel integrated passive device matching topology is shown. Enhancements achieved at the device level complement the amplifier topology efficiency and bandwidth enhancements presented in chapter three.

4.4 References

[1] Zhu, N., McLaren, R., Holmes, D., Holt, J., Rashev, P., Jones, J., "An integrated RF match and baseband terminations supporting 395 MHz instantaneous bandwidth for high power amplifier applications", IEEE MTT-S International Microwave Symposium (IMS), 2017.

[2] Barros, D., Nunes, L., Cabral, P., Pedro, J., "Impact of the Baseband Terminations on the Efficiency of Wideband Power Amplifiers Under Concurrent Band Operation", IEEE Transactions on Microwave Theory and Techniques, Vol.67, No. 12. December 2019.

[3] Takahashi, Y., Ishikawa, R., Honjo, K., "Precise Modelling of Thermal Memory Effect for Power Amplifier Using Multi-Stage Thermal RC-Ladder Network", Asia-Pacific Microwave Conference, December 2006

[4] Cripps, S.C. "RF Power Amplifiers for Wireless Communications," Artech House, pp. 203-210, 1999.

[5] Carvalho, N.B., Pedro, J.C., "Two-Tone Asymmetry in Microwave Power Amplifiers," IEEE MTT-S Digest 2000, pp.445-448.

[6] Carvalho, N.B., Pedro, J.C., "A Comprehensive Explanation of Distortion Sideband Asymmetries," IEEE Transactions on Microwave Theory and Techniques, Vol.50, September 2002.

[7] Schetzen, M., "The Volterra and Wiener theories of nonlinear systems," Krieger publishing company 1980.

[8] Vuolevi, J., Rahkonen, T., "Distortion in RF Power Amplifiers," Artech House, pp 43-68, 2003.

[9] Morgan, D.R., Ma, Z., Kim, J., Zierdt, M.G., Pastalan, J., "A generalized memory polynomial model for digital pre-distortion of RF power amplifiers," IEEE Transactions on Microwave Theory and Techniques, Vol.54, No. 10. October 2006.

[10] Wang, G, "Baseband Linearization Schemes for High Efficiency Amplifiers," Cardiff University PhD Thesis, June 2019.

Chapter Five – Power and Frequency Scaling

In this chapter, the subject of achieving both high power and high frequency operation is discussed. The chapter is divided into three sections.

Section 5.1 starts with a look at the performance of the unit cell, the building block of the power transistor. Here, the impacts of scaling at the unit cell level are explored, as this forms the building block for the higher power transistor array analyzed later in the chapter.

In section 5.2, the performance of the unit cell when operated in a power transistor array is explored. A comprehensive method to analyze the full power transistor RF characteristics, as well as the operating characteristics of each individual cell in the power array is described. The section concludes with a breakdown of the power and efficiency degradation of a very high power 12 cell GaN power device operating at 2GHz.

Section 5.3 concludes the chapter by looking at methods to reduce the performance impact of the distributed effects through the use of very high Q integrated passive devices (IPDs). A novel implementation of the IPD is described, and simulation results for the design presented. The IPD is taped out using a glass substrate, however delays in the fabrication prevented the inclusion of measurement results of the 200W GaN transistor design covering 1.8 – 2.2GHz (cellular bands 1 and 3) in this thesis.

5.1 Power and Frequency Scaling Issues

The intrinsic FET transistor comprises of a semiconducting channel of width W and length L. The 'transistor action' – or current flow through the channel is controlled by the gate voltage V_{gs}. The power capability of the transistor is determined by two parameters: -

Drain voltage swing capability, described by: -

$$vds_{swing} = vds_{breakdown} - vds_{knee} \quad (5.1)$$

And drain current swing, described by: -

$$Id_{sat} = unit\ w\ Id_{sat} \times w_{channel} \quad (5.2)$$

It can be seen from (5.1) and (5.2) that power scaling is achieved in two ways: -

- Increase in device operating voltage
    - this approach requires technology development to increase the breakdown of the transistor, or switching to a wider bandgap technology.

- Increase in device current capability
    - Achieved without technology development by increasing the channel width (often referred to as 'gate periphery') in order to increase Id_{sat}. This can be approached in two ways.
        - Increase the channel with of the unit cell.
        - Connect multiple cells in parallel, with the same channel width, therefore increasing the effective channel with of the complete transistor.
        - Examples of these two approaches are shown in figure 5.1(b) and 5.1(c)

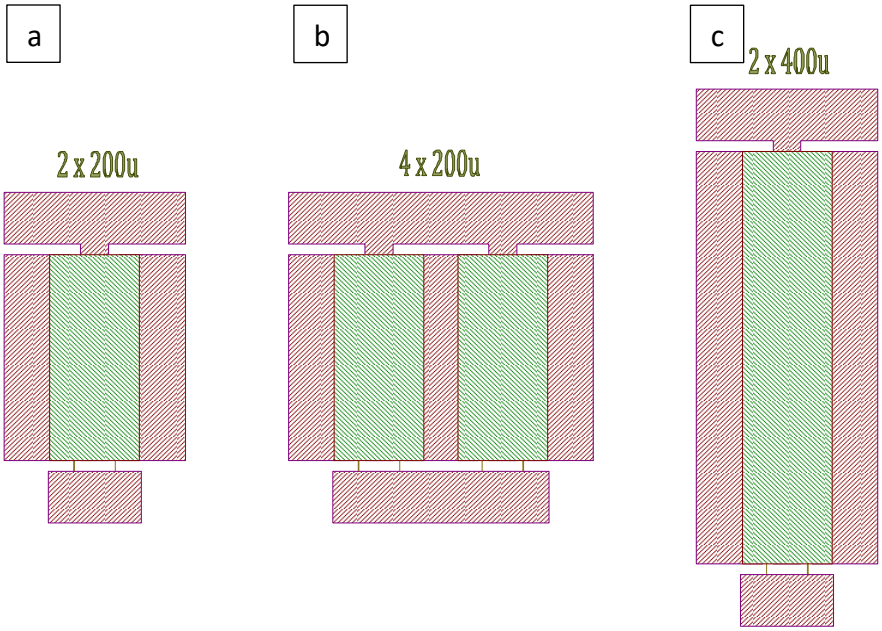

Figure 5.1 Physical implementations of device scaling: (a) 400um cell comprising 2 x 200um devices, (b) 800um cell comprising 4 x 200um transistors, (c) 800um cell comprising 2 x 400um transistors.

At low frequencies, both scaling methods provide the same performance, and the layout selection may pertain to practical considerations such as package size, number of chips per wafer, thermal performance etc. For operation at RF and microwave frequency ranges however, the subtle differences in the device characteristics start to have a large impact.

$$f_{max} = \frac{\text{ft}}{\sqrt{2((R_i + R_s + R_g)g_d + 2\pi f_t R_g f C_{gd})}} \qquad (5.3)$$

It can be seen from (5.3) that the f_{max} of the transistor is inversely related to the gate resistance R_g. Figure 5.2 shows a typical FET device layout. It can be seen that the gate is constructed of a thin metal 'finger', spanning the width of the transistor channel.

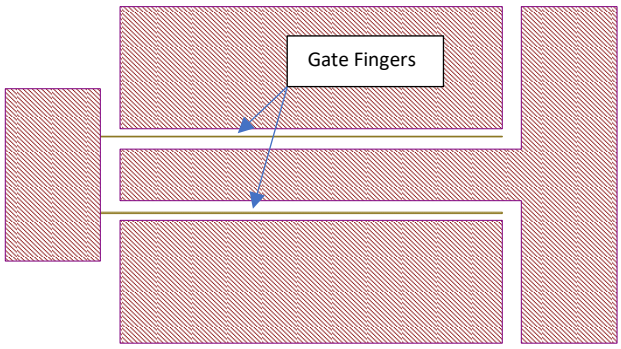

Figure 5.2 Simplified transistor metal layout showing thin gate fingers.

In considering typical dimensions, a finger length of 1mm, and the gate length (or finger width) of 0.5um would result in a series resistance of 24.4Ω when implemented in 2um thick gold.

Figure 5.3 shows how the gate resistance can be significantly reduced when the device layout uses shorter gate fingers in parallel. Simulated ESR values for a fixed 1mm total finger length, implemented in lengths of 1mm (ESR=24.4Ω), 2x500um (ESR=6.1Ω), 5x200um (ESR=0.98Ω) and 10x100um (ESR=0.24Ω) are shown.

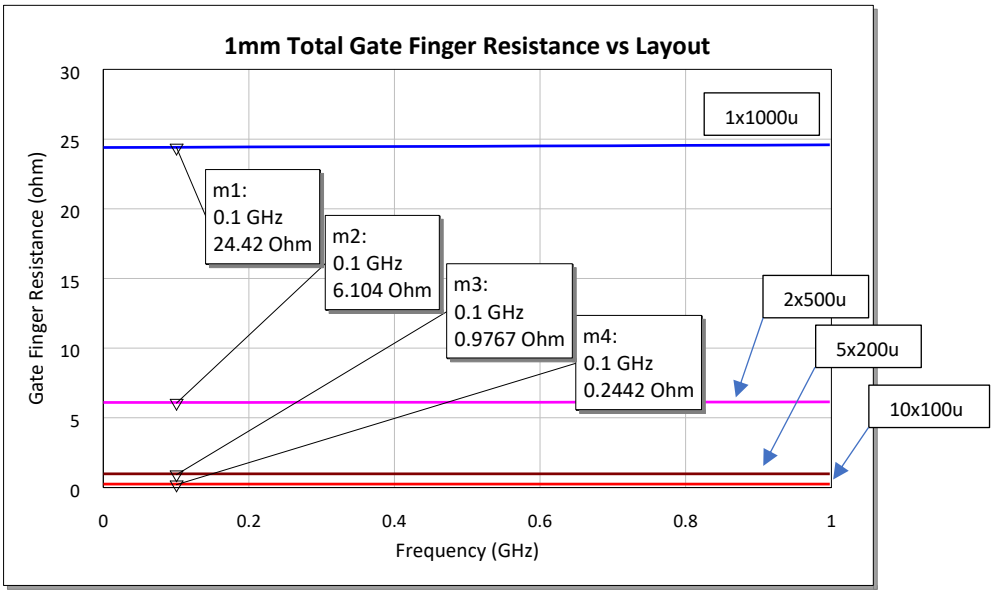

Figure 5.3 Series gate resistance of a 1mm transistor vs. gate finger configuration.

It should be noted that considerable layout optimization is performed on RF power transistors to mitigate the impact of the gate series resistance. This includes technologies such as T or mushroom gates, whereby high-level thick metals are placed above the thin gate structure, or the use of gate runners as shown in figure 5.4.

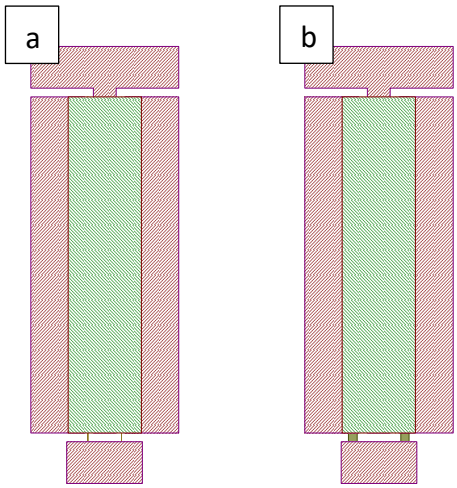

Figure 5.4 FET layout: (a) without gate runner (b) with gate runner to reduce series resistance.

The impact of the gate resistance reduction on the high frequency performance of the RF transistor is shown in Figure 5.5. The figure shows the maximum available gain (G_{max}) for a transistor with fixed channel width of 2mm, implemented with finger lengths from 1000um (x2 fingers) to 100u (x20 fingers), versus frequency.

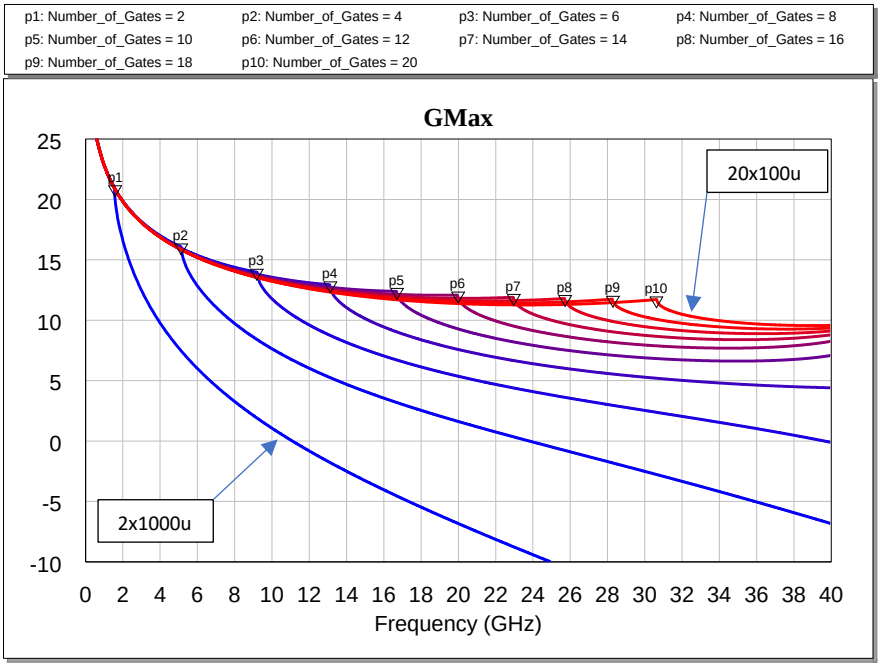

Figure 5.5 G_{max} vs. gate finger length.

Figure 5.5 shows the degradation in the high frequency performance of the transistor due to the series resistance of the gate finger with increasing length. It can be seen from the plot that the maximum finger length can be selected based around the target operating frequency of the transistor. An example of this is shown in figure 5.6, whereby the selection of finger length at the target operating frequencies is analyzed and plotted.

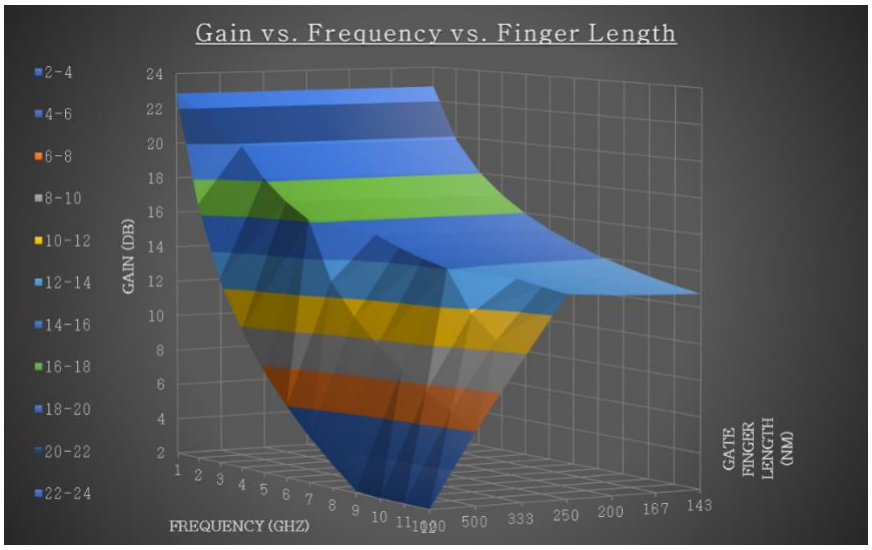

Figure 5.6 Maximum gain vs. finger length vs. frequency.

It can be seen in figure 5.6 that at 1GHz, the G_{max} of all finger lengths is consistent. However, as the frequency is doubled, a 3dB reduction in the available gain between a transistor comprising of 2x1000u fingers and 4x500u fingers is observed at 2GHz. At an operating frequency of 10GHz, it can be seen that a finger length of <= 250u is required in order to prevent a reduction in device G_{max}.

Figure 5.7 attempts to summarize the physical boundary conditions as the RF frequency and power of the transistor are scaled. As has been discussed in detail in this chapter, minimizing R_g is an essential requirement for high frequency operation. However, this trend inherently leads to a transistor with multiple parallel short fingers, and as power increases, the electrical length of the manifolds of the complete transistor become significant, and causing issues that will be discussed in section 5.2. Additionally, at the unit transistor level, a reduced gate length is well known to increase the F_{max} of the device due to reduced transit time. However, this approach also reduces the breakdown voltage of the transistor, thereby reducing the power per unit gate periphery. This leads to a higher gate periphery requirement to achieve the same output power, leading to path (2) in figure 5.7, whereby the need for a higher gate periphery and a short finger by default lead to a device with an electrically long drain manifold. A comprehensive analysis of this is investigated in section 5.2.

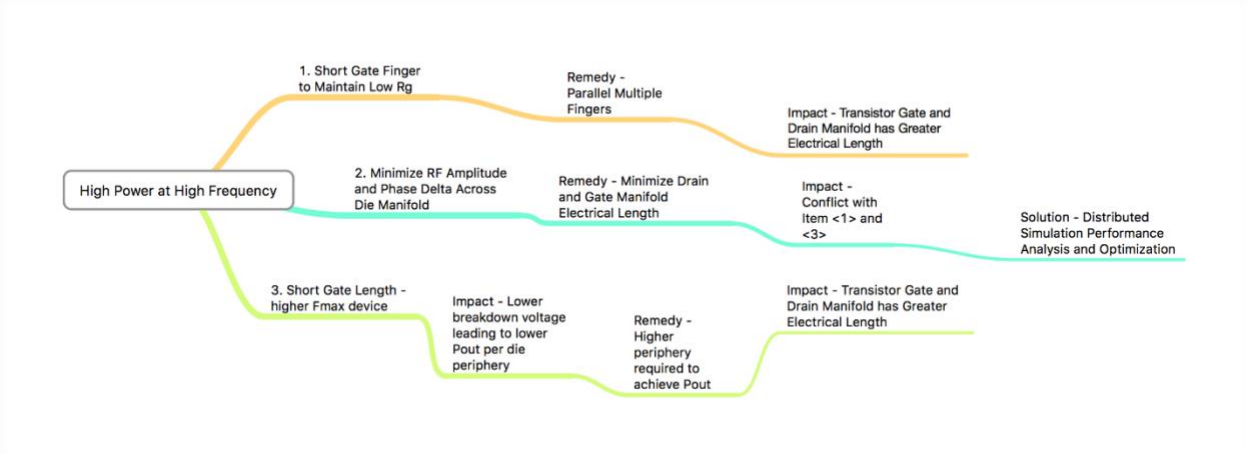

Figure 5.7 High frequency, high power device requirements and limitations.

5.2 Distributed Analysis of High-Power RF Transistor

Section 5.1 described the performance tradeoffs when scaling the power of the RF transistor. It was shown that the high frequency performance of the transistor is highly dependent on the input cut off frequency determined by the input R_g and C_g, and that maintaining a short gate finger length ensures the R_g term is minimized, therefore maximizing the high frequency performance of the device.

However, in optimizing the gate finger length of a single transistor cell for high frequency response, the electrical length of the complete power device will become significant at higher power levels. Figure 5.8 shows a power transistor with an input and output manifold width of 10mm. The corresponding frequency at which the manifold correlates to a quarter wave transmission line is also shown. Therefore, when analyzing the RF characteristics of such a transistor, the theory of a 'quantum dot' – a physical point or single impedance point – can no longer be assumed, since individual cells have significant phase delta between them, and therefore an impedance delta between cells is possible at the fundamental, and to a greater extend, at the harmonic frequencies.

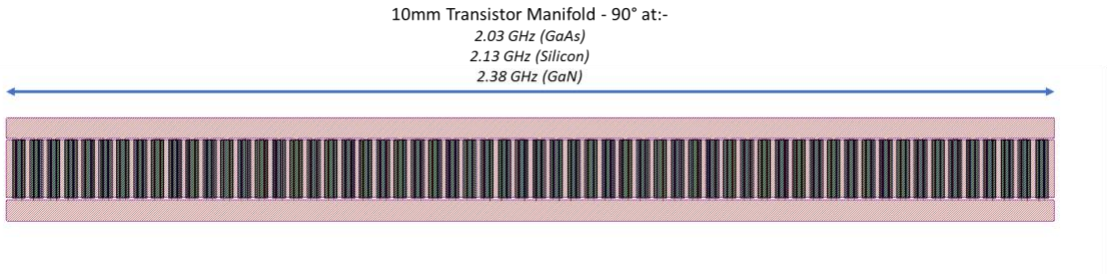

Figure 5.8 Electrical length of high-power transistor gate and drain manifold.

In this section, a method of analyzing the RF performance of a transistor with 'non-insignificant' electrical size is presented. The analysis will make use of 3D electromagnetic simulation of the die manifold, along with the surrounding package and matching elements. Harmonic balance simulation is used with multiple parallel die models to excite the EM structure and study the impact on the load conditions presented to cells along the manifold. The section concludes with the performance breakdown between the following: -

- Transistor die level performance
- Transistor die with wire bond and package losses simulated as single nodes
- Transistor die with wire bond and package losses simulated considering the distributed nature of the structure.

5.2.1 Harmonic Impedance Calibration

When analyzing and comparing the performance impact due to distributed effects of a of package or matching structures, the transfer function of the structure at both the second and third harmonic structure should be considered. Figure 5.9 shows the second and third harmonic load pull contours of a high-power RF device, specified at the transistor drain, with bond pad parasitic capacitance negated to enabling inclusion in later external electromagnetic simulation modelling. It can be seen that a reflection coefficient quadrant between +30° and -30° for both second and third harmonic loads provides high efficiency operation from this device. Therefore, when comparing matching structures, the appropriate second and third harmonic impedances should be placed externally in order to present a consistent impedance at the die level to ensure

meaningful comparisons of the transistor into a consistent operating mode. The method to achieve this is described in 5.2.3.

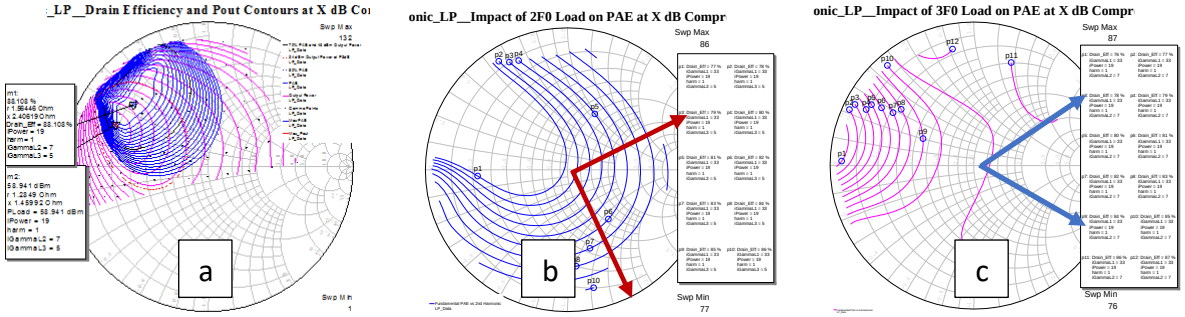

Figure 5.9 Simulated load pull contours: (a) fundamental, (b) second harmonic, (c) third harmonic

5.2.2 FEM Modelling of Die and Package

In order to capture the true distributed nature and interaction between die and surrounding matching structures, a multi technology 3D FEM co-simulation is run within a distributed load pull simulation using the Cadence AWR Analyst design environment. Figure 5.10 shows the structure used for much of the analysis in the section. It can be seen that the FEM simulation is from the die bond pads to a defined reference plane on the device package.

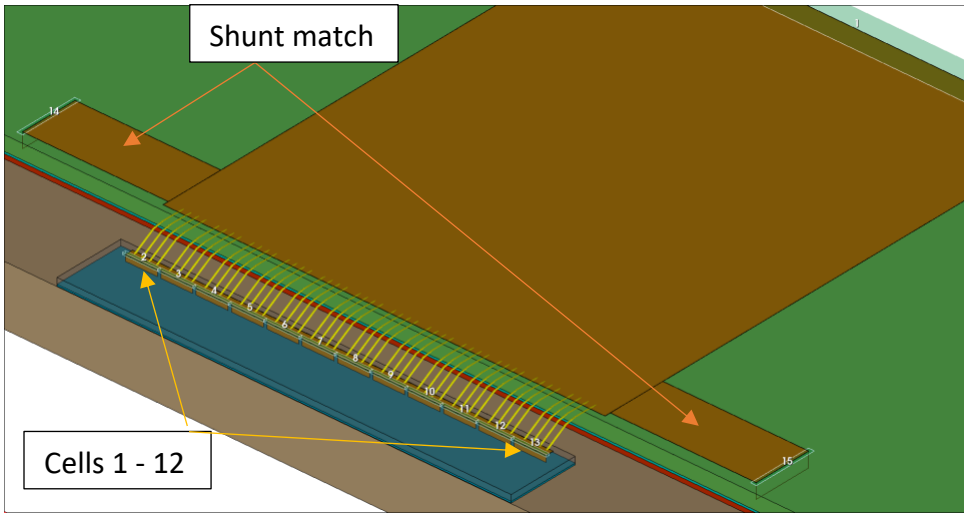

Figure 5.10 3D EM model of 12 cell RF transistor including PCB matching network.

The high Q shunt match, implemented on the device package, is depicted in figure 5.10. This element is parameterized in the EM simulation, and a multi variable sweep is run. This enables the impact of changes in the structure to be evaluated in the harmonic balance and load pull simulation domain using data from the electromagnetic simulations

5.2.3 Package Optimization for Load Modulated High Efficiency Amplifier

In order to optimize the in-package matching network for a load modulated amplifier, the network response presented at the current generator, including transistor parasitics, package response and Doherty combiner design requires consideration. The network can be broken down into the following elements: -

- *Current generator load impedance at full power level (Z_{opt}_R)*
- *Current generator load impedance at target back off power level (Z_{mod}_R)*
- *Device equivalent C_{ds}*
- *Bond wire series inductance (Ser_L)*
- *On package shunt match (Shunt_L)*

- *Doherty impedance inverter (including on package transmission line phase)*
- *Doherty amplifier summing node impedance*

In the above case, the die properties are extracted from die Id_{sat} and C_{ds} vs. V_{ds} measurements. The parasitics of the die to package interconnection is modelled as shown in figure 5.11, whereby the die bond pad, the (dominant) series inductance of the bond-wires, and the package capacitance below the bond foot is extracted.

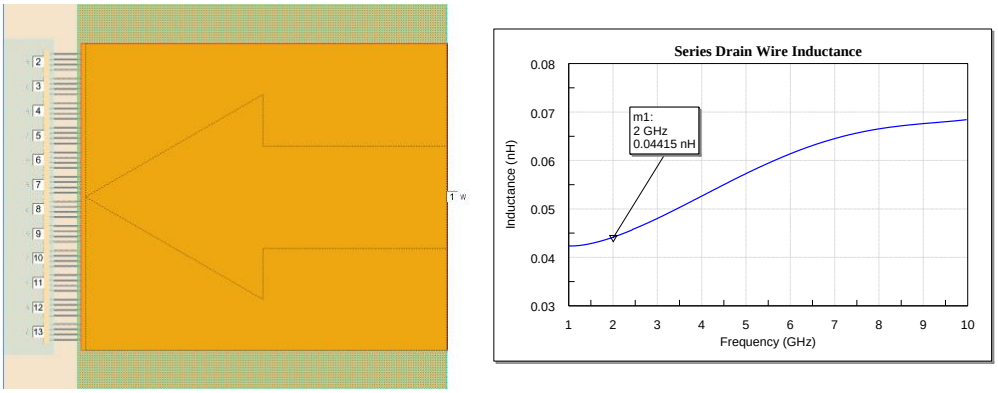

Figure 5.11 Die to package interconnection FEM modelling and extracted ESL.

In order to define the Z_{mod} resistance value, definition of the target Doherty output back off (OBO) peak is necessary. Figure 5.12 shows the relationship between target OBO, the main to peak scaling ratio and the corresponding load modulation seen by the main amplifier.

The relationships are described by: -

$$Peak\ OBO = 20.\log\left(1 + \frac{Pout_{pk}}{Pout_{main}}\right) \tag{5.4}$$

and

$$Zmain_{load} = 1 + \frac{Pout_{pk}}{Pout_{main}} \tag{5.5}$$

Where Peak OBO is the theoretical peak efficiency of the Doherty amplifier (plotted in figure 5.12 vs. Doherty asymmetry) and $Zmain_{load}$ is the load modulation ratio presented to the main amplifier prior to peaking amplifier conduction.

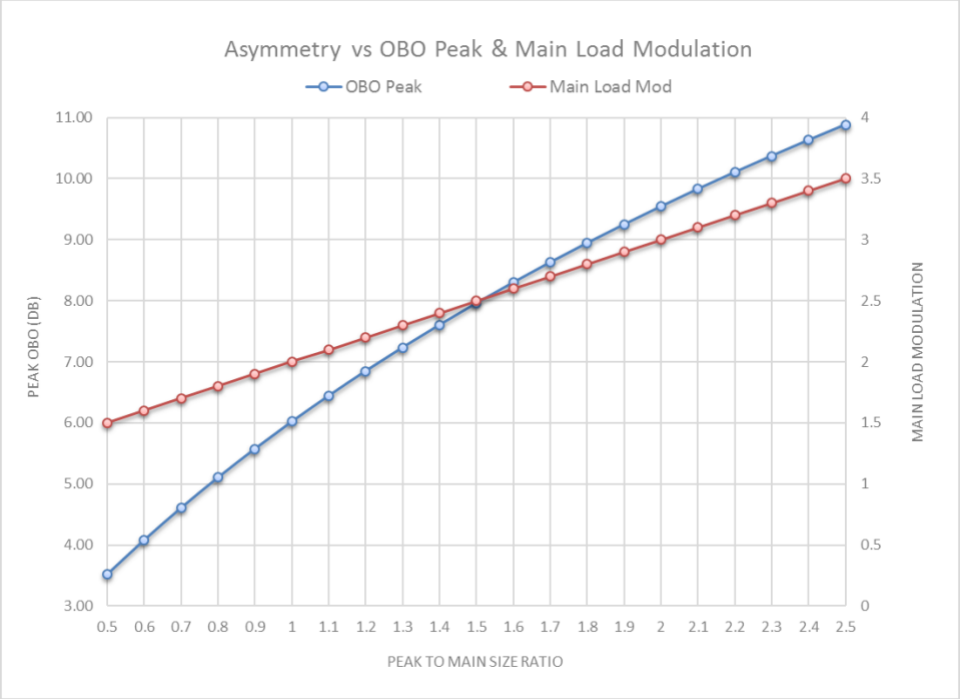

Figure 5.12 Relationship between Doherty OBO efficiency peak, main to peaking amplifier power ratio, and the main amplifier load modulation.

For a Doherty OBO peak of 9.5dB (a typical peak to average ratio for infrastructure digital modulation formats), the main amplifier load modulation of 3:1 will be presented, yielding the following target parameters: -

- *Zopt_R =* *3.5Ω*
- *Cout =* *16.6pF*
- *Zmod_R =* *10.5 Ω*
- *Ser_L=* *44pH*

High power RF transistors, particularly those based around silicon LDMOS with inherently high C_{ds} to P_{out} ratio, have traditionally utilized internal shunt matching networks to resonate with the output capacitance at the frequency of operation. Ideally, the reactance present between the device parasitic C_{ds}, and the resonating inductor should be negligible at the frequency of operation, since the fixed reactance of the series inductance will be a higher ratio of the load impedance as the load is reduced. Therefore, as the series reactance becomes non-negligible, the simultaneous matching of both Z_{opt} and Z_{mod} conditions becomes degraded. This effect is illustrated in figure 5.13, whereby the matching inductance required for the device with C_{ds}=16.6pF is plotted vs. load modulation and series inductance.

It can be seen in figure 5.13 that a progressively increasing delta in the required shunt L occurs under load modulated conditions as the series inductance between die and the shunt matching network is increased. Therefore, in order to present optimal matching for a modulated load with a fixed shunt inductance, minimizing the reactance between die and matching network is critical in order to maximize bandwidth of load modulated amplifiers such as the Doherty.

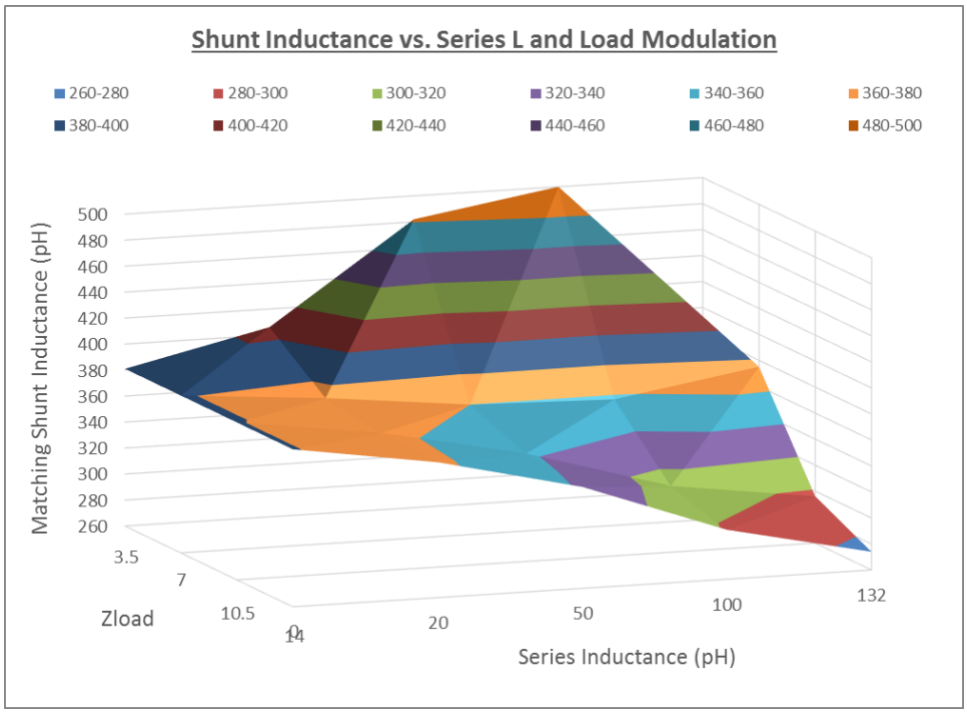

Figure 5.13 Shunt matching inductance vs. load modulation and series inductance.

The design of the on-package matching network requires wide RF bandwidth with low loss properties into both the Z_{opt} and Z_{mod} termination impedances. It is shown in figure 5.13 that the required value of the PCB shunt matching structure changes as reactance is placed between the summing nodes of the parallel resonator (comprising of C_{ds}, and the on-package shunt match). Therefore, in order to design the optimal matching structure, the bandwidth of device in the two different loading conditions is characterized. The schematic used for characterization is shown in figure 5.14, whereby the network is simultaneously terminated in two conditions: -

- Condition 1 (Doherty operating in back off)
    - network is terminated into the Doherty summing node impedance
    - current generator terminated into Z_{opt} * main amplifier load modulation. This load is referred to as Z_{mod}.
- Condition 2 (Doherty amplifier operating at full power)
    - network terminated into summing node impedance / load modulation
    - current generator terminated into Z_{opt}

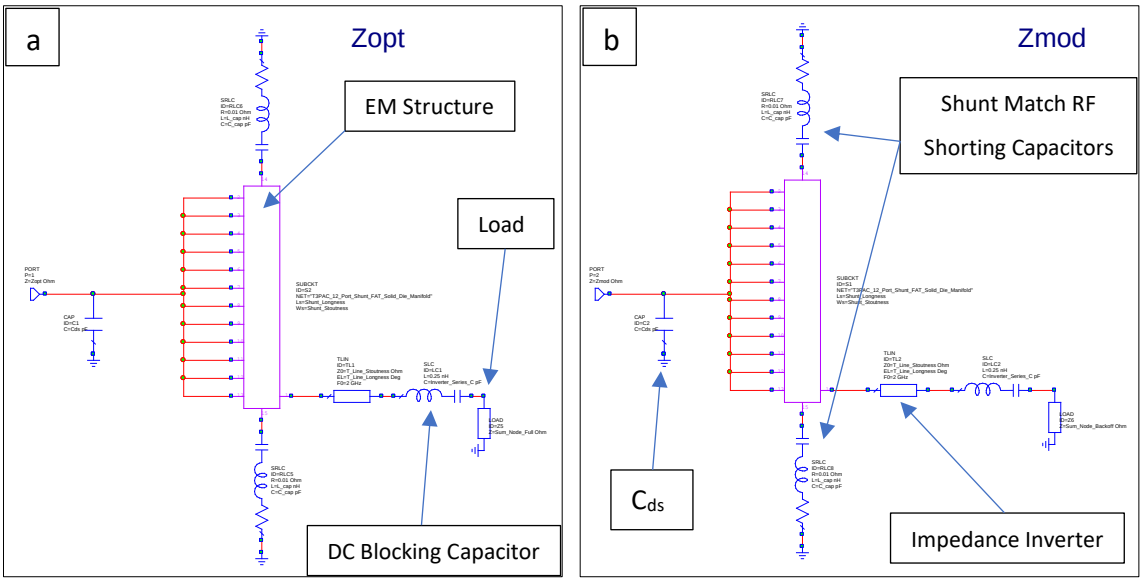

Figure 5.14 Device optimization schematic: (a) into full power load, (b) back-off power load impedances.

Figure 5.14 shows the schematic used for optimizing the PCB matching network into both full power and back off load impedances. In this case, optimization is run on the pre-computed parameterized electromagnetic data file, depicted in the center of each schematic, and containing parametric sweeps of the shunt matching branches in order to enable analysis of changes to the FEM structure in the circuit schematic.

In designing the Doherty structure, it should be noted there are multiple methods to achieve an operating goal. The design parameters used in the optimization of the above structure are as follows: -

- EM parametric data
    - shunt match electrical length
    - shunt match characteristic impedance
- Doherty impedance inverter
    - electrical length
    - characteristic impedance
- Doherty summing node impedance
- Doherty DC blocking capacitor

Upon completion of the design, optimization is run on the structure in order to evaluate all combinations of parameters. This is performed to validate the design, and ensure no alternate summing node / inverter impedance selections yield improved bandwidth.

Given the parasitic inductance of the bond wires present in the design has been simulated in isolation, and the target range of shunt inductance values is known, it is prudent to constrain the shunt matching network optimization to this range to minimize local maxima, as shown in figure 5.15.

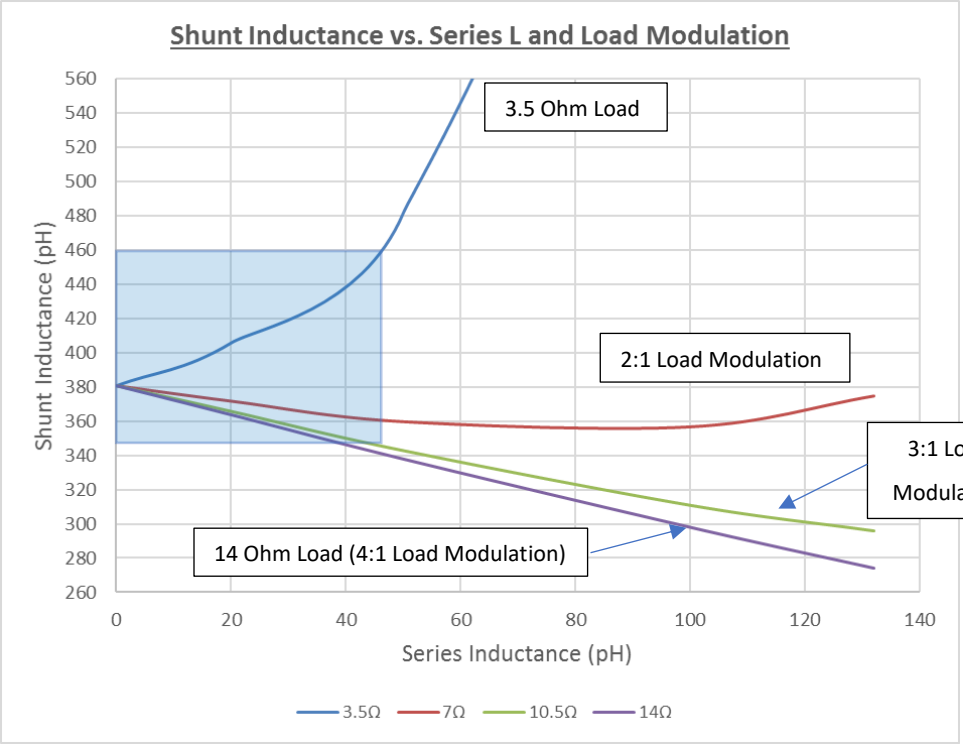

Figure 5.15 Shunt matching network optimizer constraints vs device load impedance.

Figure 5.16 shows the impedance across frequency presented at the current generator reference plane after completion of the optimization of the schematic in figure 5.14. The impedance trajectory of the full power (blue) and back off (pink), and the corresponding return loss are shown across 1GHz RF bandwidth.

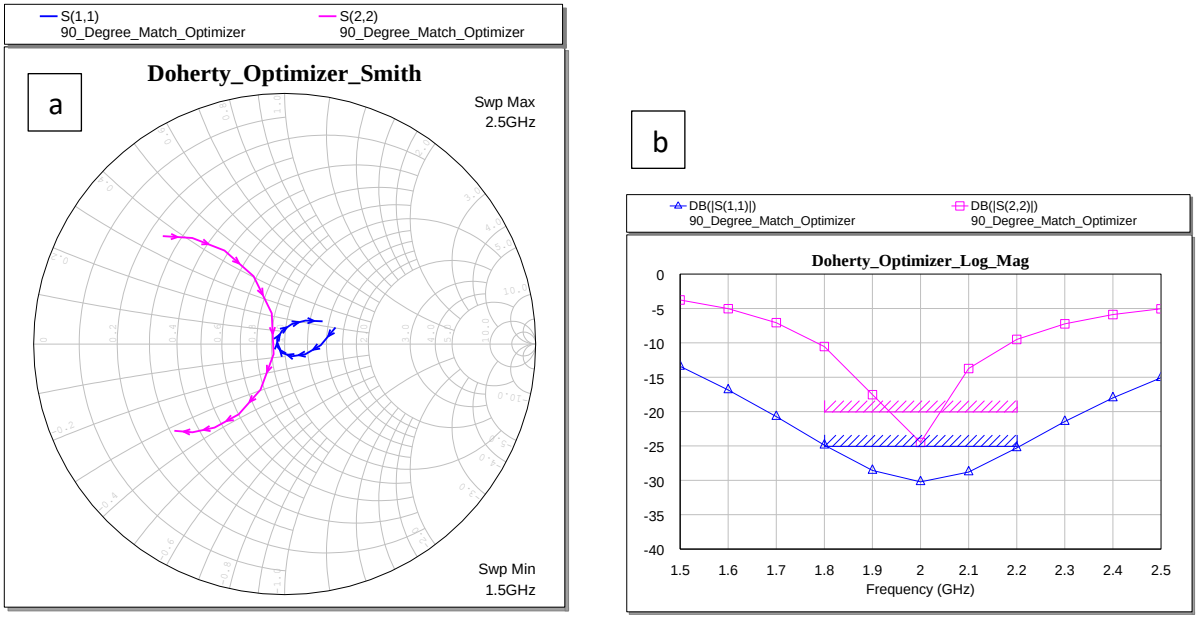

Figure 5.16 (a) Current generator load impedance trajectory at current generator, (b) return loss bandwidth.

5.2.3 Package Level 2f and 3f Harmonic Calibration

As mentioned in section 5.2.1, once the optimal parameters for the matching structure are found, calibration of external 2f and 3f impedances is required for any structure under evaluated in order to provide a consistent transistor operating mode. In this analysis, the target harmonic impedance Z parameters are recorded as a data file, and a delta function between the S parameters of this data file, and those looking into the EM structure is calculated. An optimization is then run to cover the external second and third harmonic impedances, to minimize the delta S function. The setup and results are shown in figure 5.17, with the results shown in figure 5.18.

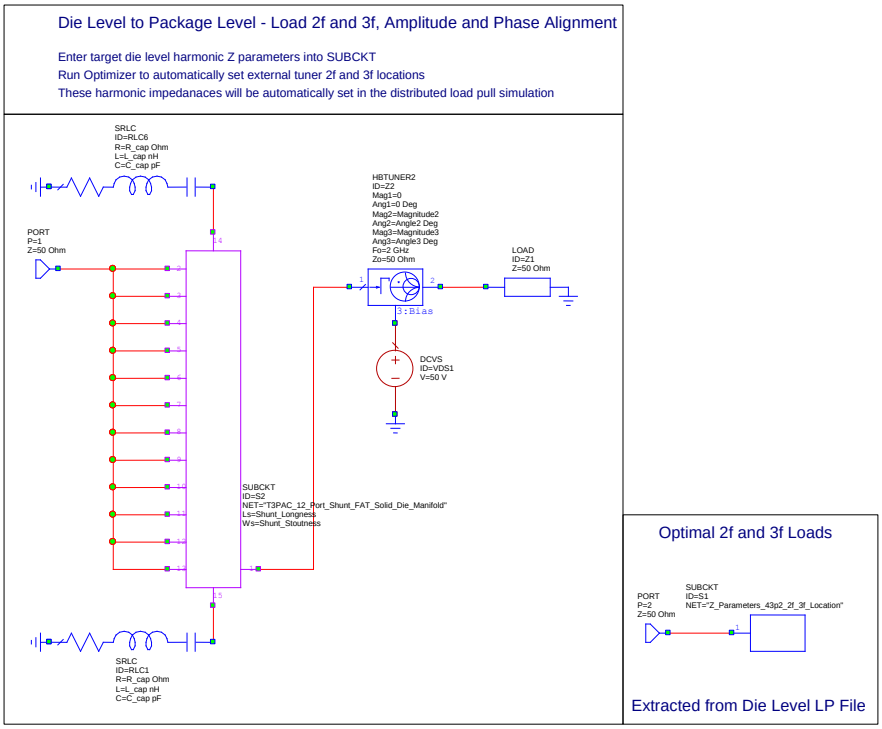

Figure 5.17 Die level harmonic calibration schematic.

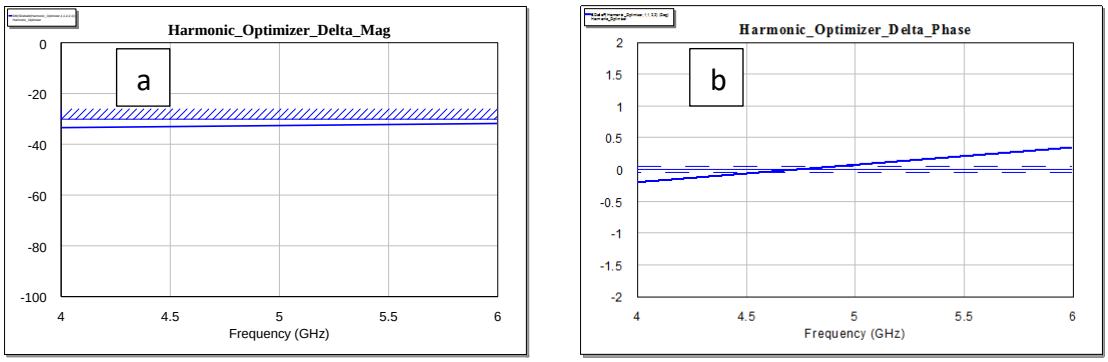

Figure 5.18 Delta function between target and presented die level harmonics: (a) amplitude, (b) phase.

Once the magnitude and phase of the external second and third harmonic impedances are found, a fundamental only load pull is run into pre-defined harmonic impedances, as is schematically shown in figure 5.19. This prevents the need for an additional fundamental, second harmonic, third harmonic load pull, as was performed at the die level in section 5.2.1. This is considered to be impractical step, since the simulation time of the fundamental only load pull on a high-end computer (12 core Intel Xeon E5-2687Wv4 @ 3GHz, 192G RAM) can take up to 30 minutes per frequency, and a nested harmonic sweep would multiply this time by the number of second and third harmonic impedance combinations measured.

5.2.4 Distributed Load Pull

Results from the distributed load pull simulation, shown in figure 5.20, enable selection of the optimal impedance for further waveform engineering analysis. Extraction of intrinsic current and voltage waveforms enables the derivation of both the fundamental and harmonic impedance delta across the cells in the array. This information provides an understanding of the practical observations seen for real high-power devices, such as imperfect power and efficiency scaling, and non-uniform thermal patterns across the unit transistor cells.

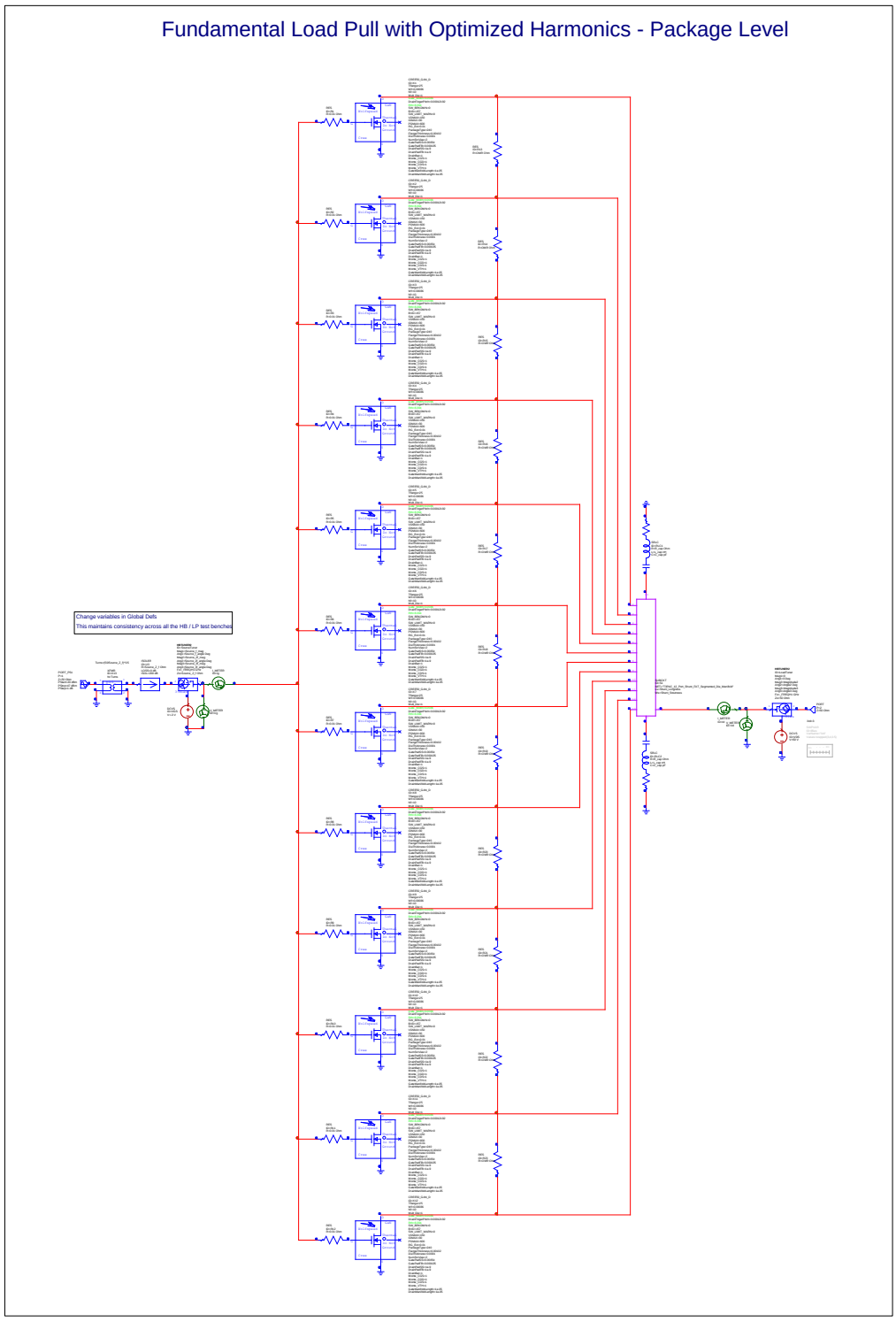

Figure 5.19 Distributed load pull schematic.

Results from the distributed load pull simulation, shown in figure 5.20, enable selection of the optimal impedance for further waveform engineering analysis. Extraction of intrinsic current and voltage waveforms enables the derivation of both the fundamental and harmonic impedance delta across the cells in the array. This information provides an understanding of the practical observations seen for real high-power devices, such as imperfect power and efficiency scaling, and non-uniform thermal patterns across the unit transistor cells.

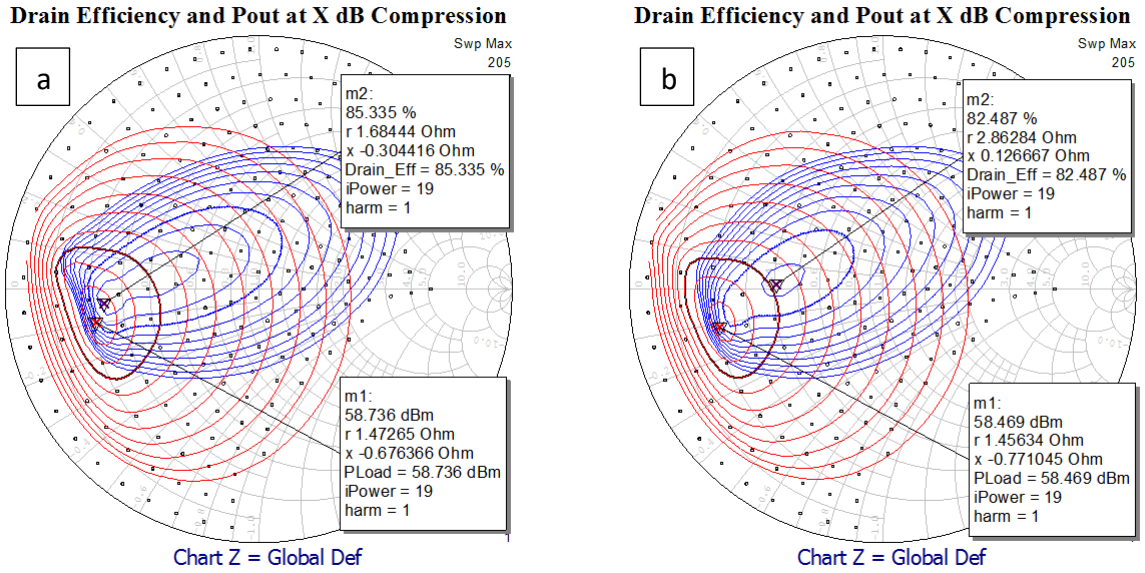

Figure 5.20 Package reference plane load pull results: (a) non distributed load impedance, (b) distributed load impedance presented to die model cells.

Figure 5.20 presents the load pull results for the schematic in figure 5.19 when run into a common load (all 12 input EM ports connected) and a distributed load (disabling the schematic shorts). In considering the peak efficiency load, a shift in both efficiency and the load impedance can be observed: -

Common Load	Efficiency = 85.3%	Z_{load} = 1.68-j0.30
Distributed Load	Efficiency = 82.5%	Z_{load} = 2.86+j0.13

5.2.5 Impact of Non-Linear C_{ds} on Individual Cell Fundamental Load Impedances

The impedances presented to the individual cells in the array are characterized into the global peak efficiency load conditions. Figure 5.21 shows the impedances presented to each cell in the array both extrinsically (die bond pad, left) and intrinsically (current generator, right) when excluding distributed effects from the simulation.

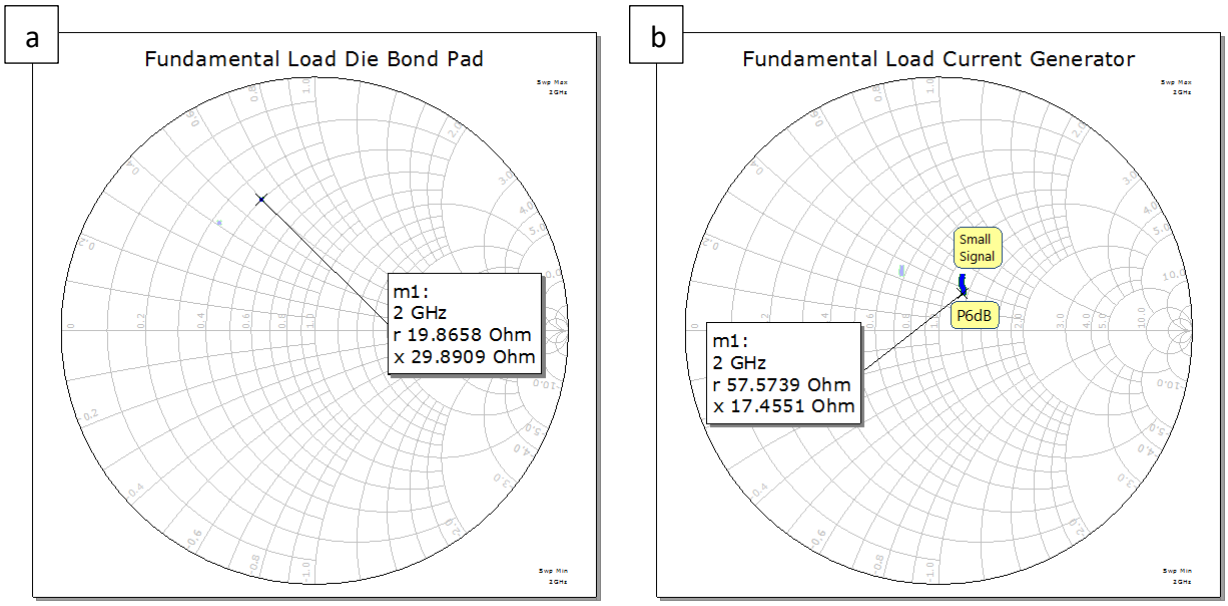

Figure 5.21 Fundamental load at peak efficiency: (a) bond pad reference plane, (b) current generator reference plane

It can be seen that as the reference plane is shifted between extrinsic level and intrinsic level, a load variation vs. drive is introduced [1]. This is seen as impedance dispersion in the current generator plots in figure 5.21 (b), and is explained when looking at the equivalent C_{ds} presented to the current generator by a nonlinear C_{ds} vs drain voltage swing. Figure 5.22 (a) shows a typical LDMOS transistor C_{ds} vs V_{ds} curve, and (b) the equivalent C_{ds} extracted vs drain voltage swing, centered around a 50v drain voltage.

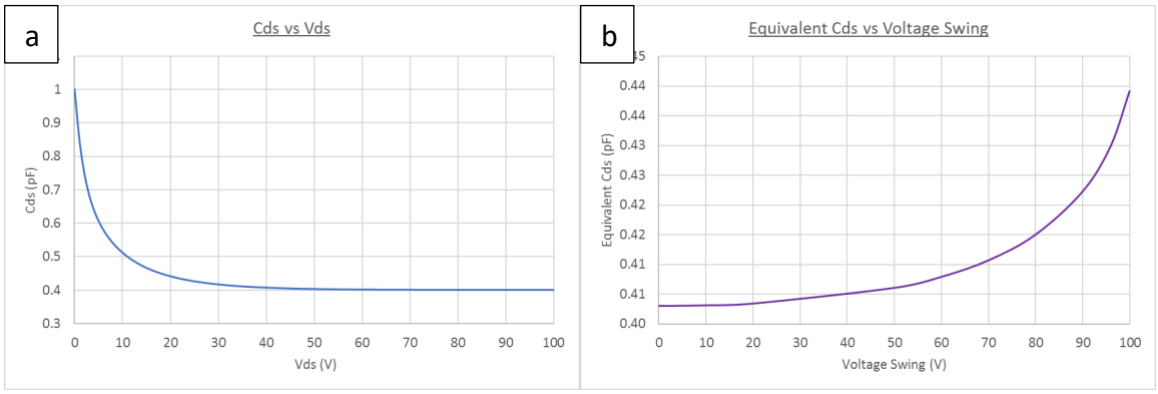

Figure 5.22 (a) Transistor C_{ds} vs. drain voltage, (b) equivalent linear device C_{ds} vs. drain voltage swing.

It can be seen by substituting the equivalent C_{ds} term from figure 5.22 into equation 5.6 that the current generator load impedance is a function of output voltage swing, and therefore output power when the device output capacitance in nonlinear.

$$Z_{load} = \frac{\mathrm{R}}{jwC_{ds} + 1} \tag{5.6}$$

5.2.6 Impact Distributed Effects on Individual Cell Fundamental Load Impedances

As discussed already, figure 5.21 shows the load impedances presented at both the extrinsic bond pad and the intrinsic current generator reference plane. It has been shown that an impedance shift vs. drive level is introduced at the current generator reference plane due to the nonlinear C_{ds} of the transistors due to the mean C_{ds} at a given drain voltage swing.

In this section, the analysis is extended to include the impact of the distributed load presented to each cell, extracted from the 3D FEM simulations. The impedance plots now reflect both the nonlinear C_{out} (as did figure 5.21), but also the impact of the amplitude and phase of the IV waveforms from the other 11 cells on the individual measured cell.

Two effects can be seen from the extracted cell impedances shown in figure 5.23: -

Firstly – it can be seen that the die exhibits cell load impedance symmetry – whereby the impedance presented to the mirror image cell about the center of the die is virtually identical. The mirror images pairs are overlaid in the 6 plots in figure 5.23, i.e. outer cells 1 and 12 on the die through to the inner cells 6 and 7, and the impedance trajectories over power can be seen to be overlaid.

Secondly, the since the load structure has significant electrical size and a distributed nature, an impedance divergence in the load presented to each cell along the transistor array, particularly at high power levels, is observed. It can be seen that the impedance presented to the outer cells (#1, 2, 11 and 12) is lower than the non-distributed load impedance (figure 5.21), whereas the impedances presented to the inner cells (#4, 5, 6, 7, 8, 9) is higher than the non-distributed load. The transition point is at cells 3 and 10, which see a load impedance close to the cumulative average (depicted as the light memory trace common to all plots in figure 5.23) from the non-distributed case.

The effect of the load impedances shown in figure 5.23 can explain the current and voltage waveforms observed in figure 5.24. It should be noted, however that the I-V waveforms will also be dependent on the harmonic impedances presented to the cells, characterized in figure 5.25 (second harmonic) and 5.26 (third harmonic load).

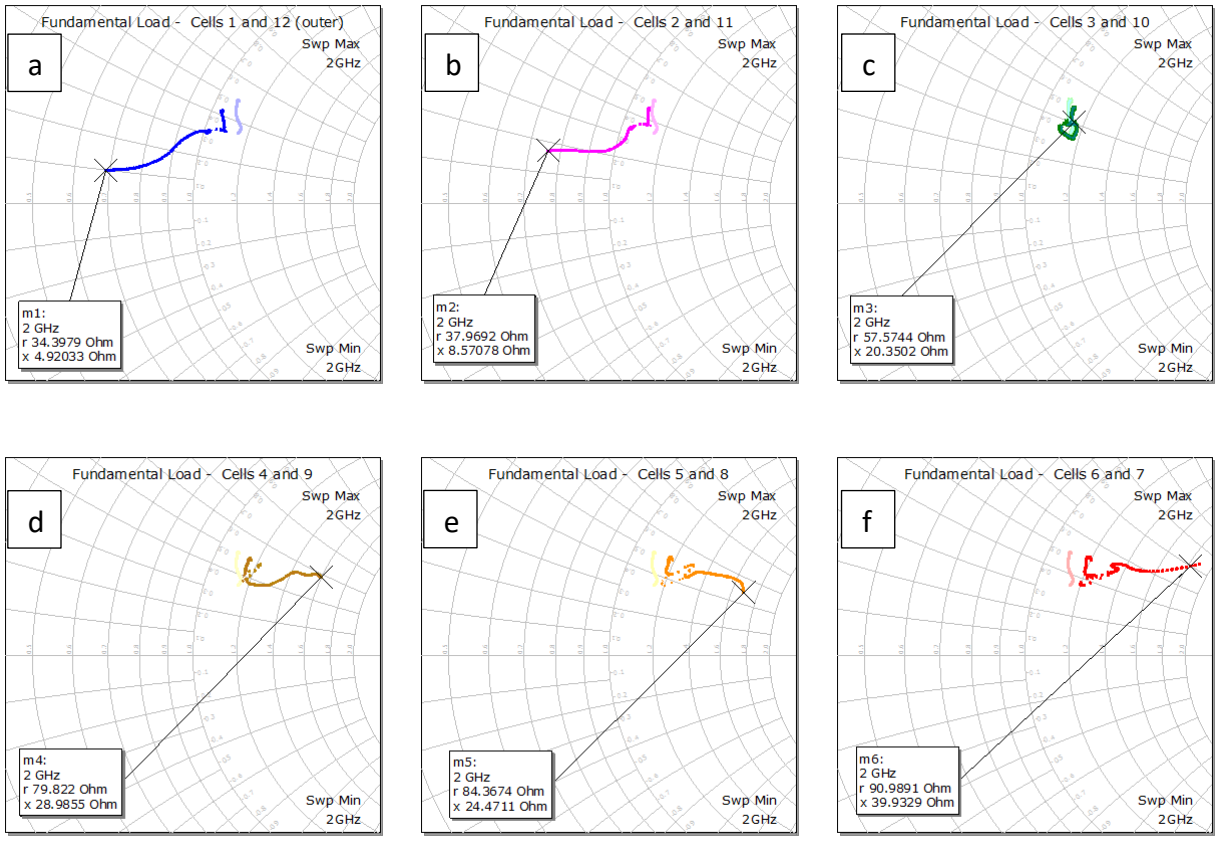

Figure 5.23 Individual transistor cells, distributed fundamental load impedance: (a) outer cells through to (f) inner cells. Marker indicated P6dB.

However, the trend is clear that the outer cells shown in figure 5.23 (a), (b) tend toward a lower impedance load at high power levels, and are seen in figure 5.24 to operate in a lower voltage / higher current condition (blue and pink traces) than neighboring cells. Conversely, the inner cells shown in figure 5.23 (e), (f) tend to a higher impedance load at high power levels, and the waveforms at P3dB show the cells operating in a high voltage, lower current mode.

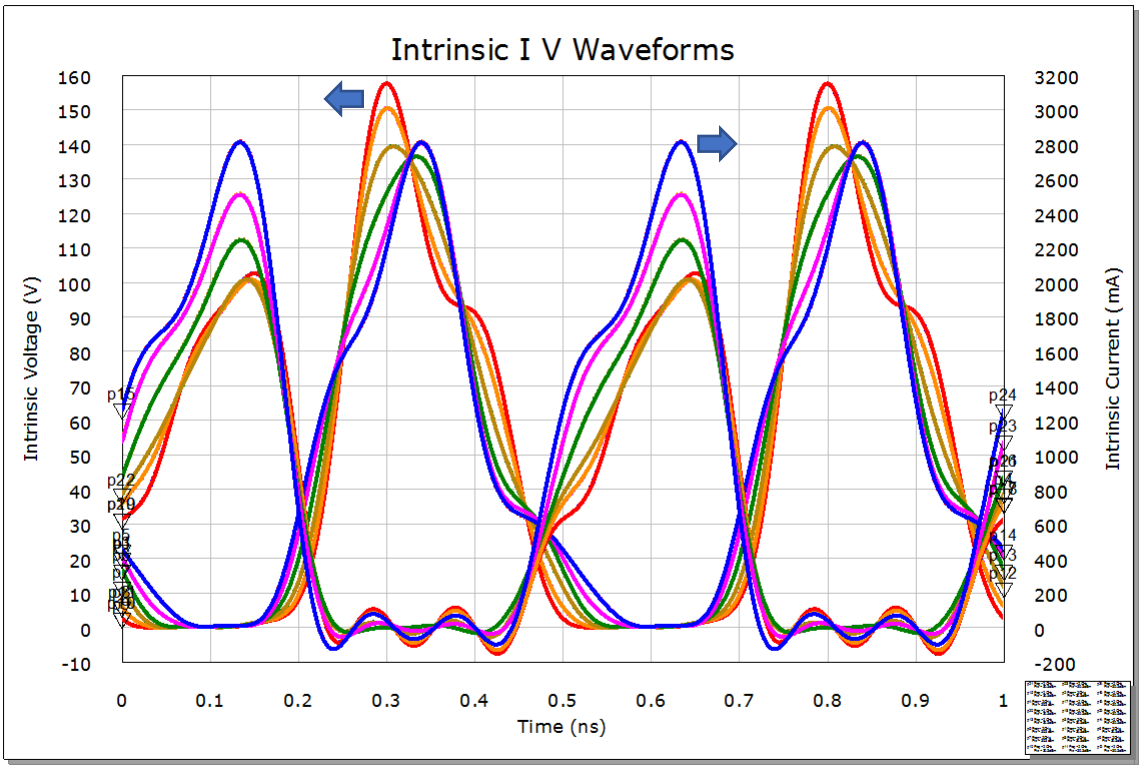

Figure 5.24 Individual cells, current and voltage waveforms at current generator into peak efficiency load @ P3dB compression.

5.2.7 Impact of Distributed Effects on Individual Cell Harmonic Load Impedances

In this section, the concept of the amplifier operating mode is considered for the high-power device. Common design practice involves selection the amplifier operating mode based on specific theoretical criteria. For example, class-F rather than inverse-F operation may be required in order to minimize drain voltage peaking when using a technology with limited voltage breakdown headroom. A design flow can start with extraction of the device output capacitance, and the target load impedance given the IV characteristics of the device (referred to as R_p/C_p values). However, even before the distributed nature of the power device is considered, the notion of a 'static' harmonic impedance is seen in figure 5.25 to provide a simplification over the real operation at the current generator plane.

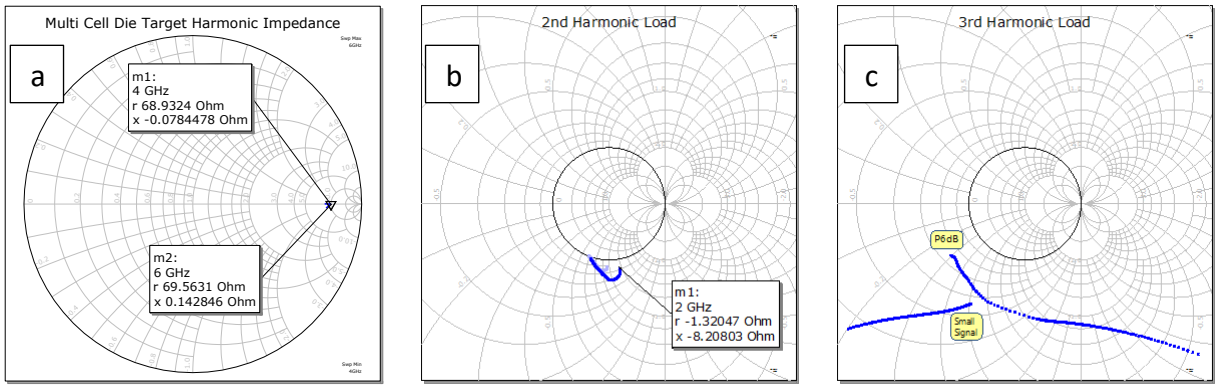

Figure 5.25 Harmonic impedance at: (a) second and third harmonic at die bond pad reference plane, (b) second harmonic at current generator, (c) third harmonic at current generator.

Figure 5.25 (a) shows the second and third harmonic design targets at the die reference plane. The impedances targeted are based on die level load pull performance data. However, in considering the harmonic impedances at the current generator, figure 5.25 (b) second harmonic and (c) third harmonic, it can be seen that the exact harmonic magnitude and phase into which the device operates become a function of the output power due to the non-linear C_{out} of the transistor.

As the physical size of the transistor becomes electrically significant, another layer of complexity is added to the plots in 5.25. Each cell in the power array generates non-linear terms at the harmonic frequencies, and these terms are a function of the fundamental load impedance presented, which, as shown in section 5.2.6, are different when a distributed load is considered. Therefore, the non-static harmonic impedance across the cells in the array is a function of nonlinear C_{out} (common to all cells), the harmonic impedance of the structure vs. cell connection location, the active load pulling from the IV waveforms of the other cells in the array, and the amplitude and phase of these to the other cells.

Figures 5.25 (second harmonic) and 5.26 (third harmonic) show the impact of the above in the context of the harmonic impedance presented to the cells along the power transistor manifold. As observed with the fundamental load impedances in figure 5.23, harmonic impedances presented show symmetry to mirror image cells in the array.

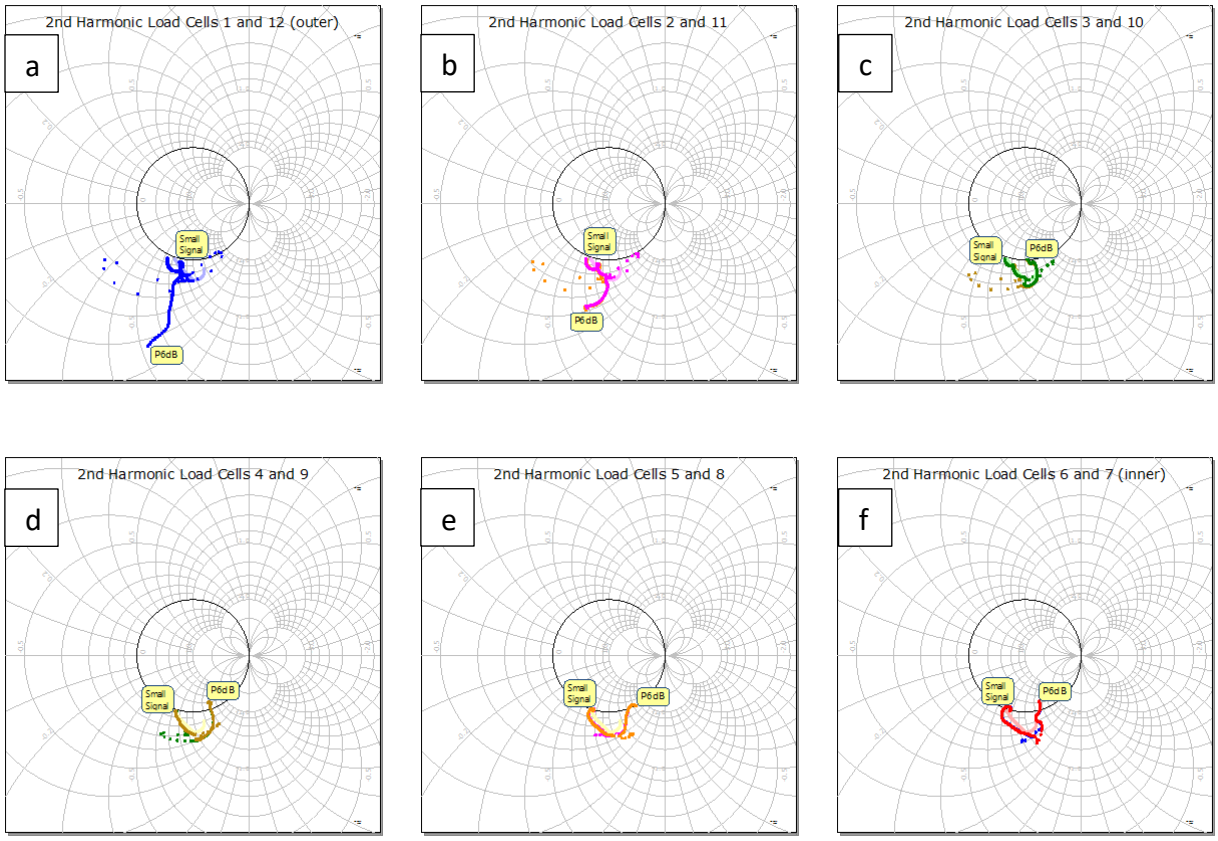

Figure 5.26 Individual transistor cells, distributed 2nd harmonic load impedance: (a) outer cells through to (f) inner cells.

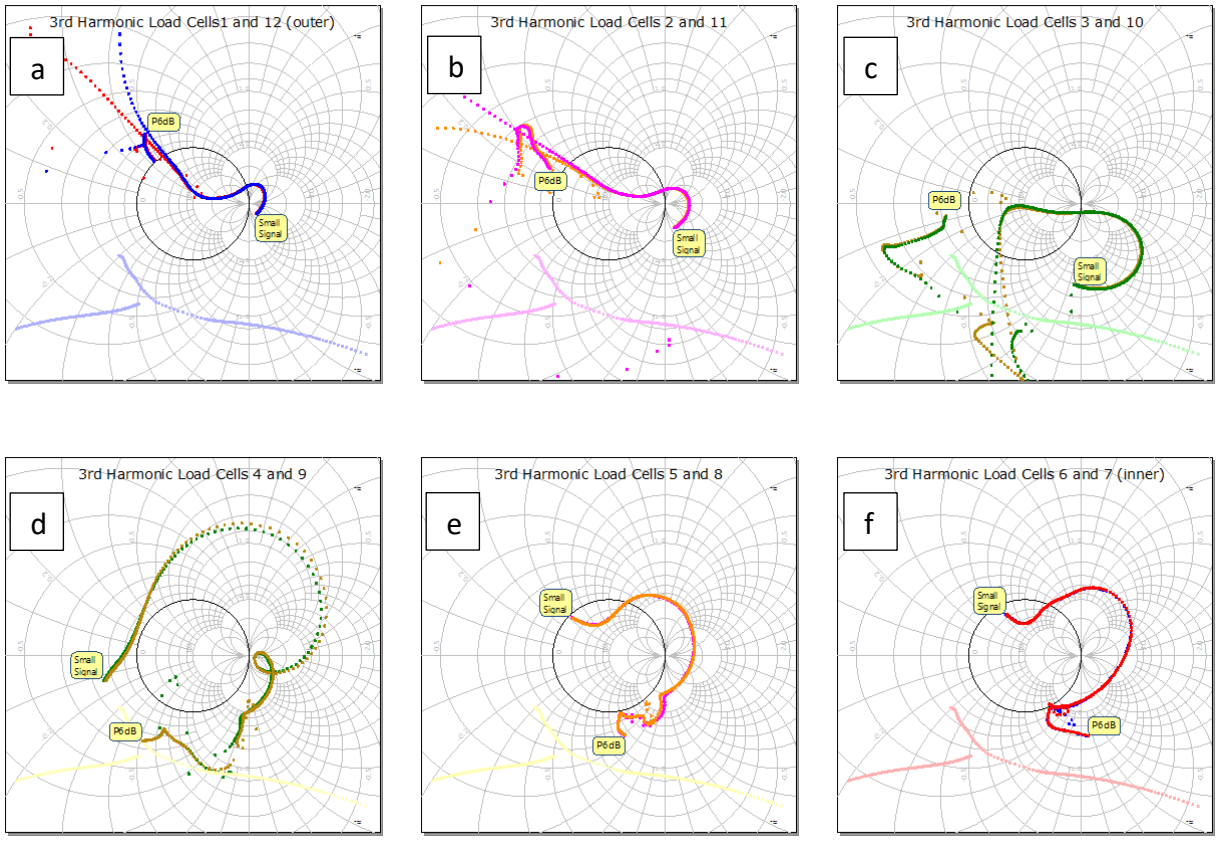

Figure 5.27 Individual transistor cells, distributed 3rd harmonic load impedance: (a) outer cells through to (f) inner cells. Marker indicated P6dB.

Figure 5.28 provides a summary of the harmonic impedance variation according to the reference plane at which it is observed, and the impact of distributed effects. Figure 5.28(a) shows that with no consideration of distributed effects, and observing extrinsically, the second harmonic load impedance is a single point across all cells and drive conditions. Figure 5.28(b) shows the impact of shifting the reference plane from extrinsic to the intrinsic plane. In this case, the harmonic impedance presented is no longer a single impedance, but a dispersive impedance that is a function of the output power. Additionally, the impedance presented is no longer limited to the unity smith chart, and an active nonlinear impedance is observed at intrinsic device level [1]. Both of these observations are due to the non-linear behavior of the device C_{ds}.

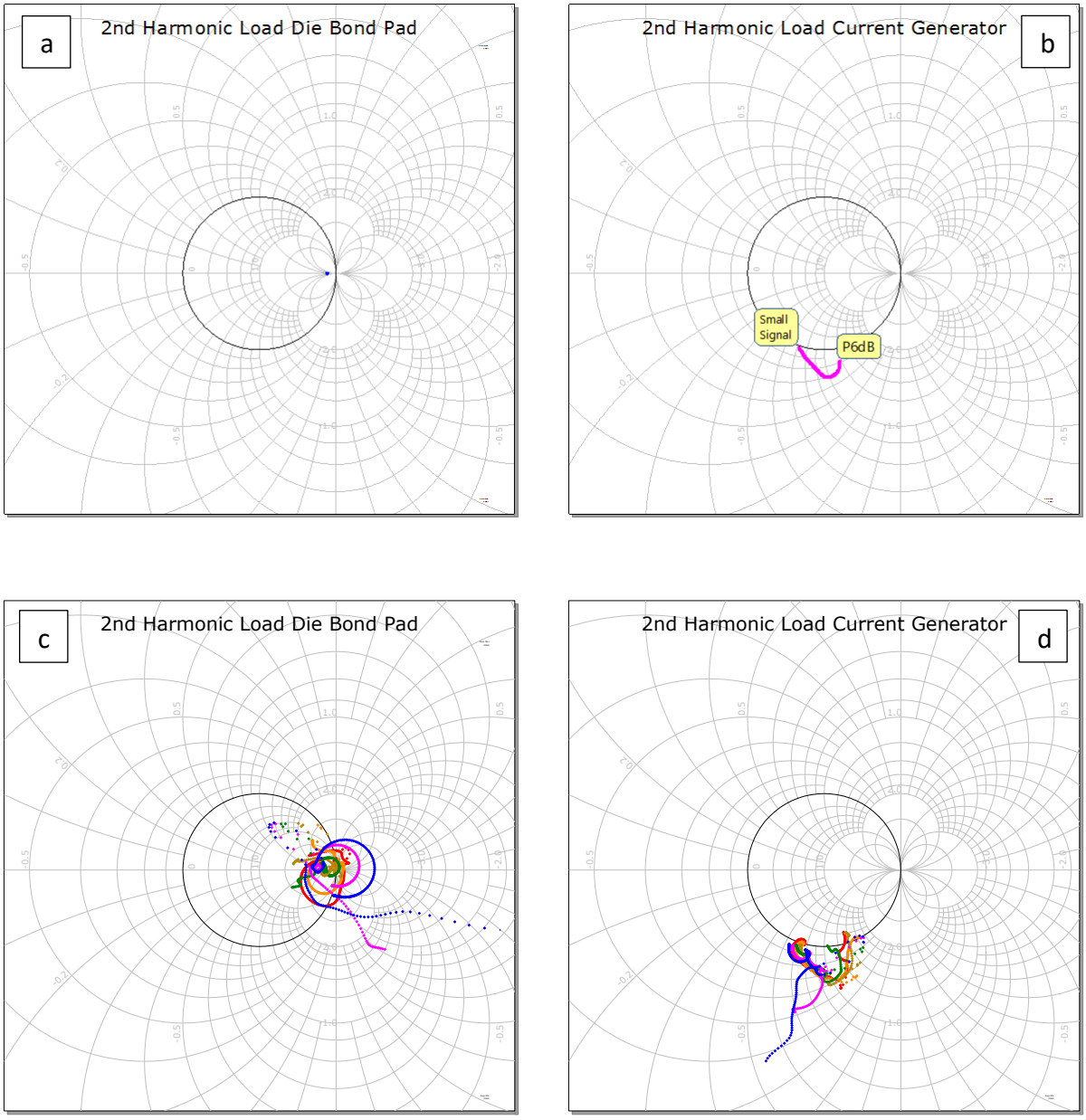

Figure 5.28 Summary matrix of 2f Impedances with non-linear C_{ds} and distributed effects: (a) die bond-pad reference plane no distributed effects, (b) intrinsic current generator reference plane no distributed effects, (c) die bond-pad reference plane with distributed effects, (d) intrinsic current generator reference plane with distributed effects.

Figure 5.28 (c) shows the impact of the distributed nature of the structure, observed at the extrinsic bond pad reference plane. At this reference plane, the impedances are not impacted by the nonlinear C_{ds} of the cell (since the observation point is extrinsic to the device C_{ds}). However, it can be seen that the distributed effects give rise to both impedance dispersion vs drive, as well as an active harmonic impedance at certain output power levels. This effect can be attributed to the active load pulling of the 2nd harmonic terms generated by the other cells in the array, and the magnitude and phase at which these injected I-V waveforms arrive at the cell under test drain manifold.

Finally, Figure 5.28 (d) shows the summation of the impact of the distributed load and the nonlinear C_{ds}, as is presented in figure 5.26 (separated into cell position). It can be seen that the harmonic load presented to the cells is both active, and varying across the die manifold.

5.2.8 Impact of Distributed Effects on Current and Voltage Waveforms vs. Output Power

Figures 5.29 and 5.30 characterize the impact on the magnitude and phase of the IV waveforms due to the fundamental and harmonic impedance variations across the cells. In Figure 5.29, it can be observed that as the device transitions from peak gain (compression reference) to P3dB compression, the current contribution from the outer cells increases significantly, transitioning from -7% to +37% relative to the center cells 6 and 7. This aligns with the impedance trajectories observed in figure 5.23, whereby the out cells transition to a lower intrinsic impedance (therefore higher peak current). The effect is mirrored by the drain voltage waveform, whereby the center cells (seeing a higher intrinsic impedance) exhibit +12% higher drain voltage than the outer cells of the device.

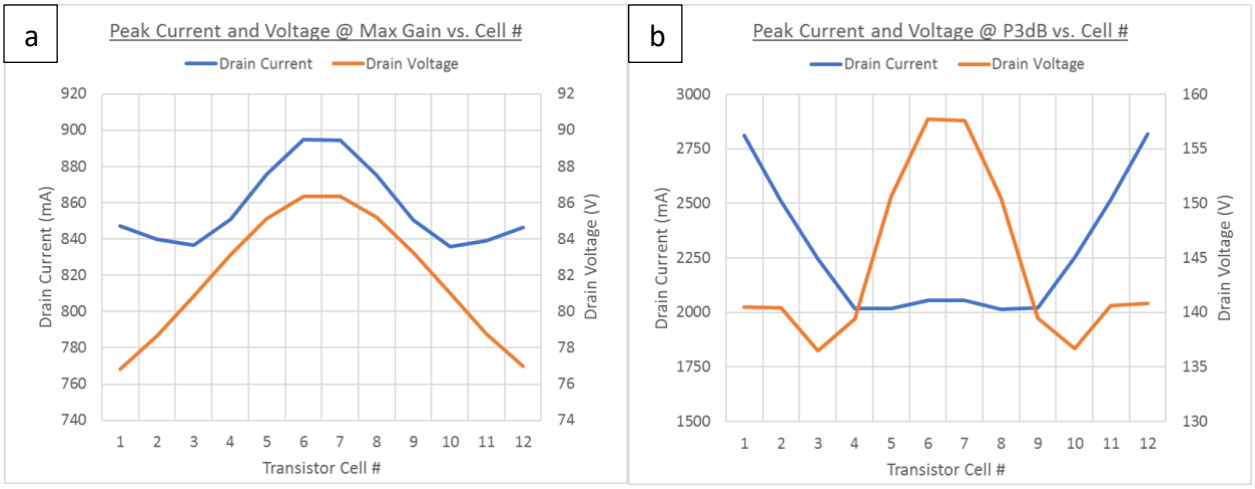

Figure 5.29 Amplitude of peak current and voltage per cell at: (a) max gain, (b) P3dB compression.

Figure 5.30 shows the impact of the diverging reactance of the presented loads to the cells. It can be seen that at the peak gain power level, the current and voltage phase deltas are contained within a tight window of 8.6° (current) and 7.2° (voltage). However, as the device enters compression and the load reactance delta increases, the phase delta windows are increased to 11.5° (current) and 28.8° (voltage).

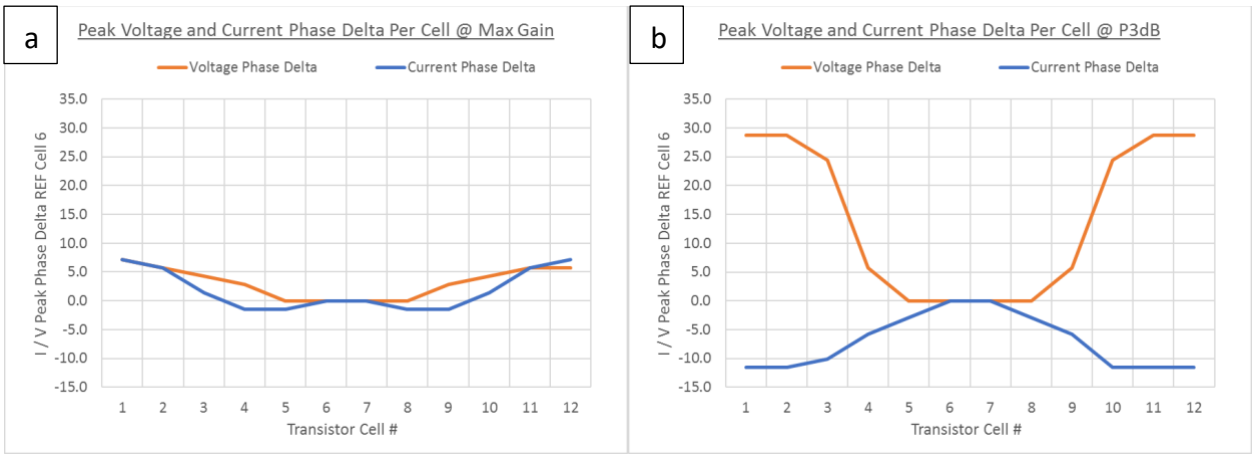

Figure 5.30 Phase of peak current and voltage per cell at: (a) max gain, (b) P3dB compression.

5.2.9 Performance Degradation Summary

Figure 5.31 shows the breakdown of the performance degrading mechanisms in the RF power device. It can be seen in 5.31 (a) that the insertion loss of the package contributes to 0.14dB loss, thus the P3dB compression of the transistor is reduced from 58.88dBm (die level) to 58.74 (package level). This power loss gives rise to a corresponding efficiency degradation from 88.39% to 85.58% (calculated) or 85.34% (simulated load pull + interpolation).

In adding the impact of the distributed die and package structure, it can be seen that the P3dB compression is degraded by 0.27dB relative to the package loss only load pull results. Correspondingly, the device efficiency is reduced from 85.34% to 82.49% (-2.85%) due solely to the addition of distributed analysis.

However, if the theoretical efficiency impact is calculated based on the power degradation, an efficiency of 80.4% is expected, thus a delta of 2.09% is noted between expected and simulated. While this phenomenon may be explored in more detail – it is 10 times greater than the interpolation seen in other cases, and therefore suggests the shift in device operating mode (whereas the device fundamental and harmonic impedances are the same when comparing die and die + package results without distributed effects).

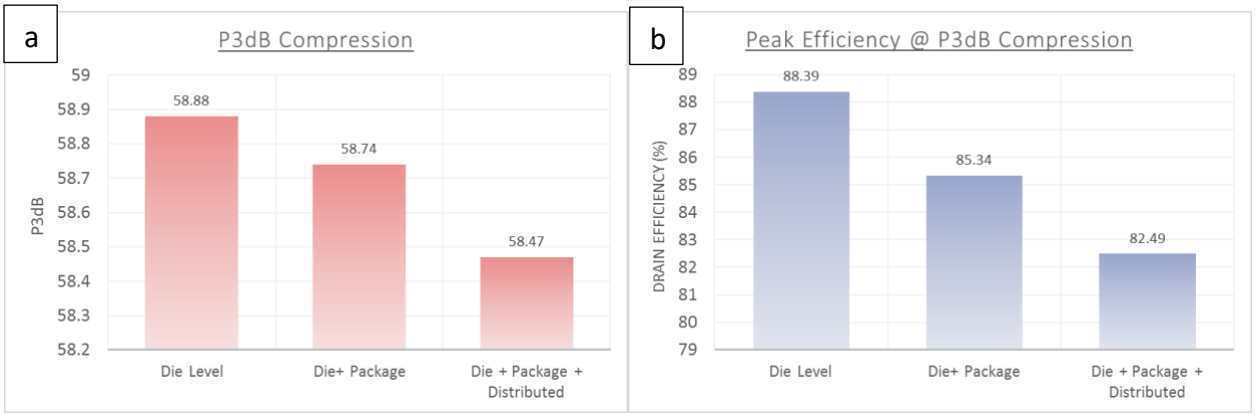

Figure 5.31 Performance loss analysis summary of 12 Cell Power Transistor: (a) Power, (b) efficiency.

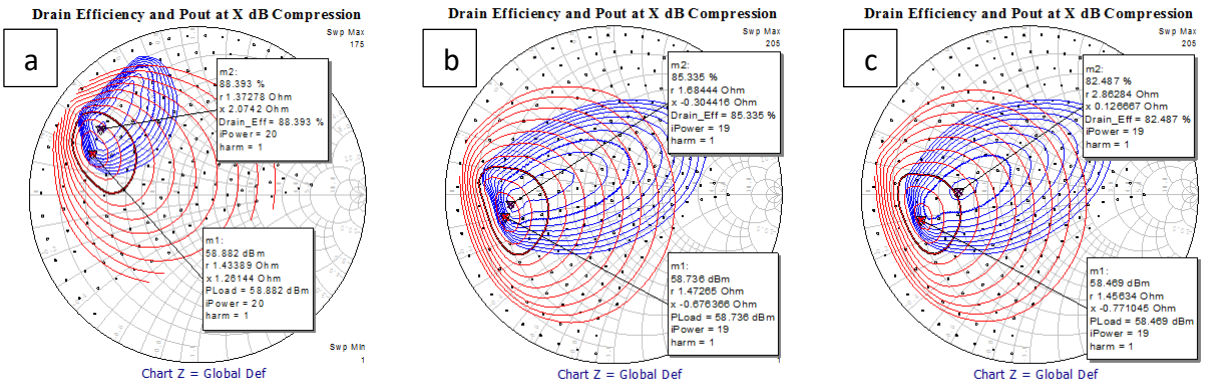

Figure 5.32 Load pull contours: (a) die level, (b) device into common load impedance, (c) distributed load impedance.

5.3 Enhanced Distributed Effects Through Novel Integrated Passive

Section 5.2 described a proposed method to analyze the impact of distributed effects in an RF power device. A power device with an on-package shunt matching network was designed and analyzed, and the performance degradation mechanisms of the device from die to package output analyzed. This was broken down into three areas: -

1. Inherent Die Efficiency into defined harmonic conditions
2. Package level efficiency, assuming constant impedance to all die cells
3. Package level efficiency, including die, wire bond and package level distributed effects

In this section, the analysis described in section 5.2 is used to optimize the degradation between (2) and (3) above, through the use of integrated passive device (IPD) matching elements. The IPD is designed to provide a high Q shunt matching network (as was the design in 5.2), but with a focus on localizing the termination to the individual cell level, rather than the entire die. The revised design is shown in figure 5.33.

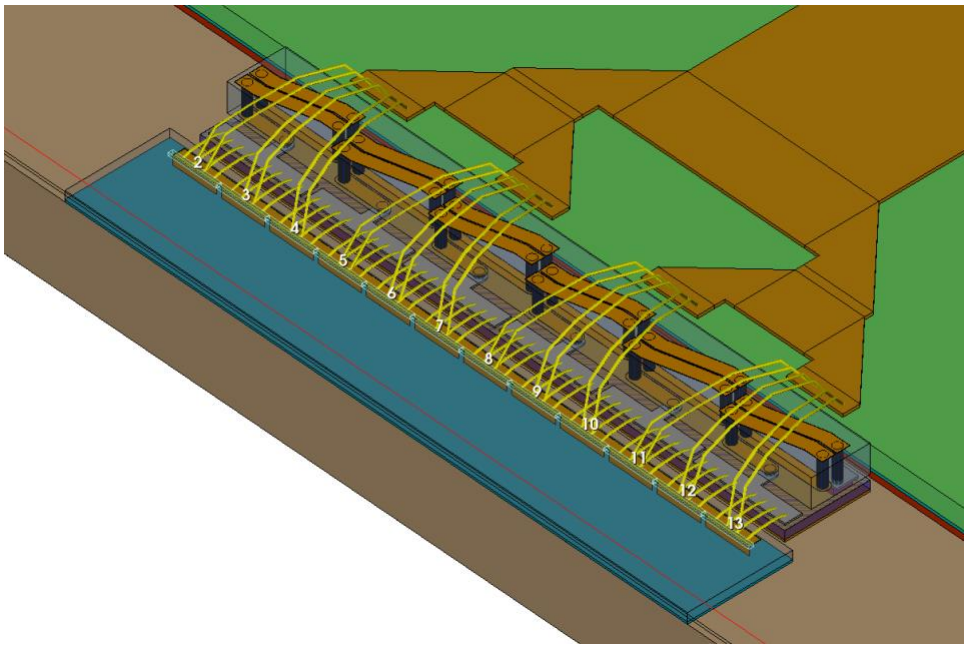

Figure 5.33 Novel multi-level IPD internal match implementation inside the packaged transistor.

The IPD utilizes a novel two-level structure, whereby the shunt match connection between die and passive is bonded to the lower layer, thus minimizing the additional inductance in series with the IPD coils. In addition, the coupling to the drain wire connections is reduced. It can be seen that the IPD comprises of six printed coils, therefore two transistor cells share one coil for impedance matching. This ratio was found to be an optimal compromise at 2GHz between passive losses (which are reduced with larger coils using fatter metal) and distributed effects (where ideally each cell would have one matching coil).

The IPD uses a second substrate layer, attached on top of the carrier substrate to which the shunt match connection is made. This serves to increase the distance between the ground plane, thereby increasing the Q of the coil due to the capacitive coupling effect, and increases the coil length (and therefore chip size) to achieve the target inductance value.

It can also be seen in figure 5.33 that the combining of the 12 transistor cells uses a 4:1 combiner in order to reduce the electrical length into which current is summed into the die combining node. This is intended to provide a more consistent impedance to the cells. The cell combining in this configuration requires 3 cells combined per combiner node. In this case, considerable

optimization was performed using 3D EM simulation to minimizing the bond wire coupling between adjacent groups of 3 cells. It can be seen that the resultant design used non parallel drain wires, since this reduces the coupling factor, and thus dependency on, adjacent wires. Due to the extra length of the outer two wires in the group of three, the height of the center (straight) wire was raised to match the inductance of the outer wires. All aspects of the design were analyzed and optimized using Cadence AWR Analyst 3D FEM simulation software. Figure 5.34 shows the simulated load pull results of the new design at the combined package output reference plane. It can be seen that with the network presented to the die as a single load impedance (no distributed effects), the peak drain efficiency of the device is 82.77%.

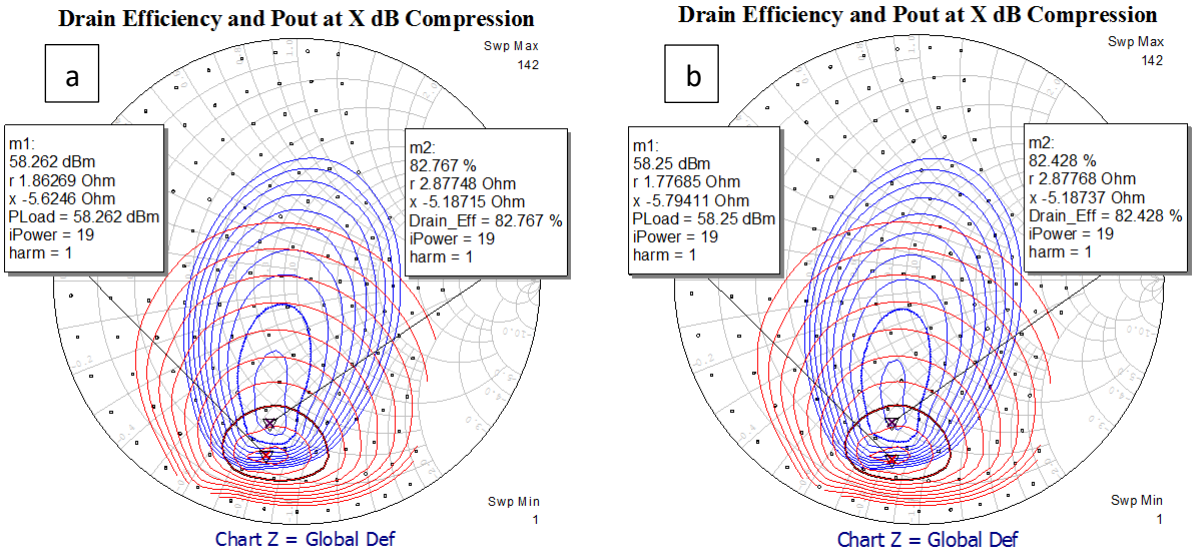

Figure 5.34 Simulated load -pull: (a) no distributed effects, (b) with distributed effects.

With the network simulated including distributed effects, the simulation is re-run and a peak drain efficiency of 82.43% is achieved, a reduction of 0.34%.

	P3dB (no distributed)	P3dB (distributed)	Degradation
Printed Match	58.74dBm	58.47dBm	0.27dB
IPD Based Match	58.26dBm	58.25dBm	0.01dB

Table 5.1 - P3dB - Performance Degradation Summary

These results can be compared directly with the previous design in figure 5.20, since the harmonic conditions at the die reference for the characterization are consistent in both cases. The performance comparison is summarized in table 5.1 (peak power at 3dB compression) and table 5.2 (peak efficiency at P3dB compression).

	Peak Eff (no distributed)	Peak Eff (distributed)	Degradation
Printed Match	85.34%	82.49%	2.85%
IPD Based Match	82.77%	82.43%	0.34%

Table 5.2- Peak efficiency at P3dB, performance degradation summary.

In comparing the performance in table 5.1, can be seen that the P3dB of the transistor in section 5.2 is degraded by 0.27dB when factoring in the distributed effects of the network. In addition, a delta in the load contour positioning (particularly noticeable being the peak efficiency load impedance) can be observed between the two cases.

By contrast the IPD design exhibits an almost negligible 0.01dB degradation in P3dB (this is small enough that it could be attributed to non-ideal interpolation between load pull points).

In addition, table 5.2 shows how the peak efficiency at P3dB is degraded when the distributed nature of the structure is considered. It can be seen that in the case of the transistor in section 5.2, the peak efficiency at P3dB is degraded by 2.85% by the distributed effects of the network, while the degradation of the IPD based design is only 0.34%.

A point to note is that design in section 5.3 has a higher electrical length in the combining / matching network to the load pull reference plane (as can be seen from position of load pull contours in 5.34 compared to 5.20). This of course leads to a greater insertion loss in the network. However, the design in section 5.2 would need to be compared at the same reference plane to compare the losses of the IPD vs. PCB based shunt matching implementation, and would ultimately need to be matched up to a working system impedance. This comparison at a fixed summing node impedance was unfortunately not possible to be made at the time of writing due to simulation

project availability. However, the 'delta' information is the metric of interest and this is independent of measurement reference plane.

Another method of observing the improved distributed effects is in comparing the intrinsic current and voltage waveforms in figure 5.35 (IPD based match) with figure 5.24 (PCB based match). Maintaining the same graph scaling, the improved current and voltage uniformity across the cells in the IPD based design is quite clear to visual inspection.

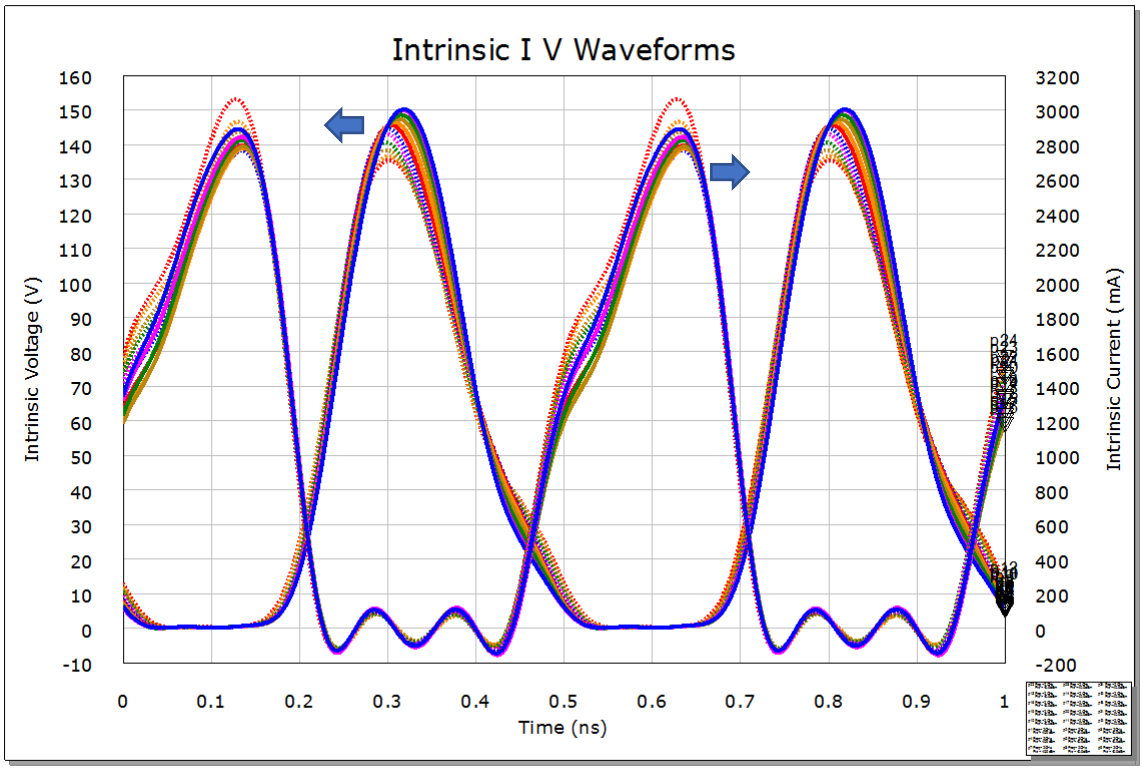

Figure 5.35 - Individual cells, current and voltage waveforms at current generator into peak efficiency load @ P3dB compression.

A final verification of the improved distributed effects can be seen in observing the intrinsic load impedances presented to each cell in the transistor array. It was shown in figure 5.23 that the device in section 5.2 exhibited an impedance shift across the array – whereby the outer cells trend

to a lower intrinsic impedance, figure 5.23 (a), (b), whereas the center cells trend to a higher intrinsic impedance vs. power, figure 5.23 (e), (f).

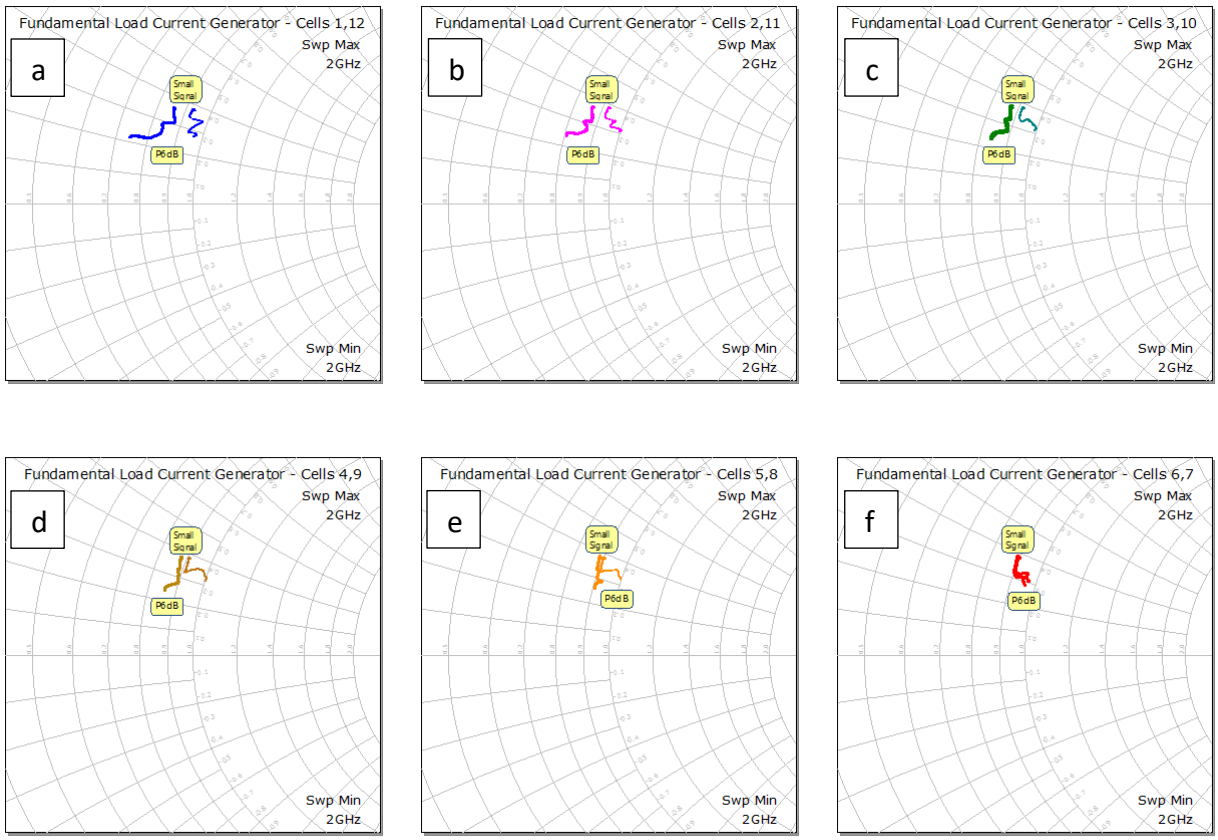

Figure 5.36 Individual transistor cells, distributed fundamental load impedance: (a) outer cells through to (f) inner cells.

It can be seen from figure 5.36 that both the impedance consistency across the array, depicted by the variation between plots (a) - (f), and the delta in impedance vs. drive, depicted by the impedance delta in the individual plots, are both reduced in comparison to the characterization in figure 5.23.

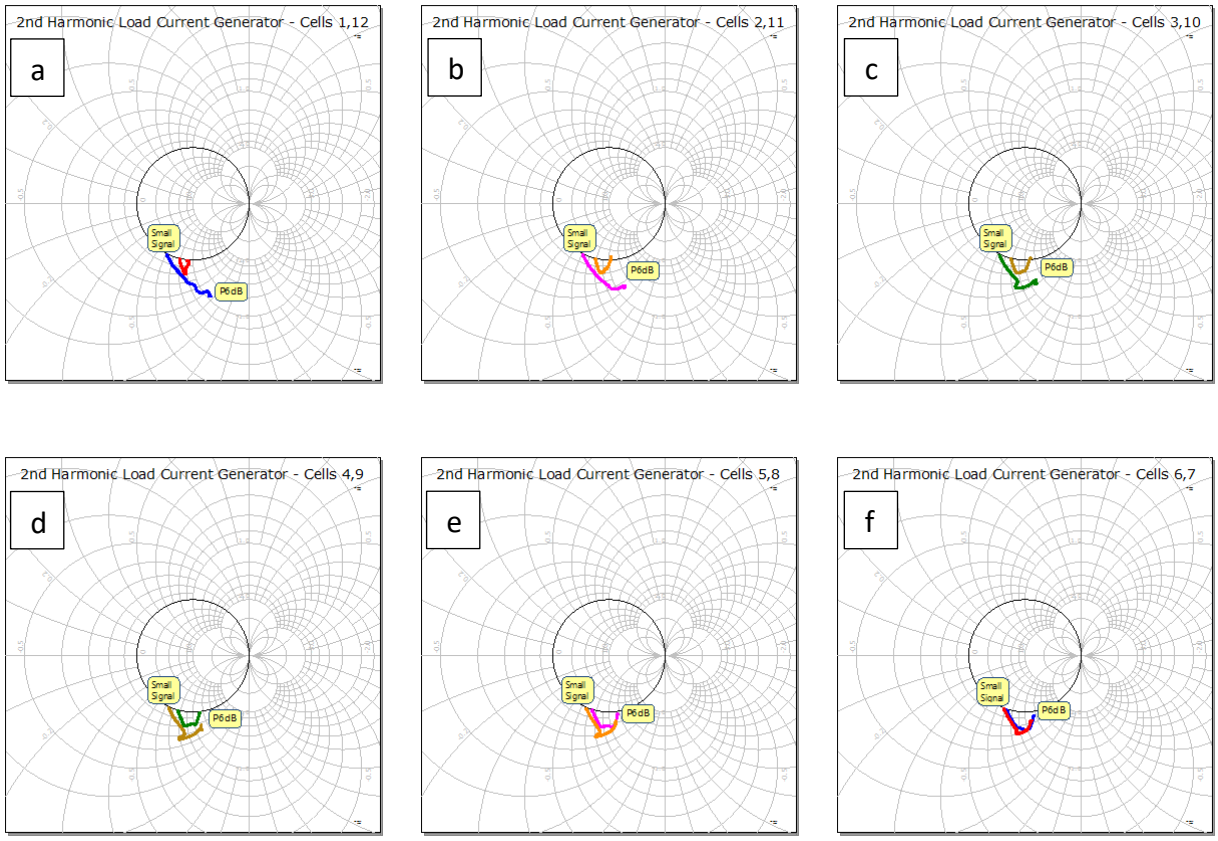

Figure 5.37 Individual transistor cells, distributed 2nd harmonic load impedance: (a) outer cells through to (f) inner cells.

Likewise, in observing the harmonic impedance presented to the cells in the array, the device in section 5.2 exhibits a rapid delta in the 2f impedance presented at the outer cells in the array, as can be seen in figure 5.26 (a), (b). In contrast, the IPD device shows tighter impedance consistency across the array, characterized in figure 5.37.

5.4 Summary

This chapter describes the limitations of scaling the RF transistor and its impact on both power and efficiency. The limitations are analyzed at two levels: -

Section 5.1 considers the effects at the unit cell level, showing limitations of scaling the high-power transistor building block.

Section 5.2 describes the issues scaling the unit cells to a high power, high-frequency transistor and proposes a detailed methodology to study the impact of the power device layout and surrounding elements (passives, package, PCB, etc.). Simulation results are presented for a high-power device using a PCB based matching topology, and the performance impact of the distributed structure is quantified.

The design methodology presented in 5.2 can be utilized for optimization at the unit cell level in section 5.1. While not the focus of this work, enhancing the layout of the unit cell used in the power array will yield improvements to the high-power device in which it is used and would be a potential topic for further work.

As part of the analysis, the concept of active harmonic impedances due to the non-linear C_{ds}, first demonstrated by Cardiff University in [2], is observed. Also, the concept of the active harmonic impedances presented to individual die cells in the array due to the distributed effects of the die and surrounding matching elements, independent of the C_{ds} characteristics, is demonstrated. To this author's knowledge, this is the first published work demonstrating this concept. The use of this finding provides the potential for further work in high-efficiency amplifier design.

Section 5.3 presents the high-power device's re-design from section 5.2, focusing on performance degradation minimization due to the distributed effects. Here, a novel implementation of an integrated passive device (IPD) is developed.

In this case, the IPD enables two different wire bondable surface heights (unlike a traditional IC with one top wire bondable surface). The lower surface, ideally at the same vertical height as the transistor die, enables a very low inductance interconnect between die and IPD. The upper surface enables the implementation of high Q printed inductors, with the lower surface providing

sufficient distance from the ground to minimize the capacitive coupling effects, increasing the Q of the inductor. A significant reduction in the performance degradation between the distributed and non-distributed results is observed between the device in section 5.2 and section 5.3. This can be observed in the load-pull results but also in characterizing the variation in load impedance across the die and in the intrinsic IV waveforms of the cells in the power transistor.

5.5 References

[1] Mohapatra, M., "Performance analysis of AlGaN/GaN based HEMT for different gate structure," IEEE International Conference Devices for Integrated Circuits, 2017, Kalyani, India.

[2] Mokhti, Z., Lees, J., Cessan, C., Alt, A., Tasker, P., "The Nonlinear Drain-Source Capacitance Effects on Continuous-Mode Class-B/J Power Amplifiers," IEEE Transactions MTT Volume 67, No.7, July 2019.

Chapter Six – Conclusions and Further Work

6.1 Conclusions

Development of high efficiency, wideband power amplifiers for high PAR signals has proved a fertile area of research over the past decade. As the frequency, power, efficiency and bandwidth boundaries increase, pushing the performance envelope requires optimization in multiple areas, including: -

1. Characteristics of the intrinsic die
2. Optimization of the die layout
3. Selection of matching network topology (on the die, inside the package, and on the PCB)
4. Optimization of die in conjunction with the device internal matching network and high-efficiency amplifier topology selection
5. PA operating mode, and implementation of harmonic termination
6. High-efficiency amplifier architecture selection
7. Implementation of selected architecture
8. Linearizability of the final amplifier

Each of the above topics contains an opportunity for much research. This thesis focuses on three main areas in this list. Chapter three presents an analysis of the bandwidth of multiple impedance matching structures, as these would pertain to application in the Doherty amplifier. The chapter shows that a power amplifier matched using traditional wideband matching structures and connected in a Doherty topology forms an inadequate design approach under load modulated conditions. In this case, the wideband structures provide wideband impedance transformation with low individual transformation ratios only in a fixed impedance environment. However, under load modulation, this is no longer true, and the transformation ratio and bandwidth of such structures reduce considerably. Also, the requirement for impedance inversion in the matching network for correct Doherty operation necessitates consideration of the electrical length of the network, shown in (6.1).

$$\emptyset_{inv\_main} = 90° + n \bullet 180° \quad (6.1)$$

Where $\emptyset_{inv\_main}$ is the phase length from the main amplifier, and n is an integer multiple of 180°.

Two concepts are then presented, whereby appropriate use of the characteristics of the structures is shown to benefit the bandwidth performance under load modulated conditions. The use of a wideband structure outside of the load modulation region of the Doherty circuit to define a system impedance at the main and peaking amplifier summing node is shown to provide the bandwidth benefits of these structures. The concept of a 'constant impedance Doherty' is presented, whereby the wideband structure transforms the external impedance, typically 50Ω for transmitters, to a system impedance equal to the Z_{mod} impedance of the main amplifier. In this case, an impedance inverter with a characteristic impedance equal to the main amplifier Z_{mod} is used, and therefore no band-limiting in the Doherty combiner occurs at the average power level. When using the constant impedance topology, a peaking amplifier matching network is required to transform the effective summing node impedance at full power to the peaking amplifier optimum load impedance - a 4:1 transformation ratio in the symmetric two-way Doherty. A hardware demonstration of this concept is presented, whereby a 100-watt LDMOS Doherty amplifier is designed at 2GHz, and demonstrating >50% efficiency over 20% fractional bandwidth, the highest bandwidth at this frequency and power level at the time of publication.

Later in the chapter, the bandwidth impact of the peaking amplifier in the off state is considered. Here, a novel 'zero phase peaking' topology is proposed, whereby the peaking amplifier impedance is absorbed into the Doherty combiner impedance. In this case, the bandwidth limitations incurred due to the peaking amplifier loading at the summing node are significantly reduced. In the case of a fixed technology design, a low summing node impedance is required (6.2), and an impedance transformation occurs in the main amplifier impedance inverter at the average power level (6.3).

$$Zsystem_{n=0} = \frac{Z_{opt_{main}}}{\left(1 + \frac{I_{peak}}{I_{main}}\right)} \tag{6.2}$$

$$Z_{inverter} = \sqrt{Z_{system} \times \left(\left(1 + \frac{I_{peak}}{I_{main}}\right).Z_{opt\_main}\right)} \tag{6.3}$$

The concept of frequency compensation techniques is described, whereby a 'compensation network' is introduced with inverse impedance vs. frequency characteristics to those of the internal device shunt matching network, ubiquitous in high power high-frequency LDMOS, and increasing in prevalence in GaN devices. A bandwidth enhancement is demonstrated compared to the shunt network in isolation. Also, the network provides zero phase shift, an essential feature in the n=0 Doherty to ensure real to real load modulation and a high impedance off-state phase angle.

The final topology presented merges the constant impedance and the zero-phase peaking topologies, using multi-voltage device technology, thereby leveraging the advantages of both topologies (no combiner band-limiting at average power, and a significant increase in off-state bandwidth). Here the main amplifier supply voltage (and corresponding selection of technology) is included as a design variable along with inverter characteristic impedance in the zero-phase peaking topology, enabling selection to form a constant impedance Doherty with appropriate V_{dd} selection. A 225-watt GaN-based Doherty amplifier covering 1.8 – 2.7GHz, achieving 45% efficiency at 7.5dB OBO, demonstrates the advantages of this topology.

In chapter four, the research focus shifts to consider the linearizability of the device. The impact of the nonlinear device input capacitance is presented, demonstrating the generation of even order intermodulation products falling in the baseband frequency range due to the nonlinear input capacitance. It is shown that these IMD products, generated at the transistor gate, are amplified

by the transistor and sum with nonlinear terms generated from the nonlinear device transconductance and the nonlinear device output capacitance.

It is shown that enhancements to the baseband impedance presented across the signal bandwidth significantly reduce the magnitude of the IMD terms generated, therefore reducing the input terms summed with the G_m and C_{out} distortion products.

A novel input matching topology, and integrated passive matching IC is developed to provide a close to ideal input baseband termination. A 60-watt LDMOS transistor is designed using the matching topology, and a baseline device is also designed for comparison using the typical internal input matching topology used for LDMOS devices on the market at the time.

A simulated DPD comparison is made of the two topologies, showing approximately 4dB improvement in the corrected performance at the same output back-off level. Reference circuits for both devices are designed and identically load tuned, and the DPD correction characterized. Enhanced DPD correction is achieved from the device with the novel input matching, typically achieving 3dB improvement in comparison to the corrected levels achieved from the standard implementation. The trend occurs across multiple DPD platforms and test signals (W-CDMA, MC-GSM, and LTE).

Chapter five describes the issues in scaling the transistor with power and frequency, and an overview of the issues at the unit cell level presented. However, the main focus of the chapter is the operation of each cell as part of a power device – whereby multiple unit cells are used in parallel to achieve a high-power RF device.

A comprehensive analysis method is presented to analyze the impact of the distributed effects across the unit cells making up the power transistor. The intrinsic IV waveforms, as well as fundamental and harmonic dynamic impedances, are presented and compared across two impedance matching implementations. A harmonic calibration technique is presented to ensure all structures are analyzed in the same operating mode, with identical harmonic impedances during fundamental load-pull, and a methodology to enable optimal device matching for load modulated operation is presented.

One novel contribution of this work is the demonstration of active harmonic impedance generation at the extrinsic device reference plane (external to the device C_{ds}) due to non-linear terms generated by other cells in distributed transistor array. This effect is demonstrated, along with the previously published work [1] whereby active harmonic impedances are generated due to the presence of a non-linear C_{ds}, clearly shown using this analysis technique.

The chapter concludes by describing a novel integrated passive matching structure development, using a multi-level IPD forming part of the internal device matching network. The distributed effects of the device are optimized in comparison to the previously demonstrated design, and the performance delta due to distributed effects is characterized, with the performance reduction summarized in table 6.1.

	Peak Power Degradation	Peak Eff Degradation
Printed Match	0.27dB	2.85%
IPD Based Match	0.01dB	0.34%

Table 6.1- Device Peak Power and Efficiency Performance Degradation

6.2 Further Work

During this research, many ideas arise that, for multiple reasons, could not be pursued. In this section, these ideas are described and recorded as goals for future research.

6.2.1 Balun or Hybrid Combiner with Impedance Transformation to System Impedance

As has been shown in chapter three, the use of wideband structures outside of load modulation enables a wideband Doherty amplifier system impedance. Two Doherty designs, presented in chapter 3, demonstrate this concept using both a Klopfenstein taper and a two-section wideband matching network.

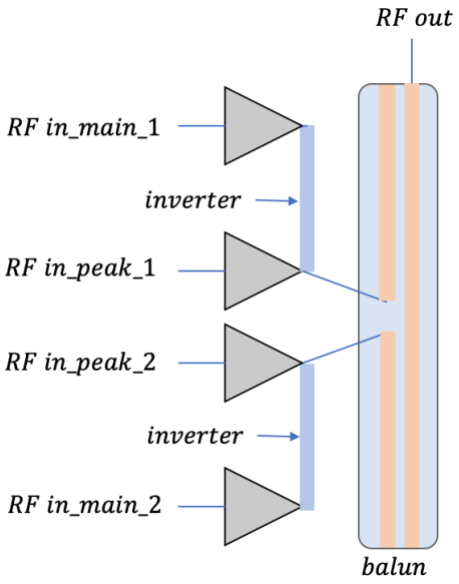

Figure 6.1 – Balun defined system impedance Doherty.

However, an alternate approach to this method is the use of a wideband balun, such as the Marchand balun [2] designed with impedance transformation to the balanced ports. This topology requires two Doherty amplifiers, as shown in figure 6.1, whereby either of the topologies

presented in chapter three (zero phase peaking or constant impedance combiner) can be implemented with the balun designed to provide the correct system impedance.

6.2.2 Load Modulation from Balun or Hybrid

The wideband properties of coupled structures can be exploited in the use of either a tightly coupled (nominally 3dB) hybrid coupler, or a balun. With proper design, impedance transformation can be implemented into the balun or hybrid structure, and load modulation introduced to the main amplifier connection port through the current contribution from the peaking amplifier. Figure 6.2 shows a typical schematic of this topology.

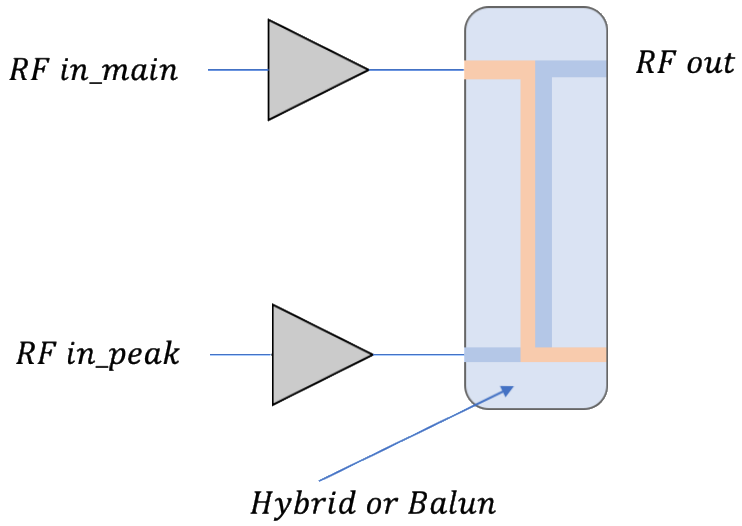

Figure 6.2 – Load modulation of balun or hybrid coupler.

6.2.3 Mixed Technology Doherty

The use of multi-voltage technology in the multi-voltage Doherty amplifier presented in chapter three enables enhanced performance, in comparison to running a higher voltage technology at a lower voltage. This is because the knee voltage of transistors typically scales proportionally with

breakdown voltage, and therefore using the main amplifier transistor optimized for the operating Vdd selected is proposed as an enhancement to the design featured in chapter three

6.2.4 Absorbing C_{ds} into Impedance Inverter for Lower C_{ds} Technologies

Much of the research presented focuses on the LDMOS RF power transistor, ubiquitously used for cellular infrastructure amplifier during the period of research. For such transistors, the use of a shunt output match is essential for high performance since the high device C_{ds} does not enable the use of external PCB matching alone. Therefore, many concepts presented here enable wideband operation of a high C_{ds} technology at high power, with enhancements in bandwidth achieved through these approaches as the maturity of GaN HEMTs, and the low C_{ds} these enable.

However, wideband results have been demonstrated at low power levels [3] by absorbing the C_{ds} into the impedance inverter. While not considered possible with LDMOS devices at high power levels, further investigation of this approach using the low C_{ds} available from high power GaN devices could enable alternative wideband solutions to those presented.

6.2.5 Folded Klopfenstein Taper

It has been demonstrated in chapter three that the use of the Klopfenstein taper provides an effective method in presenting a wideband, low system impedance to the Doherty amplifier. However, in practice, such designs are typically only seen in academic papers, where practical considerations can be overlooked in favor of performance alone. The Klopfenstein is often replaced with a multi-section quarter-wave matching transformer, largely because the latter can easily be folded into the PCB real estate available. Due to the constant impedance of each matching section, an ideal miter is easily achievable with good return loss, and with a phase that can be accurately simulated using 3D finite element method (FEM) or 2.5D method of moment (MOM) electromagnetic simulation techniques. However, the design of a folded Klopfenstein taper is considerably more complex, since the well-known functions that exist for optimally mitering uniform impedance lines do not exist for continually varying impedance vs. phase

structures. Therefore, the development of variable angled sub structures making up the Klopfenstein (or other continually variable structures such as the Heken or exponential tapers) to enable the design of folded tapered structures would be a benefit to the RF community.

6.2.6 Novel Input Matching Topology, Technology with highly Nonlinear C_{in}

Chapter four demonstrated the enhanced DPD performance of a novel input matching topology using a 60-watt LDMOS transistor. Simulations demonstrated the generation of 2nd order baseband terms through the non-linear input capacitance.

The non-linear characteristics of silicon LDMOS and GaN HEMTs are quite the opposite, LDMOS having a highly non-linear C_{ds}, but quite a linear C_{gs} (around a 10% change with drive). The GaN HEMT is the opposite, with quite a linear C_{ds}, but highly non-linear C_{gs}. Therefore, the magnitude of the input non-linear distortion terms is more significant for the GaN HEMT in comparison to LDMOS. This technique proposes interesting further work in the enhancement of GaN HEMT linearization.

6.2.7 Analysis for Die Cell Level Optimization

In chapter five, a methodology for the analysis of distributed effects in multi-cell devices has been shown and used to optimize the internal matching and on package combining for a high-power RF transistor. Using the same design approach to the optimization of the cell layout may be employed to compare combining multiple fingers, or the impact of increased finger lengths or gate finger feeder techniques.

6.2.8 Localization of Harmonic Terminations

Chapter five demonstrates the non-uniform impedance distribution across the cells of a single transistor die when operating at high frequencies. In typical amplifier designs, the harmonic termination of the high RF power transistor is usually treated as a ‘global’ implementation,

whereby a PCB structure forms a harmonic short at the appropriate electrical length from the die to present the desired harmonic phase. However, as the power, and therefore the physical size of the die increase, the impedance seen by individual cells varies significantly. The impedance variation is due to active load-pulling from the adjacent cell nonlinearities and the sizeable electrical length of the die at the harmonic impedance. Consequently, the resultant 'optimal' external harmonic impedance is the impedance that is the best average of all the unequal impedances seen by the cells. Therefore, the implementation of a localized harmonic termination at the unit cell or unit finger level enables a more consistent transistor operating mode, and more uniform intrinsic IV waveforms from each current generator, leading to improved cell power combining.

6.2.9 Using Distributed Effects as Active Harmonic Injection to Adjacent Cells.

The harmonic termination presented to each cell is a function of the location within the transistor manifold and the contribution of nonlinear terms from adjacent cells in the array. Therefore, with accurate modeling of device harmonic generation, and electromagnetic simulation of interconnecting physical structures, the cells may be configured such that the harmonic signal summation occurs consistently at each cell, and active harmonic injection occurs, giving rise to an efficiency enhancement over a passively terminated unit cell.

6.3 References

[1] Mokhti, Z., Lees, J., Cessan, C., Alt, A., Tasker, P., "The Nonlinear Drain-Source Capacitance Effects on Continuous-Mode Class-B/J Power Amplifiers", IEEE Transactions MTT Volume 67, No.7, July 2019

[2] Klopfenstein, R.W., "A Transmission Line Taper of Improved Design", Proc. IRE, Vol.44, pp.31-35, January 1956.

[3] Qureshi, J.H., Li, N., Neo, W., van Rijs, F., Blednov, I., de Vreede, L.C.N., "A Wide-Band 20W LDMOS Doherty Power Amplifier", 2010 IEEE MTT-S International Microwave Symposium (IMS).