Novel MMIC Design Process Using Waveform Engineering

A thesis submitted to the University of Wales, Cardiff

In candidature for the degree of

Doctor of Philosophy

By

Dominic Michael FitzPatrick B.Sc., M.Sc.

Division of Electrical and Electronic Engineering

School of Engineering

Cardiff University

Wales

United Kingdom

October 2012

Summary

It has always been the case that talented individuals with an innate understanding of their subject have been able to produce works of outstanding performance. The purpose of engineering science is to define ways in which such achievements can be made on a regular, predictable basis with a high degree of confidence in success.

Some tools, such as computers, have enabled an increase in speed and accuracy, whilst others have given a dramatic increase in the insight into the operation or behavior of materials; the electron microscope for instance. Still others have enabled the creation of devices on a scale unimaginable to our predecessors, Molecular Beam Epitaxy for example.

This work is the product of the availability of an understanding of complex theory on microwave transistor operation, significant increases in mathematical processing and data handling, and the assembly of a 'tool' that not only allows the measurement of high frequency waveforms, but their manipulation to simultaneously create the environments envisioned by the design engineer. It extends the operation of previous narrow band active load pull measurement systems to 40GHz and importantly facilitates the design of high efficiency modes at X band.

The main tenant of this work is to propose that rather than the linear approach of characterisation, design, test, re-iterate, that has been the standard approach to MMIC design to date, the first three stages should be integrated into a single approach which should obviate the need for design reiteration. The result of this approach should be better performance from amplifier designs, greater probability of success first time, and lower costs through less wafer real estate being consumed and fewer design 'spins'.

Acknowledgements

This work would not have been possible without the help and support of a great many people, not only during my time at Cardiff University, but during my engineering career which led me to this point; my early mentors Mike Nunn and John Lillington being key amongst those engineers who set me along the particular path I have travelled. I would like to thank Professor Johannes Benedikt for suggesting and organizing my joining the research effort at the Centre for High Frequency Engineering and Professor Paul Tasker for his inspiration and insight that has driven the work forward at Cardiff. I have been privileged to work alongside and be greatly assisted by Dr. Jonathan Lees who has probably given me the most vital assistance over the years; whilst without the patience of Dr. Tudor Williams I wouldn't have cleared those first hurdles.

The work was sponsored and supported by the Defense Technology Centre and Selex Galileo and I am particularly grateful to Mervyn Haynes and Paul Robertson for their assistance throughout the course of the project. Key assistance has been provided by Malcolm Edwards and Taisto Tinttunen of AWR Corp particularly with incorporating the Cardiff Model DLUT into the MicroWave Office™ nonlinear simulation engine.

Returning to academia was a significant step and the experience was made all the more pleasant by colleagues at Cardiff; the cooperation and assistance offered by Professor Steve Cripps, Amir Sheik, Simon Woodington, and Randeep Saini amongst many others, is very much appreciated.

This work could not have been completed without the support of my family and so I leave my last and greatest thanks to my sons Nathan, Ben, Joel and Samuel and most especially my wife, Verena, without whom this work would never have been started, let alone finished. Her encouragement and faith leave me in awe.

Introduction

It was as a young graduate engineer that I learnt my most important lesson in microwave engineering. Whilst not understanding the behaviour of a solid state power amplifier (in those days 5W at S band was considered high power!), my first mentor, in extreme frustration with me exclaimed, “It’s all $***** $ Ohm’s Law!”; perhaps I should have paid more attention at school and university? I am still surprised by how often this very simple observed relationship has come back to me as fundamental in solving or understanding complicated behaviour in microwave circuits. All too often we have sought explanations by complicated phenomena such as moding or earth loops, or have got carried away with the microwave engineers version of the perpetual motion machine; the 100% efficient amplifier, when a return to basics would have more quickly resolved the questions. It is this appreciation of the fundamental underpinning nature of Ohm’s Law and its application in sinusoidal signals and reactive impedances that has produced seminal works from Cripps [1] with loadline theory and his text books [2]. It has also been the keystone of waveform engineering, a methodology that I have had the privilege to observe develop at Cardiff University, first as a sponsor of PhD students and latterly as one myself, under the guidance and inspiration of Professor Paul Tasker.

There is a danger in the modern world of fabulous nonlinear and electro-magnetic simulators, GHz processors, and Terabyte memory that we skim over the fundamentals and get carried away with digital outputs (microwave power measurements to 4 or 5 decimal places!). Without the sanity check of Ohm’s Law we can delude ourselves or find solutions, the root of which we don’t really understand. Hence the intrinsic beauty of the measurement system developed at Cardiff University; the ability to look in detail at the actual voltage and current waveforms and their dependency on the impedance environment. I remembered from my first degree nearly 30 years ago now that square waves were made from odd harmonics, but when I could actually construct and observe these at microwave frequencies by altering the harmonic impedances it brought a real smile of satisfaction to my face. Similarly when I first de-embedded the drain capacitance from the observed RF loadline and I recognised the shapes described by Cripps in [2], the satisfaction was immense. Having been designing, or perhaps more accurately crafting, microwave amplifiers since the early 1980s, I finally began to feel a real connection between

the theory and what I could observe in the real world. Yes there are times when Ohm's Law can be stretched somewhat, particularly at breakdown in semiconductors, but experience has taught me that you ignore it at your peril.

I feel very fortunate to be working in this industry at this time, but also to have been around long enough to appreciate what we have. I did match circuits using a compass and protractor on a Smith chart,… and I wouldn't go back there. I did run CAD simulations overnight and printout a line of results for each iteration; so that we could scan through the next morning and see if there were any good combinations of variables, - and now I wish the simulations would take long enough to go and get a coffee. I did spend hours calibrating analysers and actually that hasn't changed much (but at least I get a really good screen copy of the results instead of making sketches in my log book and I can store the data in an easily readable electronic file instead of copying them out by hand). We have so much capability accessible to us, but if we don't really understand what we are doing, and take the time to make sure that our measurements are accurate, test our assumptions, model in detail, we might as well be right back in the 1980's taking wild stabs in the dark, hoping to hit the target.

{See the end of chapter 1 for reference details}

List of Publications

1. D. FitzPatrick, T. Williams, J. Benedikt, P.J. Tasker “High frequency sampling technique for microwave active and passive device characterization”, ARMMS (RF & Microwave Measurement Society) Conference, November 2007.

   Abstract: The most commonly used method for characterising microwave components is to measure the s-parameters of the device using a Vector Network Analyser. A problem with this method is that the VNAs are essentially narrow band down converters and cannot handle harmonic and distortion products. By directly measuring the incident and reflected waveforms on the device ports with wide bandwidth samplers and then processing these waveforms, not only can large signal s-parameters be extracted but also harmonic and compression characteristics. The flexibility offered by the direct access to the sampling heads allows for the construction of measurement systems which can, for example, directly measure the multi-port devices with all ports simultaneously excited (such as phased array elements).This paper describes the basic waveform measurement system developed at Cardiff University, and then some of the applications to which this system has been applied including, small and large s-parameter measurement, DC-IV device characterisation, active load-pull including harmonic, and four port antenna match. The potential for 'waveform engineering' and accurate large signal device modeling is described.

2. D. FitzPatrick, T. Williams, J. Lees, J. Benedikt, P.J. Tasker “Large signal device characterization using active load-pull for improved MMIC design”, IET seminar on RF and Microwave IC design, 28th February 2008.

   Abstract: The design of high frequency large-signal active devices depends heavily upon the accuracy of the models used by CAD programs. Traditionally, these models have either been physics or equivalent-circuit based, and this paper presents an alternative approach where measured large-signal data is directly incorporated into the design environment. This approach ensures an accurate representation of the device under large-signal operating conditions and also offers the ability to validate any matching solutions developed before the physical MMIC is realised.

3. D. FitzPatrick, J. Lees, A. Sheikh, J. Benedikt, P.J. Tasker, T. Williams “Fundamental and harmonic termination impedance considerations in the design of optimally efficient wideband MMIC PAs”, EMRS-DTC Conference, Edinburgh, July 2009.

   Abstract: The importance of harmonic terminations in achieving high-efficiency operation in narrow band amplifiers has been widely published. In wide bandwidth applications however, the emphasis has been on maintaining a broadband match at the fundamental frequencies with little regard to the harmonic impedances. The problem is further complicated by the fact that for the lower frequencies their harmonics often fall in band. This paper demonstrates a technique to measure and quantify the impact of harmonic impedance on a devices performance through the active load-pulling of a class A biased 0.3µm GaAs DpHEMT ½W device over 4 to 18 GHz.

4. D. FitzPatrick, J. Lees, A. Sheikh, J. Benedikt, P.J. Tasker “Systematic investigation of the impact of harmonic termination in the efficiency performance of above octave bandwidth microwave amplifiers”, European Microwave Conference 2009, pp. 1445-1448.
   Abstract: The importance of harmonic terminations in achieving high-efficiency operation in Narrow Band (NB) amplifiers has been widely published. In Wide Bandwidth (WB) applications however, the emphasis has been on maintaining a broadband match at the fundamental frequencies with little regard to the harmonic impedances. The problem is further complicated by the fact that the harmonics of the lower frequencies often fall within the operational bandwidth of the amplifier. This paper presents an active load-pull technique for measuring and quantifying the impact of harmonic impedance on device performance, and is achieved through the characterisation of a class A biased 0.3 µm GaAs pHEMT ½W device at 6, 12 and 18 GHz.
5. D. FitzPatrick, T. Williams, J. Lees, J. Benedikt, S.C. Cripps, P.J. Tasker, “Exploitation of active load-pull and DLUT models in MMIC design”, 2010 IEEE Radio Frequency Integrated Circuits Symposium, pp. 487-490.
   Abstract: The use of active load-pull techniques in the design of high efficiency microwave amplifiers has been well documented. This paper describes how it has been applied to the design of a wideband RFIC gain stage. The technique is particularly relevant in new and developing processes where accurate device models are not available and designers otherwise are often forced to use multiple iterations of a design to attempt to encompass the variability in the process. Often in RFIC design, components operate outside of the ideal operating impedance. A look-up table model technique based on measured data which can be used by conventional CAD programs is used to analyze behaviour. This paper shows how the design process and capabilities of the system can be combined to improve the cost effectiveness and performance of RFIC development and with a stable manufacturing process a “first pass” design methodology. The use of the measurement system as an analysis tool is described.
6. D. FitzPatrick, S. Woodington, J. Lees, J. Benedikt, S.C. Cripps, P.J. Tasker, T. Williams “Utilization of DLUT model in the design of an harmonically enhanced high efficiency MMIC amplifier”, EMRS-DTC Conference, Edinburgh, July 2010.
   Abstract: The benefits of higher efficiencies in the power amplifier stages have a multiplied effect on system performance. Traditional approaches to the design of these elements suffer from inaccurate nonlinear modeling which results in multiple iterations until satisfactory performance is achieved, with associated increased costs and delays in project completion. The approach presented in this paper not only increases confidence in delivered circuit behavior, but also allows the determination of the absolute optimum performance without relying on predetermined model constraints.

7. D. FitzPatrick, S. Woodington, J. Lees, J. Benedikt, S.C. Cripps, P.J. Tasker "The application of the Cardiff Look-Up Table model to the design of MMIC power amplifiers", ARMMS (RF & Microwave Measurement Society) Conference, November 2010.
   Abstract: The design of microwave power amplifiers has been greatly enhanced by the use of CAD; however despite the improvements in the area of E-M simulation and non-linear analysis, de-signs still often require to be 'tweaked'. While this may be a practical option for discrete amplifiers it is extremely difficult in the MMIC domain. The requirement for increased efficiency operation and hence the use of modes such as class F and J have highlighted inadequacies in current models. This paper describes the use of waveform engineering not only in the area of device measurement, but also as an integral part of the MMIC design process. The practical limits on current nonlinear models are discussed and an example design using harmonic enhancement, is shown.
8. D. FitzPatrick, R. Saini, J. Lees, J. Benedikt, S.C. Cripps, P.J. Tasker "A waveform engineering approach to the design of improved efficiency wideband MMIC amplifiers", WAMICON (IEEE Wireless And Microwave Conference) April 2011.
   Abstract: This paper describes a design methodology based on waveform engineering within a harmonic active load-pull system and utilizing the Cardiff Direct Look-Up Table (DLUT) model. The process seeks to address some of the key issues facing designers when using nonlinear simulation of high power microwave amplifiers: model validity, boundary conditions and yield. The application of this process to the design of a 0.5 W, 5-10 GHz GaAs pHEMT with good PAE is detailed.
9. D. FitzPatrick, J. Lees, T. Williams, R. Saini, S. Woodington, A. Sheikh, J. Benedikt, S.C. Cripps, P.J. Tasker "The use of harmonic active load pull and direct device modelling in the MMIC design process", Workshop on Advanced III-V MMIC design techniques, European Microwave Conference 2011.
   Abstract: This presentation will outline the importance of including harmonic load pull in the characterisation of transistors and the benefits of creating Direct Look-Up Table models in an integrated measurement – simulation-design environment. A design methodology will be presented which is applicable to both narrow and wide band applications and which delivers faster development cycles.

List of Abbreviations

AC	-	Alternating Current
ALP	-	Active Load Pull
BGA	-	Ball Grid Array
CAD	-	Computer Aided Design
CW	-	Continuous Wave
DC	-	Direct Current
DE	-	Drain Efficiency
DFT	-	Discrete Fourier Transform
DSA	-	Digital Serial Analyser
DUT	-	Device Under Test
FET	-	Field Effect Transistor
FFT	-	Finite (Fast) Fourier Transform
GaAs	-	Gallium Arsenide
GaN	-	Gallium Nitride
GSG	-	Ground – Signal - Ground
HBT	-	Heterojunction Bipolar Transistor
HEMT	-	High Electron Mobility Transistor
LPF	-	Low Pass Filter
MAG	-	Maximum Available Gain
MSG	-	Maximum Stable Gain
MDF	-	(MDiF) Measurement Data interchange Format
MMIC	-	Microwave Monolithic Integrated Circuit
MSUB	-	Microstrip Substrate
MTA	-	Microwave Transition Analyser
PAE	-	Power Added Efficiency
PCB	-	Printed Circuit Board
PDK	-	Process Design Kit
PRM	-	Phase Reference Module
Si	-	Silicon
SiC	-	Silicon Carbide
SOLT	-	Short, Open, Load, Thru (calibration technique)
tand	-	Loss Tangent (tan δ – referring to dielectric loss)
VNA	-	Vector Network Analyser
YIG	-	Yttrium Indium Garnet

Table of Contents

1 Introduction to Novel MMIC Design Process Using Waveform Engineering

"The current through a conductor between two points is directly proportional to the potential difference across the two points."

Modern form of Ohm's Law, 1827. Georg Simon Ohm, 1789-1854.

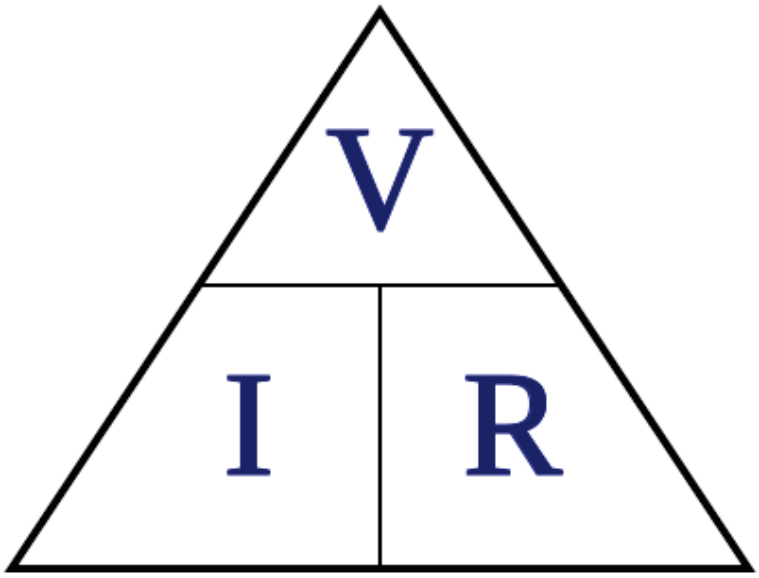


Figure 1-1, Pictorial description of Ohm's Law, the relationship between voltage, current and resistance.

1.1 Introduction

The project addresses fundamental challenges relating to further developments of RF Systems and Transduction Devices & Materials for remote sensing applications. The main objective is to improve the output power and Power Added Efficiency (PAE) of anticipated RF systems for emerging ultra-wideband radar equipment (6-18GHz) of the future. This will effectively minimise the power consumption, thermal dissipation, and the size of RF systems for a given performance, resulting in benefits to a wide range of applications with a particular relevance to mobile or field-deployable units, which are affected by significant size and power consumption constraints.

To overcome the incremental improvements in the design methodology of military RF systems the project aimed to extend techniques, developed for the optimisation of mobile communications technologies, to the higher frequencies of interest. These techniques, based on a novel measurement concept, allow complete access to and manipulation of the information contained at the interfaces between the RF system components. The approach involves the measurement and engineering (including dynamic load line control) of current and voltage waveforms and represents a major departure from currently established non-linear design and analysis methods. It is important to note that despite its novelty, the information provided still allows direct reference and comparison with past and present design techniques and thus allows the inclusion of the significant know-how which has been developed in the last few decades.

To fully utilise the additional information, which is directly obtained from this technique [3] [4] a new development methodology for broadband RF systems will be created. This will be achieved through a coherent and systematic linkage of all development stages: measurement, analysis, CAD based design and RF system testing. The resulting methodology would, for the first time, utilise the complete nonlinear information within the emergent military RF system. At present, nonlinear information is only available at few of the development stages with very limited capabilities to transfer it across development stages. This is due to theoretical limitations as nonlinear information is automatically lost when exported into present design packages; examples of which include S parameter or power spectrum measurements.

The approach described in this work allows all subcomponents and their complex nonlinear interactions to be fully accounted for. Therefore, the impact of a single design parameter on a complete RF system can be determined. This allows the investigation of design parameter sets to give the optimum performance of the complete RF system and not just its subcomponents. An example is the development of high power transmitters where power amplifiers and antennas are developed separately with little or no consideration of their complex impedance changes over the operational bandwidth leading to sub-optimal designs and performance parameters. This issue takes on additional significance in broadband RF systems, where complex interfaces between sub-components are numerous.

The methods used in the design of Microwave Power Amplifiers (MPAs) depend upon the tools available. When all that were available were S parameters, small signal

(because high output power, S parameters were difficult to measure - S22) data was used to design the best matching circuits and then the design would be tuned on the bench. One of the most difficult parts of which was knowing when to stop working on the unit, i.e. when the optimum performance been reached. In fact this approach is often still used for new processes/devices before models are available, [5]. To some extent the improvement in design tools has been a 'leap-frogging' of that of simulation, models and measurement capability. At first it didn't matter very much that the models were very crude, because the capability of the simulators was very limited. As these have improved (see Chapter 5, Nonlinear modelling) it has highlighted the deficiency of the models. These in turn put an emphasis on the measurement methods. Load pull techniques which had depended on short circuit tuners being adjusted for maximum performance, (Pout, PAE etc.), and then the tuner impedance being measured, were hugely labour intensive and required de-embedding which with the basic capabilities of hardware and software was no easy matter, began to give way to automated systems. These allowed large areas of the impedance plane to be 'mapped' thus providing the designer with evidence, not only of the location of the optimum load point but also the sensitivity of the performance characteristic to load impedance. However, as frequencies increased, the losses between the tuner and the device meant that the high reflection coefficients required, firstly with high power device fundamental loads and then with high efficiency harmonic tuning, could not be realised. Also load pull measurements on their own did not provide a complete solution as there was no way to directly incorporate this data with a nonlinear simulation package. The design engineer could construct output matching circuits to produce an impedance close to the measured optimum but the measured data could not be directly accessed by a simulator. At this point the increased computing power now available on the desktop allowed the model designers to increase the number of elements in the model and alongside new optimisation algorithms allowed them to take advantage of the increased amount of device data that was now being generated by the measurement systems. No single model has been able to capture all the nuances of device behaviour across the full range of operating conditions, as can be seen from the vast range of models developed. Some very good models were developed which were optimised to a particular narrow frequency range and specific bias conditions (such as the Motorola METModel and the Root Model), Figure 1-2, however

behaviour away from these conditions and frequencies (even by a few percentage points) could produce very inaccurate simulations.

Hence the approach successfully employed [6] was to fit trusted large signal models to measured DC and small signal behaviour and to compare the large signal performance with load-pull data measurements. For narrow band applications (<30% bandwidth) such an approach produces accurate simulations, however required DC and small signal measurements as well as the large signal to ensure that the model was accurate, and the optimisation of the model is not a trivial process in itself.

Improvements in modelling continue to be made, for example Cree included the effects of self-heating within the model [7], which assisted with one of the key differences between class A and class B biasing. When a model was received it was rare for it to include a description of the acceptable operating environment, thus it was still necessary for engineers to test devices to be used, to prove the veracity of the device models. A difficult parameter to accurately model is the transistor input reflection coefficient, partly due to its value, (high reflection coefficients) and partly due to its sensitivity (at high frequencies) to the device load impedance (due to finite S12). Although in power amplifiers the input impedance is less critical than the output impedance, it is still important in the design of the input matching network, both to avoid excessive input currents at saturation and to maximise PAE, {1-1}.

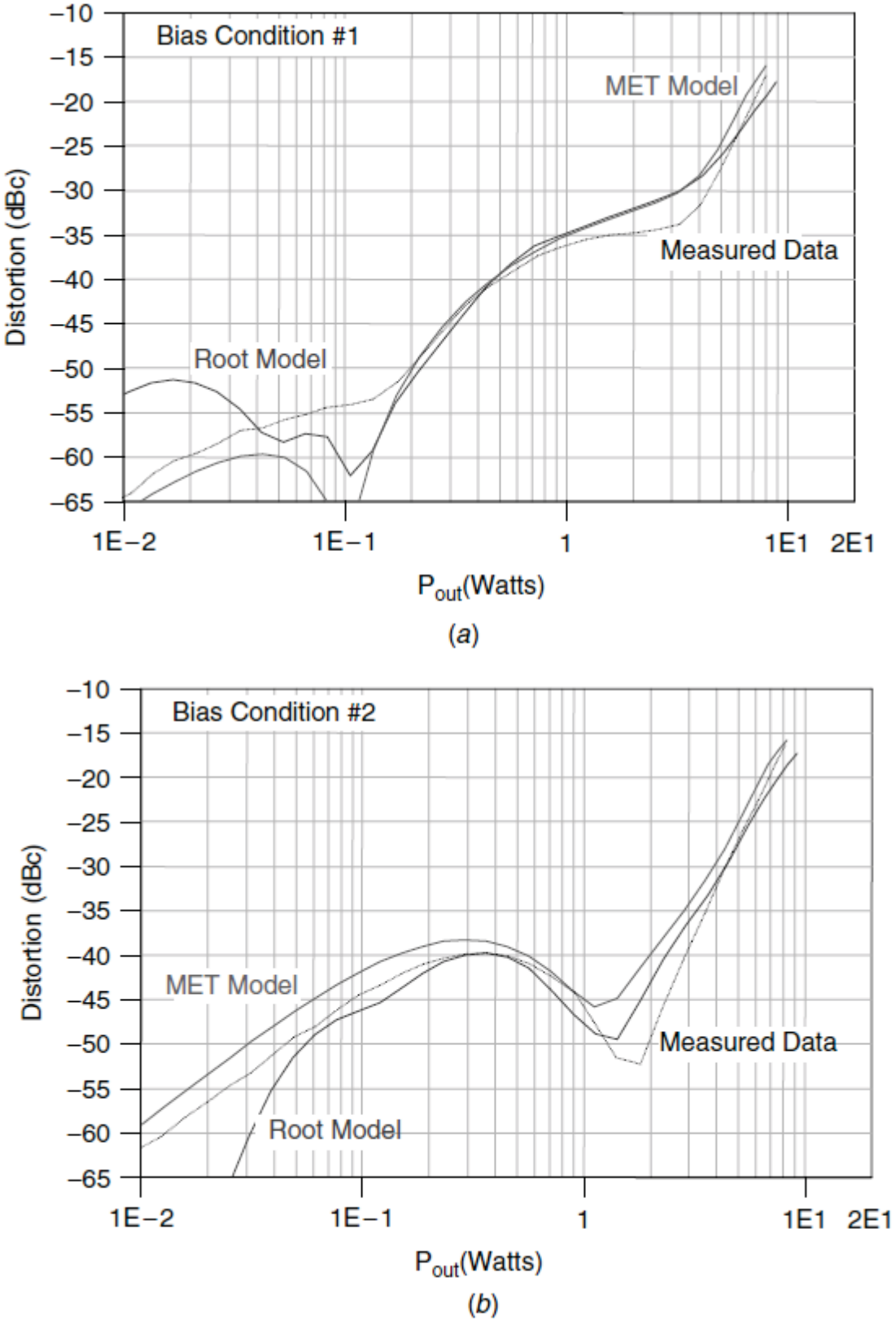


Figure 1-2, Comparison between MET Model, Root Model and measured data under 2 different bias conditions, from [39].

$$PAE = \frac{(P_{out} - P_{in})}{P_{DC}} \quad \{1\text{-}1\}$$

Amplifier efficiency can be easily improved by operation in classes 'above' class A whereby the device(s) are operated at a lower current; the RF signal itself turning the device on. The advantage of class A is increased linearity, and typically higher gain and output power. However these generalities will be shown to be only rough approximations, heating in particular greatly complicates the actual operation. A less efficient transistor will generate more heat, gain and output power are inversely proportional to temperature, thus a more efficient device may run cooler and therefore under certain operating conditions produce more power and gain than its equivalent class A configuration. Hence an important development in modelling has been the incorporation of self-heating [7], which is essential in accurate power amplifier simulation. In saturation the 'normal' biasing conditions change; the RF drive signal itself changes the quiescent bias point. As will be shown, with the correct configuration of load impedances the nominal 50% rule for maximum class A PAE operation can be 'broken'. It should be remembered that broadband class A amplifiers do not typically get close to this 50% figure for efficiency (PAE should be used as the Figure of Merit, FoM, particularly in broad band amplifiers where stage gain is often low), with results in the 20-30% range being typical, [8].

Returning to device impedances, for inter-stage matching in particular it is important to have accurate knowledge of the input and output device impedances, which has previously meant large signal S parameter measurements, [9]. It has long been recognised, [8], that to maximise the PAE over wide bandwidths there is a need for models that are not only accurate over the operating bandwidth, but at the harmonics as well. However in the classic work [2] Cripps wrote, "More and more, one is driven to the conclusion that the best way of deriving accurate models for RF power transistors is to build amplifiers, even non-optimized ones, and fit the combined circuit and device models to the measured results". Whilst possible for narrow bandwidth applications, this is obviously not suitable for wide bandwidths; however the idea of measuring actual device performance within the desired impedance environment does have considerable merit and has been reported earlier, [10]. What has been missing is the ability to fully and flexibly measure the device behaviour and capture this information in a manner applicable to modern microwave design software.

The research at Cardiff University sought to bring these various elements together, the improvement in measurement speed and capability, accurate modelling and their integration into the CAD environment. The early work, [3], [4] was centred on the narrow band frequencies for the communications industry. The research described in this work started with the creation of a wide bandwidth (1-40 GHz) Active Load Pull (ALP) measurement system, and the implementation of a Direct Look-Up Table (DLUT) Model within the MicroWave Office (MWO) design software, mirroring the developments in the narrow band application space, but in a much reduced time period thanks to the frame work defined by those who had started the investigations. During the course of the research the measurement control software was completely redeveloped and the approach to the DLUT was changed to looking up PHD model coefficients. Such progress is at the same time a great boon and a great frustration. A part of you would like to wait until “all the i's are dotted and the t’s crossed” before pursuing the new approaches; however the advantages offered by the improvements are such that they cannot be ignored. Hence some of the early work in this research has used methods that, whilst still valid as an approach, have already been overtaken by the improved models that we can use now. The proposed design methodology arising from this work will require the commercialisation of the measurement system and the integration of the measurement and simulation environments, two areas which in the industrial world have tended to be kept separate. When models can be produced quickly and viewed in the CAD tool, immediately designers will have the freedom to make changes based upon arriving at the optimum solution rather than compromising by living within the boundaries of established proven models. There will also be speed and economic benefits as any short comings in the measurements, for example impedance or power levels that have not been covered, will be quickly identified and acquired. It is much easier to acquire data whilst a measurement is in operation than to have to return to it later.

The work done by colleagues at Cardiff University will continue to develop these capabilities, whilst it can be seen that the industry is awakening to the insight and potential of the tools by the number of companies now offering large signal nonlinear measurement capability. The beauty of the Cardiff approach is that the investment in the model creation and thereby the understanding of the device truly becomes an asset of the company undertaking the work; it is knowledge that should deliver a competitive advantage, both in

the performance of the product, development cost and time to market (these latter points primarily due to reduced development time).

1.2 Applications of Broadband Amplifiers – Requirements and Performance Drivers

The term "broadband" has become somewhat over used in recent years. For example ultra-wideband (UWB) systems were defined by the United States Federal Communications Commission (FCC) as signals occupying a bandwidth greater than 500 MHz or at least 20% of the carrier frequency [11]. Associated with this decision the FCC allocated bandwidth between 3.1 and 10.6 GHz for UWB applications, thus in practice an UWB system may have a percentage bandwidth of <10% at 10.6 GHz. In the case of the research described in this thesis broad bandwidth is considered as amplifiers where the percentage bandwidth exceeds 50%. In these cases a straight forward optimal matching approach will not yield an adequate result, as will be described in section 1.3. Systems requiring such large bandwidths are found in many fields including:

- Electronic Warfare (EW) (including jamming).
- Radar.
- Television Broadcast
- Frequency hopping (covert) communications systems.
- Electro – Magnetic Compatibility (EMC) Testing.
- High data rate fibre optic communication systems.
- Multi–channel communication systems.
- Test Instrumentation.

Some of these applications require multi-octave designs, others are satisfied by having switched amplifiers still with broad bandwidths, but not covering the complete range. In EMC testing applications these amplifiers will frequently have different output power capabilities related to both the field strength required and the gain of the transmitting antenna. As with all designs there is a compromise between competing requirements, for example in a multi-band mobile phones there are a number of narrow bandwidth Power Amplifiers (PAs) which are selected as required. Thus there are switch losses associated with this approach, however it has currently been a more practical solution than using a single

broadband amplifier to cover all of the bands in one, as the individual amplifiers including the switch loss can still achieve the required output powers with better efficiency.

Efficiency is an obvious key driver in a mobile phone in order to maximise battery life. In EW systems its importance depends upon the vehicle; it may be of little importance on a naval vessel, but critical on an airborne craft. As the PA is at the end of the power train, efficiency savings at this stage have a multiplier effect. If we were to consider a very simple case of a 1 W transmit PA which had a Power Added Efficiency of 35%. If the output efficiency were increased by just 5% the impact on the overall system would be to decrease the heat generated in the PA and power supply (assuming this ran at an efficiency of 85%) by ~18% and the prime power required by ~12%. This reduction in heat generated would reduce the weight, size and power requirements of any cooling systems and thus have an additional benefit to the operation of the system, possibly allowing greater range and flight time.

Although such drivers would not carry as much weight with instrumentation systems there are still significant benefits in terms of reliability (running devices at a lower temperature), cost (smaller power supply required), noise (fewer/slower fans), besides the environmental benefit of consuming less power. It can thus be seen that in most systems the ability to improve the efficiency is an attractive proposition. The difficulty has largely been in achieving this in line with other aspects of the design.

Considerable effort has been made in making PAs more efficient but often at the expense of other parameters such as linearity. Non linearity leads to distortion of the fundamental signal, interference between communication channels and added complexity in control systems (the linear relationship between input and output power is broken). To combat this ever increasingly complex solutions have been sought, from purely analogue feed forward systems [12], to digital pre-distortion and most recently envelope tracking and dynamic biasing [13]. These however come at a price, be it of additional components (and hence weight, size and cost), detailed system characterization (time and test equipment) or sacrificing peak performance (to benefit average). Thus efforts that increase the primary efficiency of an amplifier stage without negatively impacting upon the other parameters will, if not remove the need for additional measures, at least make their requirements less stringent.

1.3 Broad Band Matching

Intrinsically, matching is the arrangement of a circuit such that a certain optimum performance is achieved. In low noise amplifier designs [14] there exists an optimum input reflection coefficient, Γ_{OPT}, which is required to be presented to the input of the device to achieve the minimum noise figure; this is not however the same impedance required to conjugately match the device for maximum gain. Maximum power transfer occurs when the source and load impedances are equal and has been understood since the early days of electric motors run from batteries [15]; in RF circuit design there is the added complexity that impedances (devices and loads such as antennas) are rarely purely resistive and for maximum power transfer the reactive element must also be taken in to account. A conjugate is the equal but opposite, for example the reactance of an inductor ($j\omega L$) that cancels out (resonates) a capacitor ($-j/\omega C$). This makes clear the fundamental problem of broadband matching, this cancellation occurs at one frequency, f_{res}, {1-2}, and away from this frequency either the inductive or capacitive reactance dominates.

$$f_{res} = \frac{\omega_{res}}{2\pi} = \frac{1}{2\pi\sqrt{LC}} \qquad \{1\text{-}2\}$$

Broad band matching circuitry seeks to resolve two key problems, (i) how to achieve a wider bandwidth with a minimum reflection coefficient and (ii) how to minimise the number of matching element sections for a given bandwidth. The issue has been around a while, having been addressed by two of the fathers of circuit theory, Bode and Fano, [16] [17]. The Bode-Fano Criterion states that for a specific load impedance there exists a theoretical minimum refection coefficient that can be achieved with arbitrary lossless matching network. The lossless element is important, in practice all matching elements will have some resistive loss, but in output circuits for power amplifier applications for maximum power transfer very low loss elements will be employed approximating to the lossless condition. In other applications where loss can be tolerated wider bandwidth matches can be achieved, indeed resistive matching can achieve very wide bandwidths (at the cost of insertion loss). For the most common RF and microwave device output equivalent circuit approximation, a shunt resistor capacitor combination, the Bode-Fano Criteria states that:

$$\int_0^\infty \ln\frac{1}{|\Gamma(\omega)|}\,\partial\omega \;\leq\; \frac{\pi}{RC} \qquad \{1\text{-}3\}$$

The minimum reflection coefficient within a bandwidth ($\omega_2 - \omega_1$) is:

$$|\Gamma|_{min} = exp\left(\frac{-\pi}{RC(\omega_2 - \omega_1)}\right) \quad \{1\text{-}4\}$$

Again these equations {1-3} and {1-4}, are idealised and rely on an infinite number of pure matching elements, and thus determines the best that can be achieved *in extremis*. Also note that the definition of bandwidth is up to the user, but needs to be consistent throughout. In filter applications it is sometimes considered as the 3dB bandwidth, i.e. a measure to the points where the insertion loss has fallen to 3dB. Alternatively it may be defined in terms of the pass band ripple. Similarly with amplifiers, although in this case the bandwidth is often defined and an allowable minimum gain or gain variation within the pass band stated.

A practical solution to what can be achieved is to consider the matching network as a Chebychev matching transformer with the ripple made equal to $|\Gamma|_{min}$; the number of elements required are available from standard filter tables, [18]. An alternative approach was suggested by Carlin, [19], rather than using idealised equivalent circuits for the output of the device the technique uses measured impedance load data, $Z_L(j\omega) = R_L(\omega) + jX_L(\omega)$ over the frequency range of interest. Matching networks can be designed by considering only the resistive elements of the source and load initially, the topography of the matching is then selected such that the parasitic element (e.g. capacitance) is absorbed into the first element of the matching network.

It is necessary at this point to introduce the concept of Quality factors, Q, as they are integral to bandwidth. There are a number of subtly, but critically different forms of Q and it is important to define clearly which form is being used. Unloaded Q, or Q_U is defined by the centre frequency ω_0 of the network and its bandwidth and independent of any other parameters:

$$Q_U = \frac{\omega_0}{(\omega_2 - \omega_1)} \quad \{1\text{-}5\}$$

$$\omega_0 = \sqrt{(\omega_1 \omega_2)} \quad \{1\text{-}6\}$$

The loaded Q, Q_L, {1-8} is by contrast dependent on the source resistance, R_S, the load resistance R_L and the Q of the matching elements themselves, Q_E. The Q of the matching elements is defined by their ability to store charge and is the ratio of the stored energy to

dissipated energy. For capacitors this is the ratio of the capacitive reactance to the Equivalent Series Resistance (ESR) {1-9}, and for inductors the inductive reactance to the series resistance of the coils, {1-10}. For distributed elements the Q is more complex, but if we consider stored charge gives a good indication as to value. An important caution is that Q is frequency dependent and thus in components usually at a specific frequency(s).

$$R_P = \frac{R_S R_L}{R_S + R_L} \quad \{1\text{-}7\}$$

$$Q_L = \frac{X_P}{R_P} \quad \{1\text{-}8\}$$

$$Q_C = \frac{X_C}{ESR} \quad \{1\text{-}9\}$$

$$Q_{IND} = \frac{X_L}{R_S} \quad \{1\text{-}10\}$$

Thus if we redefine {1-4} in terms of Q_U and Q_L, we can see the minimum reflection coefficient we can achieve:

$$|\Gamma|_{min} = exp\left(\frac{-\pi Q_U}{Q_L}\right) \quad \{1\text{-}11\}$$

Turning the expression round we can also see what bandwidth can be achieved for a particular reflection coefficient and 'quality' of matching components:

$$\frac{(\omega_2 - \omega_1)}{\omega_0} = \frac{1}{Q_U} = \frac{-\pi}{Q_L \ln \Gamma_{min}} \quad \{1\text{-}12\}$$

More commonly we define amplifiers in terms of Return Loss (RL) and this is related to reflection coefficient by {1-13} and the insertion loss, L_T, as a result of the mismatch between impedances is defined by {1-14}.

$$RL(dB) = -20\log|\Gamma| \quad \{1\text{-}13\}$$

$$L_T = -10\log(1 - |\Gamma|^2) \quad \{1\text{-}14\}$$

A common approach to matching is to use either "T" or "Pi" circuits, the name coming from the configuration of the matching elements. The Smith Chart, [20] provides a useful tool in visualising the effects of the different combinations of elements. Series elements follow the impedance contours, whilst shunt elements the admittance circles. Figure 1-3 shows how the combination of series and shunt elements can be used to 'move'

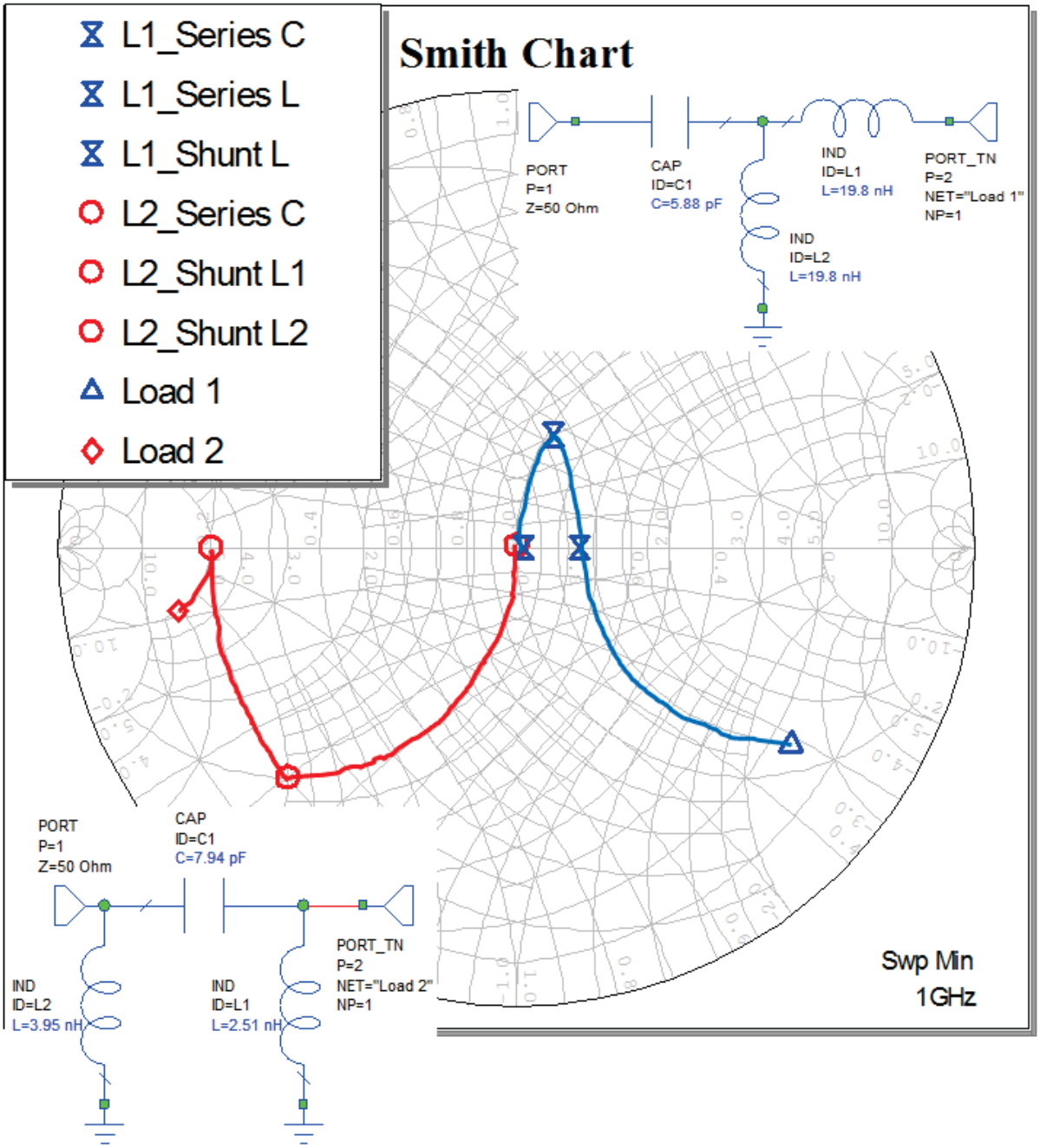


Figure 1-3, "T" (blue) and "Pi" (red) matching approaches. Load 1: 300Ω//1pF, Load 2: 10Ω//10pF

from the load impedances (Load 1 is 300Ω with 1pF parallel capacitor and Load 2 is 10Ω with 10pF parallel capacitor) to 50Ω. Of course there are many other combinations that would produce the same results.

Matching elements are not pure elements. For discrete components besides the parasitics intrinsic in the construction, inductance in leads, resistance in metallisation, unwanted capacitance between plates, etc. there are also those associated with connecting tracks (pads) and the circuit board itself. In MMIC implementations the physical proximity of the components causes cross coupling and as frequency increases the phase length of

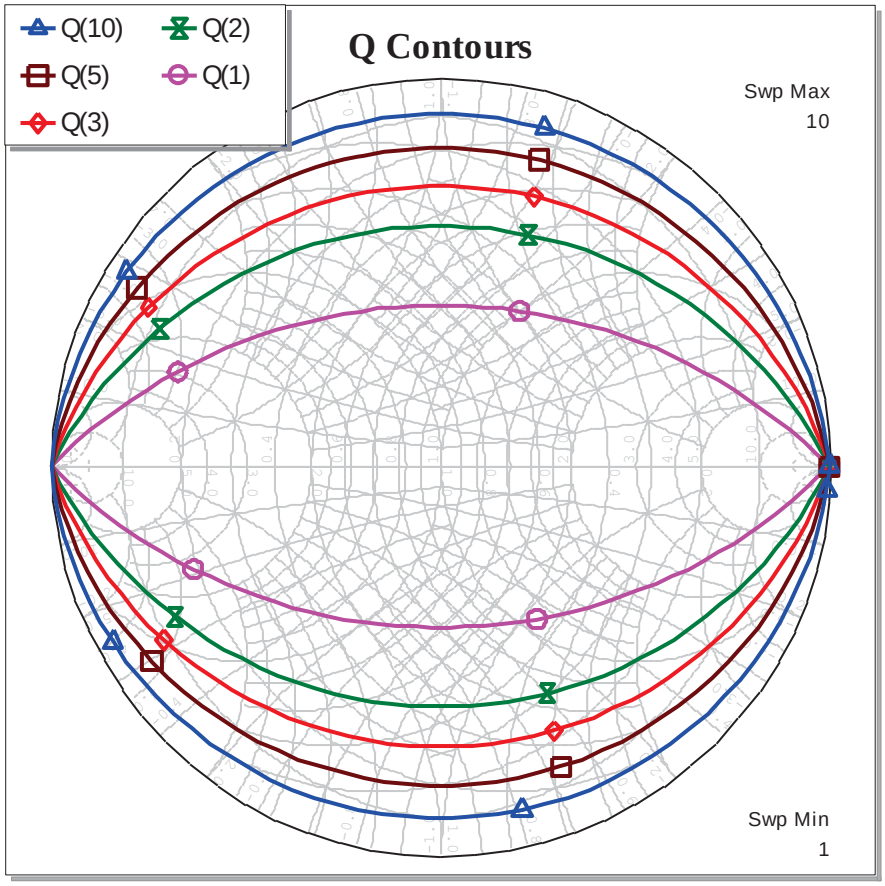


Figure 1-4, Contours of constant Q (Q=1, 2, 3, 5 and 10)

matching elements prevents them from behaving in the ideal manner. Although these factors will complicate the matching circuit design, understanding their implications is more important than the fact that they exist. In input matching and decoupling circuits losses can be beneficial in improving stability and parasitic elements can be included in matching circuit design thereby reducing the number of elements required. In the main output circuits (i.e. not bias lines) it is essential that losses are kept to a minimum, hence the use of high quality materials and components in these areas.

Plotting constant Q factors on the Smith chart produces contours where the impedance points lying on the contour all have a constant ratio of X/R from the impedance of the point R ± jX. Figure 1-4 shows contours for Q ranging from 1 to 10. For a single transformation, bounded by the resistive centre line along the diameter of the Smith Chart and a constant Q contour, the transformation ration, R_{ratio}, is given by {1-15} [21] and for multi-section (n) matching the relationship becomes {1-16}. In practice in amplifier matching circuits[1] the number of sections will rarely go above 4 due to the increasing insertion loss

[1] In filter design, particularly in low loss constructions such as waveguide, high number of sections (>10) are often used to achieve the required performance.

and reducing incremental benefit. It should be made clear that this solution does not necessarily realise the optimum matching solution, however it is probably the simplest approach and yields smaller component values.

$$R_{ratio} = 1 + Q^2 \quad \{1\text{-}15\}$$

$$1 + Q^2 = \sqrt[n]{R_{ratio}} \quad \{1\text{-}16\}$$

The Q curves can be employed to assist designing a broad band match. Starting from Load 2 of Figure 1-3, the normalised admittance of the load, G + jB, is 5 + j3.14, hence a Q≈0.63. If a successive Pi matching approach is taken until the centre of the chart is reached, staying within the Q contour of the load (Q=0.63), then after 6 shunt elements and 5 series capacitors the 50Ω, optimum match is reached, Figure 1-5. Note that the final elements, C5 and L6 do not need to move between a point on the 0.63 Q curve to reach the centre of the chart. It is also worth noting that an 'acceptable' match may be obtained with fewer elements; also drawn on the chart is a circle of constant Voltage Standing Wave Ratio (VSWR – a common term to refer to the match of a load derived from the days when standing waves reflected from a load were measured with voltage probes). If a VSWR of 1.5:1 meets the match specification then a solution can be reached with two fewer elements.

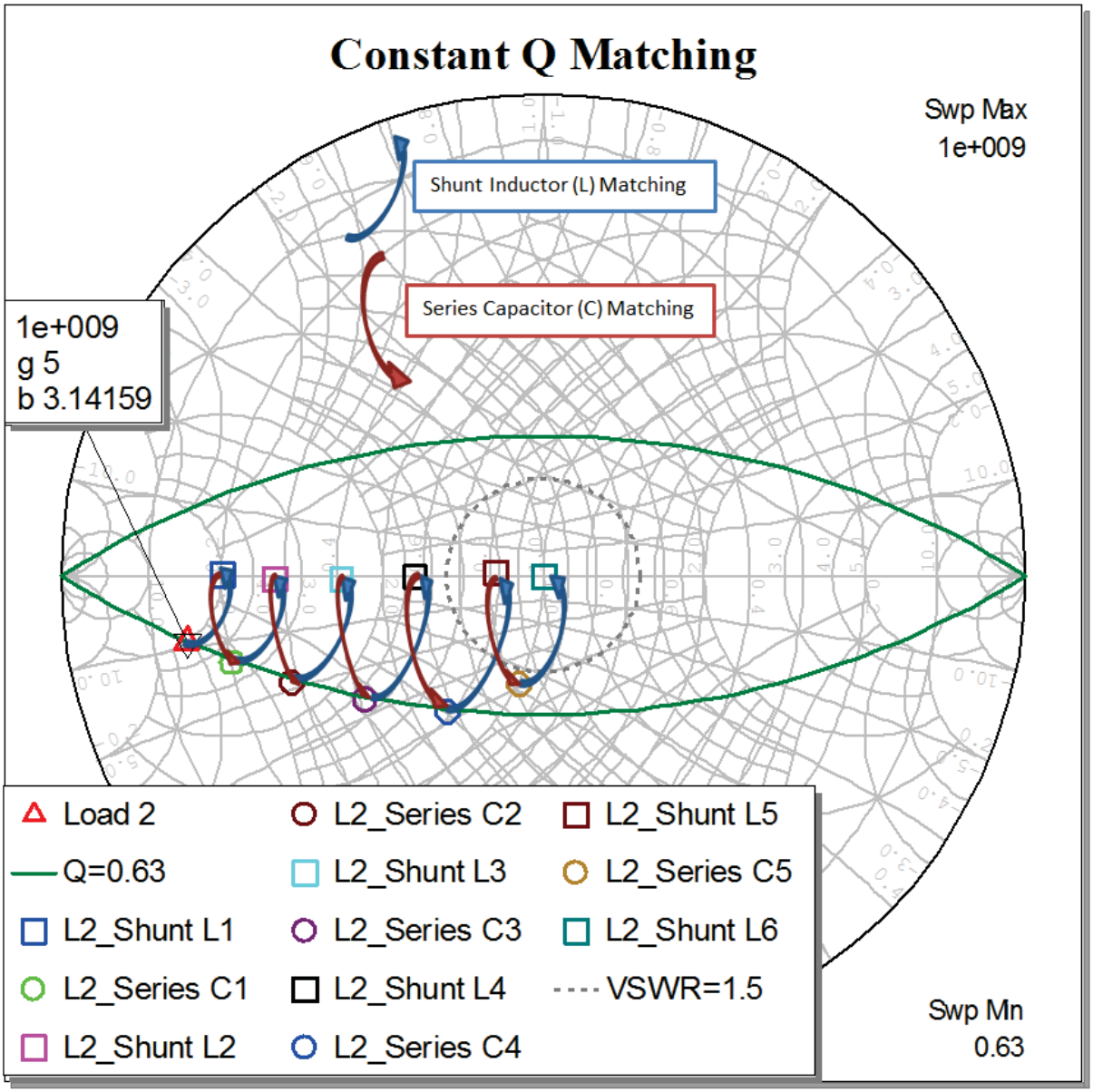


Figure 1-5, Broadband matching staying within Q of load.

Comparing the bandwidth of the 3 element match of Figure 1-3 and the 11 element of Figure 1-5, it is shown in that there is a significant increase (as would be expected) using the broadband match, Figure 1-6. It should be remembered that these results are obtained using ideal elements with no resistive loss or parasitics, in practice these low insertion losses would not be obtained.

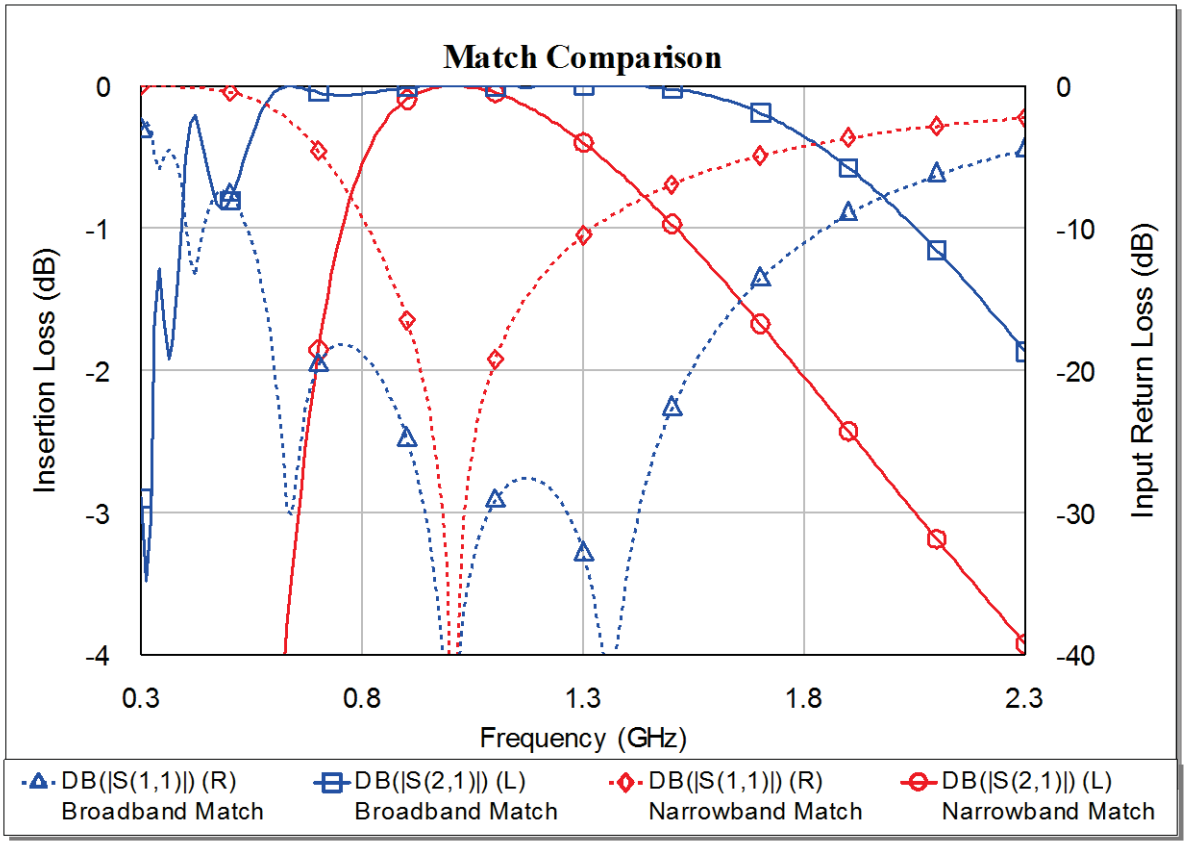


Figure 1-6, Comparison between 3 element narrowband and 11 element broadband match.

Calculating the corresponding values of return loss, mismatch loss and VSWR, it can be seen from Table 1-1 that although in order to maximise the power output we must minimise the loss due to mismatch, going beyond a return loss of 15dB (Γ=0.18) there is a diminishing return. Although this may seem trivial it is important to realise the implications of specifying a particular match for an amplifier. The better matched a transistor; the less power is required to meet a specification. Often an isolator will be added to the output of an amplifier in order to meet an output return loss requirement, but this may have 0.5 dB of insertion loss and will do nothing in terms of translating the output impedance of the device to 50Ω. Whilst there may be good system considerations for adopting an isolator (such as gain ripple on long cables) in PAs it loses hard won power. A solution based upon the best power match that can be achieved would be more efficient.

Γ	0.1	0.18	0.2	0.25	0.35	0.4	0.5	0.71	0.8
RL (dB)	20	15	14	12	9	8	6	3	1.9
L_T (dB)	0.04	0.14	0.18	0.28	0.58	0.76	1.25	3.0	4.44
VSWR	1.22	1.43	1.50	1.67	2.1	2.33	3.00	5.85	9.00

Table 1-1, Γ, Return Loss, Mismatch (Transmission) Loss and VSWR.

Any real impedance can be matched to the system impedance at a single frequency, the difficulty is doing it over a bandwidth and all amplifier circuits need to have at least a limited bandwidth to account for changes in behaviour with temperature. There are an infinite number of combinations of matching elements able to move from an impedance on one part of the Smith Chart to another [22]. However, because they are largely treated as pure lumped elements (capacitors and inductors) their use in MMIC PAs is restricted to an understanding of the theory, in practice the matching elements used are complex due to their distributed nature. The most common distributed matching elements are impedance transforming transmission lines and open and short circuit stubs. Combined with a series transmission line these can match an impedance over a defined area, this is best explained graphically as in Figure 1-7. High power transistor output impedances will typically lie within the green shaded area, (large periphery hence low resistance and high capacitance) and so often the first matching element will be an open circuit stub. In contrast lower power devices have a higher resistance and lower capacitance and can be matched with the shorted stub.

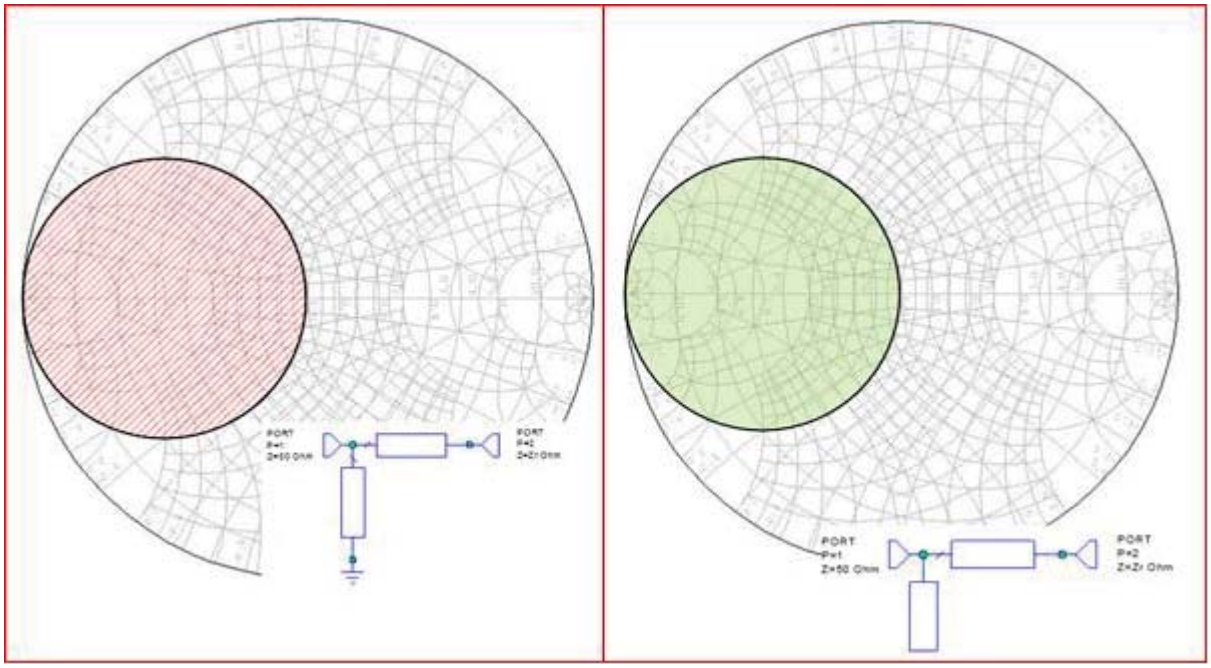


Figure 1-7, Stub matching approaches: (left) short circuit stub can match any impedance outside the red shaded area; (right) open circuit stub can match any impedance within the green shaded area. Impedance to be matched is to the right of the network.

So far, these discussions refer to matching at a single frequency and to 50Ω. As a rough guide a single quarter-wave matching structure between impedances of a ratio of 6:1 can achieve a bandwidth of ~22%, (15dB return loss). Using multiple sections this can be

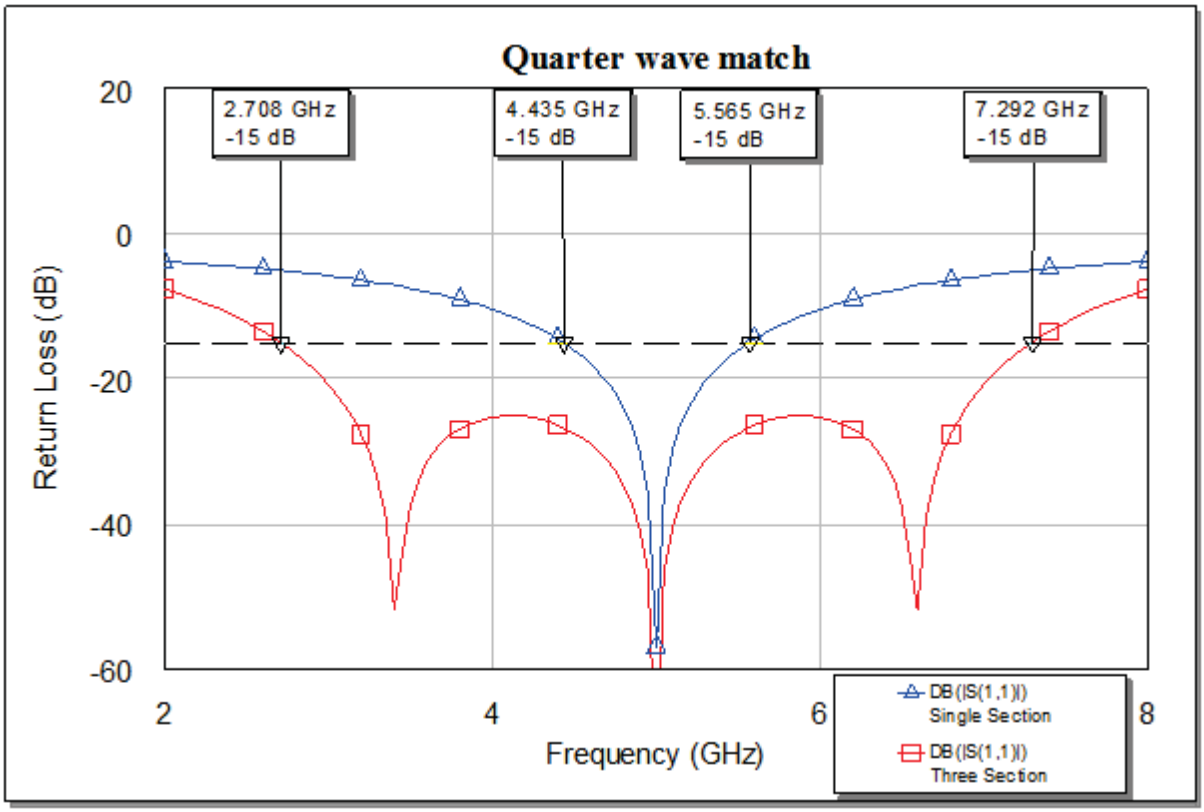


Figure 1-8, Comparison of the bandwidth achieved from 1 and 3 section quarter wave transformers between the same resistive loads.

increased, e.g. matching the same load with 3 quarter-wave transformers, the bandwidth that can be achieved increases to ~90%. Clearly, wider bandwidths can be achieved by reducing the impedance ratio; hence the benefit of using higher voltage and wide bandgap materials such as GaN with higher output impedance and, conversely, why with technologies such as LDMOS[2] with its high output capacitance, only narrower bandwidths can be achieved in the microwave region. Figure 1-8 shows the improvement from 1 to 3 sections, using ideal transmission lines and purely resistive loads. The impedance of the quarter wave matching line, Z_T, is determined from {1-17}. For multiple sections one first needs to calculate the intermediate impedance between the sections, $Z_{i(n)}$, [23], {1-18}, where N is the number of steps, and R_L the load resistance. In the case of the first step $Z_{i(1)}$ then the impedance can be found from {1-20}.

$$Z_T = \sqrt{(Z_0 \times R_L)} \qquad \{1\text{-}17\}$$

[2] LDMOS devices are made with high operating voltages (~50V) for high power operation, however to handle the power the periphery is very large which leads to very high output capacitance.

$$Z_{i(n)} = \Delta \times Z_{i(n-1)} \lim_{2} N \quad \{1\text{-}18\}$$

$$\textit{where,} \quad \Delta = \left(\frac{R_L}{Z_0}\right)^{\frac{1}{N}} \quad \{1\text{-}19\}$$

$$Z_{i(1)} = \Delta \times Z_0 \quad \{1\text{-}20\}$$

Using impedance transformers any impedance within the areas defined in Figure 1-9 can be matched. This assumes an infinite range of transformer impedances is available which, in microstrip, is typically limited to between 25 and 90Ω. If coplanar waveguide (CPW) is used [24] a wider range of impedances (~15 to 120Ω). There is much debate amongst MMIC designers over whether it is better to use microstrip or CPW circuits. The differences are not only in the transmission characteristics of the two approaches but also in the available models and the range of their application, ground connections, processing costs and density and implementation of components on the surface. In power amplifiers a key consideration is the thermal; heat sinking requires good metal connections to the backside of the substrate. In both microstrip and CPW, although the impedances can be determined mathematically, [2], it is often simpler to put approximate values in a simulator and allow the optimiser to produce the best possible result within practical constraints.

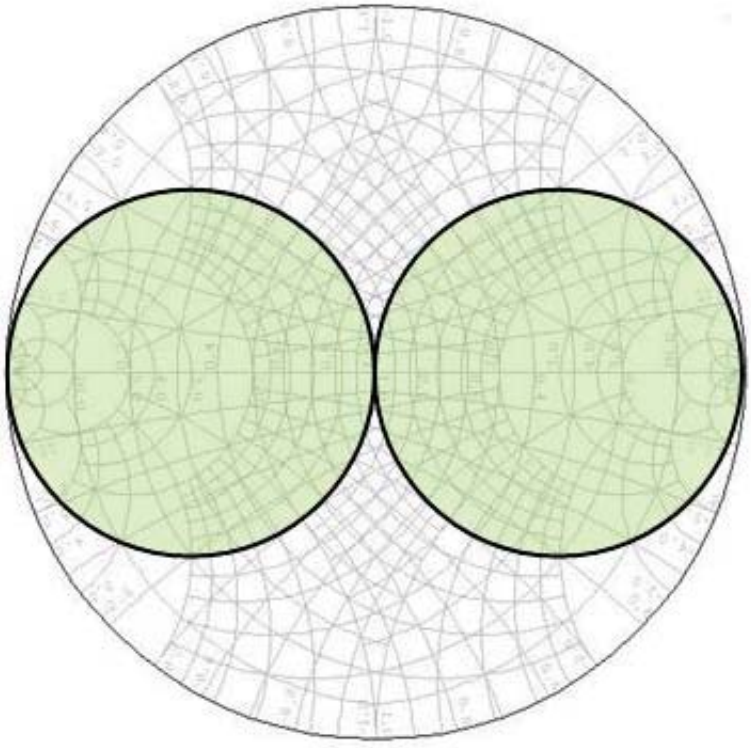


Figure 1-9, Single transformer matching space.

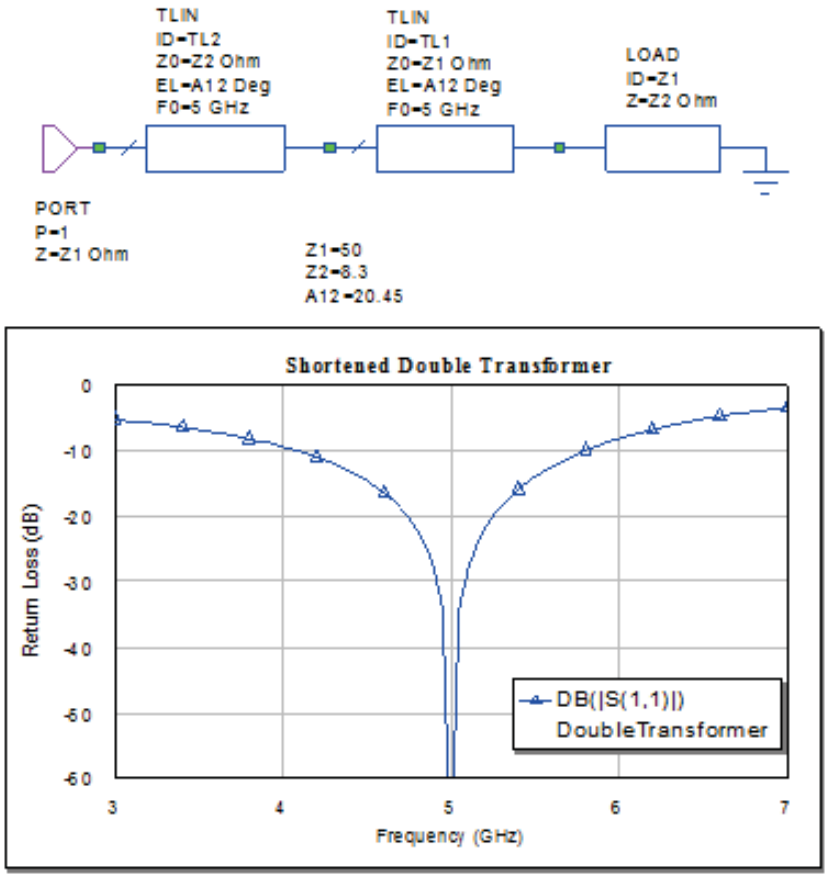


Figure 1-10, Double short transformer matching.

A number of other topologies offer impedance matching, without the size penalty of λ/4 structures. Similar performance to the λ/4 transformer can be achieved using two lines of impedances the same as those of the terminations, [23], as shown in Figure 1-10. In this case the combined length of the two lines is less than that of a single λ/4 transformer with little sacrifice to the bandwidth. A similar version to this involves using a capacitively loaded λ/8 line. This is particularly useful if the load impedance that is being matched to has a capacitive reactance as this can be absorbed into the matching capacitance. Figure 1-11 shows the circuit for a λ/8 capacitively loaded transformer used to match 12.5 to 50Ω, where the capacitors and the line impedance were allowed to vary in the CAD circuit optimiser, Figure 1-11 (a). If this approach were to be used to match to a load of 12.5 - j80.0 (a 0.4pF shunt capacitor at 5 GHz) it would be possible to replace the 0.2pF with the 0.4pF capacitance of the load and re-optimise with the result shown in Figure 1-11 (b), varying only the line impedance and the port 1 capacitor. A further advance on this is to replace the output shunt capacitor with an o/c stub, Figure 1-11 (c). This has a slightly narrower bandwidth due to the lower limit on the impedance of 25Ω, but is a printed component. In MMIC solutions the lumped element may be a better solution as it could be smaller, but

would depend upon the impact of the inductance of the via and the actual electrical performance of the MIM (Metal-Insulator-Metal) capacitor.

The discussion so far has used ideal components; capacitors, inductors and transmission lines of a fixed impedance and electrical length. In practice all components used in RF and microwave circuits have characteristics determined by their physical construction and assembly; capacitors have loss, additional capacitance including that between the terminals and ground and inductance, and these elements are distributed throughout the component. The importance of this distribution increases with operating frequency, as the size gets closer to a wavelength. Even ground connections have a physical length associated with their implementation which has an increasing impact with frequency.

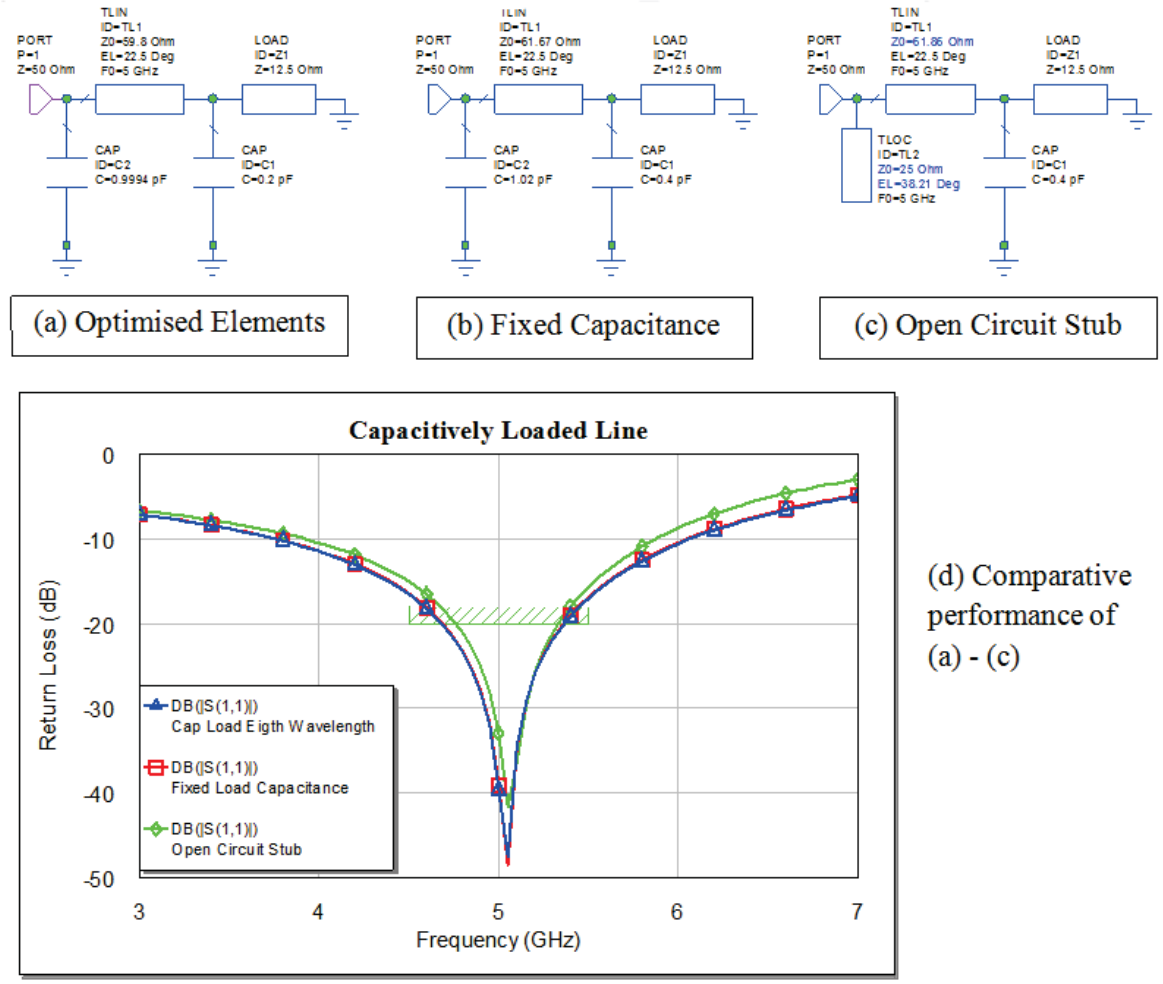


Figure 1-11, Capacitor loaded λ/8 transformer matching.

Distributed elements are those structures whose physical dimensions fundamentally determine their electrical characteristics. Hence the tolerances and repeatability of the manufacturing processes has a direct effect on the performance. On alumina circuits and

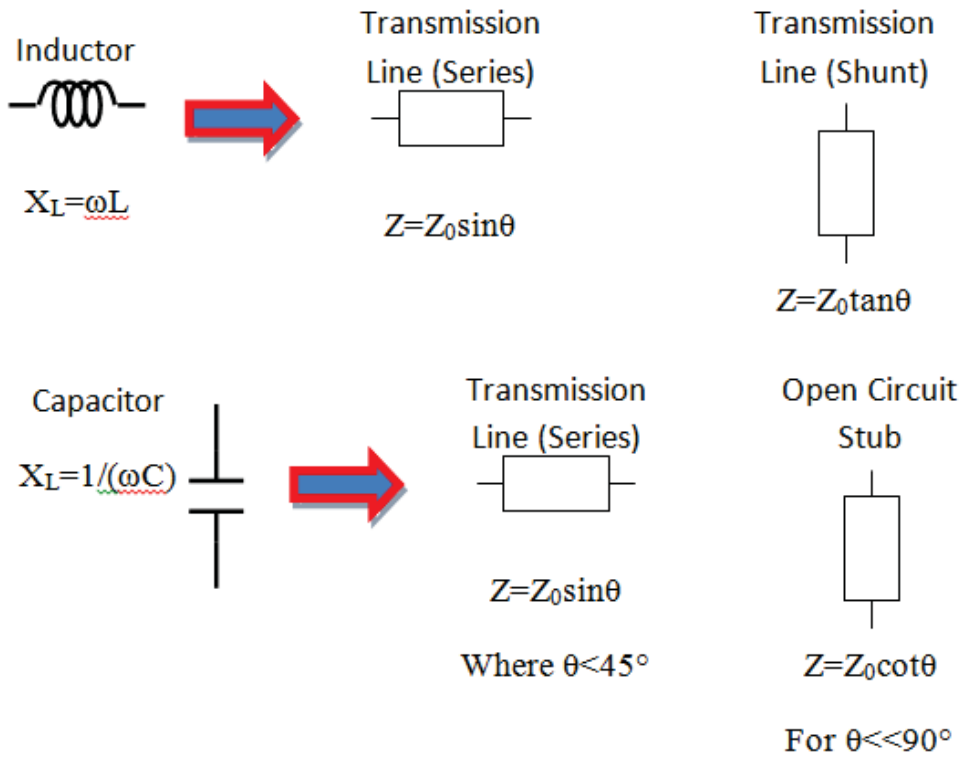
MMICs, capacitors and resistors can be incorporated directly during the circuit fabrication process. For resistors, as with their lumped element equivalents, resistive pastes can be used in thick film circuits or a NiCr layer in thin film. These resistive materials are specified in terms of ohms/square and the resistance, *R*, is proportional to the ratio of the length, *l*, to the width, *w*, {1-21}.

$$R = resistivity \times \frac{l}{w} \qquad \textbf{\{1-21\}}$$

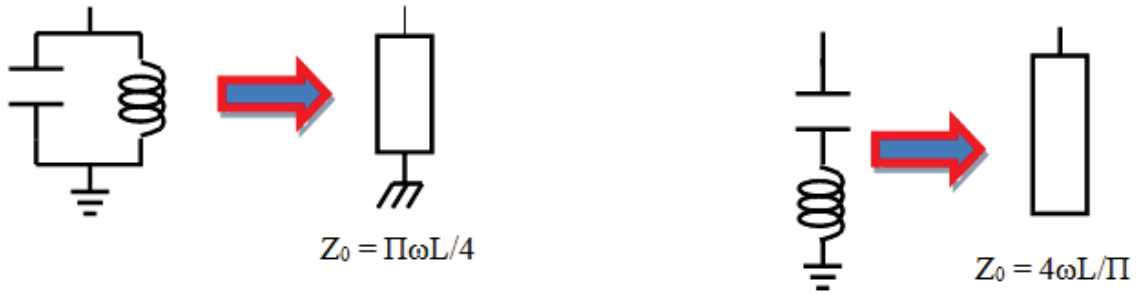
The width can be adjusted to match that of connecting transmission lines. Common resistivities are 50 and 100 Ω/sq.

Capacitors can be incorporated by putting down a dielectric on top of conductors and then overlapping with a conductive paste/plating. Although these have been proposed for use in soft substrates, particularly using conductive inks, there has not been a large scale take up as yet. Inductors on microstrip can be difficult to distinguish from high impedance lines. The exception is where the line is wound in a spiral (round or square sided) and is popular in MMIC applications. This is used less frequently in hybrid circuits due to the relatively large size and lower Q compared to wound components. Similarly, interdigital capacitors are used less often in hybrid circuits. An exception is when edge coupled lines used in some bandpass filter applications fulfil two functions, that of filtering and DC blocking. For narrow band applications where the volumes are extremely high, edge coupled lines may be justified on their own due to cost considerations.

The most common distributed components are transmission lines. As mentioned earlier, the impedance ranges that can be created are limited due to moding and etch tolerances. In practice, circuit structures are also limited by the ability to simulate them within design tools. Distributed circuits have re-entrant properties, that is, the impedances repeat (approximately) at multiples of 90 and 180°. The behaviour of distributed components with frequency differs to that of pure elements. For example, at a specific frequency a shunt inductor of impedance X_L can be replaced by a shorted transmission line of impedance, $Z_0 \tan\theta$ where Z_0 is the characteristic impedance and θ is the electrical length. However, whereas X_L increases linearly with frequency, the line impedance increases with tanθ, which is periodic. There are a variety of equivalents between lumped and distributed components, some of which are given in Figure 1-12. The realisation of distributed series



In the special cases where θ=90° (quarter wavelength) at the resonant frequency:



And when θ=180° (half wavelength) at the resonant frequency:

Figure 1-12, Equivalent Lumped and Distributed Circuits.

capacitors, as mentioned above, is very difficult in hybrid applications. An exception is where very small values are required which can be achieved by using narrow gaps.

A short circuited (S/C) transmission line or S/C stub is basically a transmission line with one end terminated in zero impedance, but this becomes more difficult to achieve as

frequency increases due to parasitics. Common methods of producing a short circuit include using a via hole, edge wrapping and a solid ground plane (with or without via holes). One of the benefits of a distributed circuit is that a short circuit at the end of a λ/4 line looks like an Open Circuit (O/C) at the other end of the line at the design frequency. Thus another way of creating a short circuit is to attach a λ/4 o/c stub at the point where a short circuit is desired. This is inherently a narrow band structure; to broaden the bandwidth a radial stub can be used. Even more effective is a double radial or butterfly stub, [25]. These solutions are often adopted as frequency increases and the inductance/phase length of via holes has more impact, or when an RF short is required but not one at DC. In bias feeds it is required that DC current be injected into the circuit but that the feed arrangement not load the RF matching network. Although microstrip impedances are typically limited to between 25 and 90Ω, it is possible to create effectively lower impedances by adding two o/c stubs in parallel. A comparison of the performance of various distributed stubs is shown in Figure 1-13. The best performance is from the via, the dispersion in this being due to the inductance/phase length of the connection to the ground plane (backside of the substrate in MMICs).

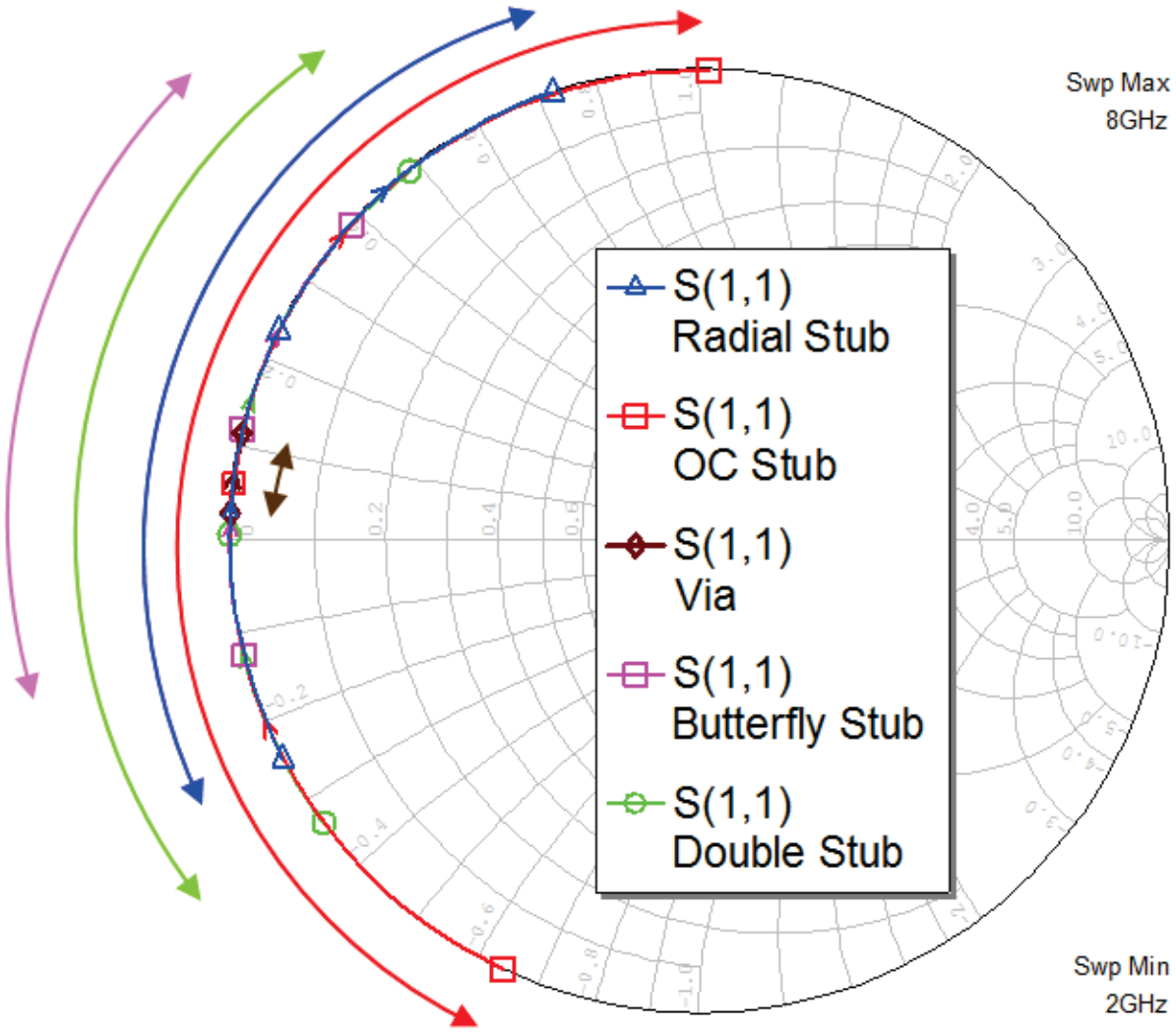
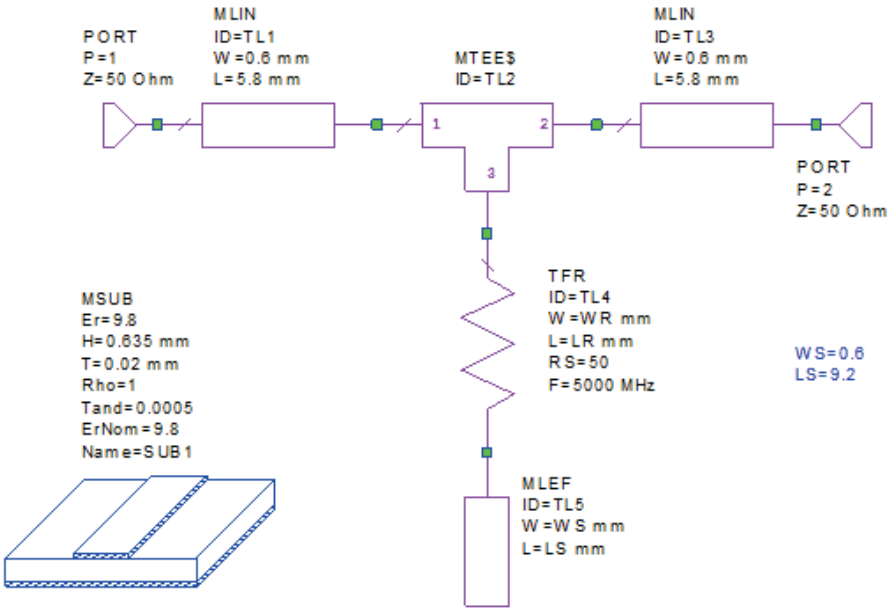


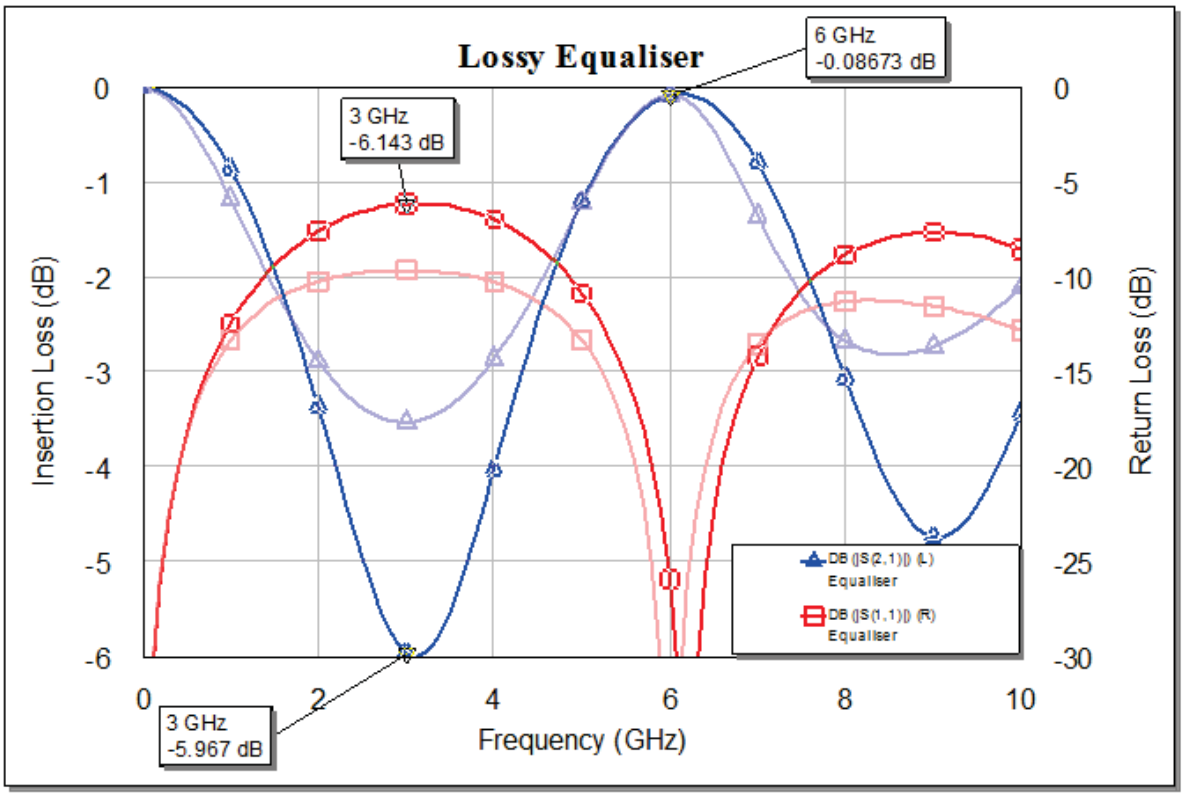
Figure 1-13, Relative performance of different short circuit elements, the lower the dispersion the closer to the ideal.

A double open circuit stub (on either side of the main transmission line) creates a wider band short circuit than a single radial or open circuit stub, however carries with it the danger of non TEM modes being generated as well as a lack of clarity as the correct modelling method – a wide double stub looks physically like a low impedance short transmission line. This presents a discontinuity and is best analysed in an E-M simulation. When designing with microstrip elements it is important to remember that the models used were developed and optimised for specific substrate thickness to line width ratios. There are often a variety of models for the same structures and it is important to choose the most appropriate one for the materials and frequency range used. Where appropriate models do not exist, the use of E-M simulation is necessary. This can be used for a specific section of the design, as simulating a whole circuit in this way can be time consuming and hence difficult to optimise.

Transistors have a natural gain slope in |S21| of 6dB/octave, which can be a particular problem in broadband amplifier design. A method for compensating for this is to use lossy stubs or equalisers, (input and interstage – not on the output). These do not



(a) Schematic of Lossy Stub



(b) Performance, Blue S21 and Red S11, as resistance is varied between 50Ω (faint traces) and 25Ω (bold traces).

Figure 1-14, Lossy stub equaliser approach.

provide a DC path to ground and hence do not upset device biasing. The basic version of this approach consists of a resistor connected to an o/c stub, as is shown in Figure 1-14 (a). By altering the resistance different slopes can be achieved as shown in Figure 1-14(b) and summarised in Table 1-2. In narrow band applications the equaliser can be used for stopping high frequency oscillation by introducing loss at the problem frequency.

Nominal Resistor Value (Ω)	Approximate Slope (dB/octave)	Worst case Return Loss (dB)
25	5.9	6.1
50	3.4	9.6
75	2.4	12.2
100	1.8	14.1

Table 1-2, Equaliser performance as a function of resistor value.

In order to simultaneously achieve improved match, stability and flat gain over wide bandwidths feedback can be employed [26]. The theory has been well documented [27], for power amplifier applications series feedback is rarely used due to the difficulty in simultaneously capacitively decoupling the source and providing a good thermal path to ground. With FETs shunt feedback must, by necessity, incorporate series capacitance as well as resistance to separate the gate and drain DC voltages. Both the capacitor and resistor will have parasitic inductance. This is can actually be of benefit as by incorporating inductance in the feedback model, it can help to increase the RF impedance and increase the gain at the upper end. In discrete designs leaded resistors, not normally used at microwave frequencies, can be specifically chosen to introduce the required inductance. Using shunt feedback, bandwidths of multiple octaves can be achieved. The feedback elements must be capable of handling the power levels of the signals travelling through them, but as they get larger to handle higher powers their parasitic components increase which makes their use more difficult.

One of the problems with introducing feedback to discrete PAs is that of incorporating the feedback elements within the available physical space. Not only are the gate and drains separated by several mm, but there is also usually a large flange. Sometimes it is possible to construct the feedback network in three dimensions, going over the top of the device rather

than remaining planar. This approach is labour intensive and not suitable for automation. It also tends to be more susceptible to performance variations due to the lack of consistency in component forming and placement. Heat sinking of the feedback elements can be an issue in both discrete and MMIC design. Feedback elements dissipating heat should ideally be thermally removed from the active device, but physical separation introduces phase shifts which must be accounted for. One approach, suitable for moderate bandwidth applications has been outlined [28]; originally this was intended for low noise applications to avoid the feedback introducing noise back to the input. However, it has the advantage for power applications of separating the feedback components and the amplifying device. The circuit incorporates two $\lambda/4$ lines between the device and the feedback resistor. Thus on a low dielectric material (~2.2) at 5 GHz, the feedback resistor can now be about 20mm offset from the main track. The feedback arrangement is shown in Figure 1-15 (a) and consists of a low pass filter arranged such that there is a 180° phase shift at the operating frequency (or towards the top of the band in wider bandwidth applications), such that the feedback has little effect on the performance. At lower frequencies, the resistor is 'in band' and adjusts the amount of feedback, thus reducing the bottom end gain. The device with feedback is matched with a simple single section impedance transformer on the input and output. Another advantage of the approach is that the bias can be incorporated within the feedback loop, Figure 1-15 (b), thus having less impact on the fundamental matching. The circuit was optimised to give a flat gain over 2-5 GHz and a stability factor >1. A further advantage of using feedback is that it reduces sensitivity to device variations, including using those from different manufacturers. In the example shown in Figure 1-15, the circuit components values were optimised for the Mitsubishi MGF0951 (red traces); the performance with the similar but still markedly different, Figure 1-15 (c), Eudyna FLC107 (blue traces). Figure 1-15 (d) shows an improvement in that devices stability and a similar gain shape. Although this approach improves stability over parts of the band, care must be taken to carry out a thorough stability analysis over a wider range as at specific frequencies the feedback can actually cause oscillation.

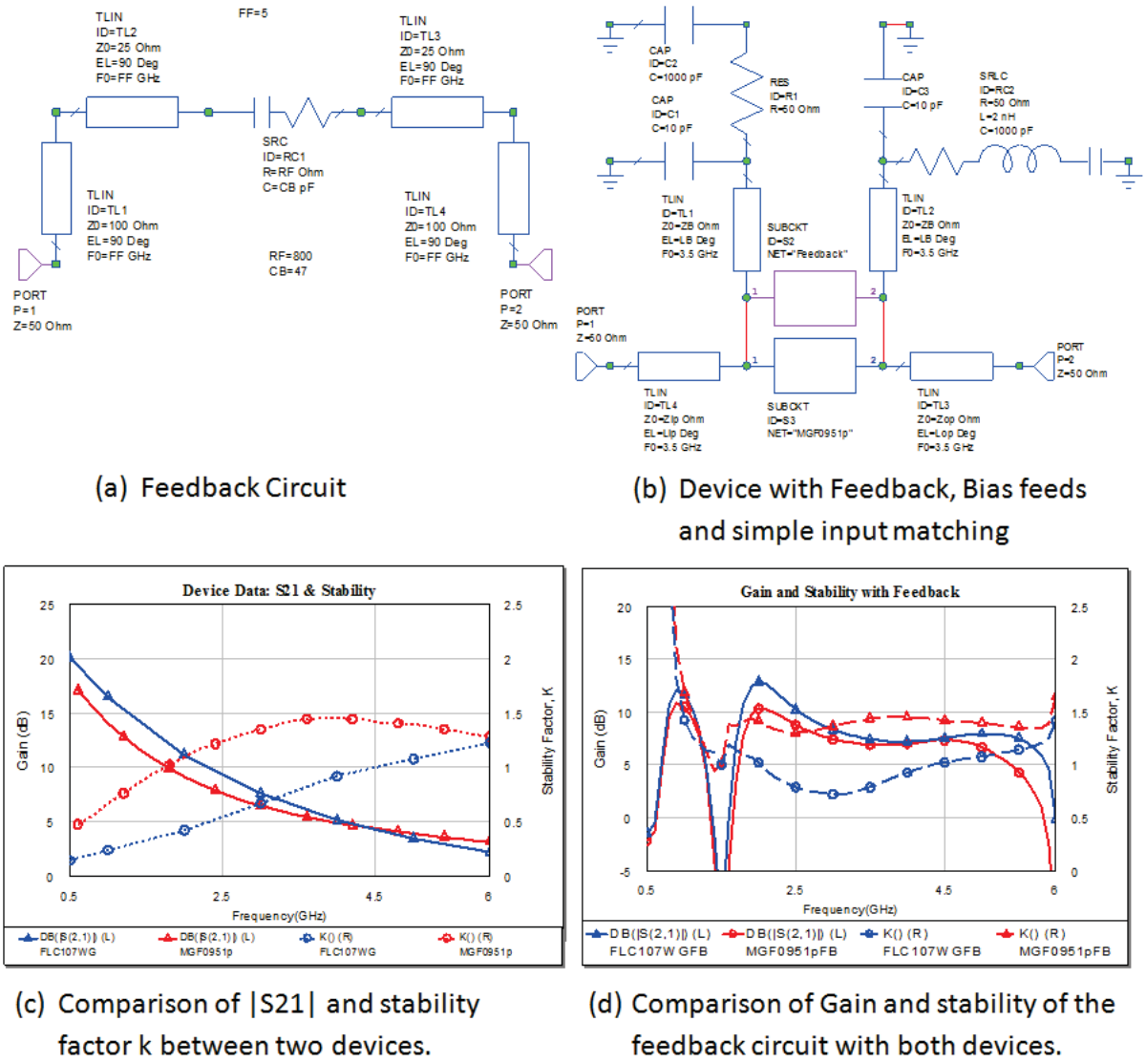


Figure 1-15, Device with feedback.

Most amplifier text books describe the use of stability factor, k, and the use of stability circles to define the unstable impedance areas, i.e. those areas where if presented to the device there is the potential for oscillation. Although stability analysis should be carried out up to the maximum operating frequency of the device, in practice for microwave devices the S parameter data available will rarely go below 500 MHz at the bottom of the range and at the top end the likelihood of oscillation decreases as the frequency increases due to increasing circuit losses. Therefore it is necessary to design circuits which will inherently ensure stability at the low frequency end and provide stable impedance terminations up to the frequency where the |S21| = 3dB. To ensure low frequency stability the main approach is to resistively terminate the bias networks. This may involve using large inductors in parallel so that the DC can still pass. It is important to remember that the AC coupling capacitors used between RF stages tend to have values in the pF range and that these will

effectively be open circuits in the MHz region where transistors will have very high gain. Feedback can be applied between the gate and drain bias feeds away from the active RF circuits [2]. Where large inductors are used it is important to contain the RF fields as these can become a source of oscillation through coupling. Ferrite beads can be particularly useful in this respect. A low value resistor in series with gate will raise the gate impedance at a cost of gain. To compensate for this a capacitor can be inserted in parallel to bypass the resistor at the operating frequency; this will be examined in more detail in section 1.4. Isolators have a DC path to 50Ω and hence are very effective for improving low frequency stability. However, care should be taken to ensure that at higher frequencies the isolator does not present an unstable impedance, as shown in the measured data taken from a coaxial X band isolator, Figure 1-16. Above the top of the operating band, at just over 12GHz the isolator presents a highly reflective impedance. This could potentially cause stability problems and shows that they cannot be presumed to be a complete solution to stability issues.

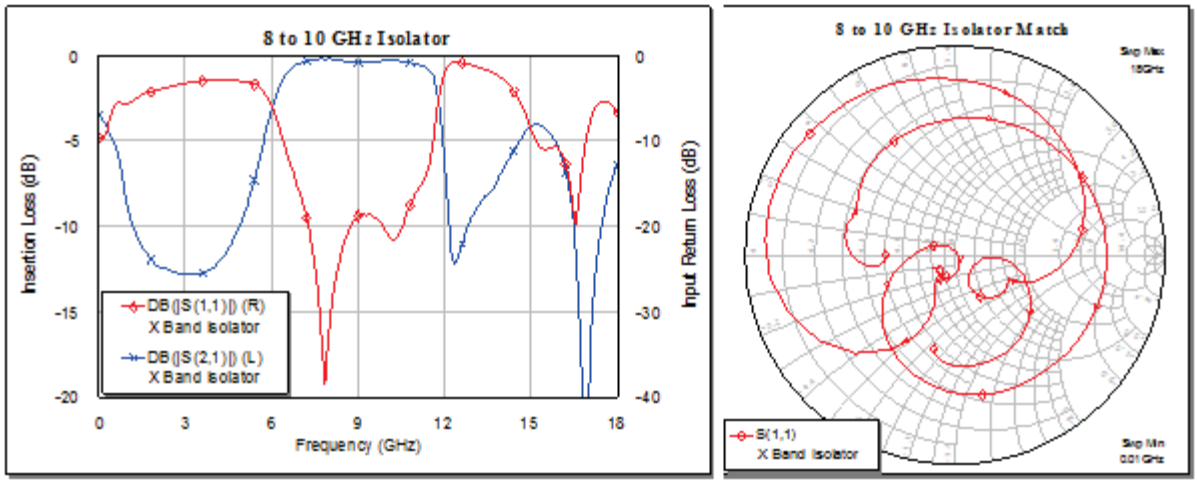


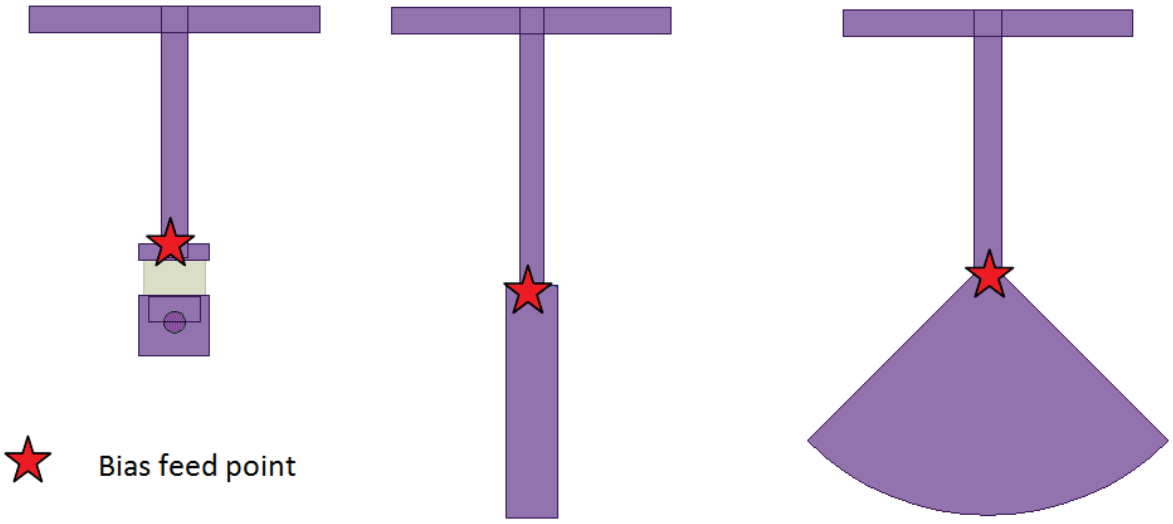
Figure 1-16, Typical X band isolator performance.

It is necessary to introduce the bias to the device without detrimentally affecting the RF performance. In the case of the drain, the DC resistance of the feed must be kept as low as possible. For a high power transistor this is especially important as the voltage drop along the feed line will reduce the output power. In contrast the RF impedance of the feed must be high, unless it is included as a part of the matching circuit itself as will be discussed later. A good solution is to introduce the bias at a low impedance point as close to the device as possible. An alternative is to use short circuit matching stubs, but replace the link to ground with a decoupling capacitor and join a high impedance RF inductor at this point. For discrete designs at lower microwave frequencies this can be done with several turns of 0.5mm wire

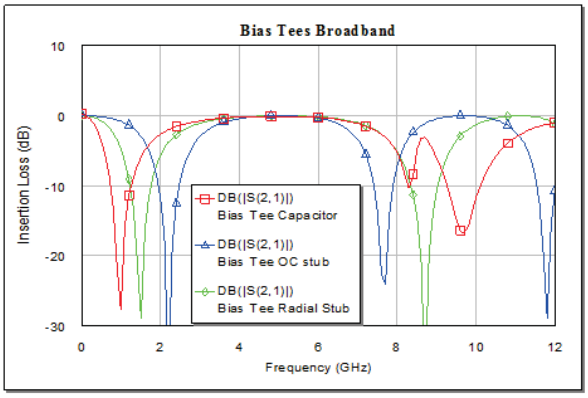
on a 2.5mm former (5 turns tightly wound is ~40nH and <0.01Ω). In MMICs printed spiral inductors can be incorporated. Increasing the spacing between the turns decreases the inductance and the capacitance between the turns.

Discrete inductors can be tested independently, e.g. using a SMA gold-plated flanged connector by soldering one end of the coil to the flange and the other to the trimmed centre pin. Observing the S11 response on a vector network analyser will show any resonances in the frequency band, altering the spacing of the coils can move these in frequency. The impedance of such a coil varies from ~250Ω at 1 GHz to over a thousand at 5 GHz (ignoring any resonances). This method is suitable for broad bandwidths, however its repeatability is poor and it is not suitable for automated assembly. Printed spiral inductors are more difficult to test as separate items but can be analysed in an E-M simulator. Due to limitations of the inductance that can be realised in a MMIC spiral inductor; 'off-chip' components are often incorporated. Air wound coils can be purchased from a number of suppliers, using standard compact footprints. As the frequency increases the number of turns required to make a high impedance inductor decreases and at X band single loops of wire may be sufficient.

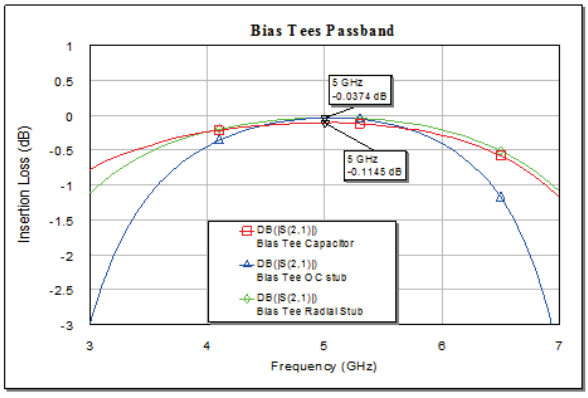
For narrow band designs a popular solution is to supply the bias through a 90° short circuit stub. This appears as an open circuit at the junction with the main line. In order to introduce a bias voltage the short must be open circuited at DC, which can be achieved by a number of methods e.g. coupling capacitor, 90° open circuit stub or radial stub, as shown in Figure 1-17 (a) – (c). The relative bandpass characteristics can be seen in the graphs in Figure 1-17 (d) and (e). The bias voltage would be introduced at the points marked with a star. In case (a) the capacitor the line length is adjusted to account for the parasitics of the capacitor. More complicated versions using several sections can be created for broader bandwidth characteristics. The impedance and length of the stubs may also be altered to assist with the device matching. Where the λ/4 lines are not wide enough to handle the DC current they can be selectively plated up or in discrete designs a thicker wire soldered to the track to increase the DC current capacity without impacting the RF performance significantly. This wire can be 'looped' off the board at the short circuit point (star) to link to the bias feed or go to extra low frequency decoupling. Not shown is the in-line DC blocking capacitor that is required to isolate the DC from the preceding and/or following stages.



(a) Capacitor (b) O/C Stub (c) Radial Stub



(d) Broadband Performance of Bias Networks



(e) Passband Performance of Bias Networks

Figure 1-17, Bias feed approaches.

Bias can be introduced at the isolated port in Branchline and Lange couplers (described in chapter 2). The isolation resistor must be AC coupled and DC blocking in the output line must be able to handle twice the RF power of the standard configuration. This is not normally done on the input unless the devices are well matched as V_g is used to set I_d.

Besides analysing the behaviour of the devices over the operating frequency range it is important to consider what happens out of band, particularly at the low frequency end. The gain of transistors increases dramatically as frequency is reduced and so it is important that signals (and noise) are correctly terminated. As discussed earlier the capacitors used for decoupling RF signals tend to be in the range of 1 to 20 pF for microwave frequencies, larger capacitors have significant parasitic inductance and multiple resonances. Thus where wide

bandwidth decoupling is required multiple capacitors of different types are used. In addition, the modulation bandwidth also needs to be considered.

The increase in modulation bandwidth has increased the complexity of the bias circuit decoupling. In order to minimise the impact on the modulating signal it is necessary to present a constant impedance to these low frequencies. However, it may also be necessary to filter out specific frequencies (such as power supply switching) and to provide a high impedance to the RF signal. For non-class A operation the supply must be able to provide the transistors with peak currents far in excess of the quiescent as the devices are turned on and off by the RF signal. This must be done without also modulating the supply voltage as change here will alter the transfer characteristics of the device and cause distortion. This usually requires large capacitors for charge storage close to the device.

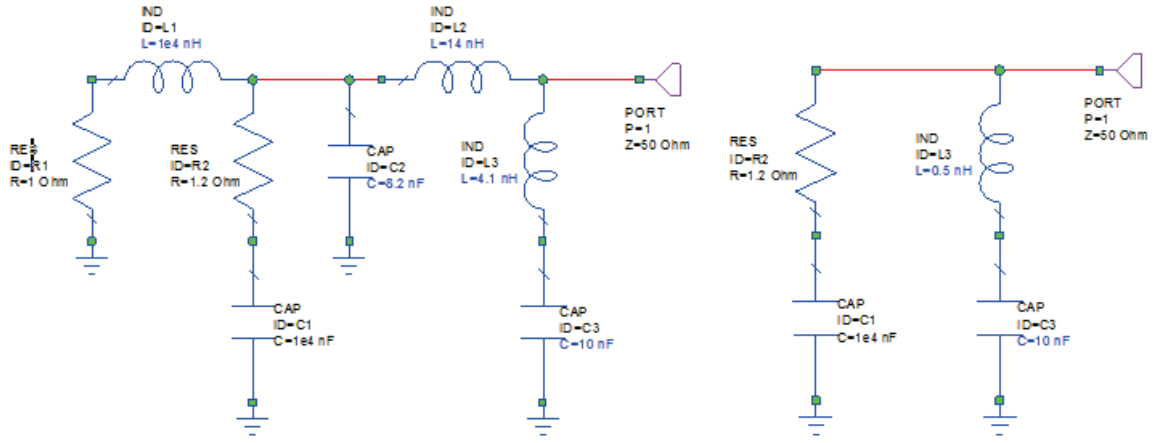
Analysis of bias decoupling has been well covered in [2]. The design of the constant impedance bias networks can be treated as a filter problem. Figure 1-18 shows how the low frequency impedance varies with frequency between a simple decoupling circuit and one designed for constant impedance. The parasitics of the components can be absorbed into the extra elements, for example the 0.5nH parasitic inductance of the 10nF capacitor shown in Figure 1-18 (b) can be included in the 4.1nH inductance in Figure 1-18 (a). The effect of the low frequency impedance can be seen in the sidebands of digitally modulated signals such as W-CDMA. An imbalance between the these sidebands is commonly referred to as “Memory Effects” and are a phenomenon in the time domain due to thermal transients and charge storage causing the bias conditions to change, [2].

It is quite common to see an arrangement of, for example, 1pF, 1nF and 10 µF capacitors on the end of a bias line. Very large decoupling values may be used to reduce spikes due to inductance in bias lines, especially when using test fixtures. For below band signals it is often advisable to include a resistive termination. In the gate bias this can be applied in series for smaller devices and incorporate the gate limiting resistor. In the drain circuit this is not practical due to the high currents. Instead, the resistor is placed in series with one of the high value capacitors.

Another factor to consider when operating devices in test fixtures run directly from laboratory power supplies is the resistance between the PSU and the device, especially if discrete bias ‘tees’ are used. A total resistance (R_{PS}) of up to 1Ω would not be unusual and this could reduce the saturated output power by up to 0.5dB. The highest power devices

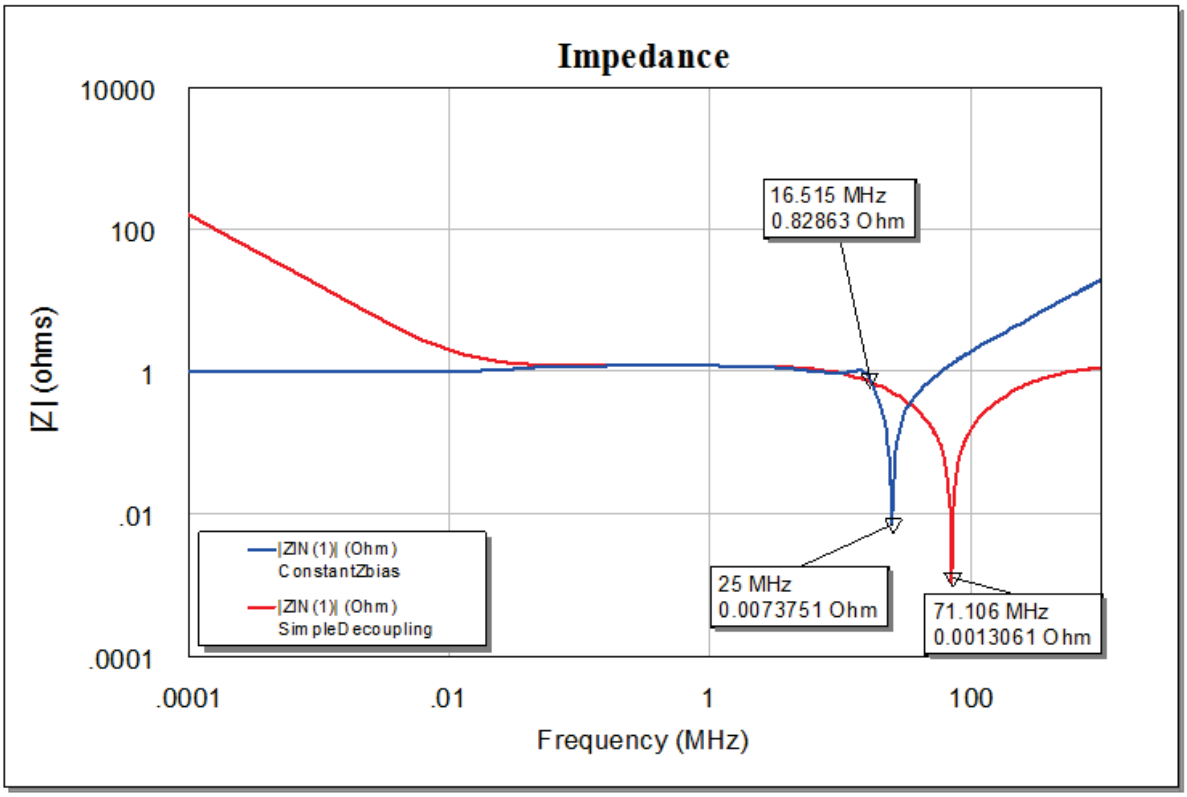
tend to be biased in higher efficiency classes than A, and as a result I_d increases with drive power, although the PSU output voltage remains constant the voltage on the device will drop by I_d x R_{PS}. This is also an important consideration when testing the devices on their own.

The changing bias voltages during switch-on may lead to oscillation. This can be seen by touching the circuit and killing the oscillation after the bias has stabilised. It may be possible to stop this oscillation by speeding up the rise time of V_d. If the oscillation



(a) Left – constant Z bias

(b) Right – simple decoupling



(c) Impedance variation with frequency.

Figure 1-18, Bias decoupling networks.

frequency is substantially below the band of operation, additional decoupling of the bias lines may help. If the oscillation occurs above the operating band, a 'lossy' stub may be required on either the input or the output. The big problem arises when the oscillation is within the operating region. This means that either the input or output matching impedance is presenting a load that is in the unstable region of the transistor's operation, which is changing as the device V_d ramps. It is important to determine whether other stages are contributing to the unstable conditions. If there are a number of stages in series without any isolating elements then these will also present changing impedances during power on. Once it has been established which stage is causing the problem, a more complicated bias sequence may be required. Instead of V_g being set to the required voltage for operation it is set to the pinch-off voltage, V_p until the drain voltage has had time to establish, V_g is then adjusted to the value for the required I_d. It is necessary to check for oscillations during switch on over the operating temperature range of the device, particularly the lower end where the gain is highest.

1.4 Input Impedance Matching

The prime consideration of this work is the output matching networks as these determine the key performance characteristics of power stages, output power and PAE. However input matching cannot be ignored, at high frequencies gain is hard won and cannot be loosely discarded, also too great a differential across the operating a frequency band in the insertion loss of the input matching circuit can lead to devices either being too heavily overdriven (into damage regions) or having insufficient power to reach compression. An important factor to keep in mind when determining the load that the input matching circuit 'sees' is the device output load impedance. Due to the finite nature of the reverse isolation (S12), mainly due to parasitic and intrinsic capacitances, the load has an impact on the input reflection coefficient. The relationship between device input reflection coefficient Γ_{IN} and the load Γ_L can be described in terms of S parameters, {1-22}. This has been expanded by Narhi [29] to describe circles of constant VSWR on the input impedance plane.

$$\Gamma_{IN} = S_{11} - \frac{S_{12}S_{21}\Gamma_L}{S_{22}\Gamma_L - 1} \quad \{1\text{-}22\}$$

As will be shown later the input impedance is also a function of the input drive power. A positive impact of this is that the match tends to improve as the device is driven in to compression, however the designer must therefore decide at which power level to target the matching circuit. For example for a harder limiting design matching to the small signal input impedance will result in a faster decrease in gain above compression. Alternatively designing to the high power input impedance will tend to improve the gain linearity. The extent of these effects will depend upon to what extent the gate source and gate drain capacitance are dependent on voltage.

In general for very broad band power stages lossy input matching circuits are used [30], [31] to improve the input match and compensate for the intrinsic gain roll-off of the device with increasing frequency. As a rule power devices tend to have a low input impedance which contributes to stability problems. A common technique employed is to connect a parallel resistor capacitor network to the input; a simplistic explanation of this is that the capacitor value is chosen to bypass the resistor as frequency increases; reducing the insertion loss (the resistor is chosen to improve gain flatness and stability). A more rigorous determination has been derived [32] that shows how the values can be calculated from the equivalent input circuit of the device to be matched. Figure 1-19 shows how a simple R-L-C equivalent circuit can be used to approximate the low power (below P1dB) input impedance. Also included in the graph are the large signal S11 at power levels of 19, 20, 21 and 22dBm and the swept input power (2-22dBm) Γ_{IN} at 4GHz.

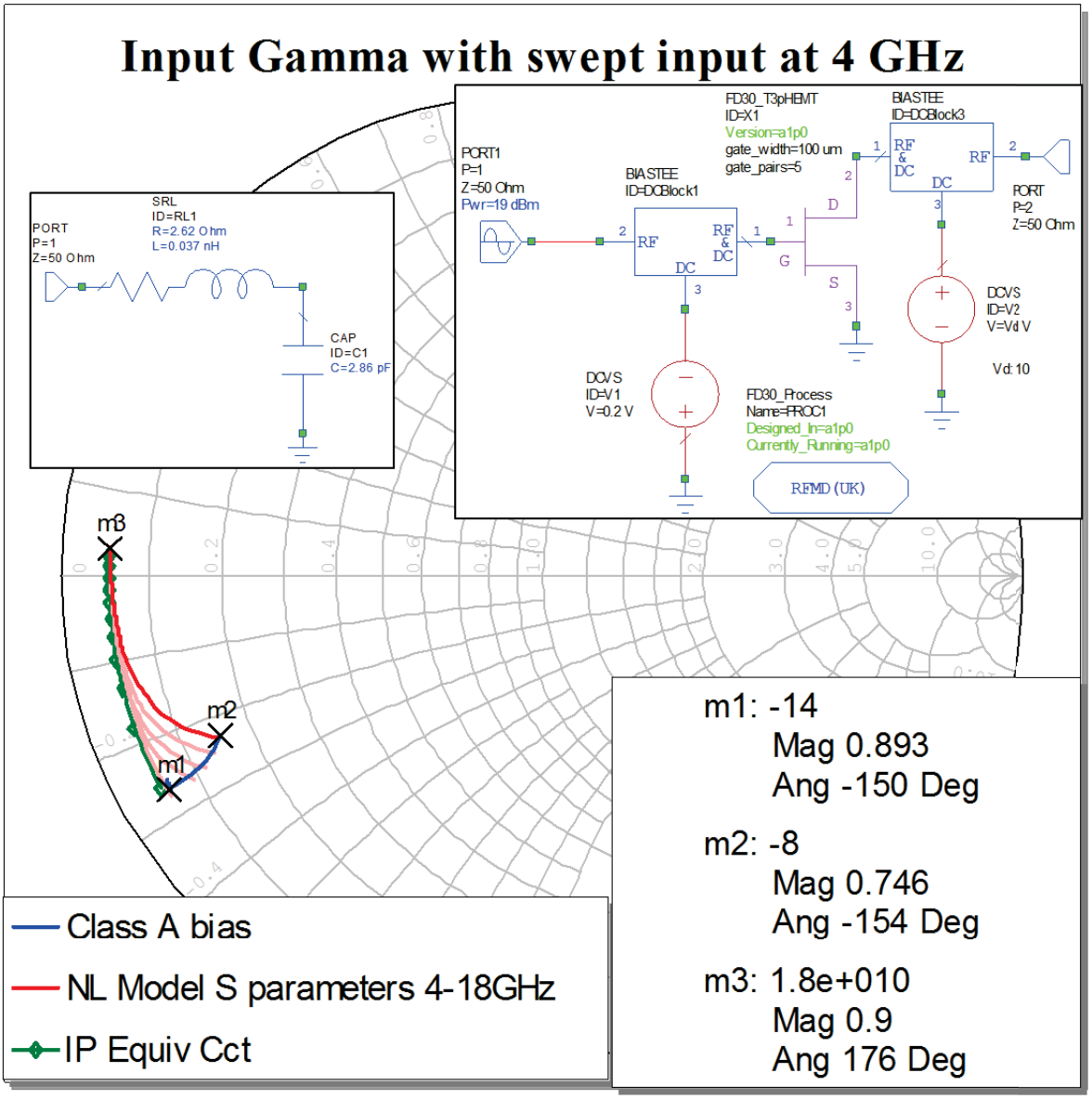


Figure 1-19, Examination of input match: equivalent circuit inset (green trace), large signal S11 (red trace 22dBm), 4GHz Γ_{IN} with swept input power (blue trace).

In this method of designing the input matching circuit, initially only the equivalent circuit inductance and resistance are included. The input impedance of the internal and external elements together is Z_{IN}, {1-23}. If the intrinsic resistance and inductance are R_G and L_G, then the external parallel resistor and capacitor, R_P and C_P, can be adjusted such that the imaginary part of the input impedance is set to 0, {1-24}.

$$Z_{IN} = R_G + j\omega L_G - \frac{R_P}{1 + j\omega R_P C_P} \quad \{1\text{-}23\}$$

$$L_G = \frac{R_P C_P}{1 - (\omega R_P C_P)^2} \quad \{1\text{-}24\}$$

(When ImZ_{IN} =0)

$$ReZ_{IN} = R_G + \frac{L_G}{R_P C_P} \quad \{1\text{-}25\}$$

In theory this means that the imaginary part of the input impedance can be cancelled out and the real part set to whatever impedance is required. In practice however, the higher R_P, the higher the loss and as has been said before gain is an expensive commodity, particularly in wideband MMICs. It is therefore important that the insertion loss of the matching circuit is monitored alongside the match. Figure 1-20 shows the results of optimising the parallel RC circuit for $Im|Z_{IN}|=0$ and $Re|Z_{IN}|=10\Omega$ over a wide bandwidth (6-16GHz). The selection of 10Ω is to some extent arbitrary; it makes the job of matching to 50Ω easier but increases the insertion loss compared to for example 5Ω (for which Rp=4.6Ω and Cp=2.9pF), Figure 1-21.

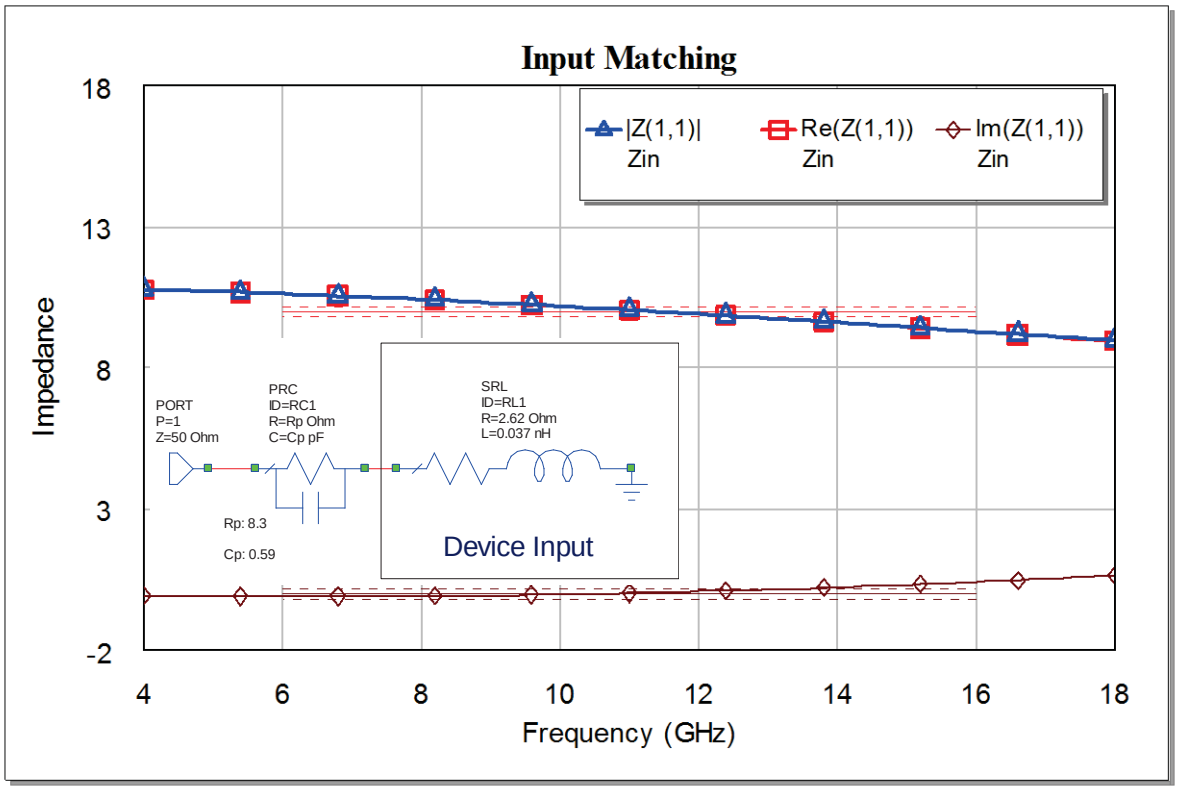


Figure 1-20, Parallel Capacitor and Resistor input matching.

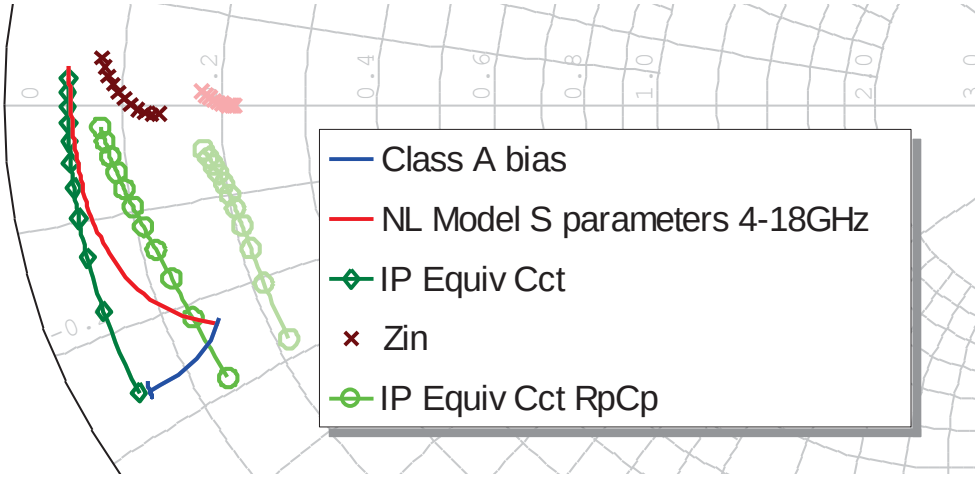


Figure 1-21, Effect of Rp and Cp on input impedance (4.6Ω and 2.9pF - bold) (8.3Ω and 0.59pF - faded) and then including dispersion from C_G, trace IP Equiv Cct RpCp.

The input capacitance also needs to be taken into account, as can be seen from Figure 1-21 it causes considerable dispersion to the match. There are a number of matching approaches that can be taken from this point, some of which have been discussed earlier. A suitable approach may be to use the λ/8 capacitively loaded line described in Figure 1-11. Incorporating the gate capacitance as the capacitor nearest the load the transmission line impedance and output capacitance can be optimised to give the best match over the bandwidth, Figure 1-22. Note that for input matching over a wide bandwidth relatively poor input matches (<-6dB) are the best that can be expected due to the low input impedances

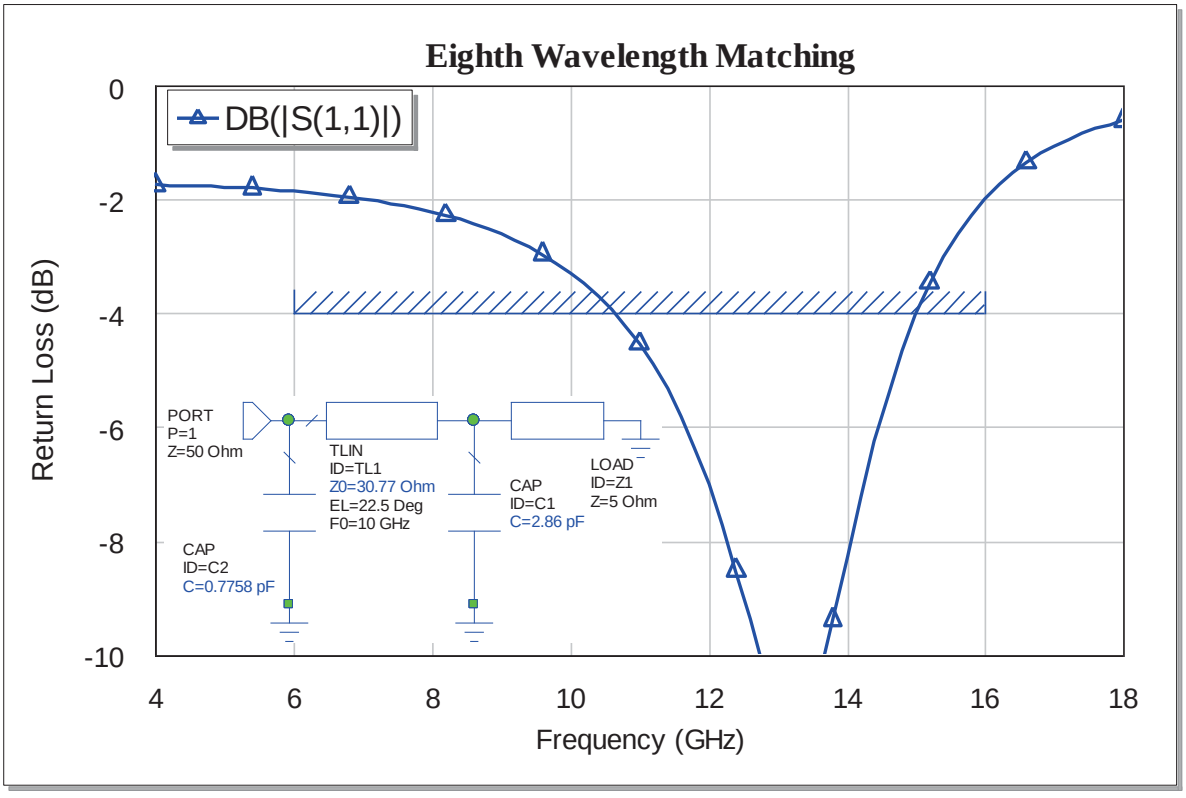


Figure 1-22, Eighth wavelength input matching to 5Ω load.

that are being addressed. To overcome this specific amplifier topologies are used as will be described in the next chapter.

Applying this matching and the parallel RC network to the equivalent circuit of the device input it can be seen, Figure 1-23 that the solution has the additional desirable attribute of a negative gain slope of 6dB/octave which will compensate for that of the device. The insertion loss at the top of the band is a minimum of 2.6dB (the frequency of the peak in the response can be adjusted as necessary, for example to the top frequency), which although low is still a significant amount of power. The insertion loss at low frequencies is predominantly due to mismatch, which may still cause stability issues. It is therefore not unusual for a lossy stub, a short circuit line joined to the main circuit by a resistor, to be added. The line is λ/4 at the top operating frequency which effectively decouples the resistor from the circuit in this region. A similar o/c approach was described earlier, see Figure 1-14.

It should also be remembered that the device impedance varies with drive power. The current design has been matched to the small signal input impedance. Combining the matching circuit with the nonlinear model and sweeping the input power, the effects of drive power on input match can be seen, Figure 1-24. It is clear that despite the 2.5dB of

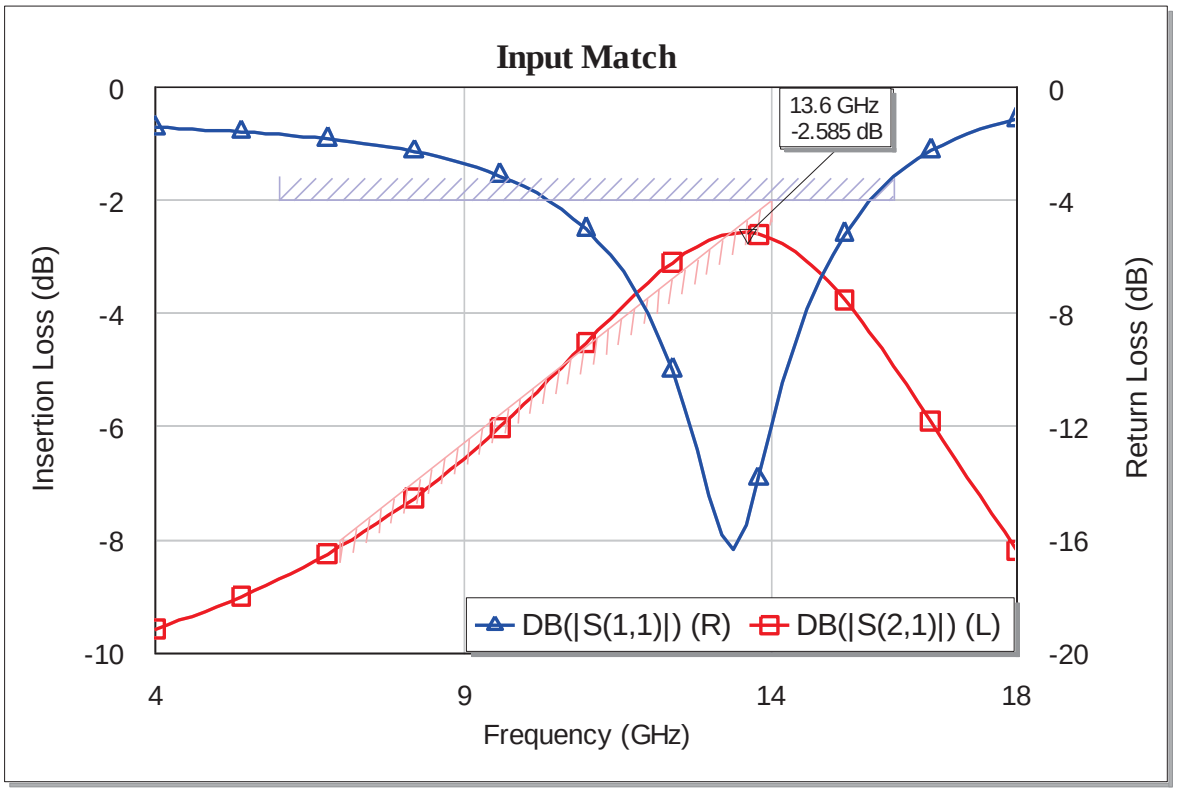


Figure 1-23, Insertion and return loss of input matching terminated with device input equivalent circuit.

insertion loss the input matching gives a 4dB gain improvement to the linear gain. As the match is optimised for the small signal impedance the compression characteristic of the device with input matching is 'soft' i.e. compression is slow as the mismatch loss is getting progressively worse as input power increases, Figure 1-25.

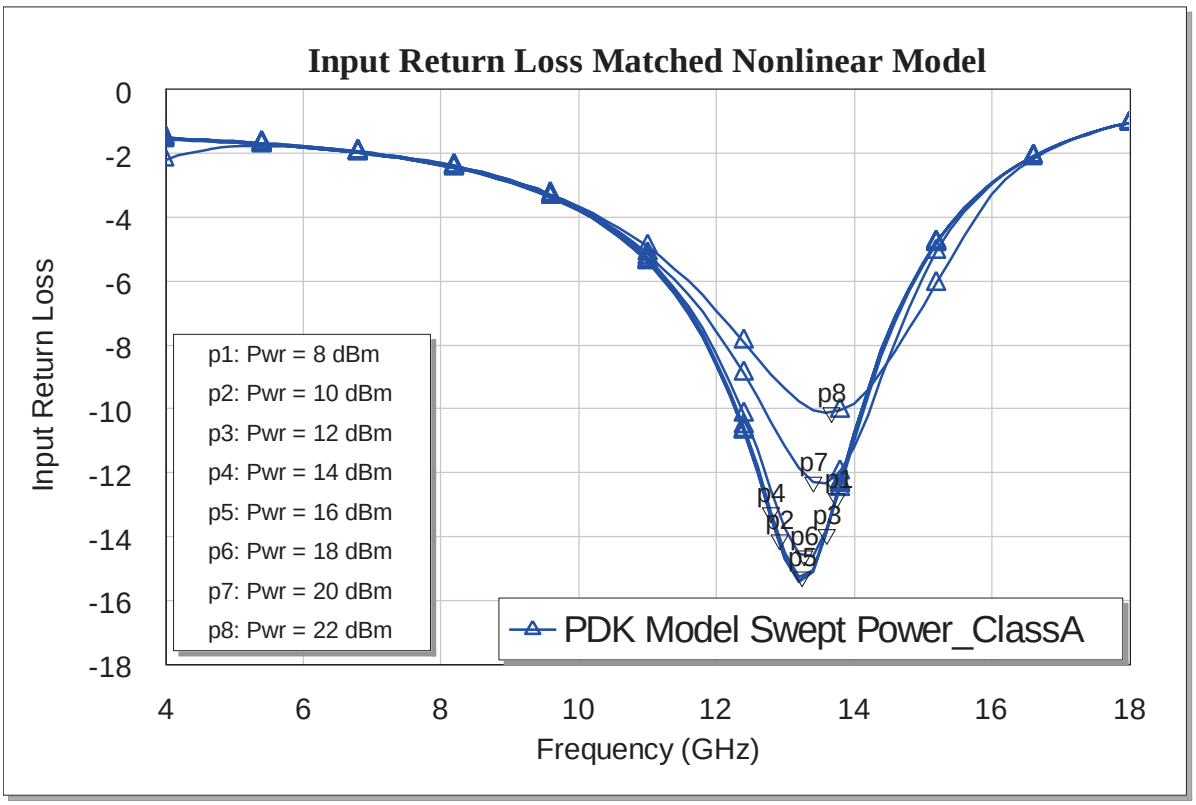


Figure 1-24, Effect of increasing input drive level on return loss.

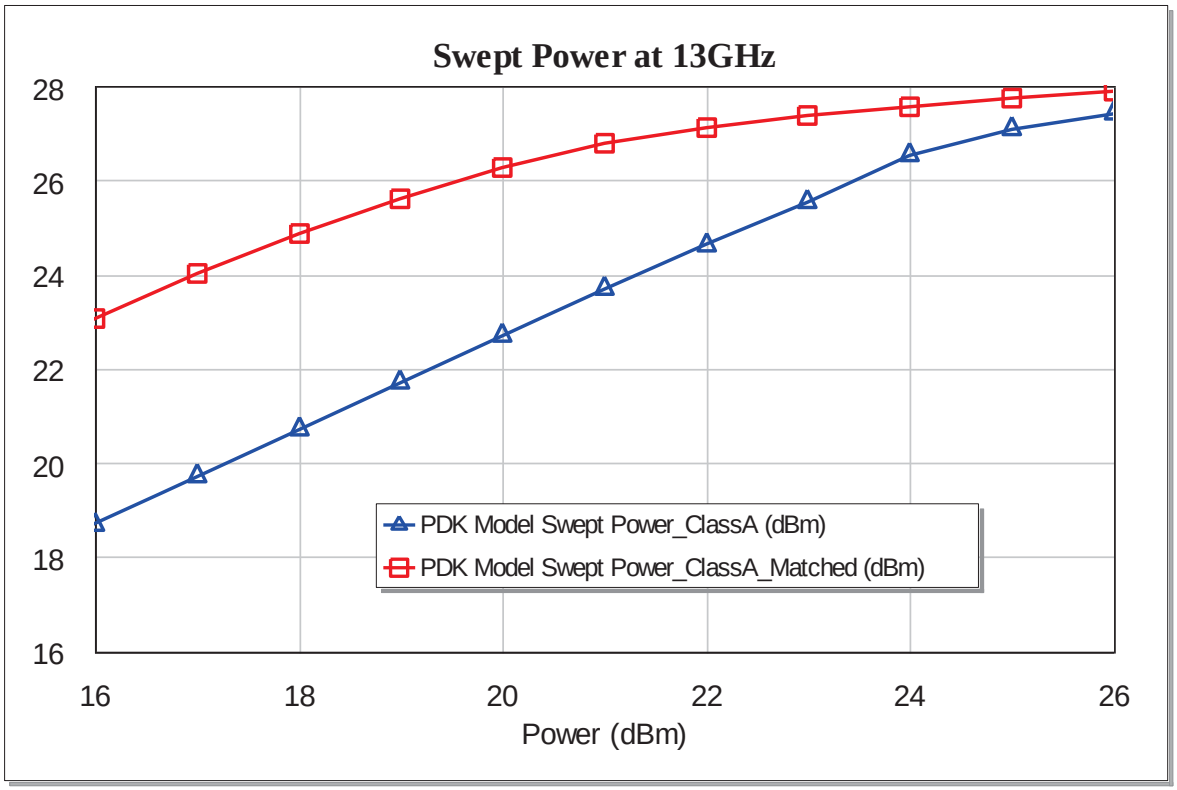


Figure 1-25, Compression curves for device on its own and with input matching.

1.5 Output Matching Impedance

It has been stated earlier that for maximum power transfer the source and load

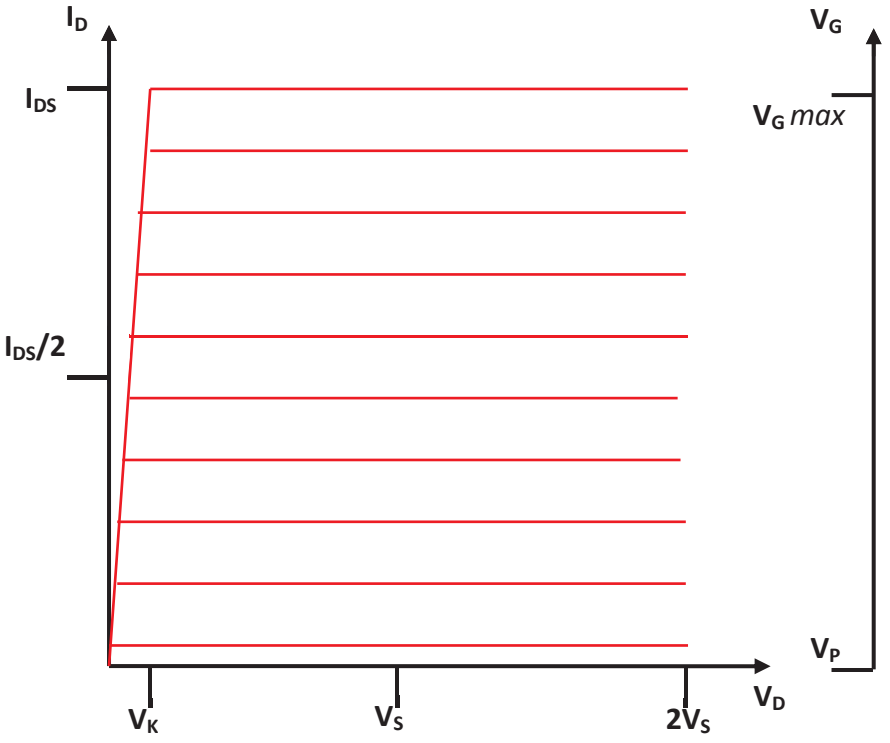


Figure 1-26, Basic DC-IV Curve Model for FET device.

impedances should be equal and that the negative reactances should be cancelled or conjugately matched. In his seminal work [1] Cripps explained a method for determining not only the optimum output power load impedance but also the curves which defined the boundaries at lower powers. The analysis of the approach commonly known as the loadline technique is based on a number of assumptions which although they may appear simplistic can actually be utilised in MMIC design [33]. The first assumption is that the FET device output can be approximated by a Voltage Controlled Current Source (VCCS) and a shunt resistance, R_D. The characteristics of the current source are such that the current I_D is determined by the drive voltage V_G (gate voltage). At pinch-off, V_P, there is no current flowing and at V_G max the maximum current or saturated current I_{DS} flows. Beyond the knee voltage, V_K, the current is assumed to be constant and, at least initially, V_K is assumed to be small enough to be ignored. It is also assumed that V_D is limited to $2V_S$. This description is summarised in Figure 1-26.

To examine how the theoretical device behaves with load resistance, the simple model is connected in a circuit as shown in Figure 1-27. The device is biased at a DC voltage, V_S, through an inductance L, which is considered to be an open circuit at RF and a short to DC. The device is connected to the load R_L via a capacitor C which is assumed to be a short circuit to RF and open to DC. The current from the drain supply, I_S, is a varying DC and so the average of the square wave is $I_{DS}/2$, where I_{DS} is the saturated (peak) drain current. Hence also V_D is a square wave, 180° out of phase with I_D as V_D is a maximum when I_D is zero and a

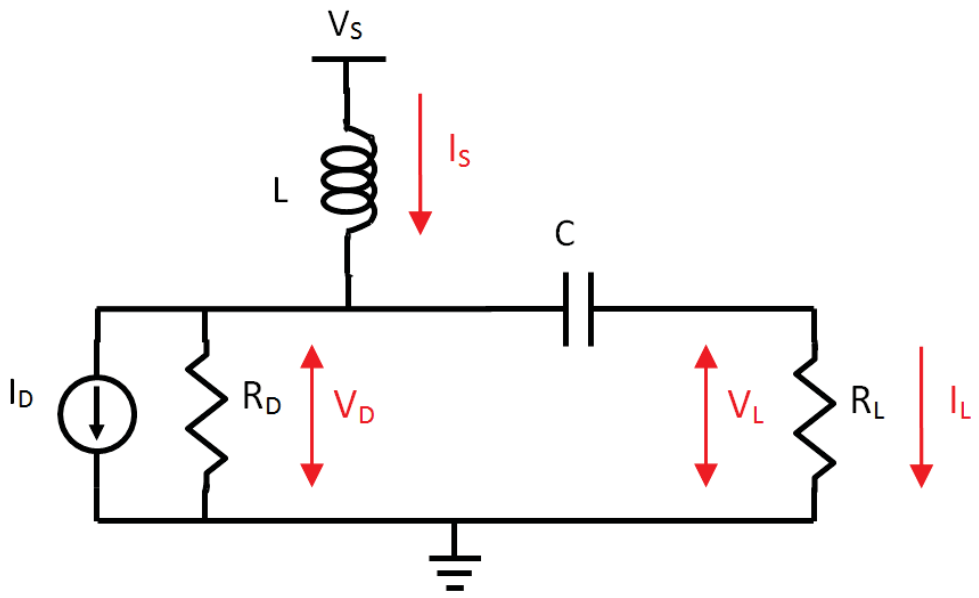


Figure 1-27, Basic device output model and load equivalent circuit.

minimum when $I_D = I_{DS}$. The value of R_D is arrived at from the device data and the operating voltage V_S. Most MMIC device processes are specified in terms of the mA/mm periphery of the device. Thus a device made on a 500 mA/mm process and having a gate periphery of 0.75mm, will have a maximum current, I_{DS}, of 375mA. The value of V_S selected is based on ensuring that the device will remain within the maximum power and voltage envelopes during operation. As the gate voltage is modulated the current will swing from 0 to I_{DS}, and this will produce a voltage swing in V_D depending on the value of R_D. From Ohm's Law we know the ratio of voltage to current is the resistance, thus if we bias the device at a voltage of V_S and set the current for $I_{DS}/2$ (so that we get the maximum current swing assuming symmetrical clipping), then we get a value for R_D, {1-26}.

$$R_D = \frac{2V_S}{I_{DS}} \qquad \textbf{\{1-26\}}$$

By the theory of maximum power transfer the maximum power is delivered to the load R_L when $R_L = R_D$. We therefore call this value of R_L, R_{opt}. From the graph we can see that when $R_L = R_{opt}$ the current swings from 0 to I_{DS} and the voltage from $2V_S$ to 0.

The equation for a straight line is *y= mx + c*, in the case of the graph Figure 1-26, a straight line can be drawn where $y=I_D$ and $x=V_D$, and c is the value of I_D when x (V_S) = 0, i.e. I_{DS}, when $R_L = R_{opt}$.

$$m = -\frac{1}{R_L} = -\frac{I_{DS}}{2V_S} \qquad \textbf{\{1-27\}}$$

This line drawn on the DC-IV curves is what is referred to as the "Loadline". In the case of the load being R_{opt} we can see from Figure 1-28 that both the current and voltage intercept their limits when driven with a hard enough signal. This is referred to as 'clipping'. For the case where $R_L < R_{opt}$ the current limits are still reached, however the voltage limits aren't, hence we have current clipping only. The voltages of the intercepts of the current limits are:

For $R_L < R_{opt}$,

$$V_D = V_s \pm \frac{I_{DS}}{2} R_L \qquad \textbf{\{1-28\}}$$

Now consider the opposite case, where $R_L > R_{opt}$, the slope of the loadline is now less steep and as a result does not reach I_{DS}. The general equation for this line is:

For $R_L > R_{opt}$,

$$I_D = -\frac{1}{R_L}V_D + I_D^* \qquad \{1\text{-}29\}$$

where I^*_D is the value of I_D when $V_D = 0$ (ignoring V_K), at the case where $V_D = V_S$, $I_D = I_{DS}/2$, hence,

$$\frac{I_{DS}}{2} = -\frac{1}{R_L}V_S + I_D^* \qquad \{1\text{-}30\}$$

$$=> I_D^* = \frac{I_{DS}}{2} + \frac{V_S}{R_L} \qquad \{1\text{-}31\}$$

These relationships are summarised in Figure 1-29. In the same way we can also investigate the power and efficiency based on these assumptions. The power can be divided into the DC, P_{DC}, and the RF (load), P_{LT}. In this idealised scenario as the voltage and current waveforms are square waves they contain harmonics. Of primary interest is the power at the fundamental frequency, P_{L1}. The basic measure of efficiency is the ratio of the fundamental load power to the DC power, and is called the drain efficiency.

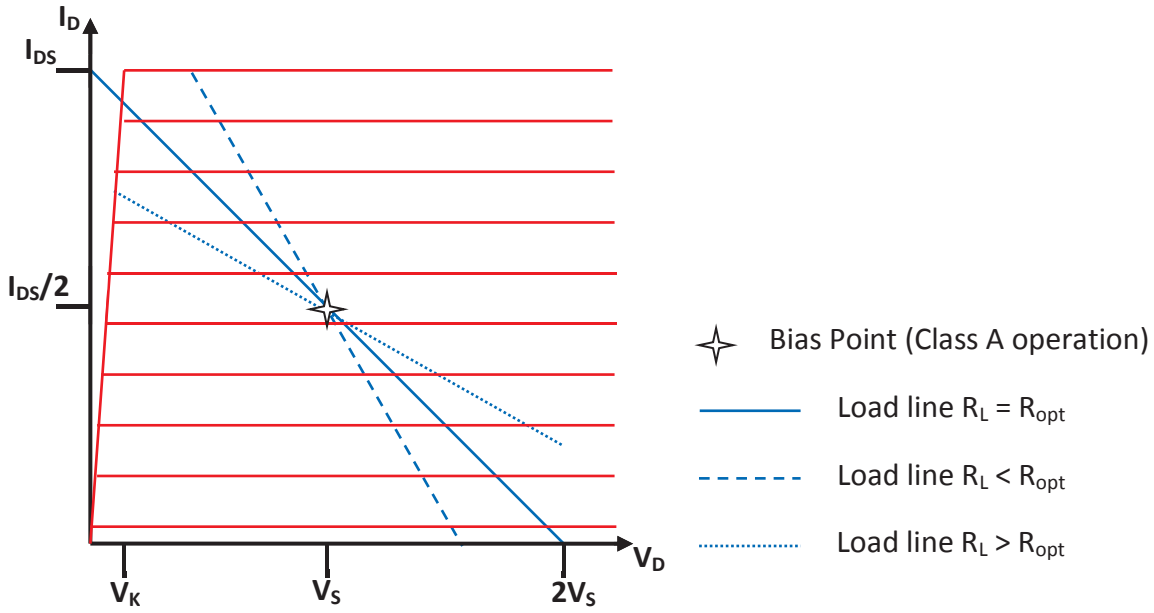


Figure 1-28, Load lines on the basic DC-IV curves.

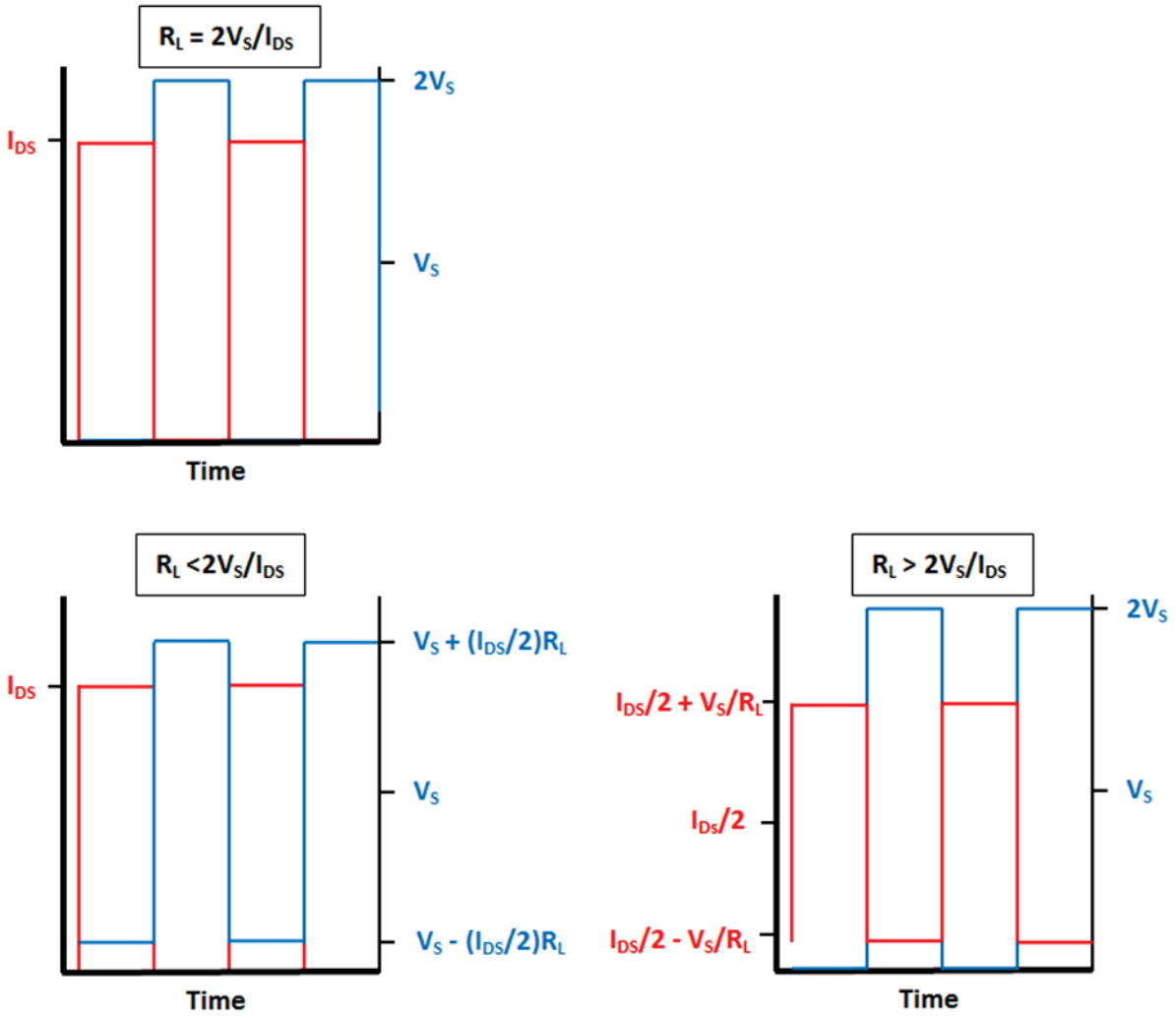


Figure 1-29, Theoretical voltage and current waveforms for RL=R_{OPT}, RL<R_{OPT} and RL>R_{OPT}.

$$P_{DC} = V_S \frac{I_{DS}}{2} \quad \{1\text{-}32\}$$

$$P_{LT} = \left(\frac{I_D}{2}\right)^2 R_L = \frac{{I_{DS}}^2}{4} \frac{2V_S}{I_{DS}} = \frac{V_S I_{DS}}{2} \quad \{1\text{-}33\}$$

$$P_{L1} = \frac{2}{\pi^2} {I_D}^2 . R_L \quad \{1\text{-}34\}$$

and when $R_L = R_{opt}$

$$P_{L1} = \frac{2}{\pi^2} {I_{DS}}^2 . R_{opt} = \frac{2}{\pi^2} . {I_{Ds}}^2 . \frac{2V_S}{I_{DS}} = \frac{4}{\pi^2} . I_{DS} V_S \quad \{1\text{-}35\}$$

Note from {1-32} , {1-34} *and* {1-26}*, when* $R_L = R_{opt}$ *the ratio of fundamental to total power,*

$$\frac{P_{L1}}{P_T} = \frac{4/\pi^2}{1/2} = 81.1\% \quad \textbf{\{1-36\}}$$

and the Drain efficiency, η_D, is

$$\eta_D = \frac{P_{L1}}{P_{DC}} = \frac{2}{\pi^2}.{I_D}^2.R_L \times \frac{2}{V_S I_{DS}} = \frac{4}{\pi^2}\frac{{I_D}^2}{I_{DS}V_S}R_L \quad \textbf{\{1-37\}}$$

and when $R_L = R_{opt}$,

$$\eta_D = \frac{4}{\pi^2}\frac{I_{DS}}{V_s}R_{opt} = \frac{8}{\pi^2} = 81.1\% \quad \textbf{\{1-38\}}$$

So far it has been assumed that the knee voltage, V_K, is negligible. If we now modify the equations for power and drain efficiency to take account of this at the peak power we get:

$$P_{L1} = \frac{4}{\pi^2}.I_{DS}(V_S - V_K) \quad \textbf{\{1-39\}}$$

$$\eta_D = \frac{4}{\pi^2}\frac{I_{DS}}{V_s}R_{opt} = \frac{8}{\pi^2}\frac{(V_s - V_K)}{V_K} \quad \textbf{\{1-40\}}$$

A number of assumptions have been made in this analysis:

- The FET is assumed to be an ideal voltage controlled current source that can sink a maximum drain current of I_{DS}, and that the drain current falls to zero when pinched off.
- The FET is modulated with a large enough input signal on the gate to produce a square wave drain current, and that the output is truly square.
- That the knee voltage V_K is constant (in the later calculation, earlier it was assumed to be negligible).
- That the drain current I_D is constant with drain voltage above V_K.
- The impact of reactance in the device.
- That the device doesn't self-bias as it is over driven.

Before examining in detail the implications of these assumptions it is useful to understand the effects of two key elements of the output impedance of a microwave transistor, the drain resistance and capacitance. Firstly looking at the drain resistance, R_D, we have already spoken of the maximum power transfer theory, that for optimum power transfer we need the load and source resistance to be equal. In practice this will not always

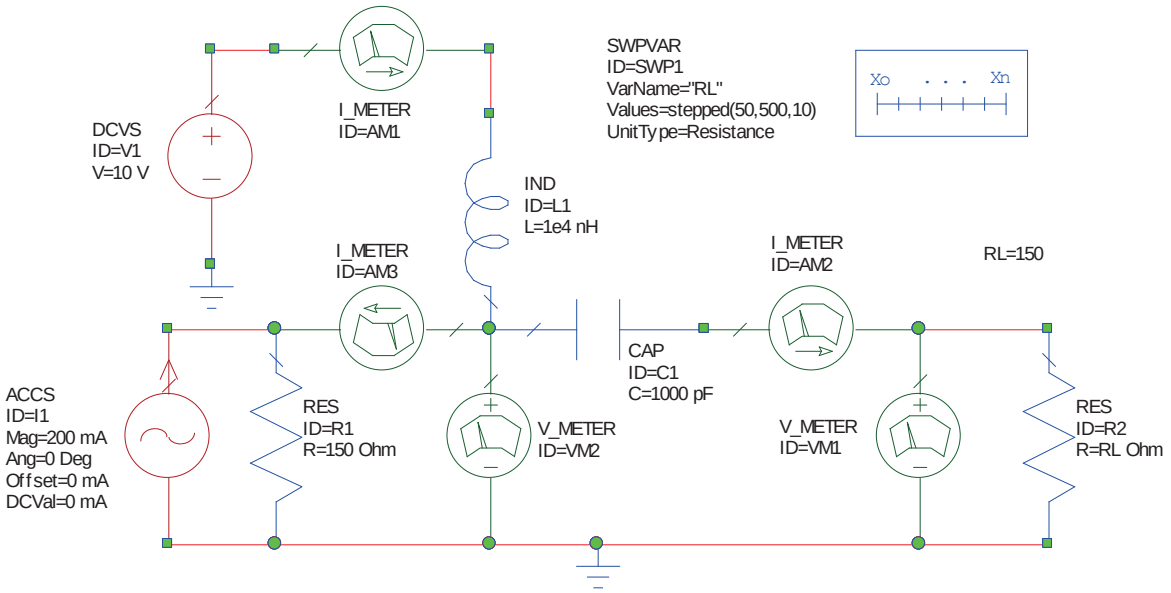


Figure 1-30, Equivalent circuit to examine power transfer to load as a function of load resistance.

be possible due device to device variations in both the source and load and also as will be seen later R_D is dependent on a number of factors which will not always remain constant. The relationship between source and load resistance and power transfer can be explored mathematically or (and as seems more appropriate in the context of this research) by creating a simple model of the circuit, Figure 1-30. The element ACCS is an AC Current Source, set to deliver a 1 GHz sine wave of peak current 200mA, thus when it is at a peak no current will flow through the ammeter AM3, and conversely when it is a minimum 150mA will flow through AM3, thus a voltage is produced across the 150Ω resistor (R_D), dependent upon the load applied R_L. In this circuit the load is swept from 50 to 500Ω. Inductor L1 and capacitor C1 ensure that the DC components are kept away from R_L and the AC is kept from the supply. In this case we will keep the source resistance fixed, whilst the varying the load and calculate the power delivered to the load by measuring the output voltage and current as shown in Figure 1-31. As expected the maximum power transferred is at 150Ω, equal to R_D. Of interest is that 95% power transfer is achieved provided the load resistance is within 95.6 to 236Ω; i.e. a range of -36% to +57%. So a consideration of what actual load value we should select would depend upon the statistical variation of R_D, and R_L, all other things being equal and assuming a normal distribution of R_L, would indicate aiming for a slightly higher nominal value of R_L.

If a shunt capacitor C_D, representing the drain capacitance is now added in parallel with R_D the effect on the power transfer can be seen (1pF was used in this instance) and as would be expected the optimum load resistance has reduced, Figure 1-32; as has the total power available to the load as the reactive element has not been cancelled. Examining the drain current and voltage waveforms, Figure 1-33, it is clear why there has been a reduction in power, the drain voltage and current are no longer in anti-phase as they were in the purely resistive circuit. Plotting the drain voltage against the current we get the loadlines for the circuits, Figure 1-34. For the purely resistive circuit the loadline is straight, whilst introducing the reactance to the source produces the ellipse (due to the phasing of the voltage and current). Adding a 25nH inductor in shunt with the load, Figure 1-35 restores the phasing of the drain waveforms and the peak power and optimum load resistance, Figure 1-36, i.e. it conjugately matches the load.

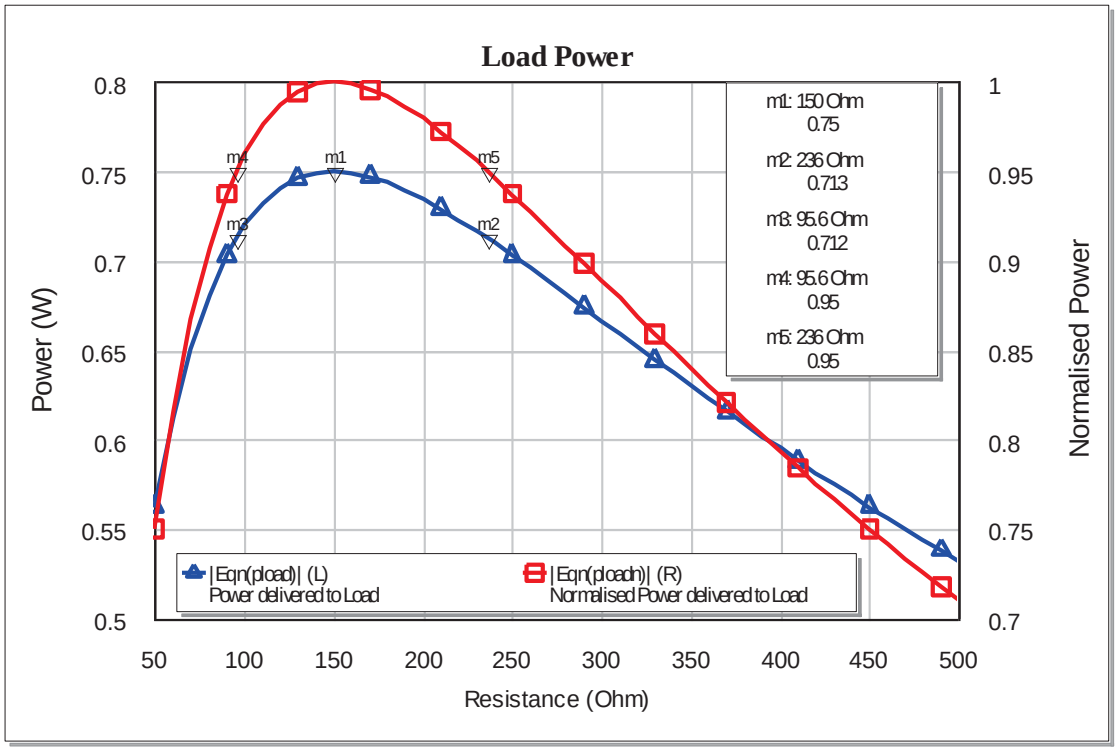


Figure 1-31, Power delivered to the load as a function of load resistance.

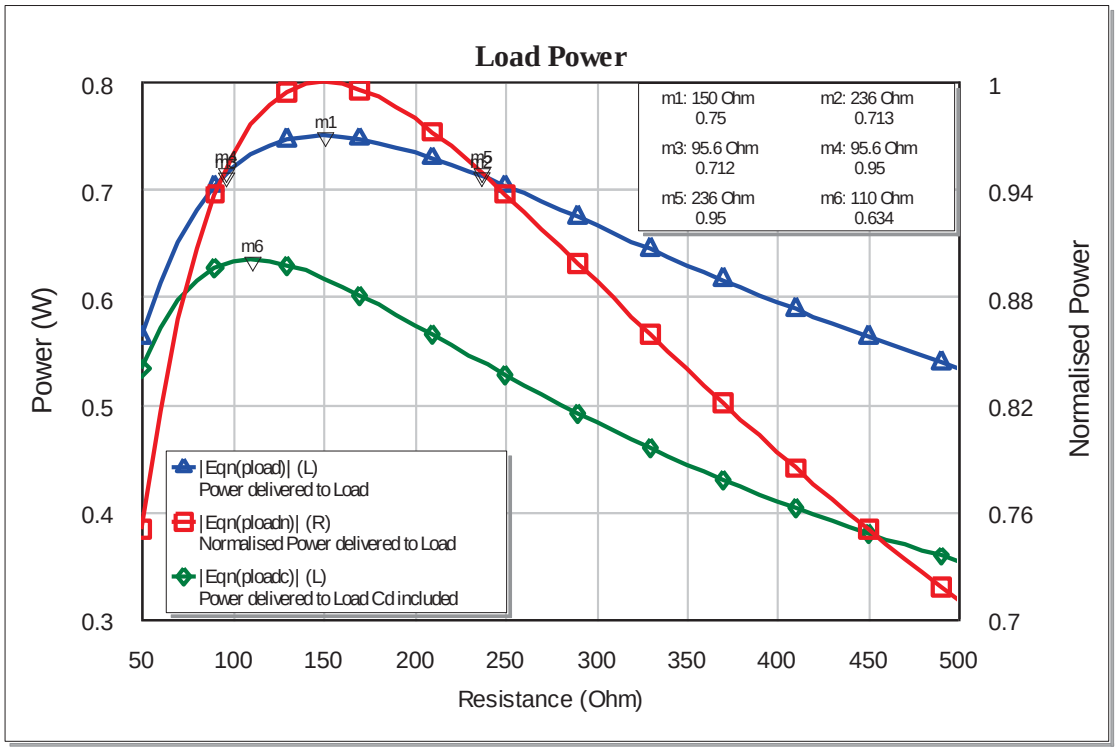


Figure 1-32, Impact of including drain capacitance (1pF) in the source.

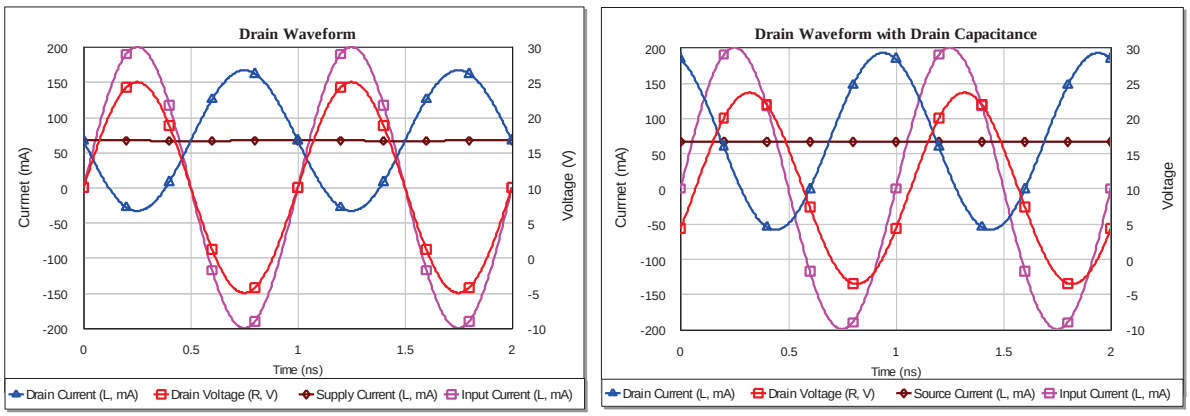


Figure 1-33, Current and voltage waveforms, purely resistive source (left) and including capacitance (right).

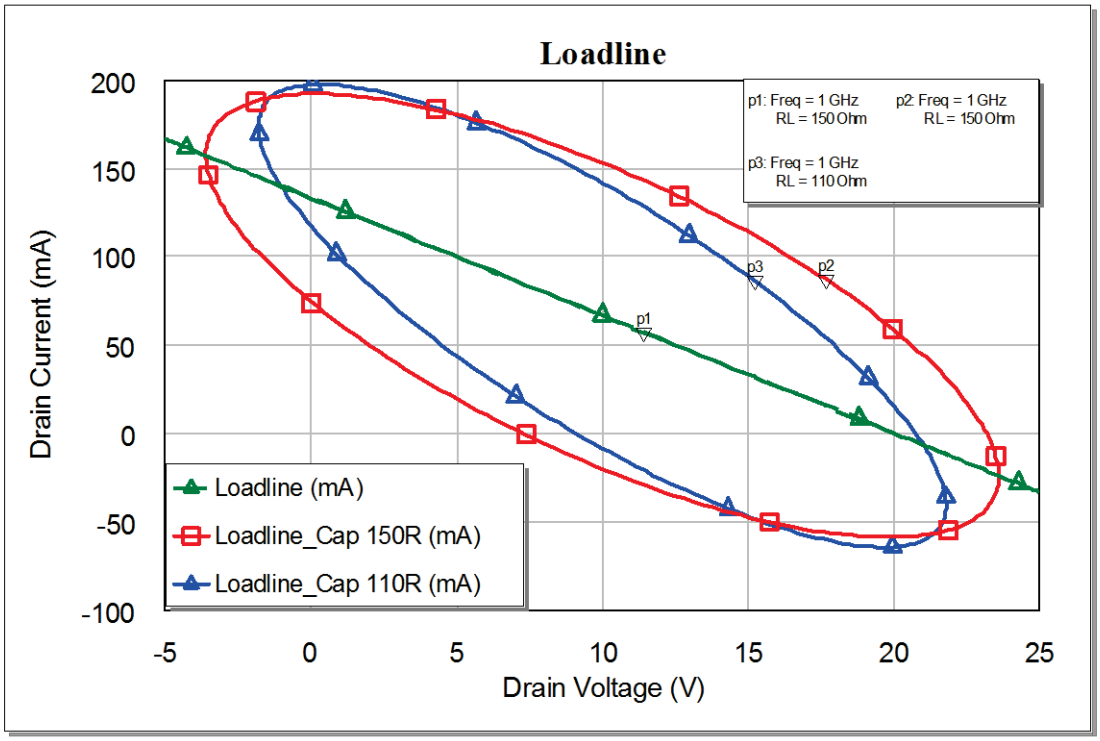


Figure 1-34, Load lines for resistive (green) and capacitive sources with fixed resistive loads.

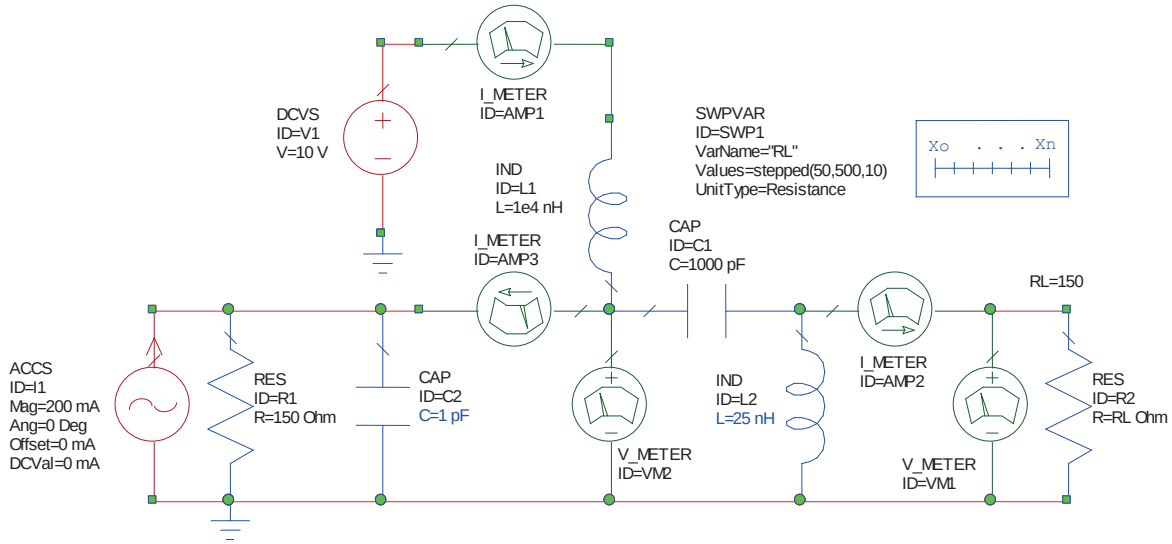


Figure 1-35, Basic output circuit model including drain capacitance and load conjugate matching inductor.

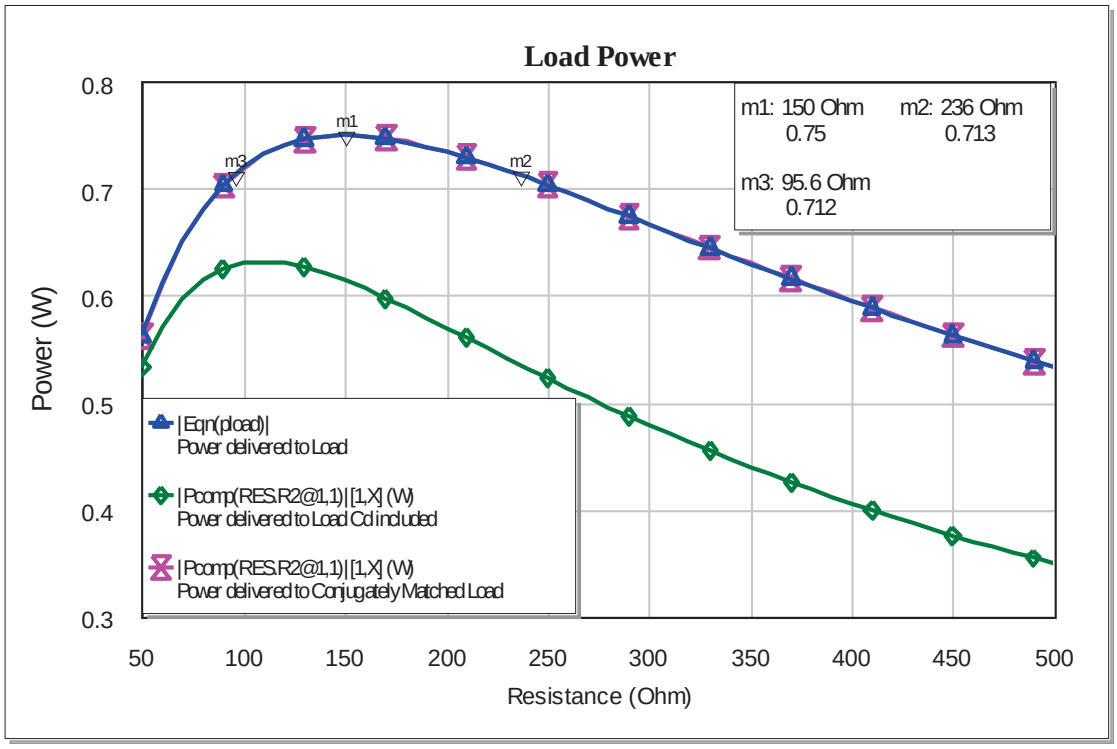


Figure 1-36, Power transferred to load as a function of load resistance for, resistive source (blue), capacitive source (green) and conjugately matched load (pink).

With the improvements to nonlinear CAD models and the speed of simulation hardware the loadline technique can be used, but with improved accuracy and without making so many assumptions, even so they are still not completely accurate due to model issues as will be discussed further in chapter 5 on device modelling. They can be used to give an insight into device operation within the boundaries of more realistic DC-IV curves, but there are differences between actual and simulated and these will have a direct impact on the accuracy of the implementation of any design. Using the nonlinear model for the RFMD (UK) Ltd. foundry FD30 process and the 10x100 (10 gate fingers 100µm wide) device, a simulation was performed and compared with the measured data provided in the Process Design Kit (PDK) information, Figure 1-37. There are several points to note about this process DC-IV graph:

- The plot is scaled, the current axis is in mA/mm, hence it must be altered to match the actual device periphery, (for simplicity of comparison a 1mm gate periphery has been chosen for the simulated device).
- The curves do not continue into the into the top right hand corner due to power limitations of ~1.2W/mm of the device.

- The curves do not show the drain source voltage breakdown, which is believed to be in the region of 22V.

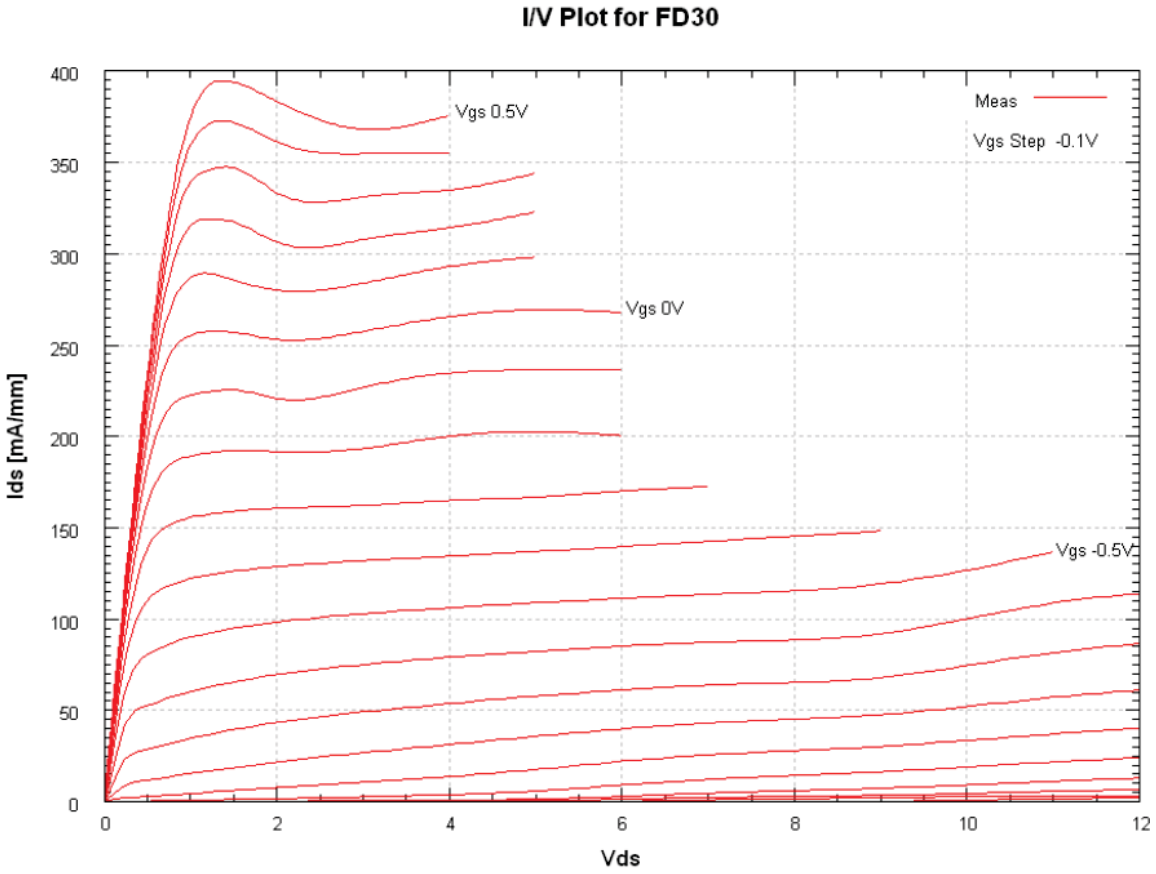


Figure 1-37, Foundry DC-IV curves for FD30 DpHEMT Process - scaled Ids.

Figure 1-38, shows the DC-IV curves from the nonlinear simulation of the foundry model. The graph shows the value of Ids, the maximum drain current which occurs at the upper limit of V_G, +0.5V and at a drain voltage of 1.8V. This is the knee voltage at this particular gate bias; as the gate bias is made more negative I_D and V_K are reduced. Also note that for a particular gate voltage I_D is not flat with V_D, this is another source of nonlinearity. Also shown is the class A bias point, V_S = 10V, I_D = 183mA. The simulation shows the drain breakdown voltage to be 28V, although information from the manufacturer suggested they thought it would be lower (~24V). This may have been a safe operating limit for the user as the model does not include a temperature parameter.

Comparing the foundry DC-IV data and the nonlinear model simulation, Figure 1-38, the points of difference to note are,

1. The difference in I_{DS} max (at Vgs = +0.5 volts); 366mA as compared to 395mA for the measured data, (Δ = 9.3%).

2. The knee voltage on the measured is ~1.4V compared to 1.8V on the simulated.
3. The measured data shows positive slopes to the DC IV curves beyond the knee, (ignoring the 'humps' that occur for Vgs > -0.2V), whilst the simulated shows negative slopes above a drain current of ~180mA. However the simulator goes into an area that the measured can't due to the power limitations, and it is quite possible that if one could go into this region (using pulsed measurements)[3] the drain current would fall due to self-heating.

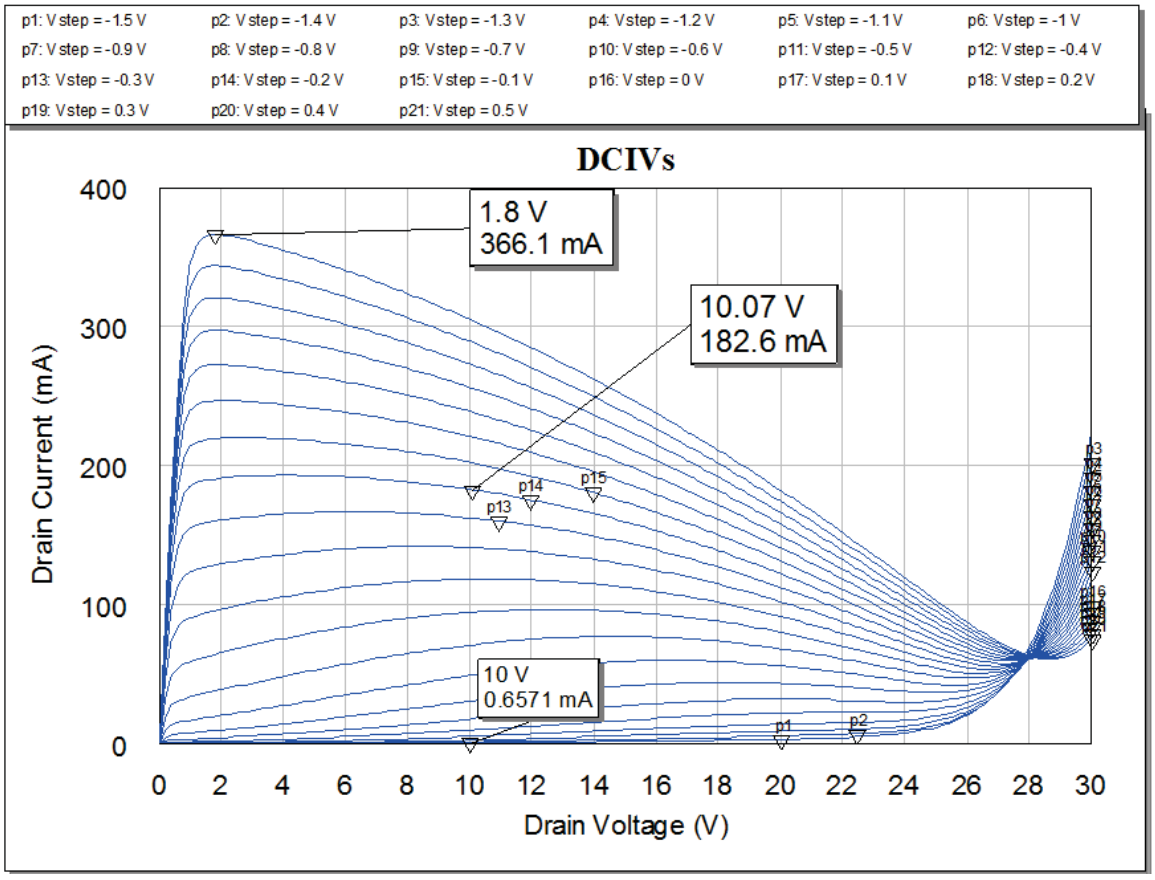


Figure 1-38, Simulated DC-IV curves from foundry nonlinear model of 10x100 device.

Following the same approach that was used for the simple model, a supply voltage V_S, of 10V is chosen. To operate at $I_{DS}max/2$ a Vgs of -0.2V is applied. The R_D is calculated, (2x10)/0.366 = 54.6Ω. Measuring the output power and the PAE with a swept input power at 3 GHz and a purely resistive load of 55Ω, it can be seen, Figure 1-40, that the PAE peaks at an input level of 19dBm, whereas the output power continues to rise with drive level. It can also be seen that above ~10 dBm input power the output begins to become nonlinear. The simulation schematic is shown in Figure 1-39.

[3] Actual DC-IV measurements on these devices is described further in Chapter 4.

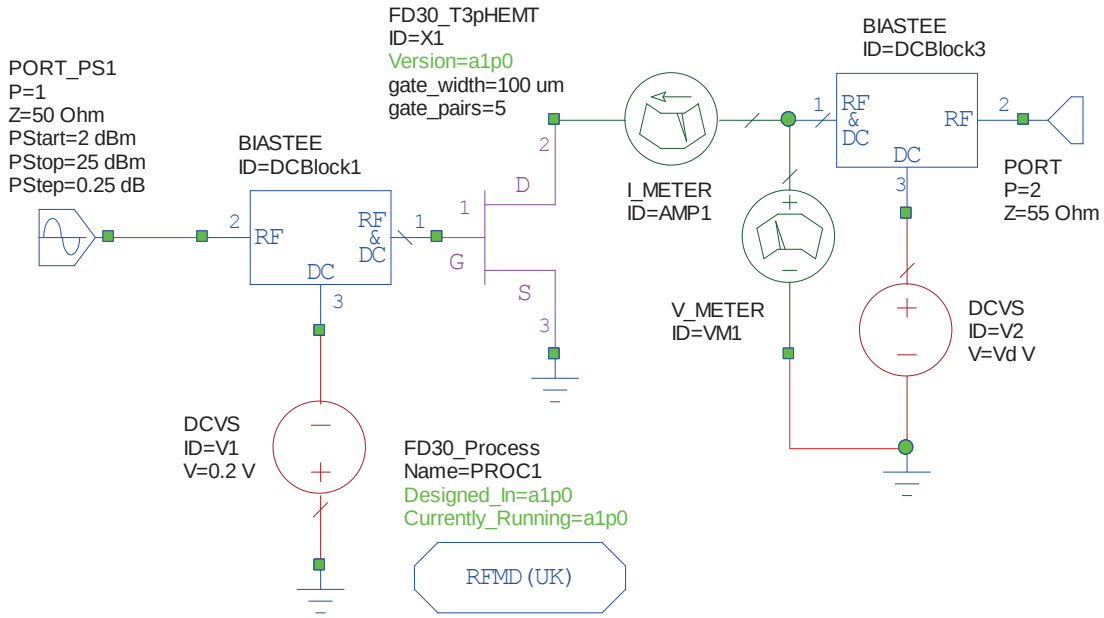


Figure 1-39, Test circuit for nonlinear model of 10x100 device.

If the load line is observed at the power levels of 10, 19 and 21dBm we can see why the observed compression occurs, Figure 1-41. At +10dBm in the left hand peak of the RF loadline, just touches the 'knee' of the DC-IV curves whilst the other end of the load line is still in the linear region of the device transconductance as can be seen from Figure 1-42; the

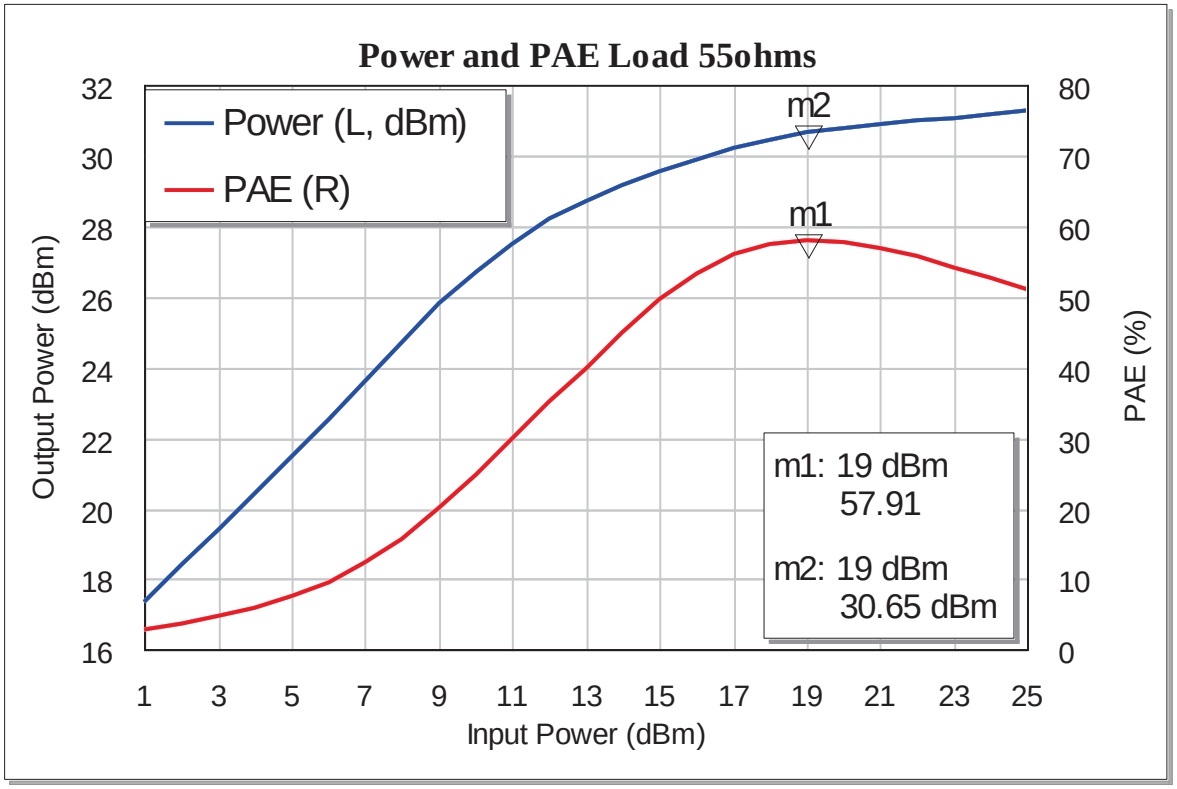


Figure 1-40, Simulation of output power and PAE as 10x100 device, biased in class A and tested at 3GHz is driven into compression.

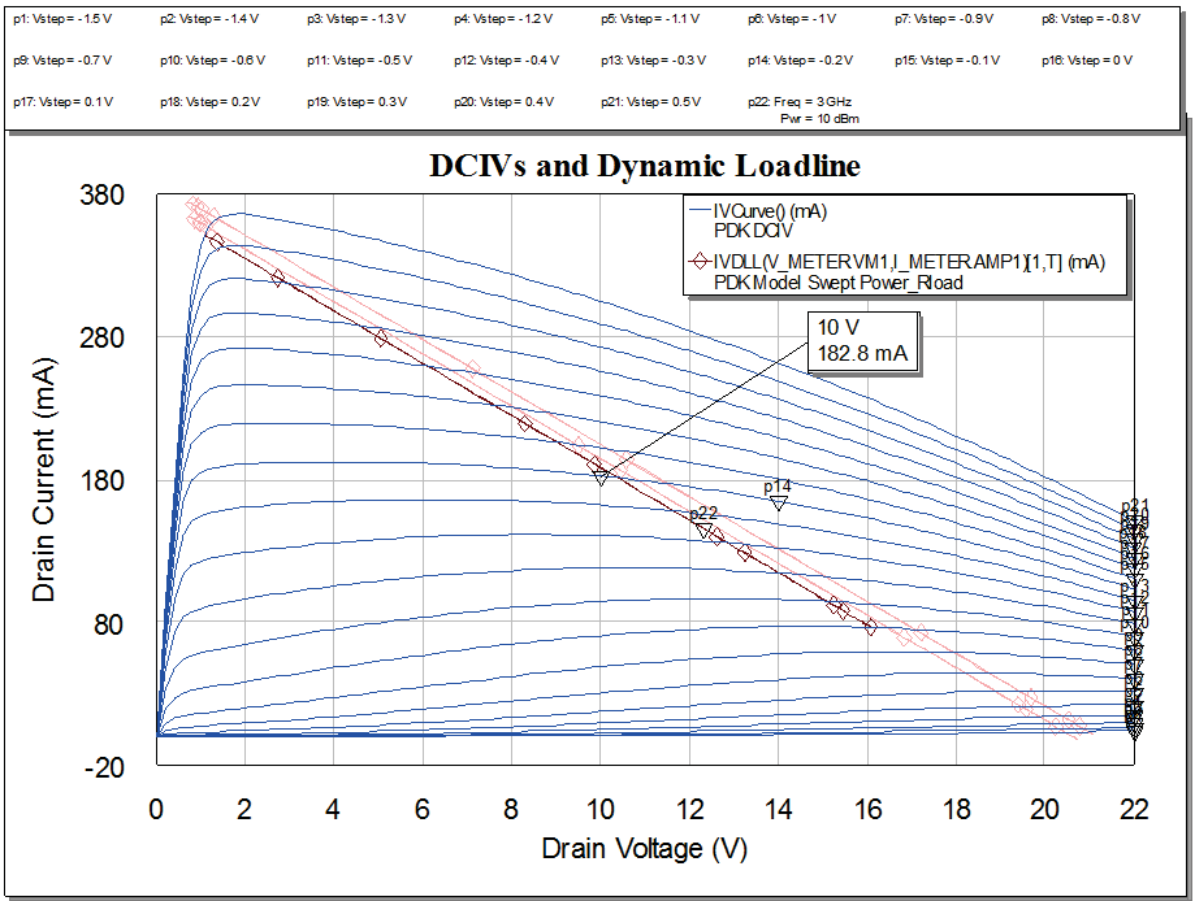


Figure 1-41, Effect on RF loadline of increasing input power, 10x100 device biased in class A at 3GHz; +10dBm in bold, 19 and 21dBm faded.

RF loadline intercepts the -0.7V gate voltage trace on the DC-IV curves and this point in Figure 1-42 is above the area where the drain current is no longer linear with gate voltage. As the input power increases the loadline (faded traces in Figure 1-41) is on the left clipped more by the knee and on the right moves into the nonlinear region of the transconductance curve. Also note how the loadline moves away from the quiescent bias point. Viewing the actual drain current and voltage waveforms, Figure 1-43, the current clipping at low drain voltages as the RF loadline intercepts the knee and the limiting of the minimum drain voltage to the knee voltage. The 'squaring' up of the current waveform suggests an increase in harmonics which can actually be observed, Figure 1-44, with the odd harmonics increasing proportionately more as would be expected.

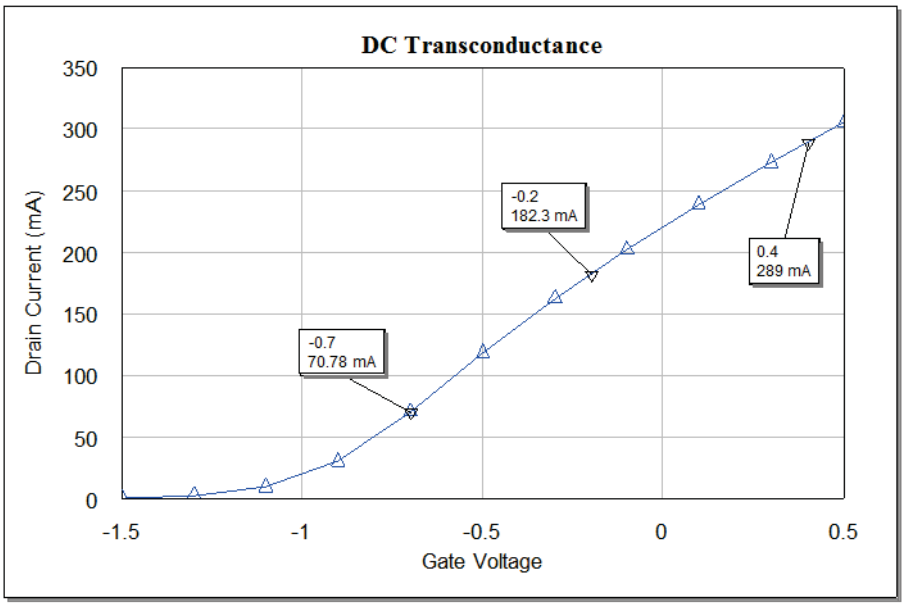


Figure 1-42, Transconductance curve of 10x100 device biased at 10V on the drain.

So far the optimum load impedance has been estimated from a very simple formula, {1-26} and one that ignores the knee voltage. Instead we could find the optimum experimentally by sweeping the load impedance at 19dBm in (the peak PAE level) and observe the how output power and PAE varied by load impedance. Figure 1-45 shows that there are different optimums depending on whether the aim is to maximise PAE or output power; note that this is maximising output power at 19dBm in, i.e. gain; power can still be increased by driving harder whilst PAE peaks, Figure 1-40. The optimum load resistance is close to that of the fundamental theory for PAE at 50Ω, whilst for output power it is markedly lower. However, as was seen in Figure 1-36, reactance acts to lower the source impedance; the device model will include drain capacitance and so to get a true picture of the optimum resistive load we need to de-embed the drain capacitance, move to the Current Generator Plane (CGP). A method to do this is to conduct a load-pull and see how far offset the optimum impedance is from the resistive line.

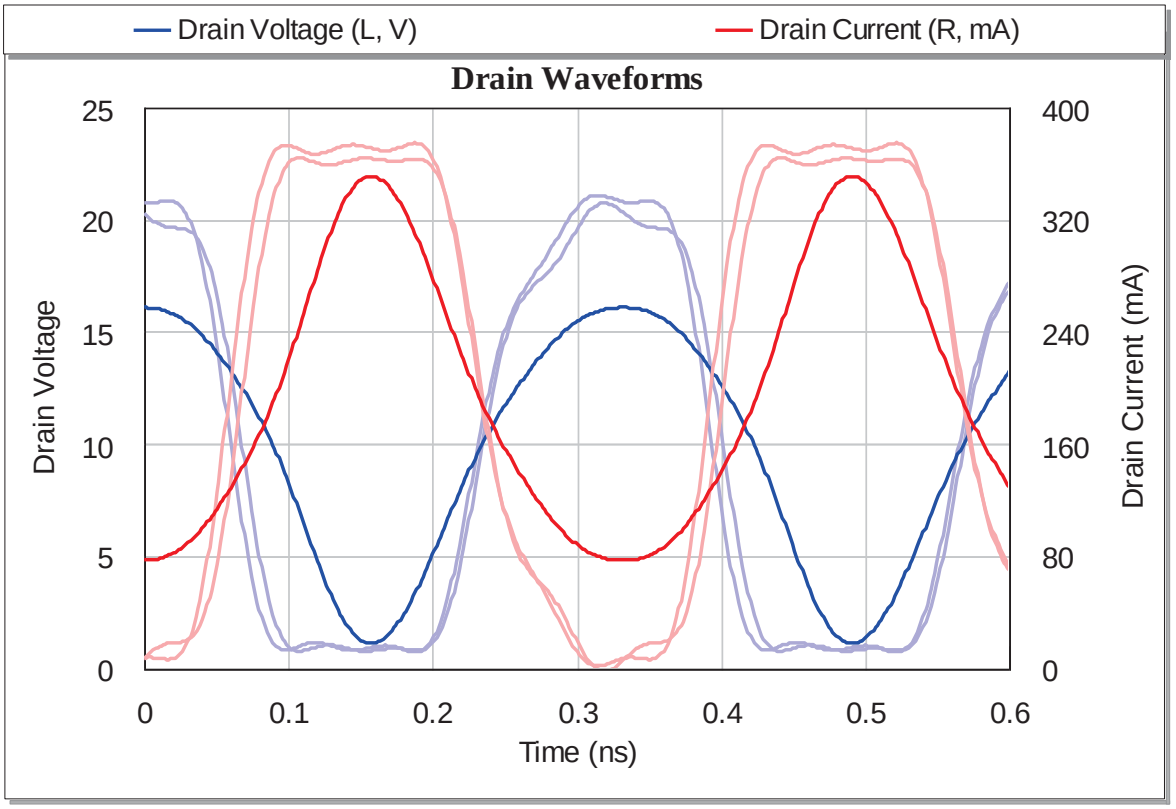


Figure 1-43, Drain current and voltage waveforms at a drain bias of 10V and gate of -0.2V with input power levels of 10 (bold), 19 and 21dBm.

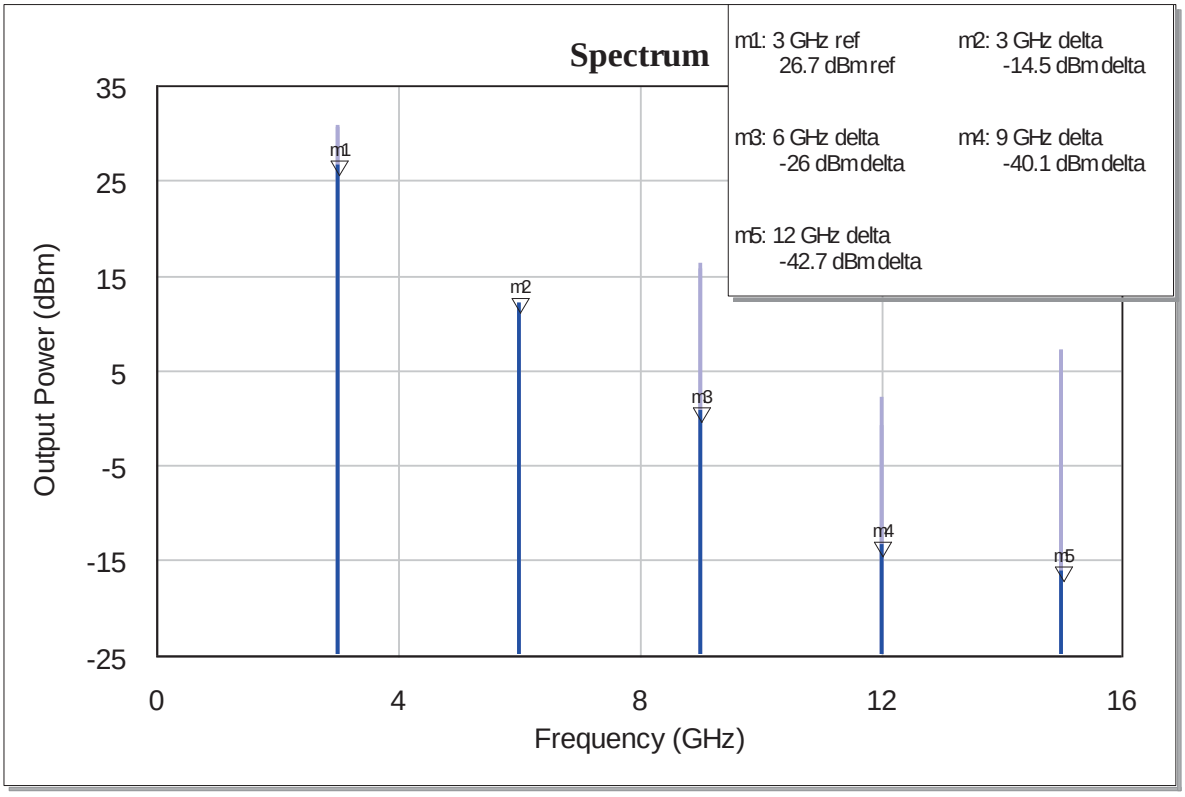


Figure 1-44, Harmonic spectrum at a drain bias of 10V and gate of -0.2V with input power levels of 10 (bold), and 21dBm.

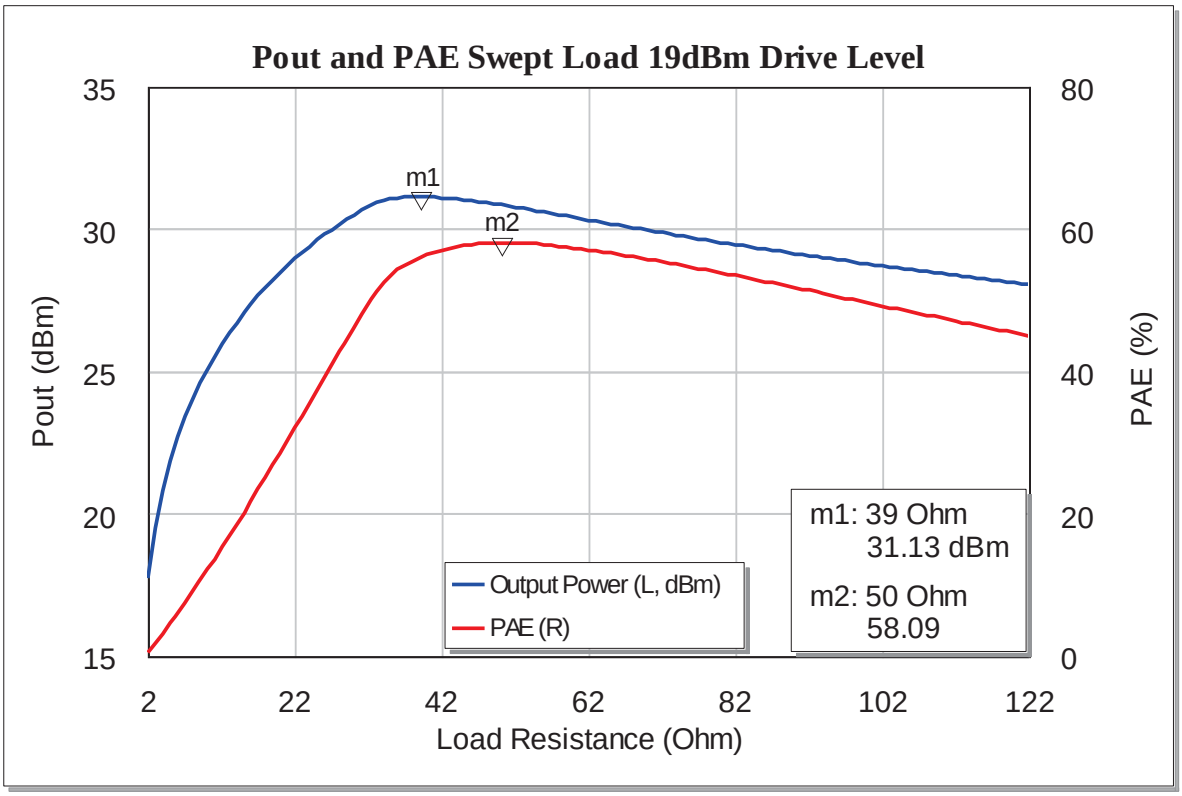


Figure 1-45, Output power and PAE as a function of load impedance at 19dBm in, class A, 10x100 device.

The load pull was conducted using the circuit of Figure 1-46. Note that the capacitance was adjusted for the maximum output power to lie on the resistive axis. The experiment was repeated with a de-embedding capacitance of 0.3pF which made the optimum PAE load be on the resistive axis. The results are summarised in Table 1-3 and the load pull contours with 0.2pF de-embedding is in Figure 1-47 and the swept load resistance in Figure 1-48.

De-embedding Capacitance (pF)	Optimum Output Power Load and Power	Optimum PAE Load and PAE
0.2	39Ω - 31.3dBm	53Ω - 60.2%
0.3	39Ω - 31.3dBm	56Ω - 60.6%

Table 1-3, Optimum load performance at the CGP.

The theory used to calculate the optimum load gives a result very close to the optimum load for maximum PAE; 54.6Ω was the calculated value and 56Ω was the simulated value with 0.3pF de-embedding capacitor. However the optimum output power

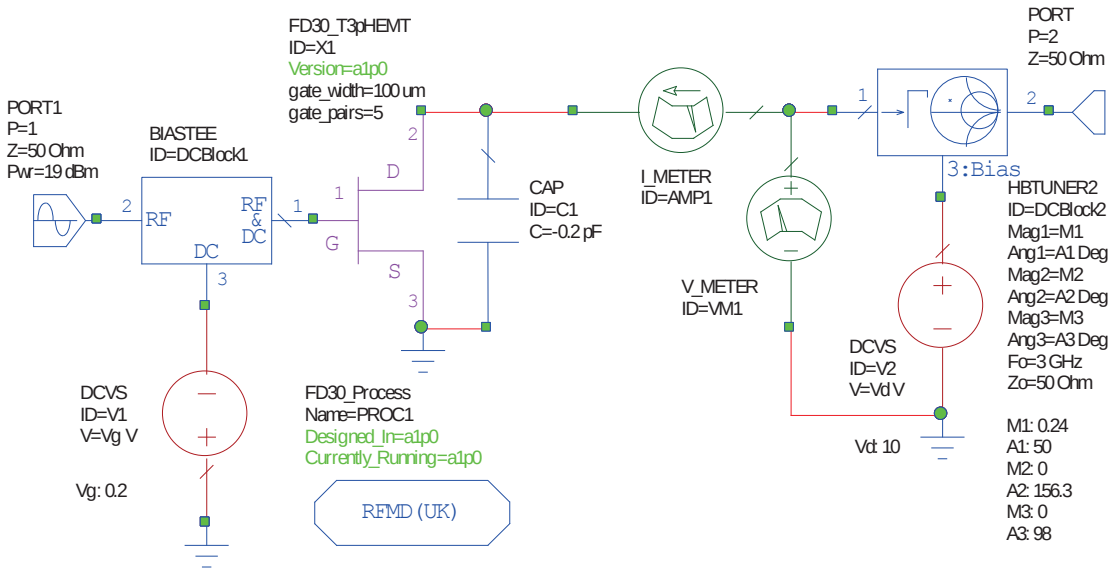


Figure 1-46, Load pull circuit with de-embedded output capacitance.

load is lower than expected at 39Ω. This is somewhat surprising as the original theory [1] was based upon maximum output power. To understand why the loads are as measured we return to the dynamic loadline Figure 1-49 and the drain waveforms, Figure 1-50. With a load of 56Ω the average current flowing is lower, closer to the quiescent bias point and peak of the current waveform flatter whilst simultaneously the trough of the voltage waveform is also flatter and closer to the minimum voltage, the corresponding low voltage when the current peaks and vice versa when the voltage peaks, explains the increased efficiency. In the optimum power case a significantly higher peak current is achieved. One of the reasons why the PAE peaks can be observed; as the input power increases the loadline shifts up and to the right, the average level of drain current increases. Another reason is that the input power level is increasing linearly whilst the output is in compression and hence is not increasing at the same rate, hence the difference between them is decreasing and so the reduction in PAE, {1-1}.

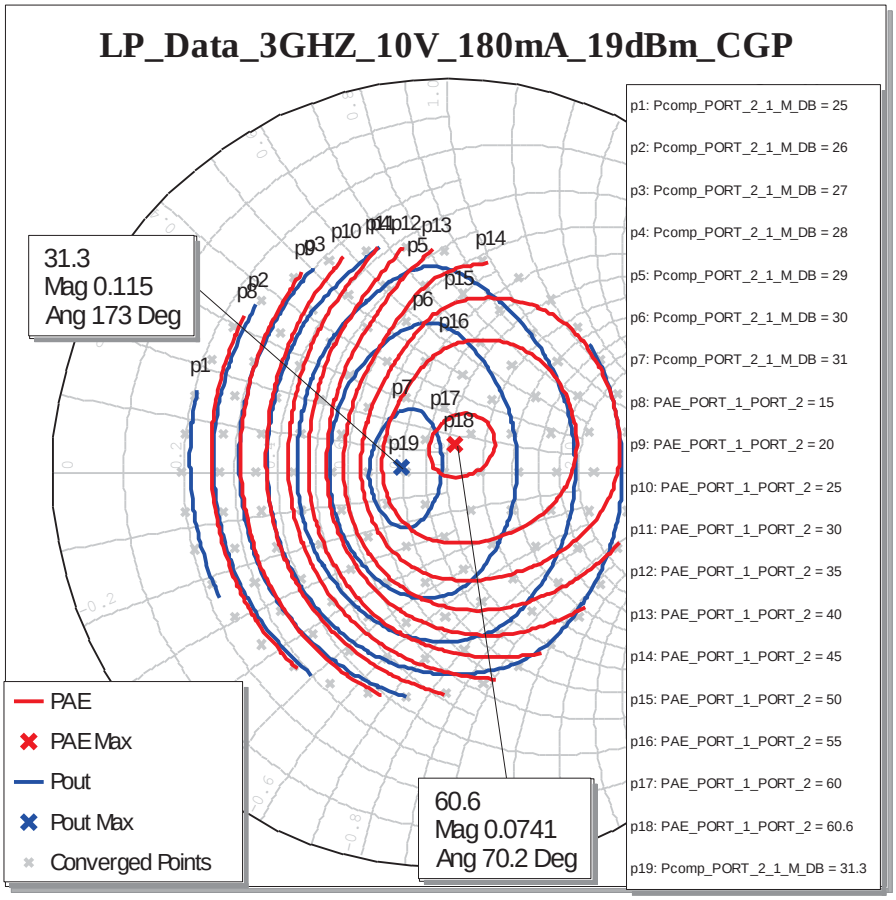


Figure 1-47, Load pull de-embedded to the CGP by adding -0.2pF shunt capacitor. Optimum output power load near to resistance axis however optimum PAE load still in inductive portion of Smith Chart.

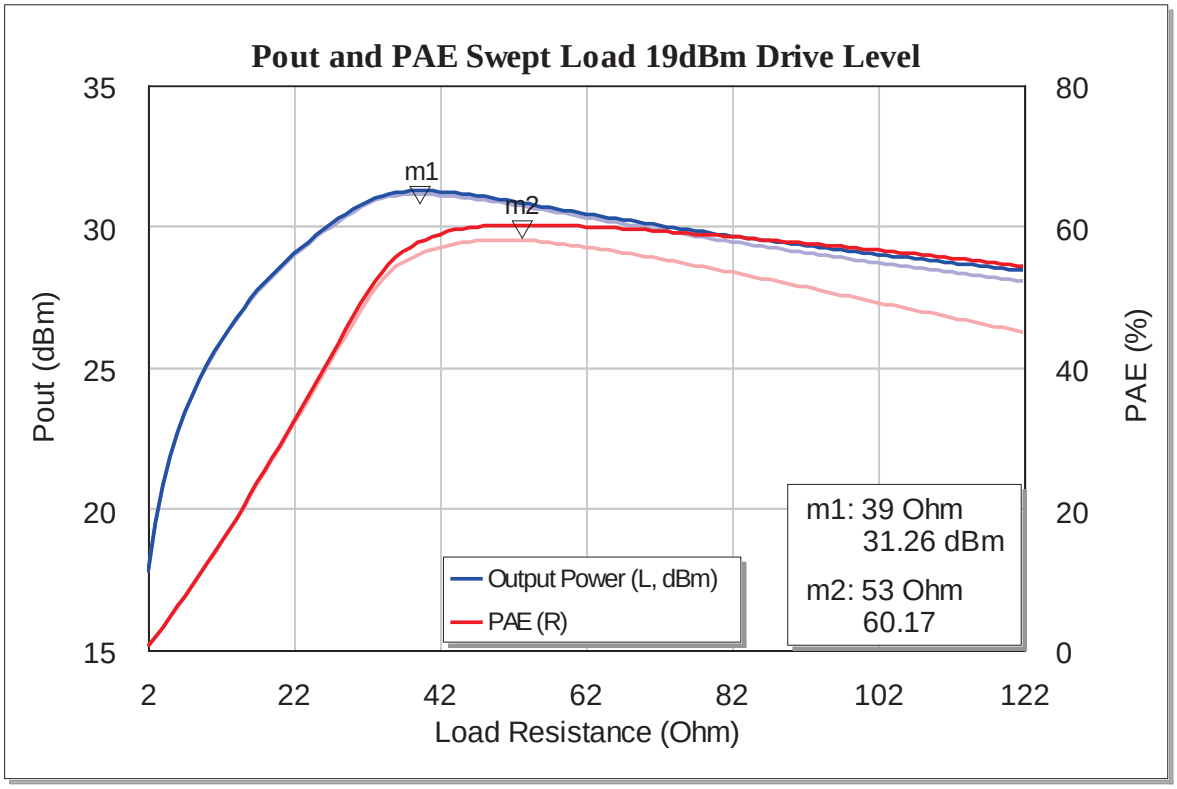


Figure 1-48, Swept load resistance at the CGP after de-embedding output capacitance (0.2pF).

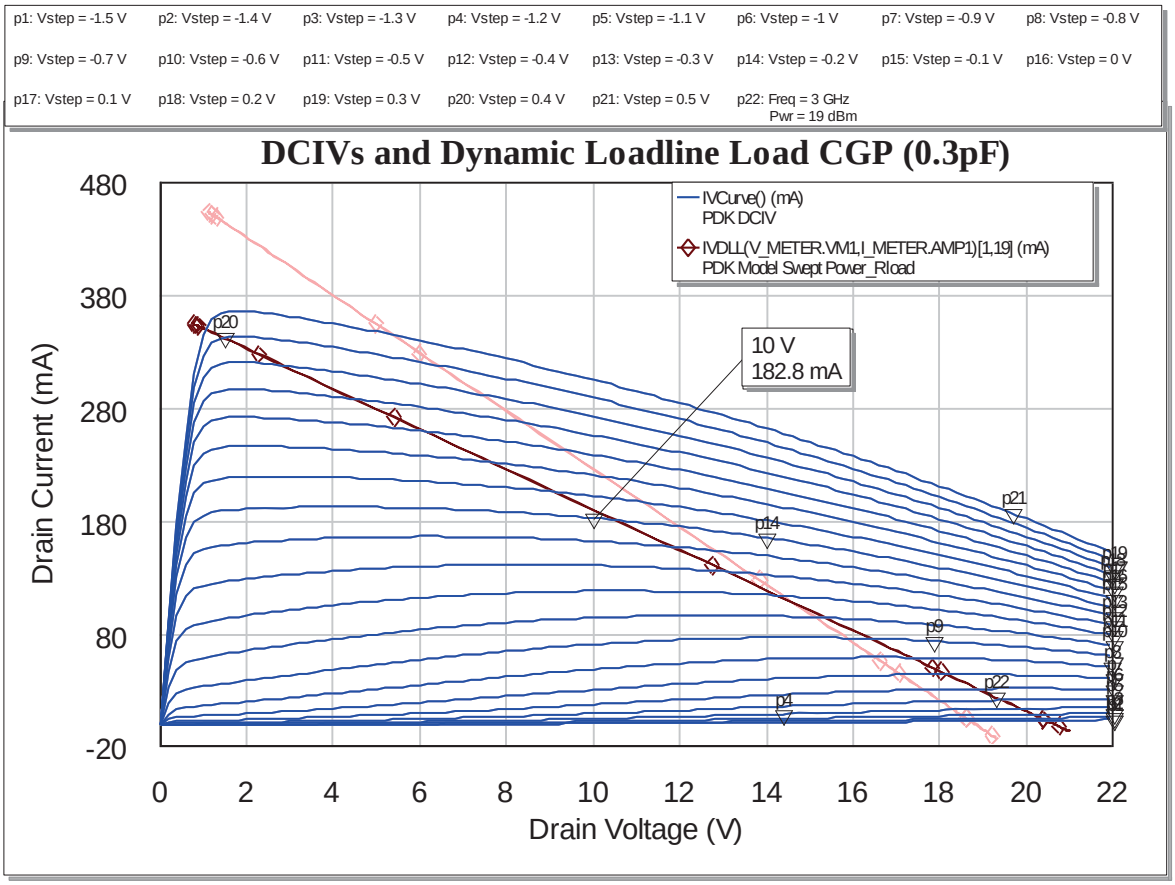


Figure 1-49, RF dynamic loadline at the CGP with 39Ω (faded) and 56Ω (bold).

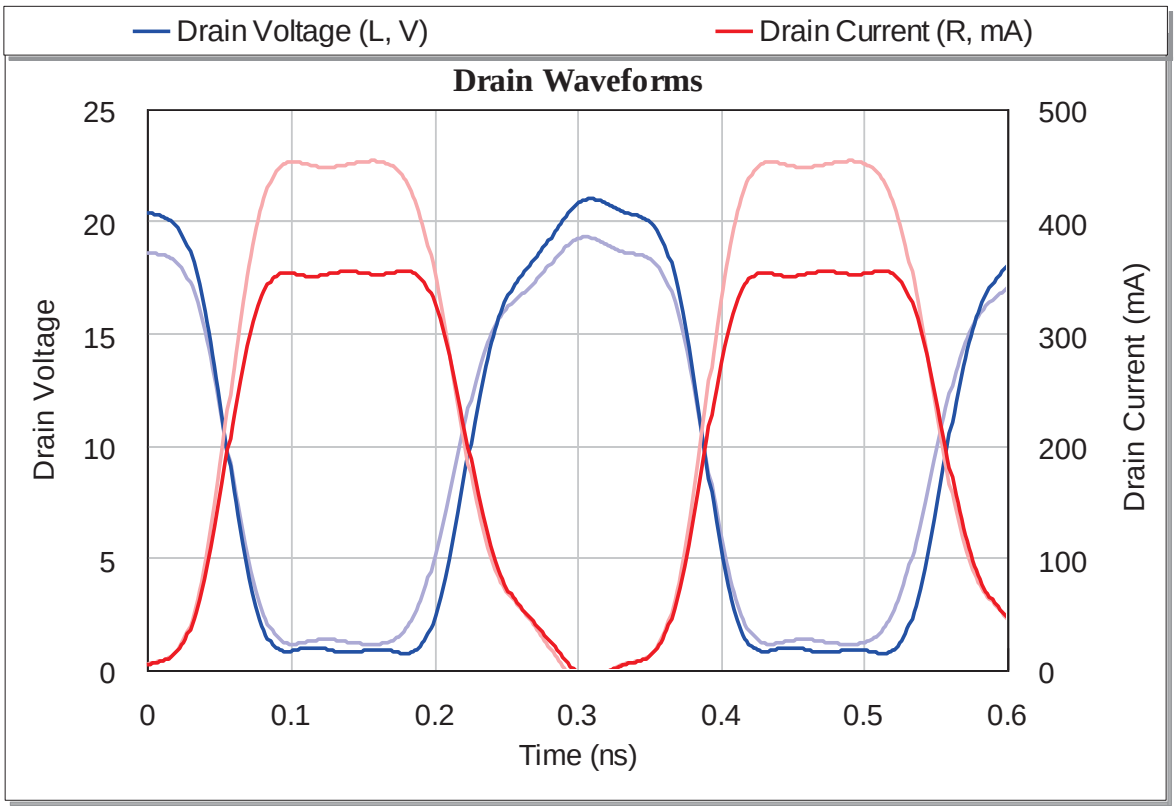


Figure 1-50, Drain waveforms at the CGP with 39 (faded) and 56Ω (bold)

The relationships between input drive level and the optimum load resistances remains constant as the PAE approaches the peak level and for output power the optimum resistance is reached a little earlier, Figure 1-51.

In an attempt to increase the efficiency of the device reduced conduction angle bias arrangements can be used as has been well documented, [34] and [2] amongst many others. The trade-off for higher efficiency is lower gain and output power and worse linearity, (although the degradation in linearity can be overcome to some extent by amplifier topologies such as balanced designs where two devices biased in class B are operated 180° out of phase and then re-combined through a phase rotating combining network). Other issues with these reduced conduction angle modes of operation are that the gain varies with drive level and the input match can vary dramatically with RF input signal. This is a particular concern with designers of radar amplifiers, frequently biased in class C. Caution should also be taken with regards to maximum power dissipation; class B devices will dissipate the most power as heat at the highest drive level, whilst class A become more efficient as they are driven harder and hence less power is dissipated as heat.

Rather than opting for a specific theoretical bias condition, designers may wish to

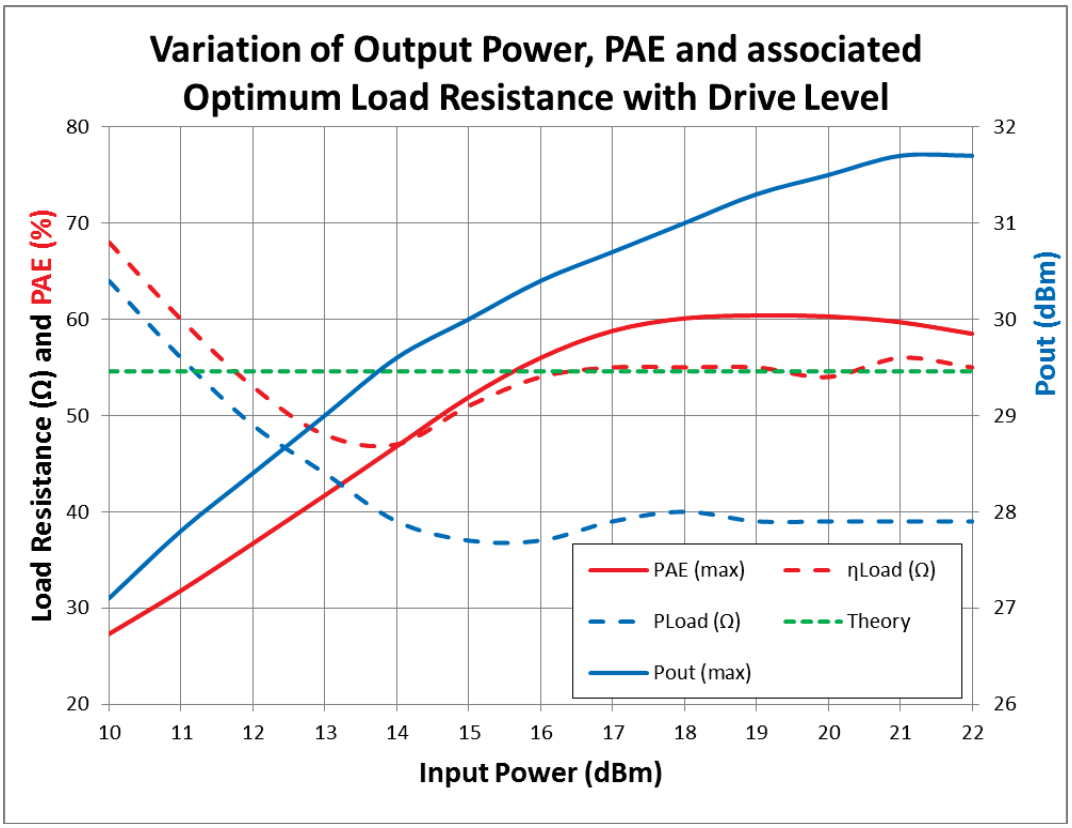


Figure 1-51, Relationship between drive level and optimum load resistance for output power and PAE.

select an intermediate or compromise condition. Again the waveform engineering and analysis technique allows the designer to obtain the optimum performance under these conditions. Consider the case of the device used earlier, but now with the bias adjusted to a quiescent current of only 3mA, which comes close to a class B condition. To compare the performance between class A and B the common approach would be to bias devices under the two conditions and perform a load pull to determine the optimum performance, Figure 1-52.

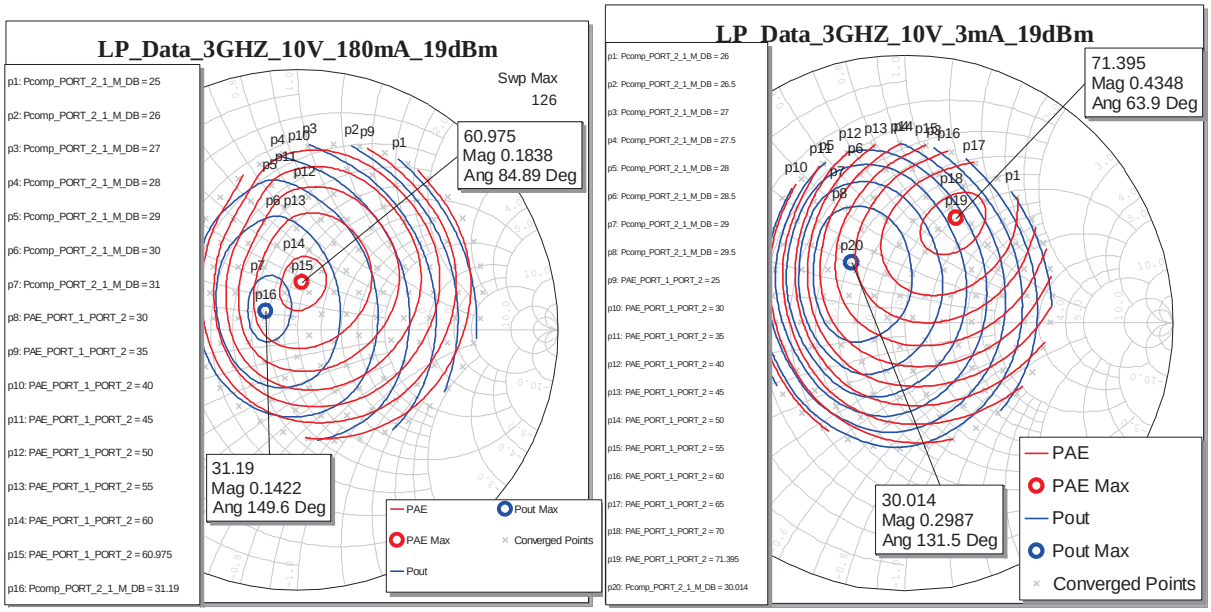


Figure 1-52, Comparison of classes A (left) and B (right) load pull contours for optimum PAE and output power.

From these measurements it is clear that the higher class produces an increase of over 10% in PAE, however not only is the maximum power in this bias lower (30.0 compared to 31.2dBm), but the output power and PAE optimum loads are further apart resulting in a significantly lower output power at the optimum PAE impedance. Loading the devices with the impedances indicated as the optimum for PAE from the load pull, the key characteristics can be compared with swept input power levels, Figure 1-53. This shows not only a substantial difference in the absolute level of the small signal gain but also the difference in gain variation with drive level. The lower gain has important implications at a system level which will be discussed in the next chapter, the gain shape shows gain expansion, which from a purely descriptive point of view creates confusion in the defining of the 1dB compression point; - from what point does one measure 1dB compression in the presence

of and amplifier with gain expansion? From a practical point of view this causes issues with amplitude control loops as the output power is not monotonic with input power.

Examination of the drain waveforms, between the two bias point conditions, Figure 1-54 and Figure 1-55, shows that although there are clear differences, they are 'clouded' by the presence of the output capacitance, and this needs to be removed by de-embedding to properly understand the differences. A further complication is due to the input match. This can be seen to be distinctly different not only in absolute value but also in trajectory with increasing drive level, Figure 1-56. The reasons for these differences will be discussed later in chapter 5 on device modelling, however suffice it to say that notwithstanding the possible model inaccuracies; input match is highly dependent upon both RF power level and device bias.

Returning to the load impedances and the need to de-embed the output capacitance, the determination of the correct amount of de-embedding capacitance depends on whether the optimum output power or the maximum PAE impedance is being considered, as is shown in Figure 1-57. If the output capacitance is completely accounted for then the terminating impedance will be purely resistive and the load will fall on the centre axis of the Smith Chart. Applying the appropriate loads and de-embedding capacitance for maximum PAE for each bias condition the drain waveforms and the dynamic load lines are plotted.

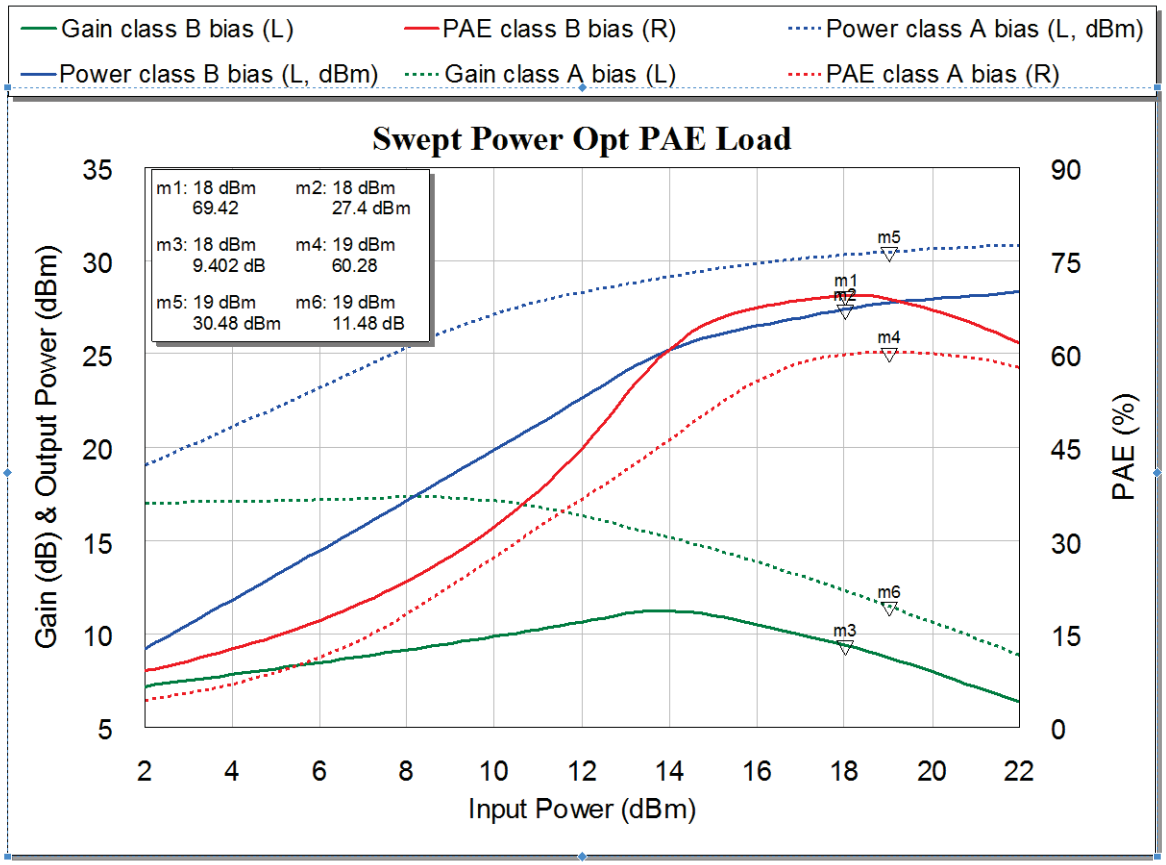


Figure 1-53, Comparison of class A and B performance with swept input power.

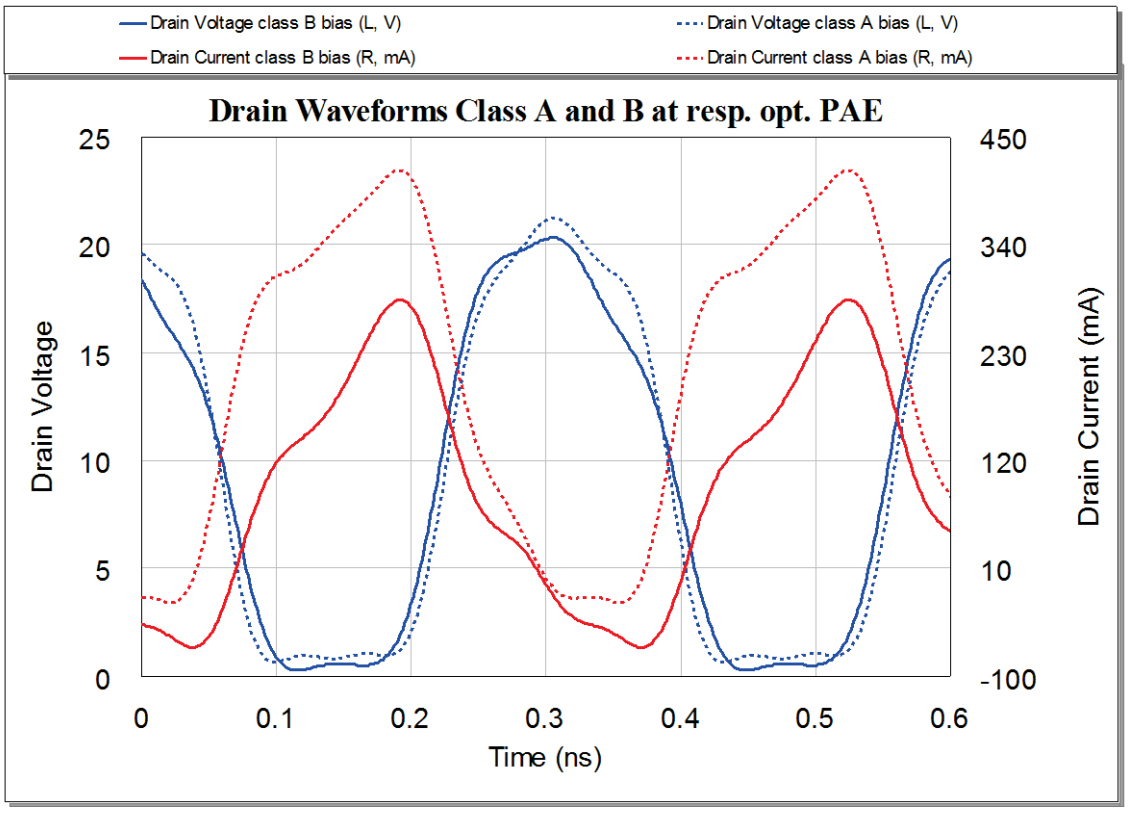


Figure 1-54, Drain waveforms of classes A and B loaded with their respective optimum PAE impedances.

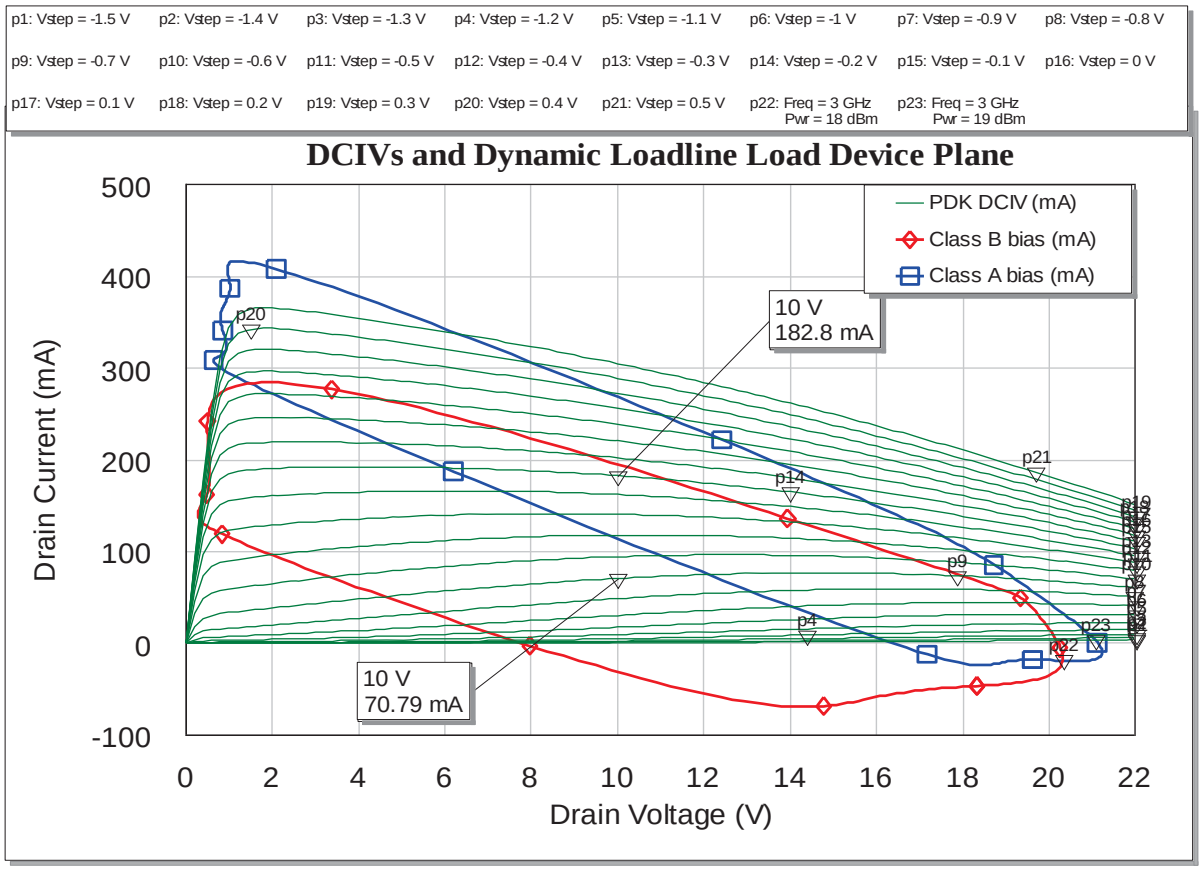


Figure 1-55, RF load line for class A ans B bias condtions at the respective optimum PAE load impednaces.

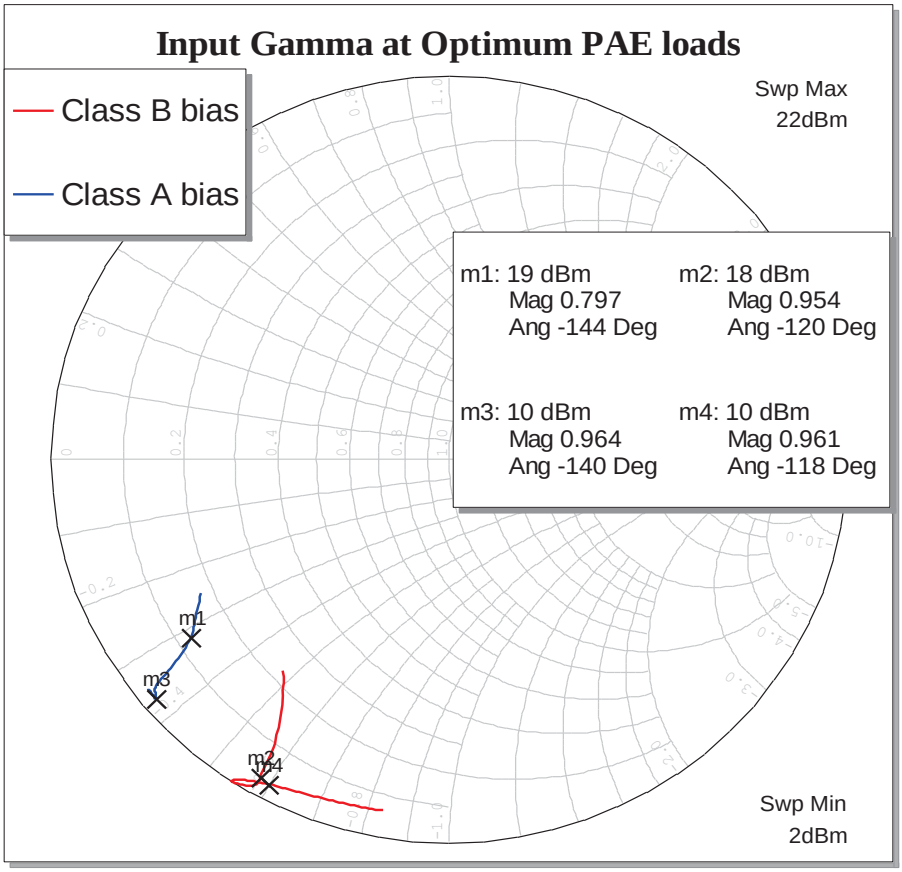


Figure 1-56, Different characteristics of input reflection coefficient between class A and B.

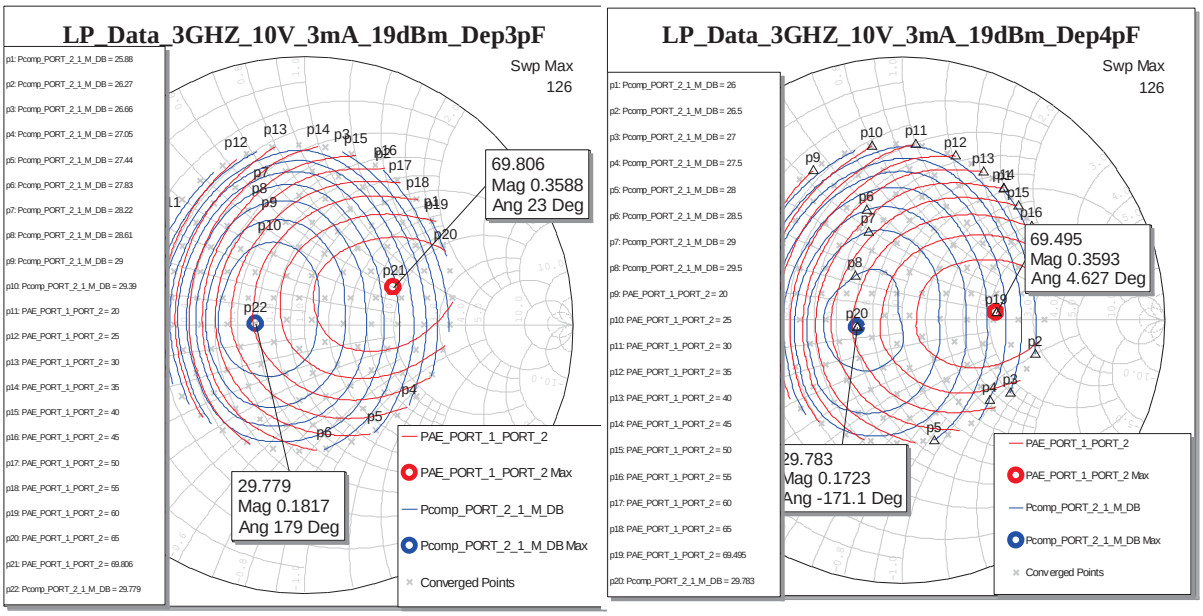


Figure 1-57, Class B load pull contours with 0.3 (left) and 0.4pF (right) de-embedding capacitance.

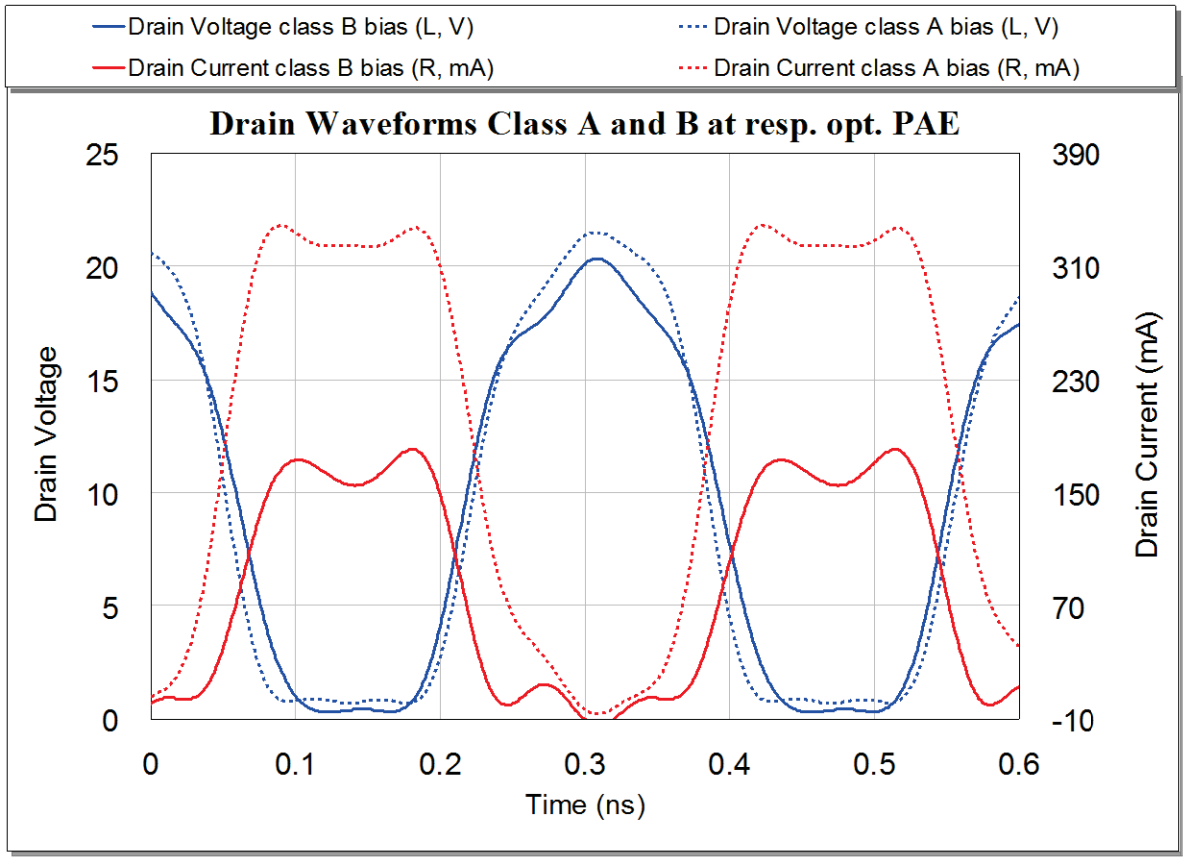


Figure 1-58, De-embedded drain waveforms at the optimum PAE load impedances and drive levels.

The higher power resulting from the class A bias is clear from the significantly higher (twice) drain current, Figure 1-58, and the higher efficiency is perhaps clearer from Figure 1-59. Although biased at a quiescent current of ~3mA at the maximum PAE input power level of +18dBm the quiescent bias has risen to ~70mA. This also agrees with the theory, referring back to Figure 1-29, where the higher efficiency load is shown to have a higher resistance.

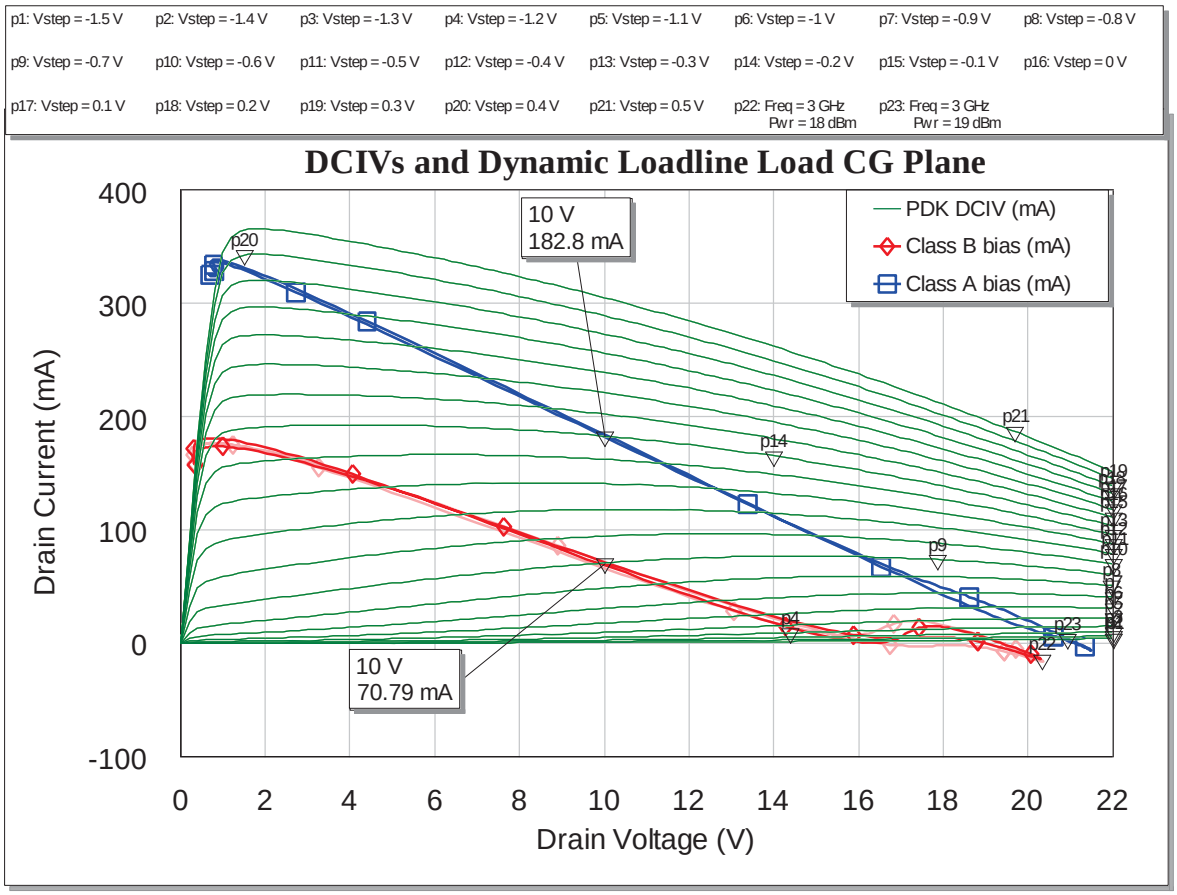


Figure 1-59, De-embedded dynamic load lines at the optimum PAE load impedance and drive levels.

Exploring further the movement of the class B load line with drive level, Figure 1-60, the quiescent current can be seen to increase directly, approaching that of half of the class A quiescent value. The two dominant peaks in the current waveform are indicative of the increasing 3rd harmonic contribution to the wave shape. The trough of the class B waveform is lower than that of the class A as it intercepts the knee at a lower voltage, (showing the impact of the finite slope of the linear region of the DC-IV curves).

Class	Cde (pF)	Opt PAE Load (Mag/Ang)	Pin (dBm)	PAE (%)	Pout (dBm)	Gain (dB)
A	0.4	0.10/_0°	19	60.3	30.5	11.5
B	0.5	0.36/_0°	18	69.4	27.4	9.4

Table 1-4, Summary of class A and B optimum PAE loads and performance.

The load for class B is given by [32] as {1-41}, in the case above the value of the load is therefore 100/(2x0.55) = 91Ω, compared with the simulated optimum of 106Ω (0.36/_0°). Thus the higher load resistance required for class B is consistent.

$$R_D = \frac{{V_S}^2}{2P_{OUT}} \qquad \{1\text{-}41\}$$

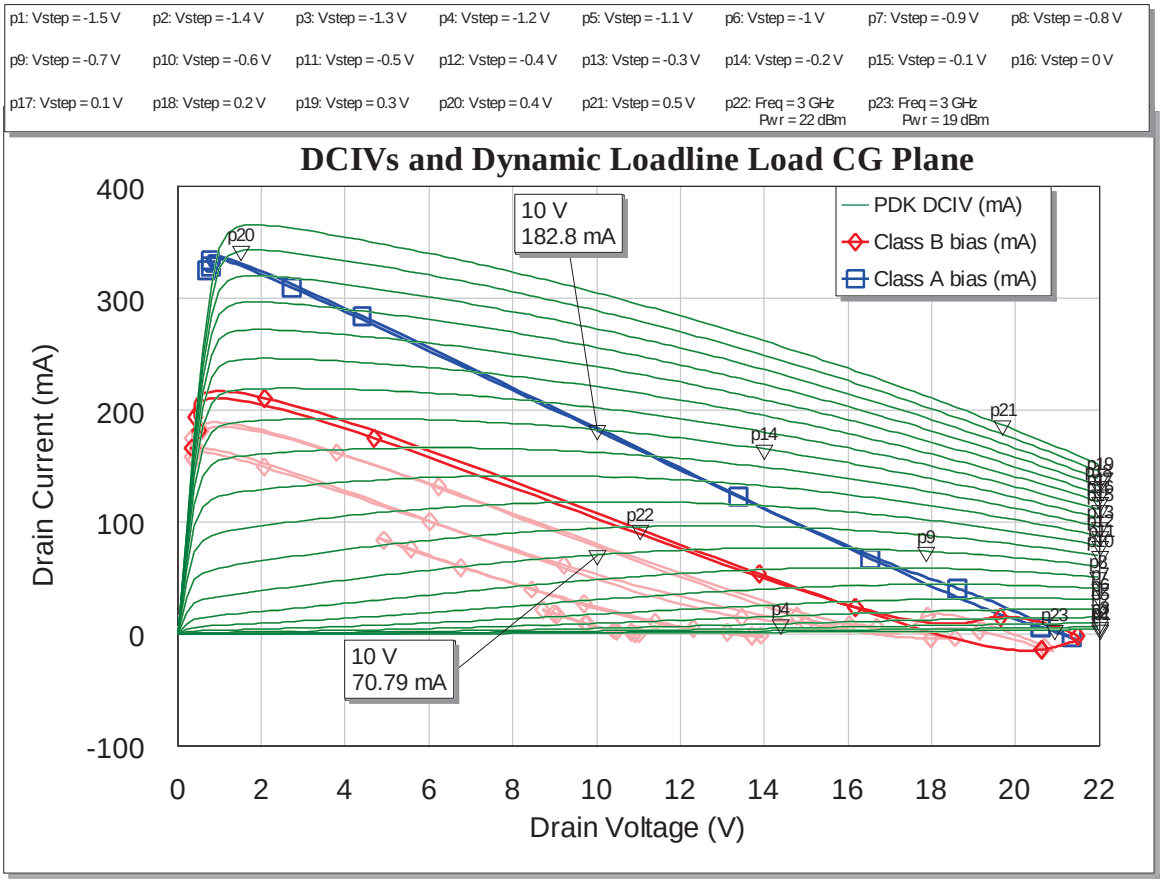


Figure 1-60, changes to class B load line with increasing drive level; input powers of 1, 10, 16, 19 and 22dBm (solid).

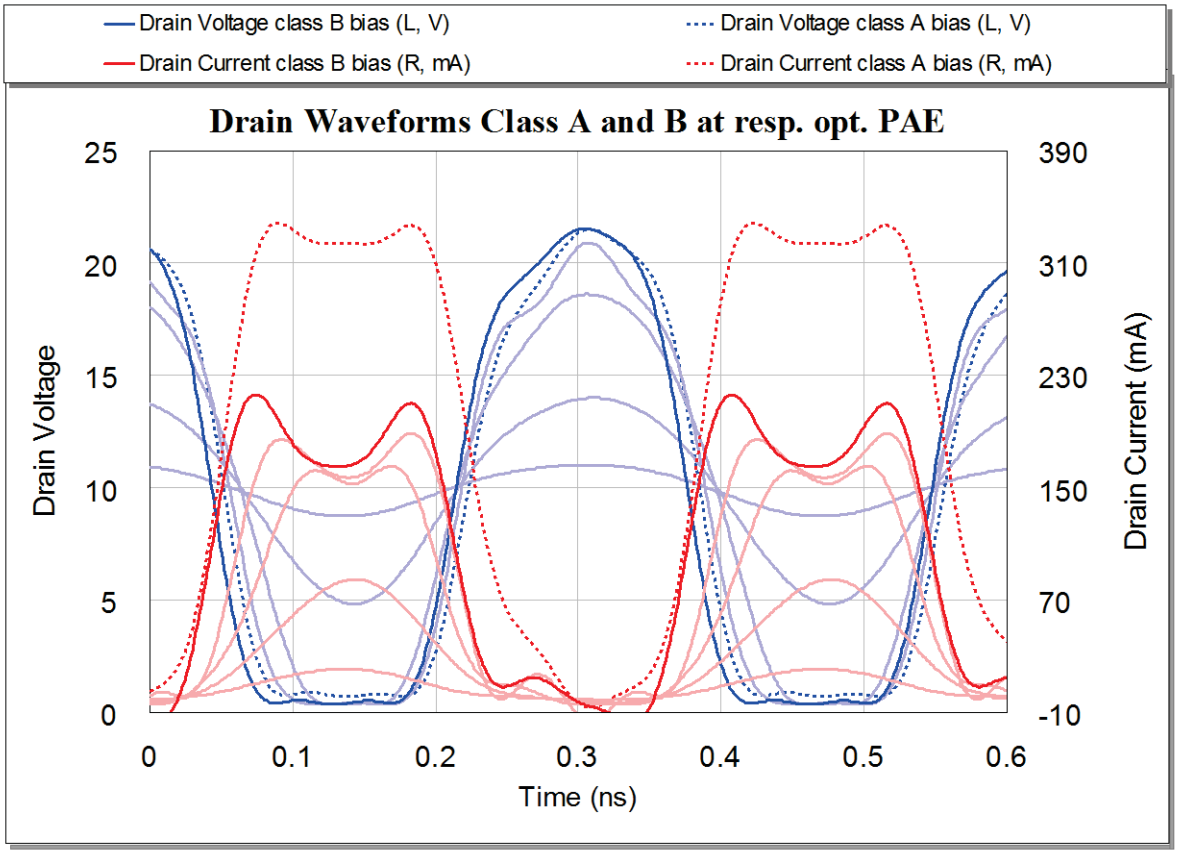


Figure 1-61, Drain waveforms with increasing drive level in the class B case, 1, 10, 16, 19 and 22 dBm (bold solid lines), compared with class A (dotted).

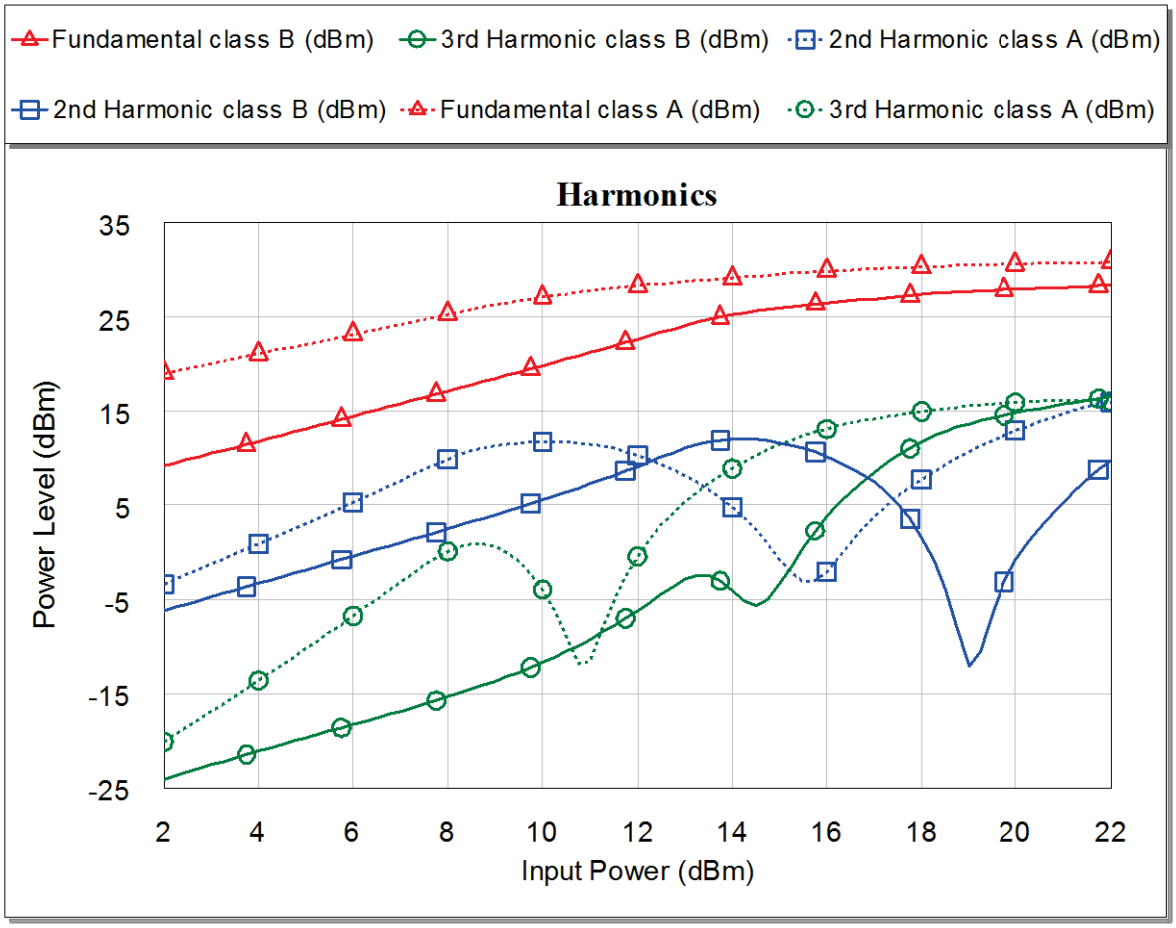


Figure 1-62, Simulated fundamental, 2nd and 3rd harmonic levels as a function of input drive power.

Finally, in this introduction to the waveform engineering approach, the impact of the harmonic impedances is investigated. It has already been noted that as the device approaches compression the harmonics increase, Figure 1-44. Simulating the harmonics from the class A and B circuits that have been analysed with the optimum PAE load de-embedded; interestingly it can be seen that the class B is predicted as having, on average, the lower harmonics, Figure 1-62. A word of caution, as will be seen in chapter 5, the accuracy of some nonlinear model harmonic predictions is suspect, but if absolute levels are ignored what is clear is that excepting the dips or sweet spots[4] harmonics increase with drive power.

[4] A term derived from determining the location of the drop in the harmonic power level, particularly the 2nd harmonic. This combination of bias and drive would be sought so as to minimise 3rd order intermodulation products which fall very close to the carrier and can interfere with adjacent channel in communication systems.

Numerous papers have now been published on the use of improving efficiency through the judicious use of harmonic impedances, such as [35] and [36], generally these have been narrow band although significant bandwidths have been achieved, [37], but still sub-octave. Little work has been done on harmonic tuning of class A amplifiers, probably because efficiency improvements are more easily obtainable by using a 'higher' operating mode. Thus a simple set of experiments are now undertaken to review the general effects of harmonic termination on the performance of a class A stage.

Returning to the class A case at the optimum PAE load, if we now add a harmonic trap at the output of the device which presents a short circuit to the even harmonics and an open circuit at the odd harmonics, and if the position or offset of this stub relative to the output of the device is allowed to be altered, there are now 3 variables, drive level, output load and stub offset that may affect the device's performance (assuming the bias is fixed in class A). Such a circuit is shown in Figure 1-63. Conducting an analysis of this circuit the performance can be displayed as a set of 3-Dimensional (3-D) graphs, Figure 1-64, and hence the optimum conditions for PAE can be determined. Setting the offset to 16° and the load to 55Ω a comparison can be made between the simple class A case and that with the additional 2nd harmonic stub termination.

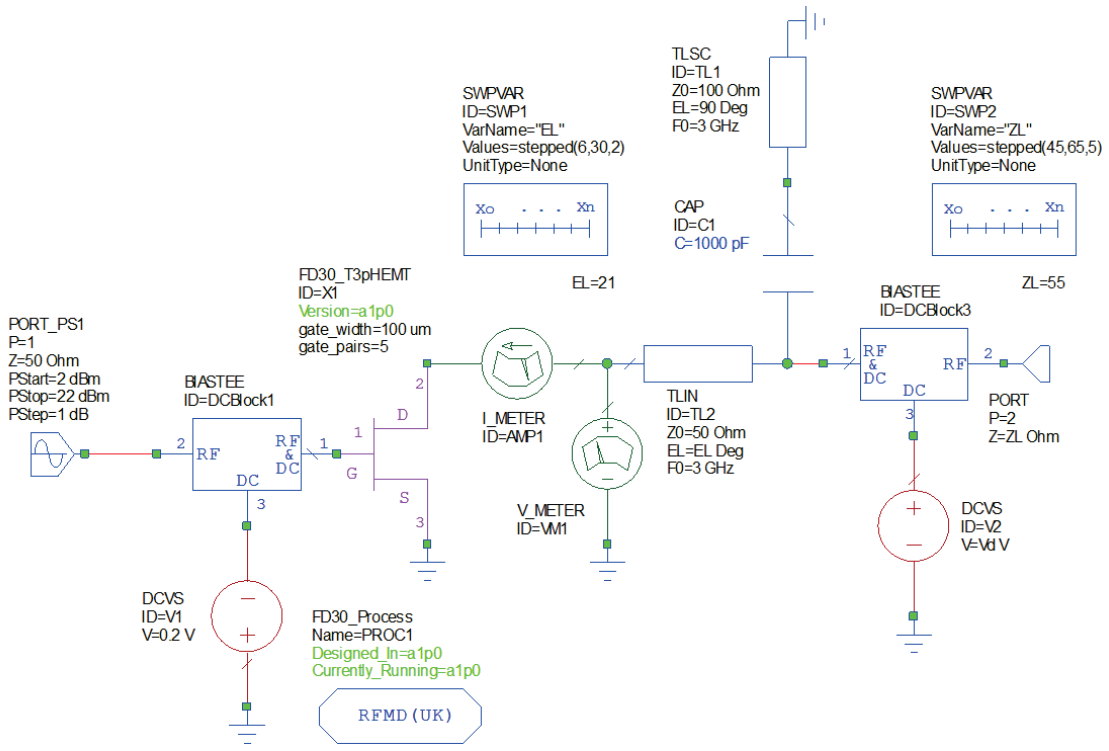


Figure 1-63, Schematic for device simulation with variables for input power, stub offset and load resistance.

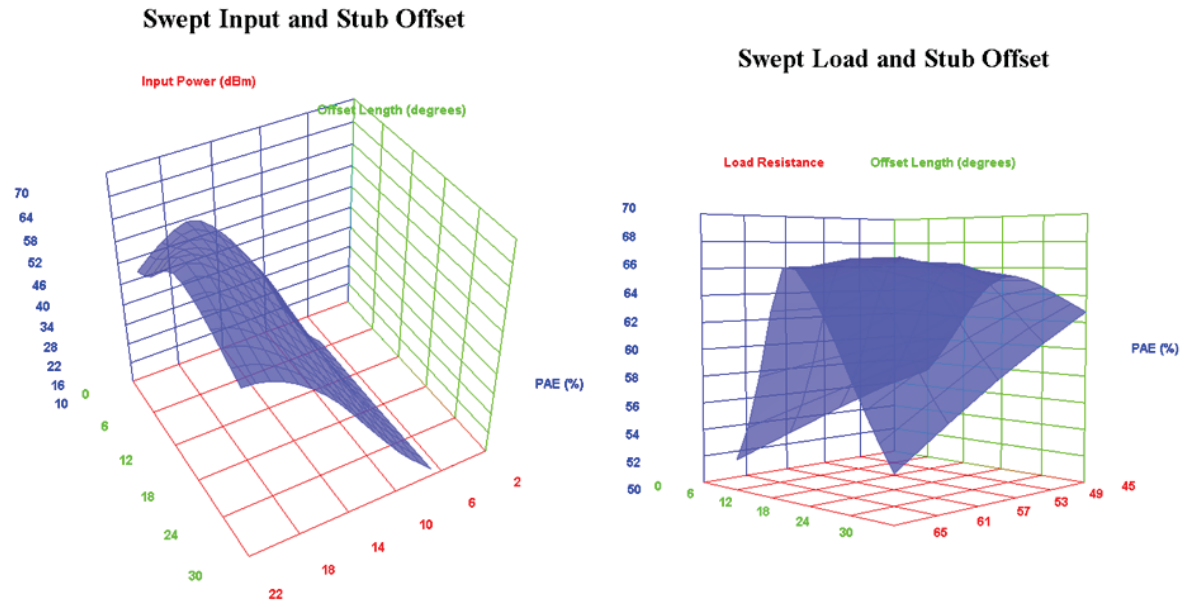


Figure 1-64, 3-D Graphs showing effect of input power, stub offset and load on PAE.

The 2nd harmonic load has made a significant difference, Figure 1-65; the maximum PAE has increased by almost 9%, although it now occurs at a 2dB higher drive level which coupled with the higher gain in compression and increased output power curve generally results in the output power at the maximum PAE being 0.9dB higher.

The impact of the harmonic load on the dynamic load line shows significant

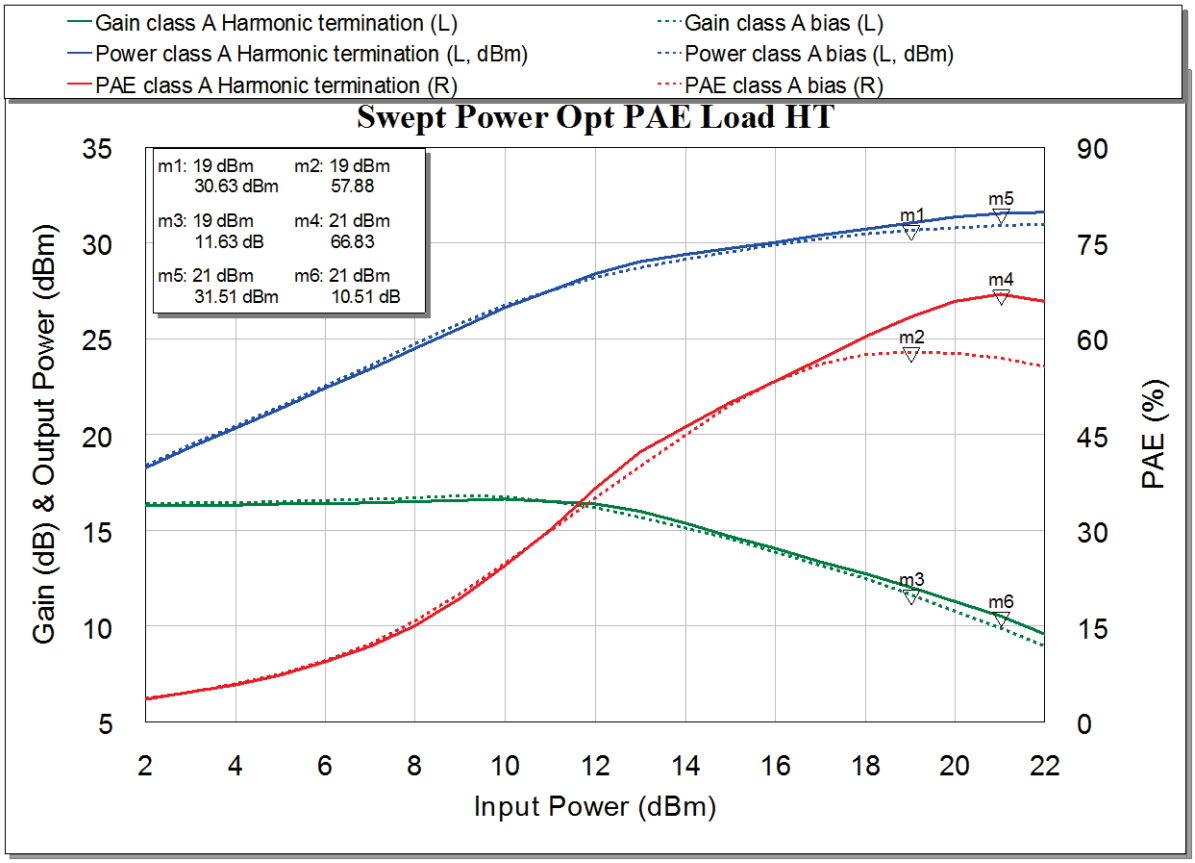


Figure 1-65, Comparison between resistively and 2nd harmonic short terminated class A transistor.

interaction with the knee, Figure 1-66, and greatly increased voltage and current swings. To understand the impact at the Current Generator Plane the intrinsic drain capacitance must be removed prior to the current meter, but then reinstated so that the impedance that the device sees remains the same. Figure 1-67 shows the waveforms at the CGP for the device operated with and without the 2nd harmonic offset short circuit stub. There is clearly a significant increase in both peak current and voltage, both waveforms do however show how some efficiency is being lost through the current and voltage waveforms not being completely in anti-phase.

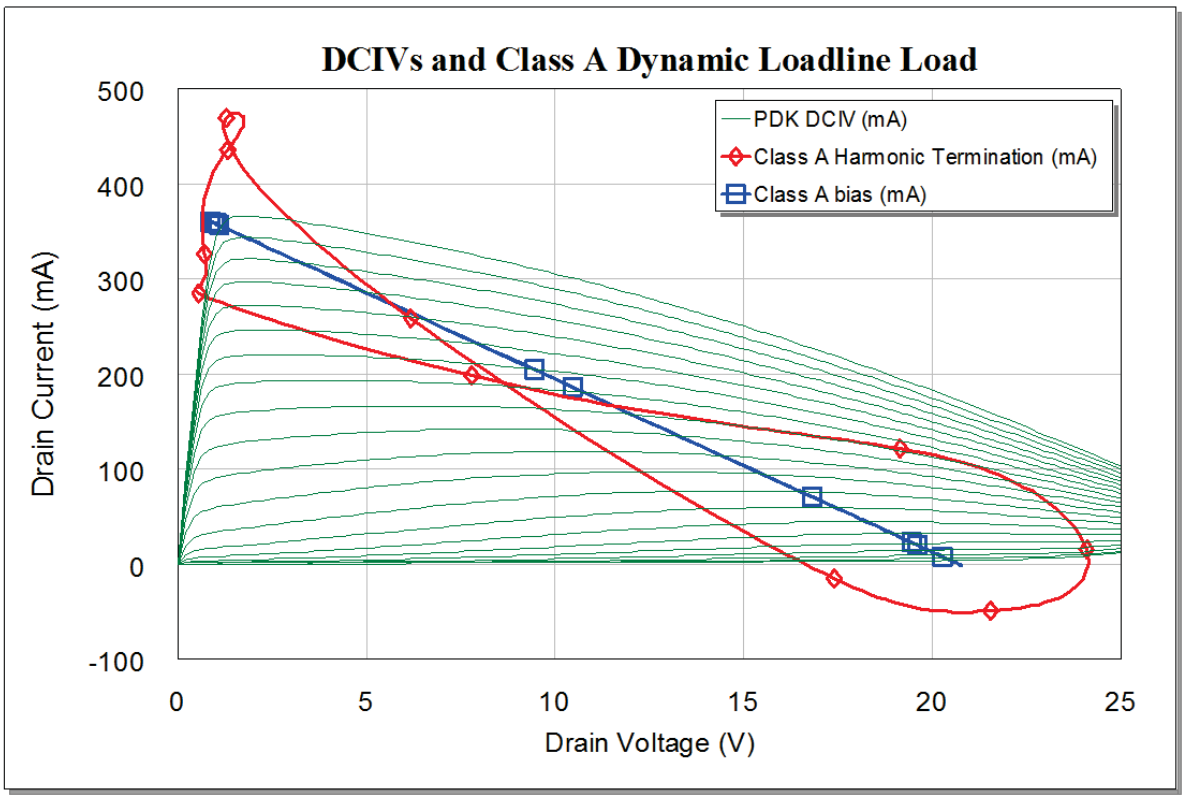


Figure 1-66, Dynamic load line with and without 2nd harmonic termination at the device plane.

This experiment clearly demonstrates that there is promise in using 2nd harmonic terminations to improve efficiency and power. Higher than 2nd order harmonics are not investigated further at this point from a purely practical perspective; in wide band amplifier designs achieving a satisfactory match over the fundamental bandwidth is a difficult enough proposition, adding the 2nd harmonic creates a significant further challenge, designing matching circuits to encompass above this is believed to be impractical. However investigation of the 3rd harmonic implications will be conducted in the measurement chapter (4) of this work. In this experiment the offset and load were adjusted as linear

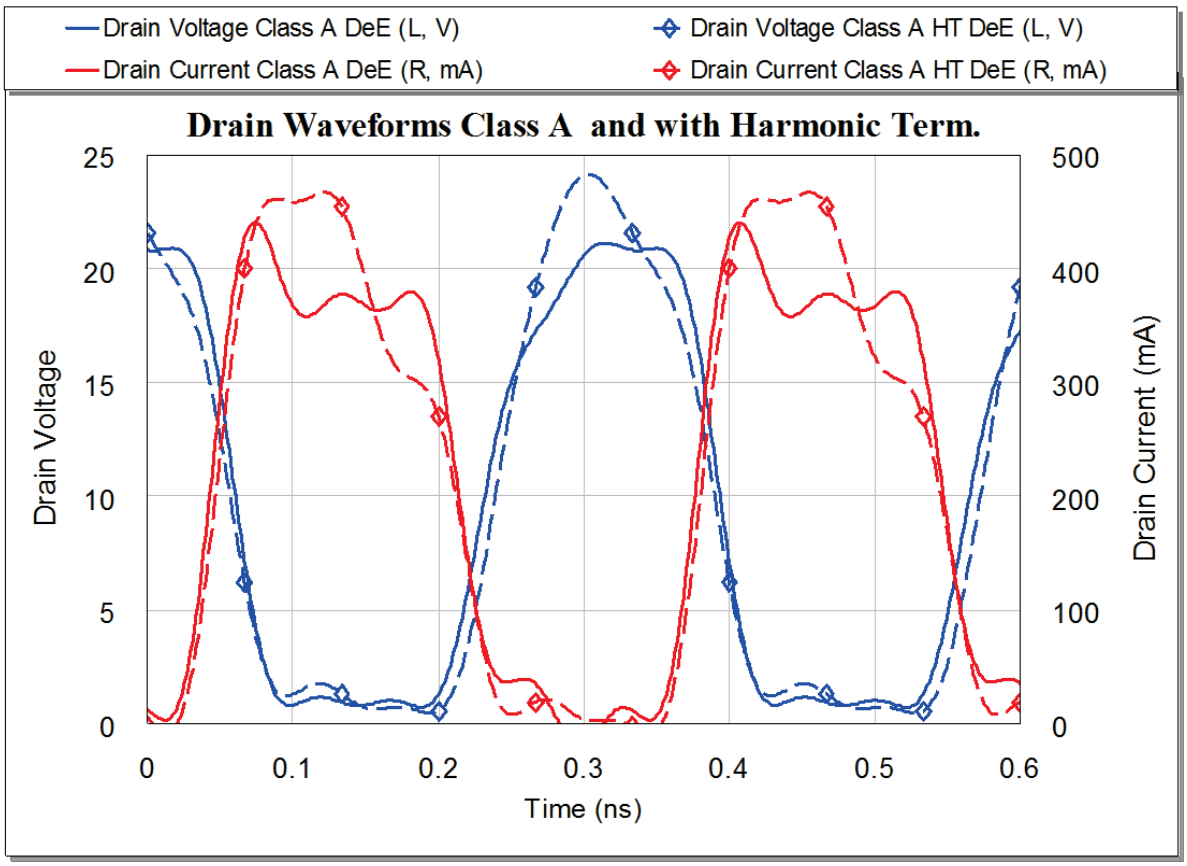


Figure 1-67, CG Plane waveforms with and without 2nd harmonic terminations.

variables to map the performance changes. In 'real world' measurements this will be done using load pull measurements (fundamental and harmonic), at different drive levels.

1.6 Class J Operation

Most high efficiency modes of operation rely upon narrow band output tuning and hence are not relevant to broadband design approaches. However operation over almost an octave has been achieved [37] using class J and therefore an explanation of this mode is thought relevant to the general discussion. Class J is based upon class B, but uses the output capacitance to provide a specific reactive termination at the 2nd harmonic.

As in the standard approaches the first step is to determine the optimum load, R_{Lopt}.

$$R_{Lopt} = \frac{(V_{DC} - V_k)}{I_{max}/2} \qquad \{1\text{-}42\}$$

The theoretical approximations start from the very beginning. Firstly what is I_{max}? Looking at Figure 1-68, the DC-IV curves from the nonlinear model for the CGH40025F, 25W discrete GaN device from Cree, it can be seen that I_{DS} is not flat with increasing drain voltage above ~1amp.

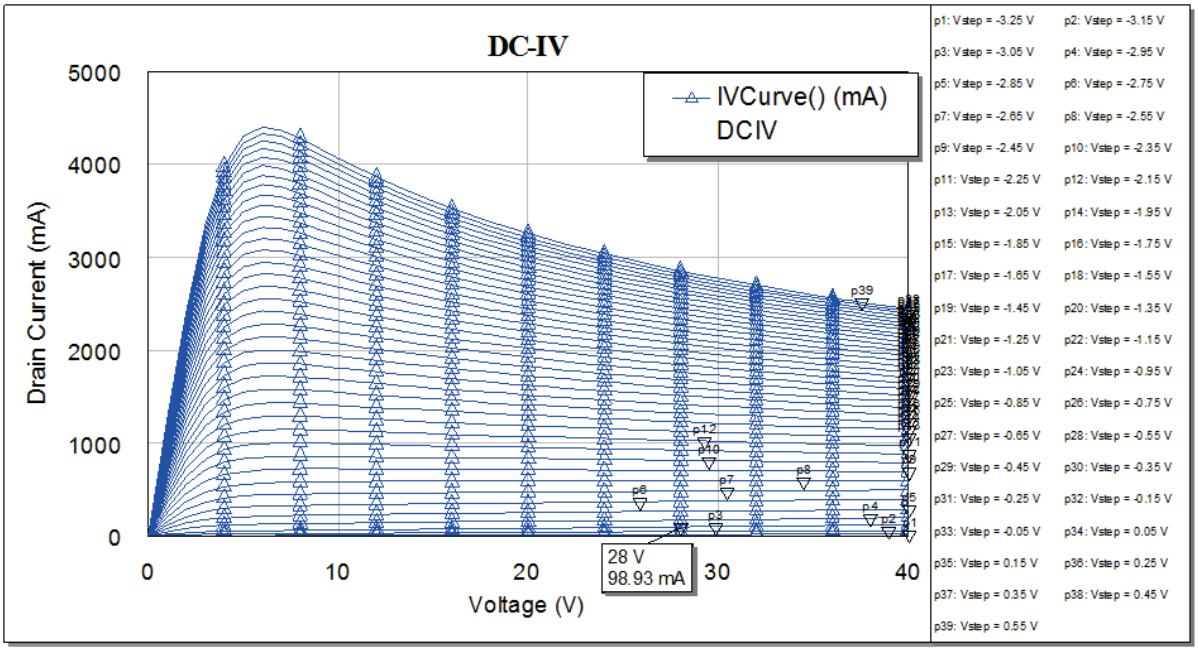


Figure 1-68, CGH40025F DC-IV curves

Cripps suggests [2] that I_{max} be taken as the value of current at a low drain voltage (where current is a maximum) and a slight forward bias on gate. In this case V_{DC}=3V and V_{GS}=0.5V, will be used, giving a value for I_{max} of 3.3A as shown in Figure 1-69. Similarly the value of V_k is not clear; the 'knee' varies with gate and drain voltage, and is not a distinct point in itself anyway. Again Cripps suggests that it is the voltage at which the current reaches 67% of its maximum value, in this case therefore the voltage for an I_{DS} of 2.2A, V_k ~ 1.82V.

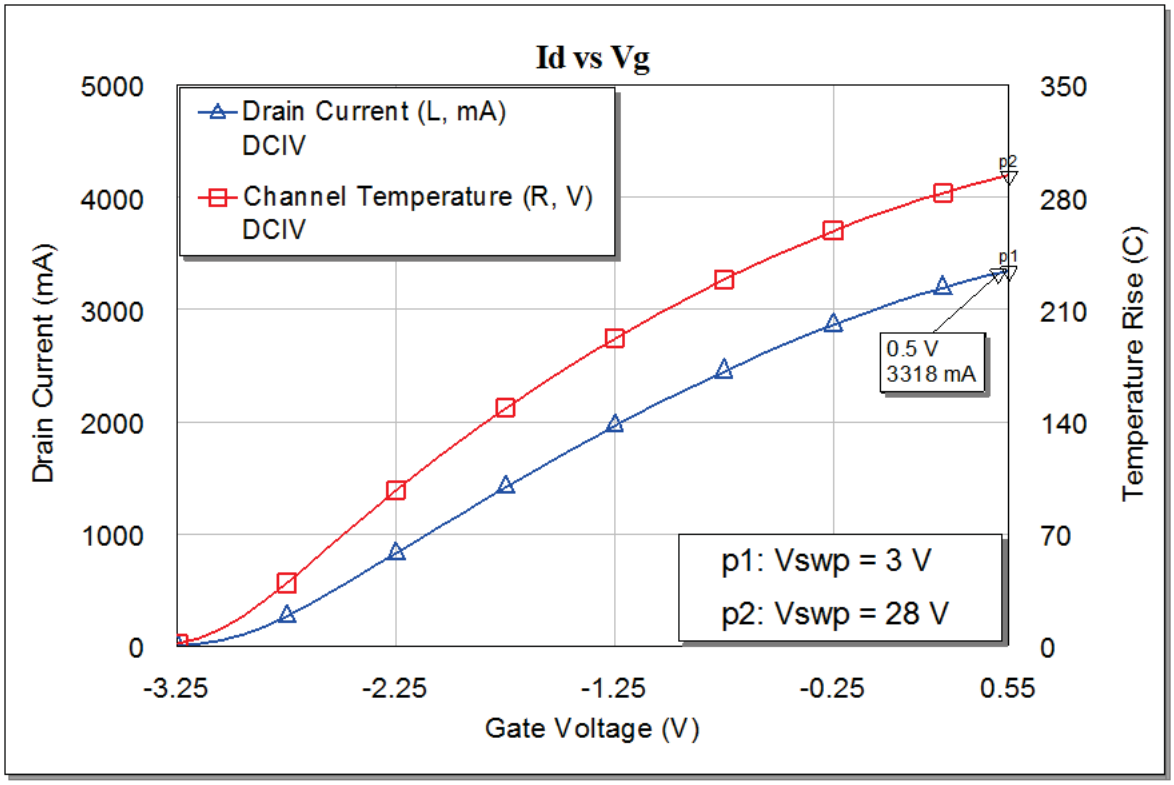


Figure 1-69, Drain Current vs. Gate voltage at 3V supply

Thus we can solve {1-42}, R_{Lopt} = (28-1.82)/1.65 = 15.9Ω

The next area of approximation is due to the fact that the theory considers the device without the packaging in which they are conventionally supplied, this does not affect the DC-IV curves but will impact upon the RF waveforms thus for the initial analysis it is necessary to de-embed the model provided by the manufacturer to the device plane, for MMICs this would consist of removing the wafer probe pads, feed lines and drain

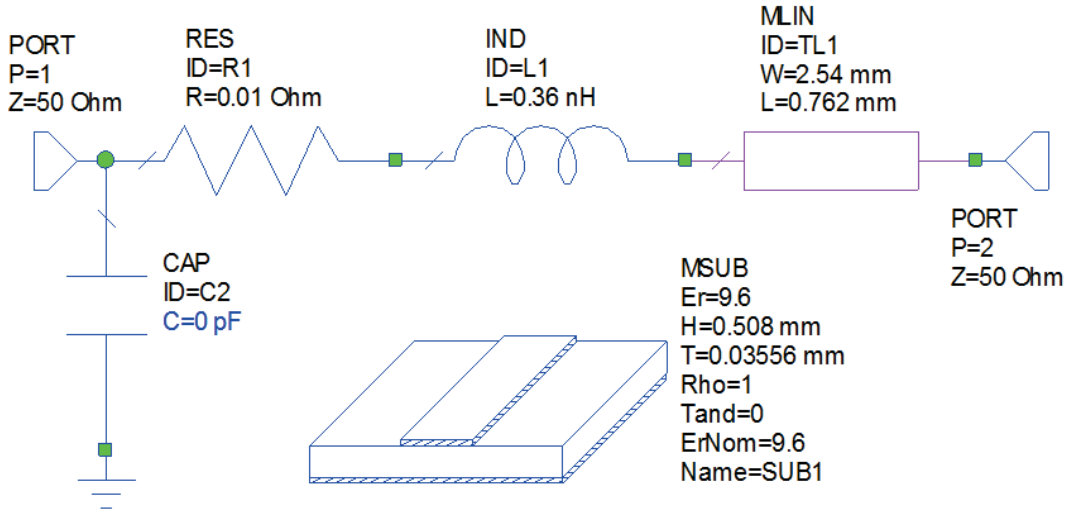


Figure 1-70, Output parasitic elements for CGH40025F

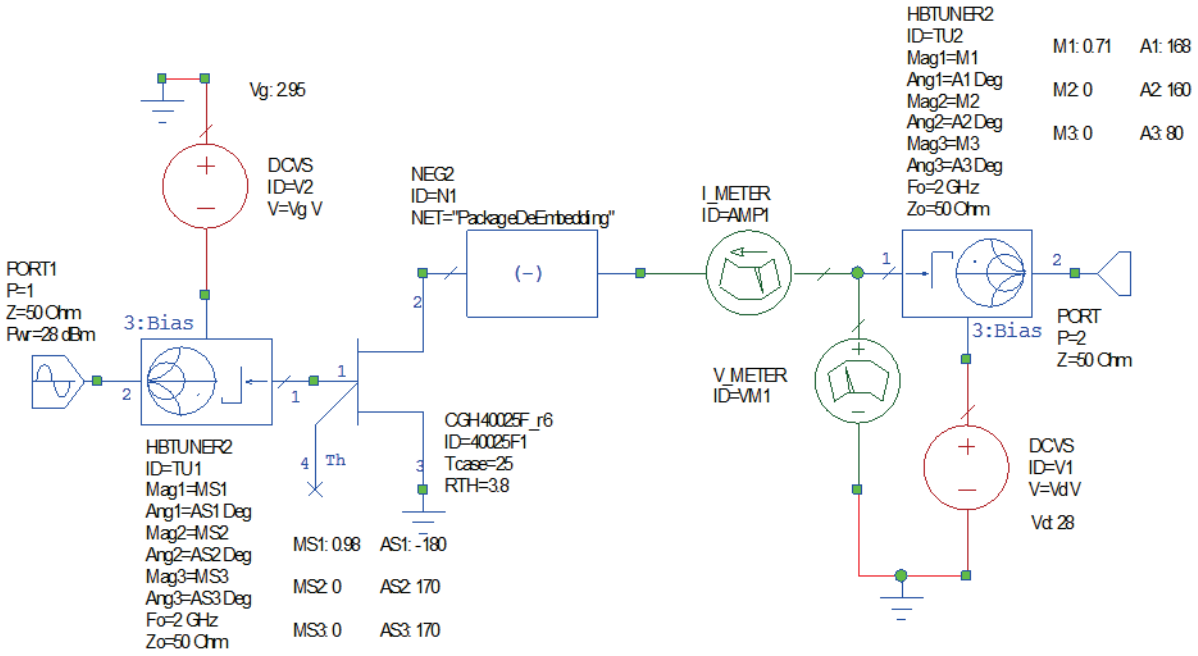


Figure 1-71, Device analysis circuit with output de-embedding

capacitance. In this case the model for the output package parasitics is shown in Figure 1-70 and the nonlinear analysis circuit in Figure 1-71. Note that within the de-embedding circuit provision is made to remove the effects of C_{DS}; however at this stage this is not implemented (C_{DS}=0pF). The correct bias current for class J operation (based on class B) is typically a figure of 5% of I_{DS}, so in this case would be ~165mA, which requires a gate voltage of ~-2.95V.

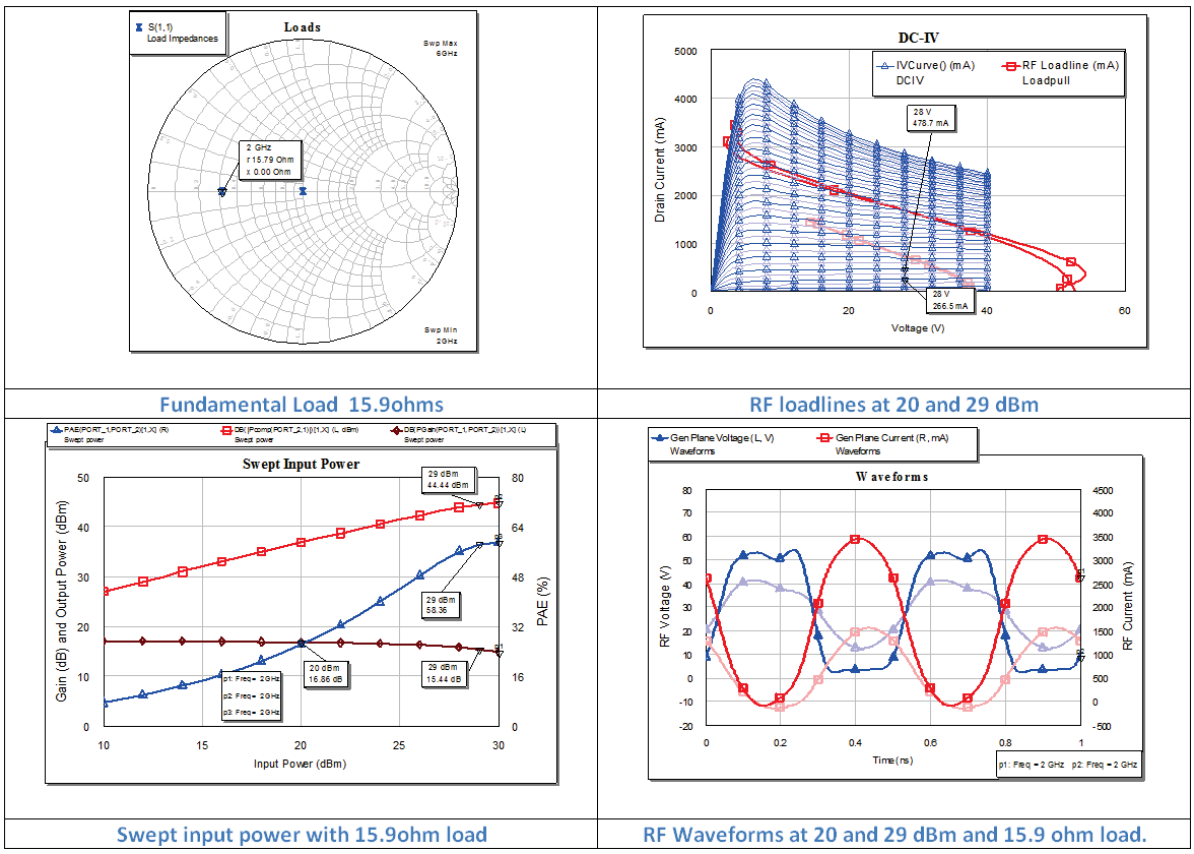


Figure 1-72, Operation of de-embedded device in class B, load 15.9Ω

Ignoring the output capacitance for the moment and looking at the device with R_{Lopt} at two input power levels, 20dBm and 29 dBm, corresponding approximately to the linear and peak PAE drive levels (with input tuning). We can see that the peak PAE is below that expected for class B (78.5%), Figure 1-72. According to [38] for class B the optimum load should be increased by a factor of √2, hence to 22.5Ω, Figure 1-73.

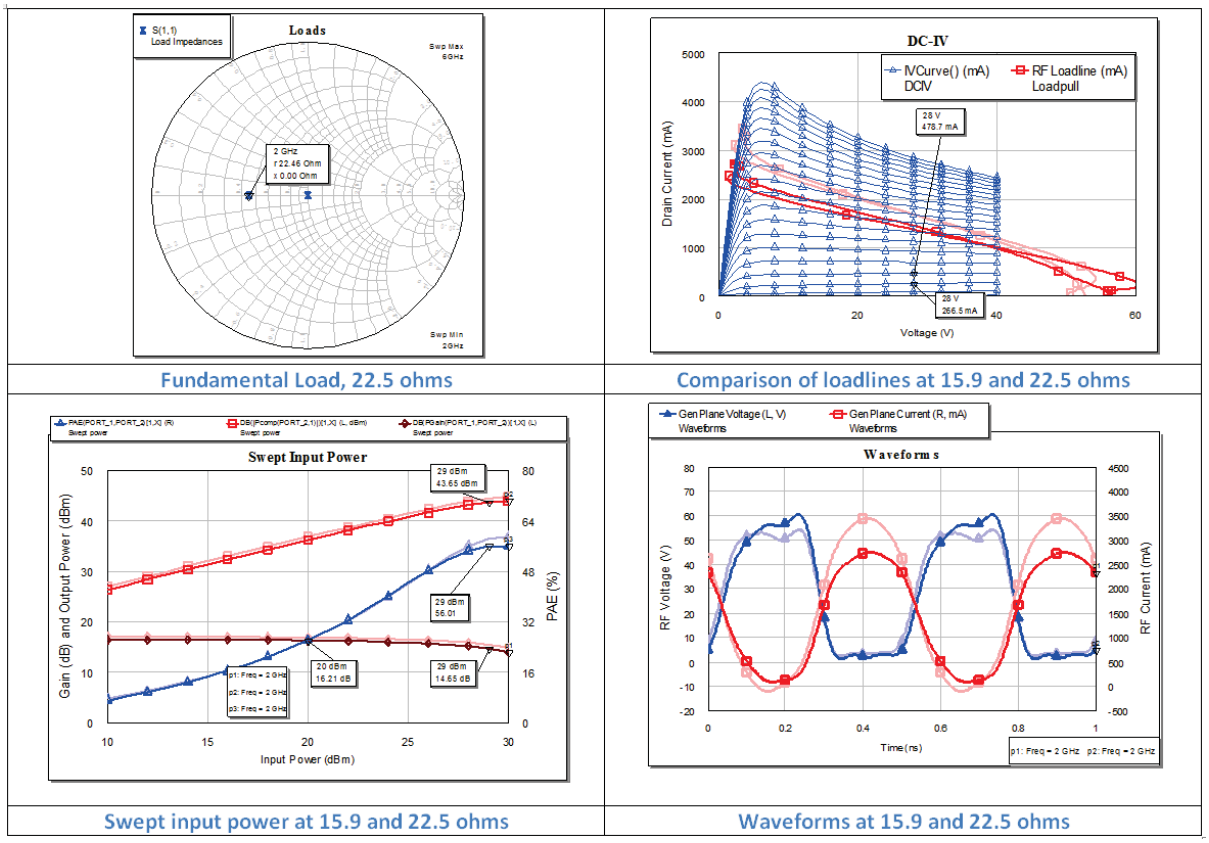


Figure 1-73, Increasing the load to 22.5Ω

Adding in the 45° phase shift for class J makes the load 22.5/_ 45° = 15.9 +j15.9 which is a gamma of 0.56/_141.4°. Further the magnitude of the 2nd harmonic load is given as $(3\pi/8)R_{Lopt}$ = 18.73, hence 2nd harmonic load is 0 – j18.73 which is a gamma of 1/_-138.9°, see Figure 1-74.

De-embedding the output capacitance, C_{DS}, which is approximately 2.2pF, the 'true' class J waveforms can be seen, Figure 1-75.

The predicted PAE is 75%, which is now close to the theoretical value for class B. Also note that the peak in efficiency is relatively flat with drive power. By appropriate phasing the 3rd harmonic the drain voltage can be 'peaked' up which will increase the PAE to over 80%, Figure 1-76.

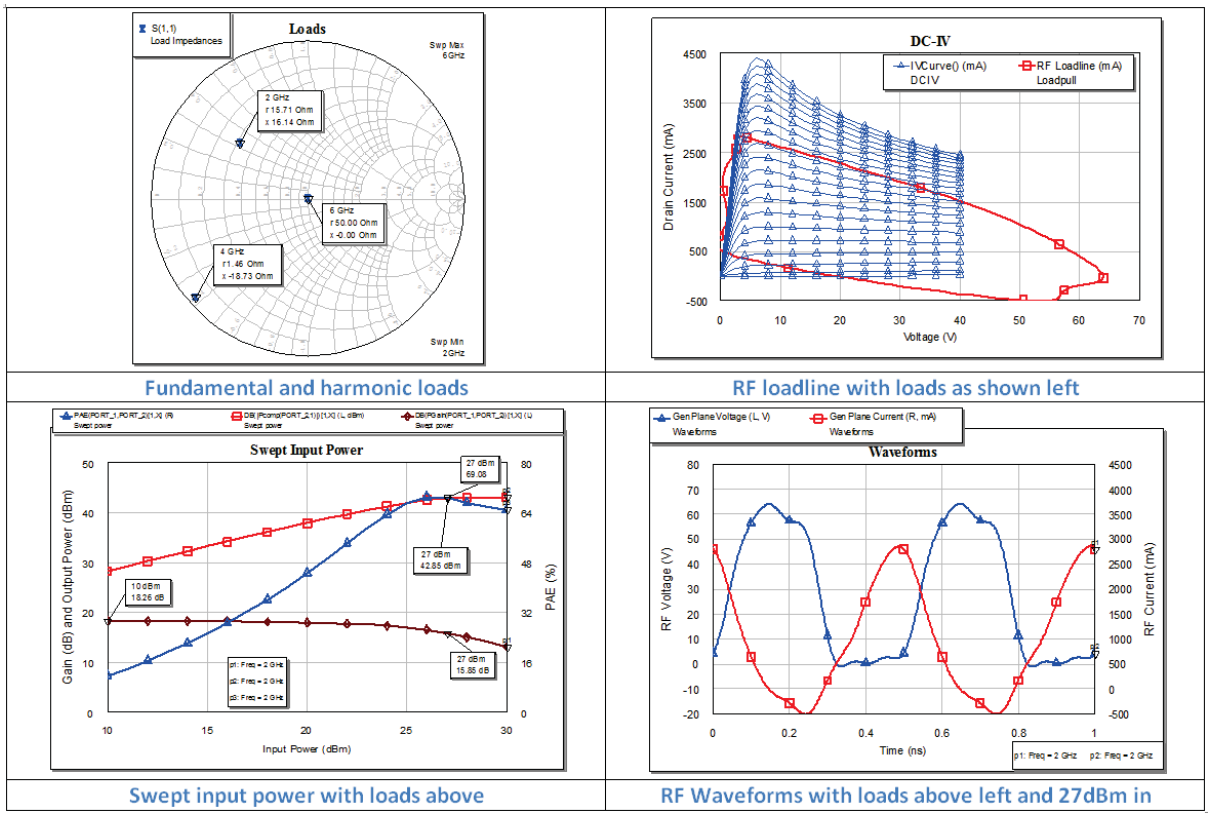


Figure 1-74, Performance with fundamental load of 0.56/_141.4° and 2nd harmonic of 1/_-138.9°

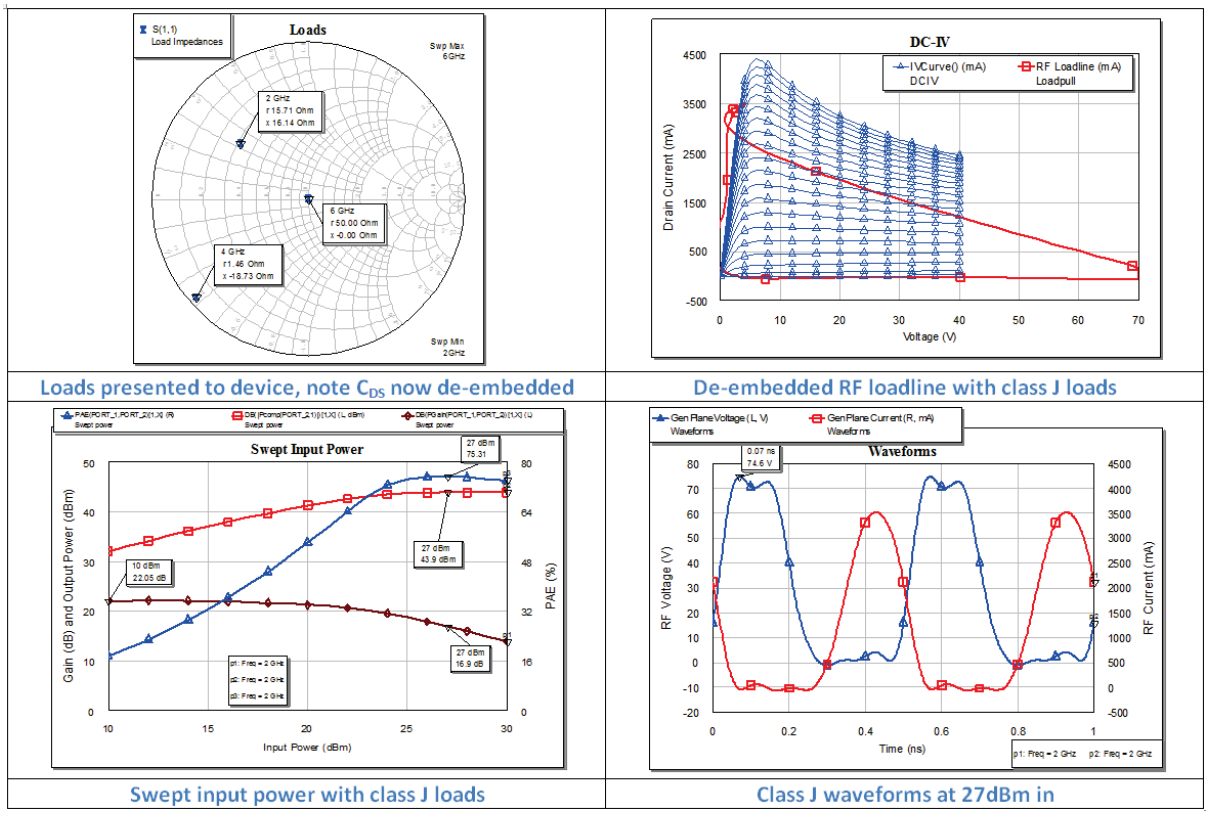


Figure 1-75, Class J waveforms and performance

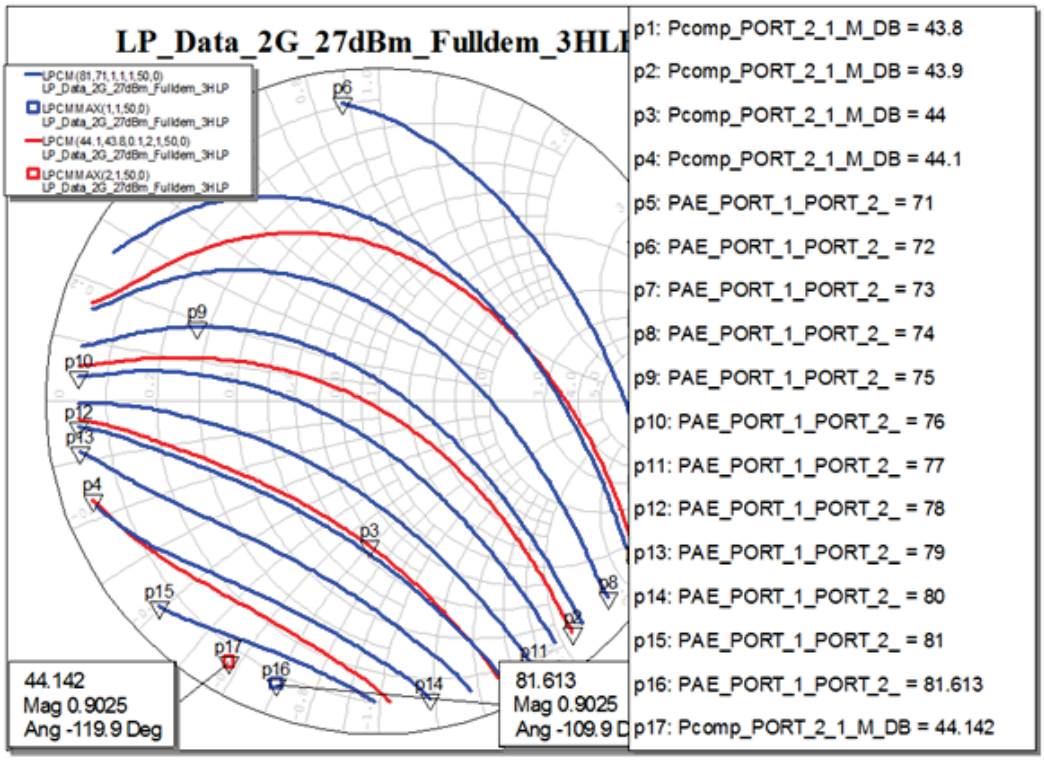


Figure 1-76, Impact of 3rd harmonic load pull on class J operation

A consequence of class J operation are the very high peak voltages ($3.2xV_{DC}$), Figure 1-77, which why this approach is suitable for GaN but few other technologies.

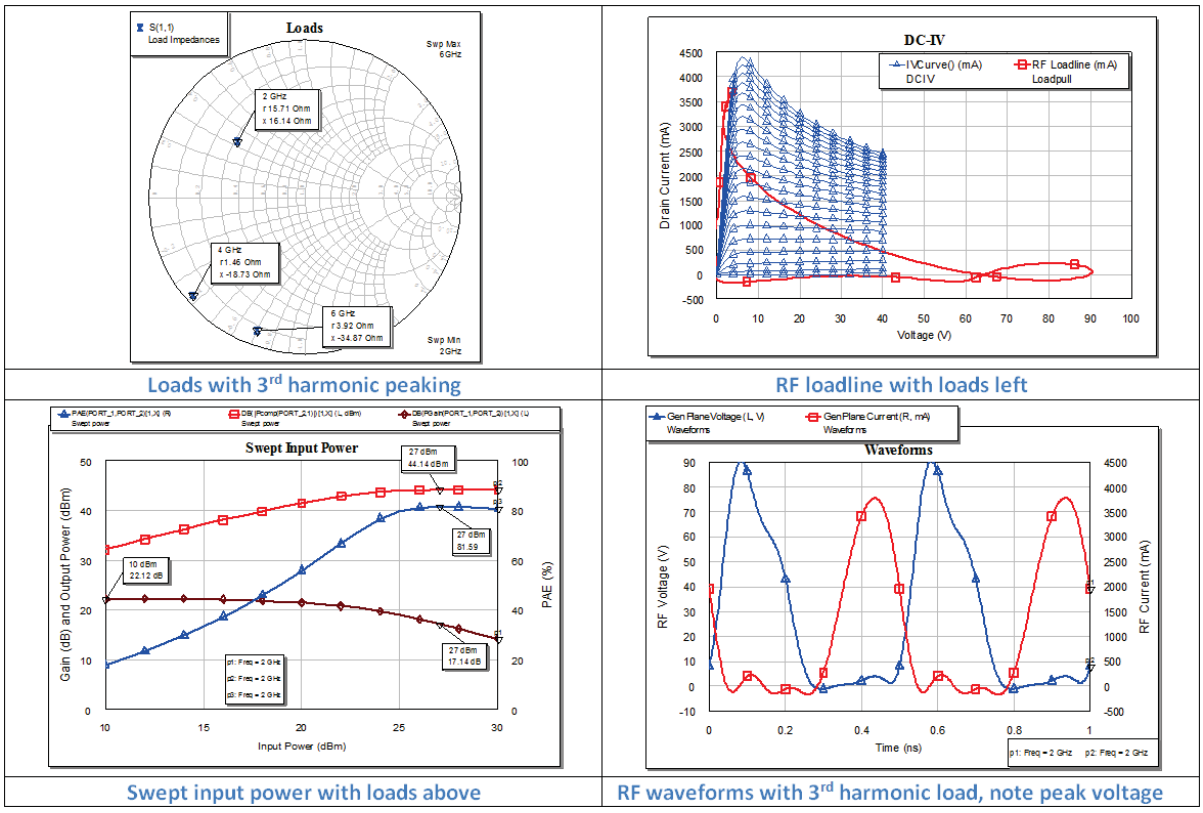


Figure 1-77, Class J operation including 2nd and 3rd harmonic loads

There are a series of loads which produce the class J operation, often referred to as the class J continuum. The 2nd harmonic is rotated around the impedance plane and the optimum fundamental impedance found for each point (by load pull), Figure 1-78. It can be seen that as the angle of the fundamental load rotates beyond an angle of ~±135° the

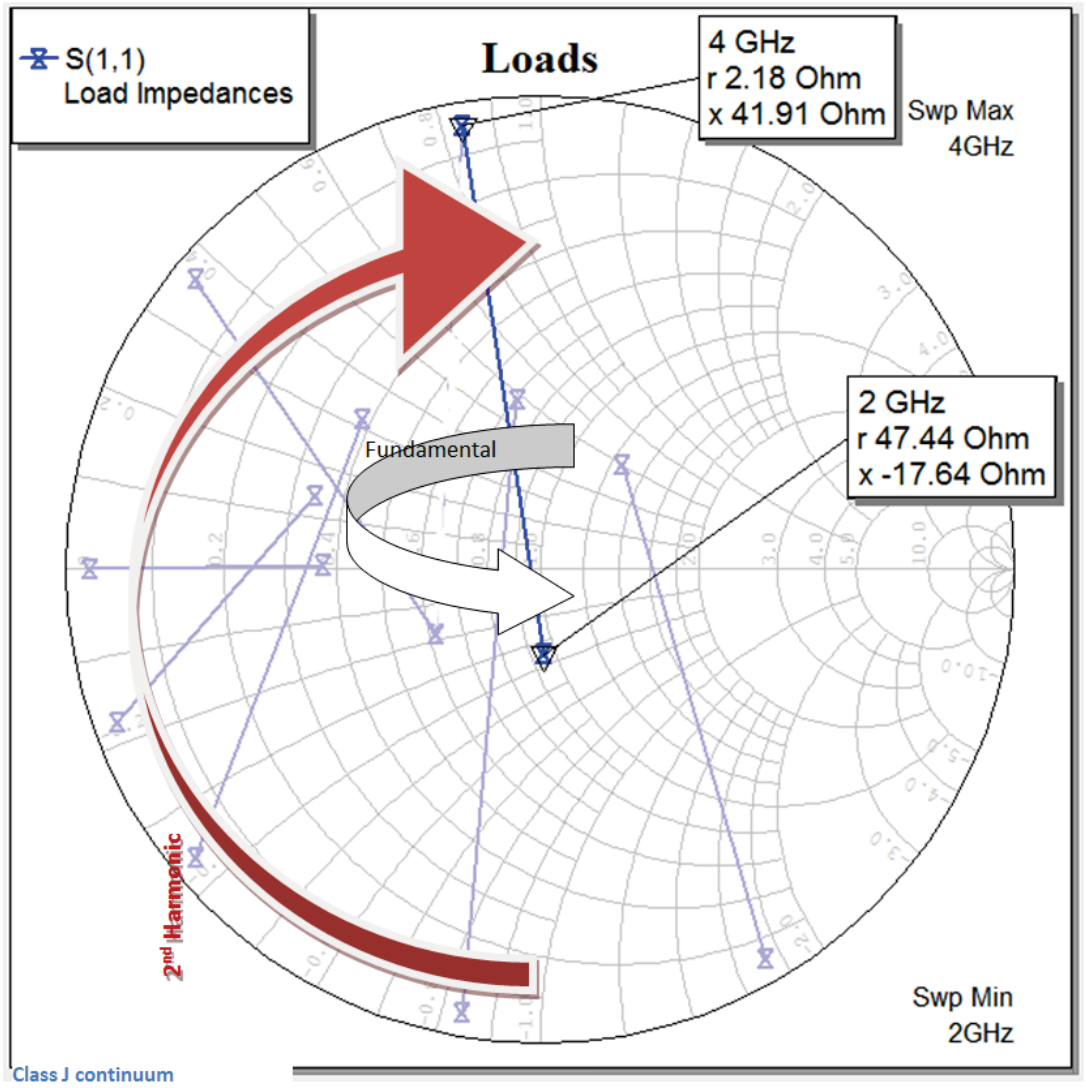


Figure 1-78, Fundamental and 2nd harmonic loads comprising the class J continuum

output power falls; this is due to voltage clipping. Plotting the output power, PAE and ratio of 2nd harmonic reactance, X2 to fundamental resistance R1, against the ratio of fundamental reactance, X1 to fundamental resistance R1, Figure 1-79, it can be seen that the power and PAE can be held relatively constant over a wide range of impedances provided the class J ratios are maintained.

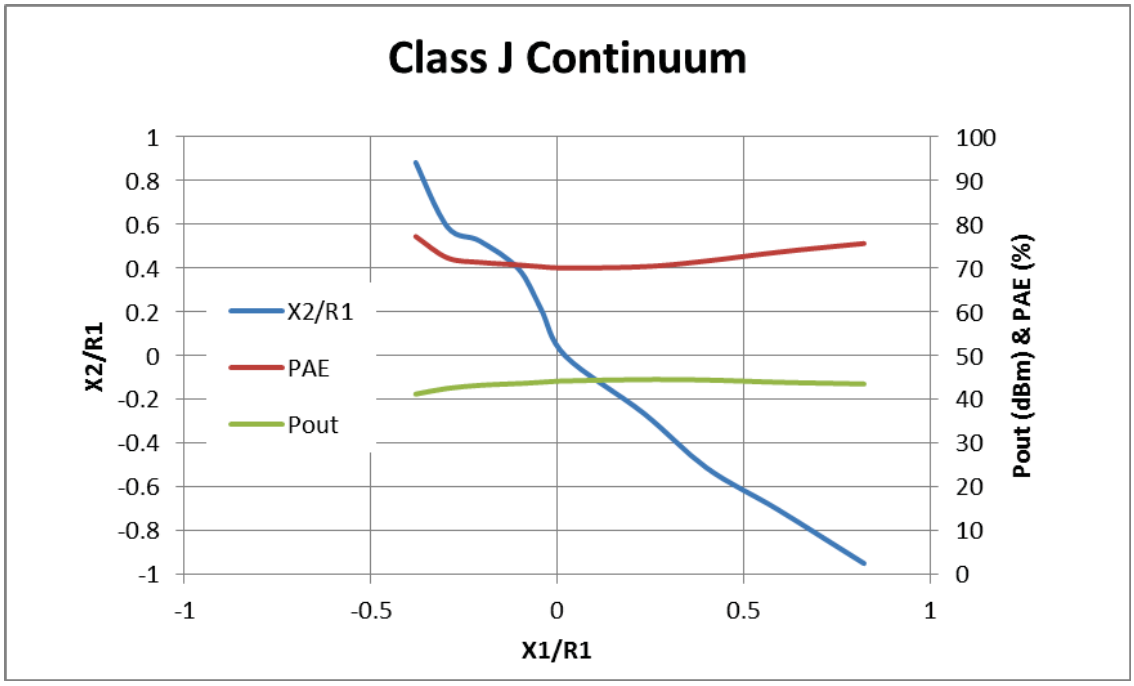


Figure 1-79, Relationship between 2nd harmonic and fundamental reactance ratios and associated performance

Repeating this process across the operating band at a number of specific frequencies, an impedance trajectory can determined for a wideband matching circuit achieving broad band power and PAE as was done in [37], Figure 1-80.

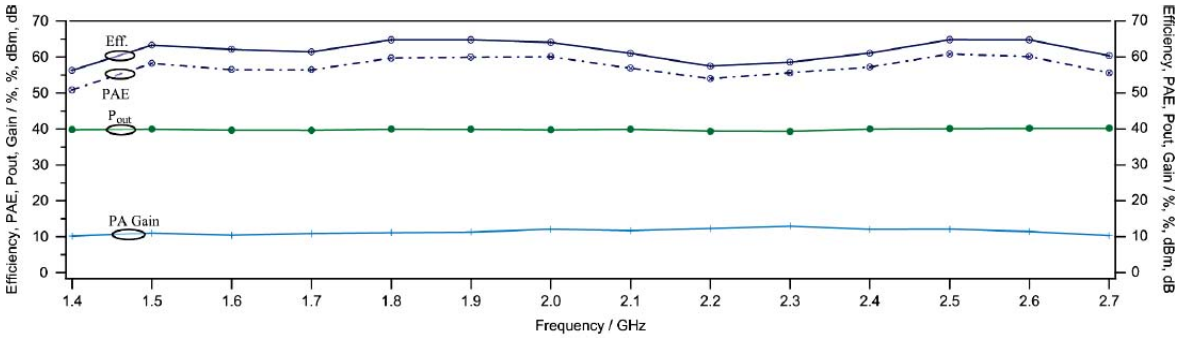


Figure 1-80, Measured broadband performance as reported by Wright et al. [37] using enhanced class J technique.

1.7 Summary

Although the most common practical approach to determining the optimum load match is to conduct a load pull on the device, the environment in which that device is operated, bias, drive level, harmonic impedances must be considered. This is a multi-dimensional problem, looking at a set of load pull contours or a swept power graph only give a very limited amount of information. On the other hand, the output waveforms

contain a vast amount of information. Amongst other things they can show is why a particular set of conditions result in a higher PAE, how close the device is to voltage breakdown, and particularly when an optimum solution is being approached.

It is important not to ignore the input match when considering the design of amplifier stages. Not only does the load impedance have a direct impact on the input impedance but also the optimum PAE is obtained at a specific drive level. PAE calculations in most software and measurement systems assume that the device is perfectly matched, i.e. the P_{IN} used in most applications of the calculation {1-1} is the power into the device, thus input matching circuits frequency response and associated losses need to take this into account.

In this chapter both MMIC and discrete hybrid designs have been considered. Although the main focus of the research has been on developing an improved design process for MMICs much of the work is equally valid in designs using discrete components. The cost implications for a 're-spin' of a design are significantly less and the time scales shorter in hybrid design, but to achieve the optimum performance the approach described here is entirely applicable. It is also often the case that for the development of new microwave semiconductor processes much of the initial testing will be done using discrete transistors. The waveform engineering approach will ensure that not only are the maximum powers and efficiencies achieved, but that there will be a better understanding of the operation of the device.

This chapter has hopefully set the scene for the research carried out in this work. The importance of observation of the waveforms in understanding the operation of microwave transistors, they are an overall summary of the effects of the multiple variables that produce the response from a stimulus. Basic theory has been shown to produce solutions that get designs in the 'right ball park' but that it is also important to consider exactly what conditions are being designed for, e.g. saturated output power, PAE, gain; as they will have different requirements.

1.8 Bibliography

[1] S. Cripps, “A Method for the Prediction of Load-Pull Power Contours in GaAs MESFETs,” in *International Microwave Symposium*, 1983.

[2] S. Cripps, RF Power Amplifiers for Wireless Communications, 2nd Ed., Artech House, 2006.

[3] J. Benedikt and P. Tasker, “High Power Time Domain Measurement System with Active Harmonic Load-Pull for High Efficiency Base Station Amplifier Design,” in *IEEE Trans MTT pp.2617-2624*, 2000.

[4] Z. Aboush, J. Lees, J. Benedikt and P. Tasker, “Active Harmonic Load-Pull System for Charaterisation of Highly Mismatched Transistors,” in *Microwave Symposium Digest, MTT-S International*, 2005.

[5] S. Azam, R. Jonsson and Q. Wahab, “Single-stage, high efficiency, 26-watt power amplifier using SiC LE-MESFET,” in *Asia-Pacific Microwave Conference*, 2006.

[6] F. Van Den Bogaart and T. De Hek, “First-pass design strategy for high-power amplifiers at X-band,” in *IEE Tutorail Colloquim on Design of RFIC's and MMIC's (Ref. No. 1997/391)*, 1997.

[7] Cree, “Thermal Optimization of GaN HEMT Transistor Power Amplifiers using new self-heating large-siganl model,” 2010. [Online]. Available: http://www.cree.com/~/media/Files/Cree/RF/Application%20Notes/AppNote%206.pdf . [Accessed 25 8 2012].

[8] I. Bahl, “0.7 - 2.7GHz 12W Power Amplifier MMIC developed using MLP technology,” *IEEE Transactions on Microwave Theory and Techniques,* vol. 55, pp. 222-229, February 2007.

[9] I. Bahl, “Low Loss Matching (LLM) design technique for power amplifiers,” *IEEE Microwave Magazine,* pp. 66-71, 2004.

[10] Y. Chung, C. Hang and e. al., “Effects of output harmoninc termination on PAE and output power of AlGaN/GaN HEMT power amplifier,” *IEEE Microwave and Wireless Components Letters,* vol. 12, no. 11, pp. 421-423, November 2002.

[11] C. Paillard and J. Wight, “Inside UWB Design: A Tutorial,” *Electronic Engineering Times,* March 2006.

[12] A. Grebennikov, RF and Microwave Power Amplifier Design, McGraw-Hill, 2004.

[13] E. Nesse, “Dynamic biasing saves power in long-haul microwave networks,” *Electronic Design,* August 2012.

[14] A. Harter, “LNA matching techniques for optimizing noise figure,” *RF Design,* pp. 20-30, February 2003.

[15] S. Marsh, “Design,” in *Practical MMIC Design*, London, Artech House, 2006, p. 90.

[16] H. Bode, Network Analysis and feedback amplifier design, New York: Van Nostrand, 1945.

[17] R. Fano, “Theoretical Limitations on Broadband Matching of Arbitrary Impedances,” *Journal of Franklin Institute,* vol. 249, pp. 57-84 and 139-154, January-February 1950.

[18] G. Matthaei, L. Young and E. Jones, Microwave Filters, Impedance-Matching Networks, and Coupling Structures, Dedham, MA: Artech House Books, 1980.

[19] H. Carlin, "A new approach to gain-bandwidth problems," *IEEE Transactions on Circuits and Systems,* vol. 24, no. 4, pp. 170-175, April 1977.

[20] P. Smith, Electronic Applications of the Smith Chart: In waveguide, circuit and component analysis., New York: McGraw-Hill, 1969.

[21] J. White, High Frequency Electronics - An introduction to RF and Microwave Engineering, New York: John Wiley & Sons, 2004, pp. 70-71.

[22] R. Rhea, "The Yin-Yang of Matching: Part 1- Basic Matching Concepts," *High Frequency Design,* pp. 16-25, March 2006.

[23] R. Rhea, "The Yin-Yang of Matching: Part 2 - practical matching techniques," *High Frequency Design,* pp. 28-40, April 2006.

[24] J. Walker, The Handbook of RF and Microwave Power Amplifiers, Cambridge: Cambridge University Press, 2011.

[25] F. Giannini, C. Paoloni and M. Ruggieri, "A very broadband matched termination utilizing non-grounded radial lines.," in *17th European Microwave Conference*, 1987.

[26] A. Pavio, "A network modeling and design method for a 2-18 GHz feedback amplifier," in *International Microwave Symposium Digest*, 1982.

[27] B. Virdee, A. Virdee and B. Banyamin, Broadband Microwave Amplifiers, Artech House, 2004, pp. 133-135.

[28] J. Pan and M. Russell, "MESFET amplifier with RF feedback gives high performance, low noise.," *Microwave Systems News,* June 1983.

[29] T. Narhi, "Use internal feedback to drop amplifier VSWR," *Microwaves & RF,* pp. 71-72, April 1986.

[30] Y. Wu, R. York, S. Keller, B. Keller and U. Mishra, "3-9 GHz GaN based microwave amplifiers with L-C-R broad-band matching.," *IEEE Microwave and Guided Wave Letters,* vol. 9, no. 8, pp. 314-316, 1999.

[31] K. Krishnamurthy, D. Green, R. Vetury, M. Poulton and J. Martin, "0.5-2.5GHz, 10W MMIC power amplifier in GaN HEMT technology," in *IEEE Compound Semiconductor Integrated Circuits Symposium Digest*, 2009.

[32] A. Grebennikov, "A 1.7 to 2.7 GHz transmission line GaN HEMT power amplifier for wireless applications," *Microwave Journal,* October 2010.

[33] E. Griffin, "Application of Loadline Simulation to Microwave High Power Amplifiers," *Microwave,* pp. 58-66, June 2000.

[34] G. Gonzalez, Microwave Transistor Amplifiers, 2nd edition, Prentice-Hall, 1997.

[35] J. Moon, J. Kim and B. Kim, "Investigation of a class-J power amplifier with a nonlinear Cout for optimised operation," *IEEE Tranactions on Microwave Theory & Techniques,* vol. 58, no. 11, pp. 2800-2811, November 2010.

[36] P. Godoy, S. Chung, T. Barton, D. Perreault and J. Dawson, "A highly efficient 1.95GHz, 18W asymmetric multilevel outphasing transmitter for wideband applications," in *International Microwave Symposium*, Baltimore, 2011.

[37] P. Wright, J. Lees, J. Benedikt, P. Tasker and S. Cripps, "A methodology for realizing high efficiency class J in a linear and broadband PA," *Transactions on Microwave Theory and Techniques,* vol. 57, no. 12, pp. 3196-3204, December 2009.

[38] M. Junghwan, K. Jungoon and K. Bumman, "Investigation of a class J power amplifier

with nonlinear Cout for Optimised Operation," *IEEE Transactions on Microwave Theory and Techniques,* vol. 58, no. 11, pp. 2800-2811, November 2010.

[39] G. Vendelin, A. Pavio and U. Rohde, Microwave circuit design using linear and nonlinear techniques, 2nd edition., New York: John Wiley & Sons, 2005.

2 Current Approaches to the Design of MMIC Power Amplifiers

"Any sufficiently advanced technology is indistinguishable from magic."

Third of Clarke's Three Laws, Arthur C. Clarke, 1917-2008.

2.1 Introduction

RF and Microwave design has persistently been described as 'black magic' by many onlookers and fellow engineers. Sadly this may in some case be as result of the approach of some of us when attempting to fix a problem. Moving your hand over a circuit and looking for the 'right spot' to affect an amplifiers performance is hardly scientific. To some extent this was due to a lack of appropriate design tools which lead to an approach which got you in the 'the right ballpark' and then required adding and removing components and tracks until the desired performance was achieved. In contrast digital designers knew what they expected to see on every pin and if it wasn't there it was very clear where the problem lay. There is obviously a considerable amount of science behind the manufacture and operation of microwave circuits; they do after all obey the laws of physics. This chapter seeks to explain the main theories behind the design and operation of RF and Microwave MMIC amplifier circuits and the current levels of performance achieved. As a starting point the technology to be used is reviewed.

The state of the art in microwave power amplification is constantly improving in some areas whilst others have remained static for some years now, as shown in Figure 2-1. Marked on this graph are three more recent benchmarks in Gallium Nitride (GaN) performance, [1] [2] [3], dating from 2006, 2010 and 2011 respectively. Although showing advances and improvements, there are no ground breaking leaps. This chart can however be misleading as it represents developments in single transistor Solid State Devices (SSDs) and Vacuum Electronic Devices (VEDs), whereas system developments such as phase arrays and spatial combining have enabled amplifiers based on solid state technology to reach further into the areas dominated for so long by the 'tube' industry. GaN has become an accepted technology below 4 GHz (although the economics of its application to the base station

market can be argued) and has made a particular impact in the very wide band (500-2500 MHz and 2.5-6GHz) Solid State Power Amplifier (SSPA) segment.

This has generally been based on 0.5 and 0.3 µm gate length technology, hence it is relatively simple to extrapolate such performance to 0.25 and 0.15µm devices and envisage the performance that should be achievable as such foundry processes become commercially available. In the radar pulsed amplifier sector, long being the last bastion of Silicon Bipolar transistors, GaN based devices are routinely breaking power and bandwidth records, [4].

The upper frequency range achievable by solid state devices is primarily dependent upon the charge carrier velocity in the semiconductor, and secondly on the size of the physical structures that can be fabricated. Fortunately for us in the Radio Frequency (RF) industry technology on this latter front is pushed ahead by processor and memory requirements driving structure sizes to 20nm processes, which in itself would permit our current transistor topologies to operate into the 10^{11} Hz range, as shown in [2]. Referring back to the prime characteristic of charge carrier velocity, at high electric fields most semiconductors achieve a saturated charge speed, v_s of ~10 cm/s. Gallium Indium Arsenide High Electron Mobility Transistors (GaInAs HEMTs) have reported performance up to 300GHz, [5], however the RF output power is low, primarily limited by voltage breakdown. Another factor is thermal management, most semiconductors are poor thermal conductors;

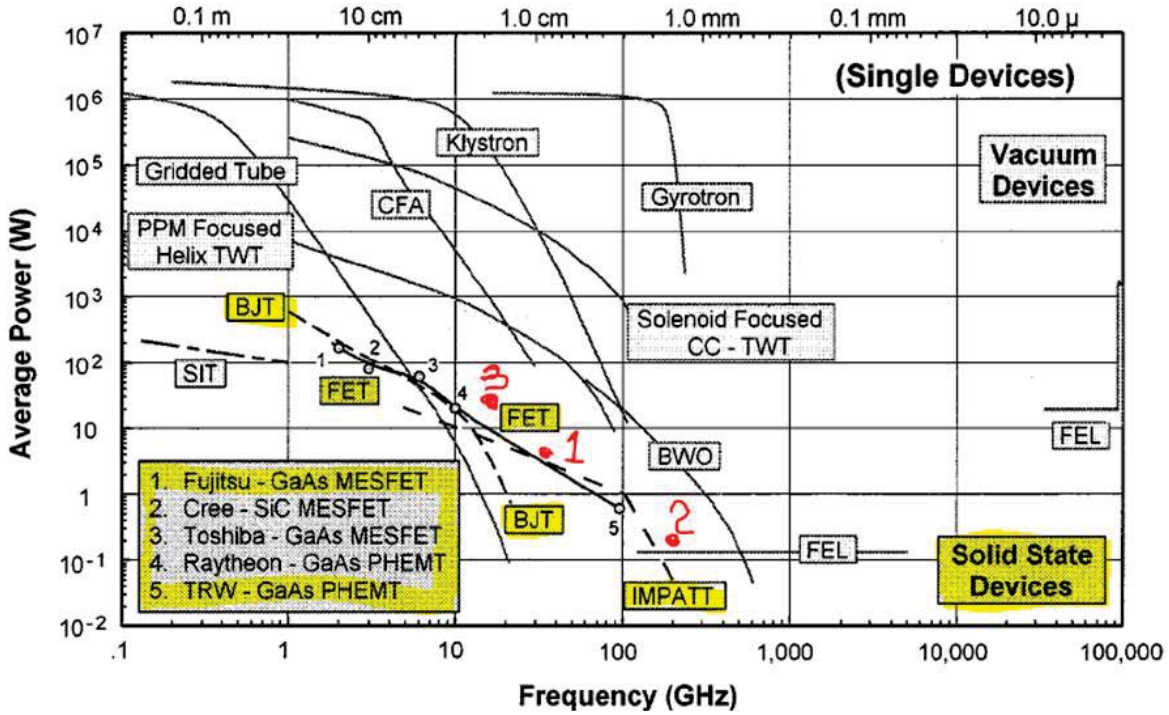


Figure 2-1, Graph showing average RF Output power vs Frequency for single SSDs and VEDs [38] with some recent additions. Numbers in red refer to references [1-3].

the exception being Silicon Carbide which is limited in frequency range by not being able to be fabricated in a way so as to produce HEMT structures, hence its focus on lower frequency high voltage high power switching applications. Traditional semiconductor materials (i.e. Silicon, Silicon Germanium, GaAs, Indium Phosphide, etc.) have relatively low barrier energy which becomes increasingly 'leaky' as temperatures rise and hence their junction temperatures are limited. In order to produce high power at microwave frequencies it is required to have high currents, this in turn necessitates large cross-sectional areas in the semi-conductor, which in turn gives rise to low input impedances which make matching difficult. Thus one of the main attractions in moving to Wide Band-Gap (WBG) devices such as GaN and SiC has been the increased bandgap and hence higher operating temperatures.

The advantages of fabricating devices from WBG materials is clear from Table 2-1, Diamond would be an ideal material except that the difficulty in adding dopants has made the production of commercial RF devices still an improbability in the near term, if and when it happens there will be a significant move in the curves of Figure 2-1. These properties have been translated into various Figures of Merit (FoMs) which try and describe their suitability to types of RF applications.

Material	Bandgap (eV)	Dielectric Constant (ε_r)	Critical Breakdown Field (MV/cm)	Thermal Conductance (W/K-cm)
Si	1.1	11.9	0.3	1.5
GaAs	1.4	12.5	0.4	0.5
InP	1.3	12.4	0.4	0.7
SiC (4H)	3.2	10.0	3.5	4.9
GaN	3.4	9.5	3.4	1.5
Diamond	5.6	5.5	5	20-30

Table 2-1, Semiconductor Material Properties [6].

Johnson Figure of Merit: based on the breakdown voltage and the saturated electron drift velocity, gives a value for the suitability for high frequency operation.

Silicon	1
GaAs	12
SiC	400
GaN	790

This figure reflects suitability with regards to intrinsic device properties. Despite Silicon's low ranking, Intel has reported a transmitter operating at 20mW on silicon at 65 GHz. This is a reflection on the maturity of silicon processing.

Balgia Figure of Merit: based on the dielectric constant, electron mobility and critical electric field, gives a value for the suitability for high power handling operation.

Silicon	1
GaAs	18
SiC	35
GaN	100

One could argue the exact values of the various FoMs; however the clear message is that WBG materials are on the face of it clearly advantageous for high frequency high power applications. Other benefits include their inherent radiation hardness, making them very suitable for space based applications, again due to the high energy gaps.

The importance of dielectric constant, ε_r, is often overlooked. GaN has about a 20% lower ε_r than GaAs which results in lower parasitics capacitances for the same size devices and hence higher frequency performance or alternatively the structures can be 20% bigger for the same frequency (associated parasitics) and hence have a larger area which enables higher currents and powers to be supported. In impedance terms the lower capacitances makes matching easier.

Early WBG devices were produced on SiC and a look at the high thermal conductivity shows one of the reasons. Thermal conductance is extremely important as the high power levels generated are in very small regions (under gates which may by as small as 0.15x50µm) and both performance and reliability are inversely proportional to temperature. Looking specifically at GaN and SiC they both offer the benefits of the high band gap intrinsic in their group, which offers the possibility of operating at high voltages. SiC has an advantage in terms of thermal conductivity, however cannot support as high a current density. Another disadvantage for SiC is that it cannot be organised in a High Electron Mobility Transistor (HEMT) structure, which is currently one of the best for high frequency operation, (SiC devices use a Metal Schottky Field Effect Transistor (MESFET) structure). Also because of its physical nature it is more difficult to add dopants to SiC than GaN, which benefits from processing which is very similar to that of GaAs, which is now a mature process.

The higher the Critical Breakdown Field the higher the voltages that can be supported within the device structure. Higher voltages allow lower currents (for the same power level) which reduce heating effects or alternatively for the same current a higher resistance, which improves impedance matching (in high power applications). The higher breakdown voltages also make the devices theoretically more robust and less susceptible to damage from high RF signals, hence the possibility of producing significantly higher dynamic range low noise amplifiers (LNAs) without the need for limiters, or at least simpler limiters with less insertion loss.

GaN not being as good a thermal conductor as SiC, has suffered from the lack of a good substrate or backing material. This is not an issue for silicon as the process is relatively easy now, so that large and 'thick' wafers can be produced. For GaN, producing the wafer by the conventional means of drawing a large ingot and slicing it has not proved economically viable. The alternative, but still an expensive process, is to grow the GaN on a dissimilar substrate (heteroepitaxy). Early GaN was produced on a sapphire substrate, but this proved very difficult to manufacture in any volume. Currently there are two schools of thought, those that recommend GaN on silicon (Nitronex), and those that have opted for GaN on SiC, (Cree, SEI, Triquint, RFMD, UMS). Whilst there may still be an application for GaN on Silicon,

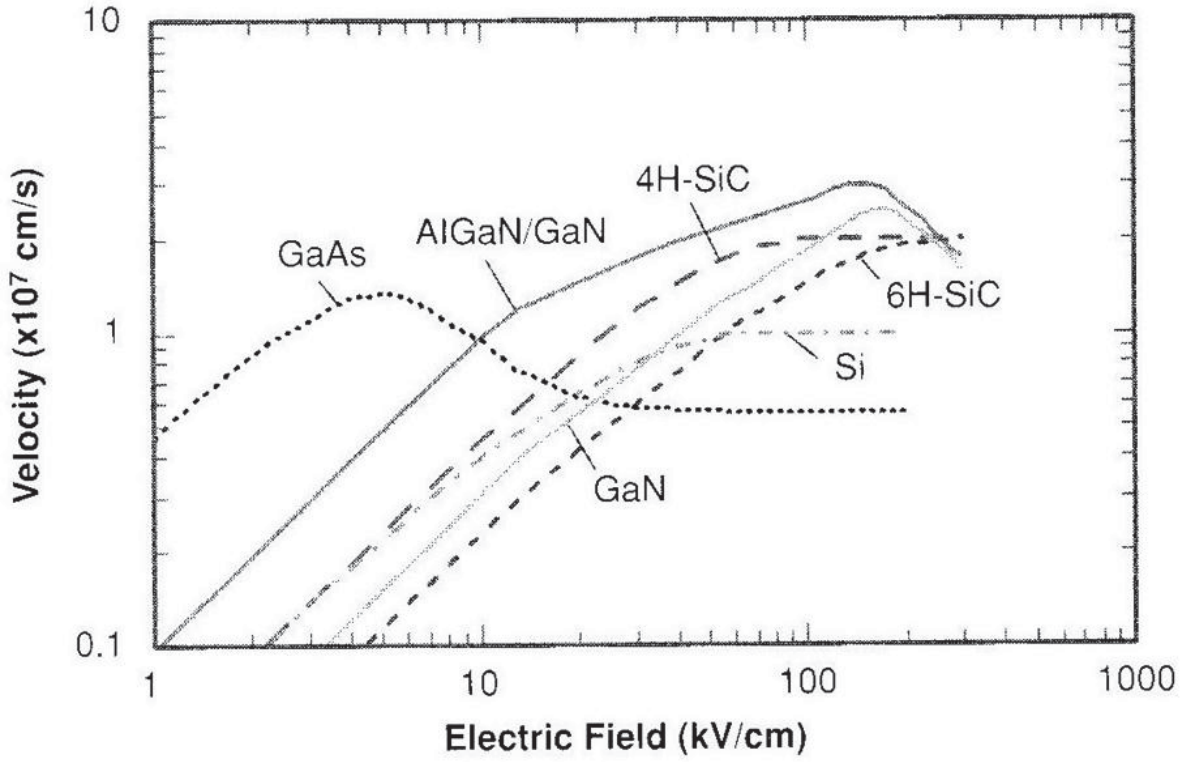


Figure 2-2, Electron Velocity versus electric field transport characteristics of various n-type semiconductors [6].

high dynamic range Low Noise Amplifiers for example, the relatively poor thermal conductivity of Silicon means that the wafers have to be lapped very thin to reduce the thermal barrier. The attraction of silicon is that very large wafers (12") can be processed, greatly reducing the costs ($/mm^2) as compared to the 3 and 4" wafers of SiC, however the process of reducing the Silicon thickness to 100µm or so over large areas is not cheap and reduces yield, hence the predominance of SiC as a substrate in the RF power device market. It is worth noting that Cree dominates the world supply of SiC substrates and although Summitomo/SEI can produce 4" wafers they still supply <10% of wafers worldwide. It can also be argued that in the earliest days of GaN devices reliability was a significant issue and every effort has been made to keep the temperature low in order to improve reliability, hence the predominance of GaN on SiC. Perhaps with improved processing and thermal materials in the future this will not be such an issue and GaN on Si will be a competitive technology.

Current is the movement of charge and is expressed as the product of charge density and transport velocity. Although initially linear, the relationship between electric field, E, and velocity, *v*, can be complex as is the case for GaAs as shown in Figure 2-2. The initial slope at lower voltages is referred to as the charge carrier mobility μ_n (cm^2/V s), the saturated velocity, v_S (cm/s) is that when the velocity levels out at high voltages, which was fine when applied to Silicon but more difficult to define for other materials. The very high value of μ_n (~5000 cm^2/V s) of GaAs shows why it was extremely suitable for high frequency devices, and shows the motivation behind the development of the more exotically doped materials such as AlGaN/GaN to try and improve the relatively poor mobility of basic GaN materials at lower voltages. The WBG materials however are able to support much higher saturated velocities as they can sustain the higher electric fields. A consequence of the low μ_n is that for GaN based HEMTs, the 'knee' voltage, V_k of the DC-IV curves is higher than that of GaAs HEMTs, by the order of about 5 times. The critical factor in device operation however is ratio of V_{max} to V_k. The higher operating voltages compensate for the higher knee.

In summary, for RF power amplification the higher the current the better, and to achieve high currents we require high electron velocity. The higher the carrier mobility, the lower the knee voltage; but also generally the higher the critical breakdown voltage, the higher the saturated velocity. These factors all show the attractiveness of WBG devices and

because of the ability to create HEMT structures in GaN it is currently the most suitable device structure (Diamond is probably 10-20 years from commercial exploitation).

2.2 Device Construction

Most of the early problems with GaN RF device implementation were due to two main issues:

a) Substrate consistency

b) High voltage charge storage

Both resulted in phenomena grouped under the heading of 'traps'. Faults or imperfections within the crystalline structure allowed for carriers to be 'captured', similarly the high voltages caused a build-up of surface charge on the top of the device. These would cause changes to the behaviour of the device in phenomena such as "knee-walk out", bias changes and memory effects. These have now largely been addresses by process improvements and modifications to the standard HEMT structure.

As mentioned earlier the majority of high power GaN devices are grown epitaxially on high resistivity SiC substrates, as shown in Figure 2-3. There is usually a GaN buffer layer to account for the lattice mismatch between the SiC and the GaN. The 2 Dimensional Electron Gas (2DEG) feature of HEMTs is created by growing an AlGaN layer on top of the GaN; this creates a 'notch' in the energy band structure which fills with carriers taken from the heavily doped AlGaN layer, as described in Figure 2-4. This thin sheet of charge supports very high current flow (the drain-source current flows in the 'X' direction on the figure). The HEMT structure itself (whether in GaAs or GaN) exhibits higher intrinsic transconductance, g_m, than the conventional MESFET structure as the electrons flowing in the 2DEG layer have a higher mobility due to their separation from the doping impurities and lattice vibrations of the normal conduction layer, hence the shift to the left of the AlGaN/GaN curve in Figure 2-2.

An additional advantage of fabricating GaN HEMTs in this way is that because GaN is a strongly polar material, the strain resulting from growing the lattice mismatched AlGaN on GaN induces a piezo-electric charge. This supplies additional electrons to the HEMT channel, up to 4 times more than in the equivalent GaAs device (GaN ~1 x 10^{13} n/cm^2).

The transconductance is described by {2-1}:

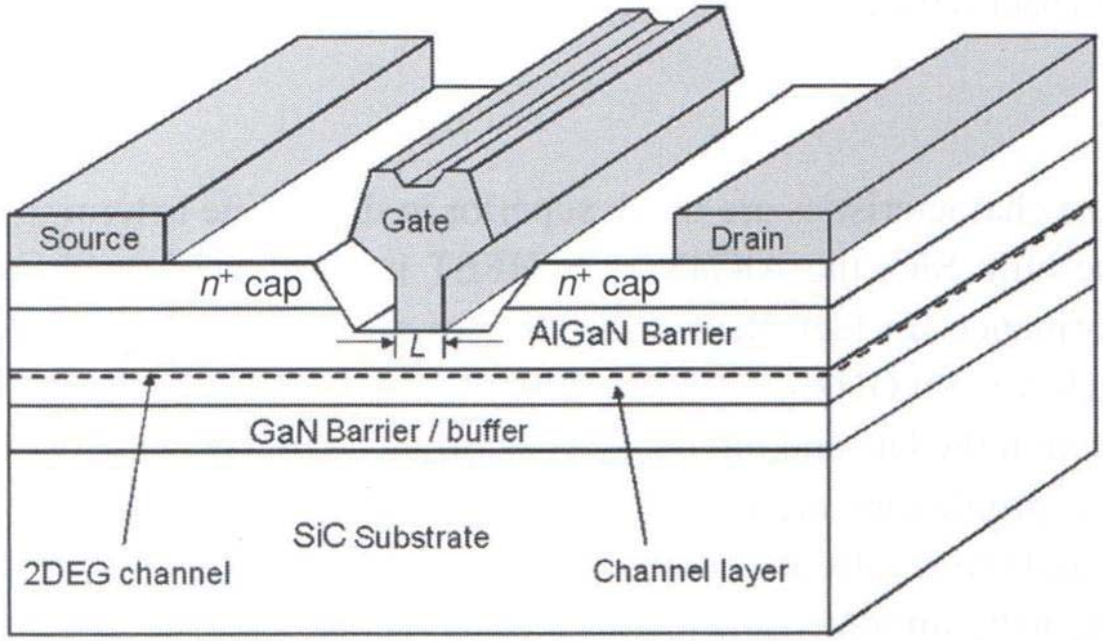


Figure 2-3, Device structure for AlGaN/GaN HEMT [6].

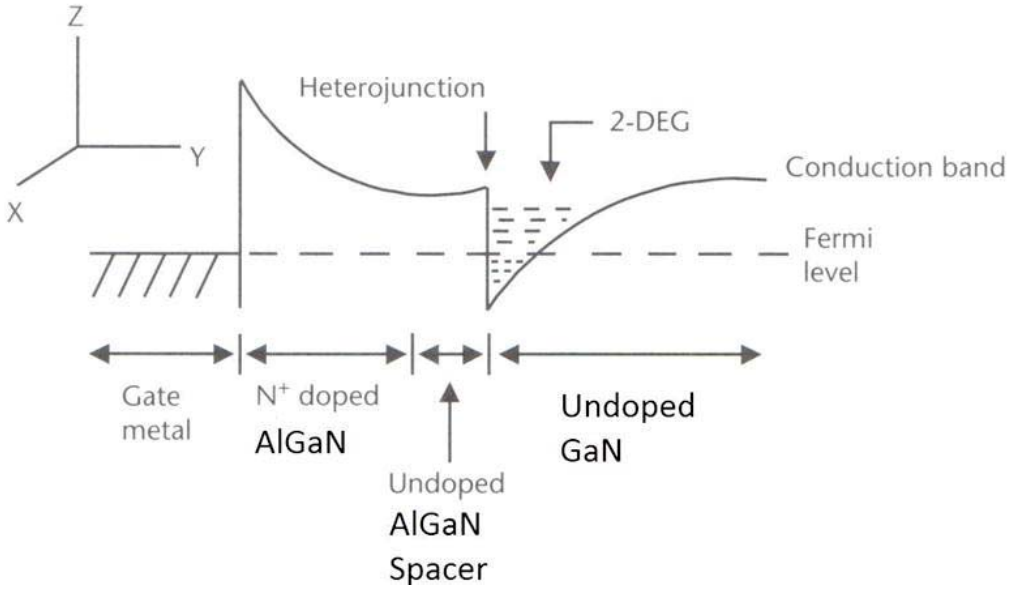


Figure 2-4, HEMT conduction band discontinuity.

$$g_m = \frac{v_s w\, \varepsilon_r}{(d + \Delta d)} \qquad \{2\text{-}1\}$$

Where *w* is the gate finger width and (*d +Δd*) represents the total thickness of the AlGaN layer, [7].

Two other physical differences are used in the design of GaN HEMT transistors and they relate to the high internal voltages. These voltages can result in a build-up of charge on the surface of the device between the source and the drain. This can effectively extend the gate region (virtual gate), which lowers output power and the maximum operating frequency. The other problem with these high voltages is that particularly where they are

concentrated around the gate they can cause reliability issues. The surface charge can exhibit a phenomenon known as current collapse, where the current in the linear part of the DC-IV curve is severely reduced. A passivation layer (SiN) is added to the device surface to reduce this charge and also the shape of the gate is modified, initially this was referred to as a Field Plate (FP), examples of which are shown in Figure 2-5, but most recent devices use the 'T' gate shape as shown in Figure 2-3, in combination with a 'top-hat' Source FP and passivation layers, Figure 2-6. These FPs however bring with them additional capacitance, which reduces gain and output power, as shown in Figure 2-7.

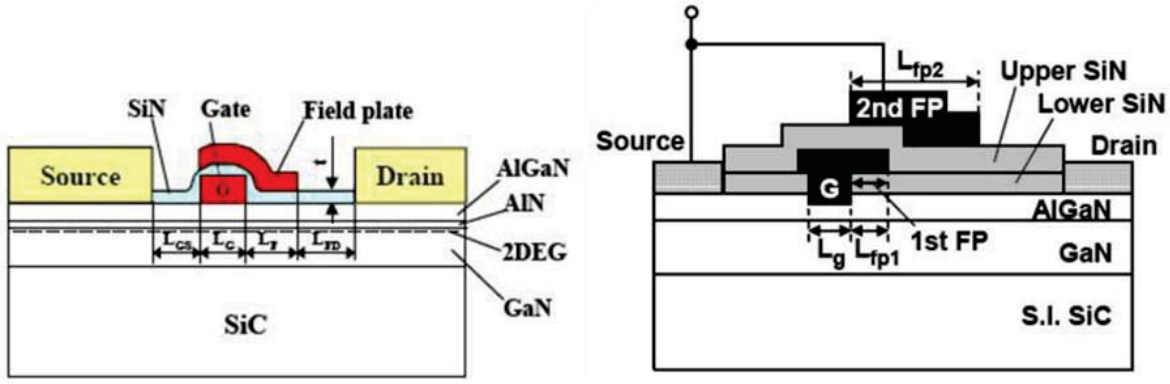


Figure 2-5, Field plate constructions to reduce surface charge and peak voltage breakdown.

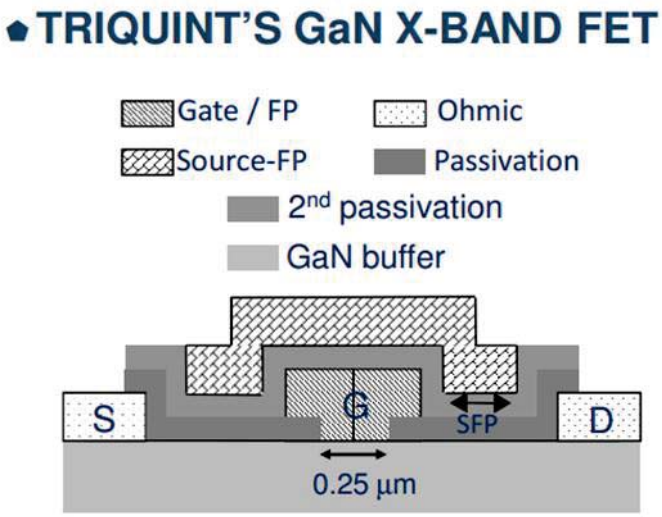


Figure 2-6, TriQuint Field Plate and passivation Construction.

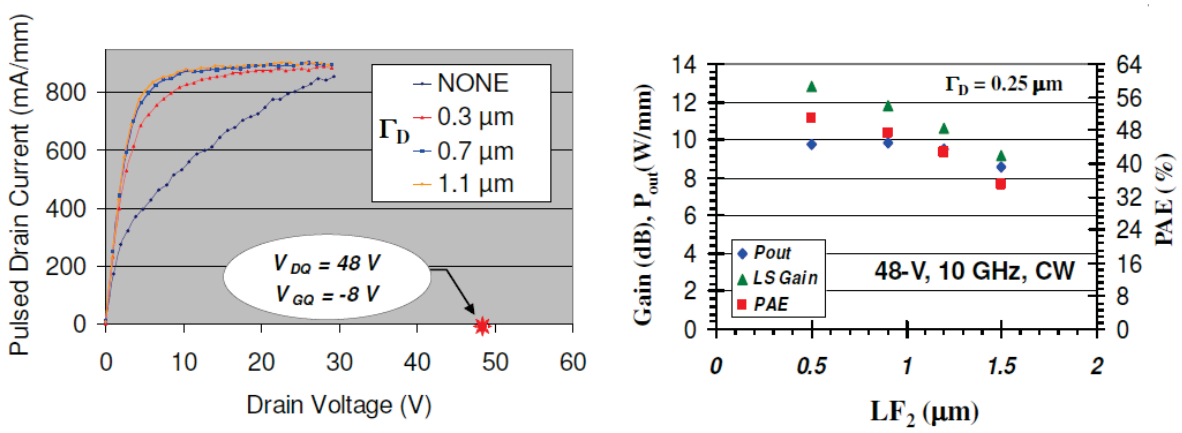


Figure 2-7, Effect of Lengthening Gate FP on Knee region and of lengthening source FP on device characteristics [39]

In summary GaN based devices offer the most suitable materials currently available for operation at high power and high frequency. They offer the increased current handling capacity, higher temperature operation, lower parasitic capacitance (for the same structure size), higher breakdown voltages, and higher output impedance (as a direct result of higher voltage operation).

2.3 Technology Selection

Having seen the evidence that GaN is the most suitable material for producing high power microwave amplifiers we now need to review how the material advantages can best be used to gain the maximum advantage. One way of doing this is to compare the performance of similar device size GaAs and GaN transistors. The devices that will be used for this analysis are the Triquint TGA2023-01 and a RFMD (UK) FD30 process GaAs DpHEMT which is one of the highest power, high frequency, GaAs processes. The comparisons made are from the manufacturers supplied nonlinear models which experience has shown have tended to be slightly optimistic in terms of output power; however the point is to examine the major differences as opposed to making detailed absolute comparisons. The main characteristics of each device are summarised in Table 2-2, and images of the devices themselves are shown in Figure 2-8 (photograph is of a 0.75mm periphery device) and Figure 2-9. The devices are biased for their recommended maximum power which in the case of the GaAs device is 250mA and the GaN device 125mA. The GaN device is biased in

class A/B, as recommended by the manufacturer, which has an effect on some of the performance characteristics that shall be observed.

Type	Gate Length	No. of Fingers	Periphery	Drain Voltage	Quiescent Current	Class
GaAs	0.3µm	10	1.25mm	10	250mA	A
GaN	0.25 µm	10	1.25mm	28	125mA	A/B

Table 2-2, Comparison between GaAs and GaN device characteristics.

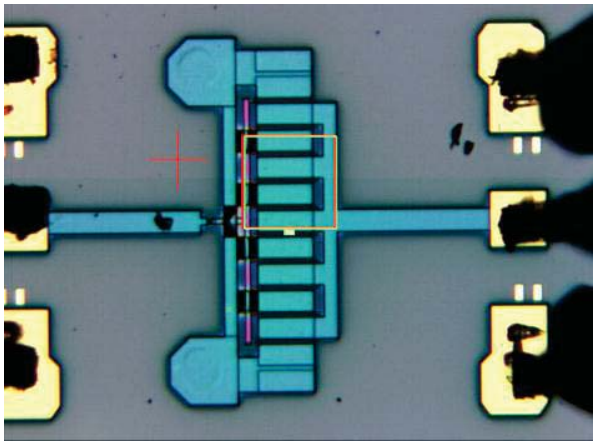


Figure 2-8, GaAs 0.3µm 0.75mm DpHEMT

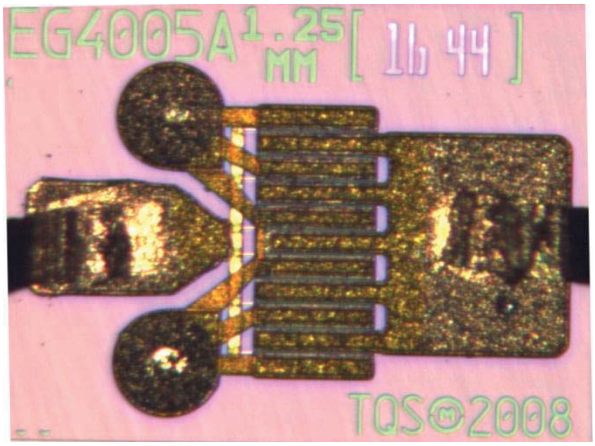


Figure 2-9, GaN 0.25µm 1.25mm HEMT

Firstly, simulating the DC-IV curves for each of the devices the difference in both the knee voltage and the saturated drain current slope, Figure 2-10 and Figure 2-11 is observed. It is worth noting the slope of the GaN device differs from other manufacturers (such as Cree) where it tends to have a negative slope. Similarly the slope of gate current flow appears to be at odds with what would be expected as shown in Figure 2-12 and Figure 2-13. The pinch off voltage required for these GaN devices is of the order of -4.2 volts whereas for the DpHEMT it is less than even the standard GaAs MESFET at >-1.5v. The different approaches to DC-IV measurements will be discussed in chapter 4, and these differences will have an impact on the base data to which the nonlinear models are optimised.

The devices have very similar maximum frequency performance, whilst below 15GHz the GaN device has superior small signal gain as shown in Figure 2-14. Of particular note is that by one measure of device stability, (Rollets K Factor) the GaN device is unconditionally stable between 6 and 15 GHz. Perhaps the clearest way to understand why this should be so is to look at the device input match, Figure 2-15 in conjunction with the reverse isolation,

Figure 2-16. The higher isolation and better match make the device less susceptible to oscillation over this frequency range.

The measure of gain GMax (Maximum Available Gain) and MSG (Maximum Stable Gain) are subtly different when the amplifier is unconditionally stable, i.e. K>1, {2-2}, but identical when K<1, {2-3}, i.e. conditionally stable.

$$G_{max} = \left|\frac{S_{21}}{S_{12}}\right|\left(K - \sqrt{K^2 - 1}\right) \qquad \{2\text{-}2\}$$

Whereas for K < 1

$$G_{max} = \left|\frac{S_{21}}{S_{12}}\right| = MSG \qquad \{2\text{-}3\}$$

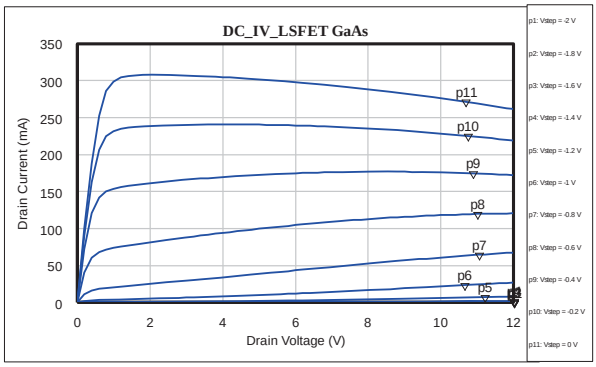


Figure 2-10, GaAs DC-IV Curves.

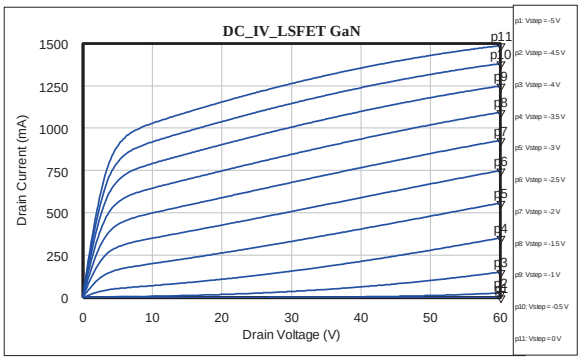


Figure 2-11, GaN DC-IV Curves.

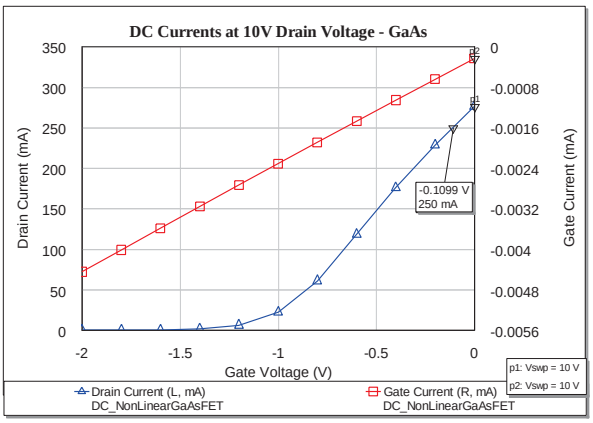


Figure 2-12, GaAs gate and Drain Currents.

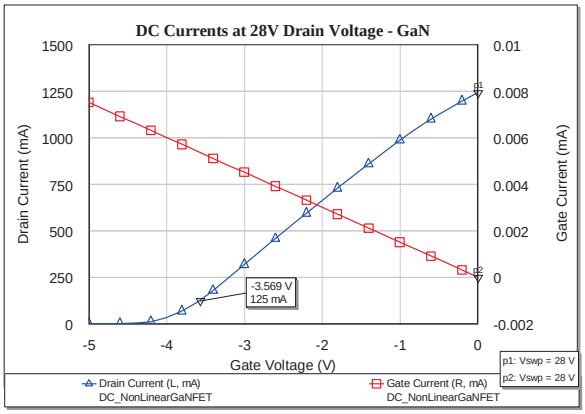


Figure 2-13, GaN Gate and Drain Currents.

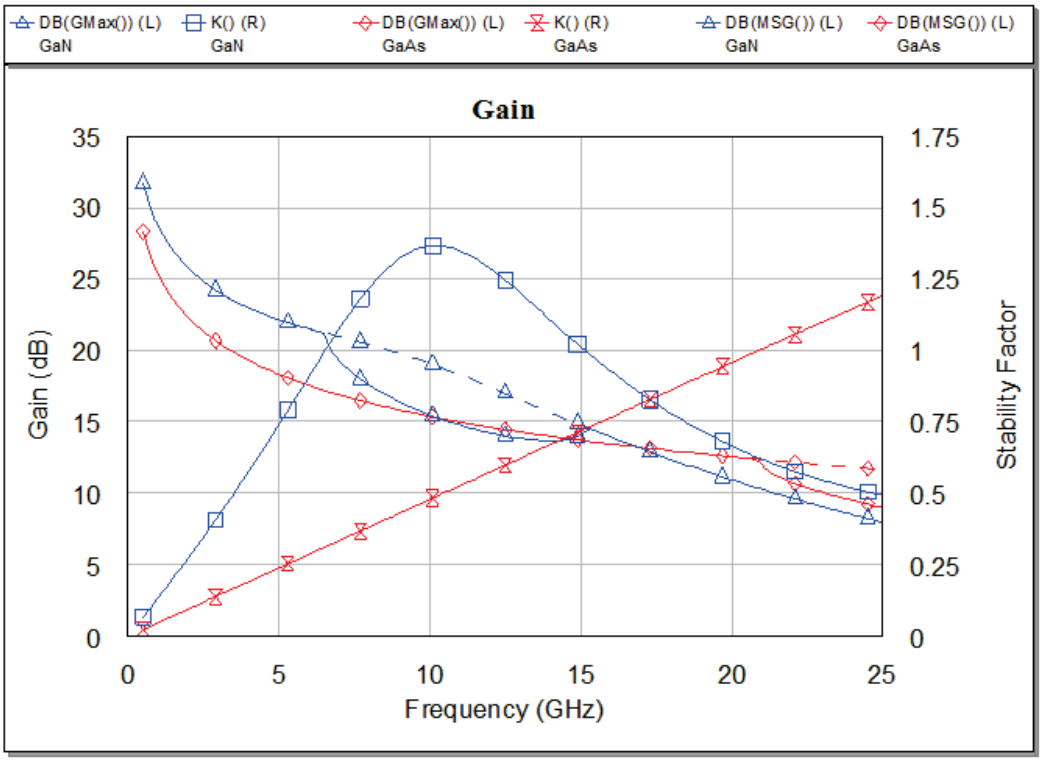


Figure 2-14, Maximum Available Gain, Maximum Stable Gain and Rollet's Stability Factor (K) for GaN and GaAs devices.

The higher gain of the GaAs process compared to the GaN above ~18GHz is perhaps surprising if the slightly longer gate length is considered (0.3 compared to 0.25µm). The increased roll-off in gain for GaN is due to increased parasitic capacitance due to the Field Plate.

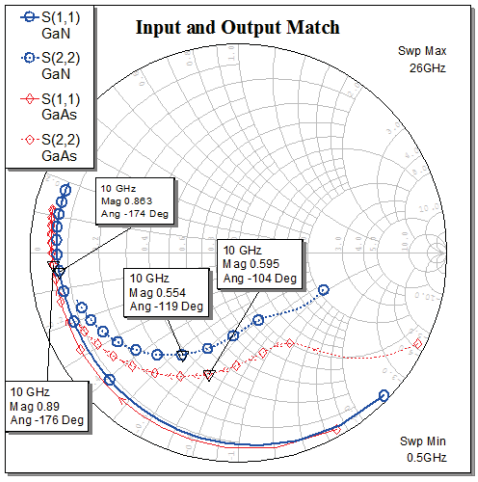


Figure 2-15, Port Match.

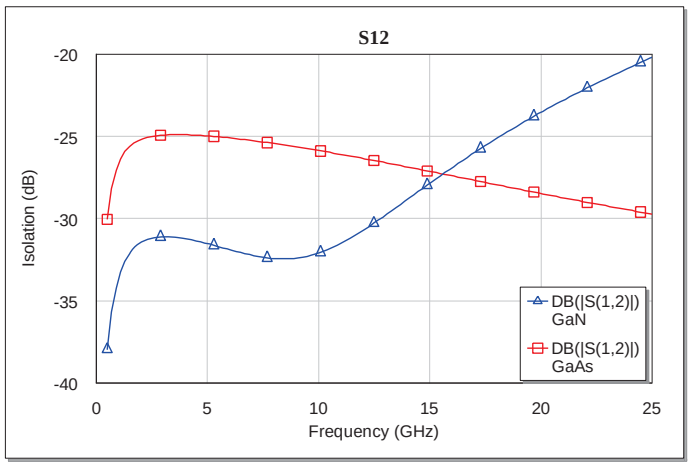


Figure 2-16, Reverse isolation.

The small signal output port match is not that required for either optimum Power Added Efficiency (PAE) or maximum output power as described by Cripps [8]. One of the simplest ways to determine these load impedances (within CAD anyway!) is to use the load

pull facilities. The user establishes a grid of impedances in the output plane at a particular drive level and device is measured at each grid point. The resulting performance contours are plotted and the optimum loads found. This was conducted at 10GHz on the devices under consideration and the results shown in Figure 2-17 and Figure 2-18.

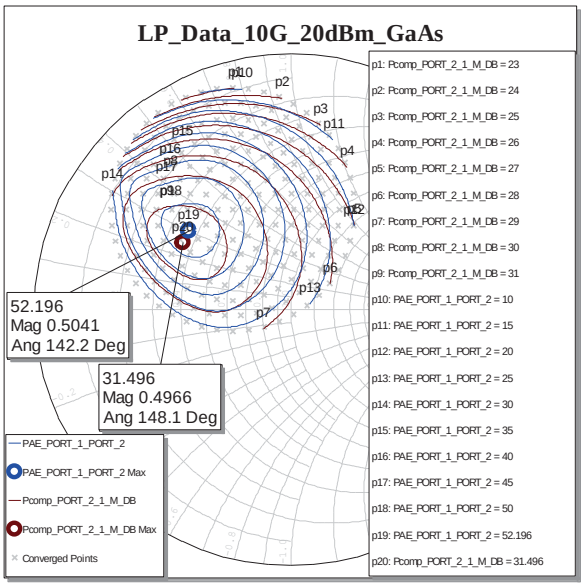


Figure 2-17, GaAs Load Pull at 10GHz 20dBm drive.

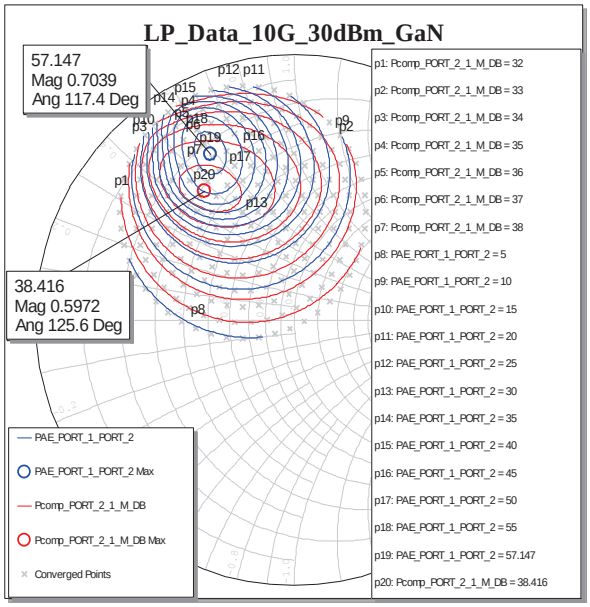


Figure 2-18, GaN Load Pull at 10GHz 30dBm drive.

The key points to take from this analysis are that:

a. The predicted maximum power from the GaAs device is 31.5dBm (1.4W).
b. The predicted maximum power from the GaN device is 38.4dBm (6.9W).
c. Hence for GaAs ~1.1W/mm and GaN 5.5W/mm.
d. The equivalent output circuit for the GaAs device is 31Ω and 0.43pF (PAE opt).
e. The equivalent output circuit for the GaN device is 85Ω and 0.47pF (PAE opt).
f. The equivalent output circuit for the GaAs device is 27Ω and 0.42pF (Pout opt).
g. The equivalent output circuit for the GaN device is 51Ω and 0.46pF (Pout opt).

Whilst both results may be optimistic, a result of ~1W/mm and ~4W/mm are commonly reported for these technologies. Also, nearly 5x more power is available from the GaN device (of approximately the same physical size) and an easier impedance to match (reactive elements are similar). The drain capacitances are of similar size which one would expect for devices of the same periphery, with the GaN being slightly higher, opposite to what you would expect from the lower dielectric constant (9.5 as opposed to 12.5), and possibly due to the field plate structure; both devices being fabricated on 100µm substrates.

As explained in chapter 1 the dynamic RF load lines are more meaningful when they are de-embedded to the current generator plane, however at this point the device plane is adequate for understanding the extent of the voltage and current swings as shown in Figure 2-19 and Figure 2-20. There is little change in the GaAs device between the two loads as they are very close in impedance. In the GaN case the expected anti clockwise rotation for higher efficiency is observed.

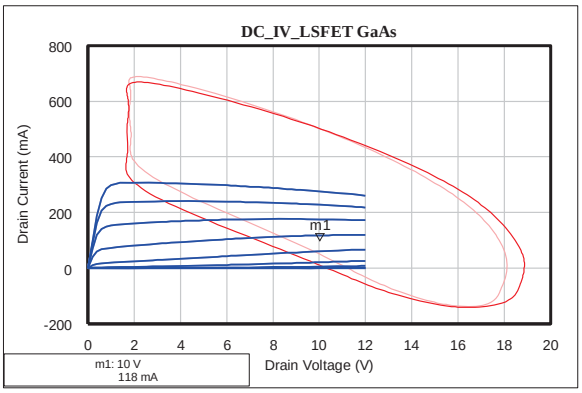


Figure 2-19, GaAs dynamic RF load line: Solid red for Opt. PAE load, faded for Opt. Pout load.

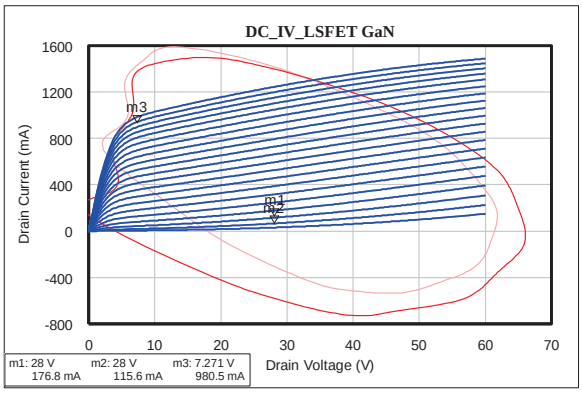


Figure 2-20, GaN dynamic RF load line: Solid red for Opt. PAE load, faded for Opt. Pout load.

A parameter sometimes overlooked is how the input reflection coefficient (Gamma, Γ) changes with drive level. From Figure 2-21, it can be seen that there is considerable change in Γ as the GaAs device is driven into saturation, whilst the GaN is considerably more 'stable'. Although the GaAs device input is more difficult to match at lower powers the saturated input impedances are very similar, (but note for a 5x more powerful device in the

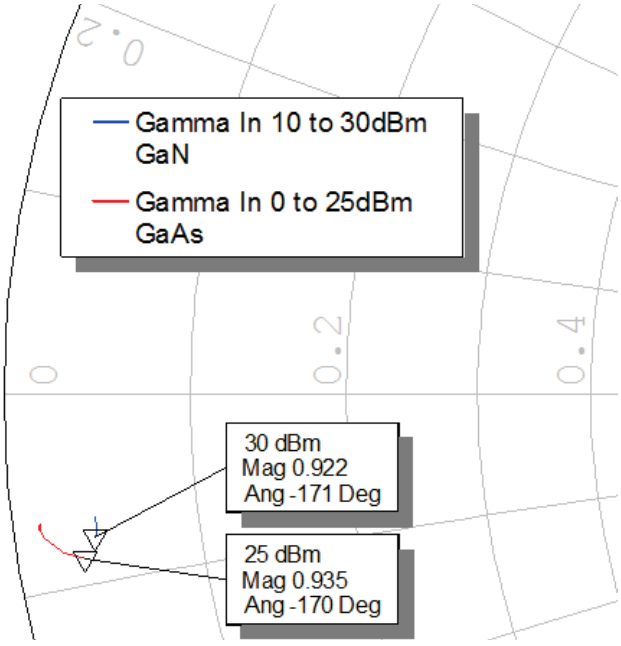


Figure 2-21, Input Gamma variation with drive power.

GaN case).

Looking at the effects of the output load impedances, we can see that the optimum PAE for GaAs occurs at a relatively high compression and hence reduced gain, Figure 2-22. The actual efficiency levels achieved are only marginally higher (~2.5%) with GaN, although it should be emphasised that these are simulated numbers and we do not know whether we are outside the envelope of the model[1]. Comparing Figure 2-22 and Figure 2-23 it is very noteworthy that whilst for GaAs the optimum output power load does not produce any more output power than the optimum PAE load (it can be seen that they are very close in the impedance plane anyway), in the GaN case almost 2 dB more power is available. Figure 2-24 shows the difference in gain shape between the two technologies. It should be noted however that this is also partly due to the biasing differences of the two devices, (class A vs. A/B), the GaAs device has a small amount of gain expansion as it approaches the compression point (which may in part be due to the reducing input Γ and to a lowering of the junction temperature as the device becomes more efficient with more power going into the RF signal). The GaN device by contrast has a less steep compression and PAE curve.

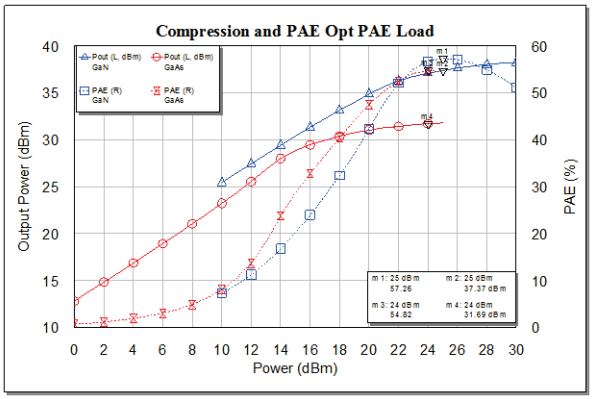


Figure 2-22, Optimum PAE Load, output power and PAE curves.

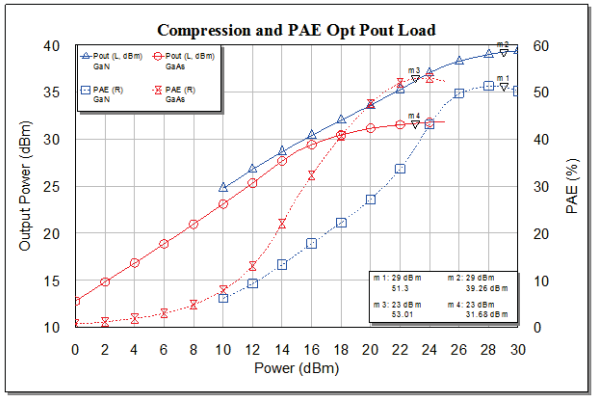


Figure 2-23, Optimum Pout Load, output power and PAE curves.

It has been observed that the GaN device can deliver over 6dB more power from the basically the same physical footprint. However we have also seen that the PAE of the two devices at the peak is very similar. Thus the dissipated power in this region (approximately 0.4x0.5mm) is over 4 times higher. At this point the thermal analysis gets complicated (which is where device models with an externally accessible junction temperature monitor

[1] It is also worth noting that PAE is calculated assuming that the input is perfectly matched. This makes a slightly unfair comparison as it does not distinguish between a device with a very high input Γ and those with lower values. To overcome this, an isolator could be added to the input of the circuit simulation.

such as Cree and Freescale are a big advantage) as the GaAs device is grown on a semi-insulating (electrically) GaAs substrate with an overall thickness of 100µm, whilst the GaN device is grown on a SiC substrate also with a final thickness of 100µm. As shown in Table 2-1 SiC has a thermal conductivity over 7 times greater than GaN, so all seems well, however there are two further aspects to consider:

1. Temperature rise due to the epitaxial layer
2. Getting the heat out from under the device.

These will be covered in greater detail later, but suffice it to say at this point the higher operating junction temperature of GaN actually turns out to be a necessity for satisfactory operation of such devices as in practice the devices end up running at higher temperatures due to the higher power densities per mm^2.

In a similar way the current generation of GaN devices do not offer a significant enough decrease in drain current to make huge differences to amplifier design; improvements yes, but it will be the availability of >48V operating devices which will make the significant impact. Typical metallisation current handling capability is ~10mA/µm. Thus a 50 µm wide line can handle 500mA. As was seen in Figure 2-19 and Figure 2-20 showing the RF load lines, we have over a 2x increase in peak drain current with the GaN device. So again on the face of it if we were comparing a similar output power device between the technologies there would be a decrease in current resulting from the increased supply voltage (which would enable us to use narrower lines and hence obtain higher impedances), in practice the tendency is to replace devices with higher power ones and so in fact the current handling requirement increases. The real advantage is in the increased output impedance, in this particular case we have the almost ideal case of having the real part of the load (for optimum output power) being 50Ω. The fact that the real part of the optimum PAE load is higher indicates that there is an even larger device with an optimum PAE resistance of 50Ω.

Finally, arguably the most important argument is that the output capacitance is almost the same despite having increased the output power so significantly. Thus in general terms the matching task that we will set ourselves is as easy (or as hard!) as that for the GaAs device with only 20% of the output power.

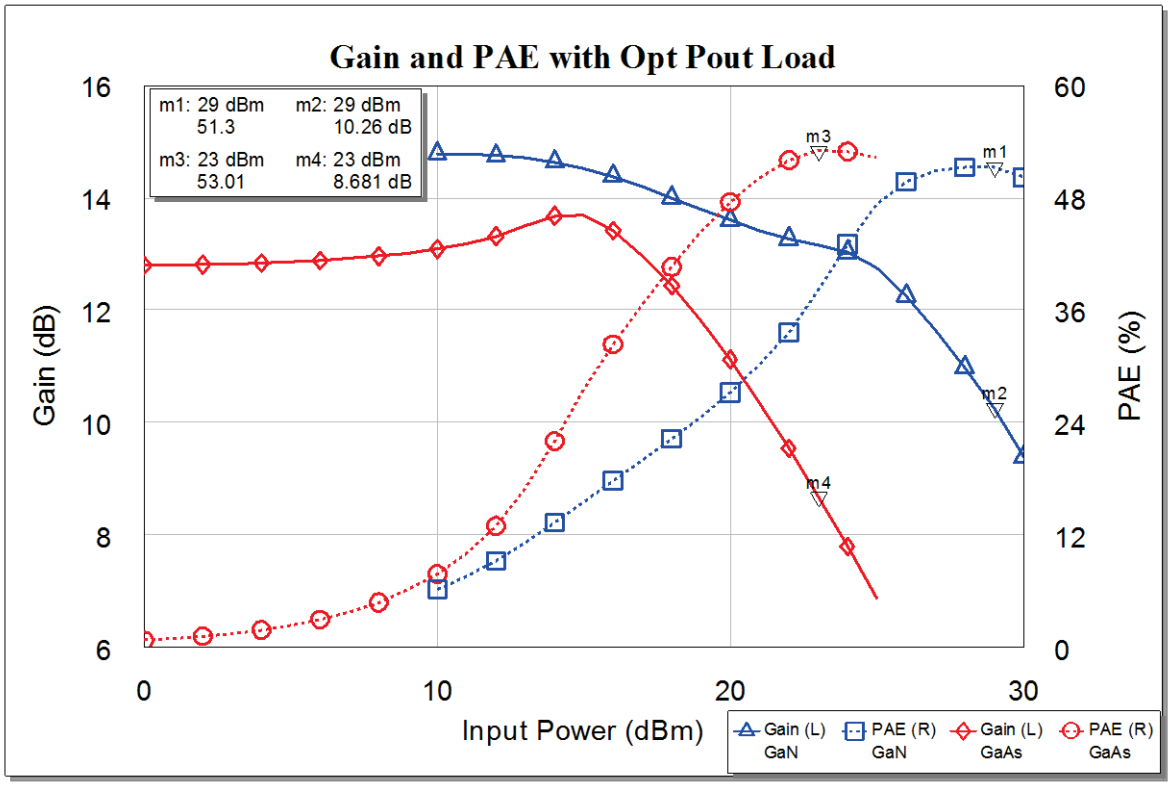


Figure 2-24, Optimum Output power load, Gain and PAE.

In summary therefore the implications of the main quoted advantages of GaN are:

a. Higher operating temperature and thermal conductivity – needed to combat higher power density.

b. Higher voltage hence lower current – not in practice as we tend to use a higher power (x5) device and increase the voltage by <3x.

c. Higher voltage - higher output resistance.

d. Lower pF/W – key to wideband impedance matching.

2.4 Amplifier Topologies

Amplification is the process of increasing the magnitude of an RF signal by converting DC power to RF. One or more active devices are required to facilitate this. The amplifier design process is dependent on the amplifier topology and the required performance but includes common elements such as device selection, biasing and matching. Having reviewed the technologies available the next step is to consider how the devices will be configured so as to meet the requirement specification.

A single stage amplifier designed with simultaneous conjugate reactive matching will always have a negative gain slope. A number of techniques can be used to flatten the gain of the amplifiers including:

- Using a multi-stage amplifier with interstage gain-slope compensation.
- Incorporating lossy matching that has higher losses at lower frequencies.
- Using a balanced topology (that incorporates reflective matching).
- Using a distributed topology.
- Using resistive feedback.

Some of these techniques are discussed in more detail below but in all cases the flat gain versus frequency response is achieved by reducing the low frequency gain, whilst maximising the top end. Interstage matching was discussed in chapter 1. This approach can be applied to many stages, but in order to provide isolation along the line-up resistive loss will need to be included. As bandwidth increases the number of matching elements required increases, but also the amount of gain slope that needs to be compensated for. Reactive compensation results in mismatching the input to a device and in a multistage amplifier this power is reflected back into the preceding stage. This is not only wasteful of energy but may be a cause of oscillations. One approach that can overcome this is to use balanced stages.

Balanced amplifiers utilise two identical amplifying circuits that operate in parallel. 3dB quadrature couplers are used to split the signal at the input and then to recombine the signal at the output. The beauty of this approach is that signals reflected back from the input of transistors are not passed back to the input but combined in anti-phase (hence cancelled) at the input port of the coupler. This is as a result of the physical properties of quadrature couplers, signals incident upon them are split equal in amplitude but shifted by 90° in phase. Provided the two devices are identical then the signals will be equally reflected and each reflected signal will undergo a further 90° phase shift and so be 180° out of phase at the input port and hence cancel. The reflected power will be in phase at the isolated port where it will be dissipated in the load resistor. Figure 2-25 shows the construction and the relative phases of the input and output signals (there is assumed to be no phase shift through the device for simplicity). At the output the two devices are

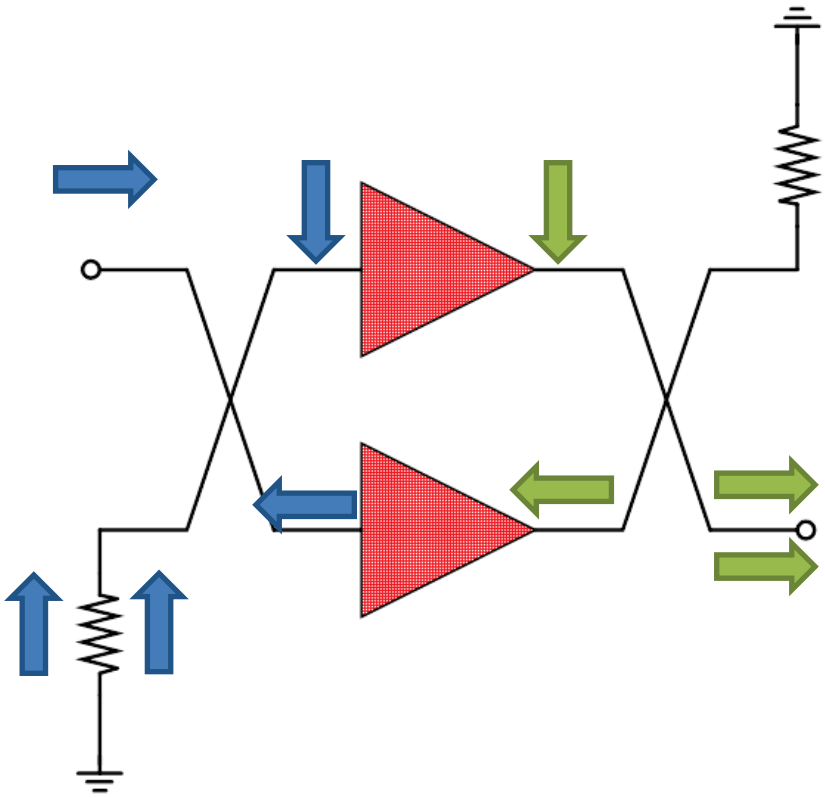


Figure 2-25, Balanced amplifier construction.

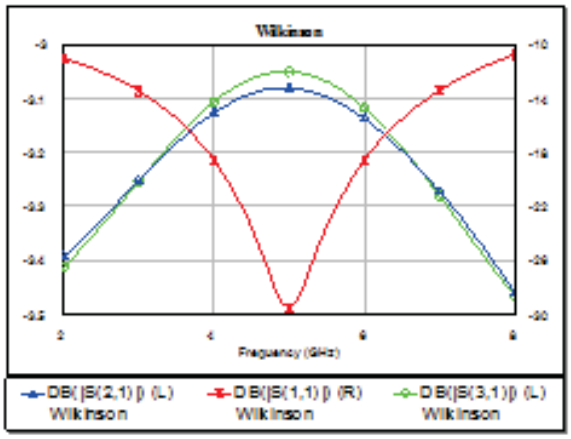
connected to a rotated coupler so that output port is in the opposite corner to the input port.

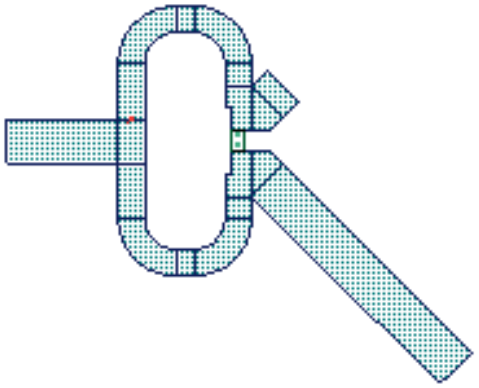
The main features of the balanced configuration can be summarised as:

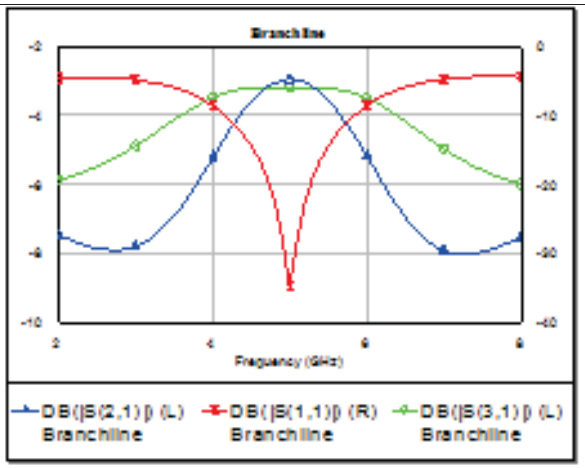
- The gain is that of one of the single amplifiers (minus the insertion loss of the two couplers).
- The noise figure is that of one of the single amplifiers (plus the insertion loss of the input coupler).
- The output Power is twice that of one of the single amplifiers (minus the insertion loss of output coupler).
- The matches are excellent (assuming the amplifying devices are identical).
- Very good gain flatness can be achieved by mismatching the input without any degradation to the load seen by preceding stages.
- Reliability is improved because a degree of redundancy is now built-in. If one amplifier fails, the balanced amplifier still operates, but with the gain reduced by approximately 6dB.
- Even harmonics are cancelled in the loads as they phase shift by ~180°.

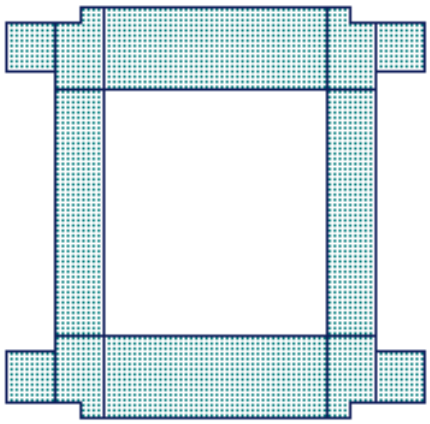
It should be noted (and this has not been found in any literature by the author) that if there is significant amplitude imbalance in the hybrid, and some wideband couplers exhibit ±1.5dB hence potentially 3 dB difference in power levels at the device inputs, then as seen in Figure 2-21, there could be a difference in the input impedance of the two devices which will lead to a poor input match to the balanced amplifier.

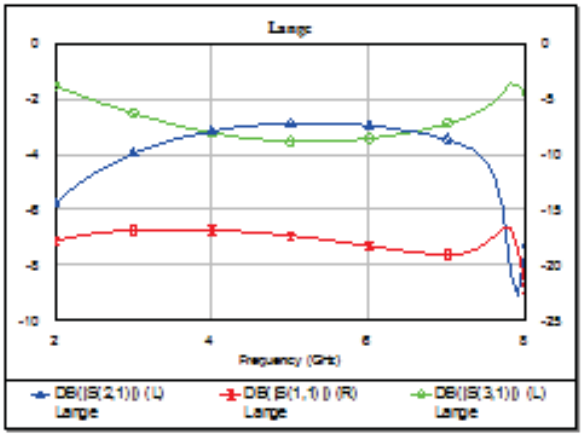
Quadrature couplers can be created directly on the substrate, the most basic approach being to use a Wilkinson splitter with a λ/4 line, to produce a 90° phase difference between the output ports. The balance between the ports is excellent over a wide bandwidth; however the phase difference is frequency dependent. The resistor between the output arms is used to dissipate any imbalance in the voltages, it may be omitted (and often is in high power applications) at the cost of degrading isolation between the output ports. Part of the problem with this resistor is the associated parasitics, and as the power increases and the resistor gets larger, so the problem gets worse. For narrow band designs the resistor can be offset by λ/4 lines reducing the parasitic effects at the centre frequency. The Branchline Coupler, has a number of advantages; the load resistor is offset from the signal path, the phase is relatively flat over up to a 20% bandwidth (this can be extended by using multiple sections, however the loss increases proportionally), and the design can be adjusted to non 50Ω output impedances which can ease the problem of matching to a transistor. For octave and greater bandwidths, Lange couplers [9] are extremely effective, the drawbacks are the small geometries that are required and the wire links between non adjacent 'fingers', although most processes can fabricate these using air bridges. A solution to the fine geometries required is to produce two more loosely coupled, 8dB, Lange couplers and connect them 'back to back' resulting in an overall 3dB coupling. Wider bandwidths can be achieved in this way, at the expense of size, complexity and higher insertion loss [10]. A comparison of the various approaches is summarised in Table 2-3 and a comparison of the phase responses in Figure 2-26.

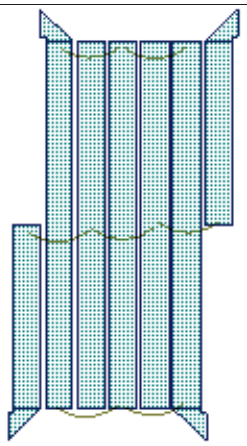


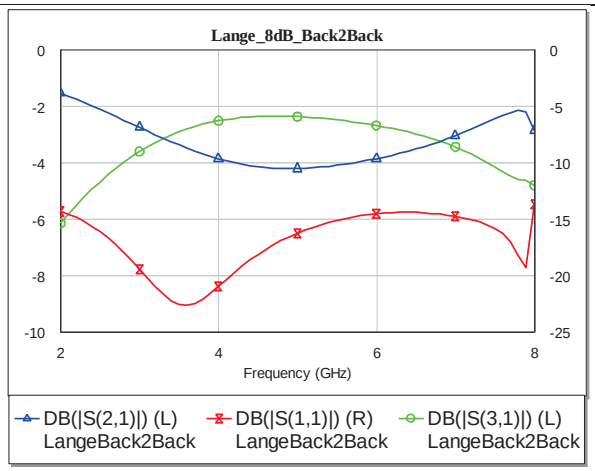












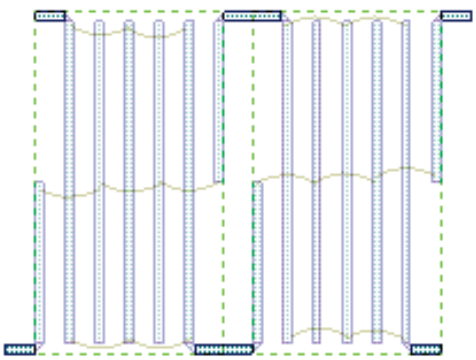


Table 2-3, Quadrature coupler approaches, from the top, Wilkinson, Branchline, Lange, cascaded Lange.

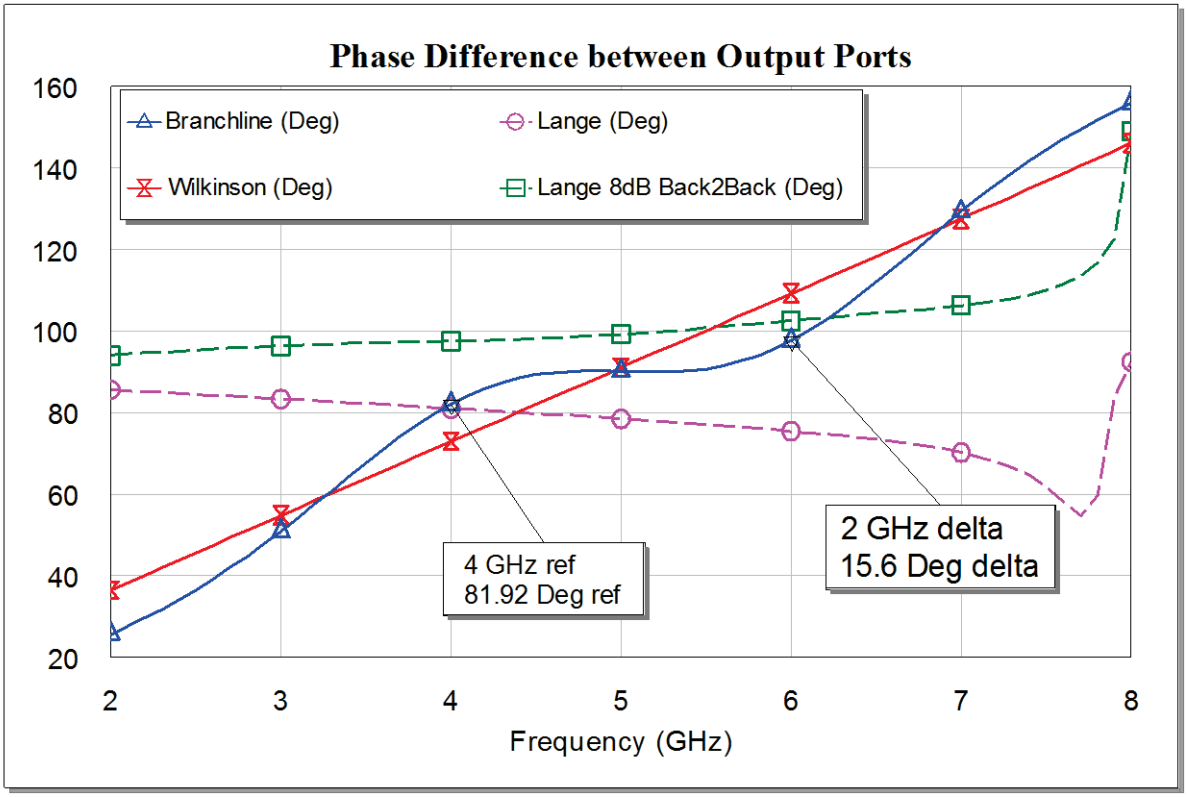


Figure 2-26, Phase variation comparison between the different approaches to quadrature couplers.

For wide bandwidths (up to about 2 octaves) Lange couplers provide the best solution for balanced amplifiers. They can either be fabricated on the device substrate, Figure 2-27, or produced separately and used to combine MMIC tiles. This is particularly attractive at lower frequencies where the size and yield[2] makes them expensive to produce on wafer.

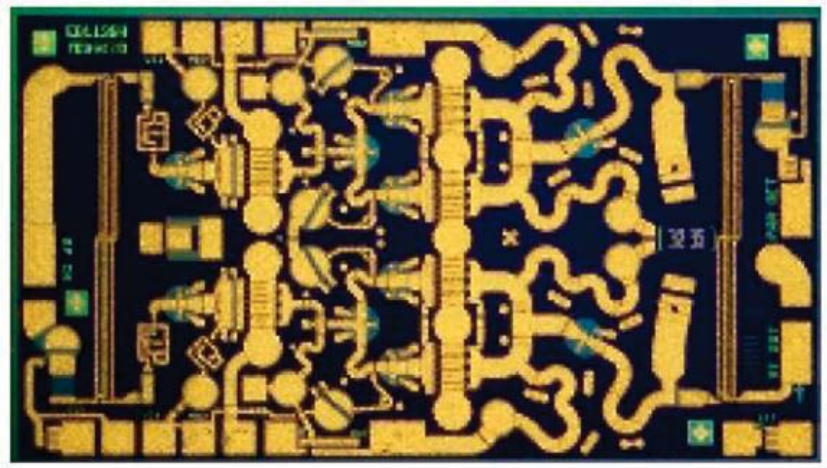


Figure 2-27, Commercially available 18 to 27GHz balanced (Lange couplers) amplifier.

[2] An on wafer balanced amplifier requires both 'halves' to be working satisfactorily, whereas they can be selected when combined externally.

An alternative approach capable of achieving extremely wide bandwidths [11] is the distributed or travelling wave amplifier (TWA). It is formed by absorbing the input and output capacitances of the transistors into a low pass filter structure. An example of a

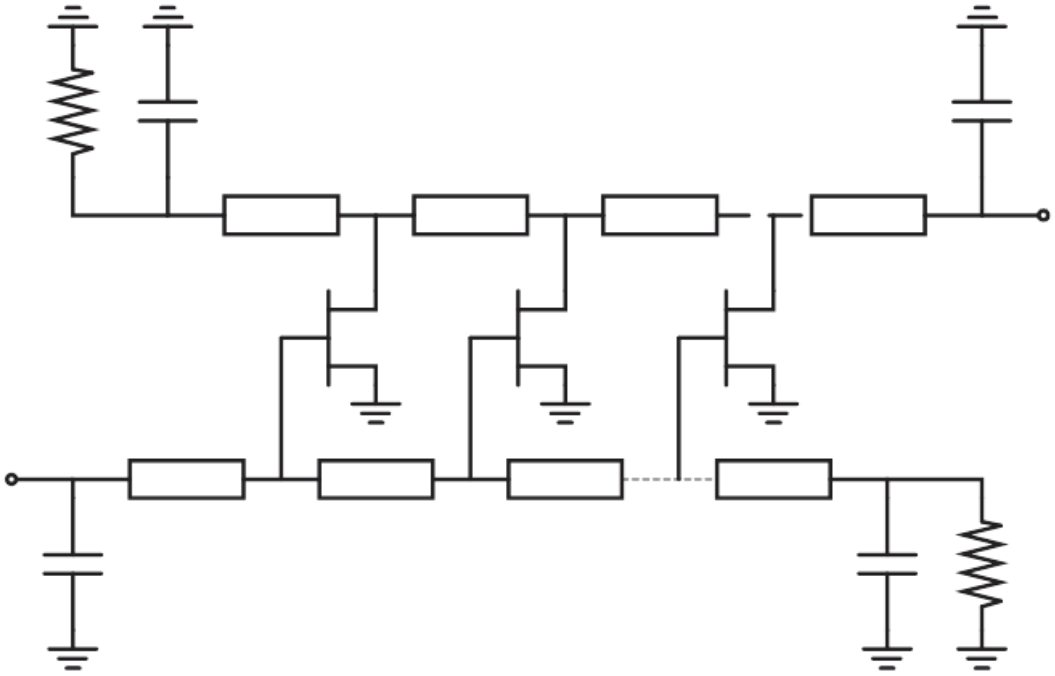


Figure 2-28, Distributed or Travelling Wave Amplifier schematic, input lower left, output upper right.

typical three-transistor design is illustrated in Figure 2-28, more sections can however be used.

The gate and drain capacitance of the transistor is absorbed into the lossy artificial transmission lines created by the inductive lines connecting the devices. As the signal travels along the transmission line from the input on the left each device is successively excited and the signal is transferred through the device and amplified on its output (drain). The signals on the drain line add in phase as they 'move' towards the output, those travelling in the opposite direction are not in phase and any that are not cancelled are dissipated in the load in the top left. The gain versus frequency response of TWAs can be tailored to provide a negative, flat or positive response, hence they are useful as input or driver stages.

TWA do not provide a solution to all problems however. The total output current is dependent upon the phase coherence of the individual current generators; hence phase velocity equalisation is required. This typically amounts to the addition of shunt capacitance on the gates and drains of the devices. Also as the signal travels along the line of devices each successive one is stimulated by a lower signal due to the gate source resistance of the

previous stage. One solution to this and the phase velocity issue is to tailor the device sizes so that each position along the line has a different gate capacitance and resistance. The gain of the overall amplifier is proportional to the number of devices used, however acting in the opposite direction is the increasing losses in the gate and drain lines which increase with the number of devices. The optimum number of devices, N_{opt} that maximises the gain at a given frequency has been derived [12] as {2-4} where A_g is the gate line attenuation and A_d the drain line.

$$N_{opt} = \frac{\ln\left(A_d / A_g\right)}{A_d - A_g} \qquad \{2\text{-}4\}$$

Shunt feedback as a solution to improving match, gain flatness and stability has been described in chapter 1. In MMIC amplifiers feedback is particularly applicable to driver stages as the gain slope can be easily adjusted. Inductors are realised using high impedance

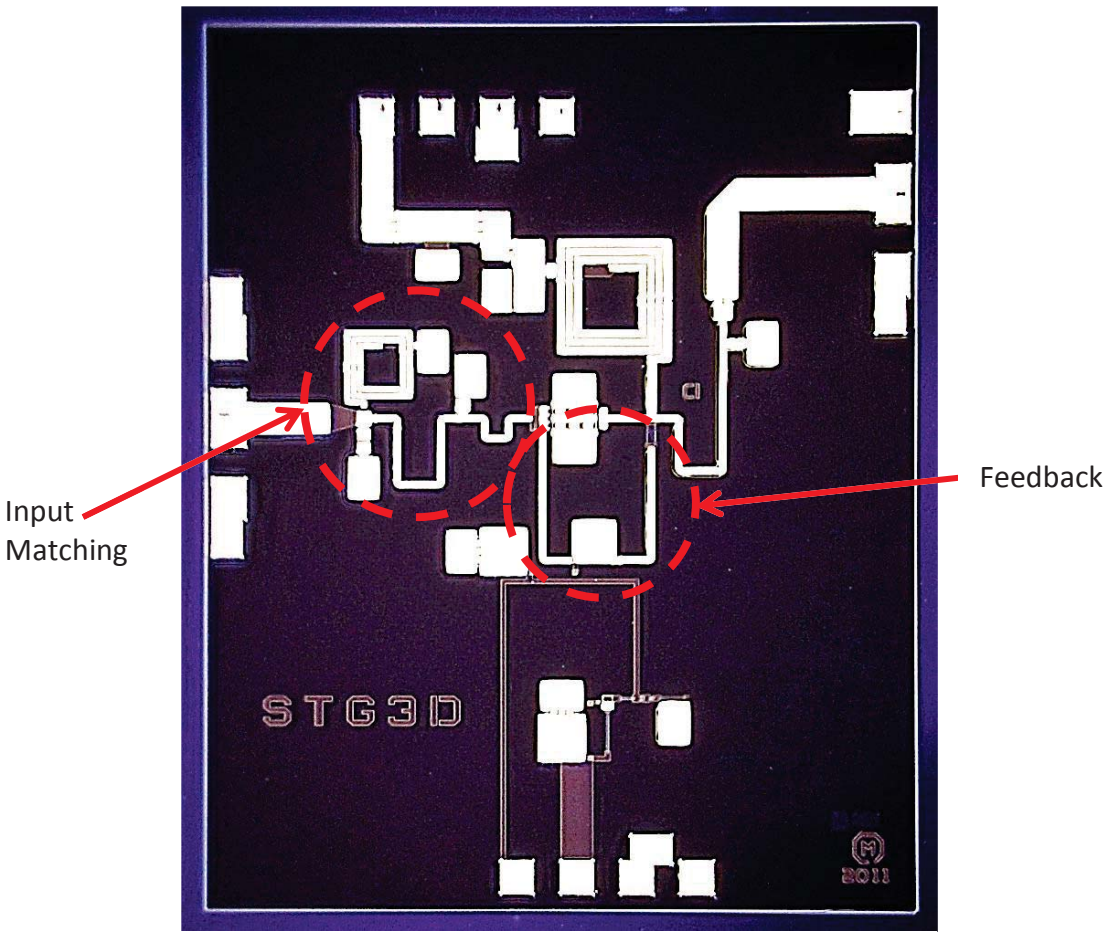


Figure 2-29, 6-18GHz feedback single stage amplifier, *{courtesy of Plextek Ltd}*.

transmission lines as shown in Figure 2-29, 6-18GHz feedback single stage amplifier, *{courtesy of Plextek Ltd}*.

Feedback amplifiers are particularly effective at low frequencies and in MMIC implementations because the parasitics have a much lower contribution. With packaged devices external feedback in wide bandwidth designs become more difficult to implement due to the additional capacitances, inductances and phase lengths that distance the circuit from the current generator in the transistor as frequency increases. In MMIC designs the limitations come as power increases and the physical size of the elements needs to increase proportionately to accommodate the increased current. It also becomes more difficult to combine transistors with feedback, due to the physical limitations of combining the devices where ideally the outputs should be close together.

Some simple approximations for the gain, G and feedback resistance, R_{fb} have been derived {2-5}{2-6} [13], but these should be treated as starting values because they are obtained from very basic circuit models of the transistors.

$$G = 1 - g_m Z_0 \qquad \textbf{\{2-5\}}$$

$$R_{fb} = g_m {Z_0}^2 \qquad \textbf{\{2-6\}}$$

2.5 Practical MMIC PA design considerations

A suggested standard approach to the design of MMIC PAs is to divide the task into 3 main sections,

1. Architecture
2. Nonlinear Analysis and Design
3. Integration

This is not to say that some elements cannot be worked on in parallel, but the main design flow should follow this path. The improvement in nonlinear models and simulation has caused a change in the approach; previously a design based on small signal data (S parameters) [14] followed agreement on the architecture, the power stages would then be 'tweaked' to allow for nonlinear behaviour. This was because large signal analysis was considered inaccurate and extremely time consuming and therefore not worth the effort.

However, as the output stages are key determinants of system performance it is essential that these elements are optimised so that the most is made of their potential.

2.5.1 Architecture Design

The basis of any design is the requirement or system specification. This should outline the minimum performance and physical characteristics, output power, efficiency, gain, size, weight, etc. as well as targets or 'nice to have' details. It is not necessary to go into system specifications in detail here, however these must form the backbone of the design and all decisions should conform to achieving the requirement defined in this document.

A typical requirement specification for a power amplifier would include the following:

- Bandwidth.
- Output power (defined in terms of P1dB or Psat).
- Gain (usually linear but may include saturated minimum gain).
- Gain ripple.
- PAE (at a specific output power level).
- Linearity (which may be expressed in a number of ways, for example third order intercept point or 3rd order intermodulation levels).
- Input match[3].
- Operating temperature range.

From these parameters a decision will be made as to the most appropriate device technology and the total area of device periphery required at the output. Note that different technologies specify the periphery differently, for HEMTs and MESFETs it is the power/mm width of the gate whilst for HBTs it is power/mm length of the emitter. Foundries are also not consistent in defining output power; some use Psat others P1dB output power, some use class A whilst others A/B or B and to add to the confusion many conduct the measurements pulsed as they are done on wafer. It is also important to allow for the insertion loss of the output matching circuit and if PAE is important to allow for the

[3] Output match is arguably unnecessary as in order to deliver the power to the load the amplifier will need to be as well matched as it can. There are also technical difficulties measuring S22 whilst the amplifier is operating at power and to measure it in its small signal state gives little indication as to the power performance.

fact that the optimum PAE load is associated with lower output power as discussed in chapter 1.

From the total area of output periphery required it is now necessary to determine the optimum cell size or how the total periphery is to be divided up into individual transistors. Although ideally it would be easier to design one large device, the output capacitance increases with device size and hence the gain decreases. Additionally it is more difficult to get the heat away from a single large device, the gate source capacitance also increases and the phase differential across the gate begins to limit efficient combining. This latter point also highlights why the maximum power of an individual cell decreases with frequency. Thus the optimum device size is dependent upon the required gain from the stage. An amplifier will typically be made up of a number of stages to meet the required overall gain level and how this gain is divided up between the stages is to some extent a matter of choice; however the gain will tend to decrease as the stages get closer to the output as the devices will get larger. There are 'rules of thumb' which say that the individual cells of the final stage should have a maximum gain of 10dB at the top frequency of the band [14]. More gain in the output stage however increases the power added efficiency by reducing the input drive level, thus besides less power being required from the driver fewer overall stages may be needed in the RF amplifier line-up. This not only increases efficiency itself but reduces wafer area and therefore cost. Figure 2-30 shows an example of a pair of balanced stages (5.6W) and how their gain affects the driver power.

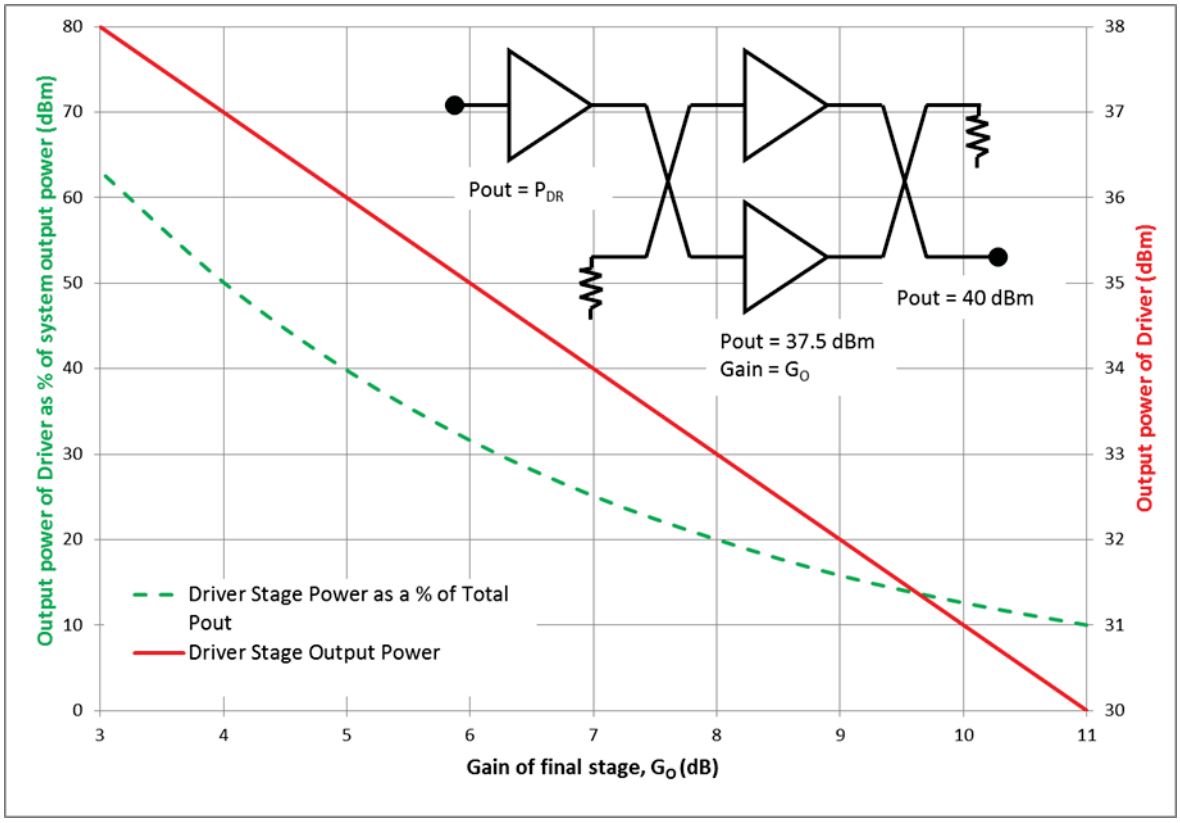


Figure 2-30, Effect of Output stage saturated gain on Driver Output Power and Contribution to Total Power drawn (assumes 0.5dB loss in combiners).

It is important that the driver stage not be the limiting factor in the output power. For linear amplifiers all of the compression should ideally occur in the output stage. For saturated amplifiers although this is less critical it is still important that the output power performance comes from the output and not the driving devices, although compression in the earlier stages can be used to improve output gain flatness. The clearest evidence of this is a very 'soft' compression curve resulting from gain being lost through power compression at more than one stage. This is partly where the 10dB gain rule comes from. If a device has a gain of 10dB then obviously the input power required is 10dB less than the output. Assuming that the saturated power compression is 3dB then the saturated drive level is 7dB below the output power. The significance of this is that theoretically the same device used as a driver could drive 4 devices in parallel. In practice the problem comes from not having a load which is a perfect 50Ω. In many systems, such as where the load is a wideband antenna, this is not the case. Similarly the driver stage is not seeing the same load that the output stage is (it is seeing the output stage's input through the splitter) which will be far from perfect. Thus mismatch losses are likely to be higher. Nonetheless if the gain is 10dB minimum across the frequency band then there will also be points where it is higher and to

some extent the law of averages helps to balance things out. Thus the increased gain of GaN can provide more margin than GaAs designs of a similar architecture.

It is important to distinguish between linear and saturated gain in the calculations. Typically the saturated gain is taken to be that of the 3dB compression level, i.e. when the gain has fallen 3 dB below the linear level. However this is not always the case and should be checked depending on the technology, bias and frequency of the device. Also note that particularly with 'soft' or 'slow' compression the power continues to increase whilst the gain continues to fall and hence saturated gain is not a fixed characteristic. Also the amount of compression is process specific, experience of different GaAs processes has shown that whilst some devices will happily survive being driven constantly 5-6 dB into compression other struggle with 2-3 dB. It is likely that GaN will be able to be driven well into compression due to the higher breakdown voltages.

When the maximum cell size has been determined based upon the required gain over the operating bandwidth, if there is a shortfall in output power compared to the system requirement, the difference must be made up by combining the cells. Multi-way combiners are used to combine large numbers of devices and it is useful to remind oneself of the relationship between the number of ways in a combiner/splitter and the coupling in dBs, Figure 2-31. This figure excludes the insertion loss of the combining structure, as this will depend upon the technique employed; it just indicates the splitting/combining power relationship. A common approach in MMIC design is to replicate the output stage as a driver. This has many advantages in reduced development time, device measurement, characterisation, modelling and risk. Using the information in Figure 2-31 a table can be made showing the how the stage gain limits the number of parallel output devices that can be driven from the same power level device, Table 2-4, it is assumed here that the saturated power level is P_{3dB} or that for optimum linear operation the driver is 3dB less compressed than the output stage.

Linear Gain (dB)	Driver : Output Stages
>7	1 : 2
>10	1 : 4
>13	1 : 8

Table 2-4, Relationship between stage gain and 'Fan-out' ratio.

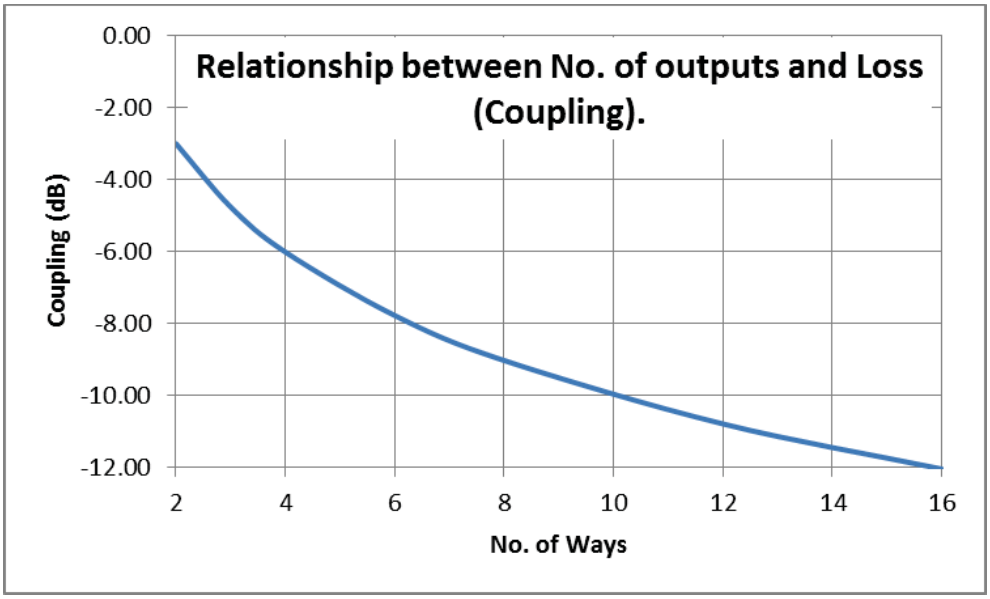


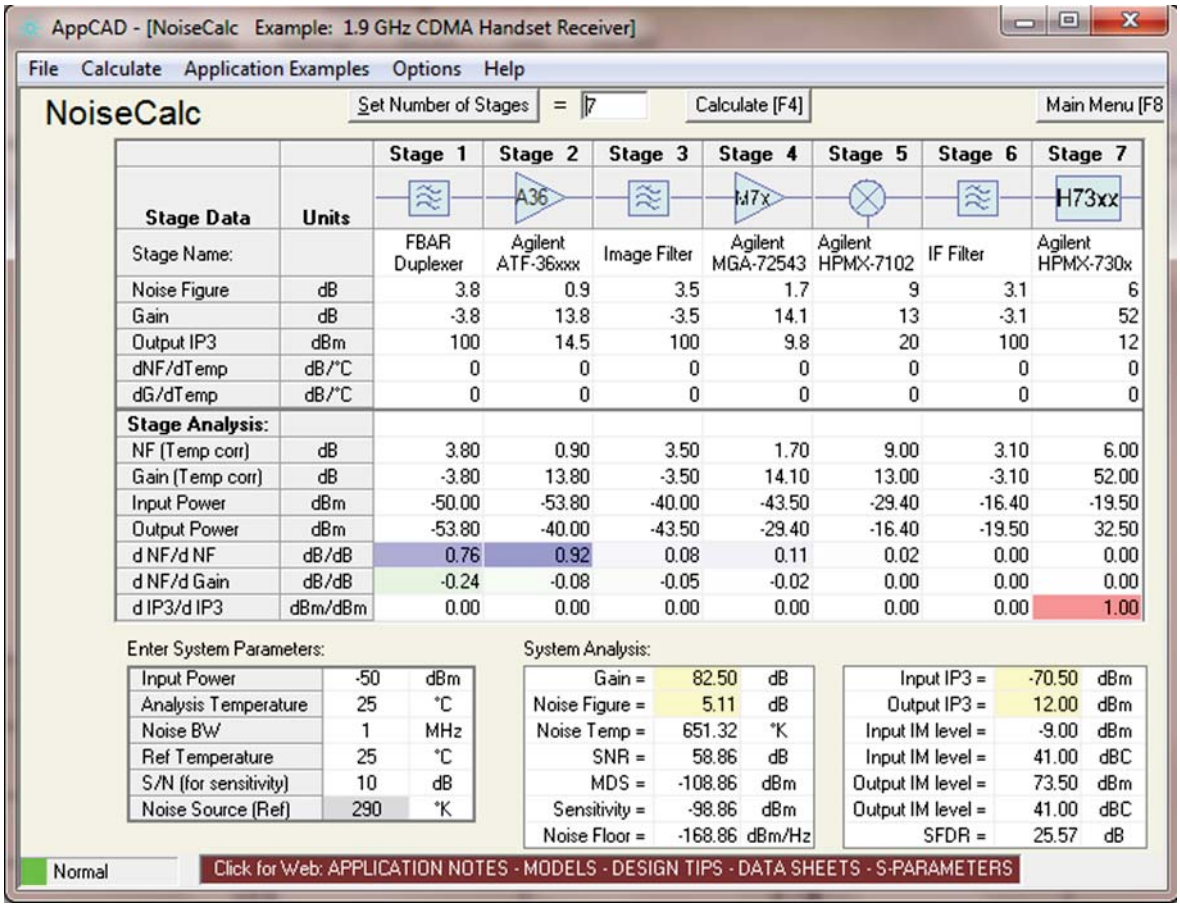
Figure 2-31, Coupling of multi-way combiners/splitters.

In wide bandwidth amplifiers there is a particular difficulty in optimising the efficiency of the driver stage. As has been said in chapter 1 the optimum PAE is achieved at specific power level, it peaks unlike the output power, which tends to increase but at an ever decreasing rate[4]. The peak PAE drive power requirement will vary across the operating frequency range due to non-optimum load matches that are achieved in practice as will be shown in chapter 6. There is thus an added constraint on the driver stage that besides also trying to get it to operate at its optimum PAE we also wish for output power delivered to follow a particular distribution with frequency. As described in chapter 1 in input matching this results in lossy approaches being taken which works against maximum efficiency.

The final step in developing the architecture of the MMIC is to create a power budget. Some CAD software includes this system analysis, including such things as compression points and intermodulation levels. There are also free versions available such as that contained within AppCAD (available from http://www.avagotech.com), a screen shot of which is shown in Figure 2-32. However these do not always meet the particular requirements, or are over complicated for this stage of the design (there is no point using a full system simulator when the exact details and performance of each of the amplifier stages is still unknown). A simple spread sheet can allow the basic calculation of the number of stages, gain, power consumption etc. to be calculated, as shown in Figure 2-33.

[4] This is a significant difference between solid state and 'tube' type amplifiers, care must be taken with travelling wave tube amplifiers not to overdrive them as this leads to decreasing output powers, hence the need for control loops and limiting amplifiers.

At this stage in the design it is important not to over complicate things. From the MMIC architecture, goals will be set for the stage designs to achieve and providing key information to others on power supply requirements, approximate size, number of stages and thermal dissipation. The complexity of the calculations can be adjusted by the user, for example more detailed thermal calculations could be added or even approximate area required (by stages and combiners), however the key is that targets have now been set for the individual stages and the relationships between stages has been defined.



AppCAD - [NoiseCalc Example: 1.9 GHz CDMA Handset Receiver]

File Calculate Application Examples Options Help

NoiseCalc — Set Number of Stages = 7 — Calculate [F4] — Main Menu [F8

Stage Data	Units	Stage 1	Stage 2	Stage 3	Stage 4	Stage 5	Stage 6	Stage 7
Stage Name:		FBAR Duplexer	Agilent ATF-36xxx	Image Filter	Agilent MGA-72543	Agilent HPMX-7102	IF Filter	Agilent HPMX-730x
Noise Figure	dB	3.8	0.9	3.5	1.7	9	3.1	6
Gain	dB	-3.8	13.8	-3.5	14.1	13	-3.1	52
Output IP3	dBm	100	14.5	100	9.8	20	100	12
dNF/dTemp	dB/°C	0	0	0	0	0	0	0
dG/dTemp	dB/°C	0	0	0	0	0	0	0
Stage Analysis:								
NF (Temp corr)	dB	3.80	0.90	3.50	1.70	9.00	3.10	6.00
Gain (Temp corr)	dB	-3.80	13.80	-3.50	14.10	13.00	-3.10	52.00
Input Power	dBm	-50.00	-53.80	-40.00	-43.50	-29.40	-16.40	-19.50
Output Power	dBm	-53.80	-40.00	-43.50	-29.40	-16.40	-19.50	32.50
d NF/d NF	dB/dB	0.76	0.92	0.08	0.11	0.02	0.00	0.00
d NF/d Gain	dB/dB	-0.24	-0.08	-0.05	-0.02	0.00	0.00	0.00
d IP3/d IP3	dBm/dBm	0.00	0.00	0.00	0.00	0.00	0.00	1.00

Enter System Parameters:

Input Power	-50	dBm
Analysis Temperature	25	°C
Noise BW	1	MHz
Ref Temperature	25	°C
S/N (for sensitivity)	10	dB
Noise Source (Ref)	290	°K

System Analysis:

Gain =	82.50	dB
Noise Figure =	5.11	dB
Noise Temp =	651.32	°K
SNR =	58.86	dB
MDS =	-108.86	dBm
Sensitivity =	-98.86	dBm
Noise Floor =	-168.86	dBm/Hz

Input IP3 =	-70.50	dBm
Output IP3 =	12.00	dBm
Input IM level =	-9.00	dBm
Input IM level =	41.00	dBC
Output IM level =	73.50	dBm
Output IM level =	41.00	dBC
SFDR =	25.57	dB

Normal — Click for Web: APPLICATION NOTES - MODELS - DESIGN TIPS - DATA SHEETS - S-PARAMETERS

Figure 2-32, AppCAD RF budget analysis software.

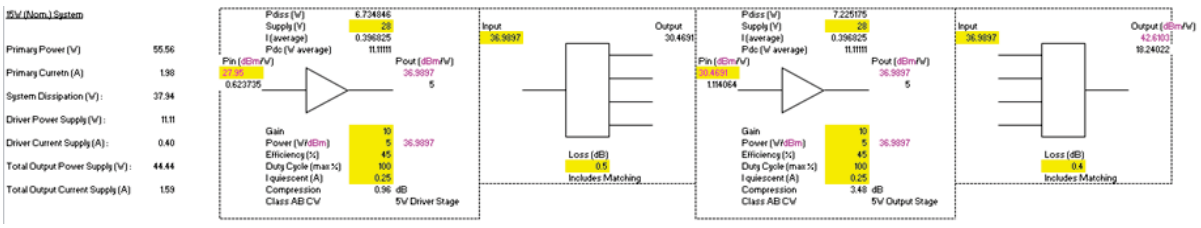


Figure 2-33, Basic spread sheet based system architecture calculator.

The normal practice for power amplifier designs is to start at the output stage and work backwards. Although it is simpler to work in logarithmic units (dBs and dBm) in the RF

sections, for the DC components mW and W are better. It is also import to display linear RF power as for example an insertion loss of 0.5 dB may not sound significant, but after a 10W device represents 1.1W in mainly thermal dissipation. It is also useful to display the amount of compression/back-off. When such a system representation has been created it is much easier to show how the effects of changes ripple through.

There are a number of combining techniques available to increase output power, in MMICs. The prime considerations, which result in the selection of the appropriate approach, are:

- Loss.
- Size.
- Bandwidth.
- Current and Voltage handling.
- Odd mode stability.
- Phase balance.

Combining can be separated between on and off chip, as has already been mentioned in the discussion of quadrature combining, but in terms of increasing the power in a single MMIC on chip combining shall be reviewed first.

Bandwidth and phase imbalance are to some extent related. The importance of phase difference between combined signals is a simple trigonometric calculation and is summarised in Figure 2-34. As seen earlier in the discussion of quadrature couplers and summarised in Table 2-3, the Wilkinson combiner/splitter has a very wide bandwidth in terms of amplitude, but is limited as a quadrature combiner due to phase. The Wilkinson splitter, is often considered a 2 way splitter, however the original paper described N-way topologies [15]. The basic Wilkinson consists of 90° lines of impedance $\sqrt{N}.Z_0$ connected to a common port of impedance Z_0 and with the other ends connected to the input/output ports with a resistor of impedance Z_0 connected to a common node. These combiners exhibit low loss and have good isolation, but their physically large size can be a draw-back, however through appropriate line meandering this can also be a useful method of connecting widely spaced devices. The other draw-back is that the size of N is limited by the realisable line widths on the substrate which can still handle the power. The technique can also be employed in off-chip radial combining. The isolating resistor can be a problem to implement

in multi-way on chip combiners, however it has been found that provided the output phases are the same the structure still works well if the resistor is omitted. The lower isolation resulting can be overcome if the output stages are quadrature balanced.

An alternative to creating a Wilkinson combiner using λ/4 lines (distributed approach) is to use lumped circuit equivalents to the transmission line, i.e. series inductors and shunt capacitors. Using spiral inductors these can be more compact, Figure 2-35, but the current handling of the inductors is not high, bandwidth is more difficult and >2 ways difficult to achieve, hence their use on the inputs rather than the outputs.

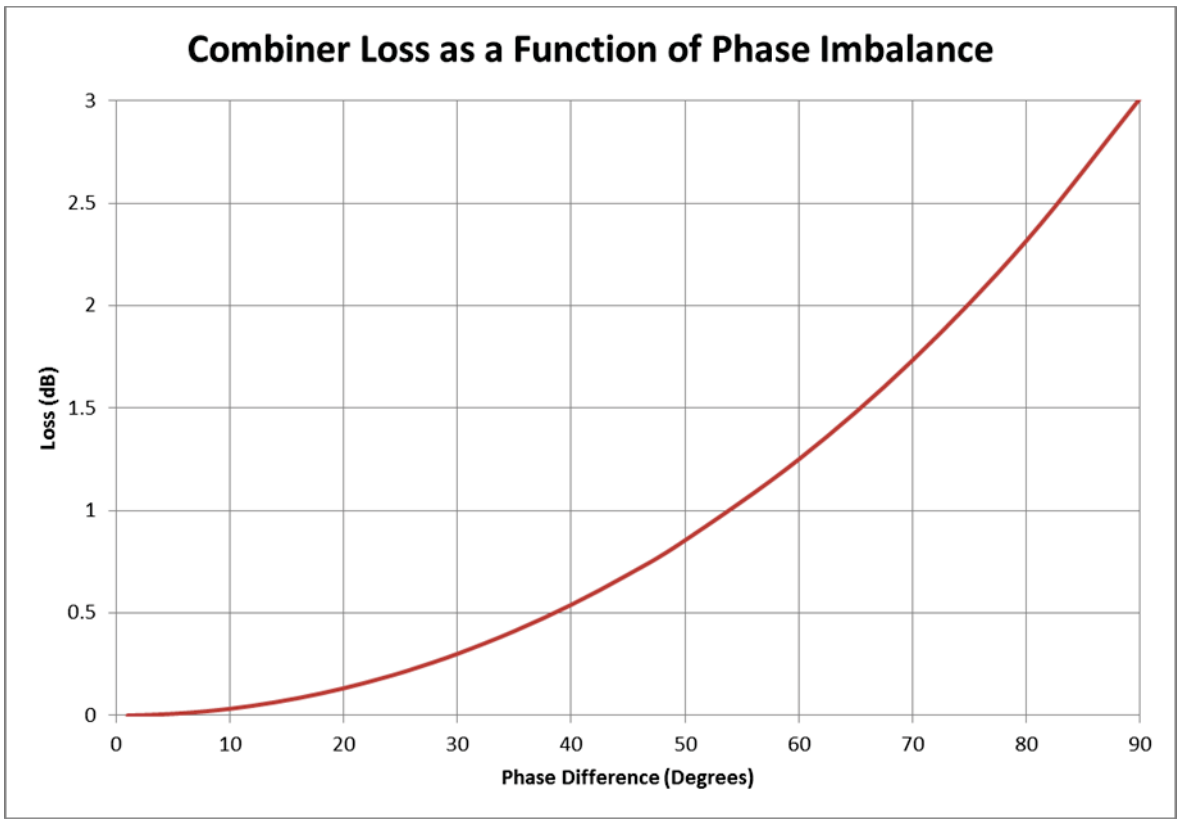


Figure 2-34, Effect of Phase difference between two combined signals and loss.

The tapered transmission line is described in detail in [16] and shown Figure 2-36. This structure transforms the input impedance to a lower impedance determined by the number of opposite ports (Zin/N). The curve of the taper determines the input match and the length of the taper the lower frequency. In the ideal case the capacitance between all of the adjacent output lines and to ground would be the same and there would be no capacitance between non adjacent lines. In practice the outer lines obviously have half the adjacent line capacitance. It is however in this way that the higher N the less significant the discrepancy from the outer lines. The structure is very broadband (100%) and has good isolation between the ports, the disadvantage is clearly size and modelling and synthesis is not trivial.

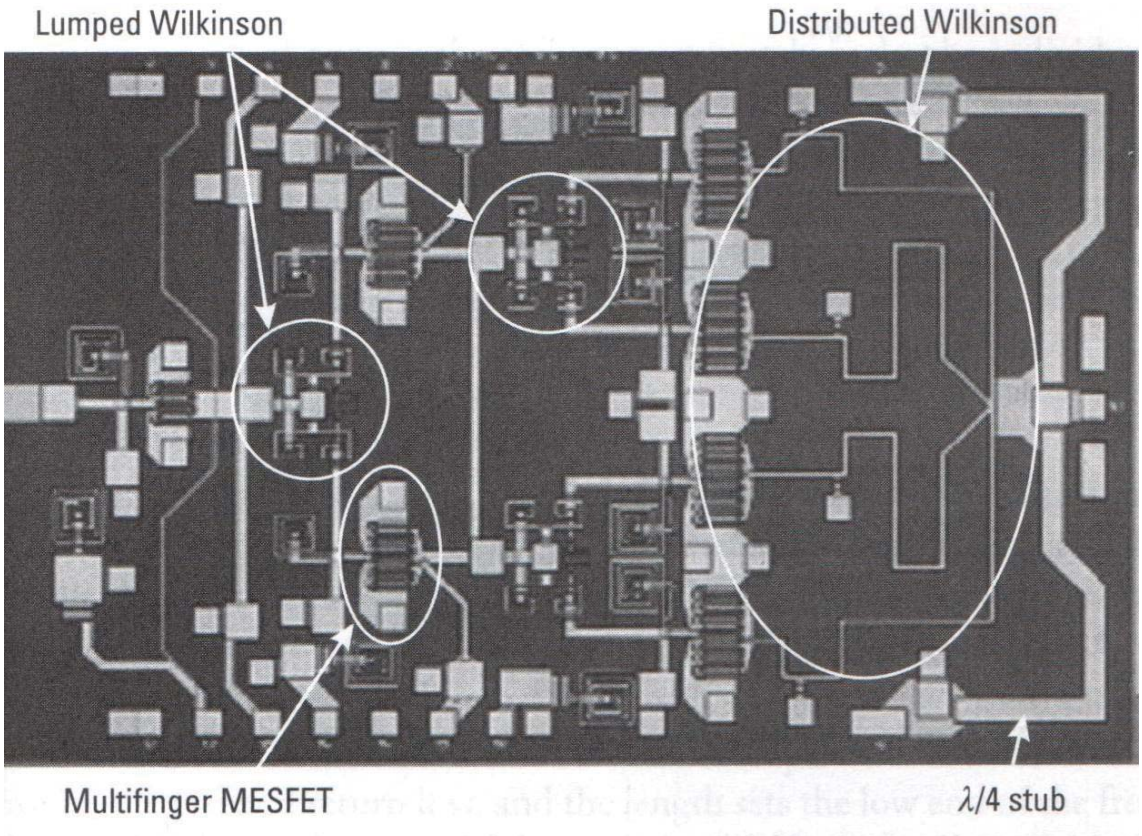


Figure 2-35, MMIC using lumped and distributed Wilkinson combiners, [14].

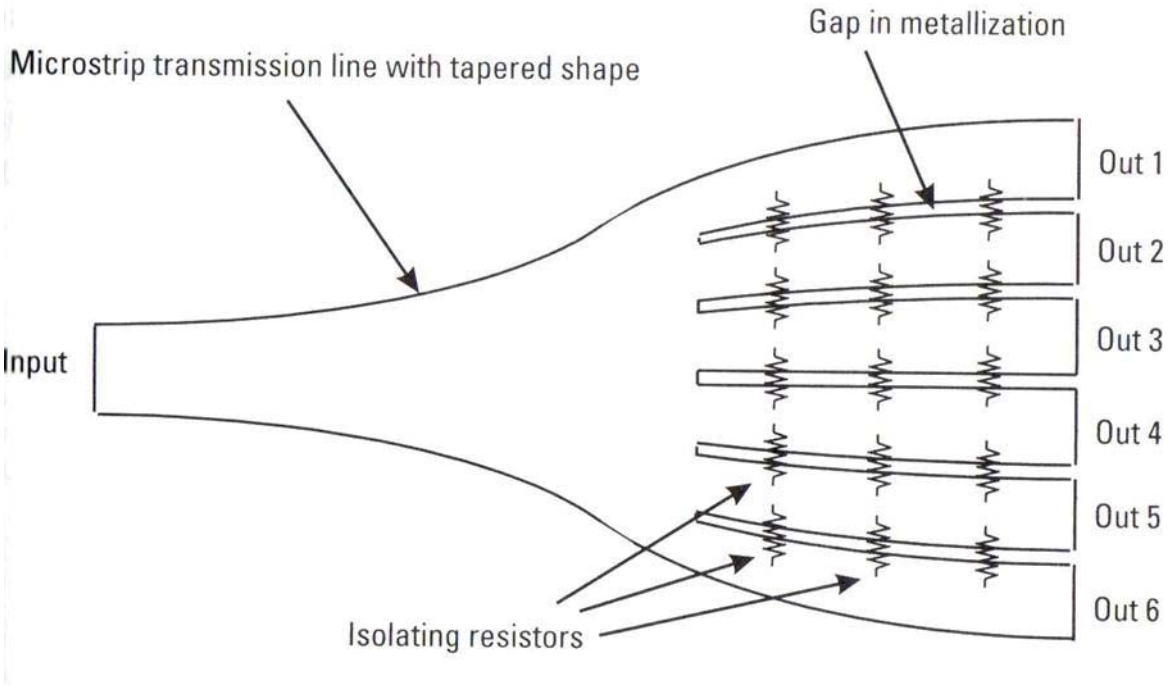


Figure 2-36, Tapered multi-way splitter, [14].

Baluns (balanced to unbalanced) are structures where the key characteristic is a 180° phase difference between the balanced ports. They are used extensively in MMIC mixers and in RF and microwave discrete power amplifiers; however their use in MMIC PAs has been limited mainly due to layout considerations. The 180° phase difference between ports has an attraction in class B amplification as two devices can be run in anti-phase and then recombined giving both efficiency and linearity. Baluns are mentioned here because they may have use in combination with other combining techniques in higher efficiency designs. Particularly with the introduction of multi-layer MMIC processes very complex baluns can be integrated [17].

The Synthesised Transformer Network, [18] consists of lumped and distributed elements. The approach uses filter synthesis methods to design a band pass circuit with multiple outputs. Network transformations are applied to the element values to get distributed circuit elements similar to that shown in Figure 2-37. The method thus has the advantage that it can include impedance transformation reducing the matching complexity. Extremely broad bandwidths can be achieved using this approach.

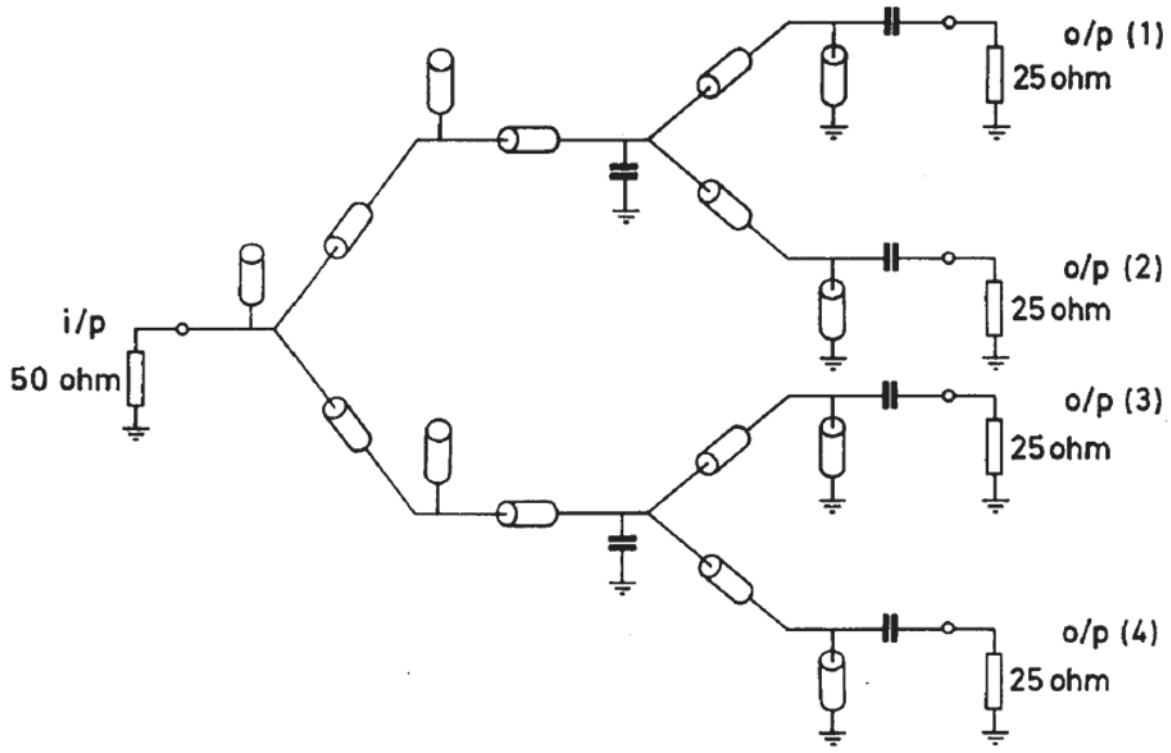


Figure 2-37, Synthesised Network of 4 way input divider with 2:1 impedance transformation, [18].

An extension of the synthesised transformer networks is to design each output device to match to an impedance of nZ_0 where n is the number of output devices. Two devices are then paralleled and their common nodes connected together. Any shunt elements are transformed into a single component thus leaving a single ended stage with two devices. The circuit should now be optimised to $(n/2)(Z_0)$, taking into account the discontinuities of the elements now combined. This process is repeated until all the devices are in parallel with single inputs and outputs matched to Z_0. Appropriate selection of circuit elements results in bias, decoupling and matching all being incorporated into the structure. The approach is limited to about 20% bandwidth and current limitations can restrict the series inductances of the matching elements.

Bus-bar combining takes its name from the bus-bar connections of electrical distribution circuits. In the MMIC case all of the output devices are connected to a wide bus-bar as shown in Figure 2-38. Note that the input feed network is of a different type. Provided that all of the devices are identical (good process repeatability), the separation is $<<\lambda$, and that they are fed by identical signals in phase and amplitude the output voltages are identical and no current flows between the devices. An advantage of using GaN with this technique is that the lower ε_r makes the effective λ on the substrate greater and hence a greater separation can be tolerated or the operating frequency extended. The mid-point

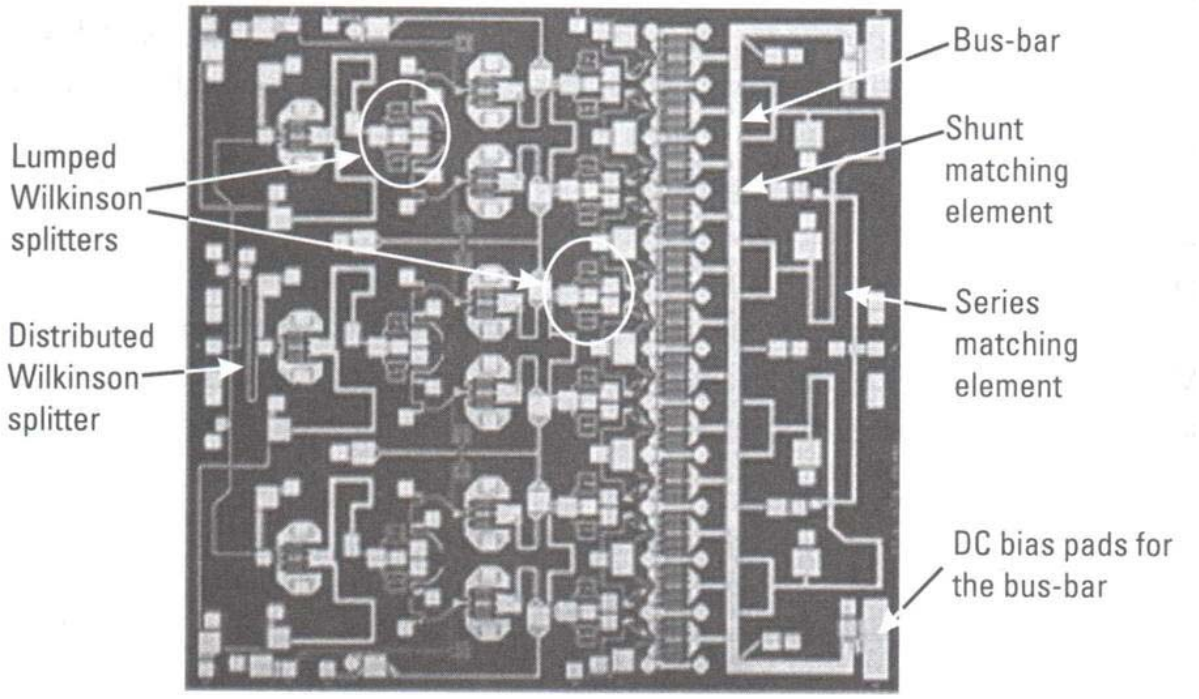


Figure 2-38, Bus-bar connected output stage 5W X band MMIC, [14].

between devices on the bus-bar is a virtual open circuit and can be used as a connection point for pairs of devices. Using impedance transformations pairs are then combined to the final output. This can be carried out in a very compact manner as shown in Figure 2-38. A key advantage of this technique is the ability to supply high DC currents to the devices and the simplicity of feeding many parallel stages, (particularly the ones in the centre of the chip). The disadvantage is the moderate bandwidths, (although it would be interesting to see what could be achieved with 50Ω output impedance GaN cells), and the necessity to use E-M and nonlinear simulation concurrently to analyse the structure.

The final technique reviewed here is an off chip method with the capability to produce not only very high power levels, but also decade bandwidths. Spatial combining was initially envisaged in waveguide, [19] where MMICs mounted on carriers mounted in waveguide would have their inputs and outputs constructed so as to match the incoming and outgoing signal. There were two main limitations to this approach, one was thermal – how to get the heat out of the devices and the other was bandwidth. These have been overcome by the use of free space or quasi-optical illumination of the MMICs input and the reverse approach on the output. The centre pin of the input coaxial connector is widened following a tapered characteristic to a point where MMIC devices mounted on carriers arranged radially about the central hub, with fin-line antenna on input and output. The output is a mirror of the input, Figure 2-39. Reported losses are less than 0.5 dB at 20 GHz and a large number (>16)

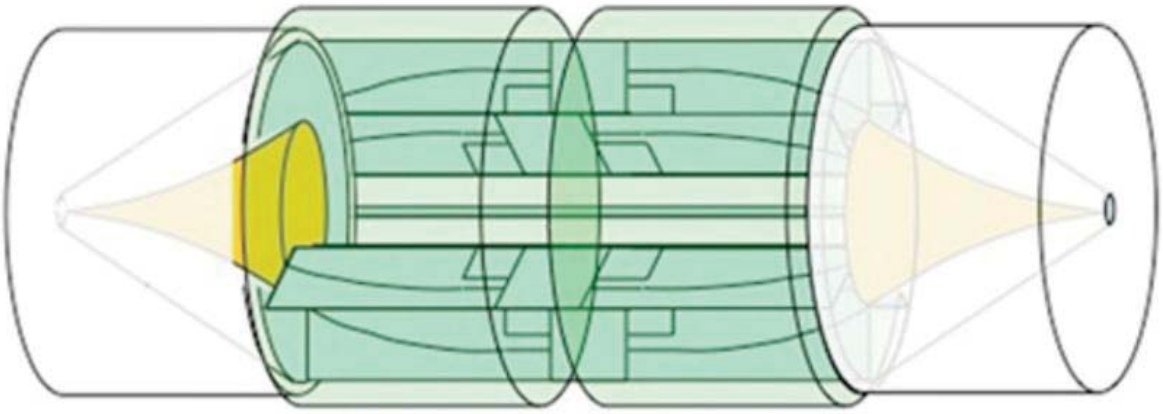


Figure 2-39, Construction of Spatial Combiner, [40]

of elements can be arranged around the hub, limited mainly by the thermal performance. Once the signals are coupled onto the MMIC circuit via the antennas the centre hub can be connected to the outer body. In this way even water-cooling can be incorporated into the structure. The approach is very attractive in that the bandwidth of the combining structures begins to approach that achievable in the amplifiers themselves but more importantly unlike most other combining structures it is not limited in the number of amplifier elements. Typically combining losses increase as the number of elements increase (either due to phasing or increased number of binary combining stages). The advantages are summarised in Table 2-5.

System	Spatium (Capwireless)	SSPA	TWTA
Bandwidth	>Decade	3 Octaves	3 Octaves
Noise Figure	<8dB	<8dB	~35dB
Harmonics at Psat	-15dBc	-15dBc	+3 to -6 dBc
Combining Efficiency	>90% independent of number of elements.	75-90%, diminishes as number of devices increases.	100%

Table 2-5, Comparison of different amplifier system performances.

2.5.2 Nonlinear Analysis and Design

The most common current design approach to power amplifiers is to consider them as linear amplifiers [14] and to 'tweak' the output matching circuits in line with measured load-pull data. Although this a relatively fast and efficient way to get circuits designed as discussed in detail in this work, it is not the way to design amplifiers with a high degree of

confidence as to how they will perform or to get the best performance from them. A standard process would be:

1. Obtain small signal S parameters of cell(s) either by measurement or from PDK nonlinear model. Conduct load pull to obtain loads for optimum power or PAE.
2. Investigate matching circuits using S parameter data from (1.), making allowances for the additional information from the load pull. *Problem: we now have a simulation that is not consistent as the load pull data is not an actual part of the simulation.*
3. Develop combining structures as necessary according to the plan from the amplifier architecture. Check physical integration conflicts.

Nonlinear analysis is the subject of much of this work and so will not be examined in more detail here other than to point out that it is highly dependent upon the accuracy of the large signal models and that if these are found to be inaccurate then this has an impact right from the very first stages of the design.

2.5.3 Integration

Having determined the impedance environments that the active devices require and decided upon the combining structures that will be implemented it is necessary to integrate the design into the available space. There are a number of elements that may not have been considered until this stage, such as how bias is to be introduced to the transistors. Some structures described earlier either include convenient bias points, for example shorted stubs, or lend themselves to relatively easy integration of bias (bus-bar). Others require a method of introducing DC currents without impacting on the RF impedance whilst still having a sufficient current capacity. A further problem is that of stabilising the device, particularly with respect to significantly reducing the low frequency gain. Due to size limitations it is always advisable to solve more than one problem with the same elements, so the bias supply lines can also be used to terminate the out of band lower frequencies. This may require resistance as well as capacitance and inductance and there is a compromise to be reached with regards to what can be done on the MMIC itself and what elements should be provided off-chip. It is usually best to split the circuit so that there are some elements on each. Although resistors can be fabricated on the substrate one of the

design aims is to minimise on chip heating, similarly inductors with high current capacity and capacitors with high values take up significant space. Thus a compromise is arrived at often determined by the space available.

As the MMIC layout takes shape it is important that the CAD circuit construction is such that impedances seen by the devices can be observed and continually adjusted so as to maintain performance. Linear circuit optimisers are significantly faster than nonlinear and E-M; however they are less accurate and require a high level of sophistication to model parasitic coupling. It is easy to lose performance, particularly at the band edges, at this stage and it is generally more successful to continually keep the design on track rather than try and optimise it all at the end.

Full EM simulation of MMICs is very complicated due to the number of layers, nonlinear active devices and the close proximity of circuit elements. However an EM simulation of the RF and DC tracking should form a key part of any design. The purpose of including the DC tracks in the EM is not to improve the accuracy of their simulation, but to ensure that any impact that they may have on the RF circuits is taken into account.

The difficulty in designing optimum performance power stages leads to the tendency towards re-use as seen in Figure 2-40. This is a 4 stage design in a series of 1-2-4-8 (number of cells in parallel) and as can be seen the basic balanced pair is repeated 7 times in the MMIC.

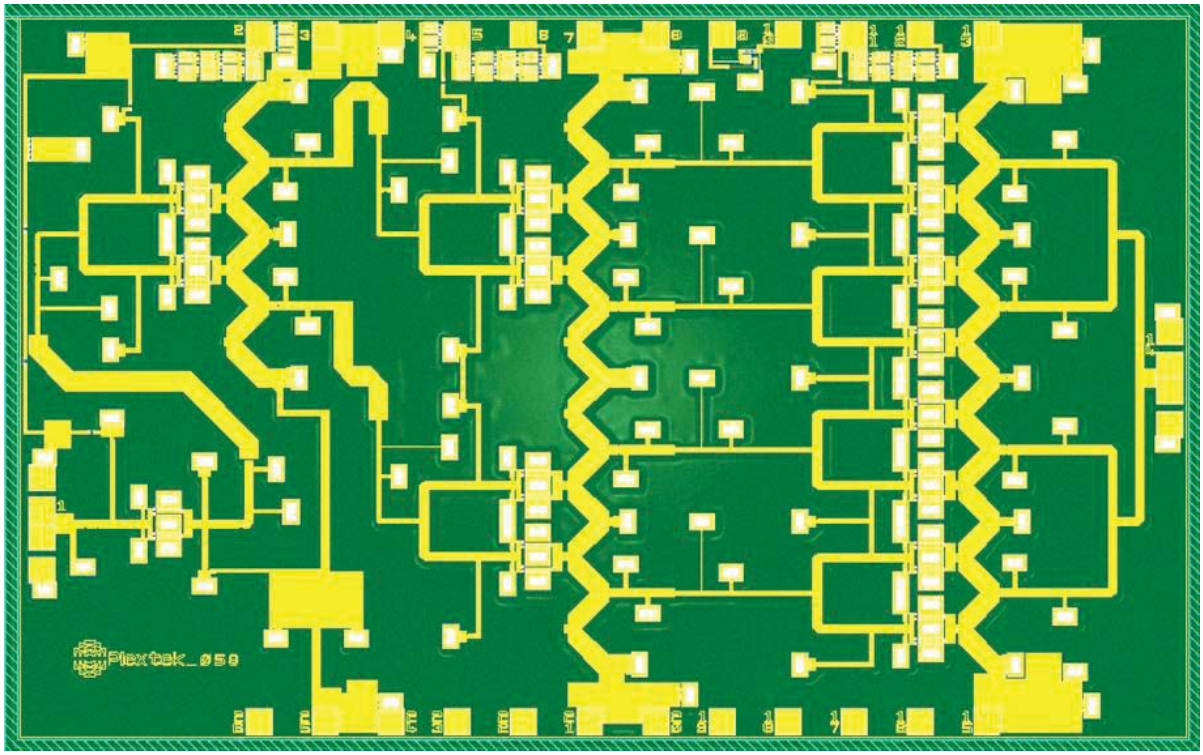


Figure 2-40, 42 GHz PA courtesy of Plextek Ltd.

2.6 Thermal Design

Thermal design is almost as critical as the RF performance in power amplifiers and in some areas the two are diametrically opposed; for high frequency performance we need to keep the device size as small as the required periphery will allow and yet this will increase the power density and hence the junction temperature. Increased temperature degrades ohmic contacts and causes the gate metallisation to migrate into the top of the substrate moving the Schottky junction and reducing maximum channel current.

Device failure is a statistical phenomenon, the likelihood of failure increasing with temperature, thus the maximum 'safe' junction temperature is dependent upon the acceptable failure rate. The GaN Arrhenius plot of Figure 2-41 shows a junction temperature of ~270°C for a MTTF of 10^6 hours, for GaAs this would be between 150 and 175°C. Failure in this case is rarely measured as a catastrophic ceasing in operation but by a degradation in performance below a specified level. Failure due to gate migration is typically observed by a decrease in channel current, so an arbitrary limit may be set to a fall of 20%. Gate migration caused by device self-heating (as opposed to external 'cooking') is actually self-limiting because as the channel current decreases so does the heating effect.

The source of the heat generation is the conduction channel underneath the gate fingers, heat will flow away from this channel in all directions, vertically through the semiconductor to the backing material, horizontally through the semiconductor, along the surface metallisation and down via holes. Thus there are numerous conduction paths each with different thermal resistances.

The measurement of junction temperature is in itself no easy task. Infra-red imaging (Micro Raman Thermography) for example suffers from resolution issues as the gate geometries approach those of the wavelength at infra-red, Figure 2-42, and therefore tend to underestimate the actual junction temperature. These difficulties mean that many people turn to thermal simulation techniques instead, as shown in Figure 2-43.

The structure is broken down into a suitable number of nodes which together form a 3-D grid or mesh. The mesh is programmed to contain the thermal conductivity properties of the various materials used. A number of assumptions are made to simplify the analysis, such as the heat source is uniformly distributed along the active GaN region, the width of the heat source is the same as the gate length and that the length of the heat source is the same as the gate width. It is also assumed that the total power dissipated is shared equally across all individual cells of the device. This assumption is considered valid for DC operation,

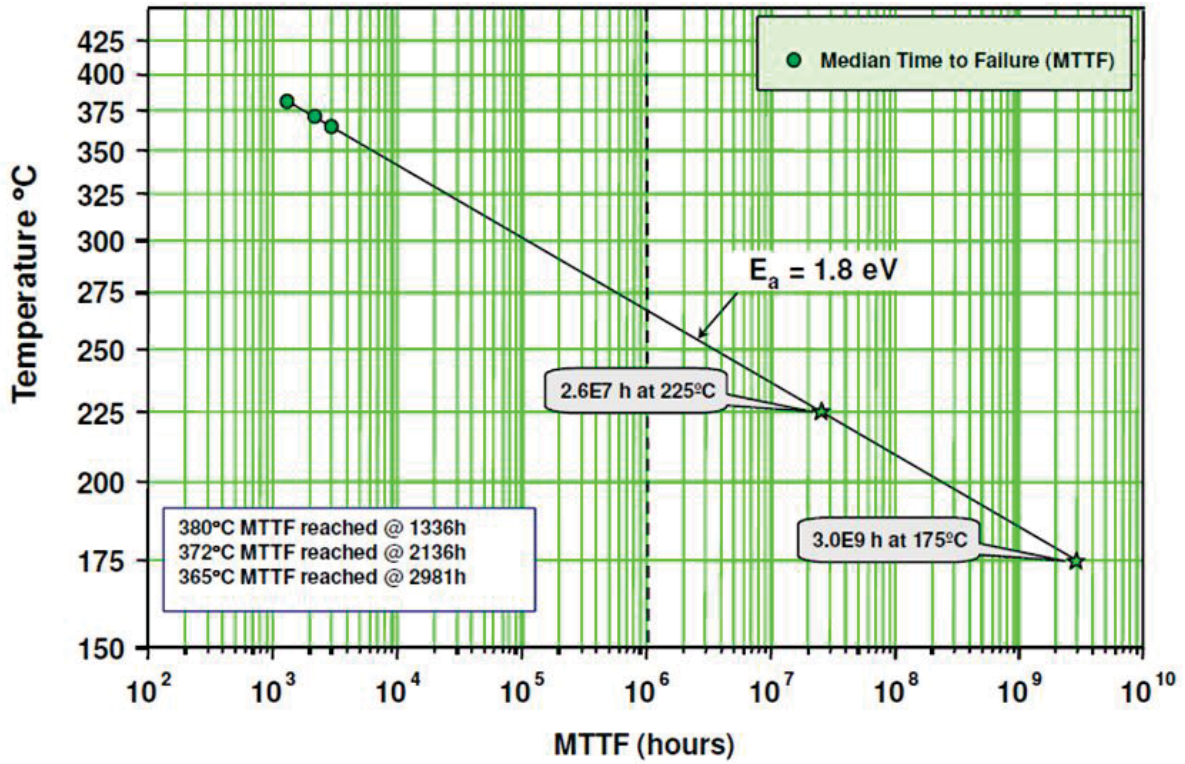


Figure 2-41, Median Time To Failure (MTTF) as a function of junction temperature for GaN (courtesy Cree, IMS 2010 Workshop).

but for RF operation will only apply in a multi-cell transistor when all cells are driven and loaded equally, which is not always the case in a practical amplifier realization. The underside of the carrier is held at a constant temperature. Heat flow is modelled away from the active region via the surface metal as well as through the bulk substrate. By measuring

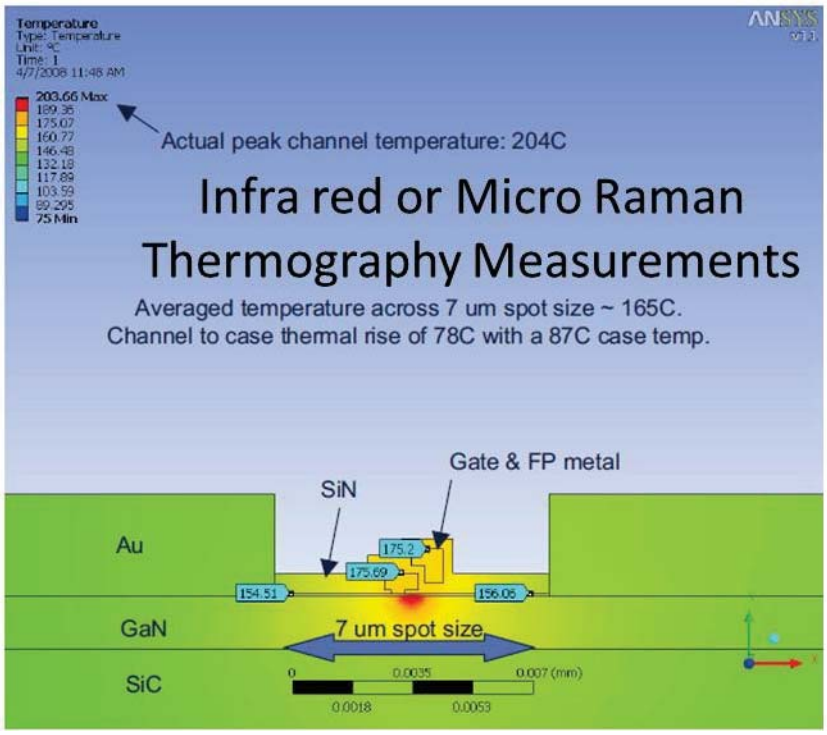


Figure 2-42, Diagram showing problem with infra-red thermal measurements [41].

the actual temperatures at the layer junctions the model thermal resistances can be adjusted so that the simulation matches the real world observations. Measurements on a 5W GaN device derived the thermal resistances shown in Table 2-6, including the overall thermal resistance of 23.1°C/W.

Material	Temperature Rise (°C) (Measured)	Thermal Resistance (°C/W) (Calculated)
Carrier	15.3	3.2
Solder	11.1	2.3
Bulk material	84.4	17.6
Total	110.8	23.1

Table 2-6, Measured temperature rise and calculated thermal resistance.

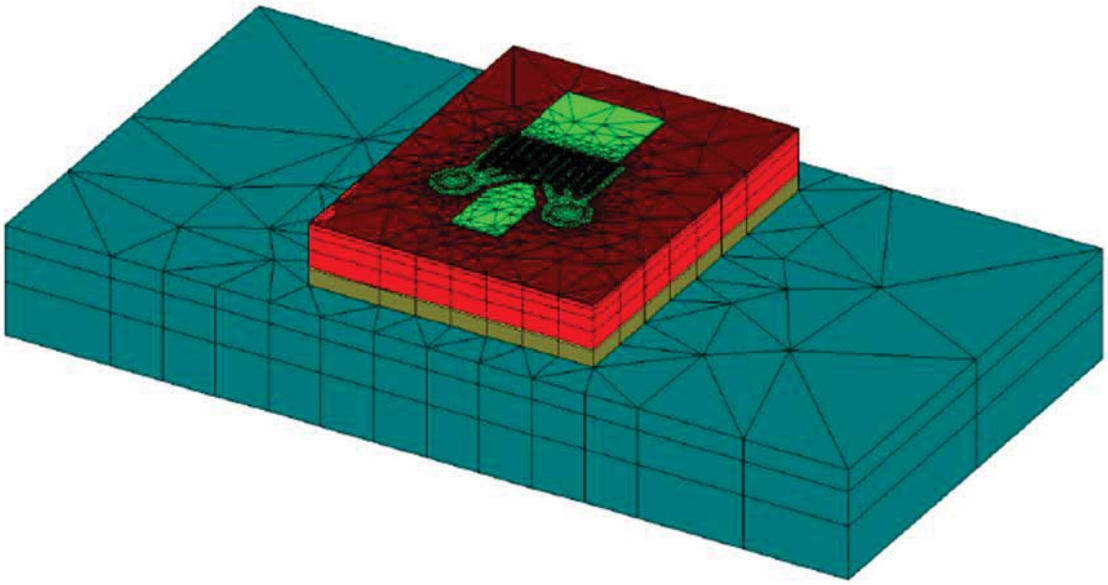


Figure 2-43, Triquint 3D thermal modelling, [41], from the bottom the layers are: metal carrier, AuSn solder, SiC, GaN and surface metallization.

It should be noted that when comparing different foundries, the thermal resistances quoted will depend upon the temperature measurement method(s) employed, (see earlier comment regarding infra-red). Thus when comparing one manufacturers MTTF, such as Figure 2-44 to Figure 2-41, it is necessary not only to ensure that the definition of failure is the same but the channel temperature measurement method.

An additional complexity in the thermal model is due to the fact that thermal conductivity varies with temperature, as shown in Figure 2-45. This is particularly unfortunate for GaN on SiC devices as the bulk of the material is the more variable SiC.

A parameter of the device that can be varied to reduce the thermal resistance is the pitch of the gate fingers. This does however work against the high frequency performance of the device if it results in a significant phase difference across the fingers or if the additional tracking joining the gate and drain fingers introduces too much capacitance. An example of the effect of varying the pitch is shown in Figure 2-46. From this it can be seen that an increase in pitch from 25 to 35µm results in a drop of 2°C/W, hence for a 5W device a 10°C drop in channel temperature, which represents a significant increase in MTTF (the exact amount depending on the actual temperature).

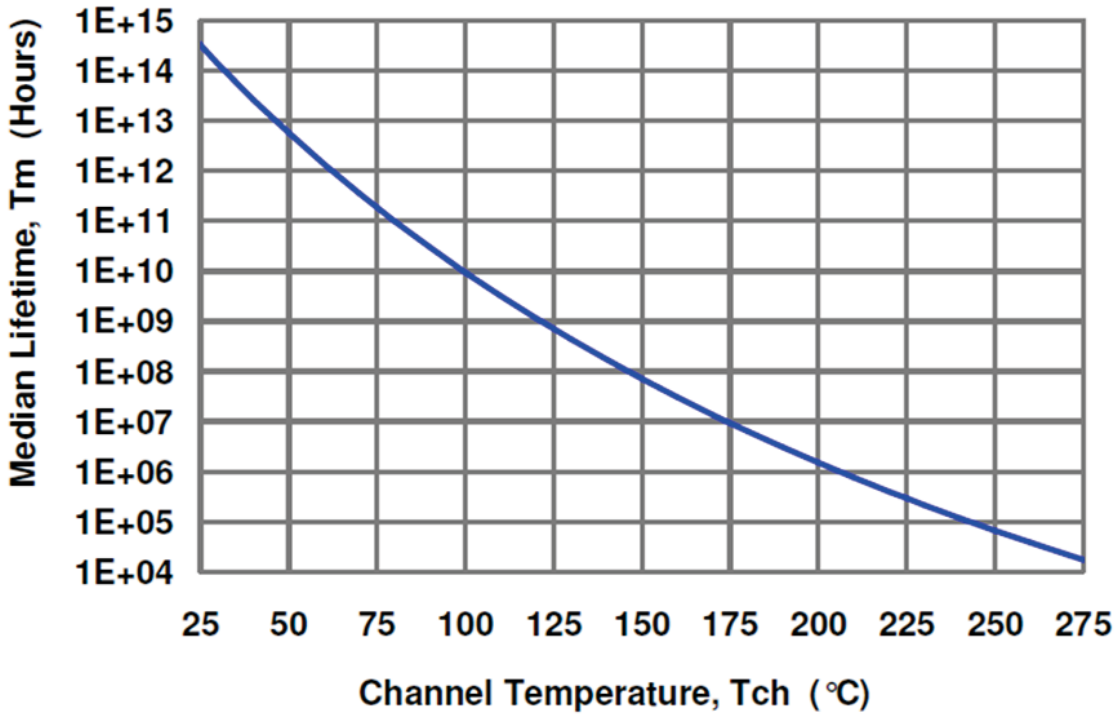


Figure 2-44, TriQuint MTTF versus temperature from [41].

For good simulation accuracy it is necessary to create thermal models for each device

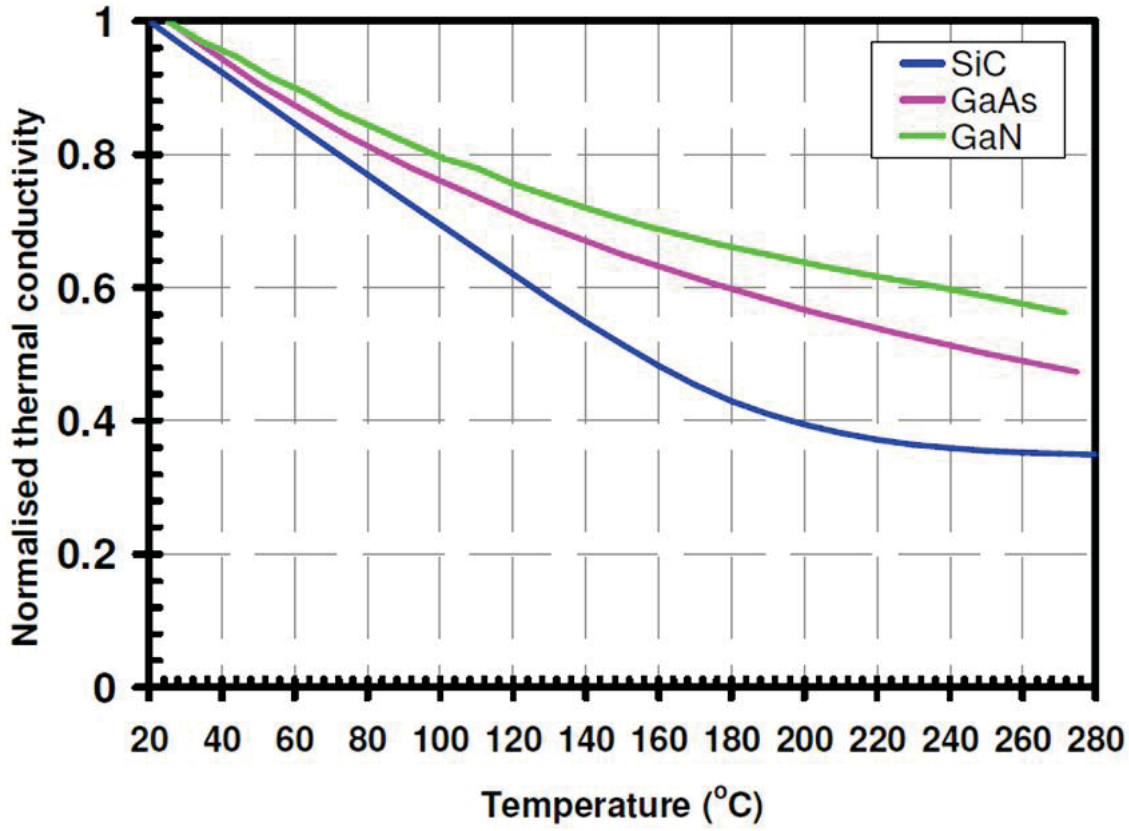


Figure 2-45, Variation in thermal conductivity (normalised to that at 25°C) with temperature, [41].

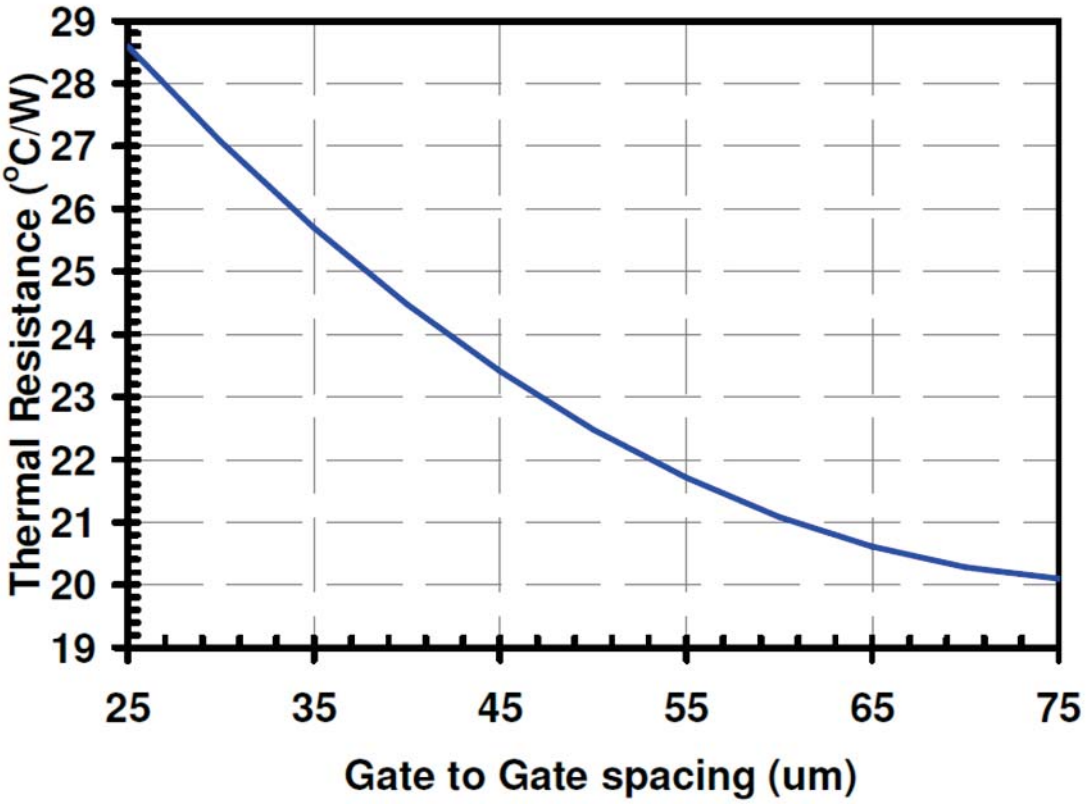


Figure 2-46, Variation of thermal resistance of 0.8mm GaN device with gate pitch, [41].

combination of finger pitch and length as the mutual heating will be different. Models scaled up from 2 finger devices (an approach often adopted on the RF modelling side) will always produce optimistic results for this reason. However this approach is rarely adopted as still many models do not even include self-heating effects.

There are a number of unknowns in the actual circuit (most modelling is conducted just looking at the actual device on its own, not including the matching and DC bias circuits) which add to the ambiguity over the junction temperature:

- Die attach/interfacing manufacturing defects (voids and air gaps).
- Proximity of vias.
- Surrounding surface metallisation.

The position and number of source vias are largely determined by the RF performance considerations. These will push for them to be as close as possible to minimise source inductance – negative feedback which reduces gain, and similarly doubling up on the number if possible. All these will have positive contributions to thermal performance. The only additional aspect that may be considered is via filling. This may have a negligible effect on RF performance due to skin depth, but will make a thermal improvement. The ability of a single via to conduct heat is given by the thermal resistivity, θ_v, {2-7}.

$$\theta_v = \frac{4h}{k\pi\left({d_o}^2 - {d_i}^2\right)} \qquad \{2\text{-}7\}$$

Where h, is the substrate thickness and d_o and d_i are the outer and inner diameters of the via. The constant, k, is dependent upon the conductor (plating) material and for copper is 384 W/m°C. This assumes that the via is unfilled, filling the via will improve the thermal conductivity, (d_i = 0).

Thermal spreading due to the heat source layout can be generally understood by using the concept of 'spreading thermal resistance'. This is inversely proportional to the geometrical size of the heat source and the effective thermal conductivity of the semiconductor and the carrier to which the chip is mounted. This can be expressed in terms of the effective thermal conductivity of the substrate carrier combination, k, and the equivalent thermal diameter of the heat source d_{th}. This latter parameter is proportional to the square root of the area of the heat source, {2-8}.

$$R_{Spreading} = \frac{Constant}{kd_{th}} \qquad \{2\text{-}8\}$$

This produces the perhaps obvious conclusion that to reduce the spreading resistance the area of the device needs to be increased. A further point is that if there is nothing that can be done about the substrate material the emphasis must be on using high thermal conductivity carrier material, unfortunately high thermal conductivity materials tend to have a significant difference in thermal expansion creating stress under temperature and are often difficult to machine and thus produce the very flat surfaces required for a good thermal interface.

In the case of GaN devices the active layer is epitaxially grown on a substrate. For most power applications this is SiC, but there are a few foundries that have opted for Si. The backside of the substrate is metallised to aid solder attachment to the carrier. Although a great deal of progress has been made in the development of high thermal conductivity epoxies, to such an extent that some GaAs HPA device suppliers were recommending their use, they still do not achieve the performance of properly applied eutectic solder (>10W/m K) and so their use for GaN HPAs cannot be recommended. Achieving a uniform die attachment, i.e. void free and of even minimal thickness, is critical for the optimum performance of the MMIC. Voids have very low thermal conductivity and if they happen to

occur directly under the active device can be fatal. They cannot be spotted by visual inspection and require specialist X-ray or ultrasonic equipment to image under the chip.

It has been said above that "if there is nothing that can be done about the about the substrate material…" and yet from Table 2-6 we can see that this is where the significant temperature rise occurs. There are in fact a number of possible ways to reduce this although they may not be available yet or at all foundries.

Diamond has a thermal conductivity of 500-2000 W/m K, compared to 460 for SiC (6H). Thus when we can commercially grow diamond crystals large enough this could be used as a carrier substrate instead of SiC. But we would need to hit the higher thermal conductivity number to make this practical. Note there is work being done on copper diamond combinations for the carrier, but although every little helps the largest proportion of the temperature rise comes in the bulk substrate. A solution employed in GaAs offers no benefit in SiC substrates but is mentioned as it may be appropriate for GaN on Si. Blind vias in the backing substrate only under the active areas and then plated up with copper thus creating heat conducting pillars under the devices, (copper 390 W/m K cf. 148 W/m K for silicon). These can also take the shape of inverted troughs and silicon processing is far more advanced than other substrates which make the technique more practical. A final approach which has been employed commercially in HBT devices where a major cause of failure is thermal runaway in the emitters; is to place a thermal shunt on the top surface of the chip as shown in Figure 2-47. These shunts are made from high conductivity materials sometimes incorporating diamond films. Obviously they can only be connected to the grounded

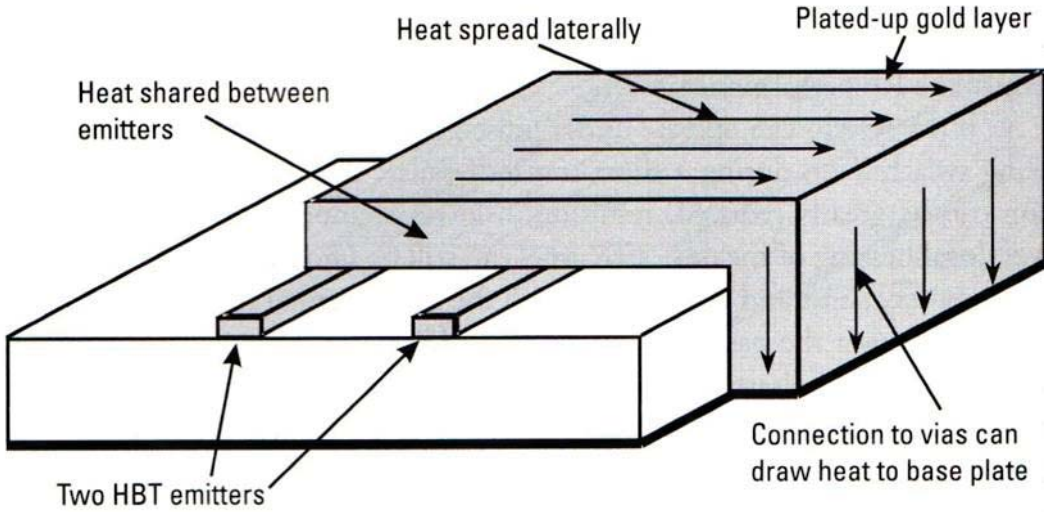
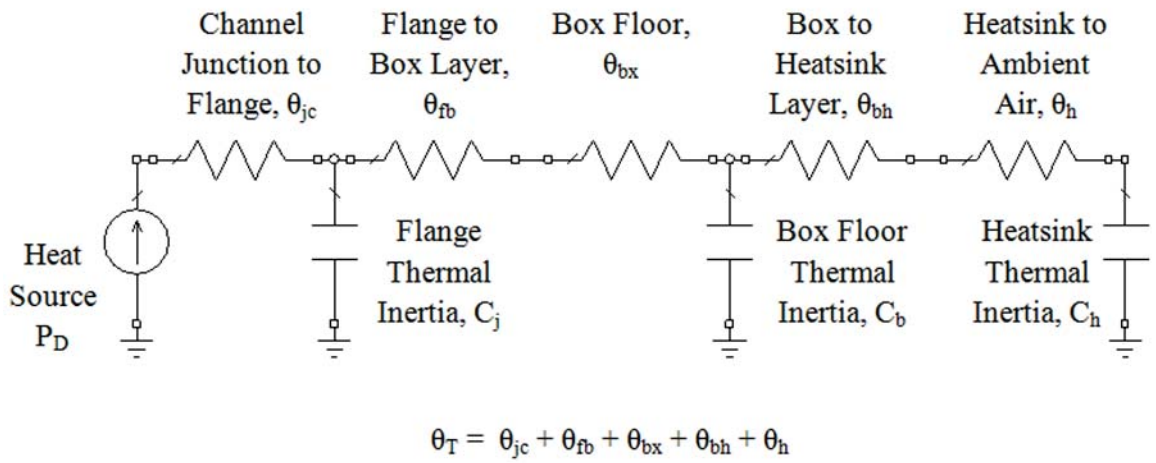


Figure 2-47, Thermal shunt implemented on HBT device, [14].



$$\theta_T = \theta_{jc} + \theta_{fb} + \theta_{bx} + \theta_{bh} + \theta_h$$

Figure 2-48, Electrical analogue of thermal resistance and inertia.

terminals (sources in GaN HEMTs) and care must be taken to ensure that additional capacitance is not added to the structure having a detrimental effect upon frequency performance. Referring right back to one of the early drawings, Figure 2-6, it is wondered whether such a shunt could be incorporated with the source field plate?

Sophisticated thermal modelling software will not always be available, but designers still need to have some guide to the thermal behaviour of their circuits. Some basic assumptions can allow crude but effective models. Under steady state conditions the resistance to heat flow is a product of two factors, the intrinsic Thermal Resistance (TR) of the material and the interface with the next layer. Initially there is also Thermal Inertia, (TI) which may be important in pulsed amplifiers, as gain and output power are proportional to the channel temperature. The thermal components can be represented by electrical analogues, resistors for TR and capacitors for TI. For example, consider a MMIC mounted on a carrier, which is bolted into a housing which in turn is bolted to a heatsink. In this case there is the TR of the device channel to carrier, θ_{jc}, that of the layer between the carrier and the box floor, θ_{fb}, the box material itself, θ_{bx}, the box heatsink junction, θ_{bh}, and finally that of the heatsink (assumed to be either in still air or a fixed air flow), θ_h. All of these TRs add to give a net θ_T. These can be summarised in a 'thermal circuit diagram' as shown in Figure 2-48.

A note of caution, the thermal resistance is not constant, it is proportional to the temperature difference, the greater the difference the greater the heat flow. Also as has been said earlier thermal resistance changes at different rates for different materials, Figure 2-45. The TI of the interface layers is typically very small and is ignored. The temperature

differential between the device channel and the heatsink is the dissipated power P_d, times θ_T. The TI can be calculated by observation of the actual temperature rise profile, which is described by {2-9}, where t is the time. A typical response is shown in Figure 2-49.

$$T_{rise} = \theta_{jc} P_d \left(1 - e^{-t/\theta_{jc} C_j}\right) \qquad \{2\text{-}9\}$$

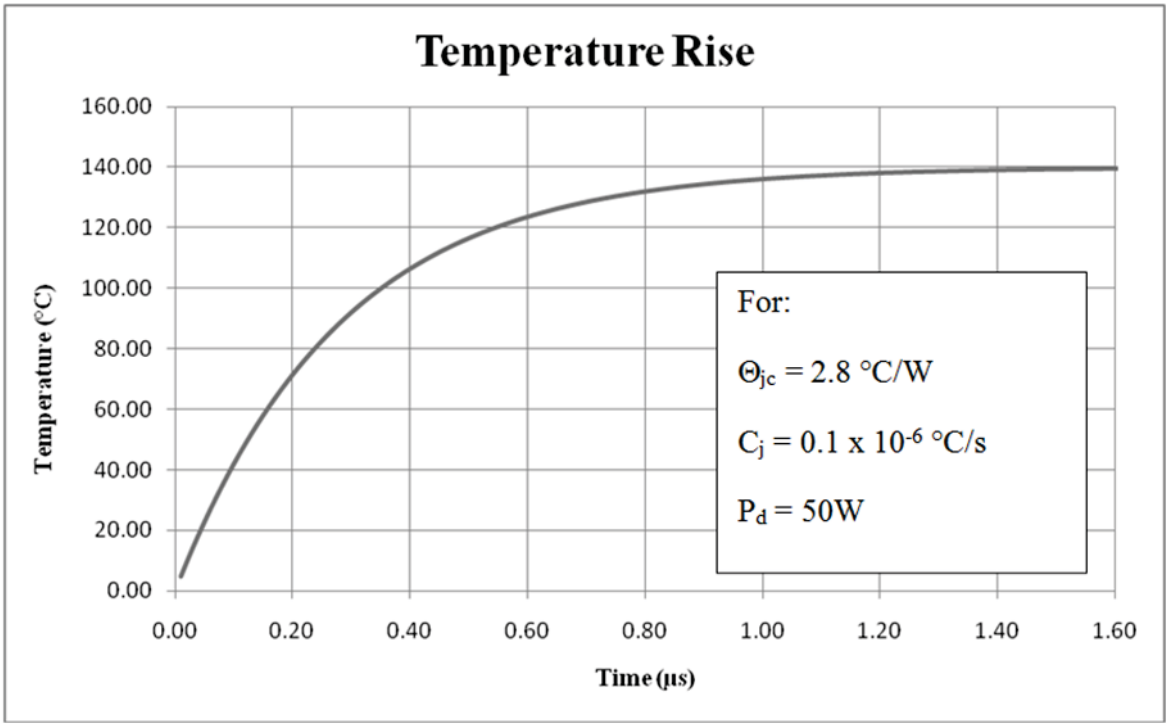


Figure 2-49, Example thermal profile showing effect of thermal inertia.

Whilst an obvious conclusion to draw would be to suggest that all materials under the active device area should be as thin as possible, this is not quite the case; in order for the heatsink to operate most effectively the temperature must be given the opportunity to 'spread', otherwise only a limited portion of the heatsink will be effective in removing the heat. This is also important in estimating how close power MMICs can be mounted to each other without having a deleterious effect. Rather than consider each individual active device for simplicity they can be considered as one, but having the heat dissipation of the sum of the individuals. The other assumption is that the bulk substrate is uniformly 'filled' with heat such that the whole of the underside is at a single temperature. In practice this may only be the case in one dimension, (across the width not along the length) output devices typically being arranged in rows, however for the purposes of this analysis it is acceptable. In terms

of heat transfer it is recommended that the heat sources be separated such that the heat 'illuminates' an area defined by a 70° angle as shown in Figure 2-50.

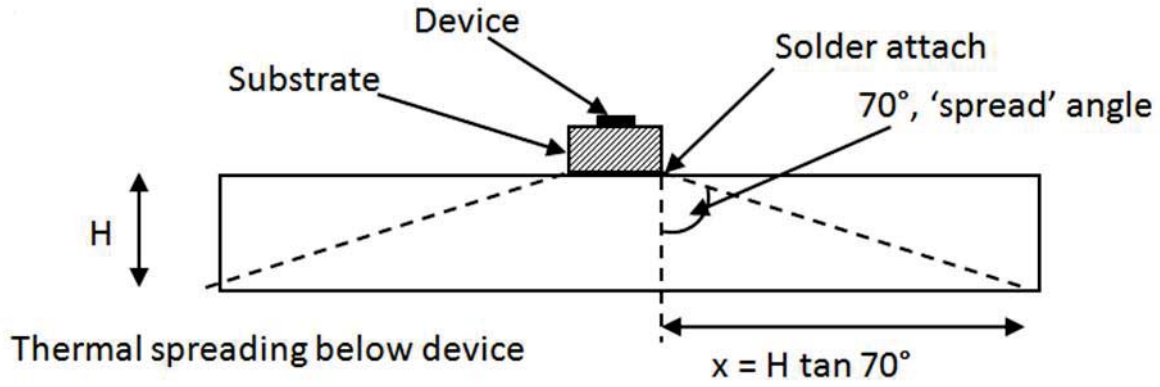


Figure 2-50, Device heat spreading (not drawn to scale).

Before leaving the topic of thermal design it is worth reflecting one last time on the effects of temperature are on RF performance. As has been mentioned some manufacturers are including self-heating functions within their models (e.g. Cree) and so the effect can be observed. Others have recorded the impact of increased junction temperature, [20] and measurements on a GaN HFET are shown in Figure 2-51, (note the x-axis is incorrectly labelled and should read "Temperature (°C)"). These effects are frequently observed when testing actual devices and comparing the results against data sheet values. In many cases a CW output power level is quoted, however the measurements have been conducted pulsed. This keeps the junction temperature low and the average power level down (relevant to

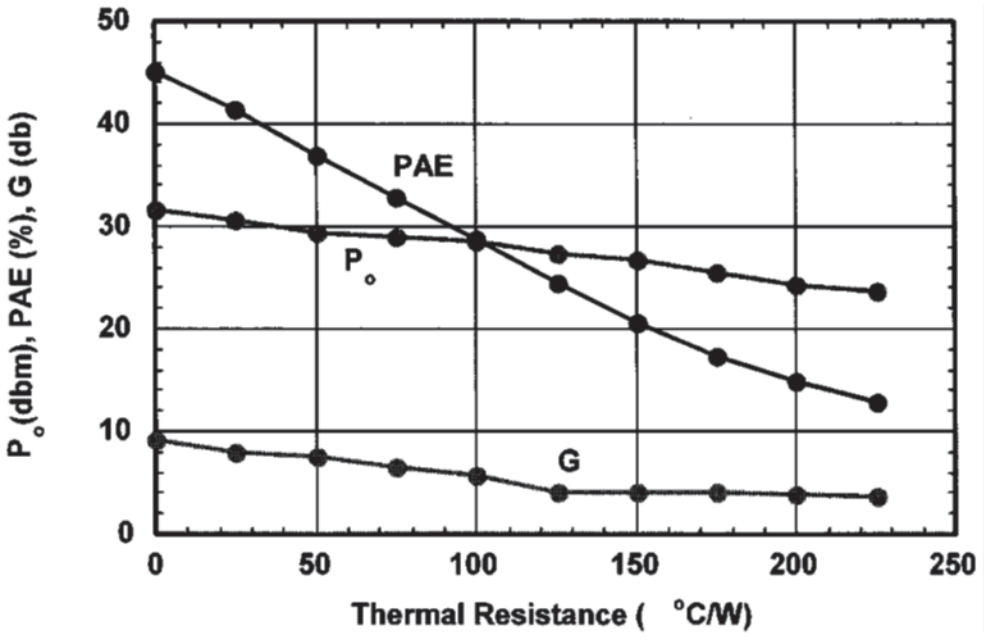


Figure 2-51, Thermal performance of GaN HEMT on sapphire, (x-axis should read Temperature) [20].

power limits on test system components such as attenuators and also enables testing to be carried out on wafer). The CW results on the data sheet are probably obtainable, if the channel temperature can be kept as low as in the pulsed case.

Clearly the ideal case is to have the device nonlinear model accurately imitate the thermal effects of the mounting and construction that will be used in the actual final application. This is not something that can be done by the device manufacturer and therefore emphasises the need for designers to be able to access to such a modelling capability, i.e. base the model on measurements made under conditions replicating the actual operating environment.

2.7 Specific Considerations for Broadband GaN MMIC design

Matching techniques have already been discussed in chapter 1. This section will briefly describe some of the considerations specific to their utilisation in GaN amplifiers.

One of the factors limiting the current handling capability of conductors is transport of material caused by the gradual movement of ions in the material. When the current density approaches 10^6 A/cm^2 [6] there is a tendency for the material to move in the direction of current flow. In thin conductors electro-migration induced damage usually occurs in the form of voids and hillocks in the metal due to the depletion and accumulation of metal grains due to heavy flow of electrons. This also occurs in transistor gates, drain and source pads and ohmic contacts. The effect of electro-migration becomes more pronounced as the temperature rises. Voids obviously increase resistance, whilst bulges can cause shorts over time. In GaN we are likely to operate at higher temperatures with increased power densities, hence this problem will get worse.

For gold conductors on GaAs the safe maximum current density is 2.22 x 10^5 A/cm^2, (chapter 8 of [6]) and from this the electro-migration requirements dictate the microstrip and inductor line widths; 10 mA/µm for a gold thickness of 4.5 µm. By increasing the metal thickness we can increase the width proportionately. Thus, for example for the output bus-bar as in Figure 2-38, if the maximum current is 3A and maximum plating thickness is 9 µm, then the bus-bar must have a minimum width of 150 µm. Although the thermal conductivity of GaN is 3x higher than that of GaAs and SiC is over 8x higher, the power

density is also much higher and the areas where damage is most likely to occur is around the actual device where geometries are smallest and the temperature highest.

Although the foundry will most likely have taken care to ensure that components such as MIM capacitors have sufficiently high breakdown voltages compared to the active

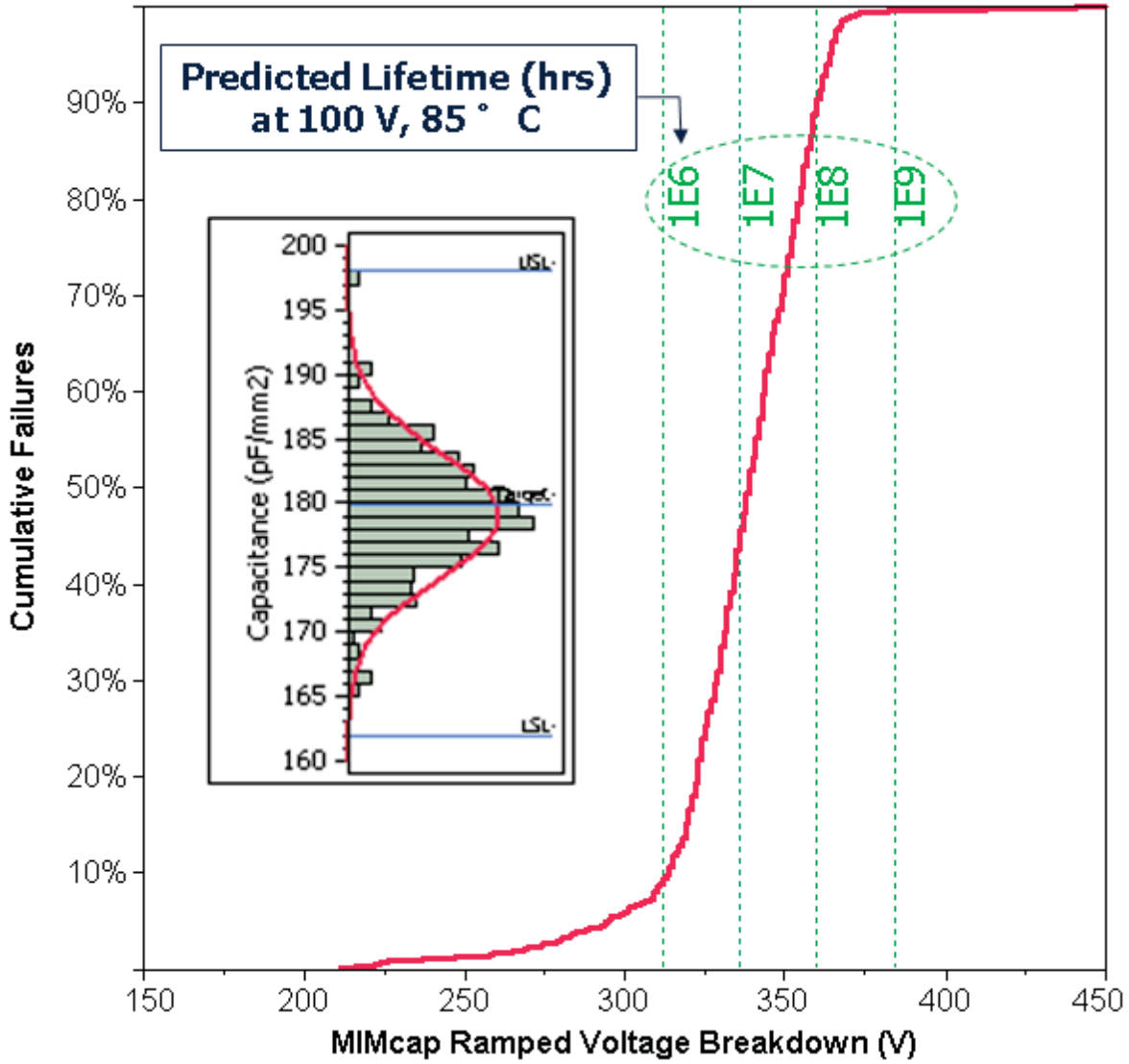


Figure 2-52, MIM capacitor ramped voltage breakdown distribution. Dashed green lines are predicted lifetime at 100V and 85 °C. [42].

devices, care and consideration should be taken especially with high efficiency amplifiers with harmonic terminations. For example a class J amplifier biased at 28V could produce a peak voltage of ~75V. In the Cree MMIC process reliability material it states that MIM capacitors have an anticipated lifetime of 10^6 hours at 85°C and 100V operation. Thus whilst we are still operating within a safe area, the reliability times are now approaching that of the transistors themselves. As the operating voltage of the devices increases, there will need to be a proportional increase in passive component breakdown voltages. Although in

some designs it is possible to physically place components away from the heat generating areas (active region) in other circuits, such as the parallel RC network discussed in chapter 1 on input matching, or capacitive compensation on the gates and drains of DPA, capacitors are deliberately placed as close to the device as possible. In the case of gate capacitors the voltages are lower; however the capacitor is directly connected to the gate finger metallisation, which is located over the hottest region of the circuit with temperatures typically in excess of 175°C.

The observed change in channel resistance due to gate migration into the channel has already been commented on. This has been observed in a number of GaN devices as a function of input drive, which as the devices were operated in a 10% I_{DS}, class A/B, is actually the result of channel heating. The effects can be hidden however if the devices are still subject to traps, as these can cause a temporary change to the gate voltage - drain current relationship. The other problem is that gate migration can take many hundreds or even thousands of hours to drift to the self-limiting point. This will cause an obvious problem as output power is proportional to current, the power is also reducing with time. Hence an amplifier which may have had a dB in hand above the specification will move towards the limit with time, the rate being dependent on operating temperature. It may be possible to compensate for this to some degree by adjusting the gate voltage, which requires a more complicated bias circuit. It should also be noted that not all devices display this behaviour and that it will depend upon the operating mode; amplifiers operating with digitally modulated signals with high peak to average power ratios will be operated backed off and only see the higher powers for a very small percentage of the time.

Similarly an application area where GaN seems highly suited is radar. The low duty cycle but high peak powers should enable GaN to be operated where its performance is best, but with fewer thermal problems. There is however a tendency of engineers to add more power until we get back to the same temperature limiting scenario.

In summary, the key advantages of GaN to broadband amplifier design are:

1. Reduced pF/mm/W, i.e. for the same output power device (in Si or GaAs) we have a reduced periphery and therefore output capacitance. This is more important to broad band matching than the higher impedance. Note that this also applies to the input.

2. Higher G_m – gain. The number of stages that are used in a broad band design is determined by the gain per stage. For interstage wide band matching lossy networks are often employed higher gain means that we can achieve a wider bandwidth.
3. Higher breakdown voltage, safely use large voltage swing topologies such as class J, but also have a more robust device, again including on the input producing higher dynamic range and input power tolerant low noise amplifiers.
4. Higher drain resistance, eases the problem matching.
5. Smaller size for the same power allows more devices to be packed into the area or chip size to be reduced.

The current higher cost of GaN will not be the case long term as more foundries producing larger wafers come online. RFMD are working on the assumption that GaN will be cheaper than GaAs ($/W) in the long term, based upon their learning curve with GaAs, Figure 2-53, and volume increases, Figure 2-54. The most significant delay in bringing GaN MMICs into production will be down to how quickly engineers can turn around reliable designs, and for this they need a reliable, repeatable processes and good nonlinear models.

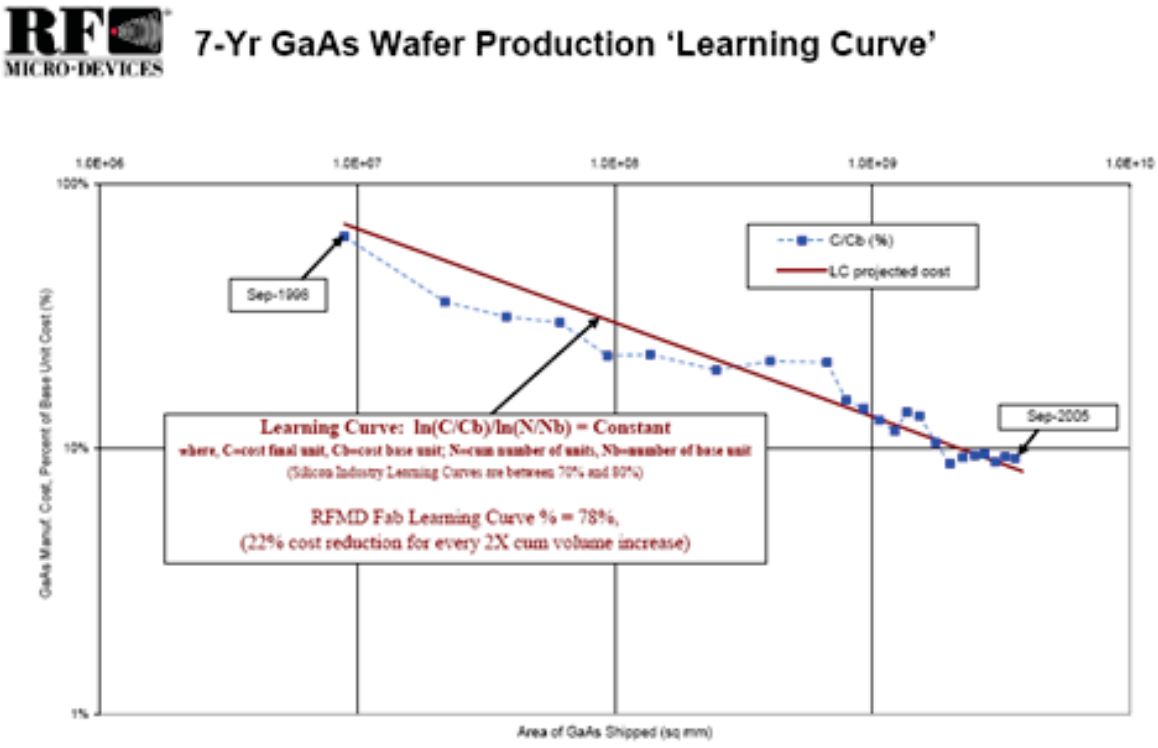


Figure 2-53, GaAs wafer price decrease as experience grows, [43] , GaAs market cost reference to baseline against number of wafers manufactured.

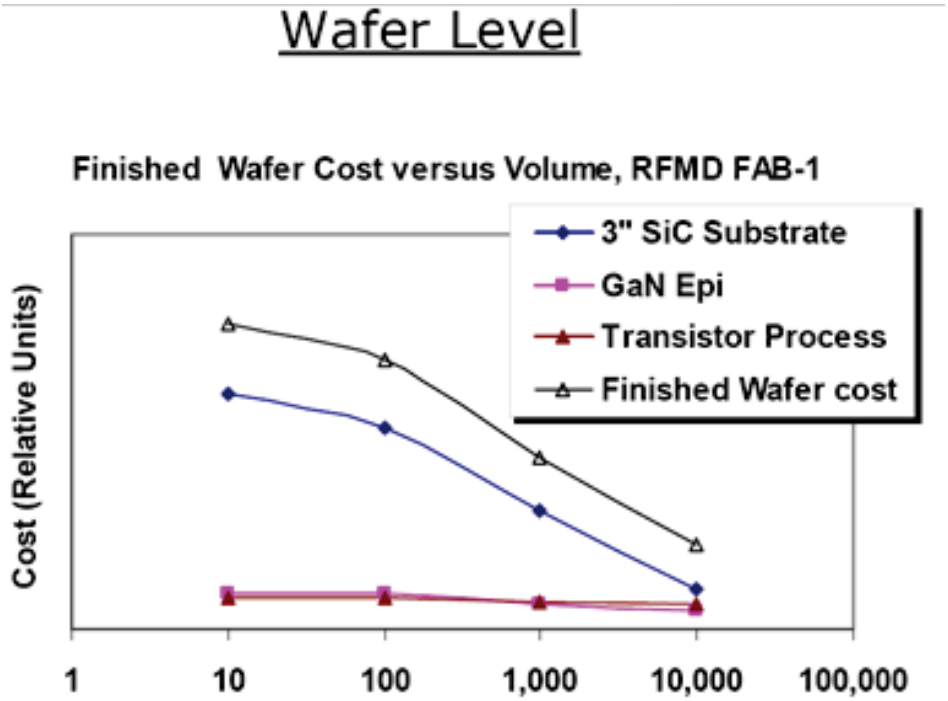


Figure 2-54, RFMD earlier years GaN wafer cost/volume profile, [43].

2.8 Performance of wideband MMIC designs

A selection of mmic GaAs, Table 2-7, and GaN, Table 2-8, based broadband power designs has been reviewed. What appears clear is that for designs approaching an octave and higher, achieving a PAE above 35% in GaAs and 40% in GaN is extremely difficult.

Ref.	Band	Power	PAE	Linear Gain (ripple)	No. of stages (devices) periphery	Type	Comment
Yu [21]	6-18	>31	>24	>11 (3)	2 (12) 6mm	Matched	Designed for balanced oper.
Komiak [22]	6-18	>36	22 av.	>12 (2.7)	2 (20) 16mm	Matched	Tested pulsed.
Bahl [23]	2-8	>37.5	>16	>12.5 (4)	2 (24) 22.6mm	Class AB LLM	LP & SS model
Bahl [24]	5-8.5	>33	>31	>17 (4)	2 (10) 6.25mm	LLM	LP SS s params
Bahl [25]	0.7-2.7	>41	>22	>20 (7.5)	2 (20) 40mm	LLM	Comments on need for models
Arell [26]	2-6	>30	>22	>18 (2)	2 (3) 4.2mm	All-pass IP matching	1992, simplified model matching.

Ezzeddine [27]	0.03-2.5	>33	>20	>20 (2)	2 (8) 22.4mm	Hi FET	DC & RF FB
van der Bent [28]	8.25-10.25	>40.3	>41.4	>19 (1-sat)	2 (20) 19.2mm	pHEMT	Harmonic Tuning
Bosch [29]	8.5-11.5	>39	>35	>19 (2-sat)	2 (20) 19.2mm	pHEMT	No mention of Harm. Tuning

Table 2-7, Sample of wideband GaAs based power amplifier designs.

Ref.	**Band**	**Power**	**PAE**	**Linear Gain (ripple)**	**No. of stages (devices) periphery**	**Type**	**Comment**
Kobayashi [30]	1-4	>34	>36	>12 (8)	1 0.8mm	Feedback MMIC	LNA/PA
Sim [31]	0.5 – 2.5	42	>30	>11 (2)	1 CGH40025F	Matched Hybrid	Based on LP
Krishna-murthy [32]	0.1-2.0	>43	>39	>10 (3)	1 2.2mm	Matched	Packaged Unit Cell.
Krishna-murthy [33]	0.5-2.5	>39.5	>39	>14.8 (3)	1 2.2mm	RLC input	Off chip op match, LC
Gassman [34]	2-15	>37	>20	>10 (5)	1 (5) 2.0mm	NDPA	
Lin [35]	0.02-3	>37	>20	>43 (3)	3 (5) 3.6mm	FB + NDPA	Hybrid
Masuda [36]	6 - 18	>40	>18	>8 (10)	1 (14) 4.2mm	NDPA + Lange	Pulsed meas.
Mougino t [37]	6 - 18	>41.1	>15	>18 (6)	3 (7) 4.2mm		Based on LP data

Table 2-8, Sample of wideband GaN based power amplifier designs.

Note that we have a useful comparison between van der Bent [28] and Bosch [29] in the effect of harmonic tuning. Although the difference in peak PAE is only ~5% (average of harmonic tuned PAE performance) the difference at 10GHz is closer to 15%, Figure 2-55 and Figure 2-56.

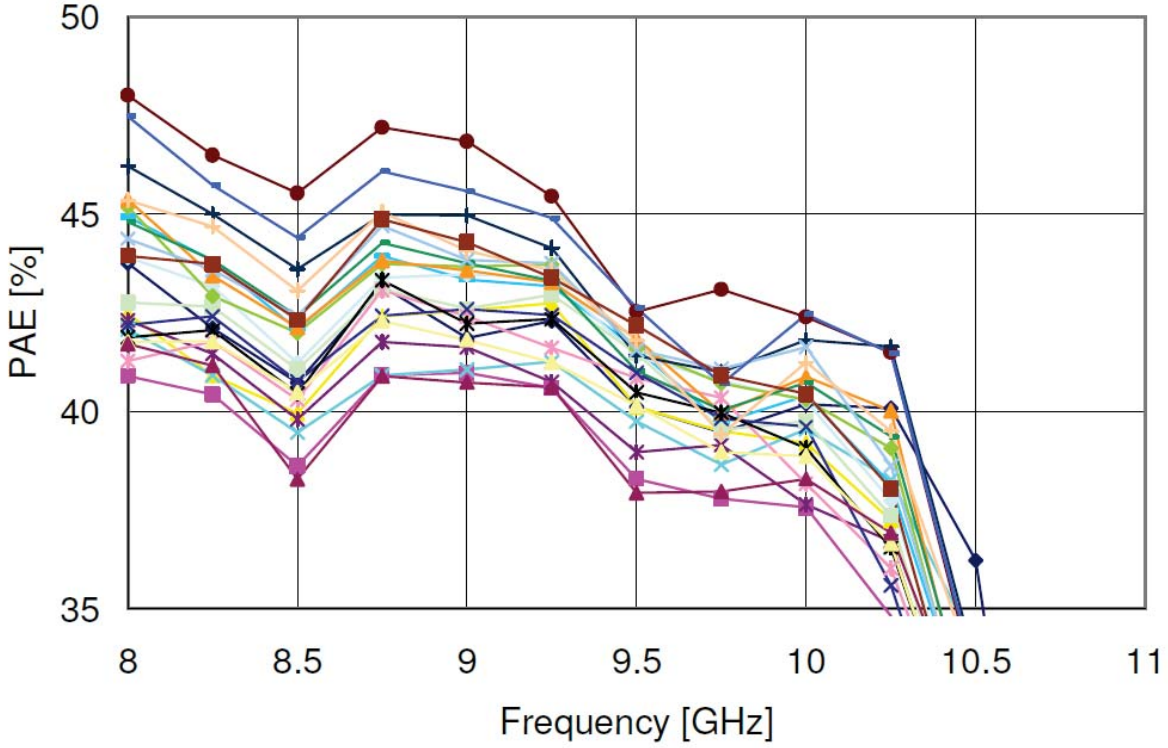


Figure 2-55, 10W X Band MMIC with harmonic termination, PAE [28].

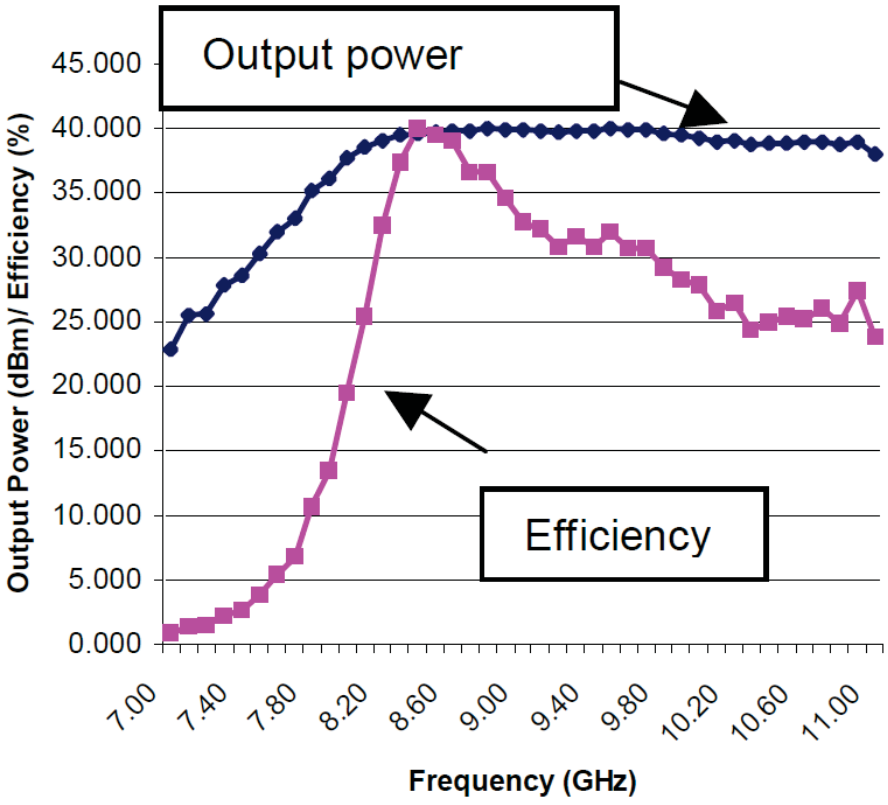
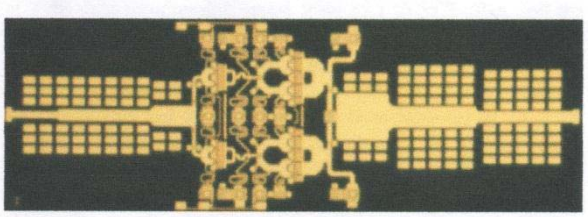
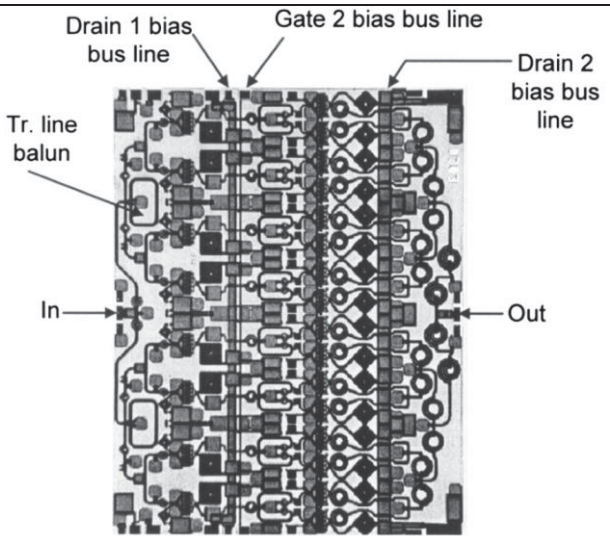
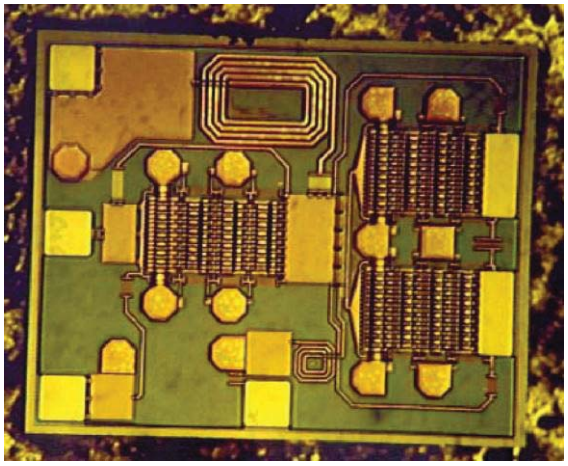
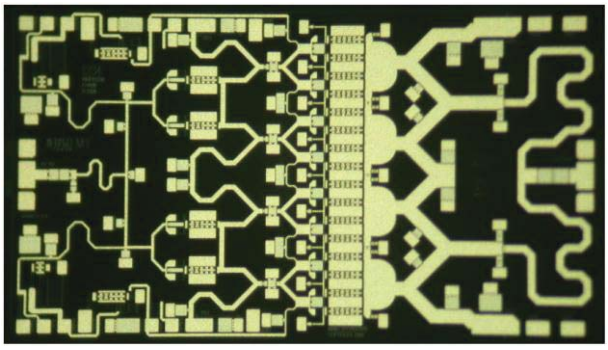


Figure 2-56, 10W X Band MMIC, PAE [29].

Yu [21] GaAs pHEMT, 0.35µm process, 60µm finger width. 300mA/mm, V_{BRDS} of 15V Multi-section impedance transformer matching. Chip size, 2.7 x 8.1mm.	
Bahl [23] GaAs, M/A Com Multifunction Self Aligned Gate (MSAG), 460mA/mm, f_T of 21GHz, V_{BRDS} of 20V. Repeated 10 fingered 94µm cell, 30µm finger spacing, thermal resistance ~90°C/W. Class AB bias, ~25%I_{DSS}. Chip size, 5.0 x 6.3mm.	
Ezzeddine [27] GaAs Hi Voltage FET: 'Stacking' the FETs increases the operating voltage, output power, gain and bandwidth: Where N is the number of series connected devices – Drain voltage = N x Vd Output impedance = N x Zin Gain increase = 10LogN Power increase = 10LogN Chip size, 2.23 x 1.82 mm Foundry process M/A Com 0.5µm GaAs	
van der Bent [28] 0.25µm GaAs pHEMT f_T of 60GHz, V_{BRDS} of 18V. Model optimised to measured load pull data. Class AB operation with harmonic tuning consisting of shunt capacitor on drain. Similar harmonic tuning on gate. Chip size, 4.41 x 2.50 mm	
(Table continued following page)	

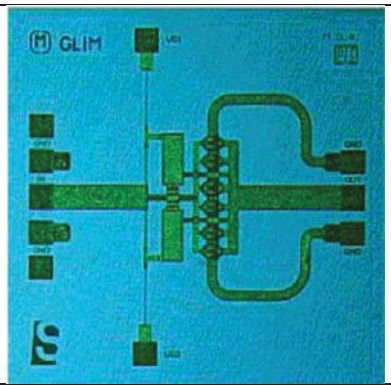
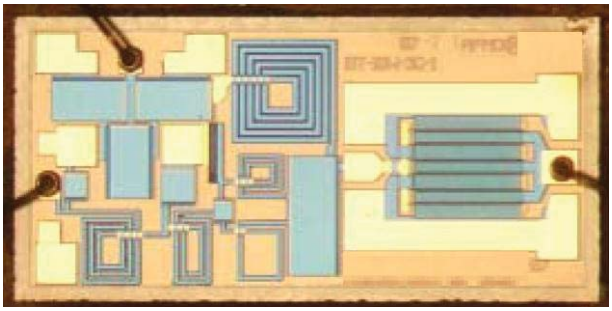
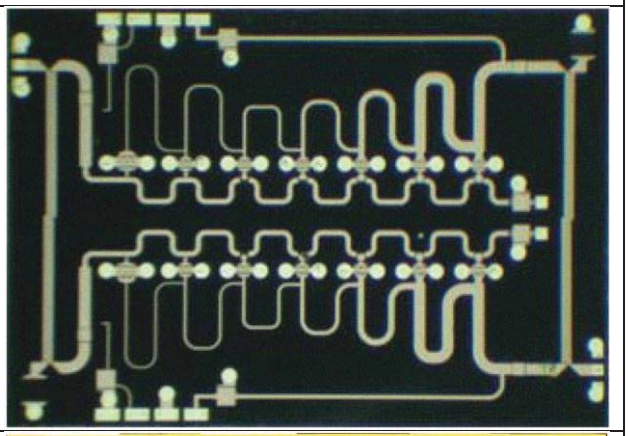
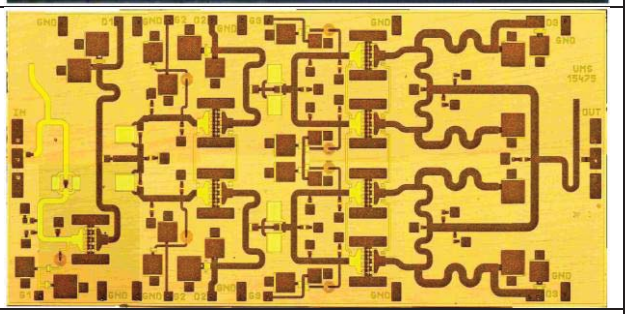
Kobayashi [30] 0.2µm GaN on SiC, 350mA/mm, f_T of ~75GHz, V_{BRDS} >60V. 1.7x1.7mm die. Series and shunt feedback; series L for noise figure matching to 50Ω, RC shunt feedback optimised for gain bandwidth, stability and output match.	
Krishnamurthy [33] Input is a 4:1 two stage LC impedance transformer, then RLC all pass network – lossy matching. Very compact with good gain flatness and input match. Off chip output matching, series L shunt C, note capacitor can therefore be high Q low loss. 0.5µm GaN on SiC process, 250mA/mm, f_T of 10.5GHz, V_{BRDS} >180V. 2x1mm die.	
Masuda [36] Non-uniform distributed line attempts to improve the power performance compared to conventional DPA. Lange coupler combining on chip. Biased at 40V. Measurements made at 10% duty using 10µs pulse. 0.25µm GaN on SiC process, f_T of 21GHz, 5.2x3.6mm die.	
Mouginot [37] UMS 0.25µm GaN on SiC process, based on 8x75µm cells, 1:2:4 corporate structure. 25V operating voltage. 6.43x3.08mm die.	

Table 2-9, Selection of amplifiers from Table 2-7 and Table 2-8, chosen to represent the different design approaches.

Most of the papers reviewed use a mixture of the Cripps Loadline method and small signal S parameters to design the output matching. There are a number of different approaches to achieving wide bandwidth operation; however ultimately they are all dependent upon the output matching that can be achieved from the output stage cell. A summary of the approaches are shown in Table 2-9. Distributed amplifiers clearly offer the

greatest potential in terms of bandwidth and they [36] appear to compare well with the more conventional approach taken [37], however the measurements on the NDPA were made pulsed (presumably as they were on wafer) whilst the UMS devices were mounted on a carrier and measured CW. Typically it is presumed that distributed amplifiers will have comparatively lower PAE as power is lost in the reverse termination. Despite its relatively new status the performance of the GaN based devices clearly exceeds that of GaAs, particularly when the output stage consists of a single device [33]. Among the GaAs based devices the high voltage FET [27] approach appears to offer considerable promise, however the technique has yet to see significant commercial success and so it is possible that there may be other issues associated with such an approach (e.g. breakdown voltage of passive components).

2.9 Bibliography

[1] M. van Heifningen and F. van Vliet, “Design and analysis of a 34dBm Ka Band GaN high power amplifier MMIC,” in *1st European Microwave Integrated Circuits Conference*, Manchester, 2006.

[2] D. Nidhi, D. Brown and S. Keller, “T-gate technology for N-polar GaN based self-aligned MIS-MEMTs with state of the art fmax of 127GHz: Pathway towards scaling to 30nm GaN HEMTs,” in *Device Research Conference*, 2010.

[3] E. Kuwata, K. Yamanaka and H. Koyama, “C-Ku band ultra-broadband GaN MMIC amplifier with 20W output power.,” in *Asia-Pacific Microwave Conference*, 2011.

[4] K. Krishnamurthy, J. Martin, B. Landberg and e. al, “Wideband 400W pulsed power GaN HEMT Amplifiers,” in *IEEE MTT-S Microwave Symposium Digest*, 2008.

[5] D. Denninghoff, S. Dasgupta, D. Brown and e. al, “N-polar GaN HEMTs with fmax > 300 GHz using high aspect ratio T-gate design,” in *69th Annual Device Research Conference*, 2011.

[6] J. Walker, The Handbook of RF and Microwave Power Amplifiers, Cambridge: Cambridge University Press, 2011.

[7] A. Ketterson, “Characterisation of InGaAs/AlGaAs pseudoorphic modulation doped field

effect transistors.," *IEEE Transactions of electronic devices,* vol. 30, no. 10, pp. 564-571, 1986.

[8] S. Cripps, "A Method for the Prediction of Load-Pull Power Contours in GaAs MESFETs," in *International Microwave Symposium*, 1983.

[9] J. Lange, "Interdigital stripline quadrature hybrid," *IEEE Transactions on Microwave Theory and Techniques,* vol. 17, pp. 1150-1151, December 1969.

[10] H. Louw and J. Nortier, "Cascaded Lange couplers," *Microwave Journal,* p. 1989, November.

[11] Y. Ayasl, L. Reynolds, R. Mozzi and L. Hanes, "2-20 GHz GaAs Travelling wave power amplifier," *IEEE Transactions on Microwave Theory and Techniques,* vol. 32, no. 3, pp. 290-295, March 1984.

[12] J. Beyer and et.al., "MESFET Distrbuted Amplifier guidelines," *IEEE Transactions on Microwave Theory and Techniques,* vol. 32, no. 3, pp. 268-275, 1984.

[13] B. Virdee, A. Virdee and B. Banyamin, Broadband Microwave Amplifiers, Artech House, 2004, pp. 133-135.

[14] S. Marsh, Practical MMIC Design, London: Artech House, 2006, p. 90.

[15] E. Wilkinson, "An N way hybrid power divider," *IRE Transactions on Microwave Theory and Techniques,* vol. 8, no. 8, pp. 116-118, August 1960.

[16] A. A. M. Saleh, "Computation of the frequency response of a class of symmetrical N-way power dividers.," *Bell System technical journal,* vol. 59, no. 8, pp. 1493-1512, October 1980.

[17] Mentor Grpahics, "www.mentor.com/electromagnetic-simulation/," December 2010. [Online].

[18] M. Healy and B. D. L. Minnis, "Wideband power combiners for GaAs MMICs," in *Proceedings of the 21st European Microwave Conference*, 1991.

[19] A. Alexanian and R. York, "Broadband spatially combined amplifier array using tapered slot transitions in waveguide.," *IEEE Microwave and Guided Wave Letters,* vol. 7, no. 2, pp. 42-44, 1997.

[20] R. Trew, "SiC and GaN transistors - is there one winner for microwave power

applications?," *Proceedings of the IEEE,* vol. 90, no. 6, pp. 1032-1047, 2002.

[21] S.-J. Yu, W.-C. Hsu, C.-S. Lee and et.al., "A flat gain/power responses 6-18 GHz power amplifier MMIC with high PAE by using transformer networks.," *Microwave and optical technology letters,* pp. 205-208, January 2008.

[22] J. Komiak, W. Kong and K. Nichols, "High efficiency wideband 6 to 18 GHz PHEMT power amplifier MMIC," in *IEEE MTT-S International Microwave Symposium Digest*, 2002.

[23] I. Bahl, "2-8 GHz 8-W poweramplifier MMIC developed using MSAG MESFET technology," *IEEE microwave and wireless components letters,* vol. 18, no. 1, pp. 52-54, 2008.

[24] I. Bahl, "Low Loss Matching (LLM) design technique for power amplifiers," *IEEE Microwave Magazine,* pp. 66-71, 2004.

[25] I. Bahl, "0.7 - 2.7GHz 12W Power Amplifier MMIC developed using MLP technology," *IEEE Transactions on Microwave Theory and Techniques,* vol. 55, pp. 222-229, February 2007.

[26] T. Arell and T. Hongsmatip, "A Unique 2 to 6 GHz 1 Watt High Efficiency GaAs MMIC ampliifier design using all-pass matching networks.," in *GaAs IC Symposium 1992*, 1992.

[27] A. Ezzeddine and et.al., "Ultra-broadband GaAs HIFET MMIC PA," in *IEEE MTT-S Digest*, 2006.

[28] G. van der Bent, A. de Hek, A. Bessemoulin and F. van Vliet, "Low cost high efficient 10 watt X band high power amplifier," in *IEEE International conference on Microwaves, Communications, Antennas and Electronics Systems*, 2009.

[29] W. Bosch, J. Mayock, M. O'Keefe and J. McMonagle, "Low cost X band power amplifierMMIC fabricated on a 0.25um GaAs pHEMT process.," in *Proceedings of the IEEE International Radar Conference*, 2005.

[30] K. Kobayashi, Y. Chen, I. Smorchkova, R. Tsai, M. Wojtowicz and A. Oki, "A 2 Watt, sub-dB noise figure GaN MMIC LNA-PA amplifier with mulit-octave bandwidth from 0.2-8 GHz," in *IEEE MTT-S International Microwave Symposium*, 2007.

[31] J. Sim, J. Lim, L. Park, W. Kang and B. Mah, "Analysis and design of wideband power amplifier using GaN," in *Asia Pacific Microwave Conference*, 2009.

[32] K. Krishnamurthy, D. Lieu, R. Vetury and J. Martin, "A 0.1-1.8 GHz, 100W GaN HEMT Power Amplifier Module," in *Compound Semiconductor Integrated Circuit Symposium*, 2010.

[33] K. Krishnamurthy, D. Green, R. Vetury, M. Poulton and J. Martin, "0.5-2.5GHz, 10W MMIC power amplifier in GaN HEMT technology," in *IEEE Compound Semiconductor Integrated Circuits Symposium Digest*, 2009.

[34] J. Gassman, P. Watson, L. Kehias and G. Henry, "Wideband, high efficieny GaN power amplifiers utilizing a non-uniform distributed technology," in *IEEE MTT-S International Microwave Symposium*, 2007.

[35] S. Lin and M. Eron, "Ultra wideband high gain GaN Power Amplifier," in *IEEE Radio and Wireless Symposium*, 2010.

[36] S. Masuda, A. Akasegawa and et.al., "Over 10W C-Ku Band GaN MMIC Non-uniform Distributed Power Amplifier with broadband couplers," in *IEEE MTT-S International Microwave Symposium*, 2010.

[37] G. Mouginot, Z. Ouarch, B. Lefebvre and et.al., "Three stage 6-18GHz high gain and high power amplifier based on GaN technology," in *IEEE MTT-S International Microwave Symposium*, 2010.

[38] V. Granatstien, P. Parker and C. Armstrong, "Scanning the technology: Vacuum electronics at the dawn of the twenty first century.," pp. 702-716, May 1999.

[39] S. Wood, "Advances in GaN based RF power transistors.," in *WAMICON*, Melbourne Beach, 2010.

[40] Capwireless, "Spatial combining technology: revolutionising the microwave power amplifier," *Microwave Journal,* September 2008.

[41] C. Suckling, "Thermal analysis of GaN devices," in *ARMMS Spring Conference*, Oxford, 2012.

[42] D. Gajewski, S. Sheppard, T. McNulty and e. al, "Reliability of GaN/AlGaN HEMT MMIC technology on 100mm 4H-SiC," in *26th Annual JEDEC ROCS Workshop*, Indian Wells, 2011.

[43] M. Poulton, "High power GaAs, GaN power amplifiers for W-CDMA base station applications," in *IWPC Conference on high efficeincy, high linearity power ampliferis for*

3G base stations, 2004.

3 Measurement System Construction and Verification

"This 'telephone' has too many shortcomings to be seriously considered as a means of communication."

Western Union internal memo, 1876.

3.1 Introduction

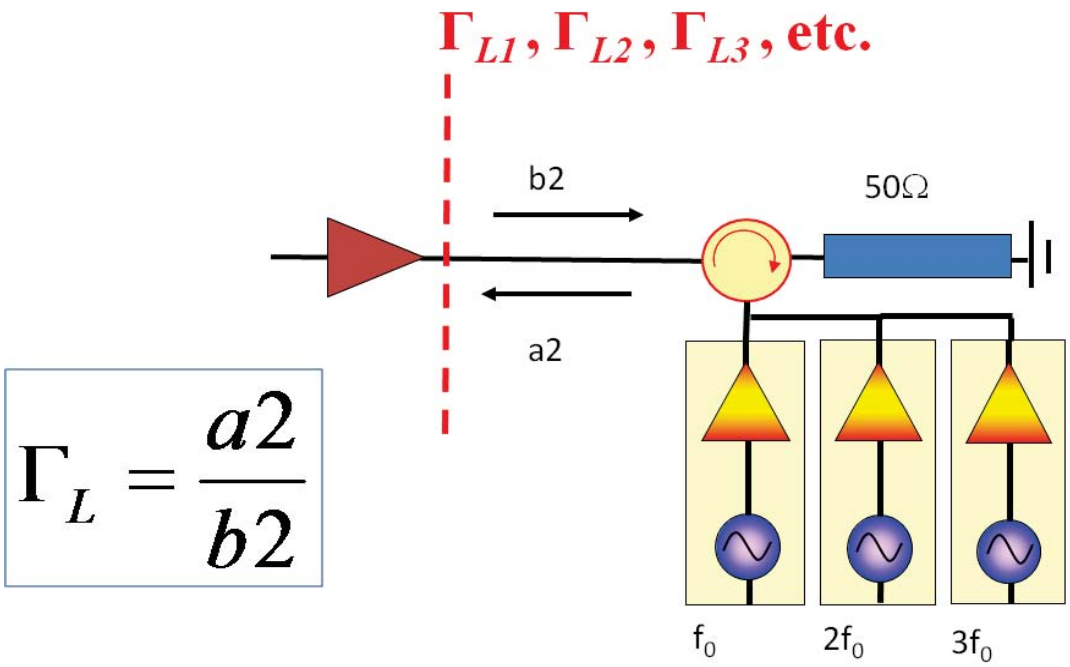
This chapter describes the design and construction of an on wafer, 1-40 GHz, active load pull, (ALP) device characterisation system. Over the period of this research, the measurement system has been described to a large number of audiences, and the quote above seemed an apt summary of the reaction of many. It is necessary to see past the large collection of signal generators and "microwave plumbing" that the various interconnecting cables appear to many. Like those early advocates of telephony, users need to recognise the need that is not being filled by current equipment and methods. Yes the current software can be cumbersome, calibration and measurement times can seem overlong (although not perhaps for those who calibrated HP8408s on HP85 desk top computers!), and the cost of the system components can seem exorbitant, (I was tempted to include the 1949 quote from Popular Mechanics, "*Where a calculator on the ENIAC is equipped with 18,000 vacuum tubes and weighs 30 tons, computers of the future may have only 1,000 vacuum tubes and perhaps weigh 1½ tons*" as the strapline to this chapter). All of these problems, or "challenges" in the popular business parlance, are solvable. A key problem that will be identified, phase coherent sources, isn't actually a technical problem; it just wasn't recognised as a need until recently. Now such sources are being introduced to the market. If anything, the need for the measurement system described here is increasing as data rates and system frequencies move higher and higher and new technologies such as Gallium Nitride come on stream. A key industry driver is time to market and facilities that can improve this can justify the investment required.

The measurement system is controlled by software whose development started with earlier systems at Cardiff University [1], and has continued throughout this project. The addition of higher frequency signal sources (20 & 40 GHz) and a 60GHz Digital Signal Analyser (DSA) allowed the techniques developed for the narrow bandwidth mobile communications frequency ranges to be extended to cover the higher frequency bands,

particularly those of interest to the defence industrial sectors, (e.g. 6-18 GHz). The key issues overcome with this 'scaling' of the system were concerning the triggering of the DSA and the phase locking of the sources. The system was proved by measuring the power contours with active fundamental load pull of a known Heterojunction Bipolar Transistor (HBT) device, and by comparing the measurements of a passive antenna on the system with those on an 8510 Vector Network Analyser (VNA). Further work on the measurement system included the reduction of cable lengths to reduce losses and the development of a wideband adjustable frequency Triplexer for the harmonic load pull.

3.2 Harmonic Active Load Pull

The load 'seen' by a device can be described by the reflection coefficient Γ_n, the ratio of the forward and reflected power waves, b_{2n} and a_{2n}, where '2' refers to the device port number and 'n' refers to the harmonic number. In active load pull systems the forward power wave b_{2n}, is absorbed via the isolator and an a_{2n} power wave is injected from a signal generator (and drive amplifier) as shown in Figure 3-1. If the system impedance is known the voltages and currents can be calculated from the knowledge of the power waves as will be described later. The active load pull system has the advantage that at all frequencies, other than the test frequency, the impedance seen by the device is 50Ω (within the bandwidth of the isolators and bias tees), thus enhancing the stability and therefore repeatability of the device measurement. The independence of the magnitude and phase of the a_{2n} wave allows the entire real impedance plane (Smith Chart) to be covered, and indeed opens up the possibility of harmonic injection [2]. The amount of drive power required by the driver amplifiers is dependent upon the insertion loss of the measurement system between the amplifier output and the device plane, the ratio of the device output impedance to the system impedance [3], the output power of the device and the extent of the impedance plane to be covered. The losses arise from a number of factors, the dual directional coupler, bias 'tees', isolator, associated cabling, and the net effect of all the mismatches in the line-up. All of these factors increase with frequency, as does the cost of driver amplifiers, hence the necessity in high frequency measurement systems of keeping cable lengths to a minimum and ensuring that all connections and transitions are of a high quality.



$$\Gamma_L = \frac{a2}{b2}$$

Figure 3-1, Harmonic Active Load Pull

Whilst some passive load pull systems, in an attempt to reduce the losses between the tuner and the device, and hence increase the measurement area of the impedance plane, place the directional coupler on the opposite side of the tuner to the device, in our system this approach would reduce the dynamic range by reducing the measured b_{2n} wave. At high frequencies and with very wide bandwidth directional couplers directivity is an issue and therefore all measures that can be taken to increase signal levels should be taken. Hence the directional couplers are mounted as close as possible to the measurement plane.

The choice of an ALP system over a passive system for this project becomes clear when the effect of system losses is considered on the impedances that can be replicated by the system. Figure 3-2 shows two boundaries on the impedance plane determined by an insertion loss of 1 (blue trace) and 2dB (red trace) prior to a short circuit. These boundaries encompass approximately 62% and 39% of the impedance plane respectively, and the losses are not unrealistic at X band and above. The question that should be asked of course is as to how critical the area outside these boundaries is in the measurements and investigation of device performance. As has been described [4] and [5] the area around the outer perimeter of the impedance plane, particularly at the harmonic and therefore higher frequencies, is the key to achieving the highest efficiencies.

A question that may be asked is why is it necessary to map the impedance plane, a stated advantage of the measurement system developed. In the ideal world we are looking to find the optimum impedance points at the fundamental and harmonic loads. A good first

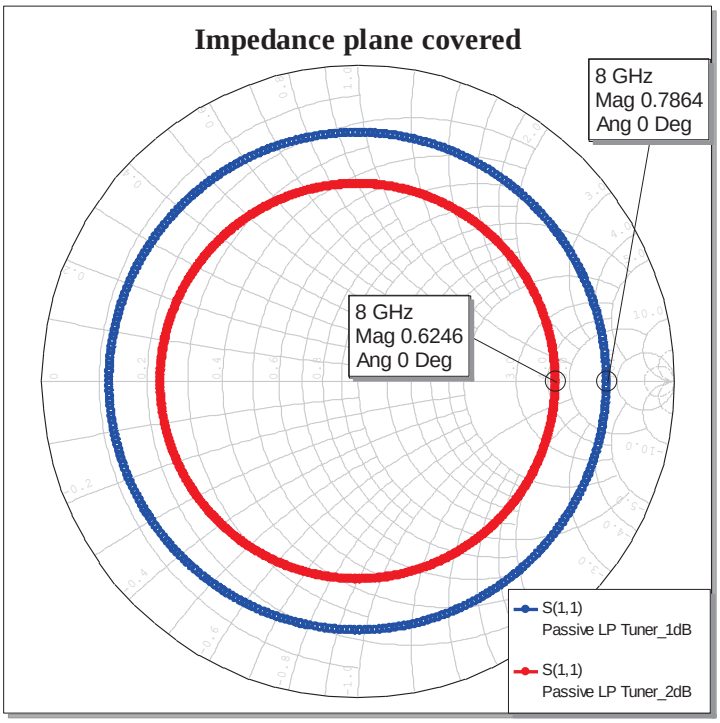


Figure 3-2, Impact of loss on impedance range of passive load pull

approximation can be found from basic theory [6], and a reasonable search algorithm could find the optimum impedance points [7], thereby producing the information we require. However in reality, and particularly as frequency increases, the theoretical models deviate from practice. Also we do not get an idea of the sensitivity of the device performance to load impedance that can be seen from the load pull contours, and finally, especially in broadband design, we cannot always achieve the impedances we require and we need to have a picture of how our compromises will affect performance.

Methods, which can determine specific appropriate load impedances [8], still require that load pull measurements are carried out in order to check that an optimal result has been achieved and determine the performance away from the ideal load. We need to know not only which areas to target, but also which to avoid.

3.3 System Description

The block diagram for the basic calibration of the measurement system is shown in Figure 3-3. The fundamental calibration was carried out using a 40 GHz source. The sampling heads of the Digital Serial Analyser (DSA) had a 3dB bandwidth of 60 GHz. The frequency limiting elements of the measurement system were thus the couplers which had a nominal bandwidth of 1-40 GHz. The combination of sampling heads and directional coupler is shown in Figure 3-4. Load pull measurements in the system were practically

limited by the availability of high frequency sources. Although fundamental load pull could be conducted over the range 1-40 GHz, as the number of harmonics to be measured/load pulled increased so the fundamental frequency range was decreased, as shown in Table 3-1.

Load pull	Input Source	Load Source	2^{nd} Harmonic	3^{rd} Harmonic
Fundamental	40 GHz	40 GHz	N/A	N/A
Fund & 2^{nd}	20 GHz	20 GHz	40 GHz	N/A
Fund, 2^{nd}, & 3rd	13.33 GHz	13.33 GHz	26.67 GHz	40 GHz

Table 3-1, Harmonic Measurement and Frequency Limitations.

S parameters are measured in fundamentally the same way as with a standard Vector Network Analyser, (VNA) by sampling the forward and reflected waves using a directional coupler. The difference is that these sampled waves are measured in the time domain by a fast sampling scope; in the earlier narrow band systems using the Agilent Microwave Transition Analyser (MTA) [1] and with the system described here using the Tektronix DSA 8200. A major advantage of this latter system is the provision of 4 sampling detectors allowing for the simultaneous measurement of all four coupler ports. The waveforms captured by the samplers are processed in software to provide magnitude and phase information at each constituent frequency. Provided the waveforms captured are repetitive, the software can, through performing a Finite Fourier Transform (FFT) on a cycle, determine the magnitude and phase of each frequency in the spectrum. Thus not only can the linear characteristics based upon S parameters be determined but also the nonlinear and distortion effects.

The error correction procedure [9] is similar to the 6 term per port correction applied with VNAs. The purpose of the calibration is to improve the measurement accuracy by removing the systemic errors, to establish known measurement reference planes and increase dynamic range. In order to improve the measurement accuracy averaging is used when acquiring the waveforms, this reduces the random errors in the measurement system that cannot be removed by calibration. The standard S parameter calibrations only provide relative measurements of the a_n and b_n travelling waves. In order to be able to determine the absolute value of the voltages and currents another calibration step is required, referred to the as the power calibration. Ideally a power measurement of the a_1 wave would be conducted at the now calibrated measurement plane, however as this plane is usually in

microstrip or at wafer probe tips the measurement must be conducted at another point and referred back, thus requiring the establishment of another reference plane.

The method adopted was for a thru-line to be connected between the wafer probe tips and a coaxial cable connected to the unused port of the left hand switch in Figure 3-3. A 1-port short, open, load, coaxial calibration was conducted at the end of this cable thus determining the reference plane. A power meter was then connected to this port to provide an absolute measure of power at each of the calibration frequencies. This was then repeated using one of the samplers from the DSA which thereby provided the ability to calibrate the sampler head sensitivity with frequency. The measured power waves could thus be referenced back to the wafer probe tip measurement plane. Knowing the impedance and the power waves the voltages and currents at the measurement ports can be calculated:

$$v_n = \sqrt{Z_0}(a_n + b_n) \qquad \{3\text{-}1\}$$

and,

$$i_n = \frac{(a_n - b_n)}{\sqrt{Z_0}} \qquad \{3\text{-}2\}$$

In order to conduct the measurements it is necessary to synchronise the sampling heads and the signal sources. This can be an issue with active load pull systems based around sampling oscilloscopes. The triggering of the DSA is possible via one of two ports on

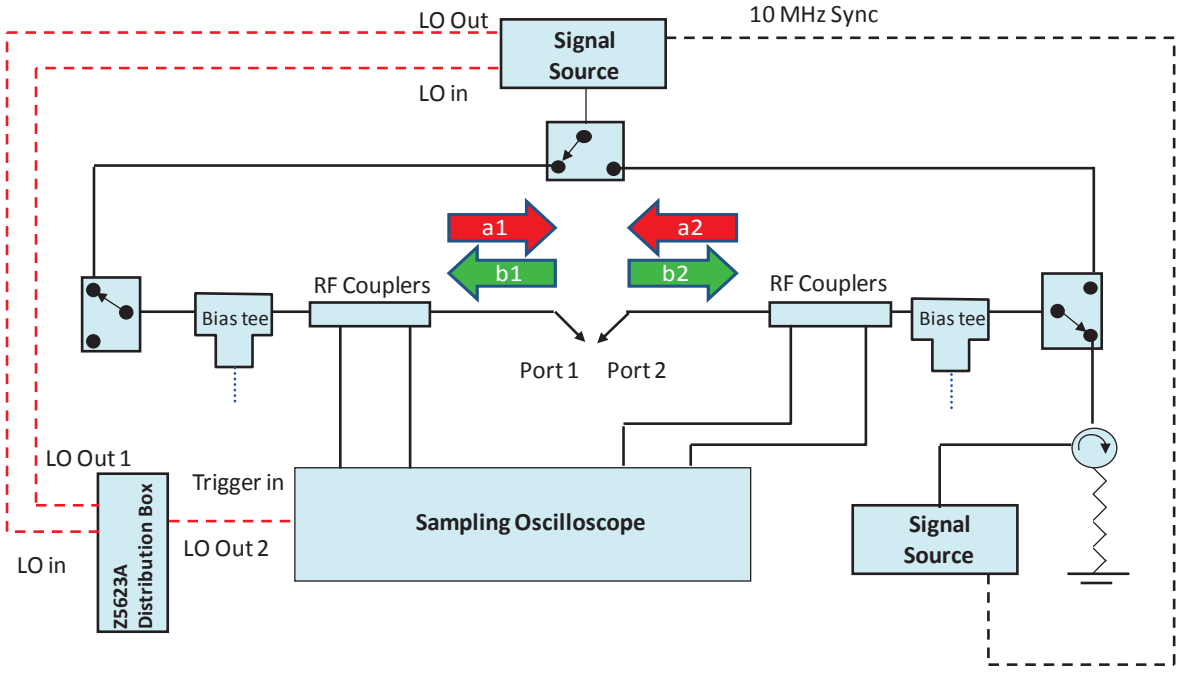


Figure 3-3, Block diagram of the realised Large Signal fundamental waveform measurement and engineering system

the front panel; selected dependent upon frequency range, or by phase locking through the Phase Reference Module (82A04). This latter option was abandoned in part due to the difficulties of achieving phase coherence of the system over the full frequency range, as will be discussed later. Even across the limited frequency range over which the system could be coherently phase locked it was found that using the Phase Reference Module increased signal acquisition time dramatically. The trigger bandwidth of the Tektronix DSA was limited to an upper frequency of 12.5 GHz, (part of the original reason for acquiring the Phase Reference Module was to obviate the need for a trigger and above 12.5 GHz to lock directly onto the fundamental input signal), hence in order to achieve correct triggering with signals up to 40 GHz a new triggering architecture was required.

An alternative, novel, method of triggering was devised using the Distribution Box (Agilent Z5623A Option K08) originally provided for phase locking two sources together [10]. The signal sources procured for the measurement system have option HCC which routes the fundamental frequency of the internal YIG oscillator via the back panel. This path can be broken and taken to the Distribution Box where it is amplified and split providing two coherent outputs. One output is supplied back to the original source (the Master) and the other is intended for the second source (the Slave). However in the devised configuration the second output is routed to the trigger input of the DSA. Importantly the output of the YIG oscillator is always in a frequency range of 3.2 to 10 GHz, as the output occurs before the frequency doublers and dividers within the signal source, Figure 3-5. This means that

Figure 3-4, Directional coupler, attenuators, sampling heads and bias 'Tee' (test port is on the left hand side).

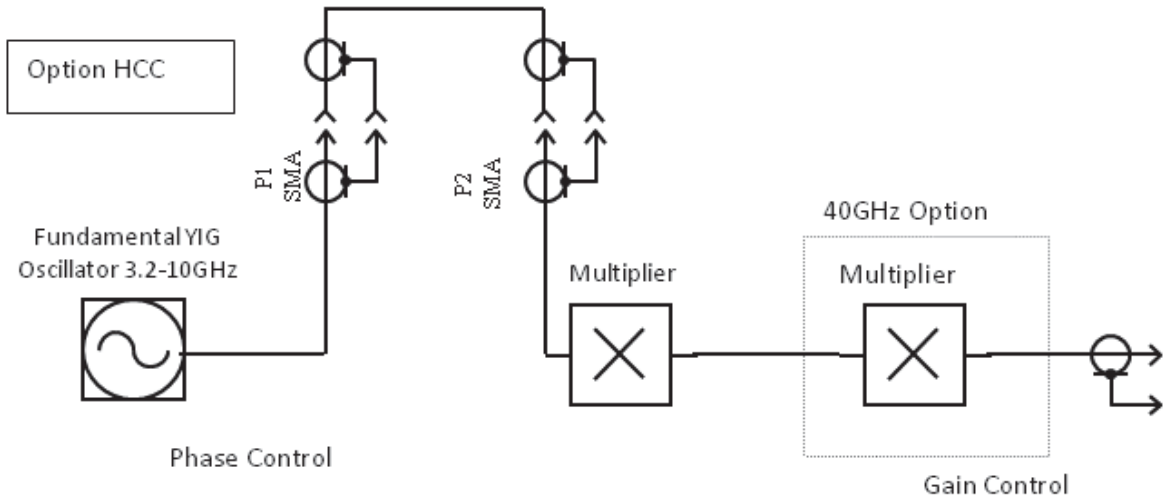


Figure 3-5, Signal Generators with Option HCC for external phase locking.

the trigger frequency is always either at exactly the same frequency or some sub-harmonic of the stimulus signal with perfect phase coherence and is always in the relevant frequency range of the trigger circuitry. On the DSA 8200 there are 2 trigger inputs, one for below 3GHz and the other for 2-12.5 GHz. The rear panel connections for connecting a signal generator, the distribution box and the trigger cable to the DSA are shown in Figure 3-6. In order to still be able to use the distribution box to phase lock two sources together the trigger signal was coupled off (10dB coupler) the output from the distribution box. This also helped reduce the signal level closer to that specified for the trigger input. The final schematic for a phase coherent fundamental load pull system (two signal sources) with digital oscilloscope triggering is shown in Figure 3-7.

A consequence of this method of locking sources together is that an alternative method of controlling the phase of the output signal is required as normally the phase is adjusted at the fundamental YIG oscillator. Fortunately the sources used are equipped with I & Q modulators in the output chain and these can be used to alter the phase of the output signal. Note that the output power level could also be adjusted in the same way; however as the power control is after the multiplier stages this is not strictly necessary.

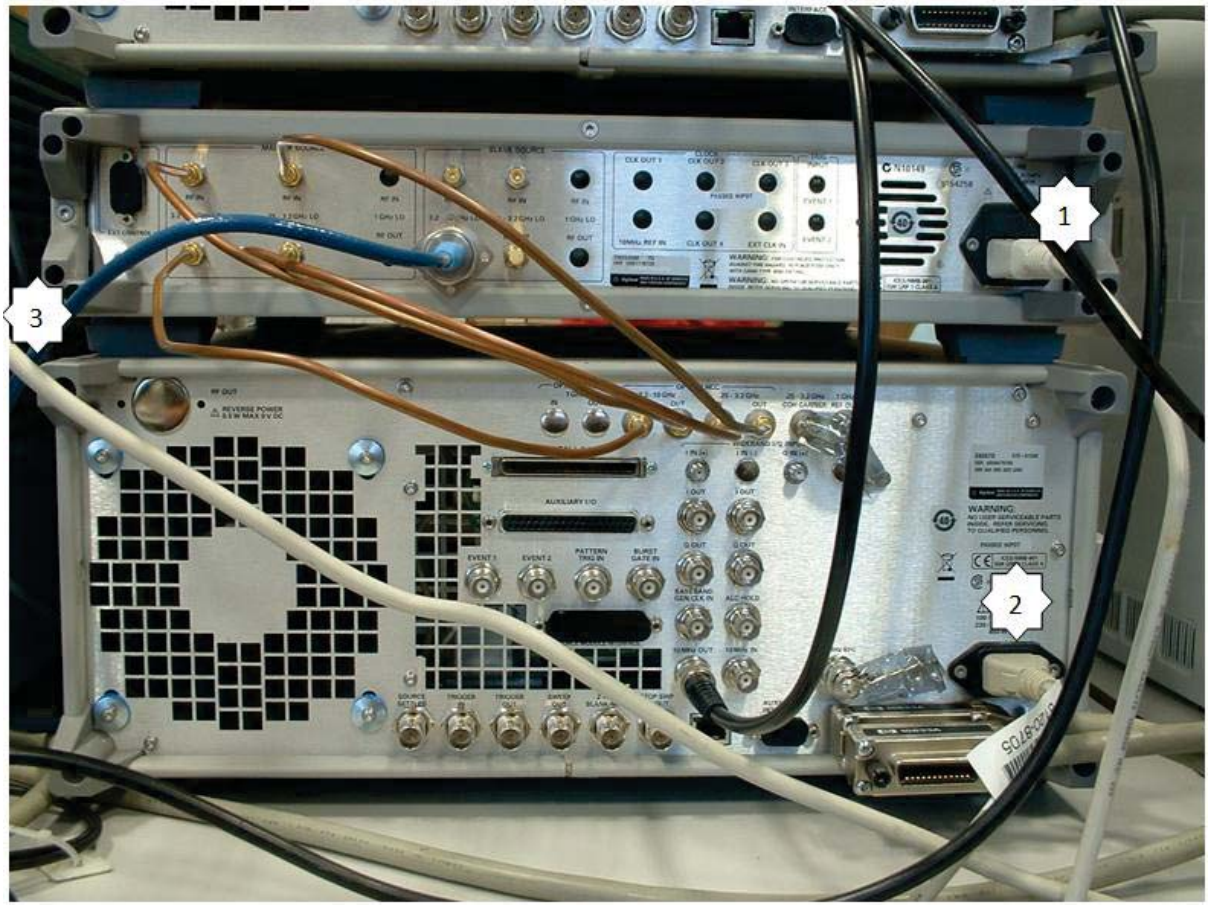


Figure 3-6, Rear panels of the distribution box (1), Signal Generators with Option HCC (2), trigger cable (3) for sampling oscilloscope external triggering.

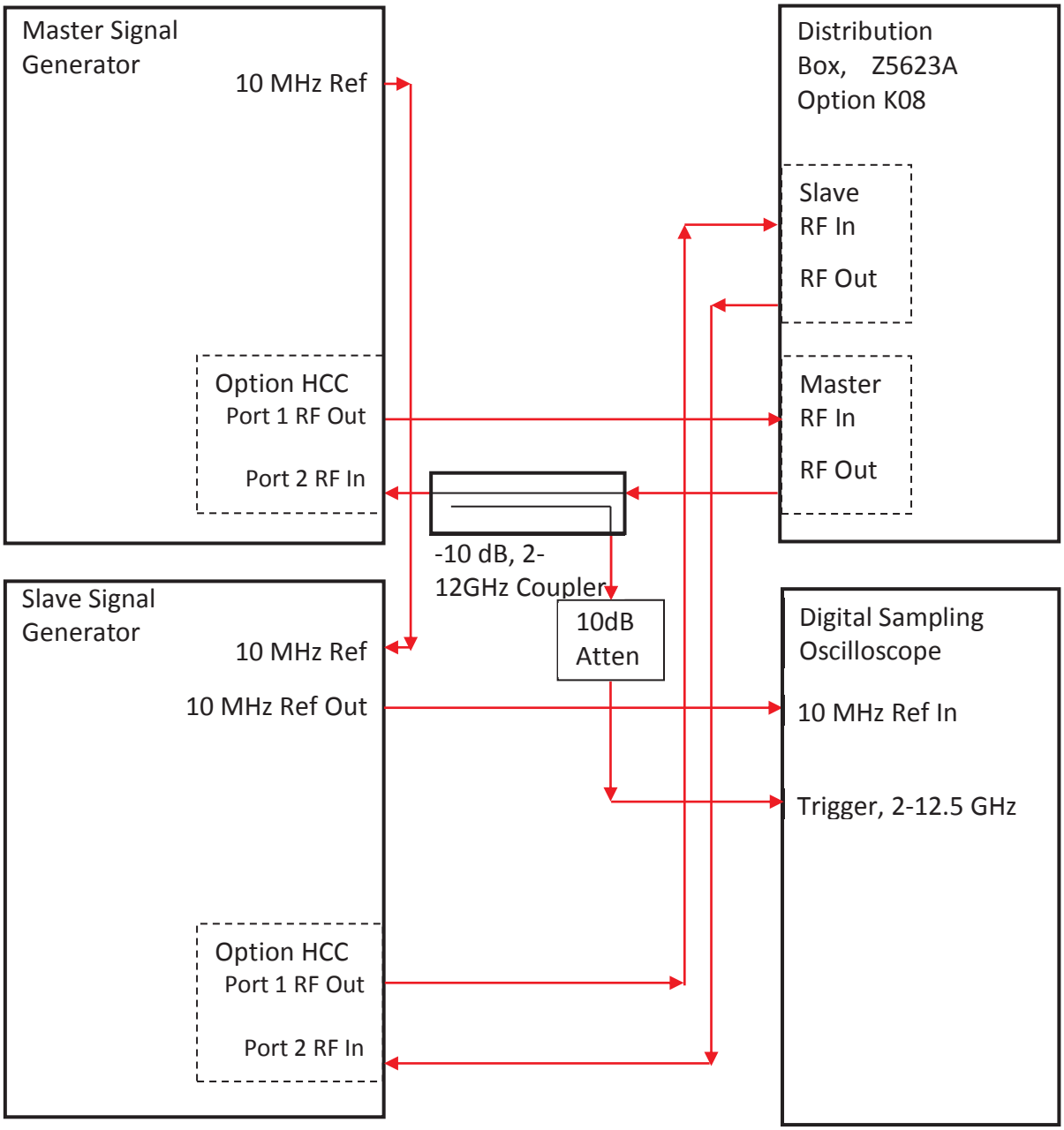


Figure 3-7, System connections for fundamental active load pull (two signal sources) with phase coherence and DSA triggering.

The system described so far is able to phase coherently lock the fundamental input signal and the second source, which is used to provide the active load pull, a_{21} wave. At first glance it may appear that to control more sources for harmonic load pull it is only necessary to replicate the 'slave' circuitry of the distribution box. However a fundamental problem exists with the architecture of the signal generators such that phase coherent harmonic control is only possible over a limited range of frequencies. This can best be understood from the frequency bands for the signal generators as described in Table 3-2 and the frequency plan, Table 3-3. As mentioned, the fundamental source in the signal generator is a wideband YIG oscillator covering the frequency range 3.2 to 10 GHz. To achieve higher frequencies multipliers are used. This works very conveniently for 7 GHz (for example); 7

GHz is in the fundamental band, the 2nd harmonic, 14 GHz is in the 2x band and the 3rd harmonic, 21 GHz is in the 3x band. However for 6 GHz, whilst the fundamental and 2nd harmonic are derived in the same way, the 3rd is in the 2x band and hence uses a fundamental of 9 GHz and so it cannot be derived from the same source. Unfortunately there is no way to override the multiplier control and produce 18 GHz from 6 GHz using the 3x multiplier (the multipliers are band specific as they include filters to maintain the signal generators harmonic and spurious specifications).

Band	**Operating Frequency Range**	**Derivation**
Heterodyne Band	< 250 MHz	Down-converted from 750 - 1000 MHz band
Low Band	0.25 - 3.2 GHz	Divided from Fundamental band
Fundamental Band	3.2 - 10 GHz	
2 x Band	10 - 20 GHz	Doubled from Fundamental
3 x Band	20 - 28.5 GHz	Tripled from Fundamental
5 x Band	>28.5 GHz (to 44 GHz in E8267D)	Quintupled from Fundamental

Table 3-2, Agilent PSG Signal Generator Frequency Bands.

Fo	Fund	Multiplier	2Fo	Fund	Multiplier	3Fo	Fund	Multiplier
4	4	1	8	8	1	12	6	2
5	5	1	10	10	1	15	7.5	2
6	6	1	12	6	2	18	9	2
7	7	1	14	7	2	21	7	3
8	8	1	16	8	2	24	8	3
9	9	1	18	9	2	27	9	3
10	10	1	20	10	2	30	6	5
11	5.5	2	22	7.3	3	33	6.6	5
12	6	2	24	8	3	36	7.2	5
13	6.5	2	26	8.7	3	39	7.8	5
14	7	2	28	9.3	3	Yellow highlight – 3 harmonic phase coherent control.		
15	7.5	2	30	6	5			
16	8	2	32	6.4	5			
17	8.5	2	34	6.8	5	Red border – 2 harmonic phase coherent control.		
18	9	2	36	7.2	5			

Table 3-3, Agilent PSG Signal Generator Frequency Derivations.

Thus we are limited to 2nd harmonic load phase coherent active load pull of 5.1 GHz (10.2/2) to 10 GHz (20/2) and 3rd harmonic 6.67 GHz (20/3) to 9.5 GHz (28.5/3). Outside of

these frequency ranges it is possible to operate the measurement system, but with the sources only locked using the 10 MHz reference. The consequence of this non coherent operation is reduced phase stability between the fundamental and the harmonics and thus in the load pull the reflection coefficient 'wanders' as the relative phase drifts. As said earlier the load pull reflection coefficient, Γ_L, is defined by:

$$\Gamma_{nL} = \frac{a_{2n}}{b_{2n}} \qquad \{3\text{-}3\}$$

where n is the harmonic number.

b_{2n} is produced from the device under test, but originally generated by the fundamental source, a_{11} , whilst a_{2n} is from the load pull sources. In the system as described, a_{21} can always be phase coherent with b_{21}, however for other values of *n*, this is only true for the limited cases described above. The extent of the phase drift of the harmonic load signals will depend upon a number of factors, including frequency, the particular sources used, temperature, etc., but for an idea of the relative phase behaviour a comparison of phase coherent and 10 MHz locked sources was made. The distribution box has the ability to switch in or out the master source. Thus in "calibrate" mode the signals from the sources are fed back to themselves and hence the sources behave independently. In "distribution" mode the master source fundamental frequency is fed to the slave source as well as back to the master. For this experiment, the master source was set to 9 GHz and the slave to 18 GHz, thus the two signals are within the second harmonic phase coherent band. The sources and the DSA were locked to a 10 MHz reference generated by the master and the outputs of the sources were connected to two sampler heads of the DSA. With the distribution box set to "distribution", i.e. phase coherent, there was no discernible phase drift between the 9 and 18 GHz signals observed on the DSA over a 40 minute period. The distribution box was then set to "calibrate" and the traces observed. In this case the 18 GHz signal drifted by 14.5° relative to the 9 GHz signal, over a period of 30 minutes. This equates to almost 0.5°/minute. For some measurements this appears to be acceptable, however in automatic load pull operations at some high reflection coefficients it can take a 10-15 iterations to reach the desired load point and this phase variation helps to create a 'moving target', the more harmonics that are trying to be controlled the more impact this variation has.

It is important to remember with 2nd harmonic load pull we are dealing with 3 sources, fundamental input, fundamental load and 2nd harmonic load.

To examine the phase variation as measured by the DSA itself, a single frequency source was connected to two samplers by splitting the main output from an E8267D 20GHz signal generator, set to 5 GHz and 0 dBm output power. The 6 dB 2-way resistive splitter had one port connected to the channel 2 sample head and the other to 2-12.5 GHz trigger prescaler input, thus ensuring synchronous sampling. Measurements of signal phase where taken every 5 minutes. The results are summarised in Figure 3-8, and show a 0.5° average drift over 105 minutes (0.8×10^{-3} °/min) with a variation of ±3.5°. Drift in the fundamental frequency itself is not a problem, provided the other sources can be locked relative to it. The phase variation did limit the number of averages that could be taken when measuring the traces on the DSA, to less than 512 averages [11].

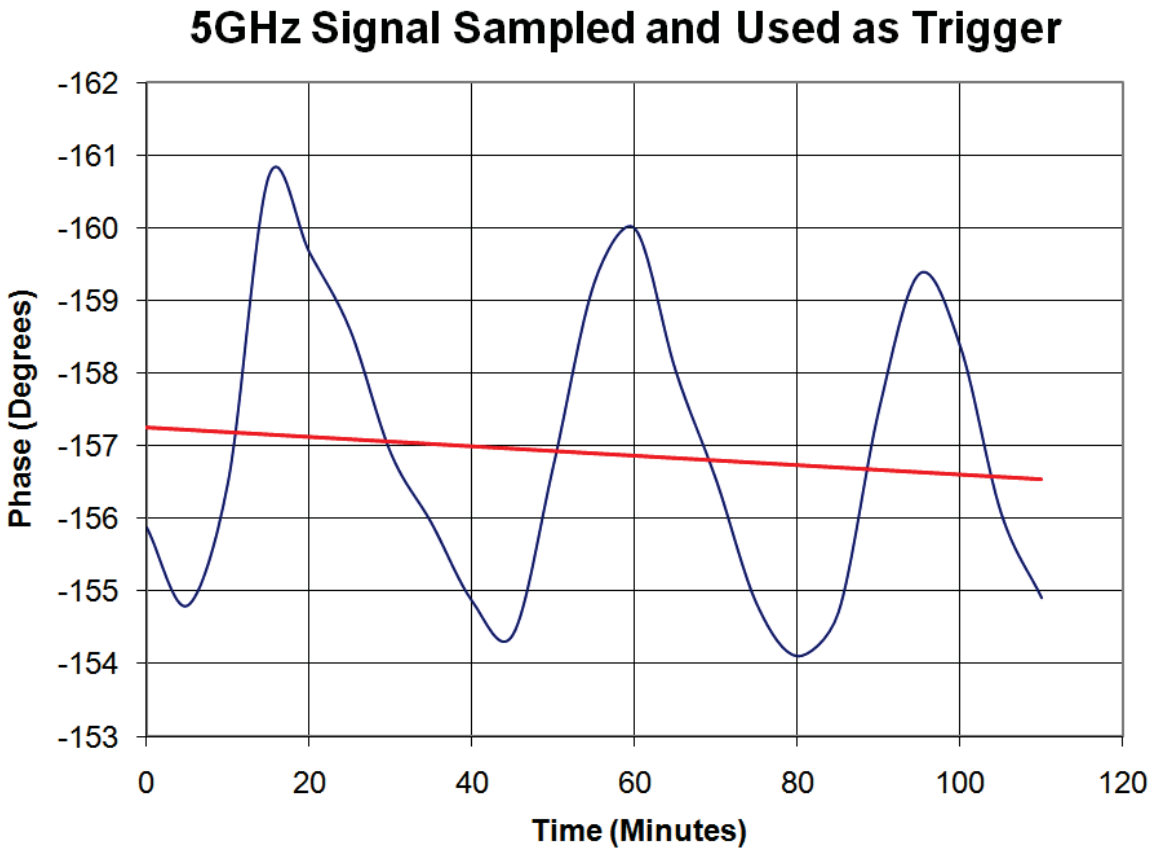


Figure 3-8, Measured phase variation (blue) of a single 5 GHz signal on the DSA 8200, showing short term phase variation (blue) and long term trend (red) of the DSA sampling heads.

The measuring of the drift with a single sampler provides one part of the puzzle when it comes to the measurement system phase errors. More important is the relative drift between samplers. It is possible that the sampler heads themselves introduce some error. To check the possibility and magnitude of this potential problem the previous test

was modified to follow the two way splitter with a four way splitter, with each port of this device connected to a sampler head, as shown in Figure 3-9, the samplers having previously been calibrated into 50Ω loads. A software procedure was written to measure all 4 channels at regular intervals over a defined period; magnitude was recorded as well as phase. The absolute values of magnitude and phase are not critical, and hence the ports of the 4 way splitter and the adapters have not been separately measured. The magnitude and phase are calculated by carrying out an FFT on the captured waveforms. The resulting magnitude is scaled by half the number of points in the FFT to give the peak magnitude of the signal. The phase is converted from radians to degrees.

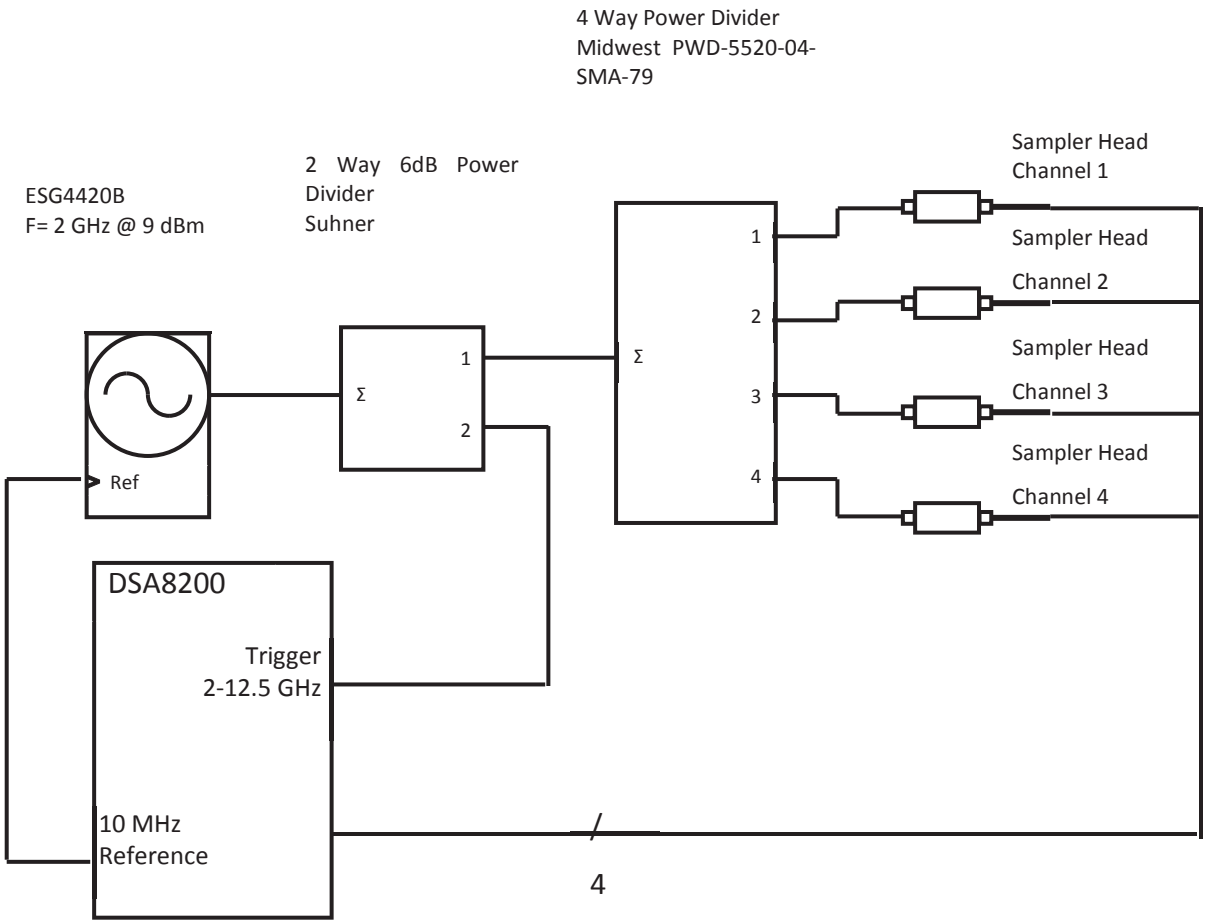


Figure 3-9, Sample Head magnitude and phase variation measurement.

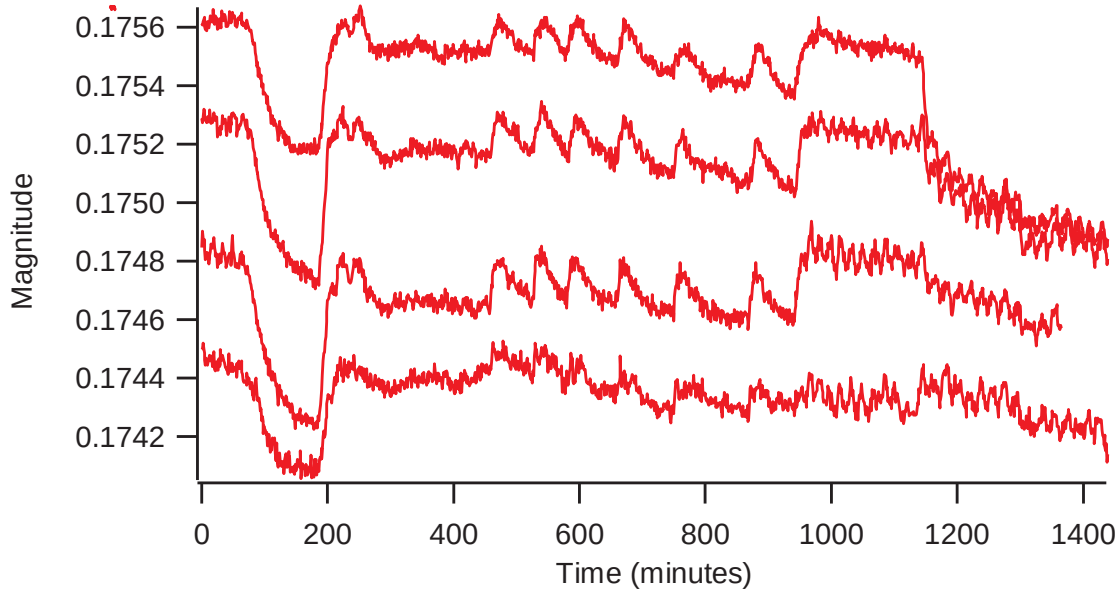


Figure 3-10, Variation of measured Voltage with Time.

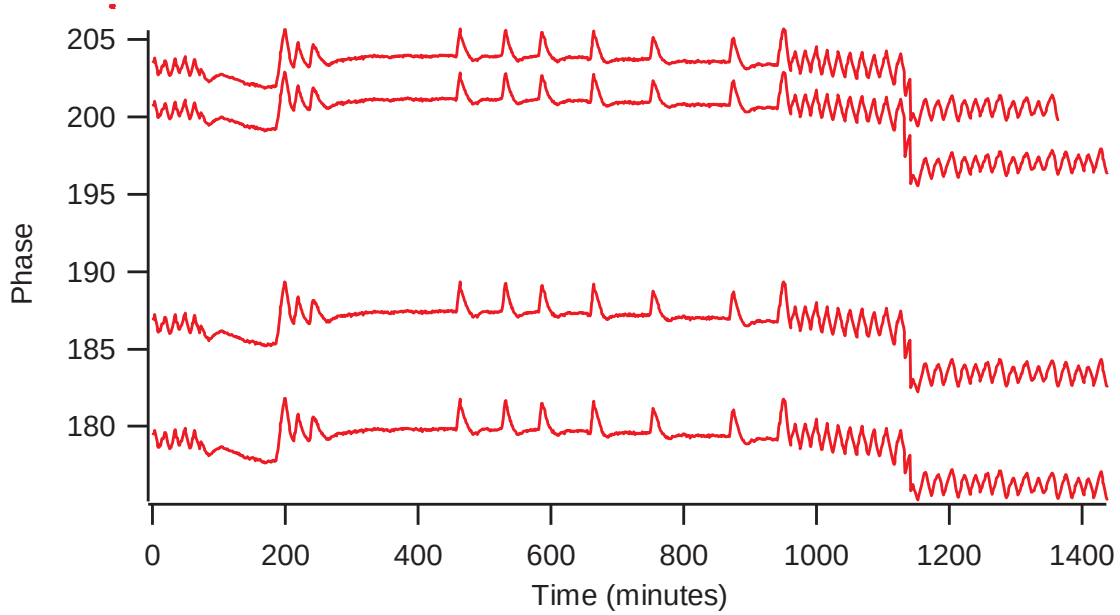


Figure 3-11, Variation of measured Phase (°) with Time.

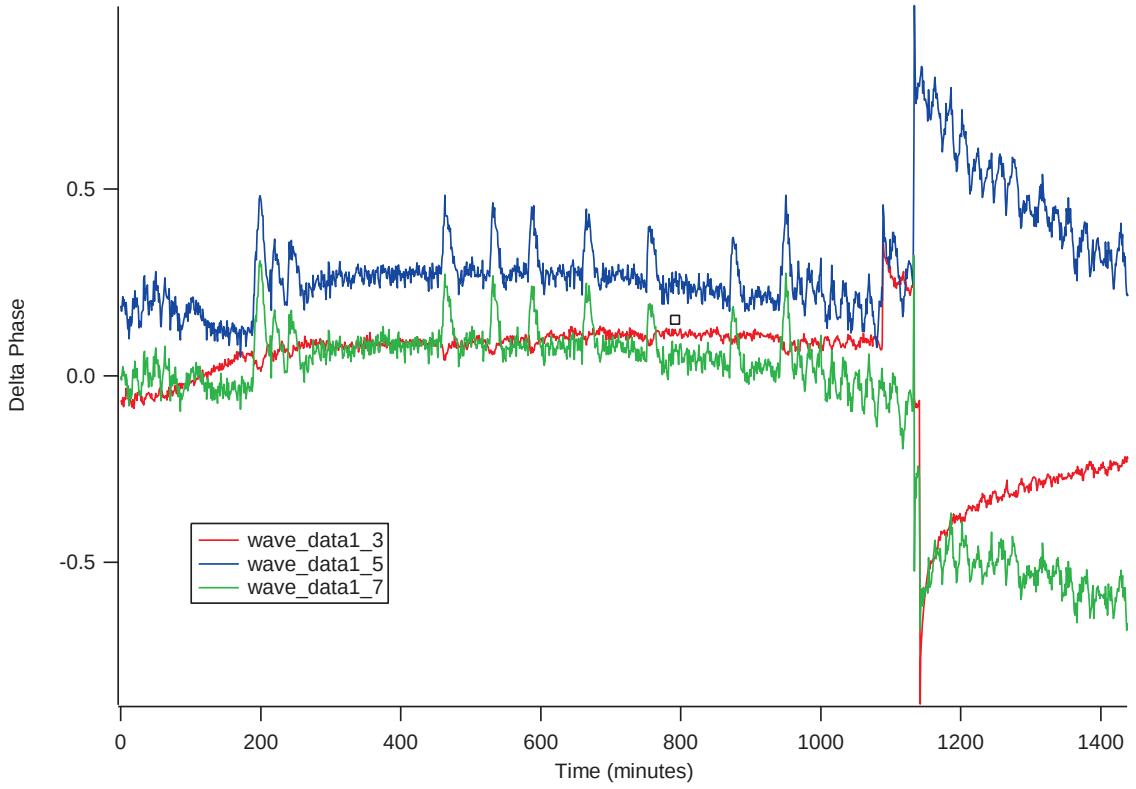


Figure 3-12, Relative Variation in Measured Phase (°) between Sample Heads

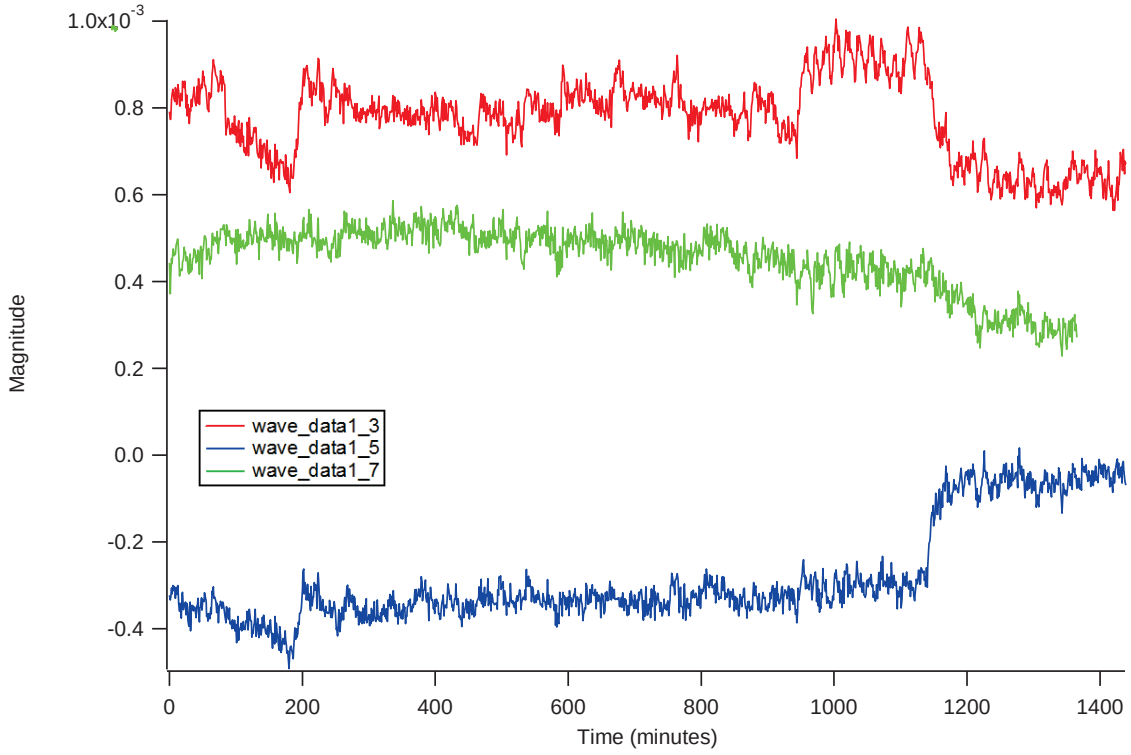


Figure 3-13, Relative Variation in Measured Voltage between Sample Heads.

Making comparative measurements, using channel 1 as the reference is shown in Figure 3-12 and Figure 3-13.

Data Set Name	Information
wave_data1_3	Phase comparison of Channel 2 to Channel 1
wave_data1_5	Phase comparison of Channel 3 to Channel 1
wave_data1_7	Phase comparison of Channel 4 to Channel 1

Table 3-4, Traces used in relative phase measurement, Figure 3-12 and Figure 3-13.

The magnitude variation over the first 15 hours (900 minutes) of each channel is within ±120µV and there is no clear trend. The phase variation is more interesting. There is a clear difference between the two sets of sample heads (Channel 1&2 are connected to one sampling module, 3&4 to the other), including relative phase spikes. However the relative phase variation between the channels over the first 15 hours is less than 0.5°, and within the magnitude of the phase spikes between the sampling modules. An incident occurs at about 19 hours (~1150 minutes), which causes a phase 'hit' of over 1°, however this is still within acceptable bounds for the measurement system. The cause of the change is unknown, but the DSA does occasionally self-calibrate and it could be that this occurred at this time.

As said earlier absolute phase drift in either the DSA or the frequency sources is not an issue, however relative drift will alter the measured phase of the a_n and b_n waves. The phase drift is of the same order as the noise between different sampler modules and therefore is no more of a limiting issue than the absolute dynamic range of the DSA.

3.4 Measurement System Switching

The switching arrangement used in the system allows for the two port calibration via the transfer switch, whilst the other switches (switch_1/2/3) allow the transfer switch and cables to be by-passed and the driver amplifiers incorporated in measurement mode, thus reducing the losses and maximising the driver power at the test ports. Referring to Figure 3-14, during the forward calibration (S11 and S21), "switch_1" is set to the "cal_state(0)" position and the transfer switch is set as shown in the diagram. This connects the primary source to the input port of the DUT position (during calibration the DUT is replaced with the appropriate calibration standards). The output port of the DUT is terminated in the 50Ω load connected to the transfer switch, which during calibration is

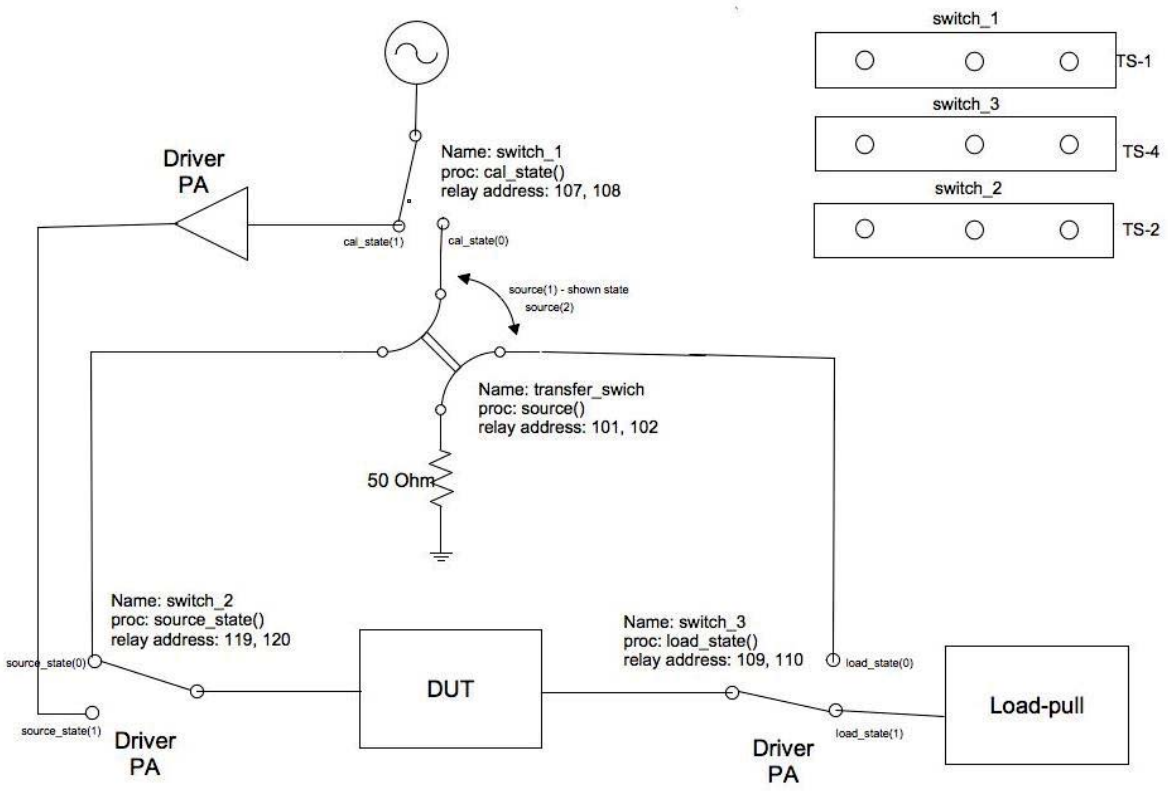


Figure 3-14, Measurement system switching arrangement

only strictly needed for the through line measurement. For the reverse calibration (S22 and S12) the transfer switch toggles such that the primary source is now connected to the output port of the DUT position and the input port is terminated with the 50Ω load. For the power calibration the transfer switch is left in the reverse state and the calibration standards, sampler and power meter are connected to "switch_2" in "source_state(1)" position (usually via the cable that normally connects to the driver amplifier output for convenience).

For calibration the primary source is typically the highest frequency unit available, as we wish to calibrate over the maximum frequency range. However, in operation, this source will normally be used for the highest harmonic load pull and a lower frequency source used for the input driver as defined in Table 3-1. For devices requiring higher drive power than available from the signal generators (typically ~23dBm) then a driver amplifier is incorporated.

3.5 System Verification

A practical measurement system [11] must have sufficient directivity (ability to distinguish between a 100% reflect and a perfect match) and also to be able to consistently

repeat measurements. To verify the completed system a number of investigations were made:

1. Repeatability of reflection measurement at frequencies in the range 5-40GHz.
2. Dynamic range measurements.
3. Comparison of the S parameters of a non-ideal 1 port measured on the system and a traceable VNA.

It is only necessary to test a single port (sampler and coupler combination, Figure 3-4) as this exercises the forward and reflected wave measurement system. The investigations carried out were designed to show the consistency, range and accuracy of the set-up.

3.5.1 Repeatability of Reflection measurements

A non 50Ω impedance was measured at 5, 10, 20 and 40 GHz, 2000 times, with no averaging. Figure 3-15 shows the repeatability of a 1 port reflection measurement between 5 and 40GHz in terms of impedance spread on the Smith Chart. The measurement data was then averaged and the magnitude of the standard deviation in phase and magnitude plotted as shown in Figure 3-16 and Figure 3-17. As the number of averages increases the standard deviation decreases, until above 500 averages it can be seen that the impedance measured drifts, shown by the increase in the standard deviation. This is due to the relative frequency drifts between the signal source and the DSA, as discussed previously.

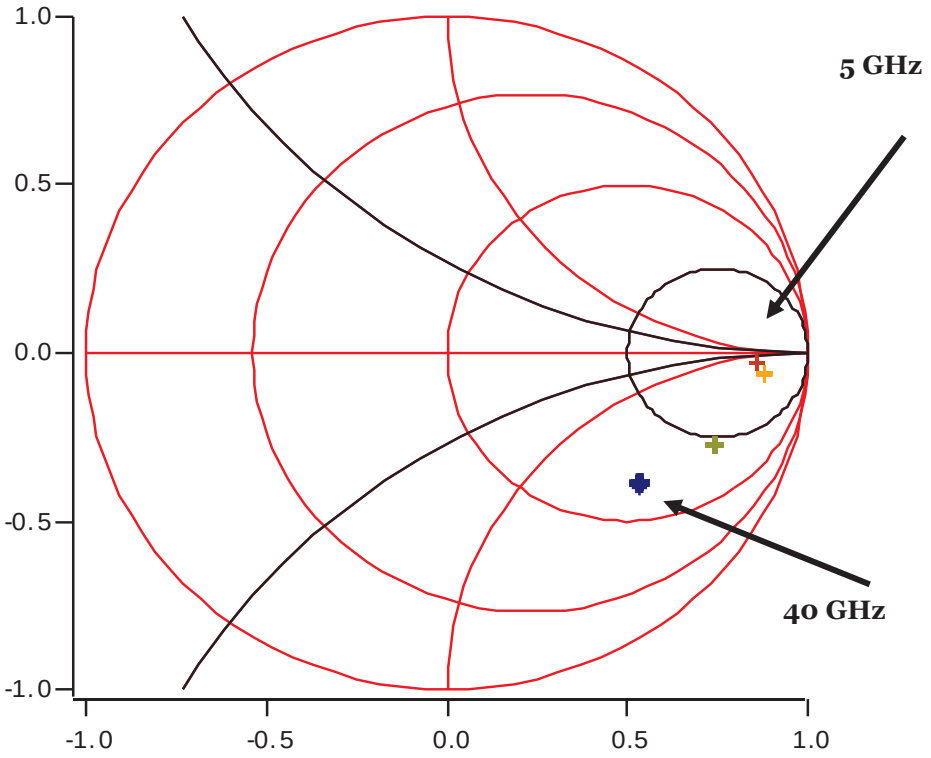


Figure 3-15, Repeated Measurements of S11 on non-50Ω impedance at 5, 10, 20 and 40GHz.

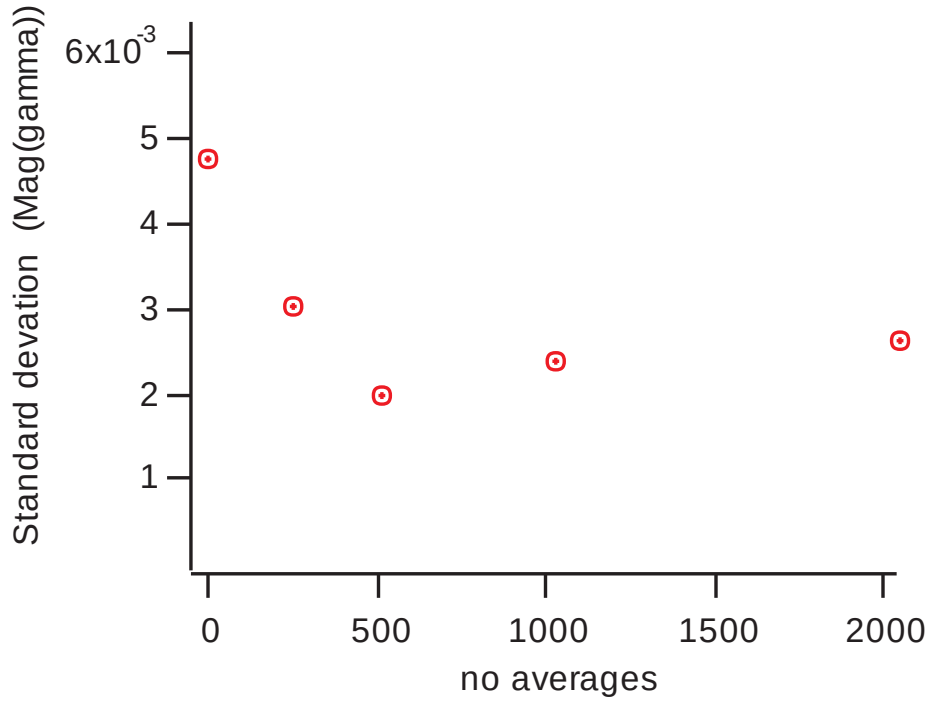


Figure 3-16, Variation of Magnitude of Reflection Coefficient with number of measurements (time).

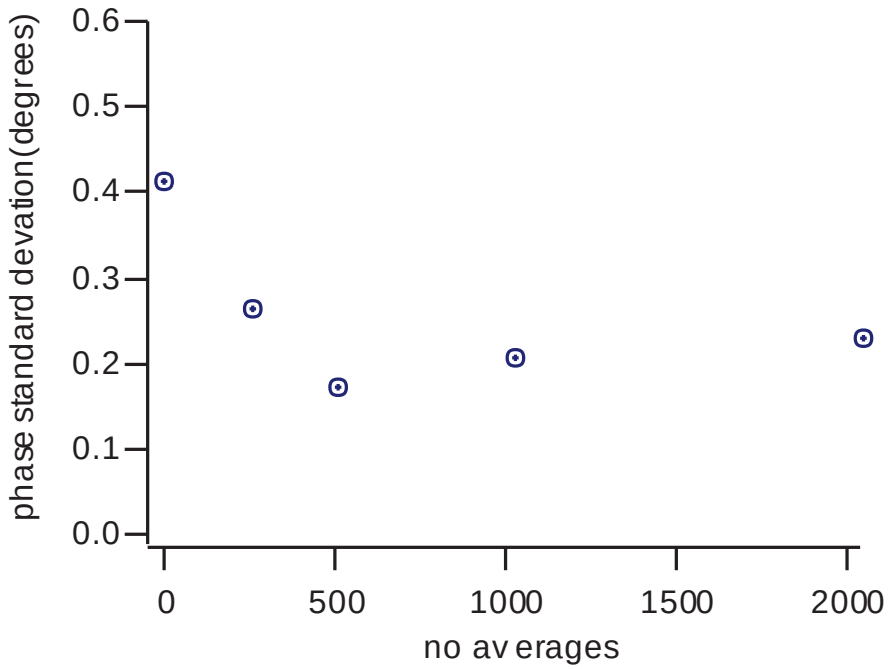


Figure 3-17, Variation of Phase of Reflection Coefficient with number of measurements (time).

3.5.2 Dynamic Range of measurements

This characteristic of the measurement system is ultimately limited by the minimum directivity, the ability to distinguish between a perfect reflection and the matched system impedance. The system is calibrated with open and short circuits on the measurement ports; these provide the maximum return signal levels. For a 1 port calibration a good 50Ω termination is connected to the measurement port. The signal level returned from this termination provides the lower reference level. However any power measured at the test port at the measurement frequency will be considered a returned signal, thus the 'breakthrough' of the main incident signal to the coupler input port (the directivity) must be calibrated out. The raw data measured by the system at the coupler output ports (i.e. coaxial calibration) is shown in Figure 3-18; these show that without calibration a directivity of >13 dB can be achieved up to 40 GHz. Once calibrated the system measurement performance can be improved and a reference plane established. The results of a Short, Open, Load (SOL) 1 port calibration are shown in Figure 3-19 and Figure 3-20, re-measuring each of the standards after applying the error correction, shows an excellent directivity of >50 dB to 40 GHz, for a coaxial calibration.

The majority of the device measurements were however conducted on bare die and hence calibration at the wafer probe tips is required. This necessitated the incorporation of additional cables, and hence losses, between the calibrated wafer probe tip reference plane and the couplers. The initial system configuration is shown in Figure 3-21. Measuring the calibration standards over the frequency range 4 to 40 GHz shows a slight decrease in the raw directivity compared with the coaxial calibration to >8 dB to 40 GHz, Figure 3-22, however the calibrated performance is still >50dB, Figure 3-23 and Figure 3-24, which is more than adequate for the active load pull and large signal device measurement.

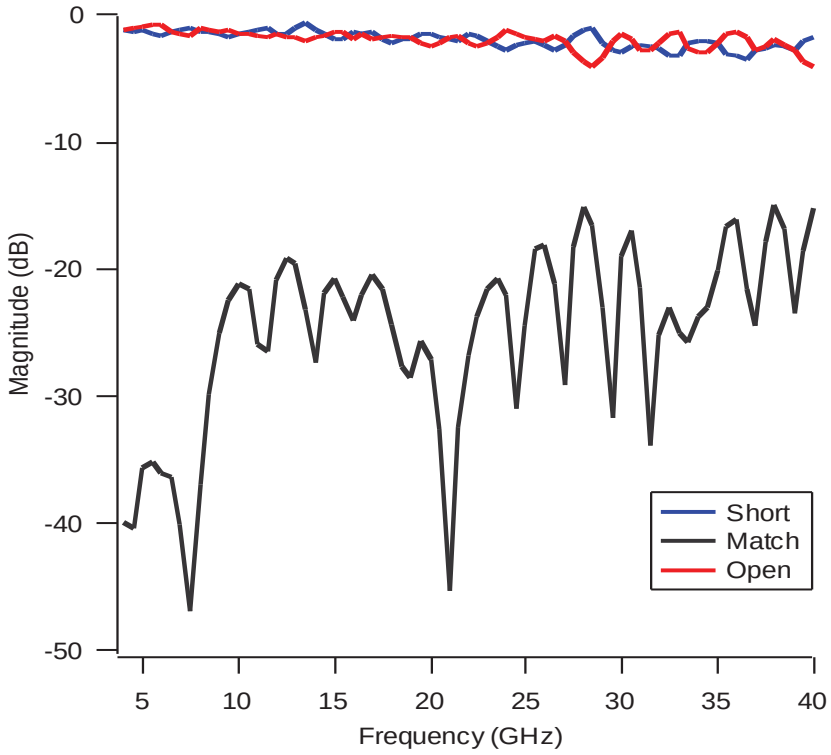


Figure 3-18, Raw data from coaxial 1 port measurements.

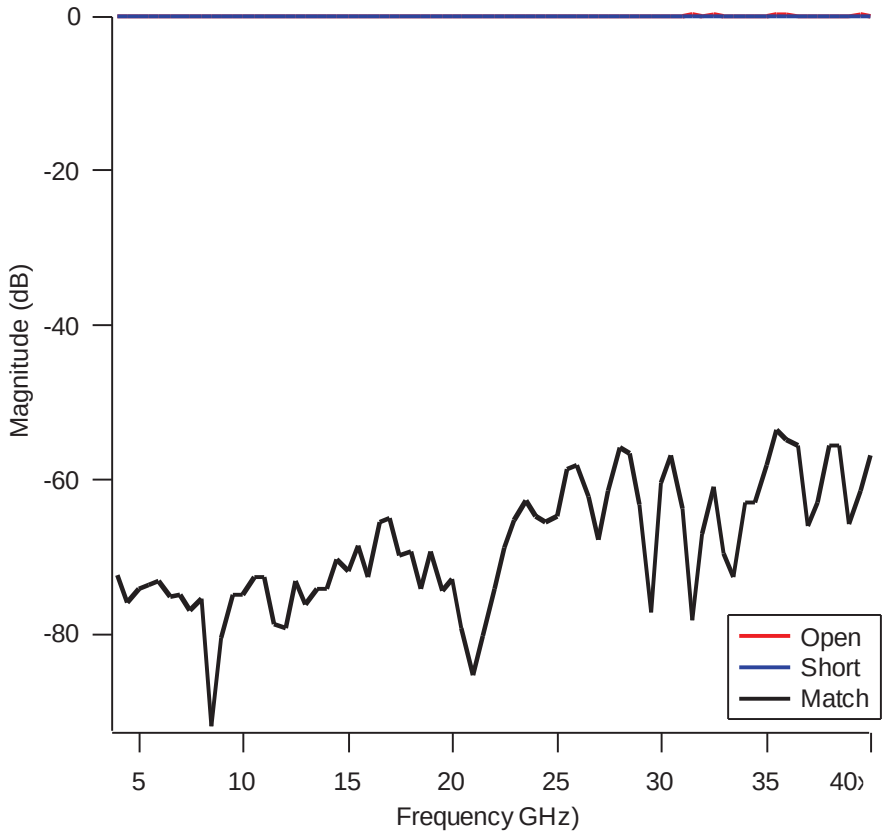


Figure 3-19, dynamic range of error corrected coaxial standards.

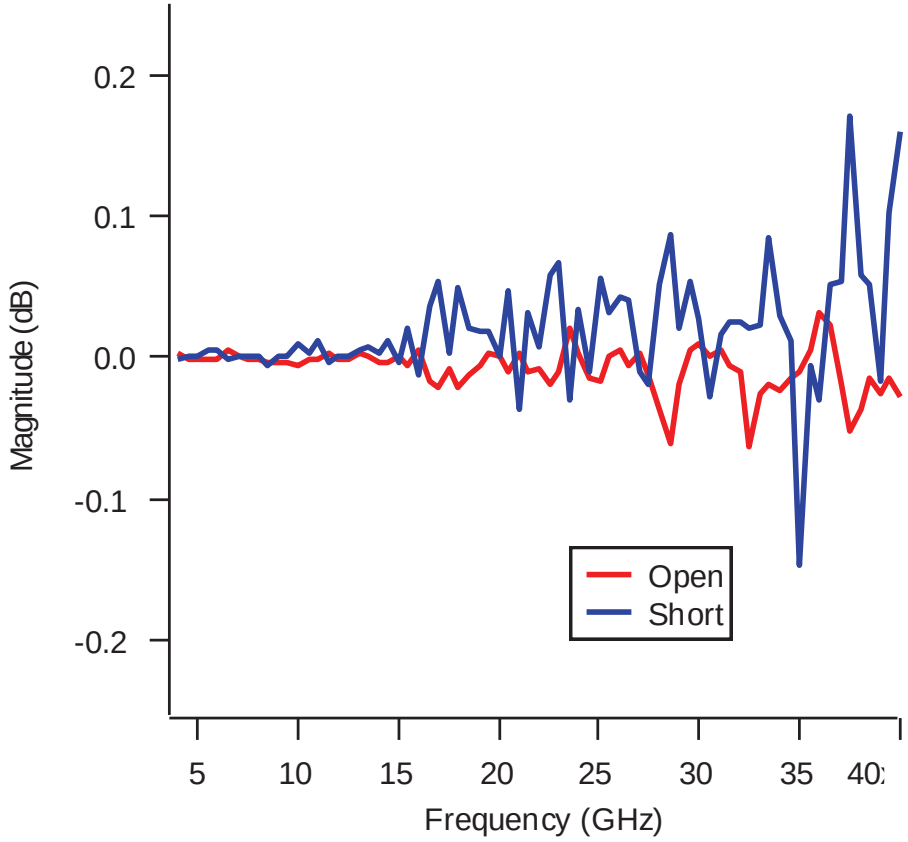


Figure 3-20, calibrated high reflection coaxial measurements

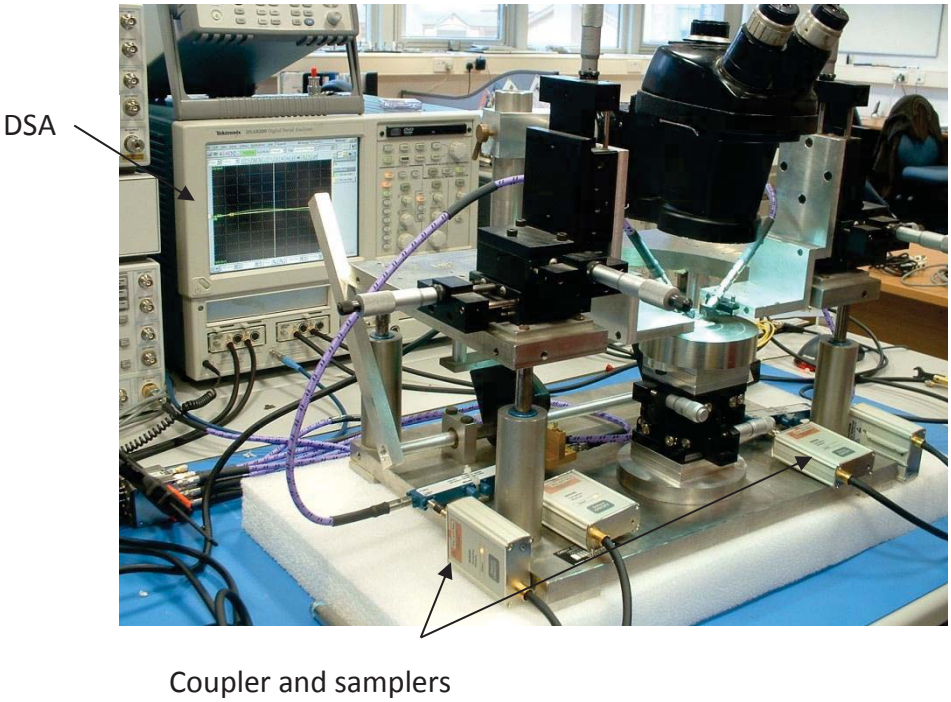


Figure 3-21, Initial High Frequency Sampling Measurement System.

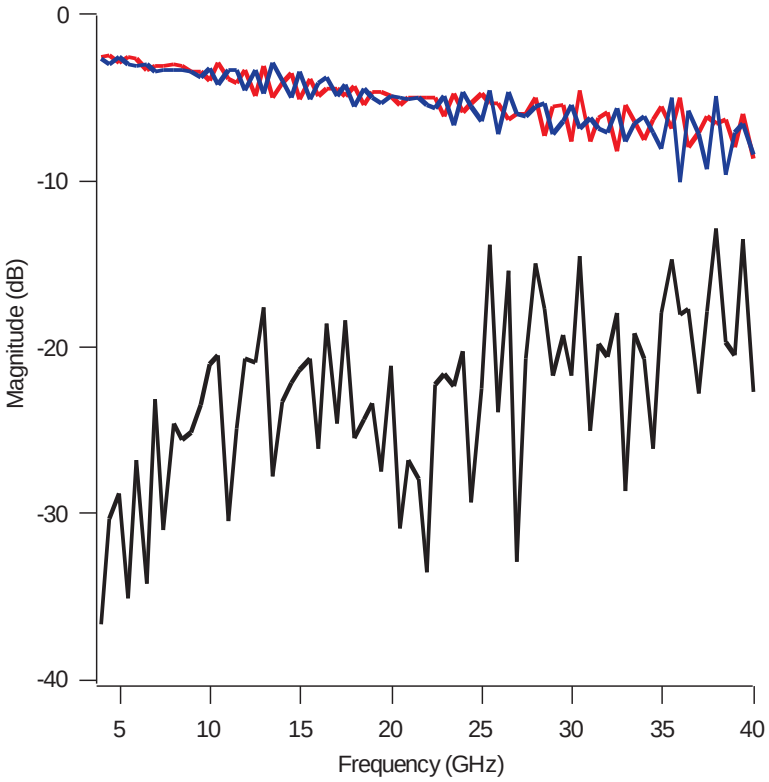


Figure 3-22, Raw data from wafer probe 1 port measurement of calibration standards.

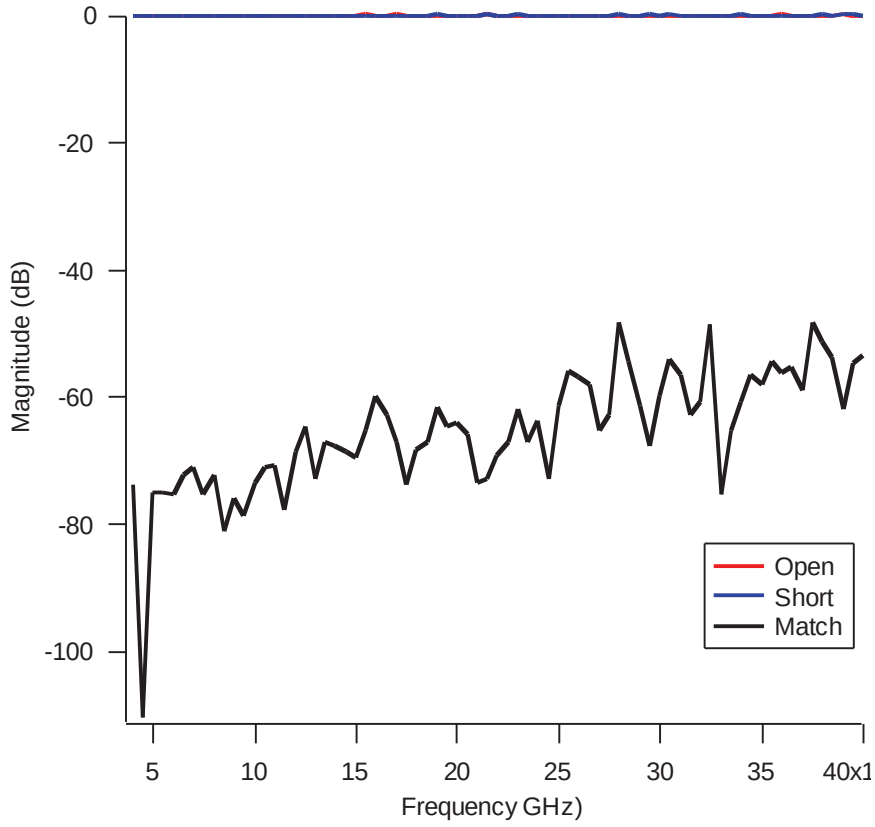


Figure 3-23, dynamic range of error corrected measurement of wafer calibration standards

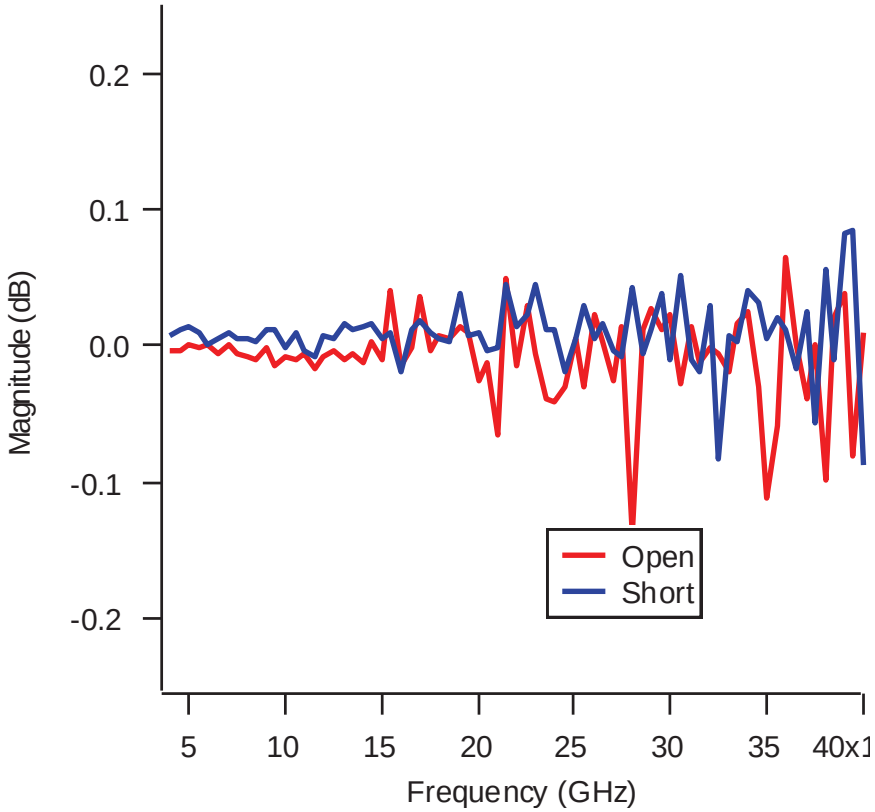


Figure 3-24, measurement of wafer probe open and short circuits after calibration.

3.5.3 Comparative Measurement with VNA

The final verification test was to compare a wide bandwidth non-50Ω impedance measured on the system and a traceable measurement system, in this case an HP (Agilent) 8510 VNA. A 6-18 GHz, 4 port antenna (3 ports were terminated in 50Ω) was chosen. Excellent agreement between the two systems is show in Figure 3-25; it should be noted that the measurement system used frequency steps of 500MHz, whilst the VNA used 180MHz, which accounts for some of the differences due to interpolating between points.

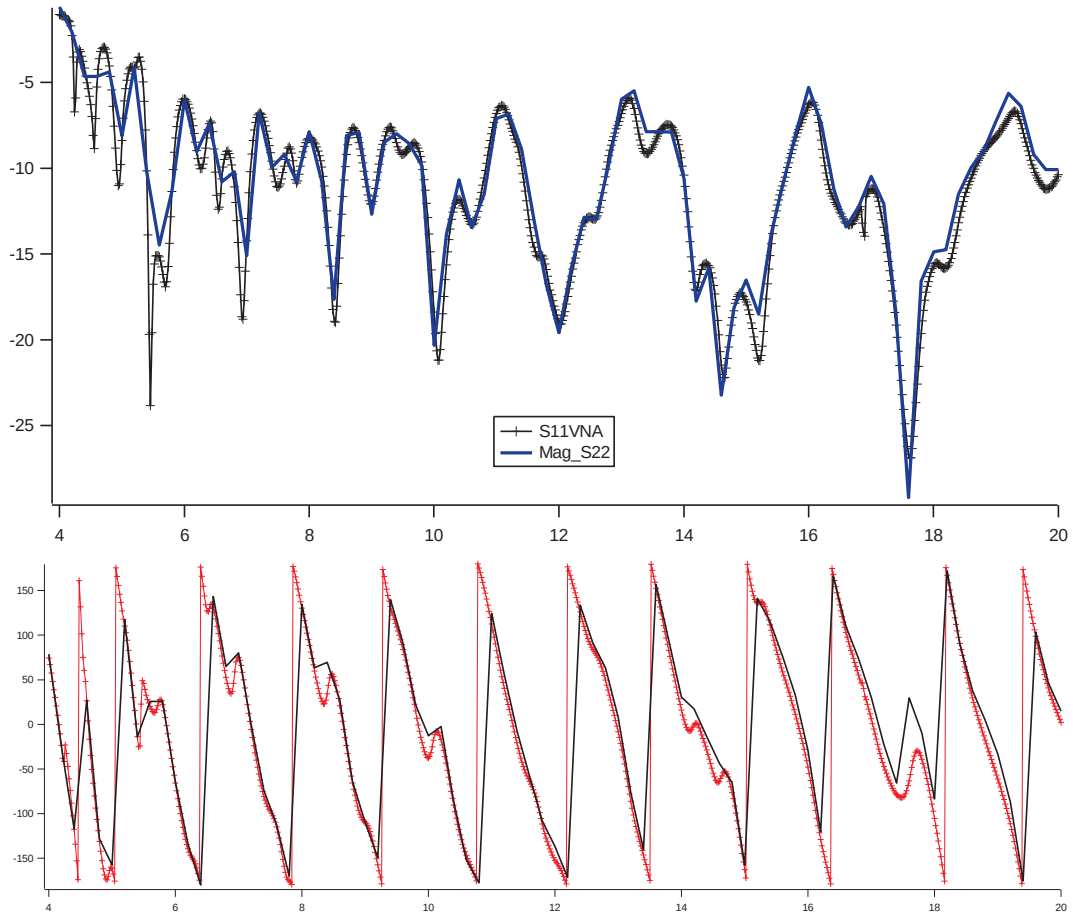


Figure 3-25, comparative measurement of wideband antenna port between Measurement System and HP8510 VNA. X-axis for both graphs are in GHz.

3.6 System Measurements

For the purposes of this research project the prime objective of the measurement system is to characterise on wafer active devices and then using active load pull to simulate the impedance environments produced by the designed matching circuits, including at the harmonic frequencies. Thus the measurement system is required to operate over varying power levels, determined by the device input reflection coefficient, gain, and output power,

(assuming the whole output impedance plane is required to be measured). The measurement system is independent of power level up to the linear limits of the sample heads (~+13 dBm). By fitting attenuators prior to the sampler (as shown in Figure 3-4) the maximum signal level can be adjusted such that power levels up to the limits of the coaxial components can be measured [12], thus power sweep measurements can be made.

To demonstrate the active device measurement capability a small signal chip transistor was measured in the system. Voltage and current waveforms measured on the output of the transistor with increasing input power levels are shown in shown in Figure 3-26, the squaring of the current waveforms shows the increase in harmonic levels with drive power. The system captures successive input and output waveforms at each power level, typically between 256 and 512, and averages them. From this data a large number of device parameters can be determined; for example by performing a Fast Fourier Transform on the measurements the information can be viewed in the frequency domain. In this way the spectral components can be measured. The fundamental, 2nd and 3rd harmonic power levels of the waveforms in Figure 3-26 are shown in more conventional form in Figure 3-27. Note the upper harmonic that can be measured is limited by the coupler response and the system noise floor. The system performance was cross referenced with the measurement

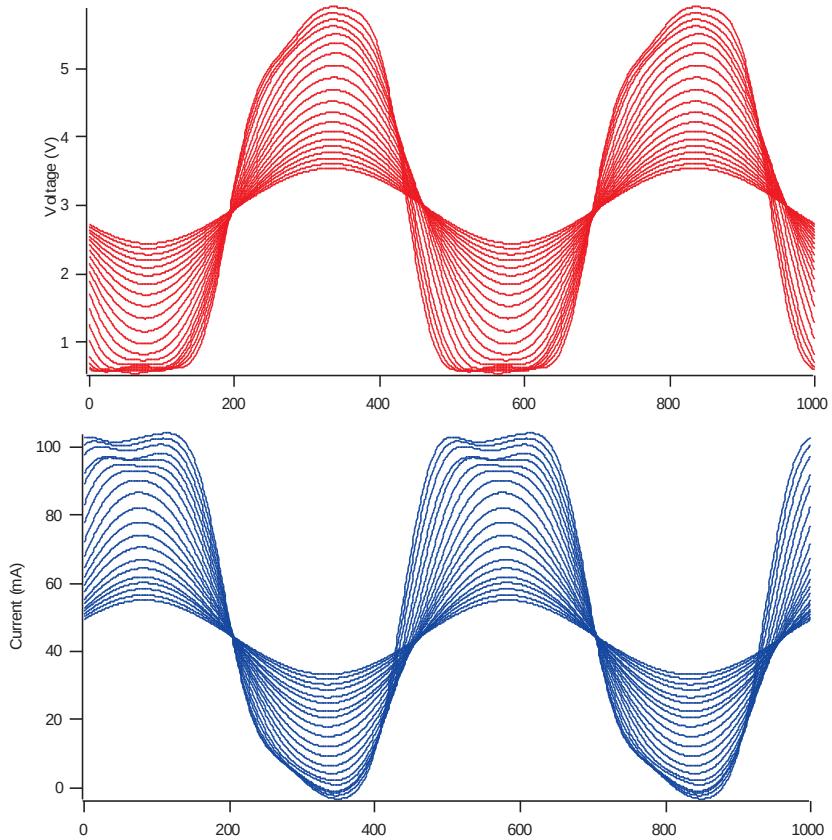


Figure 3-26, Output Voltage & Current Waveforms from test transistor with increasing Input Power.

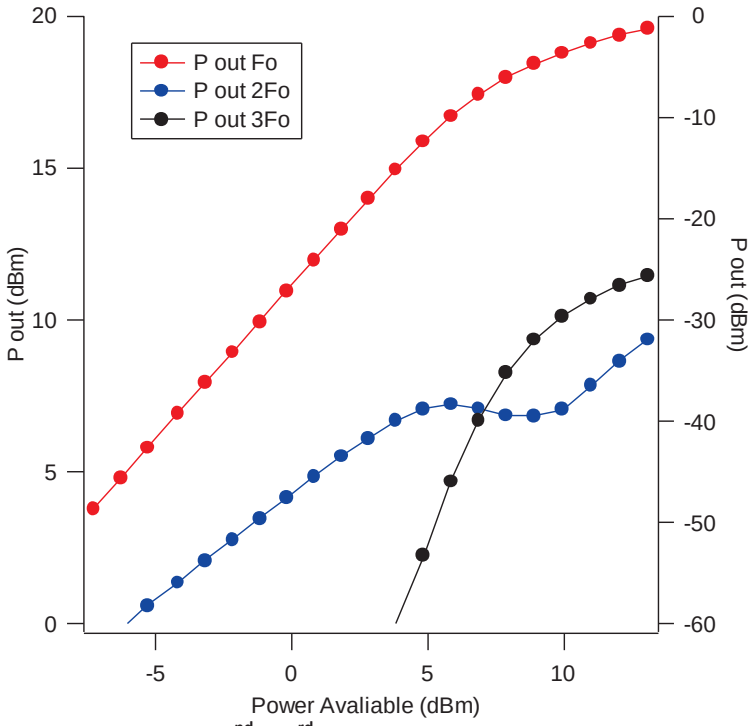


Figure 3-27, Fundamental (left axis), 2nd & 3rd harmonic output power (right axis) with increasing Input Power.

of the relative fundamental and harmonic levels using a signal generator and a spectrum analyser.

Key to the development of high power microwave amplifiers has long been the use of load pull design techniques, [13]. This method presents the appropriate impedance to the device (traditionally using mechanical tuners) and the output power is measured. Thus by searching the load impedance plane the optimum load impedance can be determined (for saturated power, linearity, efficiency, etc.). With active load pull systems, typically the output from the device is passed through a circulator and 'dumped' in a load. A new coherent signal of controlled phase and magnitude is passed through the circulator back to the device. Thus by adjusting this signal any load impedance across the Smith Chart can be simulated. Further signals can be injected at the harmonics so that the fundamental and harmonic impedances can be separately controlled. In this case as the output power of the device was well below the maximum reverse power that the load pull signal generator can tolerate and thus this can be used as the device load, (although care should be taken that reverse intermodulation products are not generated – a good signal generator will be protected from this effect however it is always advisable to check). The magnitude and phase of the injected signal was varied and the output power of the device under these conditions was measured. The results were plotted on the Smith Chart, Figure 3-28. This

shows not only the optimum load impedance for maximum power but also the sensitivity of output power to the load.

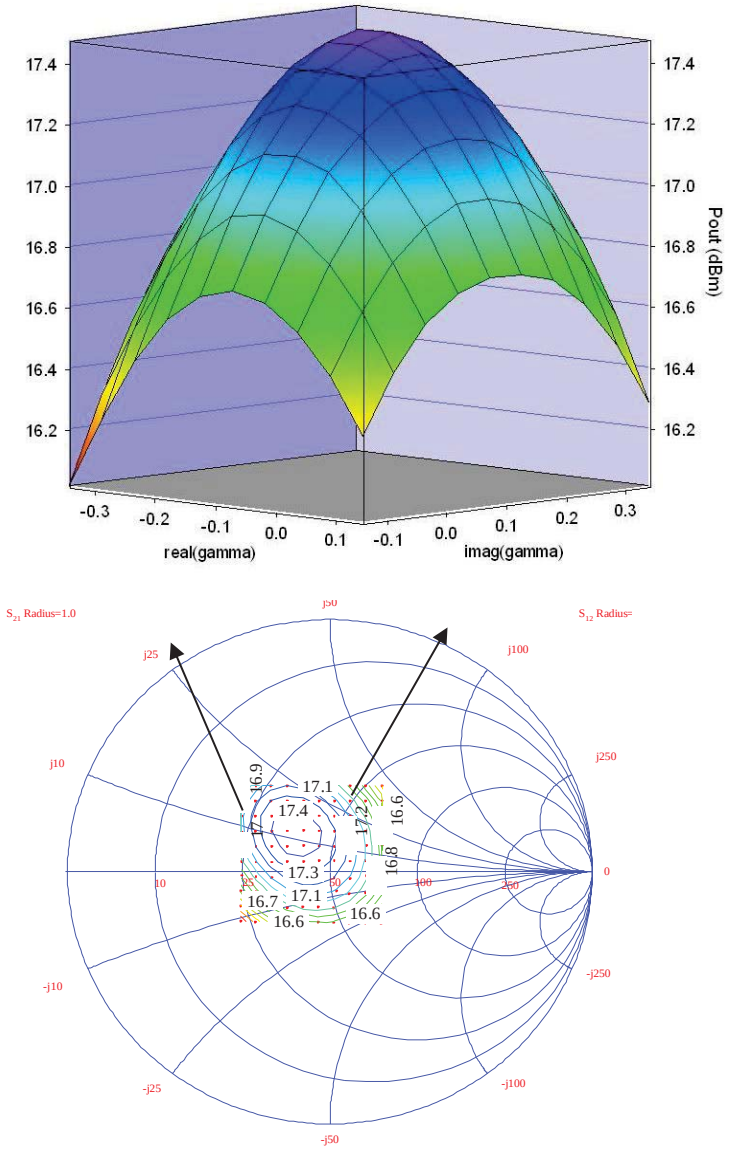


Figure 3-28, Active Load Pull Power Contours.

The size of the measurement grid and the number of power levels measured depends upon many factors, not least of which is time. For the power sweep, ideally only a few data points are needed in the linear region of the device, however as the device moves into compression a finer step size is required as interpolation becomes more difficult. Often a compromise step size is used which results in more data than necessary at low power levels and less than ideal at higher levels. The area of the Smith Chart mapped is to a large extent determined by application. For narrow bandwidth high power applications the immediate area around the optimum PAE and Output Power loads may be all that is

necessary, enough to cover the likely variation in load impedance presented by the practical output matching circuit, whereas for a wide bandwidth driver device that may be presented with non-optimum impedances by succeeding stages, a larger impedance area will need to be covered. Another aspect that will be discussed in more detail later is the use of the system to map the effects of the harmonic terminations whilst holding the fundamental impedance constant at a particular level. Indeed it is even possible with an active load pull system to exceed the boundaries of the Smith Chart and thus examine the effects of harmonic injection, [2].

As a conclusion to the initial system proving measurements, an opportunity arose to configure the system in an unusual way to measure a coaxial 4 port antenna that required all the ports to be stimulated simultaneously. Multi-port devices, which require more than one port to be stimulated simultaneously, can pose particular problems for standard test equipment; differential device measuring systems are available, however for more ports bespoke systems are generally required. Using the high frequency sampling technique a 4 port antenna requiring all the ports to be driven in phase quadrature was tested as shown in Figure 3-29, and a photograph of the couplers, samplers and antenna are shown in Figure 3-30. Due to the symmetry of the structure it was only necessary to measure 2 of the ports. To measure more ports 2 additional samplers (the DSA8200 can accommodate up to 8 samplers) and a directional coupler per port is required. The results, Figure 3-31, appear to show a positive return loss on port 2 at some frequencies, which is not possible from a purely passive structure. The probable explanation for this is that there is limited isolation between the ports and hence at some frequencies the breakthrough from the other port will add constructively to produce a greater returned signal to the port than the incident signal. Rotating the phase of the unmeasured ports by 180° alters the levels of these breakthroughs and although the absolute level of the return loss changes the general shape of the match is similar, Figure 3-32.

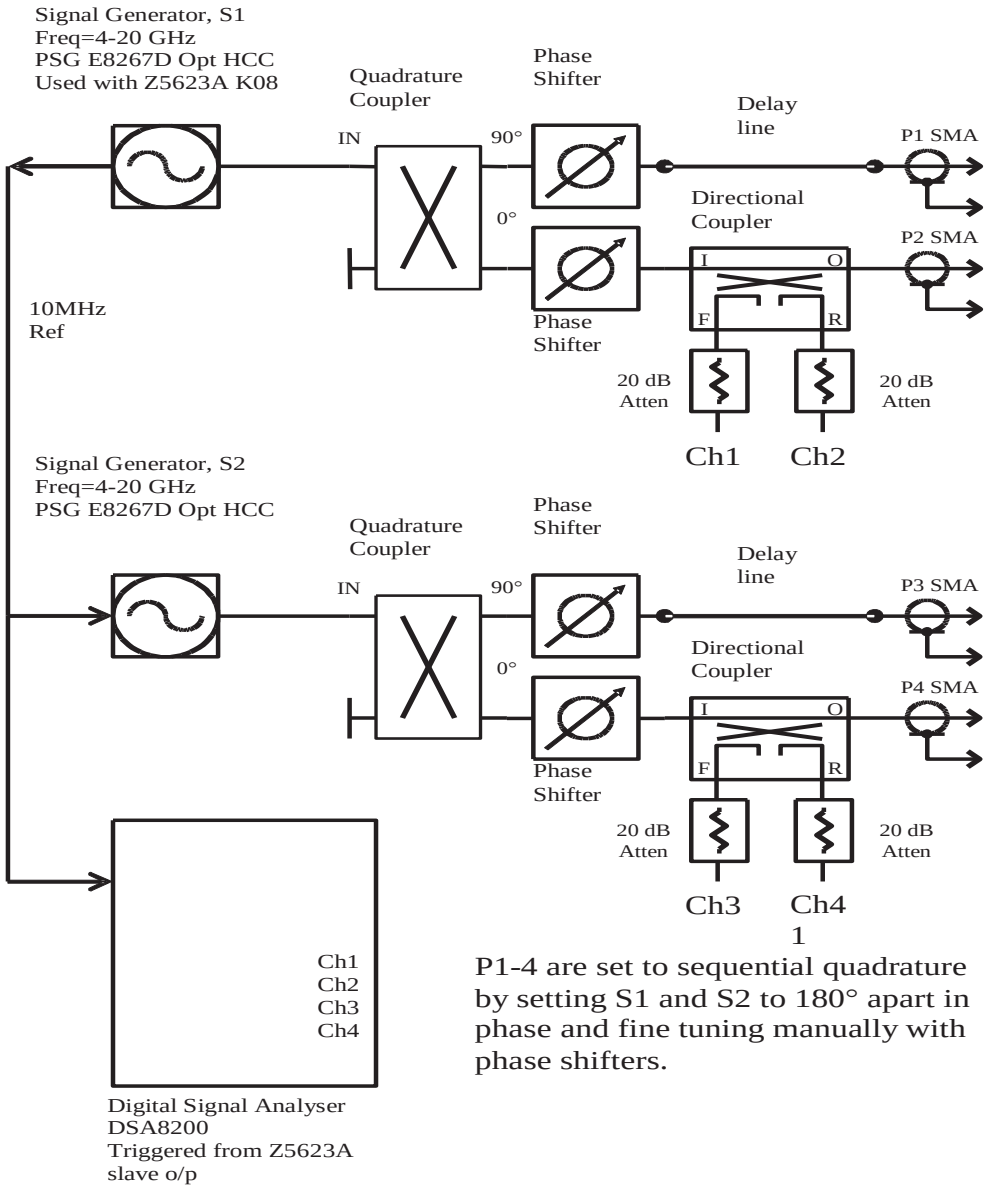


Figure 3-29, Four port antenna measurement system.

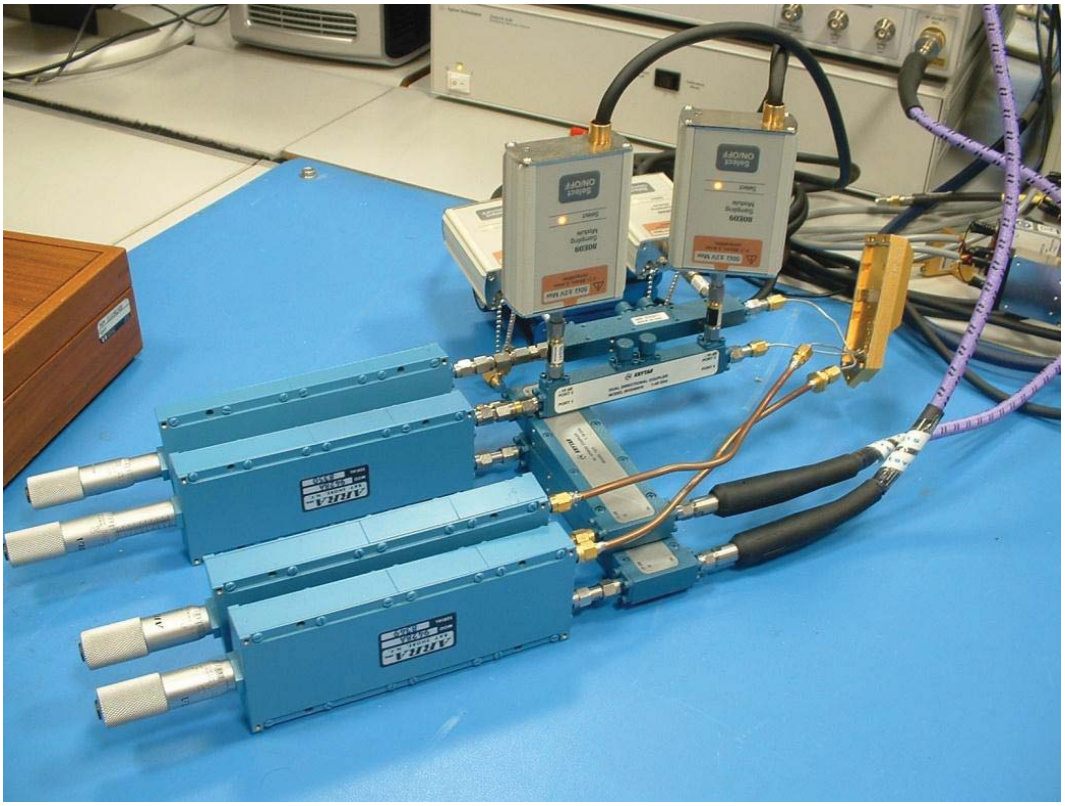


Figure 3-30, Coupler, sampler and phase shifter arrangement for 4 port antenna measurement

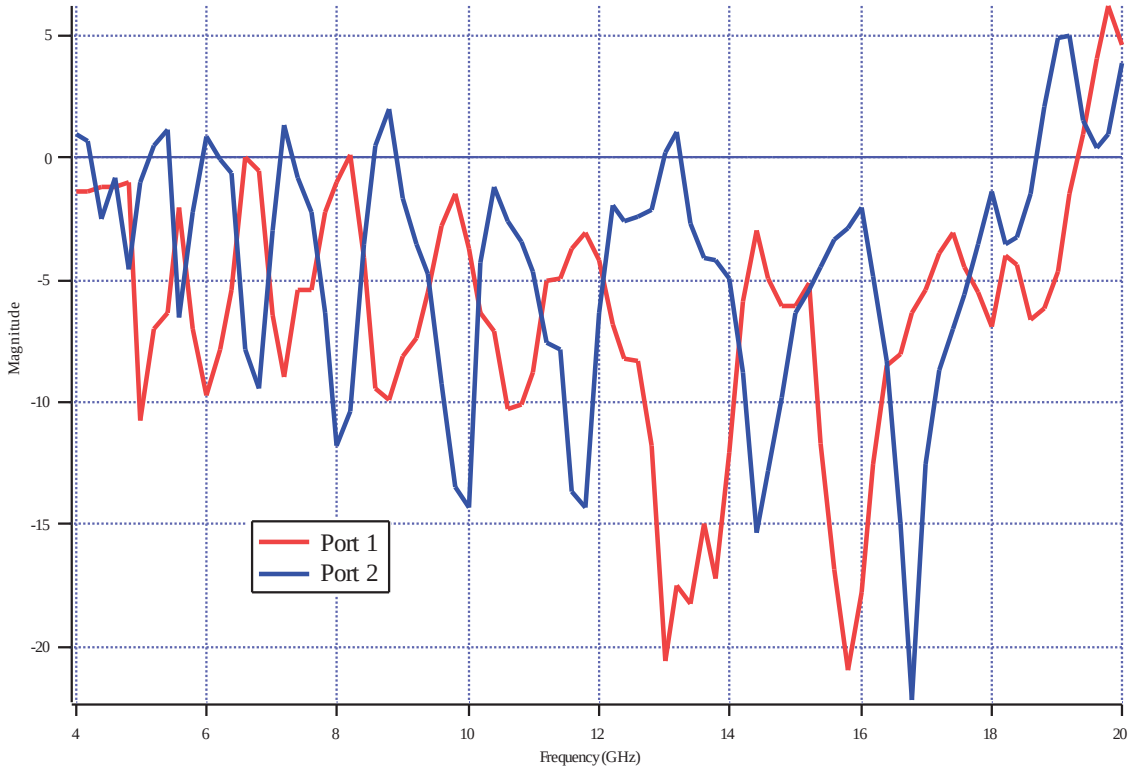


Figure 3-31, Match presented by 4 port antenna - all ports driven in quadrature.

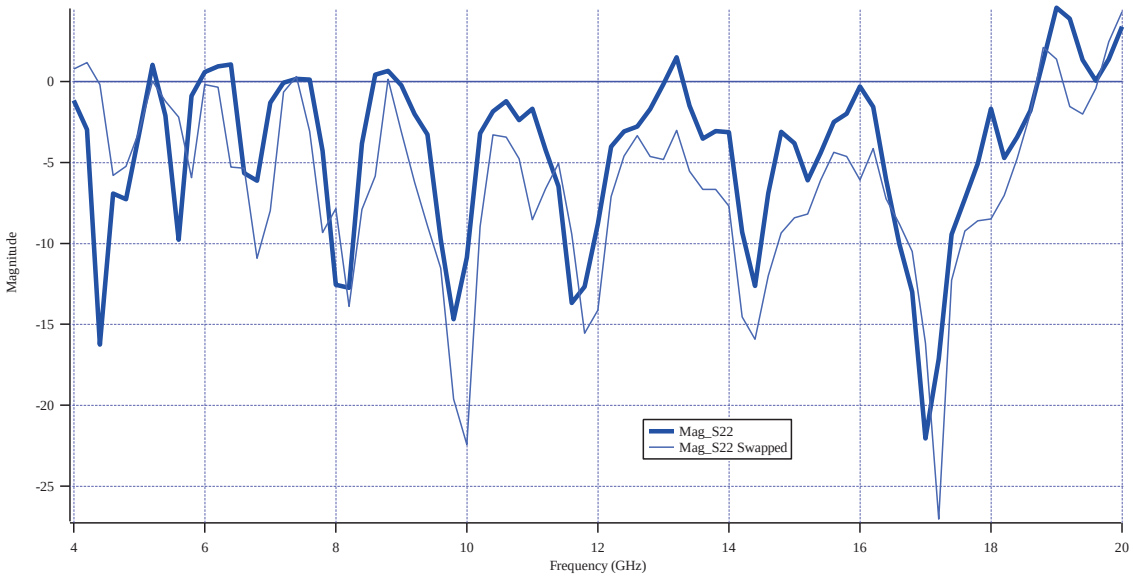


Figure 3-32, Effect of rotating the phase of the unmeasured ports by 180°

3.7 System Improvements

This work was the first attempt at creating a measurement system capable of very wideband measurements, previous systems, [3], [9], had been targeted at specific communications bands. Although some of the issues had been anticipated based on this earlier experience, there were still some surprises and some areas where the initial precautions were not sufficient.

3.7.1 System Losses

Whilst measuring the small signal device it was found that the losses of the in-line components (switches, coaxial lines, etc.) became a considerable limitation on the systems' ability to inject power into wafer probed devices. This was due to the high input reflection coefficients such devices presented. In order to be able to drive these devices into the non-linear region it is either necessary to include pre-matching (either laid out with the device or using a passive tuner) or increase the injected power levels. For the size of devices currently under consideration input power levels of <0.5W (27 dBm) are required. As will be described later in chapter 4 on device measurements, some broadband matching can be incorporated on the input to the device in the layout cell. This does however require some knowledge of the devices input impedance and complicates the de-embedding to get at the actual device plane. For most measurements undertaken in this research the input power

levels were achievable with the broadband amplifiers available in the laboratory and with some measures taken to minimise the path losses.

On a previous system it was possible to mount the couplers adjacent to the wafer probes, as shown in Figure 3-33. However, in this case the samplers were inside the measurement instrument and connected to the coupler via cables. In the new system the samplers were able to be connected directly to the measurement coupler which consequently increased the weight and size and made it impractical for direct connection to the wafer probes. Further, the coaxial connectors of the 40 GHz components are of the 2.4mm variety, whilst the wafer probes were 3.5mm. The solution used was for the couplers to be mounted into custom made Perspex blocks fixed to the measurement system platform as close to the probe stage as possible. A custom semi-rigid cable with 2.4mm connectors on one end and 3.5mm on the other was connected between the coupler and probe. This had sufficient flexibility to permit the limited movement of the probes. The final system is shown in Figure 3-34.

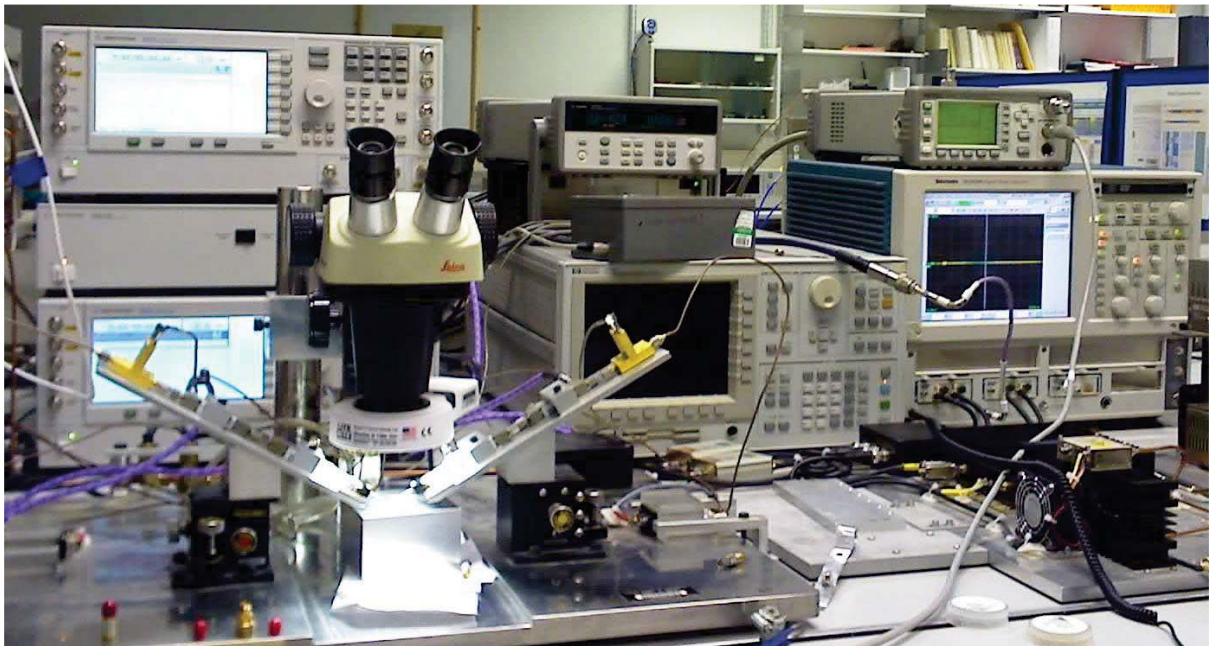


Figure 3-33, 20GHz remote sampler measurement system, note the couplers connected directly to the wafer probes.

The power levels needed to achieve the required load termination impedances necessitated the provision of high power amplifiers. The power required by the load pull sources is determined by:

a) Extent of impedance plane coverage.
b) Maximum device output power.
c) Losses in the system.
d) Differences between the device output impedance and the measurement system impedance.

To quantify the impact of this last point (c); where there is a difference in impedance, a mismatch will occur which causes power to be reflected. This mismatch loss can be calculated as follows:

Consider two impedances Z_S and Z_L,

$$\Gamma = \left|\frac{Z_S - Z_L}{Z_S + Z_L}\right| \qquad \{3\text{-}4\}$$

and mismatch loss L_M,

$$L_M = 10\log(1 - \Gamma^2) \qquad \{3\text{-}5\}$$

Furthermore the power delivered by the device will depend upon the load presented to it.

Figure 3-34, Final 40GHz measurement system. {1} Coupler & Sampler Block, {2} Probe, {3} Driver amplifier, {4} Bias Tee, {5} Diplexer, {6} Switch Box.

The power delivered by the device, P_d is dependent upon the available power from transistor, P_{av}, the devices optimum load reflection coefficient, Γ_{LD}, and the load presented to the device at the measurement plane Γ_{LP}, [3].

$$P_d = \frac{(1 - \Gamma_{LP}{}^2)(1 - \Gamma_{LD}{}^2)}{(1 - \Gamma_{LP}\Gamma_{LD})^2} \times P_{av} \qquad \{3\text{-}6\}$$

and the power required from the load pull source, P_{LP}:

$$P_{LP} = \frac{\Gamma_{LP}{}^2}{1 - \Gamma_{LP}{}^2} \times P_d \qquad \{3\text{-}7\}$$

Previous [12] work with high power devices has shown how the drive power requirements escalate with increasing output power and the corresponding fall in output impedance as shown in Figure 3-35. In the case shown in order to produce the required 100W of device power the load pull source needs to provide 688W. Note the positive and negative values of reflection coefficient refer to the case where Z_S is < Z_L (negative) and Z_S > Z_L (positive). Primarily in this research we have been concerned with measuring in a 50Ω system and at power levels <10W, hence looking at the load pull power requirement encountered for a 0.5W 0.3µm GaAs pHEMT [14], Figure 3-36, we find that we only require 186mW (ignoring system losses) to achieve the optimum match. In practice the system losses from bias tees, diplexer and cables at these frequencies amount to about 3dB and

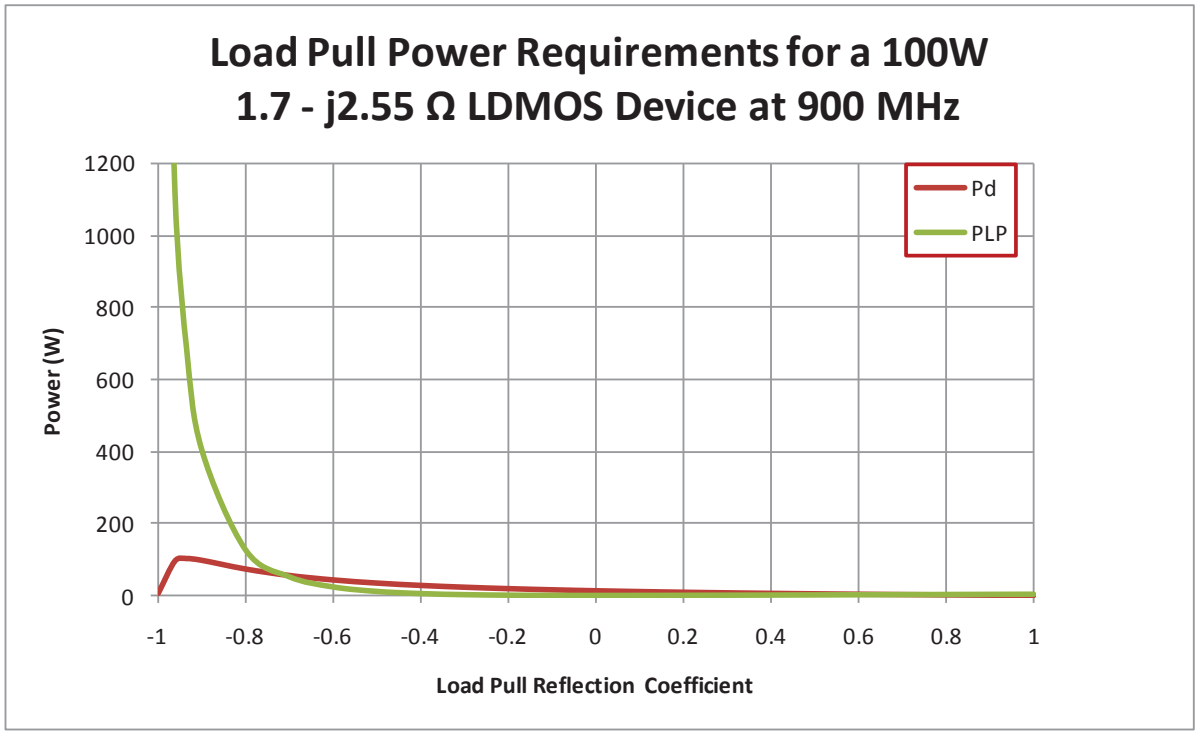


Figure 3-35, Load pull power requirements for a 100W LDMOS FET at 0.9 GHz

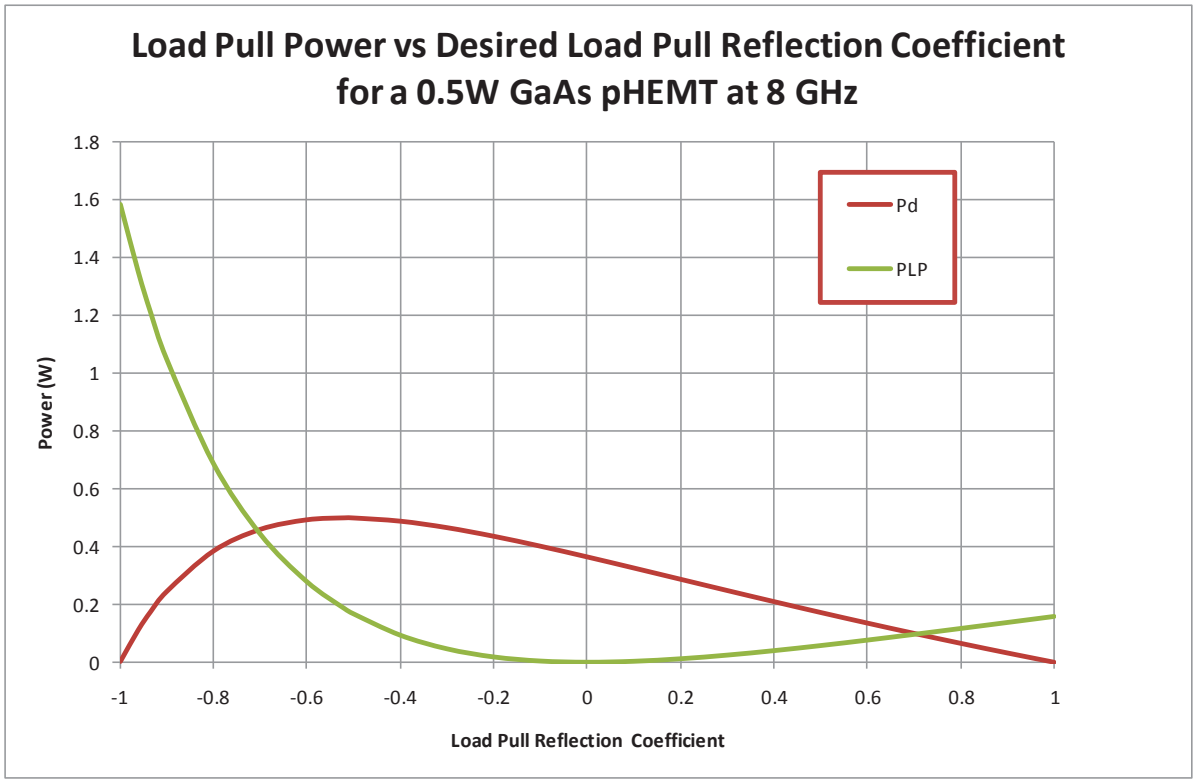


Figure 3-36, Load pull power requirements for a 0.5W 0.3µm GaAs FET at 8 GHz.

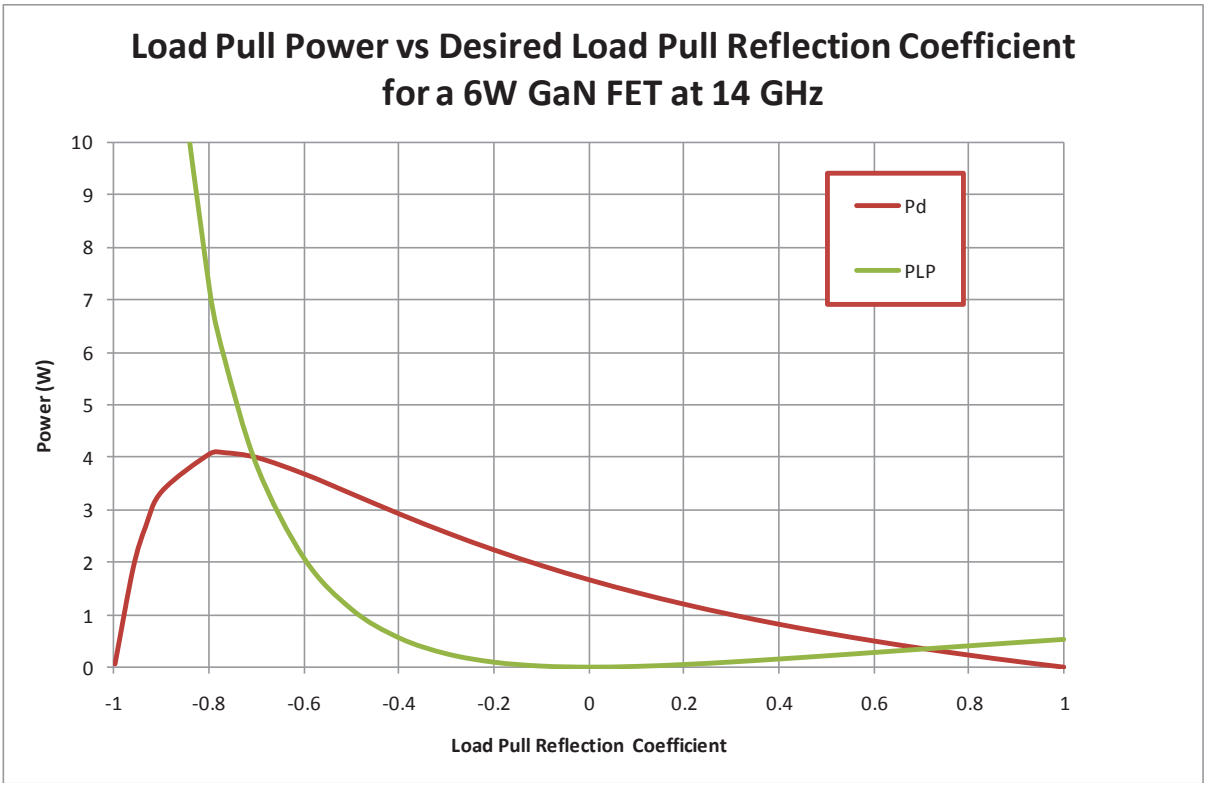


Figure 3-37, Load pull power requirements for a 6W 0.25µm GaN FET at 14 GHz.

thus the actual load pull power requirement is ~370mW (25.7 dBm).

With the advent of Gallium Nitride (GaN) devices higher powers, higher frequencies and increased system losses combine to substantially increase the load pull power requirements [15] as shown in Figure 3-37. Again, in practice to these power levels must be

added the system losses, which increase the power requirement at the peak device power from 5.9W to nearly 15 W.

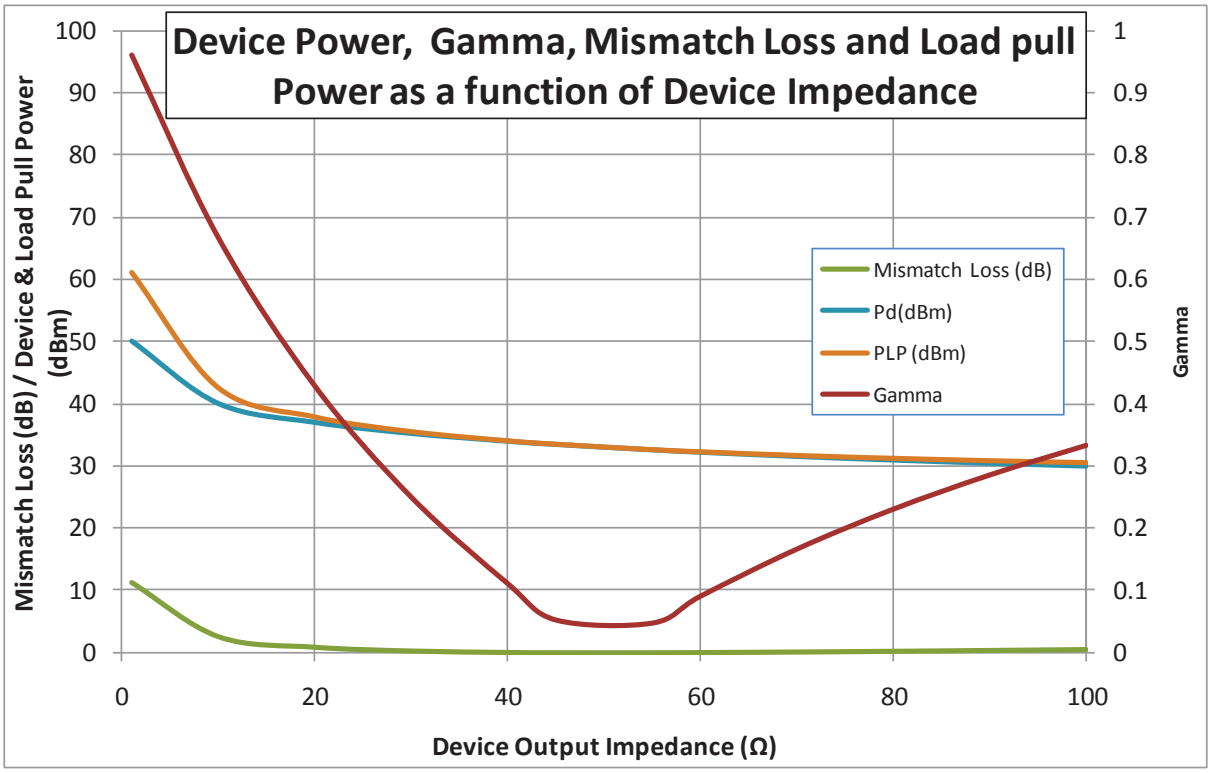


Figure 3-38, Impact of Device Output Impedance on Mismatch Loss, Gamma and Load Pull Power assuming Output power is inversely proportional to device impedance.

Assuming that there is an inverse relationship between device output power and its output impedance a graph can be plotted that shows the relationship between these key components in a 50Ω measurement system, Figure 3-38. From this we can see that as the device impedance falls below 10Ω the load pull power required increases dramatically.

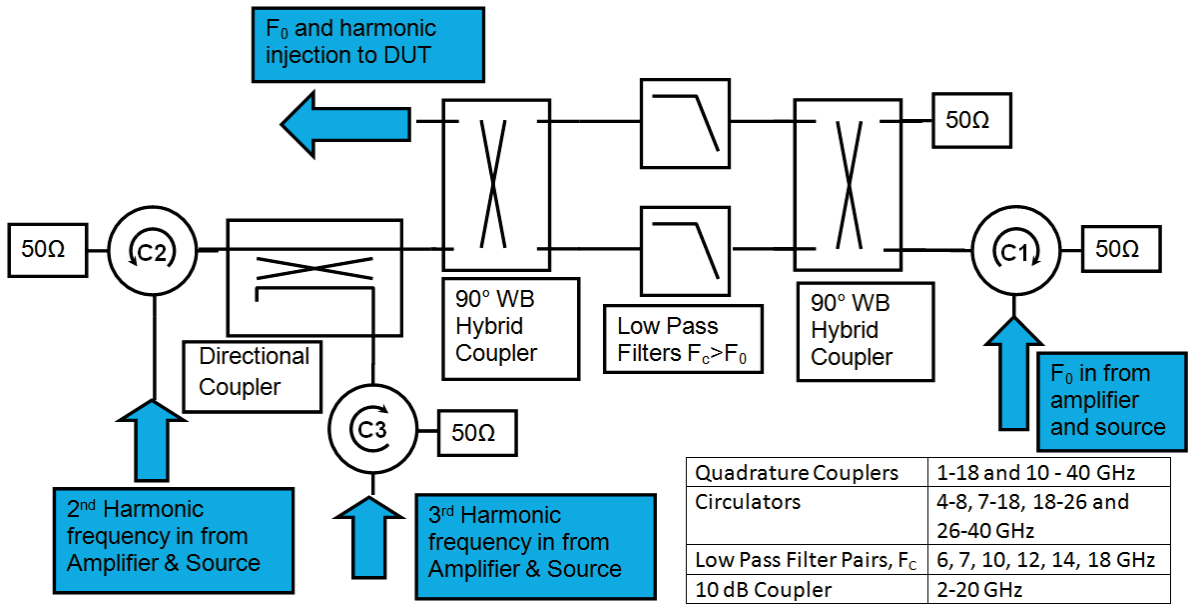
3.7.2 Broadband Triplexer

In order to obtain the maximum efficiency from the device it is necessary to control the harmonic terminations as well [6] as the fundamental. Although the power levels at the harmonics are typically >10dB below the fundamental at compression, they are also often not only required to produce higher reflection coefficients than the fundamental, but also the system losses increase with frequency. Key to being able to produce fundamental and harmonic a_{2n} signals at the device is a multiplexer, which allows the signals from multiple sources to be combined into a single path. From the discussions above it is clear that this component of the system must also be as low loss as possible. In narrow band systems it is possible to procure dedicated triplexers (for up to 3rd harmonic control) as the communication band frequencies are well defined, e.g. 800, 900, 1800, 2100MHz. For a

broad band measurement system it is necessary for the multiplexer to be broadband so that it can be used for multiple frequencies, but also flexible so that it can be re-arranged to accommodate an even wider range of frequencies, (a single Triplexer solution is not possible as for example, a frequency may be the 2nd harmonic in one measurement and the fundamental in the next). In the measurement system used in this research only up to the 3rd harmonics were considered.

The Triplexer was constructed from wideband building blocks so that each configuration could be used over a reasonable bandwidth, but also so that by changing a few components the operating range of the Triplexer could be adjusted. The basic construction is shown in Figure 3-39, and a photograph of the implemented arrangement for 9 GHz fundamental frequency is shown in Figure 3-40. The component combinations used for various measurements are shown in Table 3-5.

It should be noted that the load for the circulator C1, must be able to handle the forward power from the device. For the devices measured in this work the standard loads available were marginal on power; however the additional loss of the diplexer arrangement was of benefit in this respect. For devices with output powers much above 0.5W higher power loads will be necessary, (a low cost method of providing this attenuation is to use semi-rigid cable; for example a 1 metre length of 2.2 mm outer diameter semi-rigid cable will have a nominal attenuation of 2.6 dB at 10 GHz and 3.9 dB at 20 GHz [16], and this also allows the load to be positioned remote from the circulator).



Quadrature Couplers	1-18 and 10 - 40 GHz
Circulators	4-8, 7-18, 18-26 and 26-40 GHz
Low Pass Filter Pairs, F_C	6, 7, 10, 12, 14, 18 GHz
10 dB Coupler	2-20 GHz

Figure 3-39, Triplexer arrangements for wideband 3 harmonic active load pull.

The basic operation of the wideband Triplexer is as follows; the b_{21} wave from the DUT is incident upon the left hand side (LHS) hybrid coupler of Figure 3-39 and is split between the two paths on the right hand side. As the frequency F_0 is below the cut-off frequency, F_C, of the Low Pass Filter (LPF), the b_{21} wave passes through and is recombined in the hybrid on the right. By the quadrature nature of the hybrid the combined signal exits at the lower port. The b_{21} wave passes through the circulator C1 and is dissipated in the 50Ω load. The generated a_{21} signal at F_0 is injected into the 3rd port of C1 and travels round to the arm connected to the hybrid. This wave travels through the LPF and hybrids in the reverse manor to the b_{21} wave and leaves the Triplexer towards the DUT. For the harmonic frequencies of F_0, the b_{2n} waves, pass through the LHS hybrid and are reflected back by the LPF, as they are above F_C. They exit the hybrid via the lower arm and pass through a circulator (C2) chosen for the harmonic frequency band and are dissipated in the load.

F_0 (GHz)	F_2 (GHz)	F_3 (GHz)	Quad (GHz)	LPF (GHz)	C1 (GHz)	C2 (GHz)	C3 (GHz)
4	8	12	1-18	6	4-8	7-18	7-18
5	10	15	1-18	6	4-8	7-18	7-18
6	12	18	1-18	7	4-8	7-18	7-18
7	14	21	1-18	10	4-8	7-18	18-26
8	16	24	1-18	10	7-18	7-18	18-26
9	18	27	1-18	10	7-18	18-26	26-40
10	20	30	10-40	12	7-18	18-26	26-40
12	24	36	10-40	14	7-18	18-26	26-40
14	28	-	10-40	18	7-18	26-40	-
16	32	-	10-40	18	7-18	26-40	-
18	36	-	10-40	18	7-18	26-40	-

Table 3-5, Triplexer components used for harmonic active load pull measurements. (Labels refer to Figure 3 39).

To combine the a_{22} and a_{23} waves the a_{22} is injected via C2, whilst a_{23} is coupled into the path using a 10dB directional coupler and circulator C3. Although this is a lossy solution for the 3rd harmonic, the power levels required for these frequencies were significantly lower and hence this provided an acceptable solution, minimising the loss at the 2nd harmonic. A lower loss solution would have been to replicate the hybrid LPF arrangement but at a frequency of $2F_0$.

Measured results for the individual LPF (13 and 15GHz) and then the combination creating the Diplexer are shown in Figure 3-41 to Figure 3-44, the 13 GHz Diplexer uses the 1-18GHz quadrature hybrid and the 15GHz the 10-40GHz. An important point to note is that

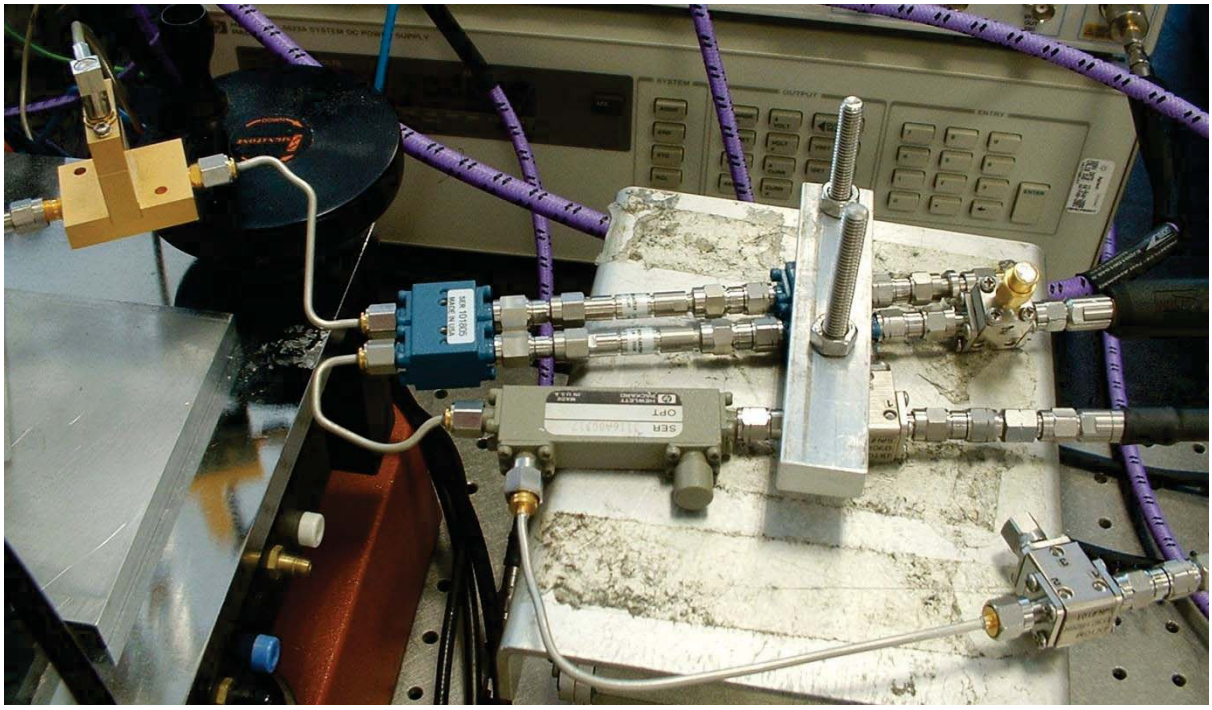


Figure 3-40, 9 GHz Triplexer for 3 harmonic active load pull.

the when in Diplexer arrangement the frequency response becomes bandpass rather than the low pass of the filters.

This characteristic can be utilised to make the Diplexer a bias 'Tee' as well. The low pass path is to the opposite port to C1 on the RHS of the Diplexer of Figure 3-39. The 50Ω load is removed and drain bias can be introduced through this point. In order to prevent low frequency oscillation due to the LF impedance presented at this point, damping should be included on the port; this can simply be a parallel R-L in series with the bias and a series 10nF capacitor and 50Ω resistor.

Initially it was thought practical to design and build bandpass filters for the system, and if successful then attempt to design our own diplexer. Two prototypes were built, using a short circuit stub configuration as described in [17] and [18]; the starting values (before optimisation) for the two circuits are given in Table 3-6. Two microstrip filters were designed, a 6-12GHz and a 9-18GHz, on Rogers TMM3 (0.762 mm thick) substrate. The circuit schematic, E-M layout and simulation results of the 6-12GHz filter are shown in Figure 3-45 to Figure 3-47. Although the E-M prediction showed that the circuit response would be shifted down in frequency the design was manufactured (on the in-house PCB milling machine) as there was no reference data available indicating the accuracy of either simulation. The manufactured circuit was measured on the HP 8510 Vector Network Analyser and was seen to roll-off at the top end of the frequency range as predicted by the E-M simulation. Two circuits were manufactured and the results are shown in Figure 3-48.

When the 9-18 GHz version was designed the difference between the E-M simulation and the linear was even more marked. This gave rise to the thought that the simulation of the Short Circuit stubs may be an issue, particularly the ground transition.

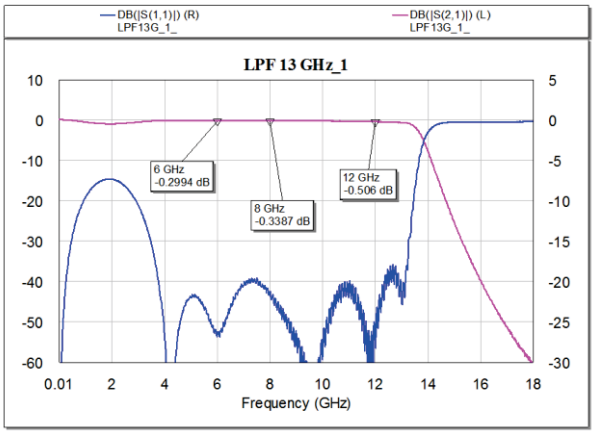


Figure 3-41, Measured 13GHz low pass filter performance.

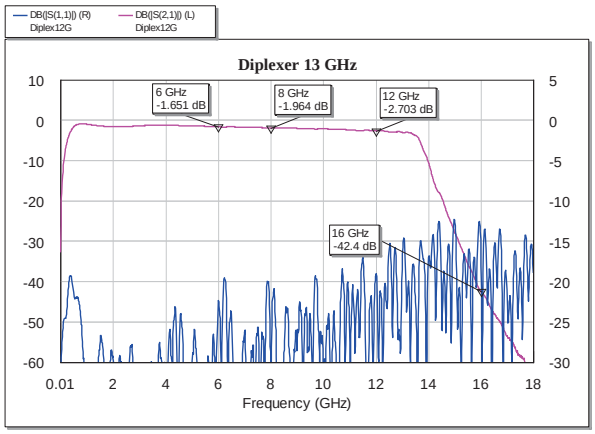


Figure 3-42, Measured 13GHz Diplexer performance.

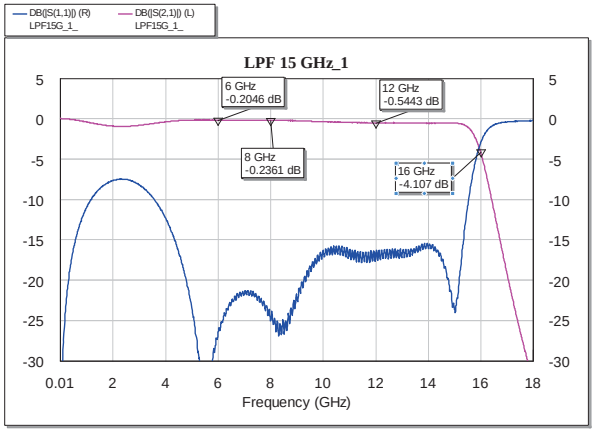


Figure 3-43, Measured 15GHz low pass filter performance.

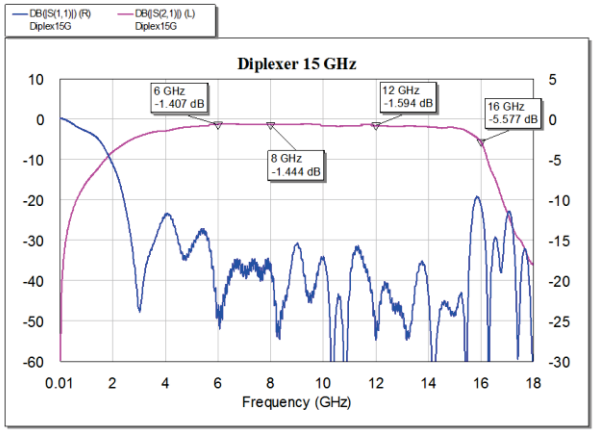


Figure 3-44, Measured 15GHz Diplexer performance.

Element Label	Z or ∅	Dimension @ 9GHz (mm)	Dimension @ 13.5GHz (mm)
W50	50Ω	1.80	1.83
L50	90°	5.11	3.38
WO	115Ω	0.3	0.31
LO	46.3°	2.8	1.86
WI	63.4Ω	1.2	1.22
LI	31.1°	1.8	1.19
W1	63.4Ω	1.2	1.22
L1	102.1°	5.9	3.9
W2	45.3Ω	2.1	2.13
L2	108.1°	6.1	4.03
W3	104.4Ω	0.4	0.41
L3	75°	4.5	3.0

Table 3-6, Starting element parameters and dimensions on TMM3, Er=3.27 H=0.762, for 6-12 and 9-18 GHz Bandpass Filters.

The filters had been intended to be grounded using a 'wrap-around' technique, whereby copper tape would be folded over the edges of the board and soldered to the top and bottom. Looking at the layout in Figure 3-46 it can be seen that before a signal reaches the grounded edge there is an impedance step when the short circuit stub meets the ground plane. In the schematic these stubs were modelled simply as shorted lines. For the 9-18 GHz filter the ground plane changes were included in the linear circuit model and via holes (pins in practice) added, as shown in Figure 3-49. This improved the correlation between the linear and E-M simulations in the lower half of the band; however the E-M simulation indicated that there were resonances in the upper half. One structure that looked like it could be an issue was the open circuit stub on the input and output, which was close enough to cause a coupling effect with the first short circuit stub. However due to time constraints it was felt that the circuit should be tried. The circuit was manufactured on the milling machine and instead of wrapping tape around the edges of the board; spare pieces of circuit board were soldered perpendicularly along the edges. This made an 'H' section which it was hoped would give the filter more rigidity. Pinned vias were also inserted into the board. The response was similar to that predicted by the E-M, but with poor match in places and hence a high degree of pass band ripple. However with some tuning the insertion loss at 18GHz was minimised on one of the filters, Figure 3-53. It should be noted that such tuning was extremely difficult due to the size and the fragility of the assemblies.

Although the 6-12GHz filter was used on the input to the measurement system to reduce the harmonic levels out of the input driver amplifier and to remove a noise 'hump', at ~200 MHz which was believed to be giving stability issues, the filters were not used in the diplexer. It was found that the construction was not robust enough for practical use and that for best performance the filters would need to be placed in a metal housing. In view of the cost and timescales involved with this redesign it was decided to instead use the balanced low pass filter technique. It should be noted that the insertion loss of the bandpass filters compares very favourably with the quad hybrid based Diplexers, (e.g. at 10GHz in the 6-12GHz unit 0.5dB compared to 2.4dB in the 13GHz Diplexer) which potentially gives a significant increase in the dynamic range of the load pull arrangement. It is therefore felt that there is potential for this route in the future, but with due provision for housing the substrate and minimising the ground inductance.

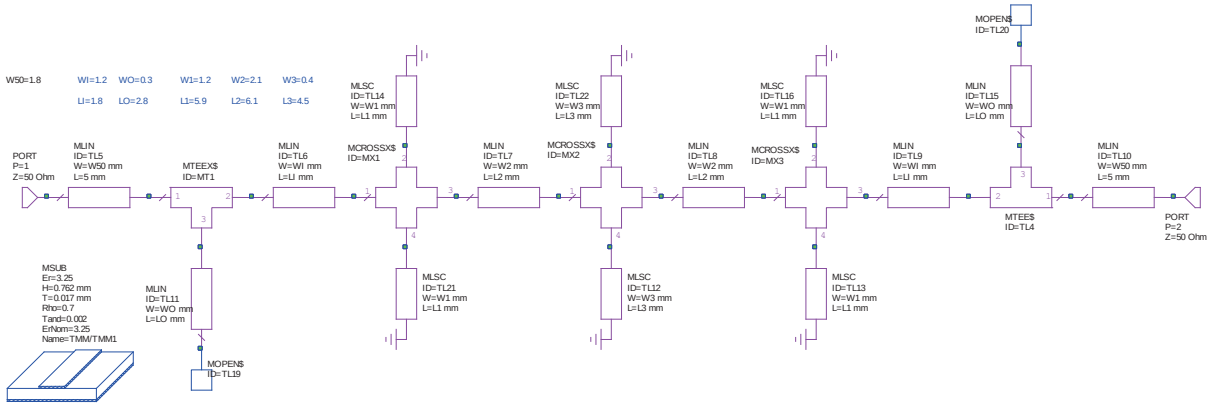


Figure 3-45, Schematic for 6-12 GHz Bandpass Filter

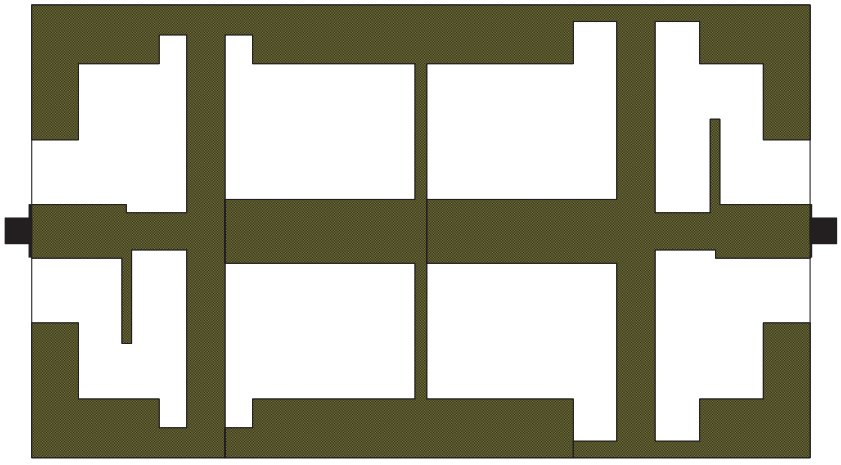


Figure 3-46, Layout for 6-12 GHz Bandpass Filter

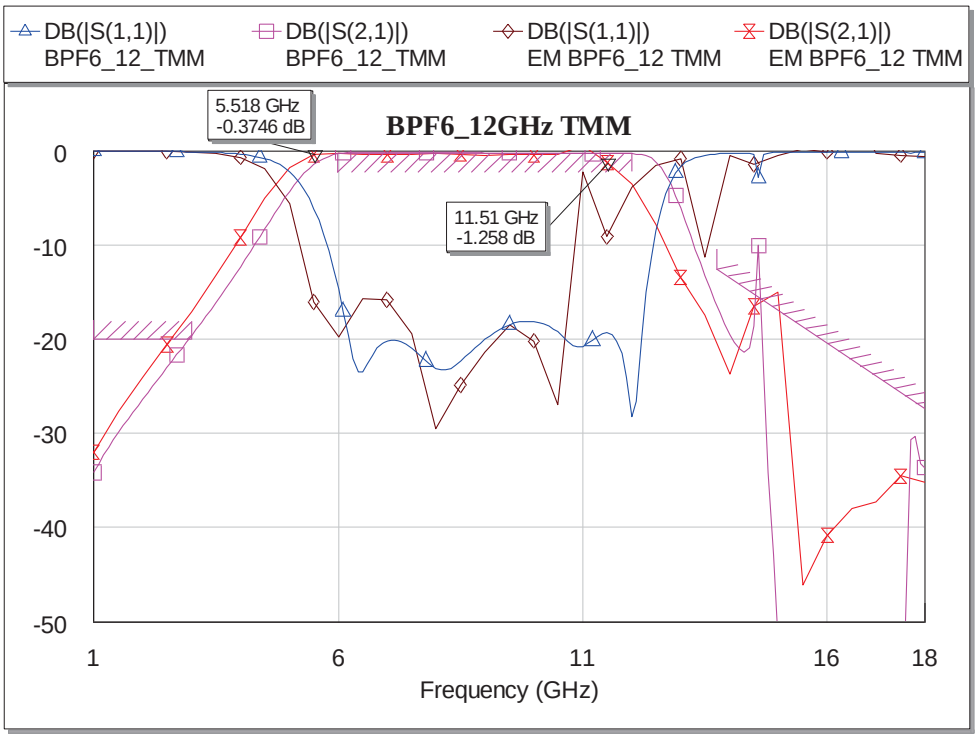


Figure 3-47, 6-12 GHz Bandpass Filter linear and E_M simulation.

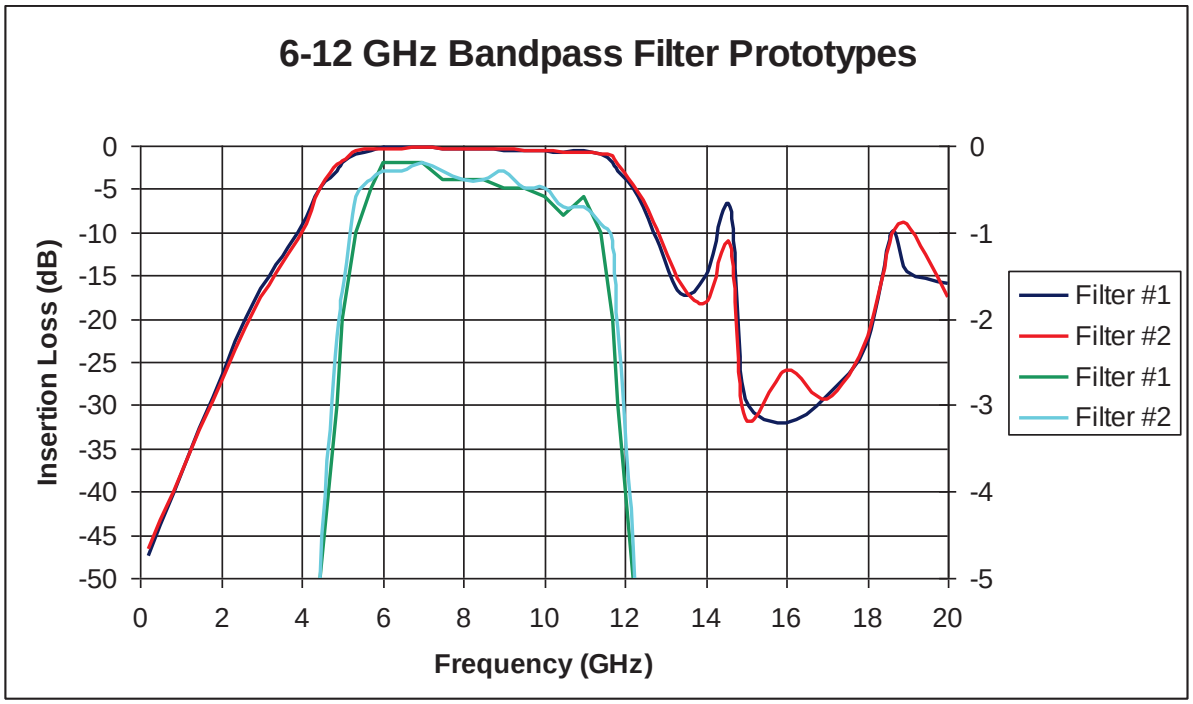


Figure 3-48, 6-12 GHz Bandpass Filter measured results, two units. Lower traces pass band loss (dB) – right axis.

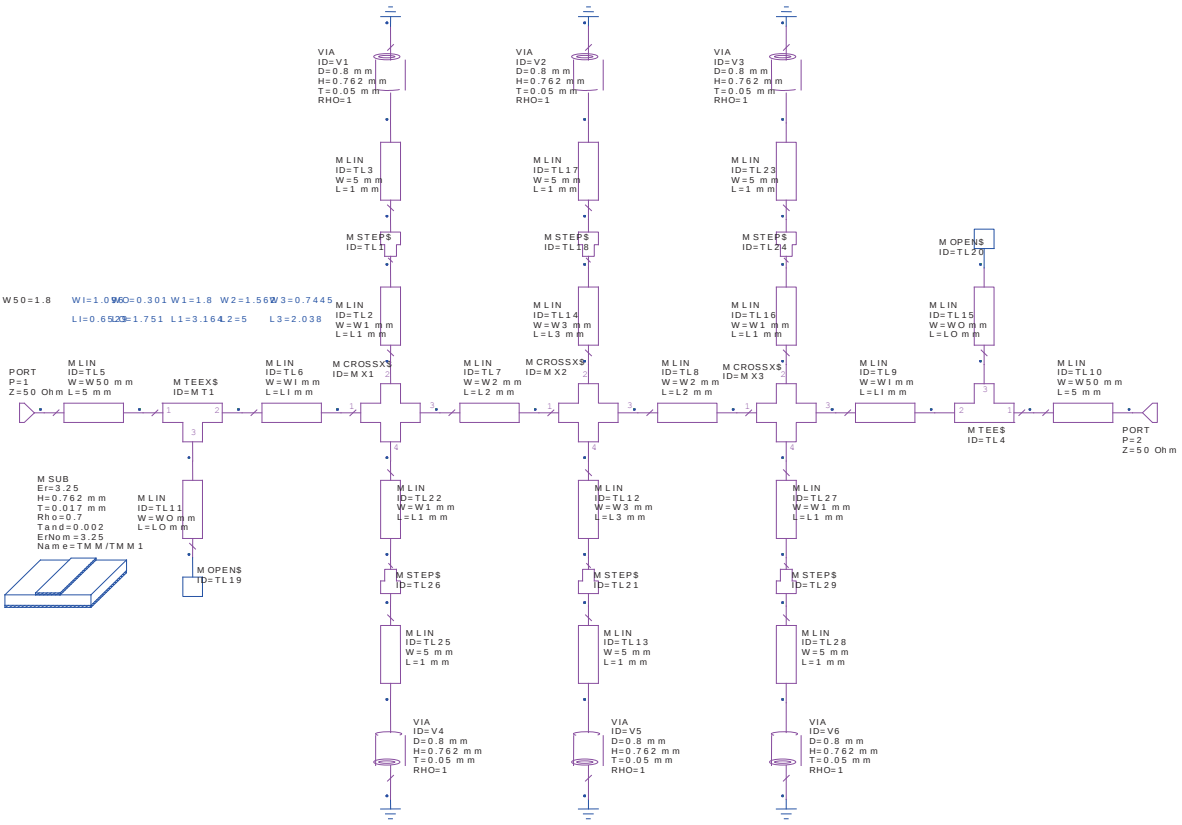


Figure 3-49, 9-18 GHz Bandpass Filter schematic including more detailed grounding of stubs.

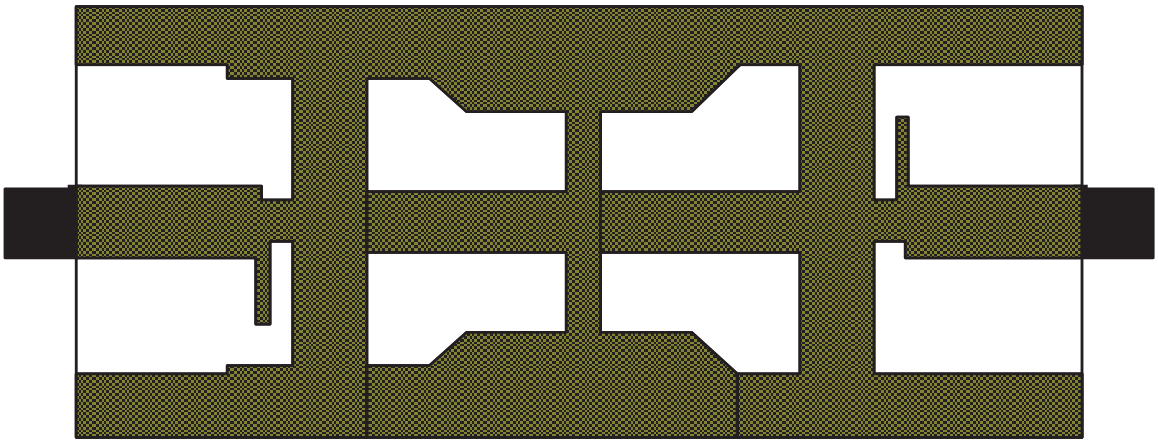


Figure 3-50, 9-18 GHz Bandpass Filter Layout incorporating via holes.

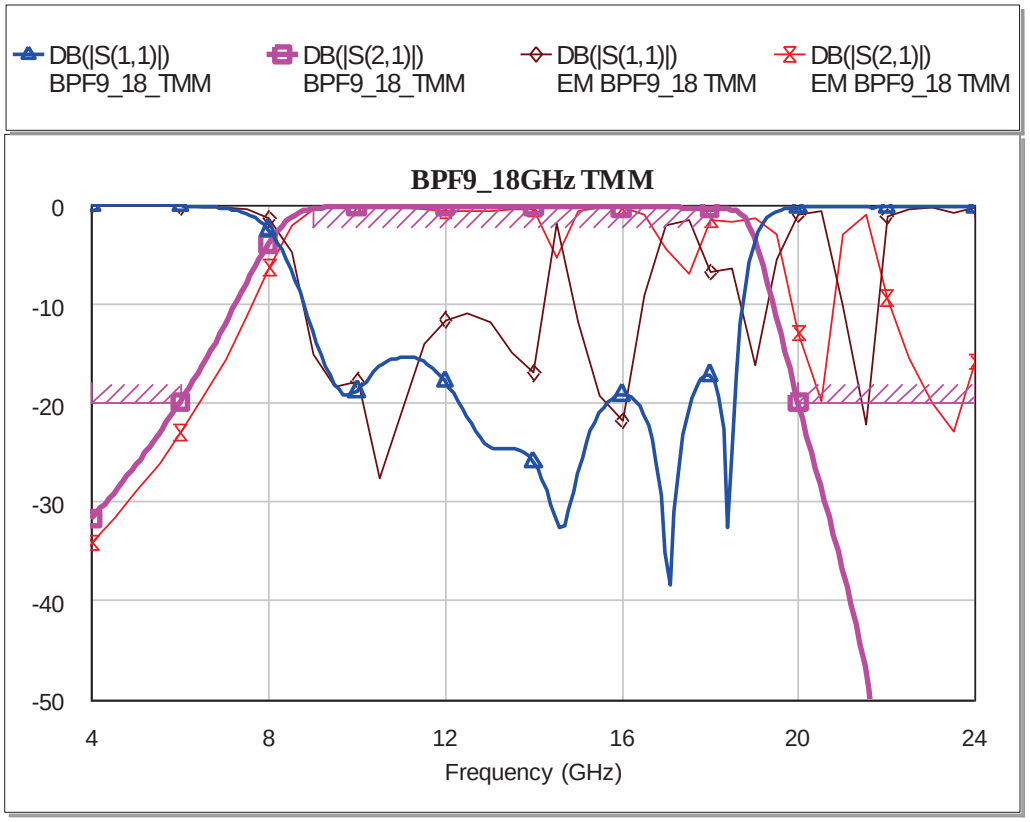


Figure 3-51, 9-18GHz Bandpass Filter Linear and E-M simulation results, E-M showing problems above 14GHz, possibly in part due to open circuit stubs on input and output.

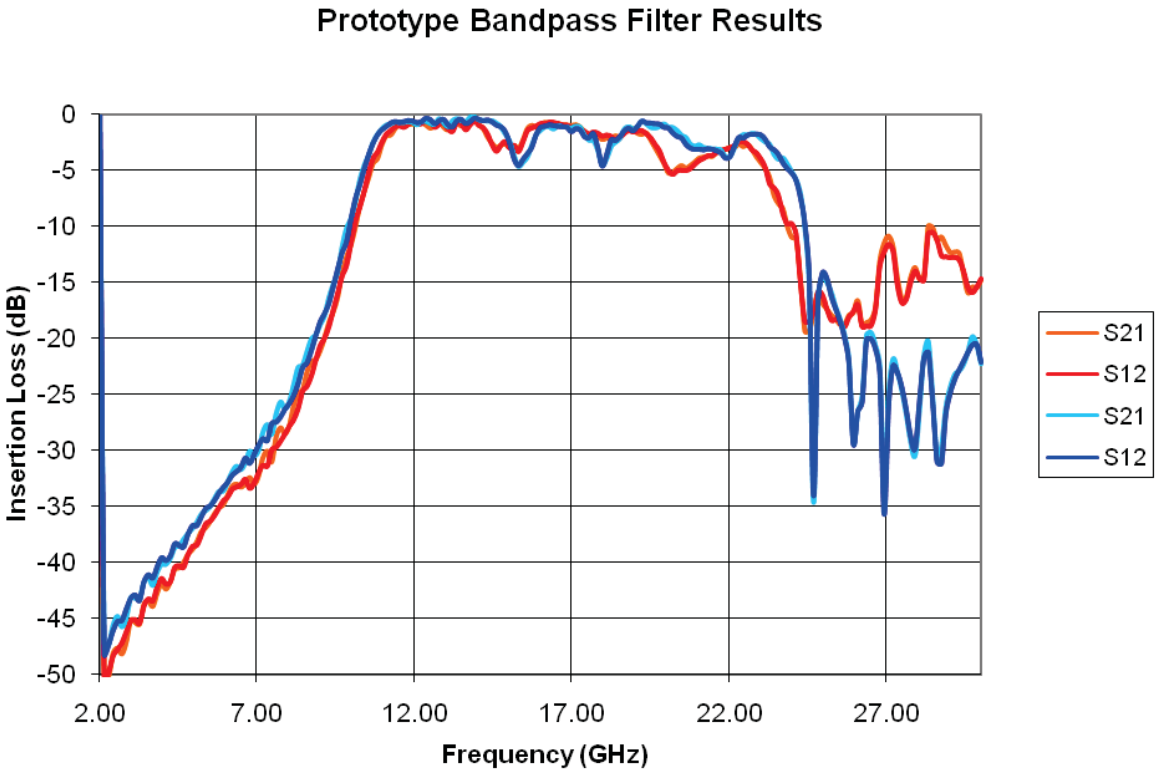


Figure 3-52, 9-18GHz Bandpass Filter, tuned (blue trace) and un-tuned (red trace).

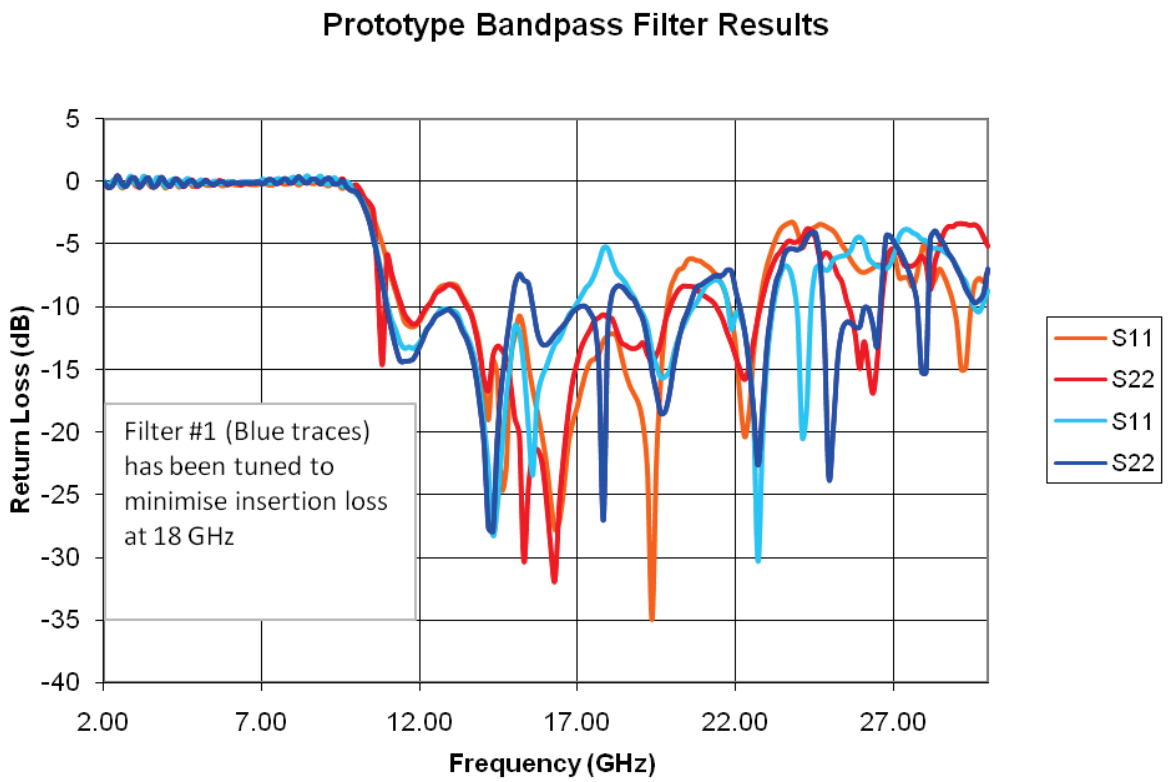


Figure 3-53, 9-18GHz Bandpass Filter measured results, one of the filters (blue traces) having been tuned for optimum performance at 18GHz.

3.7.3 Driver Amplifiers

The output power from the signal generators used in the measurement system was between 20 and 25dBm depending on frequency. The output power of the devices measured on the system as part of this research was up to 30dBm, (1W). The insertion loss of the cables, switches and bias Tee between the measurement plane and the signal generators was typically between 2 and 4dB again depending on frequency. In addition, the optimum fundamental output impedance of the devices was typically between 50 and 100Ω. Thus from Figure 3-38, it is clear that driver amplifiers would be required in order to adequately load pull the devices. Laboratory units were available with saturated output powers of 27dBm, which could be used in some circumstances, however ideally output powers in excess of 2W (33dBm) were needed. An initial search for bench cased wideband amplifier models, or multiple narrower band units, indicated that the cost would be prohibitive (>£16,000). As an alternative discrete packaged MMIC devices were found that

data indicated might meet the requirements, if the high power units could be connected in parallel. The datasheet performance of these devices is summarised in Table 3-7.

Device	Bandwidth	Gain	P1dB	Supply	Cost
AMMP-5612	6-20 GHz	13 dB	+19 dBm	5V, 110mA	£27
AMMP-6408	6-18 GHz	16 dB	+27 dBm	5V, 650mA	£26

Table 3-7, Outline performance characteristics and guide costs for Avago MMICs

It was decided to attempt to design a balanced pair of the 6408 devices driven by the 5612, in order to get a saturated power level of >1W within the available budget at the time. Each device was firstly laid out on its own PCB (Printed Circuit Board) and then a unit was designed using two 6408s in parallel. The PCBs were manufactured off site due to the need to have a plated through hole via process. A coplanar approach was used in order to minimise radiated fields from the transmission lines and coaxial connectors were soldered directly to the PCBs. In the case of the 6408 a small (40x40 mm) heat sink, (intended for mounting onto the top of BGA packages) was attached to the underside of the board. The assembly instructions and board layouts are given in Figure 3-55.

The combined MMICs provided in excess of 2W over 7-8GHz but rolled off quickly

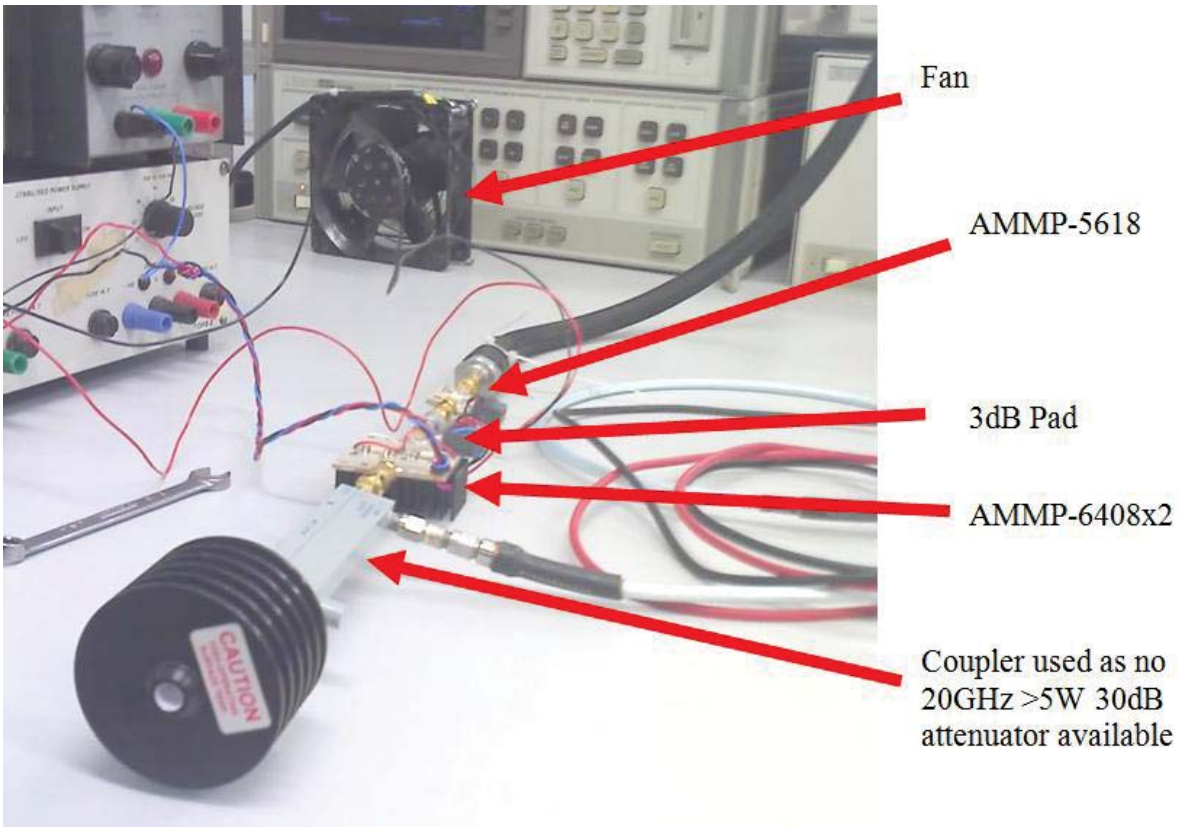
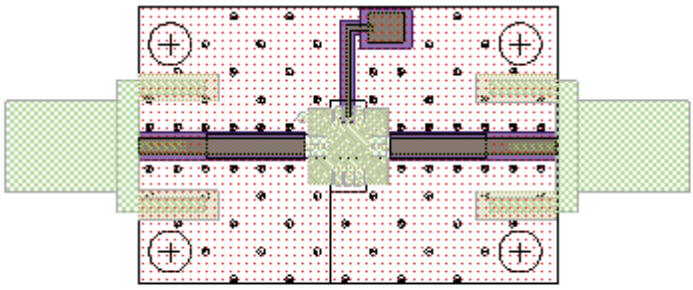


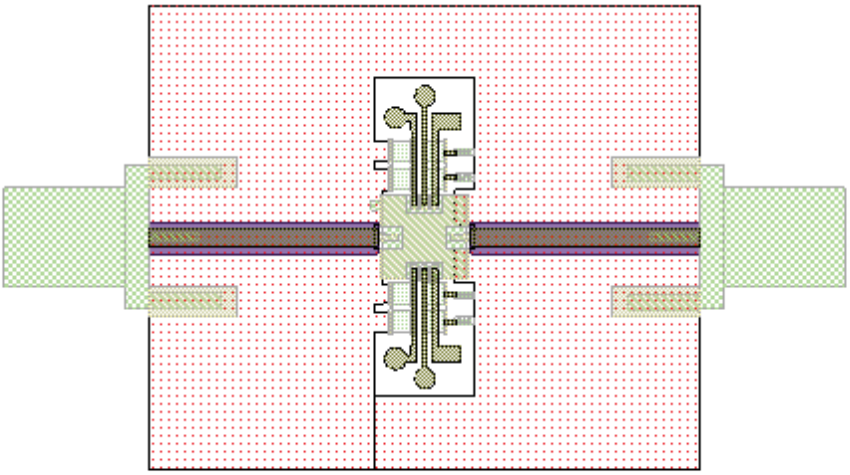
Figure 3-54, Testing prototype driver amplifiers.

after that. The performance plots are shown in Figure 3-56, and the devices being measured in Figure 3-54. It can be seen from the input return loss plot for the AMMP-6408 pair that this arrangement could give loading problems for the AMMP-5618 driving it; hence a 3dB attenuator was added between the stages. Also from the comparison plot between the single ended and balanced pair of AMMP-6408 at the centre of the band the single ended version actually gives slightly more power, hence showing the unbalanced Wilkinson splitter has probably not been ideal. An improvement to this design might be to layout two stages, but add connectors to both and use external wideband quadrature couplers for the combining. Nonetheless the amplifiers provided an improvement in power above that of the signal generators and their small size meant that they could be located close to the measurement device, thus minimising cable losses. A problem with this solution was however that both negative and positive power supplies are required to bias the 6408 devices, and these supplies needed to be sequenced so that the negative was switched on first and off last. Unfortunately during the dismantling of the test station by another student this procedure was not followed and the amplifier was permanently damaged.

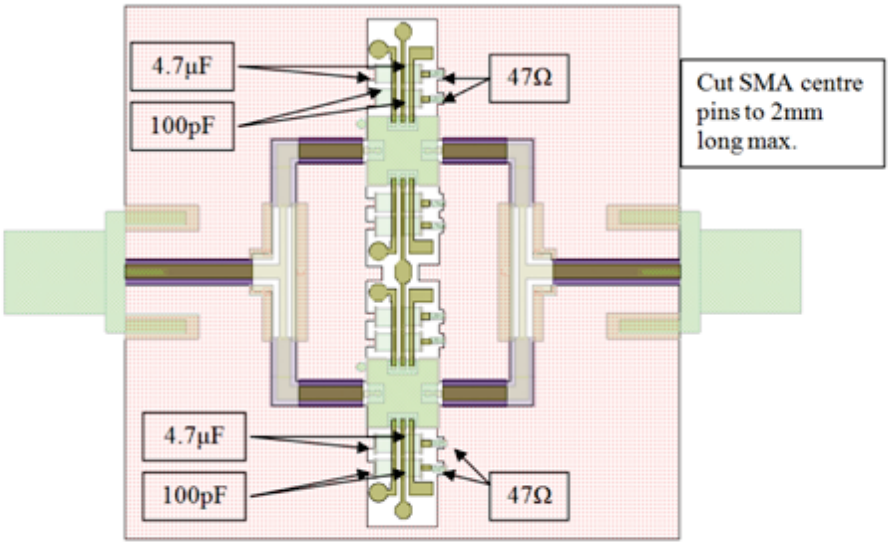


AMMP-5618 assembly.

Dot: (top left hand corner of device) in same corner as marked on circuit board green layer.
Static sensitive – take anti-static precautions. Ensure that care is taken not to short pads to ground.
See device data sheet, however although data sheet says hand soldering is not recommended we may do so if reflow oven is not available.
Ensure that solder paste is applied to the central area under the device but take care to limit the amount so that pins are not shorted. Fit SMA connectors.

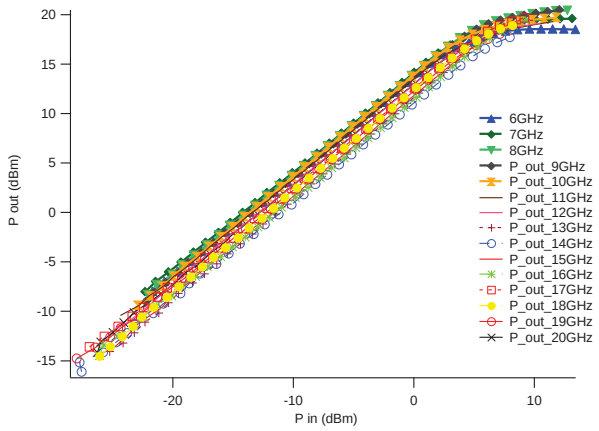


AMMP-6408 assembly.

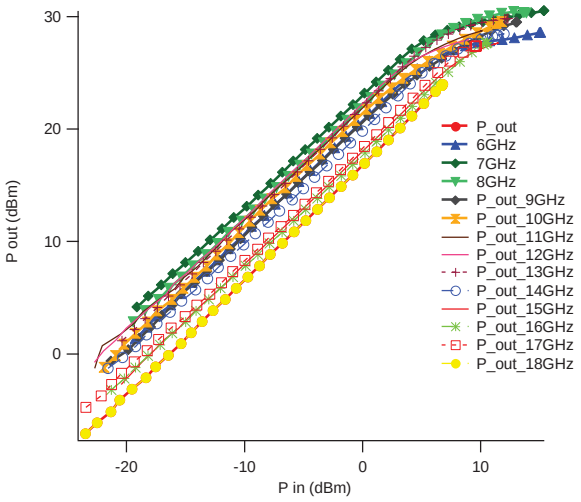


Balanced AMMP-6408 assembly.

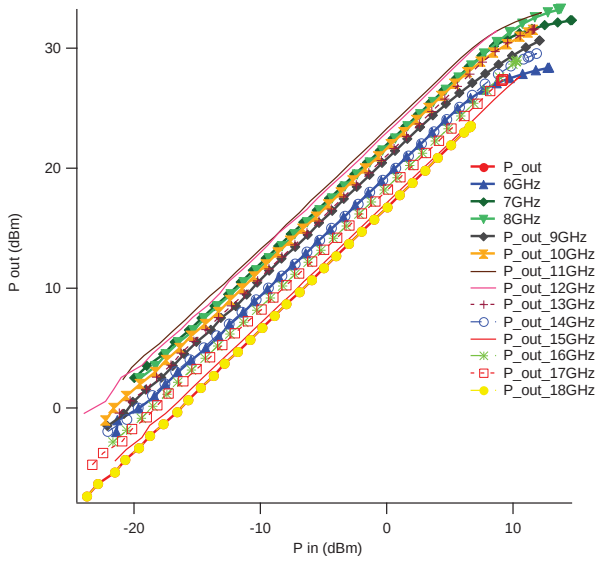
Figure 3-55, Assembly instructions and layout for MMIC driver amplifier prototypes.



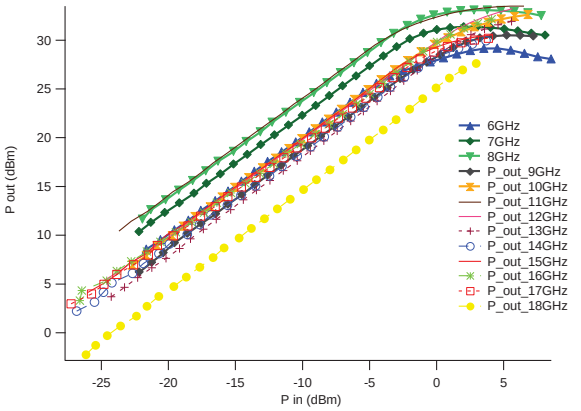
AMMP-5618 Compression Curves



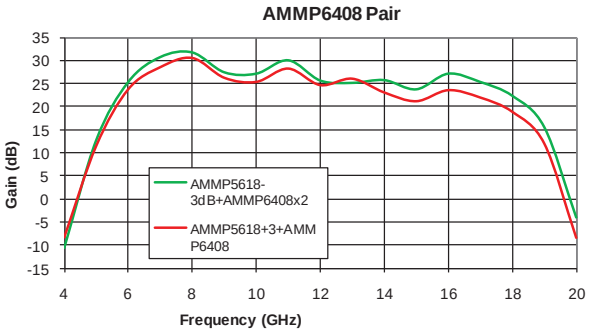
AMMP-6408 Single Ended



AMMP-6408 balanced



AMMP5618 +3dB Pad +AMMP6408 Balanced



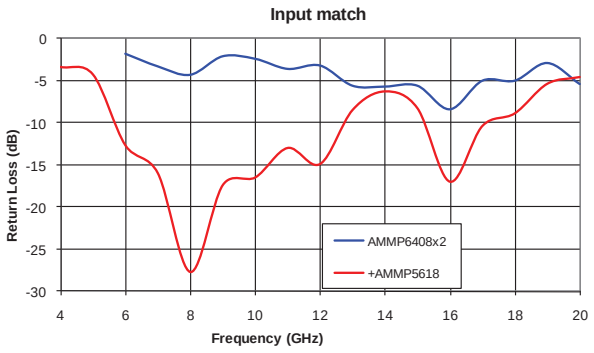


Figure 3-56, Performance curves for the prototype driver amplifier(s).

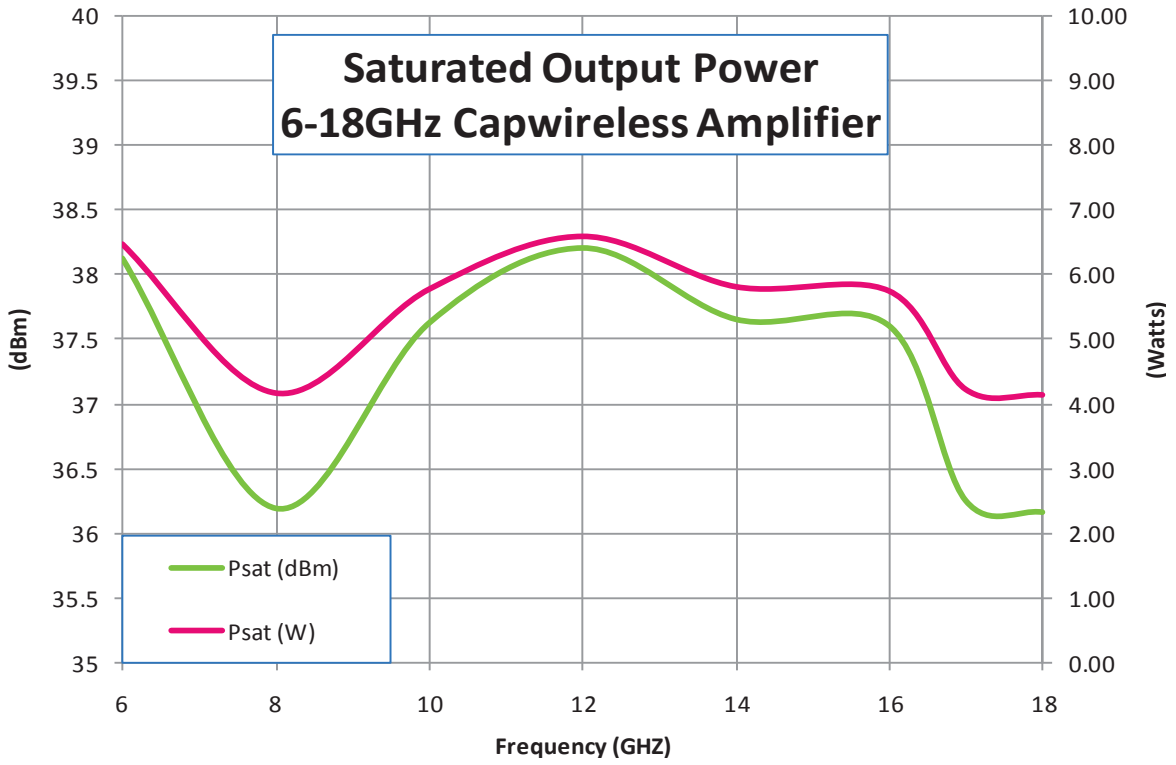


Figure 3-57, Performance of Capwireless KS5388 5W amplifier module.

At the end of the first year of the research a sufficient amount was left in the project budget to allow for the purchase of a 5W 6-18 GHz amplifier module (Capwireless KS5388). Initially there were some problems with this unit drawing excessive current at switch on at some frequencies and drive levels. After some weeks of operation the unit failed altogether. The unit was returned to the manufacturer where the bias supply was modified and the failure found to be caused by a solder void under one of the output devices. The returned unit has been operating successfully since then, without the current draw issues. The output power from this unit is shown in Figure 3-57. In the final extension to the project a budget was allocated for the purchase of two 2-20 GHz 10/5 W bench mount amplifiers, (Gigatronics GT-1000A). These greatly simplified operations and performance of the system, as well as increasing the operating envelope. A lesson learned from the use of these modules was that units incorporated in the test system had to be extremely robust. Systems were regularly re-assembled to meet other research requirements and the possibility of switch on sequences not being followed or amplifiers being left open circuited were high.

3.7.4 Signal Generation

The problem with the limited frequency range over which the measurement system could be operated with the load pull signals coherent with the input signal, as shown in Table 3-3, causes a severe limitation on the speed and repeatability of the active load pull. As part of the final 6 month extension to the research project, a method of improving the bandwidth of phase coherent operation was investigated. A block diagram of the connections of the various source and load pull sources is shown in Figure 3-58. As outlined earlier the problem with this arrangement is that the way that the multipliers are arranged the 3 harmonic load pull system only works over a fundamental frequency range of 6.67 to 9.97GHz. An alternative solution was proposed whereby a separate source was used to provide the fundamental signals to the multipliers as shown in Figure 3-59. Note the external source has 4 independent, phase coherent frequency outputs. The frequencies sent to each of the multiplier chains would need to be controlled from a Look-Up table based on the information in Table 3-3 to give the desired output frequency. The gain control would be implemented at the output of each multiplier chain as before, however the phase control could be carried out either at the fundamental frequency (adjusting appropriately for the multiplier ratio) or via the I & Q modulator at the output. This system is based on still using the output stages of the existing signal generators. Although this may seem wasteful, the development of a multiplier chain to 40GHz with the associated filtering and amplification is complicated and time consuming and outside of the scope and budget of this research project.

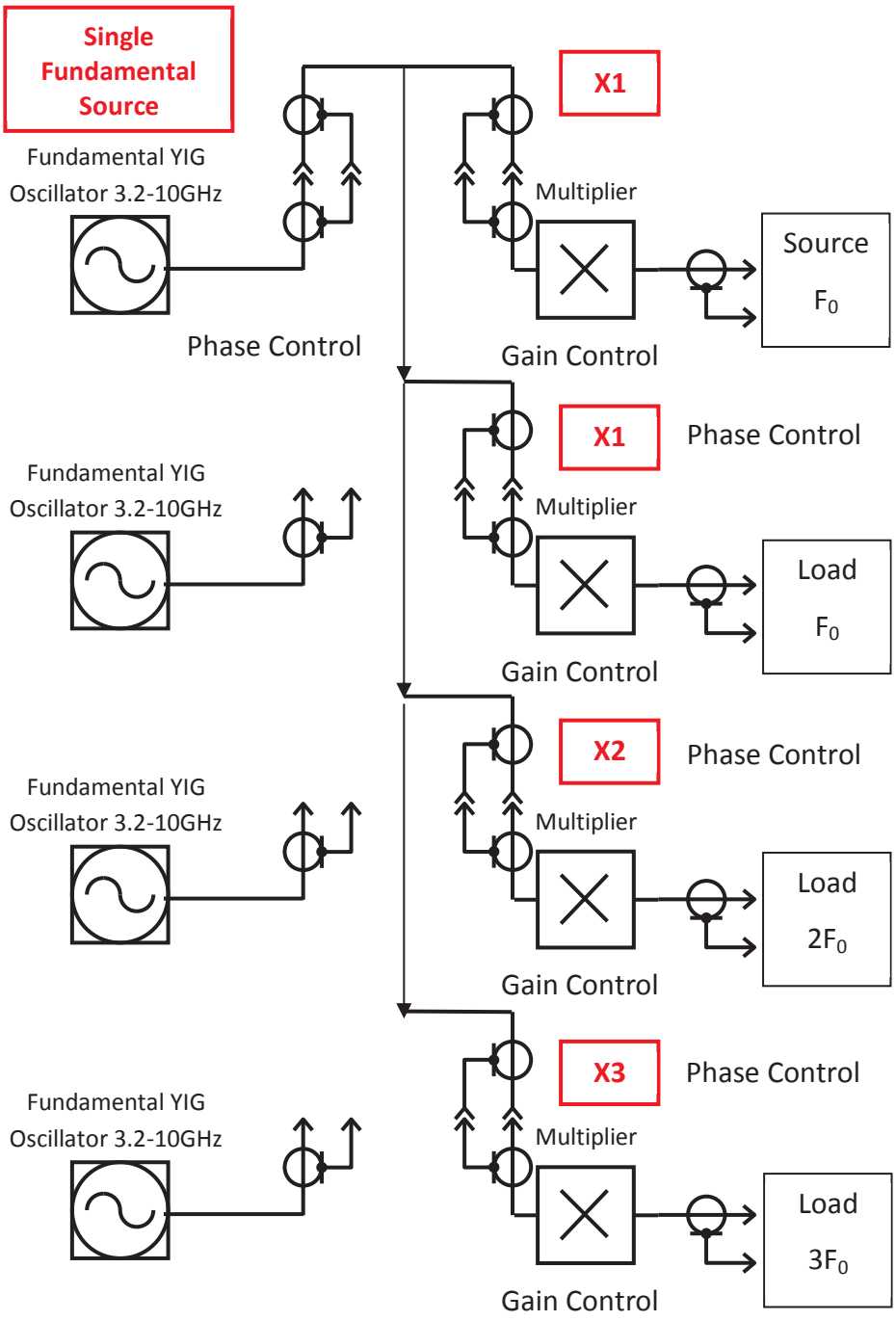


Figure 3-58, Block diagram of Implemented Phase Coherent Narrow Band 3 Harmonic Active Load-Pull.

The outline requirement for the phase coherent source can be summarised as follows:

- 3 frequency independent but phase coherent outputs.
- Frequency range 3-10 GHz minimum.
- Phase control to 0.5° increments (if phase control implemented at fundamental).
- Output power level +15 dBm ±2 dB.
- Harmonics -40 dBc.

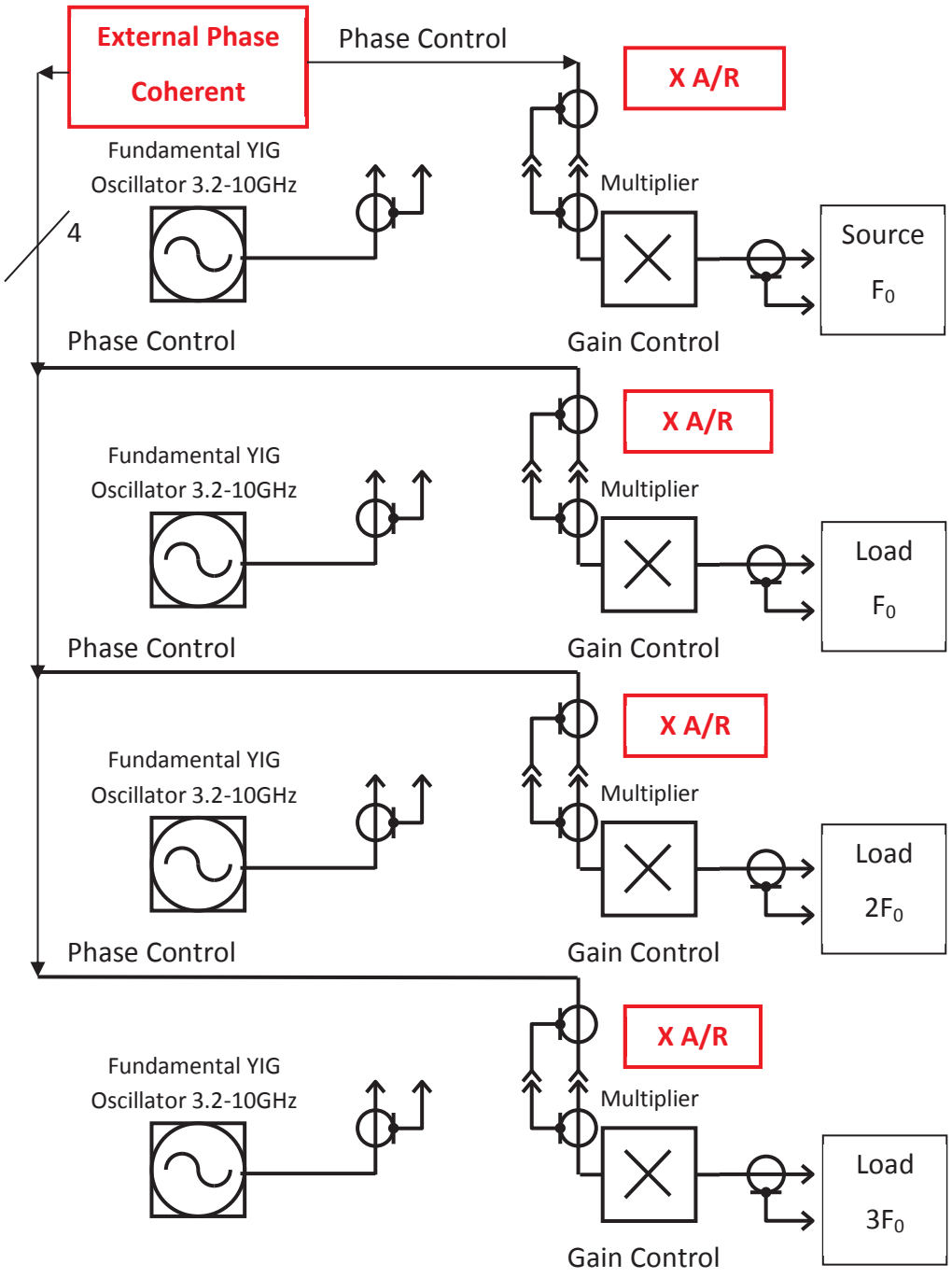


Figure 3-59, Block diagram of the Proposed Phase Coherent Wideband 3 Harmonic Active Load-Pull.

Only 3 channels are actually required for the measurement system as the fundamental source and load can be provided using the option HCC in the conventional way (the source and load signal generators always have the same output frequency and hence primary YIG oscillator frequency) from the same signal, but split before going to the multiplier chain if the phase control is executed by the I & Q modulator. Note that if the frequency of the phase coherent source is high enough (e.g. 18 GHz) then it may not be necessary to use an ESG for the fundamental source, although in this case amplitude control would be required of the alternative source. It was thought that such a signal generator had

been found. Holzworth Instrumentation Inc. (www.holzworth.com) have an 8MHz to 6GHz 4 channel source [19] and had an 18GHz unit in development, due for delivery in September 2010 (provisional date due to product not having been released for manufacture). Although an order for this unit was placed it had to be cancelled at the end of 2010, when the company could not provide a firm date for completion of the development. A 6GHz unit was however provided on another project at the university, and worked successfully.

A suggestion has been made for the utilization of the 6GHz unit as the necessary source by the addition of a doubler on the output of the synthesiser and adjusting the frequency accordingly. Figure 3-60 shows a block diagram of the doubler, including possible filters and amplifiers. As can be seen it is necessary not only to include a tracking filter but also additional amplifiers to increase the signal level to that required by the P2 port of the option HCC on the signal generators. A suitable bandpass filter would be the MLFP-22018PD, 2-18GHz Yttrium Garnet (YIG) digitally controllable module from Micro Lambda [20]. The suggested doubler [21] has a conversion loss of 12dB, and a fundamental and 3rd harmonic rejection of 15 and 20dB respectively. The suggested amplifier [22], has a gain and power slope of ~3dB, hence the attenuator should be designed with the opposite slope, such as the one shown in Figure 3-61. The most costly item in this approach is the YIG filter at about £2,600 each including driver. If the driver were produced internally the cost of the filters is about £1,500 each. Thus a doubler is approximately £2,000 per channel for the system.

Freq. (GHz)	1.5 – 5	3-10	3-10	3-10	3-10	3-10	3-10
Power (dBm)	15	3	-1	-4	7	18	15

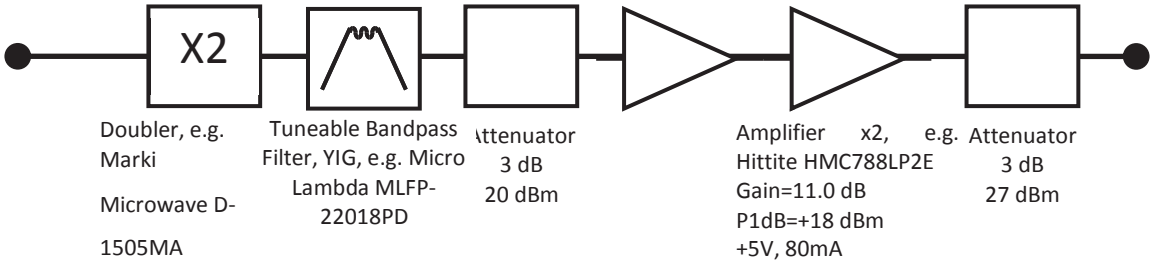


Figure 3-60, Block diagram of proposed frequency doubler for phase coherent source.

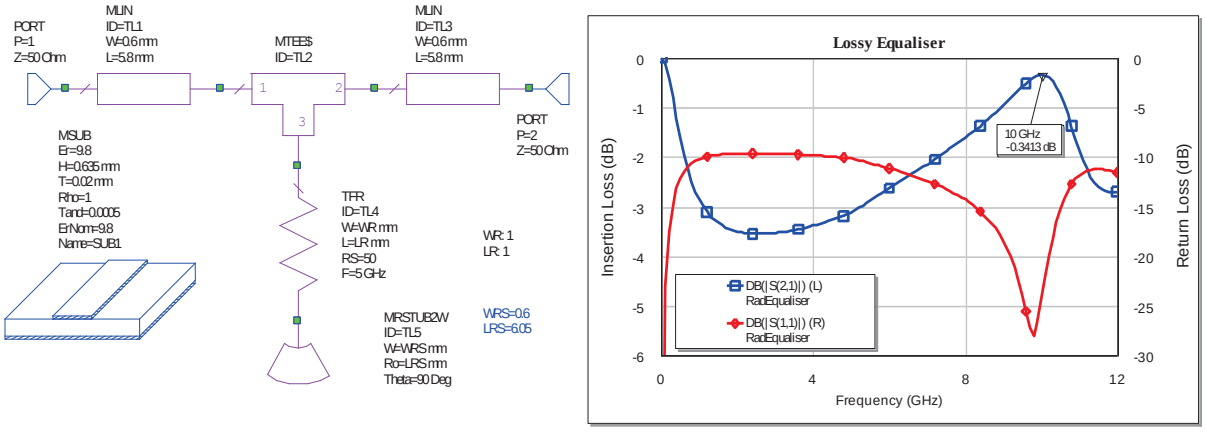


Figure 3-61, Possible compensation network for amplifier power and gain roll-off.

3.8 Summary

A measurement system has been constructed which enables the characterisation of high frequency passive devices and the full non-linear behaviour of active devices with output power levels into the region of several watts and frequencies currently up to 40GHz. The operation of the system has been shown with control of up to 3 harmonics up to 40GHz and the linear measurement results show good agreement with other standard measurement systems (VNA). The calibration of the system shows that a useful dynamic range in excess of 40 dB can be achieved up to 40GHz, which is adequate for the characterisation of devices and active load pull operation. The versatility of the system to be reconfigured so as to be able to encompass unusual measurement requirements has been demonstrated.

A novel method of reliably triggering the digital oscilloscope used for the acquisition of the waveforms was developed and the benefit of using phase coherent sources shown. A system capable of phase coherent active load pull on up to 3 harmonics was produced and tested over a limited fundamental frequency range (6.7-9.9GHz), and a proposal for wider bandwidth version capable of operation over the full current system bandwidth (1-40GHz) has been made.

The requirement to be able to control the harmonic impedances presented to devices in the design wideband high efficiency amplifiers has been discussed in chapters 1 and 2. Furthermore it is necessary to be able to create high reflection coefficients which as frequency increases becomes more difficult (due to system losses). Thus the only practical solution is to use active load pull. Even so such systems themselves are limited by the power levels that can be generated by the amplifiers in the system. The input amplifier needs to

have sufficient power so as to drive the DUT into compression/saturation. Using passive input tuners can significantly reduce the power required. On the output of the device, the amplifiers in such wideband systems are limited by the maximum power transistors, by other components such as the wafer probes and the ability of the DUT to dissipate heat. The current system only operates CW, when a pulsed version is developed this will increase the power that can be handled, however looking at practical MMIC implementations (chapter 2) even with the development of GaN transistors it is unlikely that there will be a need to characterise devices of more than 10W. The real driver will be to increase the operating frequency of the system. There are already many applications in the 20-40GHz region (such as links for communications backhaul networks), and the current system would only be able to conduct fundamental measurements in these applications. As the gate length of commercial GaN devices decreases and f_T of ~75GHz [23] become more common there will be a need for nonlinear models for new devices and hence characterisation systems.

3.9 Recommendations

Not addressed at this stage has been the application of bias voltages within MMICs themselves. Unlike discrete transistors MMICs typically have bias networks incorporated into their circuits. The bias is introduced at separate pads to the RF input and output, which often incorporate AC coupling. Thus the bias tees used in the discrete device measurements cannot be used. The standard method is to use specific bias probes to provide the bias voltages and currents to the devices, however the devices that are under consideration for future projects present two distinct problems:

1. The currents involved with a 5W power device are of the order of 1 amp which is greater than the standard bias probes can handle.
2. MMICs usually require decoupling of the order of 100pF close to the bias pad for stability.

Two possible options to resolve this are, (a) a custom design bias probe (including built in de-coupling) and (b) the mounting of the MMICs on a carrier which would include decoupling and bias points (feedthroughs screwed or soldered into the walls). An advantage of this latter solution is that the mechanical construction can be made similar to that of the intended application. Temperature has a significant effect upon device performance and so replicating the likely thermal scenario is an important factor in anticipating actual device

performance. This would highlight potential problems at an early stage, something that pulsed operation may mask.

A deficiency in the current active load pull measurement systems (generally, not specifically to that described in this research) is the lack of a reference standard to show consistency in large signal measurements. Some form of gold standard device which could be reliably re-measured to clearly validate system operation and calibration. The ideal device would produce harmonics with a consistent magnitude and phase relationship to the fundamental. A problem is that probed devices tend to have a relatively short life time, they are both delicate in their nature and there is a limit to the number of times a pad can be wafer probed due to the inherently destructive 'landing' operation. It has been suggested that some form of multiplier diode could be used. If so it would need to be able to be removed from the circuit and mounted in a replacement as the connecting pads wear out, or come from a highly repeatable process. More investigation on the best solution for this problem is required.

3.10 Bibliography

[1] D. Williams and P. Tasker, "An automated active source and load pull measurement system," in *6th High Frequency Postgraduate Student Colloquium*, 2001.

[2] A. AlMuhaisen, P. Wirght, J. Lees, P. Tasker, S. Cripps and J. Benedikt, "Novel wide band high-efficiency active harmonic injection power amplifier concept," in *International Microwave Symposium*, Anahiem, 2010.

[3] Z. Aboush, J. Lees, J. Benedikt and P. Tasker, "Active harmonic load-pull for characterising highly mismatched high power transistors," in *International Microwave Symposium*, 2005.

[4] P. Wright, A. Sheikh, C. Roff, T. P.J. and J. Benedikt, "Highly efficient operation modes in GaN power transistors delivering upwards of 81% efficiency and 12W output poer," in *International Microwave Symposium Digest*, 2008.

[5] D. FitzPatrick, J. Lees, A. Sheikh, J. Benedikt and P. J. Tasker, "Systematic investigation of teh impact of harmonic termination in the efficiency performance of above octave bandwidth microwave amplifiers," in *European Micorwave Conference*, 2009.

[6] S. Cripps, RF Power Amplifiers for Wireless Communictions, 2nd ed., Artech House, 2006.

[7] R. E. Leoni, S. A. Harris and D. G. Ries, "Active Simultaneous harmonic source and load pull assisted by local polyharmonic distortion models," in *International Microwave Symposium*, Anahiem, 2010.

[8] S. Di Falco, A. Raffo, F. Scappaviva, D. Resca, M. Pagani and G. Vannini, "High-efficiency broadband power amplifier design technique based on a measured-load-line

approach," in *International Microwave Symposium*, Anahiem, 2010.

[9] J. Benedikt, P. Tasker and et. al., "High Power Time Domain Measurement System with Active Harmonic Load-Pull for High Efficiency Base Station Amplifier Design," in *IEEE Trans MTT pp.2617-2624*, 2000.

[10] Agilent, "Agilnet E8267D PSG Vector Signal Generator with Option HCC Data Sheet," [Online]. Available: http://cp.literature.agilent.com/litweb/pdf/5989-0697EN.pdf. [Accessed 2008].

[11] T. Williams, D. FitzPatrick, M. Haynes, C. Whittle, J. Benedikt and P. Tasker, "A High Frequency Time Domain Measurement System and its application to Robust and Scientific Ultra-Wide Band MMIC design for Military Applications," in *EMRS-DTC Conference*, Edinburgh, 2008.

[12] Z. Aboush, C. Jones, G. Knight, A. Sheikh, H. Lee, J. Lees, J. Benedikt and P. Tasker, "High Power Active Load-Pull System for Characterisation of High Power 100W Transistors," in *IEEE MTT-S International Microwave Symposium*, 2005.

[13] F. van Raay, R. Quay and et. al., "A coplanar X band AlGaN/GaN power amplifier MMIC on s.i. SiC Substrate," *Micorwave and Wirelesss Component Letters,* vol. 15, no. 7, July 2005.

[14] D. FitzPatrick, R. Saini, J. Lees, J. Benedikt, S. Cripps and P. Tasker, "A Waveform Engineering Approach to the Design of Improved Efficiency Wideband MMIC Amplifiers," in *WAMICON 2011*, Clearwater, 2011.

[15] TriQuint Semiconductor, "TGF2023-01 6W Discrete Power GaN on SiC HEMT," 2011.

[16] Belden, "Technical Data Sheets," [Online]. Available: http://www.belden.com/techdatas/metric/1671A.pdf. [Accessed 26th April 2011].

[17] G. Matthaei, L. Young and E. Jones, Microwave Filters, Impedance-Matching Networks, and Coupling Structures, Dedham, MA: Artech House Books, 1980.

[18] R. W. Rhea, HF Filter Design and Computer Simulation, Norcross, GA: Noble, 1994.

[19] Holzworth Instrumentation, "http://www.holzworth.com/Spec_sheets/HS6004_Web_Datasheet.pdf," 7 April 2011. [Online]. Available: http://www.holzworth.com/Spec_sheets/HS6004_Web_Datasheet.pdf. [Accessed 7 April 2011].

[20] Micro Lambda Wireless, "Products: YiG Filters: Bandpass: 2 Stage," [Online]. Available: http://www.microlambdawireless.com/drivers_control/filters/pdfilter.htm. [Accessed 7 April 2011].

[21] Marki Microwave, "Datasheets: Marki Microwave," 8 April 2011. [Online]. Available: http://www.markimicrowave.com/menus/datasheets/D-1505.pdf. [Accessed 8 April 2011].

[22] Hittite, "Documents: Datasheets: Hittite Corp.," [Online]. Available: http://www.hittite.com/content/documents/data_sheet/hmc788lp2.pdf. [Accessed 8 April 2011].

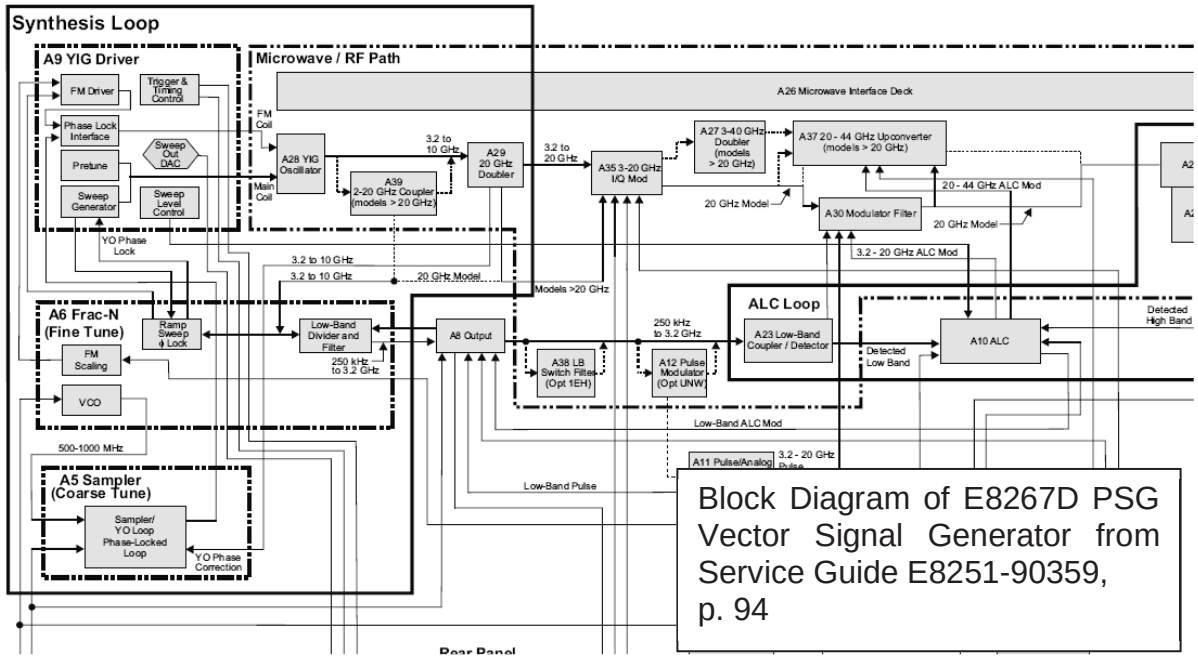
[23] K. Kobayashi, Y. Chen, I. Smorchkova, R. Tsai, M. Wojtowicz and A. Oki, "A 2 Watt, sub-dB noise figure GaN MMIC LNA-PA amplifier with mulit-octave bandwidth from 0.2-8 GHz," in *IEEE MTT-S International Microwave Symposium*, 2007.

3.11 Appendices

Agilent Approach to Phase Locking Multiple Sources

Option HCC takes the fundamental 3.2-10 GHz YIG Oscillator signal (see figure A-1) and the divided down 250 kHz to 3.2 GHz signal and routes them via the rear panel. These paths can be broken, split and fed back into the master and a slave source to lock them together.

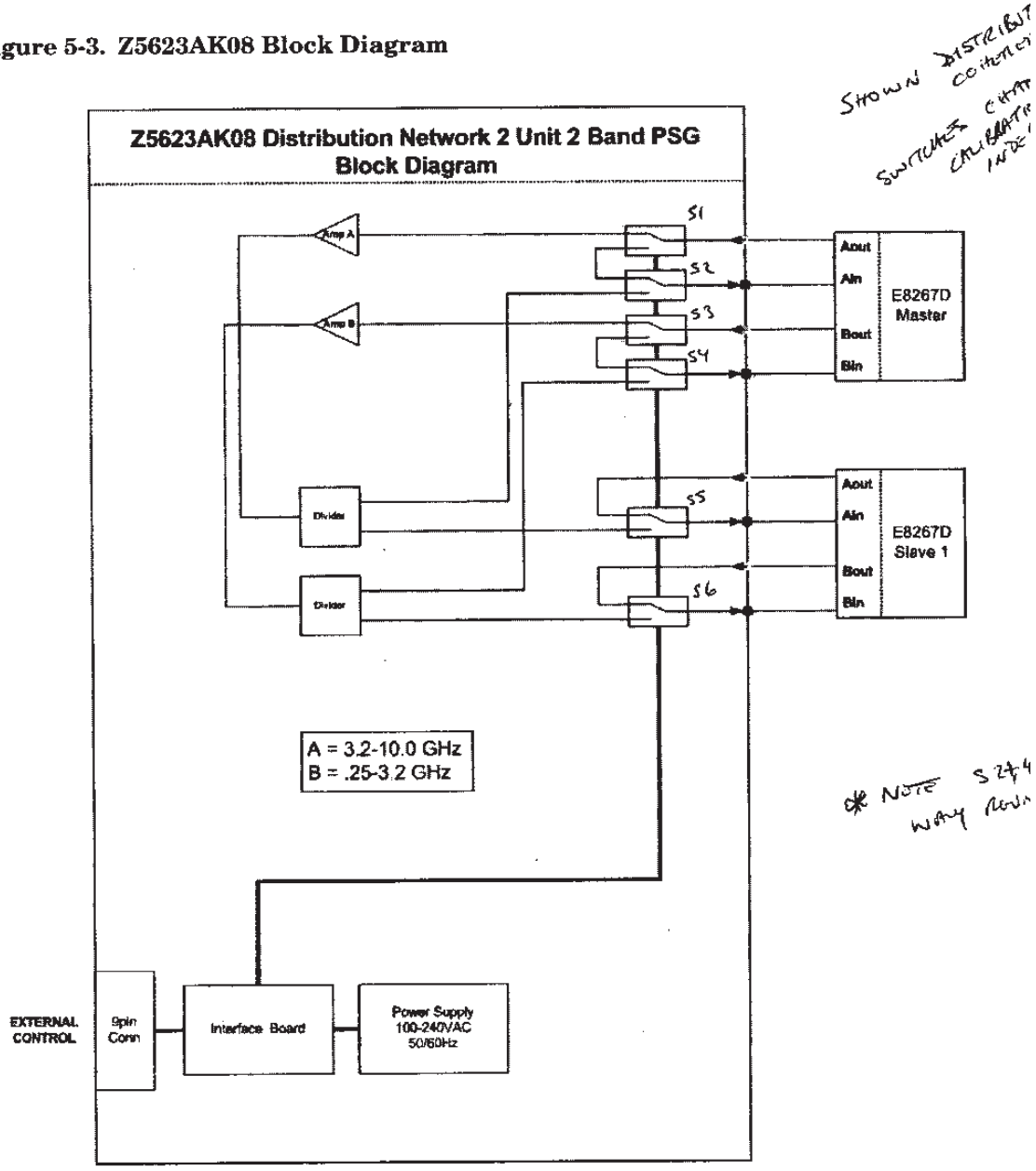
The distribution box is simply a method for splitting the LO signal and bringing the level back up. On an application note (5989-6850EN) on locking PSG 8267D together Agilent state, "The Z5623AK07 Tri band lock box provides necessary circuitry to ensure phase coherence to within 3 degrees".



A 1: Synthesiser Loop from Agilent ESG signal generator operation manual

.

Figure 5-3. Z5623AK08 Block Diagram



A 2: Distribution Box Block Diagram

4 Active Harmonic Load Pull Device Measurements

"Ex umbris et imaginibus in Veritatem.' - 'Out of the shadows and images and into the Truth'"

Cardinal John Henry Newman, circa 1890

4.1 Introduction

A dilemma faced by device manufacturers and designers is how to introduce new technology to the market as quickly as possible. From the manufacturers point of view they wish to start getting a return on their capital invested, from the designers so that they can take advantage of the advances the new technologies offer. However, balancing this desire the semiconductor supplier needs to supply the designer with information on how the device(s) perform, and to do this the manufacturing process must be stable so that the characteristics are repeatable, which leads to an inevitable time lag whilst the processes are being optimised. In order to justify the massive investment in a semiconductor manufacturing facility a number of different processes will be run, some more suited to low noise devices, others switching, high power or high frequency. For each of these processes, besides the active transistors, there are many passive structures such as capacitors and resistors as well as the distributed components for which models are required. The devices themselves will have a number of different geometries based on:

- Gate width.
- Number of gate fingers.
- Finger spacing.

These parameters are adjusted to provide a variety of gain/power/frequency options. The number of variants of the basic device cell makes it expedient to characterise only a few of the options in detail and to then use 'scaling' to infer the performance of those devices not measured. For example, for the semiconductor process used to manufacture the devices characterised in this project, a single set of DC-IV curves is produced, Figure 4-1. Note that the vertical axis, I_{ds}, is in mA/mm. This refers to the gate periphery (gate width x number of fingers) and therefore needs to be scaled for a particular device. In general terms the cost of pulsed DC-IV measuring equipment is relatively low (compared to Vector Network Analysers

for example) and therefore more attractive to device manufacturers. Another reason for this approach can be the difficulty in measuring some of the larger devices, in terms the required power levels and bias currents, on wafer. Finally, most wafer measurements are conducted in a pulsed mode to reduce the heat generated (it being very difficult to remove heat from complete wafers). When scaling it is normal to split the transistor between its active area and passive feed structures. The signal distribution networks (to the gate fingers or from the drain contacts) are more easily modelled using linear or E-M simulations, or even direct measurement (two identical networks connected back to back) – similar to the de-embedding structure as described in section 4.3. Figure 4-2 shows a 1 gate finger device and Figure 4-3 a 10 finger version of the same gate width. Note that one of the issues with scaling is the ratio of via connections to gate fingers changes. It is also a fact that transistors may be operated at different bias conditions, for example mobile applications generally prefer to use low drain voltages primarily due to battery considerations, (a problem for GaN devices with their relatively high knee voltages). Base station and instrument markets can use higher drain voltages and here the consideration may be more related to other circuit parameters such as minimising current to allow smaller (with lower current carrying capacity) structures to be used – for increased frequency of operation or more efficient power combining. Further as has been discussed earlier there are a number of operating modes in which transistors can be biased; from the well-established class A to more exotic high efficiency modes. What has also been shown is that these more efficient modes produce high order harmonics which also require modelling, thus the frequency range now needed for the model creation is two or three times greater than the fundamental frequency only model.

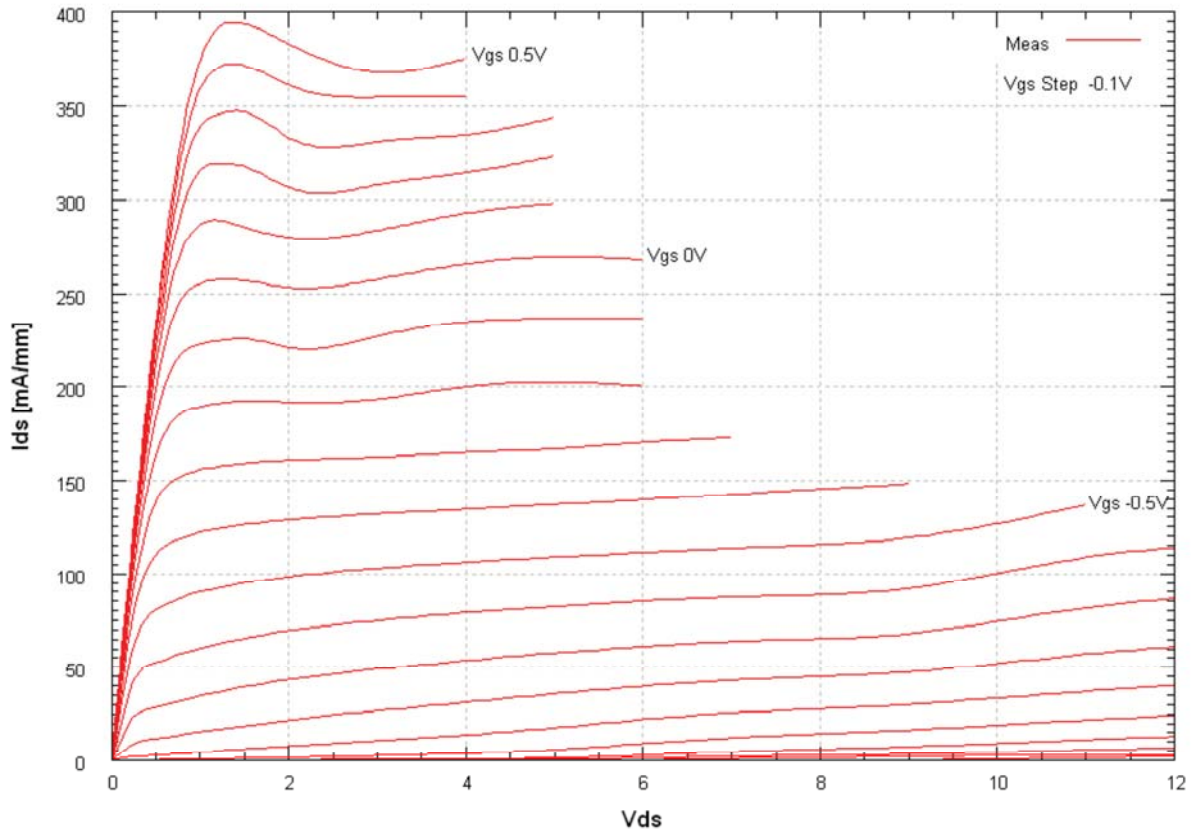


Figure 4-1, General DC-IV curves for the GaAs pHEMT process used.

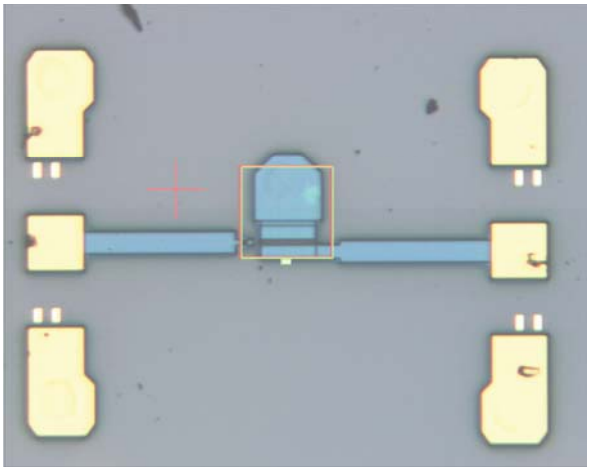


Figure 4-2, Single cell, 1x75µm device.

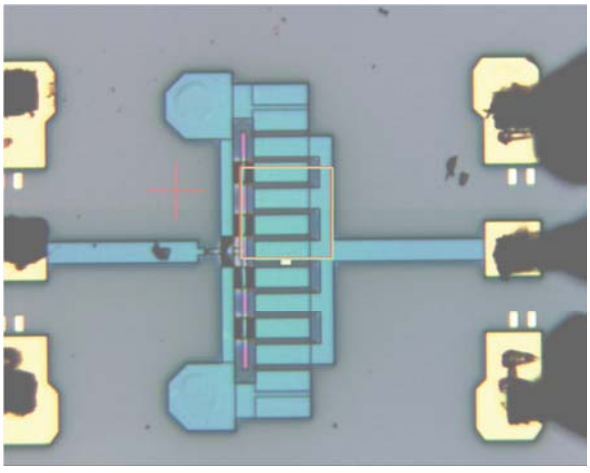


Figure 4-3, Multiple cell with feed structure, 10x75µm device.

Finally, the most convenient and cost effective way of testing devices at the foundry is before the wafer is diced; this is more suited to automatic measurement and if there is a problem with the wafer it can be spotted at this stage (this can not only save the cost of splitting the wafer up but also alert the processors to any issues at as early a point as possible). On wafer measurements pose a severe limitation for power devices; there is very little heat sinking on wafer and hence measurements must be conducted pulsed with short pulse widths, in the order of µs, and low duty cycles, generally <10%. This again introduces

another error factor; the vast majority of the data available to the manufacturers modelling group will be on-wafer pulsed, whilst the designer will use devices mounted in some way, and for power devices with an emphasis on lowering the thermal resistance coefficient; and even if they choose to operate pulsed it is unlikely that it will have the same pulse profile as the measurement data set.

So as can be seen the challenge facing the modelling group at a semiconductor house is mammoth; whilst they are under pressure to have all the modelling information, Process Design Kit (PDK), available for the commercial process launch. For the designer, on whom a project possibly involving a substantial investment is depending, a key consideration is risk management, and essential to this is defining the unknowns and putting in place strategies to contain them. Hence there is a natural tendency to stick with known components with which one has accumulated wide experience. For a company to remain competitive however, it needs to be developing new products which will be competitive in the market and to do this they must be using the latest devices available – or else their competitors who take advantage of new developments, will have a technical advantage. This is particularly the case for the power amplifier designer. This component is often the key to determining system performance, be that in terms of actual output power or linearity. Hence it is this stage which commonly can give a company the edge in the market.

The flip side of this last argument is to know when not to use a device. Traditionally this has meant conducting trial runs on a process, which means paying for a wafer run (or sharing in a 'pizza' mask). This not only requires the investment in cost and time of the wafer run, but also the engineering effort required to design the trial circuits. For this same investment a substantial number of device investigations could be carried out on an active load pull measurement system, for which it would only be necessary to have individual devices supplied by the manufacturer, something that they should be prepared to do at little or no charge if they want them to be seriously considered. The ability to 'play' with a device not only shows what it is really capable of, but also can point out some of its shortcomings at an early stage. Such problems as a tendency to oscillation or a lack of robustness to input overdrive are important to know as early as possible. It is much better to be able to rule out a device before investing too much into it. The author has direct experience of using one manufacturer's GaAs devices which were perfectly happy being operated 4-5 dB into compression, whilst another's failed over time at this level (this would

have been seen on the measurement system as it involved a gradually increasing gate current). In another instance one suppliers GaAs FETs were very reliable when the drain voltage was increased from the nominal 10V to 12V whilst a competitor's failed in the field, despite both parts having identical data sheets. During experimentation it is quite normal to overdrive, over voltage, short circuit the output, do any number of things which stress the device, in a controlled or accidental manner. One thereby gets a feel for the ruggedness of a device, which can lead to protective measures being put in place, the device being selected or ruled out.

The measurements recorded in this chapter were made on devices from only a few wafers. This is not sufficient to be able to assess the variability of device performance; a key requirement of the design process is to be able to estimate the yield of a particular solution and yet this has been poorly addressed in the past with regard to amplifier design. Passive components and structures, (such as substrate materials) will often be supplied with at least some basic information on the key characteristic's standard deviation or tolerance. For transistors one is often supplied with a single S parameter data set. Where variations are given, manufacturers typically provide very wide limits with no guidance as to how the performance may be spread. For example, a well-used, standard GaAs FET 10W device is the SEDI (formerly Eudyna, formerly Fujitsu) FLL120MK, the data sheet, a section of which is shown in Figure 4-4, quotes a typical I_{dss} as 4000mA with an upper limit of 6000mA; a 50% variation. Who would use a resistor in a circuit with a nominal value of 100Ω, but that could be as high as 150Ω? The variation in pinch-off voltage V_P is even greater, from a minimum -1.0V to a maximum -3.5V. There is bound to be a certain amount of 'back covering' going on, and experience has shown that the actual variation even over years of supply has been much smaller. But this doesn't help the designer; what is the point in design centring the output matching circuit if the actual variation of the required output load is not known?

ELECTRICAL CHARACTERISTICS (Ambient Temperature Ta=25°C)

Item	Symbol	Test Conditions	Limit			Unit
			Min.	Typ.	Max.	
Saturated Drain Current	I_{DSS}	V_{DS} = 5V, V_{GS} = 0V	-	4000	6000	mA
Transconductance	g_m	V_{DS} = 5V, I_{DS} = 2400mA	-	2000	-	mS
Pinch-off Voltage	V_p	V_{DS} = 5V, I_{DS} =240mA	-1.0	-2.0	-3.5	V
Gate Source Breakdown Voltage	V_{GSO}	I_{GS} = -240μA	-5	-	-	V
Output Power at 1dB G.C.P.	P_{1dB}	V_{DS} = 10V I_{DS} = 0.55 I_{DSS} (Typ.), f = 2.3GHz	39.5	40.0	-	dBm
Power Gain at 1dB G.C.P.	G_{1dB}		9.0	10.0	-	dB
Power-added Efficiency	η_{add}		-	40	-	%
Thermal Resistance	R_{th}	Channel to Case	-	3.3	4.0	°C/W

CASE STYLE: MK

G.C.P.: Gain Compression Point

Eudyna

Edition 1.1
July 1999

1

Figure 4-4, Part of data sheet of FLL120MK 10W GaAs FET.

The question therefore arises; how, using the process described in this research, does one establish the tolerance on device performance? Firstly, more devices than were measured here are needed. There are mathematical principles concerning probability that can be used to determine how many transistors need to be measured to achieve a certain level of confidence in the results. This isn't an approach one would typically use on, for example, a spectrum analyser or a vector network analyser when reducing the noise on a display trace. If the response is 'noisy' the standard approach is to use averaging and adjust the number to get an acceptably smooth response. The measurement, modelling and design process described in this document is entirely practically based and it would therefore seem appropriate to advocate that the same approach should be made to the assessment of variability, i.e. measure between 5 and 10 devices (from different wafer runs) and see how much, for example, the optimum load impedance varies. If the results are very consistent then that is probably sufficient, if not, increase the number of units measured until a clear pattern emerges. It is of course necessary to monitor results over time, particularly as new wafers are processed, but this just good quality control practice. Obviously if the spread of results (in this case impedance) is too large, then it suggests that perhaps this isn't the best device for the application or that the circuit design will need to be able to encompass the variability with implications for tuning, Select on Test (SOT) components and test time. It

should be noted that this is a recognised problem in amplifier design, [1], suggests using four S parameter data sets, associated with the lowest and highest gain, and lowest and highest current, and analysing the circuit with these to see the 'spread' of performance.

The measurements that follow where made both to investigate how the measurement system could be used in a design process and to prove the feasibility of using the measurement system in the higher frequency range (previous work had concentrated on communication band frequencies below 3 GHz). Some of the data obtained has been used in the design of MMIC stages as detailed in chapter 6.

4.2 DC-IV Measurements

The conditions under which DC-IV measurements are conducted affect the results obtained as explained in detail by Ladbroke [2]. Measurements were made on the devices shown in Figure 4-2 and Figure 4-3 using the Accent (now Auriga) DC-IV measurement system under static and dynamic conditions, with a number of different pulse widths in the dynamic measurements. DC-IV measurements can be made on the ALP system itself, the power supplies are under control of the software, but there is no pulsing capability itself. The results are shown in Figure 4-5 to Figure 4-12 and show how the DC-IV curves will vary depending upon the measurement methods employed. There are a number of interesting factors to note regarding the DC-IV measurements and the type of device measured. It is first worth noting that these devices are Double heterojunction pseudomorphic High Electron Mobility Transistors (DpHEMTs) [3] which have an additional current carrying layer in order to increase power handling. This results in a significantly higher I_{dss} (drain current at V_{gs}=0 V_{ds}=V_k) over a conventional HEMT (310:220 mA/mm) and I_{dssm} (maximum saturated drain current at maximum gate bias voltage) (600:380 mA/mm) [4]. A notable characteristic of these devices is that I_{dssm} is almost twice I_{dss}. This has important implications when selecting the loadline impedance and quiescent bias point.

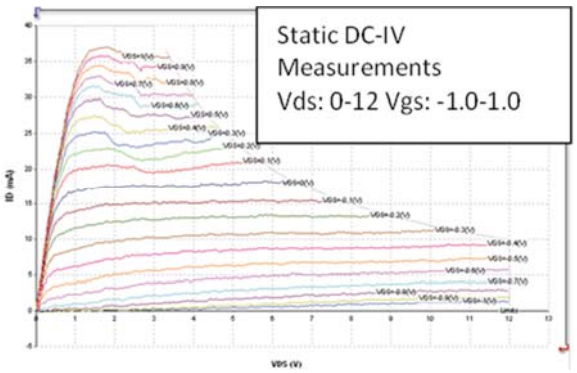


Figure 4-5, T1x75 Static DC-IV Curves

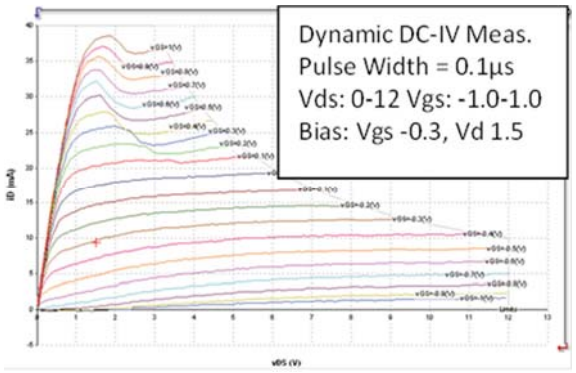


Figure 4-6, T1x75 Dynamic DC-IV Curves 0.1µs

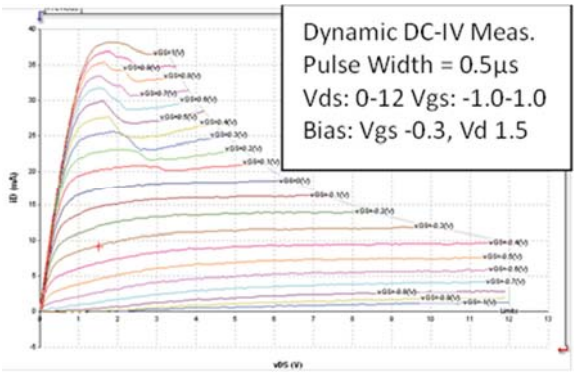


Figure 4-7, T1x75 Dynamic DC-IV Curves 0.5µs

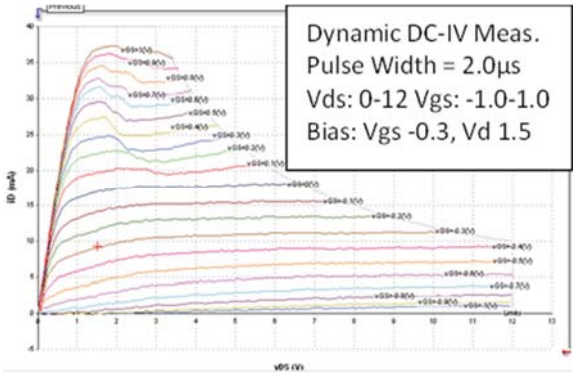


Figure 4-8, T1x75 Dynamic DC-IV Curves 2.0µs

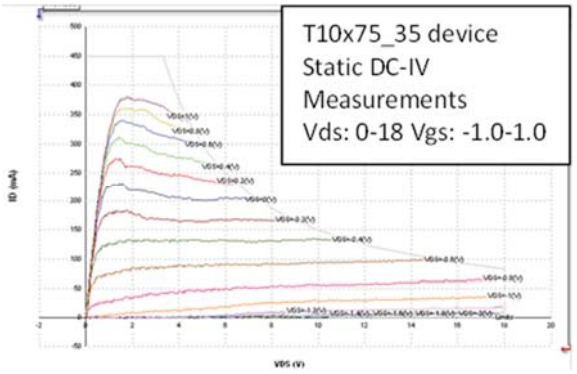


Figure 4-9, T10x75 Static DC-IV Curves

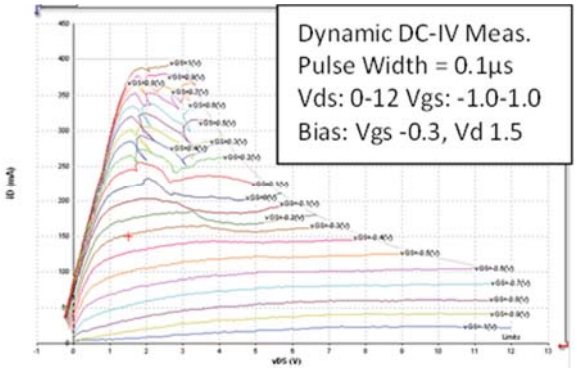


Figure 4-10, T10x75 Dynamic DC-IV Curves 0.1µs

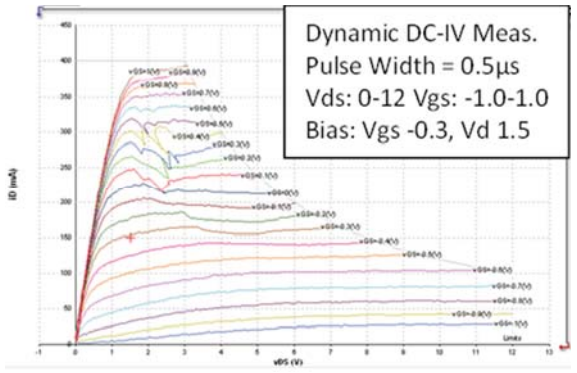


Figure 4-11, T10x75 Dynamic DC-IV Curves 0.5µs

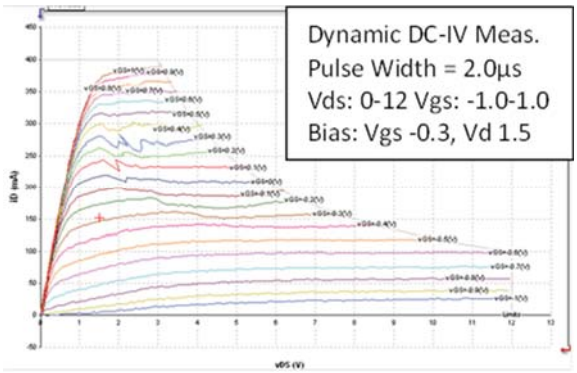


Figure 4-12, T10x75 Dynamic DC-IV Curves 2.0µs

As can be seen from Figure 4-5 to Figure 4-12, above a V_{gs} of 0 volts the drain current peaks, falls off, and then climbs again. In the case of Figure 4-10 to Figure 4-12 a further instability in the curves is evident. This is due to 'hot electron' tunnelling [5] which is a low frequency effect which does not occur at RF as the voltage collapses before the electrons can move through the layers. However as Figure 4-13 to Figure 4-16 show these effects also depend upon the initial bias conditions. The grey (feint) line in these figures represents the power limiting envelope for the device.

The general DC-IV curves for the FD30 process, shown in Figure 4-1, do not say under which bias conditions they are made, nor whether they are static or dynamic (and in which case what are the pulse conditions) and finally on which device size they were performed. This exemplifies one of the basic premises of this research; when supplied data from manufacturers, it is rarely clear under what conditions they were generated and therefore how applicable this data is to a particular design environment. The DC-IV curves measured also indicate another problem, the data from these measurements forms the basis of most nonlinear model generation, the question therefore arises which data set should be used for creating the model? Which most accurately represents the behaviour of RF waveforms?

When using the T10x75 device the industrial partner had found a preferred bias of V_d=9V and V_{gs} adjusted for a fixed 150mA I_{ds}, which from Figure 4-9 to Figure 4-16 ranges from -0.3v to -0.4v. Referring to Figure 4-1, the left axis must first be scaled for a 0.75mm device, or alternatively the current can be increased for the equivalent 1mm periphery:

$$\frac{150}{0.75} = 200mA$$

There are no traces at 9V that reach this level so it is necessary to extrapolate them to this voltage, and this suggests that a V_{gs} of -0.2V is required.

The DC-IV measurement system used for acquiring the curves in Figure 4-5 to Figure 4-16 was not in calibration and so the absolute accuracy of the measurements cannot be certain, however the point that is trying to be made is the differences between the performance based upon measurement conditions; pulse width and quiescent bias. Further it was observed during measurements that the gate bias voltage necessary for 150mA changed after the device was exercised, (driven hard into compression) during the active load pull. This aging effect suggests that before detailed measurements are taken it would be best to 'soak-test' the devices under bias and RF drive conditions so that performance

has stabilised[1]. It was observed after testing a number of devices that the actual gate bias voltage required for 150mA at 9V and after aging was taken into account, was in the region of -0.15V.

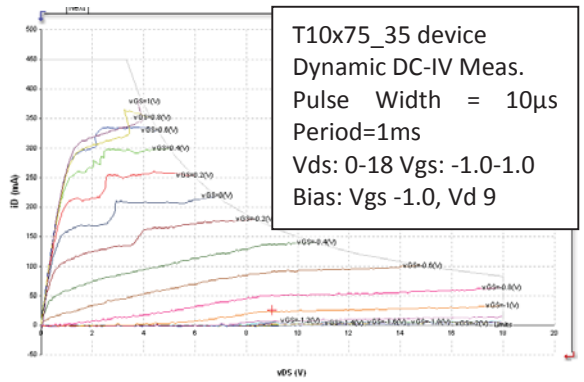


Figure 4-13, Dynamic DC-IV from a bias point of Vds=9v and Vgs=-1v

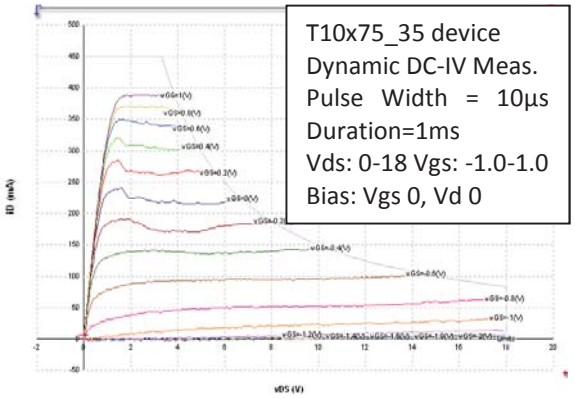


Figure 4-14, Dynamic DC-IV from a bias point of Vds=0v and Vgs=0v

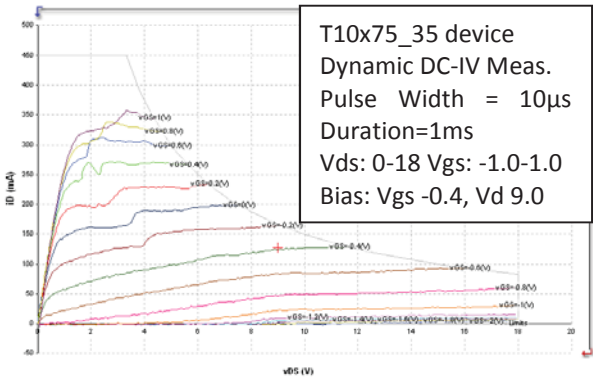


Figure 4-15, Dynamic DC-IV from a bias point of Vds=9v and Vgs=-0.4v

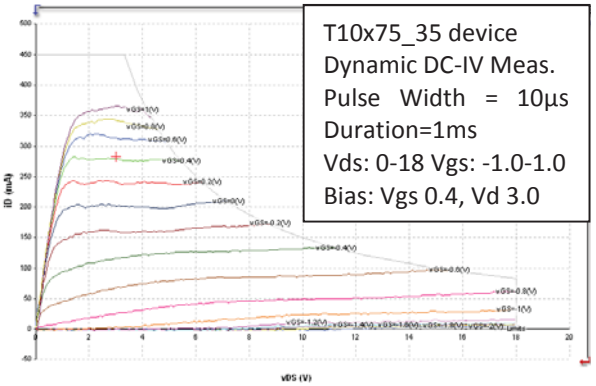


Figure 4-16, Dynamic DC-IV from a bias point of Vds=3v and Vgs=0.4v

[1] Gate degradation is a recognised phenomenon [13] and results from a migration of the gate into the substrate, as discussed earlier in this work it tends to be a self-limiting effect as the increased resistance reduces a primary cause – drain current.

4.3 Device De-Embedding

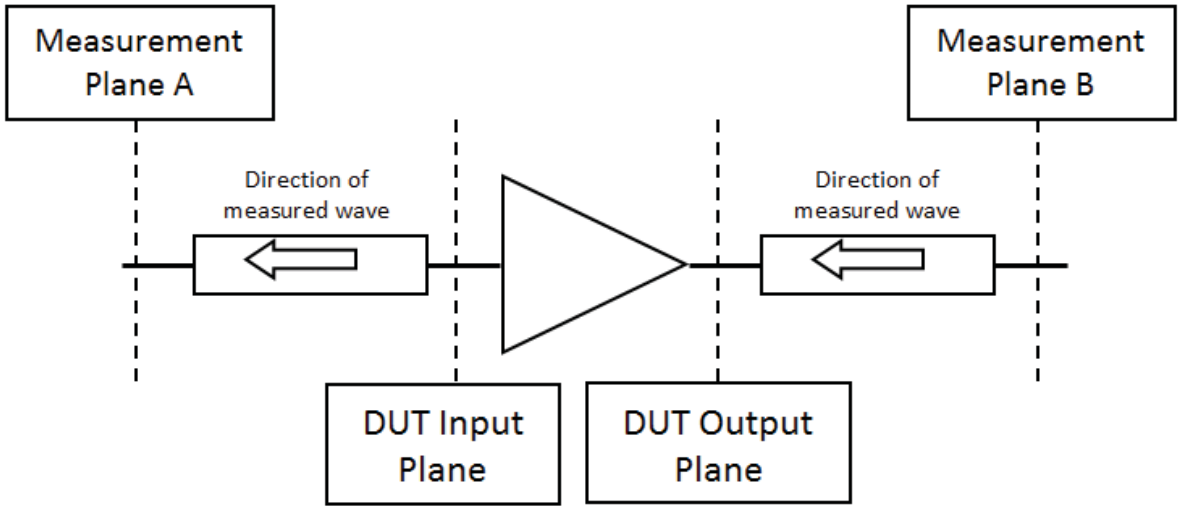


Figure 4-17, Measurement (Calibration) and Device Planes

The DC-IV measurements described previously, although using wafer probing do not require de-embedding as they are DC or Low Frequency. However to measure the high frequency performance of the actual device to be used in circuit designs it is necessary to remove the effects of the probe landings and the feed lines that can be seen in Figure 4-2 and Figure 4-3. Conventional on-wafer calibration standards were used in the calibration of the system, which provide a reference plane at the tips of the probes (Ground Signal Ground – GSG 200µm in this case). The devices measured however are 'embedded' via pads for the probe tips to land on and 30 x 200 µm lines to connect to the actual transistors. These lines are used so that there is sufficient separation between the probes so that they don't couple and that their presence doesn't affect the Device Under Test (DUT) behaviour. These features would not exist in the actual circuit realization and their effects therefore need to

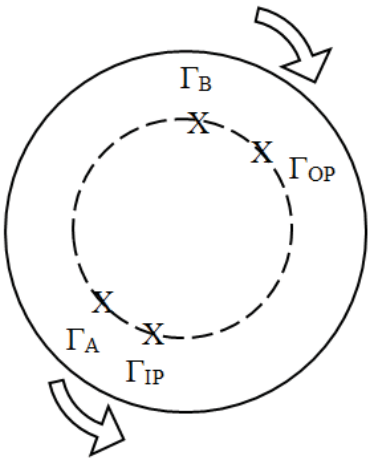


Figure 4-18, Rotation of Measurement planes through de-embedding, referring to Figure 4-17.

be removed. This process of moving the calibration or reference plane up to the device plane is referred to as de-embedding. Referring to Figure 4-17, in the case of the input reflection coefficient the wave at the measurement plane has travelled along the embedding line and so we must negate this effect in order to display the reflection coefficient at the device input plane. For the Load reflection coefficient we are measuring the load presented at the end of the embedding line. To translate this load to the device plane we must add the effect of the feed line. Note: this is for the case where we are looking at the impedance presented to the device. When we are looking at the measured S parameters we must negate the line length on both input and output.

The method of de-embedding used was to measure the S parameters of a pair of back-to-back lines (also included on the DUT substrate). The lines and probe pads are assumed to be symmetrical about the centre. Optimising a model of these lines within a linear CAD package, Figure 4-19, allows a set of S parameters to be produced from a 'half set', which equate to one set of probe pads and a 30 x 200 µm line. This information is then

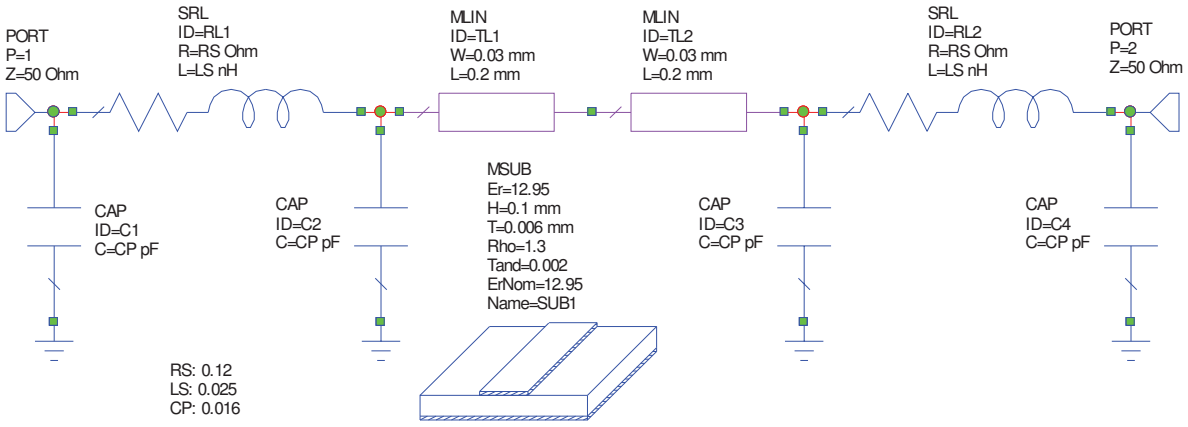


Figure 4-19, Equivalent Circuit for De-embedding structure.

used to modify the error coefficient files, generated when the system is calibrated, and shifts the reference plane to the edge of the device structure itself.

The nominal length of the transmission line in the launcher is 200µm and the width is 30µm. The following MSUB properties are appropriate for this line on a Gallium Arsenide, GaAs substrate are: Er=12.95, H=100µm, T=6µm, Rho=1.3, Tand=0.002. The pads are described by a simple PI network of two shunt capacitors and a series resistor-inductor, Figure 4-19. The values are chosen to get the correct fit to the measured through line for the magnitude and phases of S11 and S21. Note that the influence of the pad parasitics is to move the effective line impedance to a lower value than the microstrip line itself. The series

inductor includes a resistive term as it is often observed that there is an additional loss over and above the line loss and that it arises from radiation, or higher order mode generation, at the probe/pad interface. This is usually only significant above 20GHz but is useful to include it for curve fitting below this frequency. The individual values for the C's and L's will depend on the particular system set up (probes used, size of landing pads, etc.), but as "ball-park" figures, shunt capacitance is about 16fF and series inductance is about 25pH in the Cardiff measurements system with 200µm probes. The performance of these values and two sets of measured data are shown in Figure 4-22 and Figure 4-23. There is also an extent to which the frequency range must also be considered. As mentioned, radiation and higher order modes contribute to the measured results, but are not included in the models. Thus depending on whether the E-M simulation takes these effects into account, it may be of more use as a confidence check rather than for use in the de-embedding. Optimising the equivalent circuit parameters it is important to get the weighting correct between the characteristics. Key is the insertion (S21) phase as this affects the waveforms to a more significant degree than the return loss, (when return losses exceed about 16 dB). The actual balance that is achieved can be debated, but in the calibration used and for the frequency range of the measurements recorded in this chapter, values determined are: Cp=0.01pF, Ls=0.02nH and Rs=0.15Ω and the results as shown in Figure 4-24 and Figure 4-25. The emphasis has been on matching the phase over the full 40GHz bandwidth and the insertion loss and return loss up to 18GHz.

After a full on wafer calibration, a test, through 50Ω line is measured to demonstrate the directivity of the measurement system. As can be seen from Figure 4-26 the system achieves a dynamic range in excess of 22dB to 40GHz. Another key indicator of correct calibration of the system is to observe the input and output voltage and current waveforms. At the input these should be in phase and at the output in anti-phase as shown in Figure 4-27 and Figure 4-28. Similarly this can be applied as a check on the de-embedding. After the de-embedding has been applied to the measurements to move the calibration to the centre of the structure, the input waveforms should be in phase and the output in anti-phase.

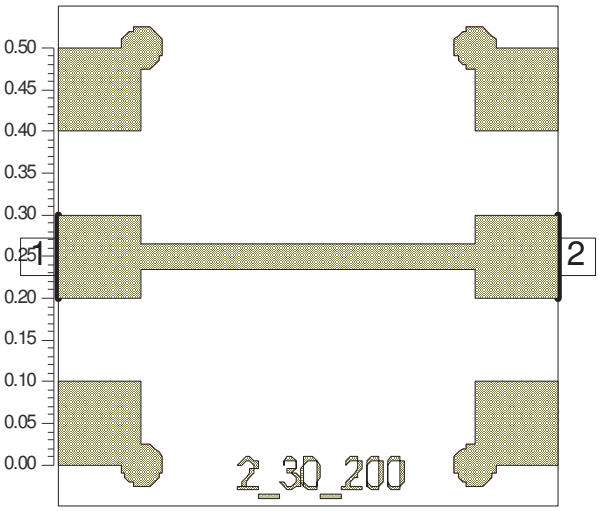


Figure 4-20, E-M Simulation structure

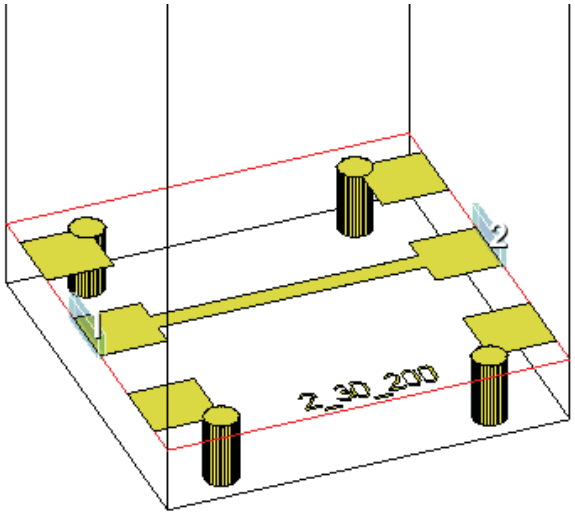


Figure 4-21, 3-D E-M Structure View

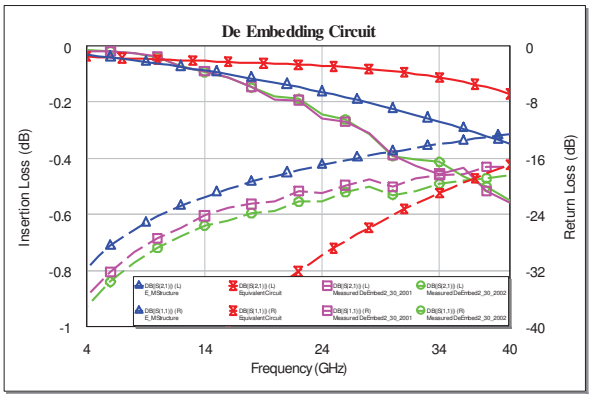


Figure 4-22, Comparison of E-M and Linear Simulations with 2 sets of measured data of de-embedding structure.

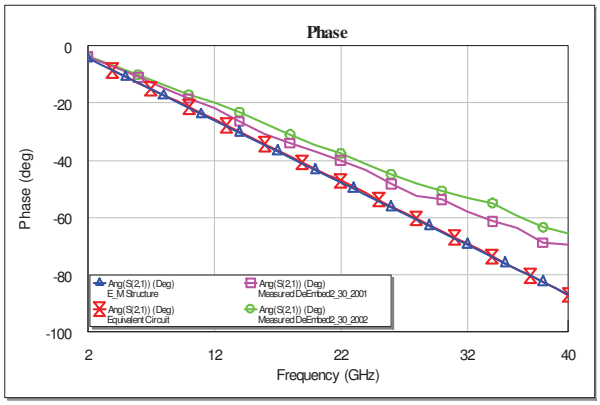


Figure 4-23, Comparison of E-M and Linear Simulations with 2 sets of measured data of de-embedding structure.

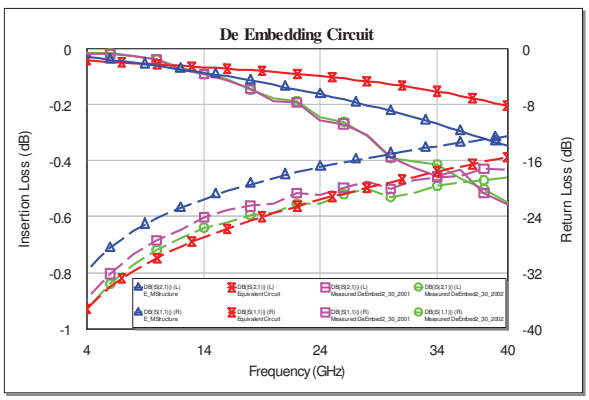


Figure 4-24, Optimised Equivalent Circuit parameters for Fundamental Frequency up to ~18GHz

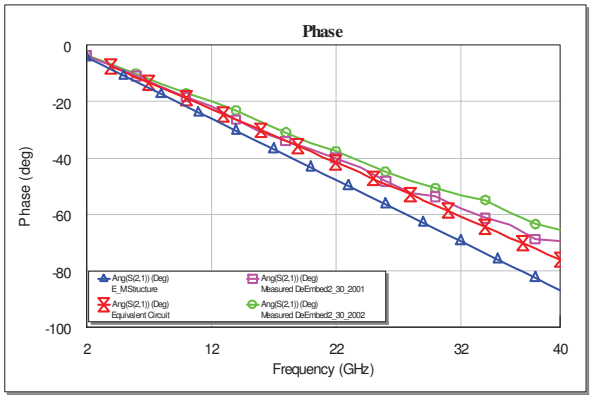


Figure 4-25, Optimised Equivalent Circuit parameters for Fundamental Frequency up to ~18GHz

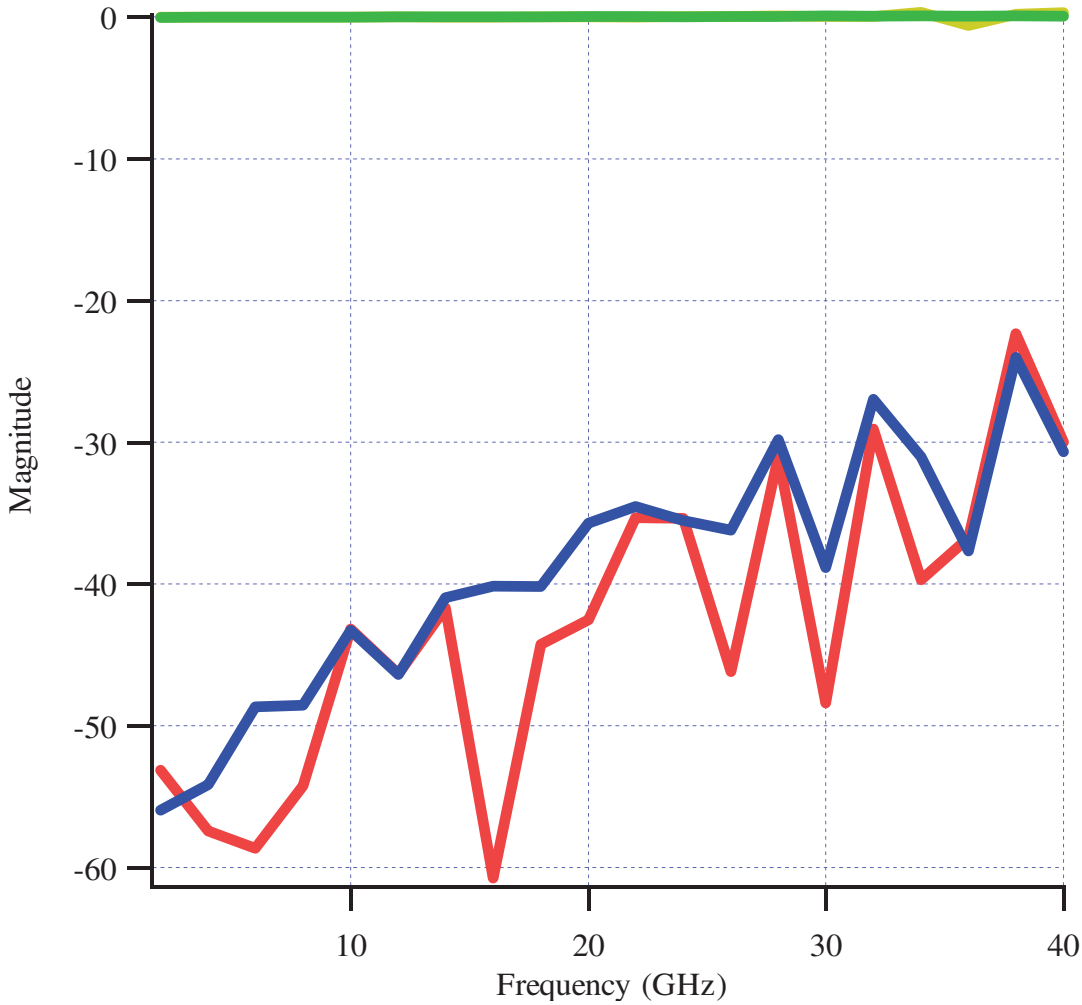


Figure 4-26, dynamic range of measurement system, S21 & S12 (green & yellow) S11 & S22 (red and blue).

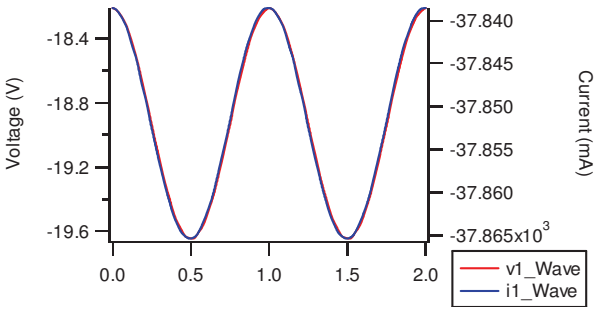


Figure 4-27, Correctly Calibrated input voltage waveforms

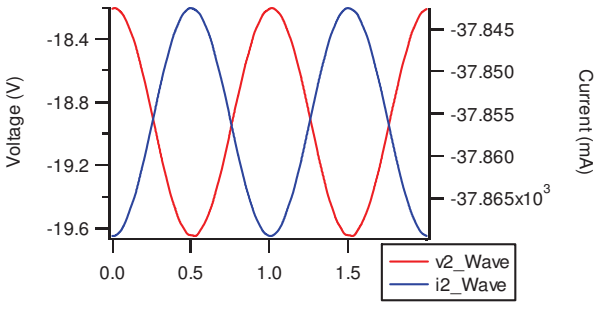


Figure 4-28, Correctly calibrated output waveforms

4.4 Device Measurements

As a first step in the design process a device was chosen that would typically be used in the industrial partner's products. The selection was also based upon the measurement capabilities of the system at the time. A major consideration is the drive power available; to evaluate the optimum Power Added Efficiency (PAE) it is necessary to be able to supply a sufficient amount of drive power to the device, which is typically high enough for the device to be several dBs into compression. The signal generators incorporated into the system have output power in range of 21 to 25dBm depending upon frequency. Between these and the device input are cables, switches, coupler, bias tees and wafer probes. Despite efforts to keep losses to a minimum the system insertion loss of this line-up amounted to just over 4dB at 18GHz. On the input side the other significant factor is the reflection coefficient of the transistors gate. Figure 4-29 shows the effect input reflection coefficient has on mismatch loss, and the typical range of Γ_{in} presented by microwave devices in the frequency range 1-18GHz. This shows that this is often the dominant feature affecting drive level requirements, and that typically >6dB of input is lost due to input mismatch, (and hence the considerable benefit arising from input tuning.

In order to provide the required additional test signal power, external amplifiers are added to both input and output. Initially the amplifiers on either side of the bias "Tees"

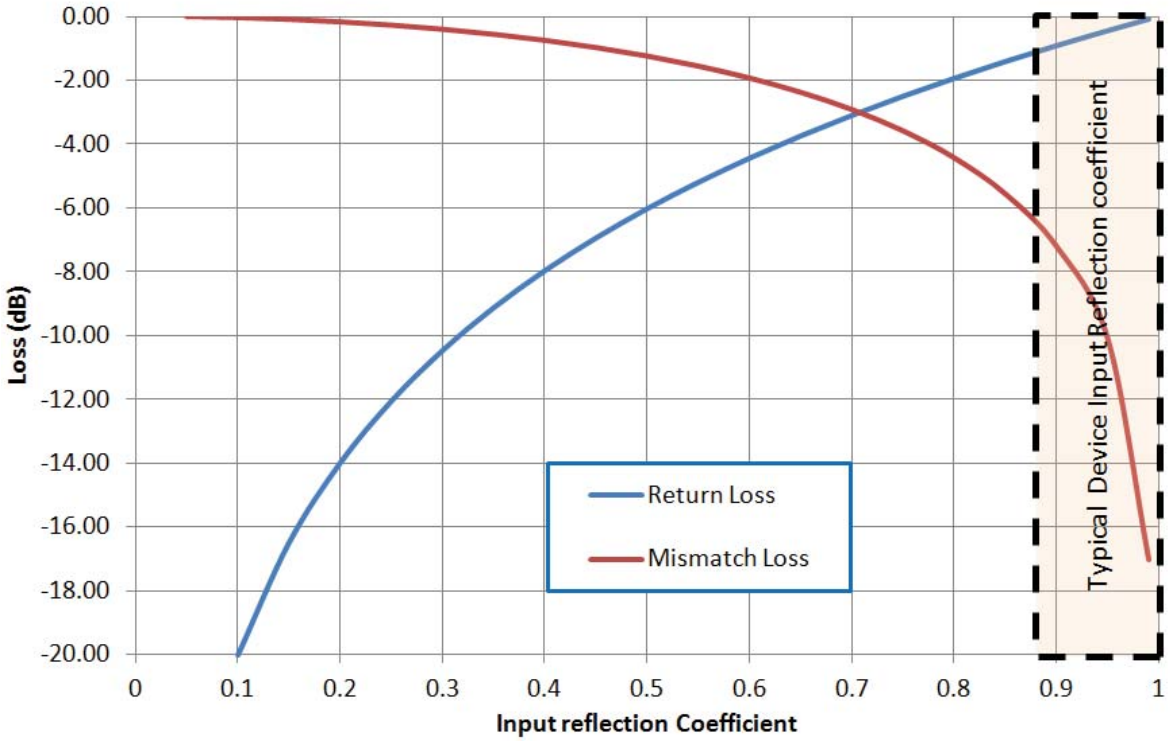


Figure 4-29, Impact of Input Reflection Coefficient on Mismatch Loss.

provided up to 1W of drive signal for input and load-pull, and were limited to 26GHz maximum frequency, whilst this was sufficient for the fundamental both higher powers (due to increased losses and higher reflection coefficients) and frequencies are needed for the harmonics. However as the amplifiers are outside of the calibration planes they can be substituted as required (and as funding allowed). The power required for load pulling was discussed in more detail in chapter 3. In the communication bands measurement systems, specific narrow band amplifiers could be used for the fundamental and each harmonic. For the broadband applications considered here it is impractical to do this and hence wide bandwidth units are preferred which greatly add to the cost of the system.

The operating mode required of the devices to be characterised was primarily a form of overdriven class A, the theory of which was described in chapters 1 and 2. Where the bias is not prescribed the process of evaluating a device will often start with examining what performance can be achieved under different bias conditions, however this should always be done in the light of particular applications; radar amplifiers are typically biased in class C as duty cycles are low, heat management and efficiency are high priorities and linearity is of little importance. Conversely instrumentation amplifiers are most commonly biased in class A as linear gain and high output power are the key parameters, efficiency tends to be less important.

For the purposes of this project the objective was to determine the optimum output power and efficiency loads in the frequency range 4-18GHz. Initially only fundamental loads were considered, but later harmonic impedances were included. Similarly the device performance in other operating modes was investigated. In parallel with the measurements, nonlinear simulations using models from the Process Design Kit (PDK) provided by the manufacturer were conducted. These not only gave an interesting comparison and analysis of the accuracy of the model, but acted as a guide as to what could be expected from the device.

In determining the optimum loads for output power and efficiency in a measurement based environment an iterative process is necessary, although estimations of the load and drive power can put the user in the right ‘ball-park’. For example, the output capacitance of GaAs FETs, although highly process and manufacturer dependant, is typically in the order of 0.3pF/mm. The optimum load comprises of a resistive and reactive

component (although as we are looking at a shunt connection it is more useful to calculate the susceptance) and is given by {4-1} and {4-2}:

$$R_{Lopt} = \frac{(V_{DC} - V_k)}{I_{dssm}/_2} \quad \textbf{\{4-1\}}$$

$$B_{Lopt} = 2\pi f C \quad \textbf{\{4-2\}}$$

Thus in the case of the T10x75 device we have;

$$R_{Lopt} = \frac{(9 - 1.5)}{.33/_2} = 45\Omega$$

and,

$$B_{Lopt} = 2\pi \times 6 \times 10^9 \times (0.3 \times 0.75) \times 10^{-12} = 8.68 \times 10^{-3}\, mho$$

so the approximate optimum load admittance is 1.11 –j0.434 (normalised in a 50Ω system), which is equivalent to a reflection coefficient of 0.208/_116°. Note that this theory applies to the optimum output power load. Also note the approximations in the theory, from Figure 4-1, it can be seen that the exact value of the knee voltage V_k is not easily defined. Similarly the measured DC-IV plots suggest a wide tolerance in the possible values attributed to V_k and I_{dssm}.

The input reflection coefficient (Γ_{In}) from the small signal S parameters and gain can be used to estimate the required drive power, working back from an estimate of the output power based on the foundry process, in this case 1W/mm, hence for this device 0.75W (28.8dBm). The magnitude of the input reflection coefficient at 6 GHz in 50Ω is 0.89, which from Figure 4-29 results in a mismatch loss of ~6dB. Conversely there will be an increase in gain from the output matching, which with a small signal |S22|of 0.59, will be about 2dB. So the drive level that can be expected will be of the order of:

Pout – (|S21| - Input Mismatch + Output Match Gain – Compression)

28.8 – (11.8 – 6 + 2 – 2) = 23dBm

It should be remembered that these are starting values to be used as a guide; the point of the measurements is to determine accurately the device characteristics.

The S parameters are also useful in calculating the input and output stability circles; these determine the stable load impedances. One of the most common measures of stability is Rollet's Stability Factor, (K). This is derived from S parameter measurements as shown in {4-4}. The magnitude of K in itself only gives an indication of how stable the transistor is at the particular frequency; when K>1 the device is unconditionally stable with any real source and load terminations. Arguably more useful, is a graphical representation of K, whereby the source or load impedances that could cause instability are plotted on the Smith Chart, to this end new variables C1 and C2 are calculated, {4-5} and {4-8} which are used to determine the location of the input and output stability circles, rS1 and rS2 and the radius of the circles pS1 and pS2 are calculated using {4-5} to {4-10} as shown below:

$$D_S = S_{11}S_{22} - S_{12}S_{21} \quad \textbf{\{4-3\}}$$

$$K = \frac{1 + |D_S|^2 - |S_{11}|^2 - |S_{22}|^2}{2.|S_{21}||S_{12}|} \quad \textbf{\{4-4\}}$$

$$C_1 = S_{11} - D_S S_{22}^* \quad \textbf{\{4-5\}}$$

$$r_{S1} = \frac{C_1^*}{|S_{11}|^2 - |D_S|^2} \quad \textbf{\{4-6\}}$$

$$P_{S1} = \left|\frac{S_{12}S_{21}}{|S_{11}|^2 - |D_S|^2}\right| \quad \textbf{\{4-7\}}$$

$$C_2 = S_{22} - D_S S_{11}^* \quad \textbf{\{4-8\}}$$

$$r_{S2} = \frac{C_2^*}{|S_{22}|^2 - |D_S|^2} \quad \textbf{\{4-9\}}$$

$$P_{S2} = \left|\frac{S_{12}S_{21}}{|S_{22}|^2 - |D_S|^2}\right| \quad \textbf{\{4-10\}}$$

The input stability circle is a contour in the source plane that indicates source termination values that make the output reflection coefficient have a unity magnitude. Conversely, the output stability circle is a contour in the load plane that indicates load termination values that make the input reflection coefficient have a unity magnitude. A reflection coefficient less than unity indicates a stable device, while a reflection coefficient

greater than unity indicates a potentially unstable device. In the simulation software used[2] the display of the stability circle indicates the unstable region using a circle drawn with a dashed line in the unstable region. If the dashed circle is inside the solid circle, then the outside of the circle indicates the stable region, whereas if the dashed circle is outside the solid circle, the inside of the circle represents the stable region, Figure 4-30.

The stability circles are approximate as they are based upon the small signal S parameters, however they do indicate potential problem areas, particularly low frequency instability; note in Figure 4-30 the output load stability circle at 100MHz (lowest red curve), encompasses most of the top half of the Smith Chart, and at 6 GHz it is not too far from the optimum load just calculated (0.208/_116°). The dangers of instability are that the device could be damaged through drawing excess current, or that the oscillation modulates the fundamental to such an extent that it becomes difficult to converge on a load impedance. Low level oscillations in themselves do not necessarily cause a problem with the measurement system as such, but they will affect the performance of the device and so need to be avoided. There is the added problem that as the oscilloscope sampling acts as a filter, only oscillations that are synchronous with the sampling frequency (sub or harmonically related) will be detected, hence it is unlikely to show up in the waveforms. Thus the operator needs to be on the lookout and check for an oscillation (for example using a spectrum analyser).

The main approach in ensuring stable operation is to ensure resistive terminations at low frequencies in the bias “tees” and elsewhere presenting a reasonable 50Ω load. The measurement system itself will (barring the failure to converge on a load during active load pull) be able to operate in the presence of instabilities, and a typical indication of this behaviour is the ‘collapse’ of the load pull grid.

A recommended procedure for the device measurement is therefore:

1. Calculate optimum output power load from bias, device size, frequency and capacitance (process and device topology dependant).
2. Calculate approximate optimum drive level.
3. Check stability circles (from measured or simulated small signal S parameters).

[2] Microwave Office, from AWR Corp.

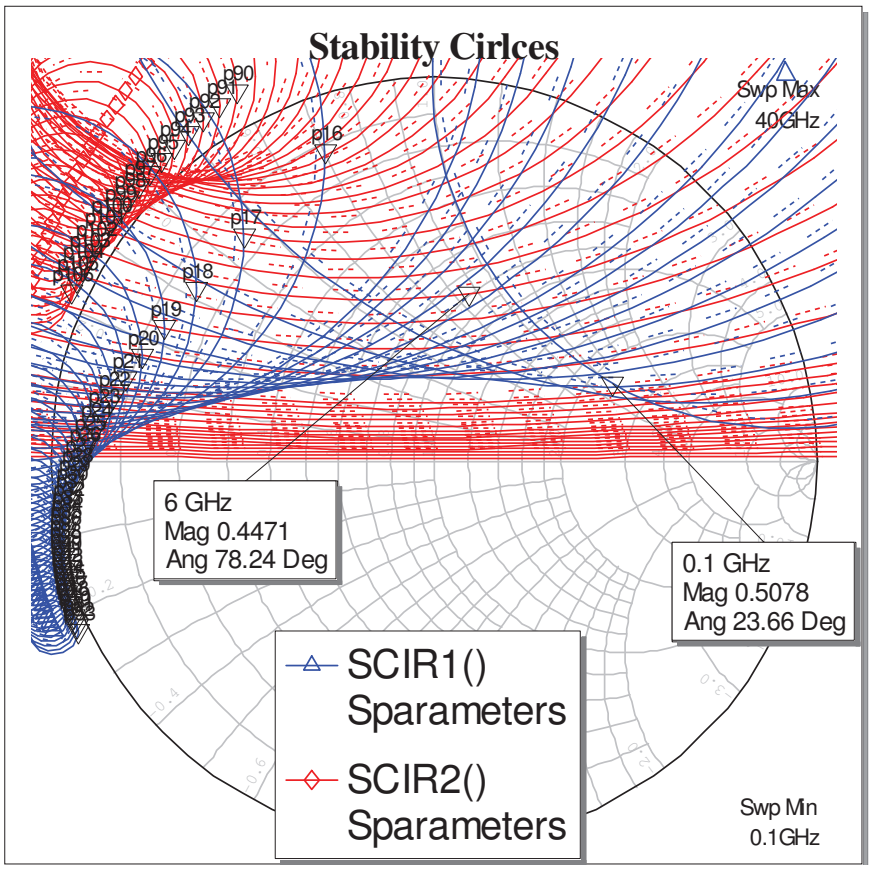


Figure 4-30, Input (Blue) and Output (Red) Stability Circles from small signal S

4. Conduct load pull, establishing a grid around the theoretical optimum and at the calculated drive power.
5. Conduct power sweep at the optimum load found in step 4. and hence determine the optimum PAE power level.
6. Conduct load pull at the power level from step 5., possibly moving the grid further from the centre of the Smith Chart; the optimum PAE load has a higher resistive component than the output power.
7. Conduct power sweep at the optimum PAE load, if the optimum drive level is different from that used in step 6., then repeat from this step until the load pull has been conducted at the optimum drive level.

Of course if a large enough grid is used and multiple drive power load pulls are conducted automatically then all the data (and more!) outlined above can be obtained in one measurement set. Experience has shown however that using the steps suggested above, at least for the initial measurements, ensures that problems are ironed out. There is nothing quite as frustrating as setting up a weekend long set of measurements only to find on Monday that there was insufficient load pull drive power. Once a measurement set has

been well defined then a semi-automatic measurement routine can be followed, which is particularly useful when comparing a number of devices.

An initial set of measurements were conducted on the 10x75 device, biased at V_{ds} of 9V and V_{gs} adjusted to give an I_{ds} of 150mA (no RF present). The results are shown in Figure 4-31 to Figure 4-34. Conducting similar investigations, but using the PDK model it can be seen that the theoretical optimum load and the simulated optimum are very close, Figure 4-35.

Comparing the measured and simulated results in Table 4-1 shows the difference between what was expected and what was observed with this device. Initially only devices from a single wafer were available. The last data row on the table includes the results from a device manufactured on a new wafer run, D_1309.

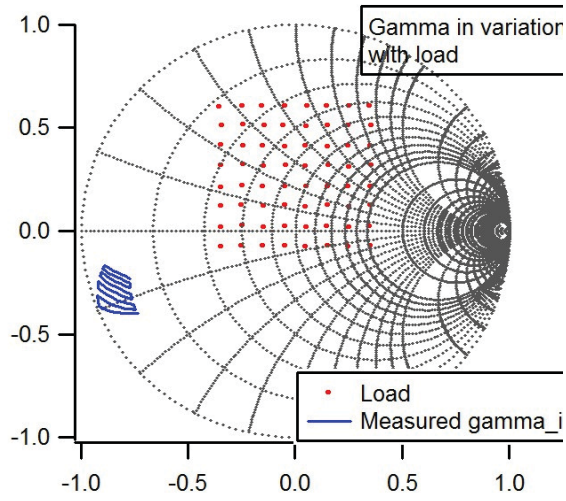


Figure 4-31, Fundamental Load Grid and impact of load impedance on input reflection coefficient

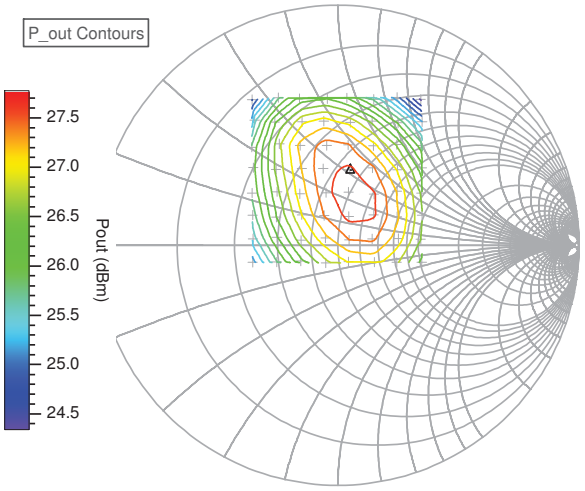


Figure 4-32, Output Power Contours at 6 GHz 9V and 150mA

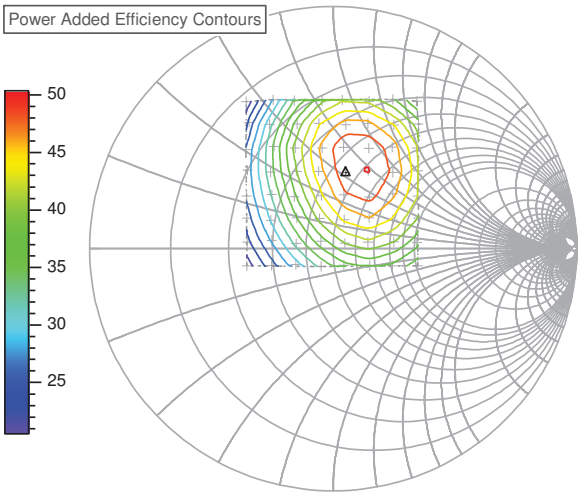


Figure 4-33, PAE Contours at 6 GHz 9V and 150mA

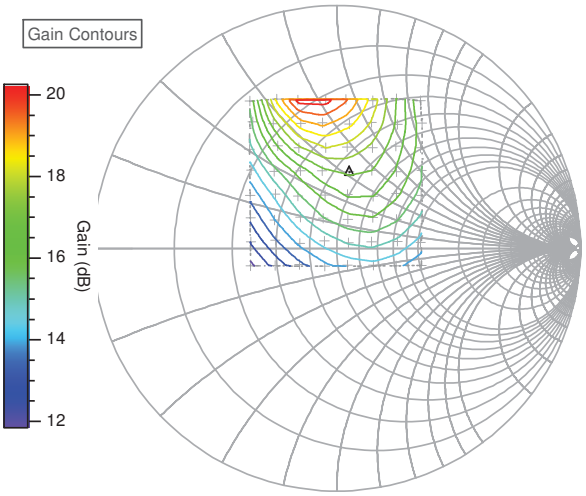


Figure 4-34, Available Gain Contours at 6 GHz 9V and 150mA

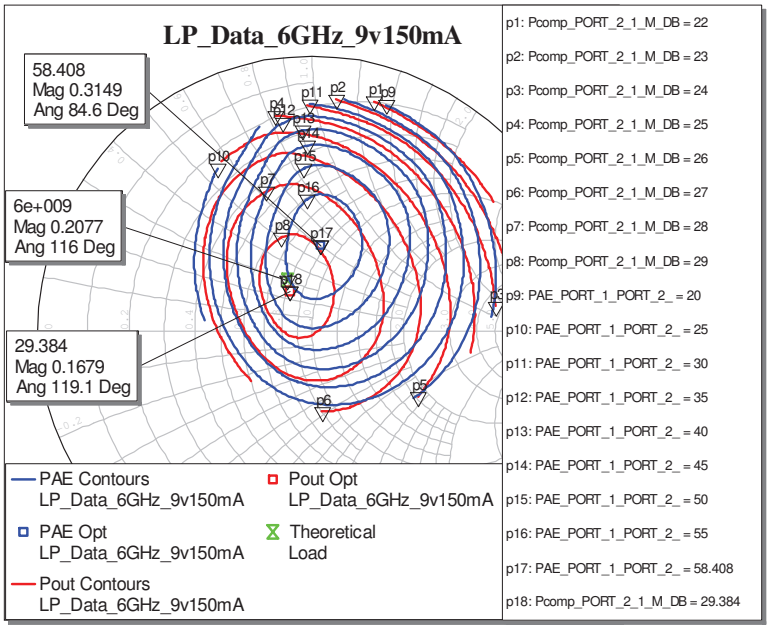


Figure 4-35, Simulated Output Power and PAE contours and Theoretical Optimum Output Power Load

Method	Max. PAE (%)	Load Mag/Ang	Max. Pout (dBm)	Load Mag/Ang	Equivalent Circuit Load (Pout)	
					R (Ω)	C (pF)
Measured Original	51.4	0.35/_65.9°	28.0	0.22/_78.1°	60	0.20
Simulated	58.4	0.31/_84.6°	29.4	0.17/_119.1°	45	0.18
Theory	-	-	28.8	0.21/116	45	0.23
Measured D_1309	55.1	0.43/_70.0°	28.2	0.16/_91.2°	52	0.17

Table 4-1: Comparison of Measurement and Simulation Optimum Loads

Besides the optimistic nature of the PDK model predictions for both PAE and output power, the phase of the optimum output power load is about 30-40° from the original measured value. The PDK model is however close to the theoretical load, this suggests that factors not included in the 'simple' model (of the theory) come into play. Looking further at the differences between simulated and measured at the optimum output power load and then at the optimum PAE load, is therefore required.

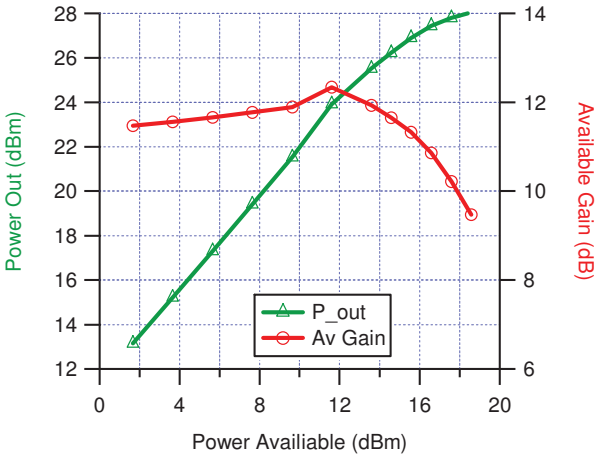


Figure 4-36, Measured Output Power and Available Gain at 6 GHz

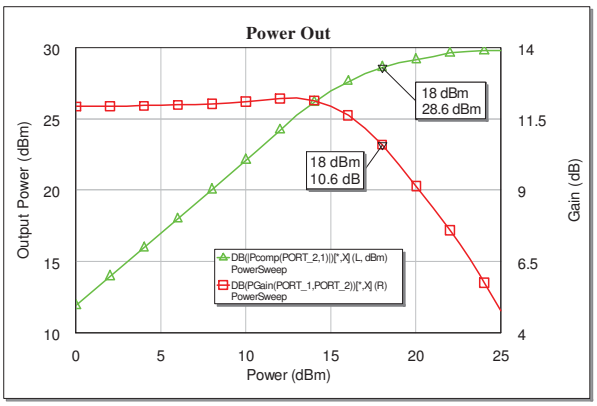


Figure 4-37, Simulated Output Power and Available Gain at 6 GHz

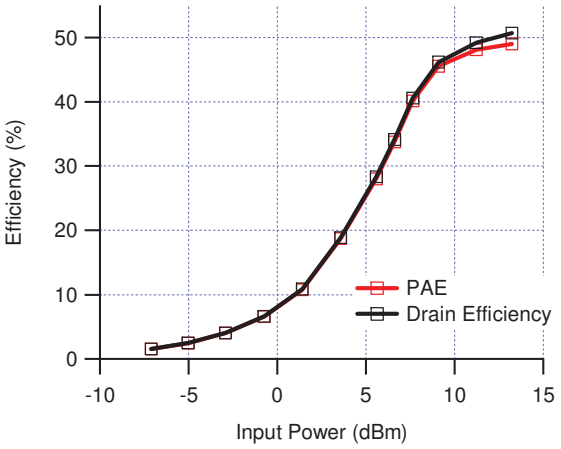


Figure 4-38, Measured Drain Efficiency and PAE at 6 GHz

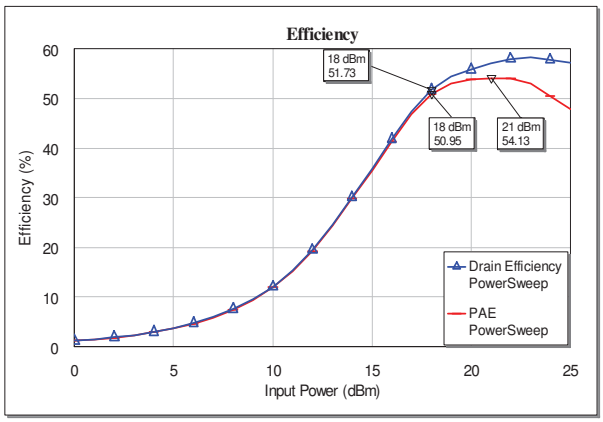


Figure 4-39, Simulated Drain Efficiency and PAE at 6 GHz

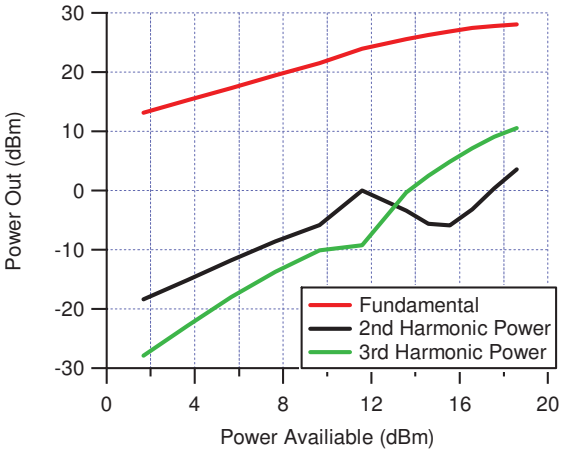


Figure 4-40, Measured Fundamental, 2nd and 3rd Harmonic Output Power at 6 GHz

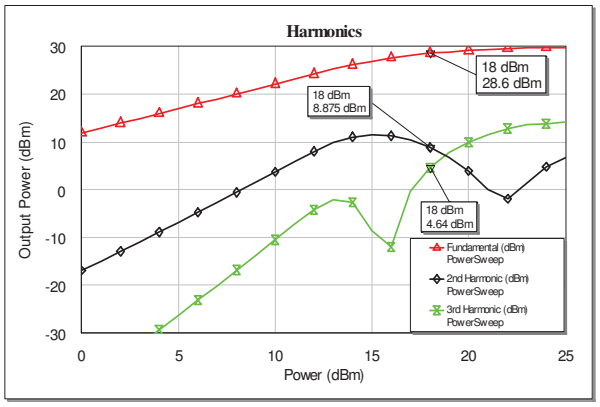


Figure 4-41, Simulated Fundamental, 2nd and 3rd Harmonic Output Power at 6 GHz

From the results of Figure 4-36 to Figure 4-41 there appears to be a good correlation between measurement and simulation, with the notable exception of the predicted harmonic output power levels. The shape of the efficiency curves of Figure 4-38 and the results of Figure 4-39 suggest that the peak efficiency occurs at about 2 dBs higher than the

maximum power used in the simulation, note that available power on the measured data is the same as the input power of the simulation. The previous measurements and their agreement with both fundamental theory and the nonlinear simulation (within the limitations described) gave confidence in the system construction. The next stage was to repeat the measurements at other frequencies, in the range 4 to 18GHz.

T10x75_35 at 9v 150 mA - Measured									
Freq. (GHz) & Compression (dB)	4 @ -2.2			6 @ -3.7			8 @ -3.8		
Maximums	Value	Γ mag	Γ pha	Value	Γ mag	Γ pha	Value	Γ mag	Γ pha
PAE (%)	48.8	0.26	44.6	50.2	0.35	65.9	50.8	0.52	80.8
Pout (dBm)	27.5	0.21	65.1	27.8	0.23	78.8	27.9	0.31	94.3
Freq. (GHz) & Compression (dB)	12 @ -3.4			16 @ -2.7			18 @ -2.7		
Maximums	Value	Γ mag	Γ pha	Value	Γ mag	Γ pha	Value	Γ mag	Γ pha
PAE (%)	44.2	0.54	104.1	36.3	0.70	118.0	34.8	0.70	118.6
Pout (dBm)	27.5	0.43	108.1	26.7	0.48	118.1	26.5	0.57	116.0

Table 4-2, Optimum Load Impedances for key performance characteristics over the range 4-18 GHz, heavily compressed.

T10x75_35 at 9v 150 mA - Measured									
Freq. (GHz) & Compression (dB)	4 @ -2.2			6 @ -1.6			8 @ -2.2		
Maximums	Value	Γ mag	Γ pha	Value	Γ mag	Γ pha	Value	Γ mag	Γ pha
PAE (%)	48.8	0.26	44.6	45.5	0.33	99.5	49.0	0.42	78.8
Pout (dBm)	27.5	0.21	65.1	27.5	0.14	113.4	27.7	0.31	95.9
Freq. (GHz) & Compression (dB)	12 @ -2.2			16 @ -2.0			18 @ -2.1		
Maximums	Value	Γ mag	Γ pha	Value	Γ mag	Γ pha	Value	Γ mag	Γ pha
PAE (%)	44.2	0.54	104.6	34.0	0.56	115.6	34.2	0.62	121.7
Pout (dBm)	27.4	0.43	108.2	26.4	0.48	119.1	26.2	0.56	115.2

Table 4-3, Optimum Load Impedances for key performance characteristics over the range 4-18 GHz, ~2dB compressed.

For the measured results the compression levels were as calculated. For the simulation the data is recorded at 2dBs compressed and the peak PAE level. One feature that would either require an 'intelligent' power sweep algorithm, or the use a finer power increment (and hence many more data points), would be the ability to display the measurements at a specific compression level.

T10x75_35 at 9v 150 mA - Simulated									
Freq. (GHz) & Compression (dB)	4 @ -4.0			6 @ -3.6			8 @ -4.5		
Maximums	Value	Γ mag	Γ pha	Value	Γ mag	Γ pha	Value	Γ mag	Γ pha
PAE (%)	59.3	0.22	71.2	58.4	0.31	84.6	57.2	0.39	95.2
Pout (dBm)	29.1	0.11	111.5	29.4	0.17	119.1	29.3	0.24	116.3
Freq. (GHz) & Compression (dB)	12 @ -2.7			16 @ -1.8			18 @ -1.4		
Maximums	Value	Γ mag	Γ pha	Value	Γ mag	Γ pha	Value	Γ mag	Γ pha
PAE (%)	53.5	0.48	117.5	49.7	0.58	131.3	45.7	0.58	131.5
Pout (dBm)	29.0	0.38	125.2	28.8	0.51	134.5	28.5	0.43	121.3

Table 4-4, Optimum Load Impedances for key performance characteristics over the range 4-18 GHz, at peak PAE drive.

T10x75_35 at 9v 150 mA - Simulated									
Freq. (GHz) & Compression (dB)	4 @ -2.0			6 @ -2.0			8 @ -2.0		
Maximums	Value	Γ mag	Γ pha	Value	Γ mag	Γ pha	Value	Γ mag	Γ pha
PAE (%)	47.8	0.22	71.2	47.3	0.31	84.6	47.2	0.39	95.2
Pout (dBm)	28.0	0.11	111.5	28.9	0.17	119.1	28.9	0.24	116.3
Freq. (GHz) & Compression (dB)	12 @ -2.0			16 @ -2.0			18 @ -2.0		
Maximums	Value	Γ mag	Γ pha	Value	Γ mag	Γ pha	Value	Γ mag	Γ pha
PAE (%)	51.3	0.48	117.5	49.4	0.58	131.3	45.4	0.58	131.5
Pout (dBm)	28.9	0.38	125.2	28.9	0.51	134.5	28.6	0.43	121.3

Table 4-5, Optimum Load Impedances for key performance characteristics over the range 4-18 GHz, ~2dB compressed.

Plotting out these impedances (for optimum PAE and Pout Γ), Figure 4-42 and Figure 4-43:

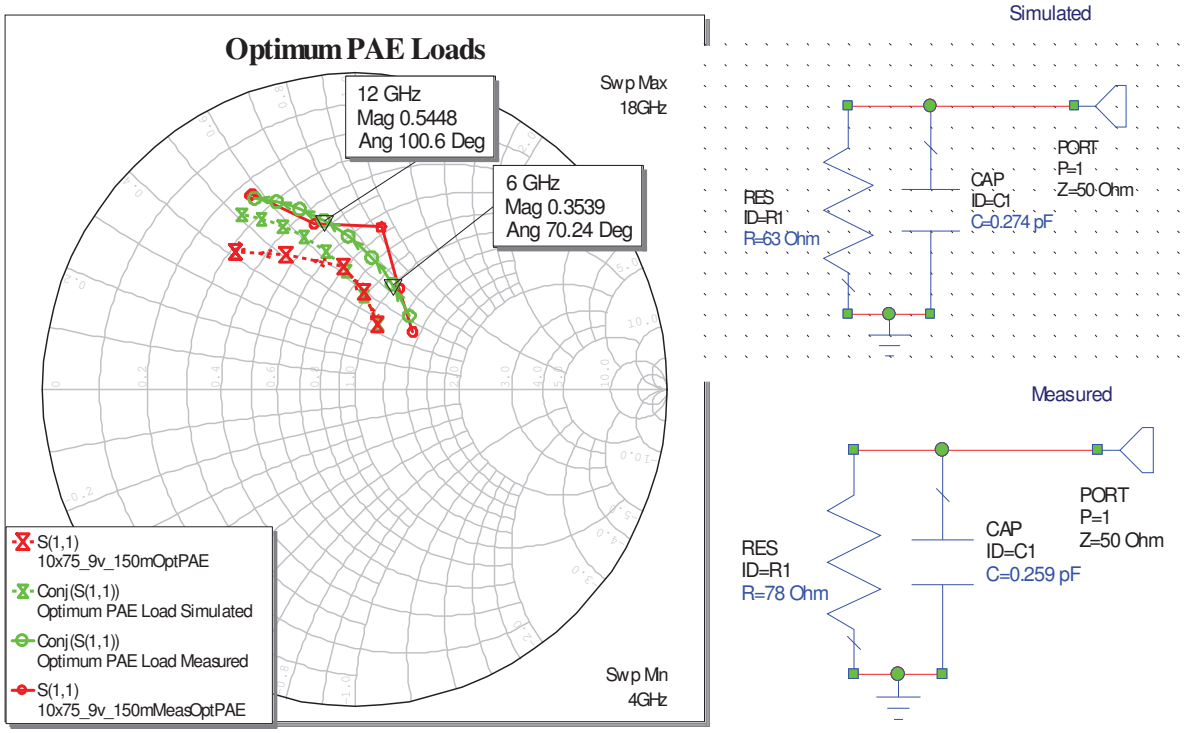


Figure 4-42, Optimum PAE loads 4-18 GHz, Measured and Simulated (Red) and Simple Model (Green)

Comparing the swept frequency load points with the equivalent circuit model of the output allows a certain amount of averaging to be done, that single frequency data doesn't

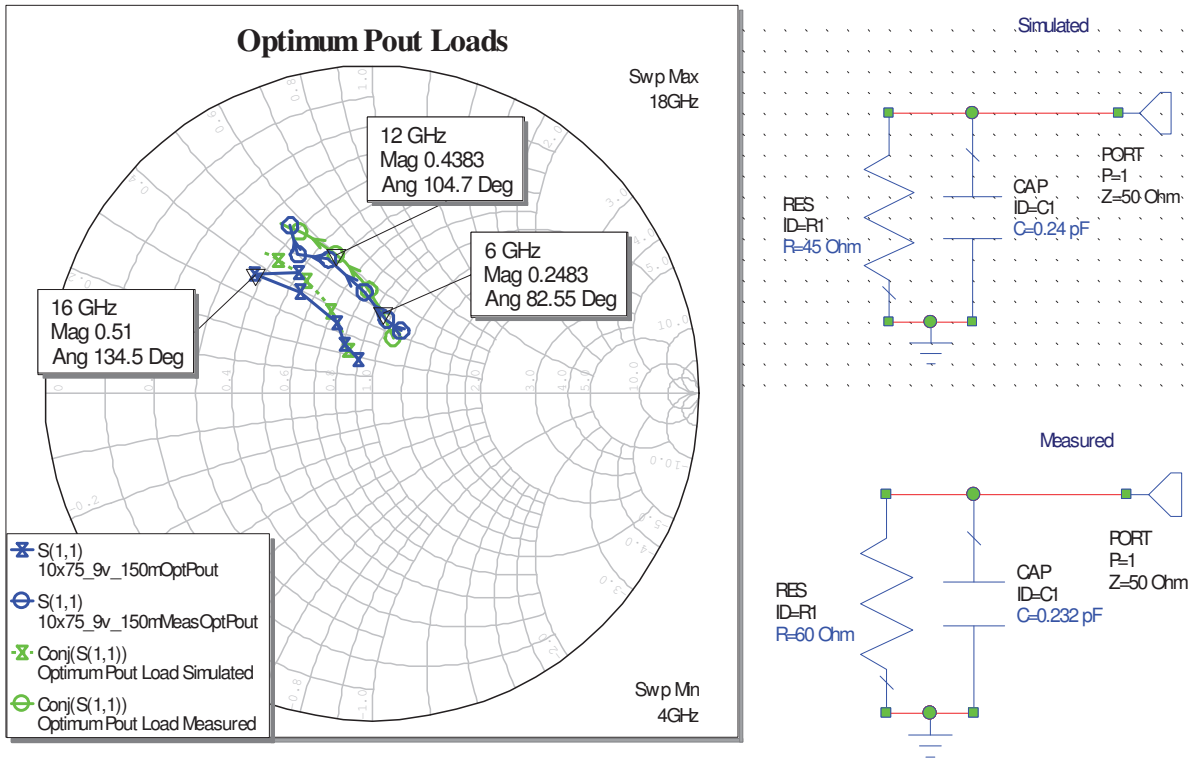


Figure 4-43, Optimum Pout loads 4-18 GHz, Measured and Simulated (Blue) and Simple Models (Green)

show. Observing the optimum PAE loads Figure 4-42, there are a number of noteworthy points:

- The PDK simulation predicts a lower drain resistance than the measurements (63:78Ω = 0.81)
- The PDK model deviates away from the simple shunt drain capacitance model, which the measurement data shows holds up to at least 16GHz.
- The drain capacitance for PAE appears from the two equivalent circuit models to lie in the range 0.26-0.27pF.
- The phase at 18GHz of the measured optimum PAE load (red with circle marker) has not changed significantly from 16GHz.

As can be seen from Figure 4-43 and the subsequent discussion the drain capacitance for optimum PAE is different from that for optimum output power. The device is driven at the same level and has the same DC bias conditions; however the different loads produced different voltage waveforms which would appear to produce different output capacitance. The relationship is however the opposite to that recorded in [6], in that higher V_{ds} corresponds to a higher C_{ds}. This is borne out in both the simulated and measured waveforms, Figure 4-44 to Figure 4-47. The ratio of voltage swing to drain capacitance between the optimum PAE and output power loads is constant for both the measured and simulated results.

Again it can be seen from Figure 4-43 that:

- The measured drain resistance is higher than the PDK model (45:60Ω = 0.75).
- The phase of the PDK model at 18GHz appears out of line with the lower frequencies and brings into question the model at the upper end of the frequency range.
- The drain capacitance for the optimum output power load appears from the simple equivalent circuit models to be in the range 0.23 to 0.24pF (the higher capacitance is difficult to predict as it depends where the 18GHz response truly lies, however increasing it further would move the lower frequency Γs further apart).
- Although less pronounced the effective drain resistance in the PDK model also appears to decrease with frequency.

Simulated results:

$$\frac{Vswing\ at\ opt.\ Pout}{Vswing\ at\ opt.PAE} = \frac{17.0 - 1.038}{18.87 - 0.705} = 0.88$$

$$\frac{Cds\ at\ opt.\ Pout}{Cds\ at\ opt.PAE} = \frac{0.24}{0.274} = 0.88$$

Measured results:

$$\frac{Vswing\ at\ opt.\ Pout}{Vswing\ at\ opt.PAE} = \frac{17.0 - 1.464}{18.248 - 0.876} = 0.89$$

$$\frac{Cds\ at\ opt.\ Pout}{Cds\ at\ opt.PAE} = \frac{0.232}{0.259} = 0.90$$

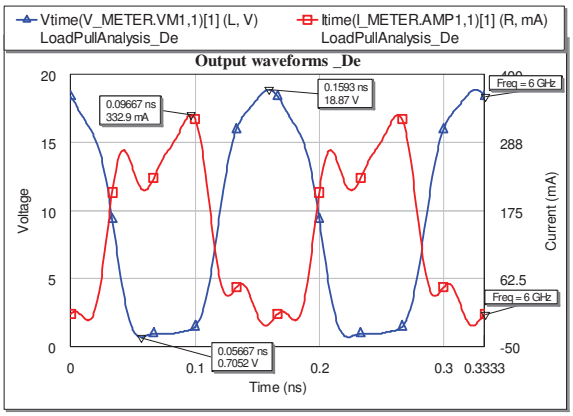


Figure 4-44, Simulated output waveforms CG plane optimum PAE load, (harmonics terminated in 50Ω)

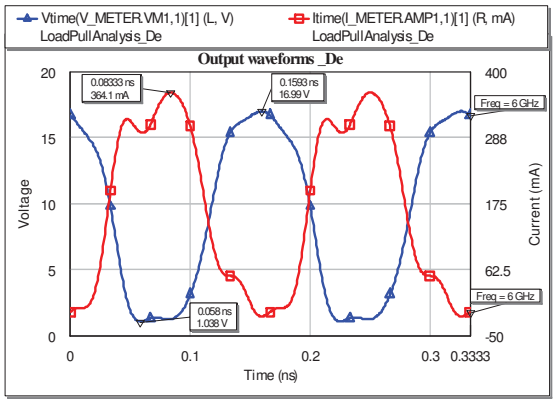


Figure 4-45, Simulated output waveforms CG plane optimum Pout load, (harmonics terminated in 50Ω)

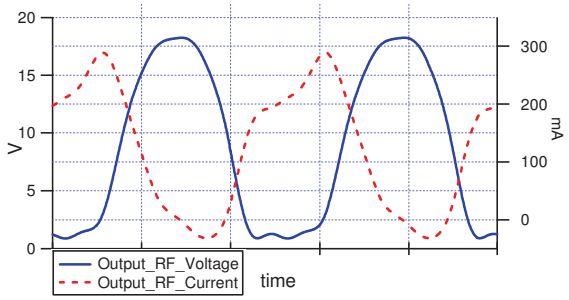


Figure 4-46, Measured output waveforms, Device plane optimum PAE load, (harmonics terminated in 50Ω)

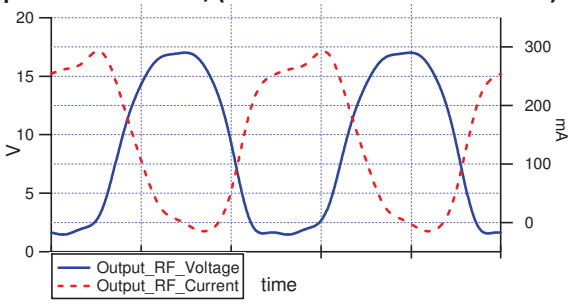


Figure 4-47, Measured output waveforms Device plane optimum PAE load, (harmonics terminated in 50Ω)

If instead of looking at the overall power and efficiency performance of the device, the output waveforms are observed, an improved understanding of the device behaviour can be obtained. However, it is necessary to move the measurement plane from the device, (which we obtained by effectively removing the feed lines), to the effective Current

Generator (CG) plane. Unlike the first de-embedding that was carried out, this is not the removal of the physical structure, but that of the effective output capacitance, which in this instance is considered as a lumped element. Note that this movement of the analysis plane doesn't affect the voltage swing but has a significant effect on the current; hence it didn't matter in Figure 4-44 to Figure 4-47 that we were comparing the voltages at two different planes. Observing the RF loadlines in Figure 4-49, the solid blue line shows the dynamic load line at the de-embedded measurement or device plane. This contains a large negative current swing due to the displacement current stored in the effective drain capacitance. If we remove this capacitance we obtain the dotted blue loadline, which if one takes the 'average' of the hysteresis, begins to approximate to the class B loadline described in chapters 1 and 2. This is an alternative way of estimating the drain capacitance, adjusting the value until the amount of negative current is minimised (as opposed to matching the measured reflection coefficients as in Figure 4-42 and Figure 4-43), or repeating the load pull with different amounts of de-embedding capacitance, until the optimum impedance lies on the resistive axis. As can be seen from Figure 4-48 the simulation follows a similar pattern to the measured data, the greater current swing is to be expected as Table 4-1 records the simulation predicts significantly (1.4dB) more power than the measured results.

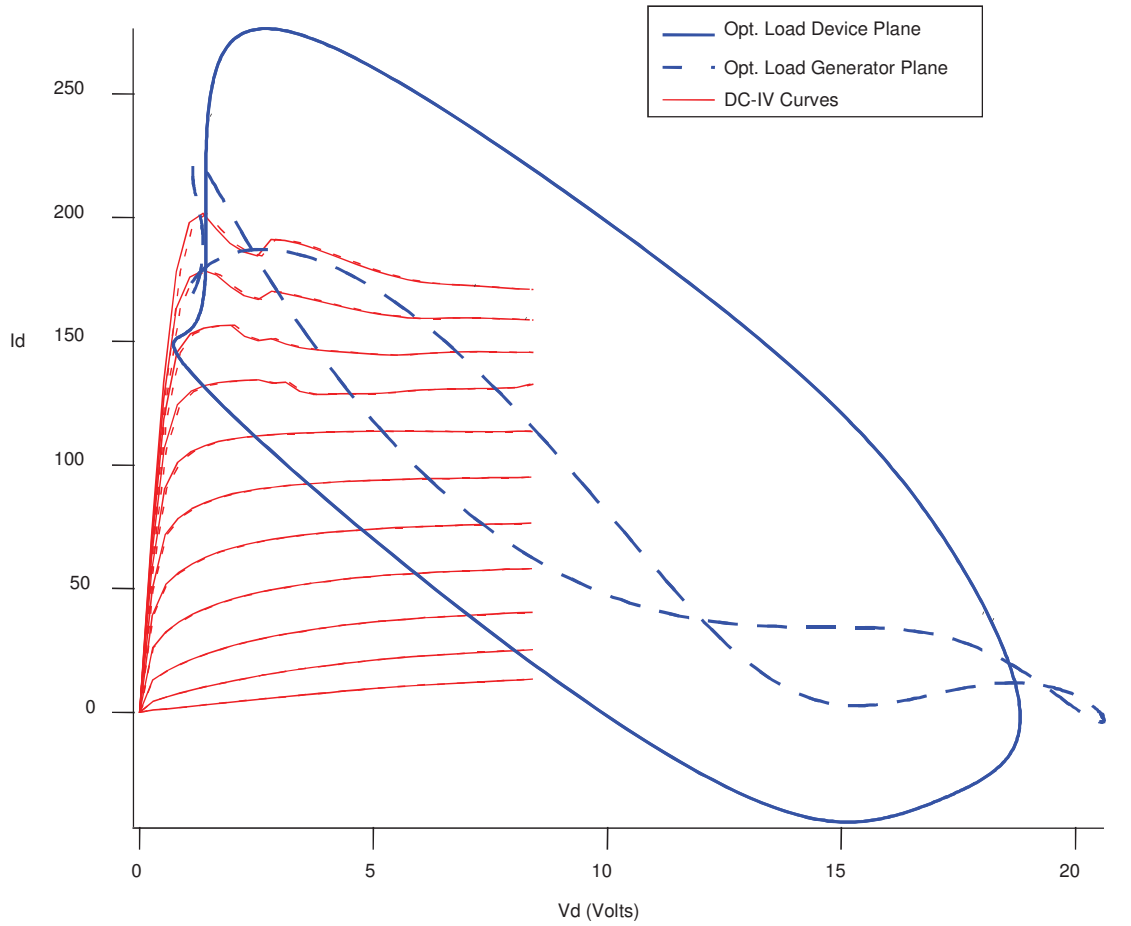


Figure 4-49, Measured DC-IV and RF Load lines at 6 GHz with fundamental load 0.35/_65.9°

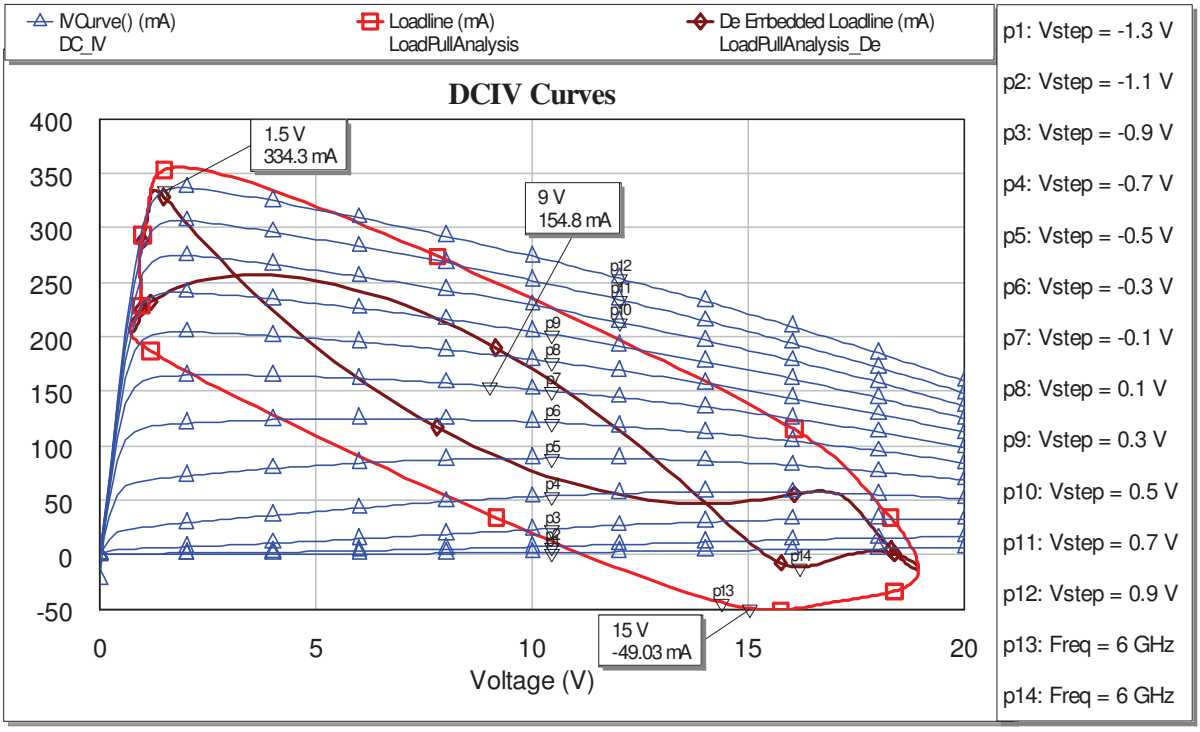


Figure 4-48, Simulated DC-IV and RF Loadline at CG plane at 6 GHz with fundamental load 0.31/_84.6°

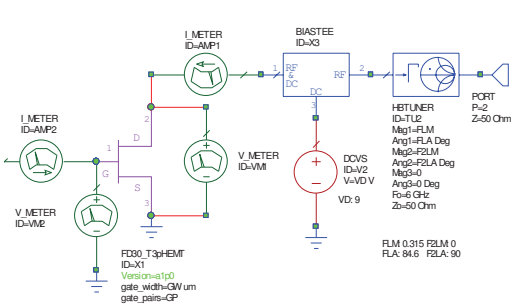


Figure 4-50, Load pull output wave form measurement Device Plane

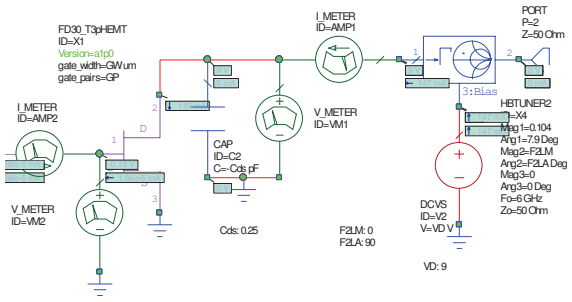


Figure 4-51, Load pull output wave form measurement CG Plane

Figure 4-50 produced the red trace in Figure 4-48, whilst Figure 4-51 produces the brown trace. The voltage and current waveforms are measured at the CG plane by adding a 'negative' capacitance before the voltage and current meters, the load presented at the fundamental is taken from the load pull conducted at the CG plane which is as one would expect, primarily resistive as shown in Figure 4-52. However at the tuner the harmonics are set to 50Ω, hence the harmonic loads presented to the device, at the CG plane, is a negative capacitive, i.e. inductive.

Interestingly, looking at the loadline corresponding to this point, we get a result which is much closer to the theoretical straight line of chapter 1 as shown by the first two loads in Table 4-6, thus to present a purely resistive load to the device (at the CG plane) the fundamental is a mainly resistive load, whilst the harmonic impedances contain a reactive component equal and opposite to the drain reactance, despite the fact that this has already been de-embedded (its effect removed). Unfortunately this effect was only noticed after the measurements were complete and thus this investigation is only made on the PDK model and could therefore be a characteristic of the model itself. Experimenting with the load impedances and observing the effect on the load line it can be seen that the 3rd harmonic appears to have the most significant effect, Table 4-6. Of particular note here is that the 2nd harmonic load has little impact on the dynamic loadline and as can be seen in the output voltage waveforms this is because the current and voltage remain in anti-phase. The impact of the harmonic load impedances will be looked at in more detail later, and using measurement data rather than the simulated – whose validity we have no knowledge of.

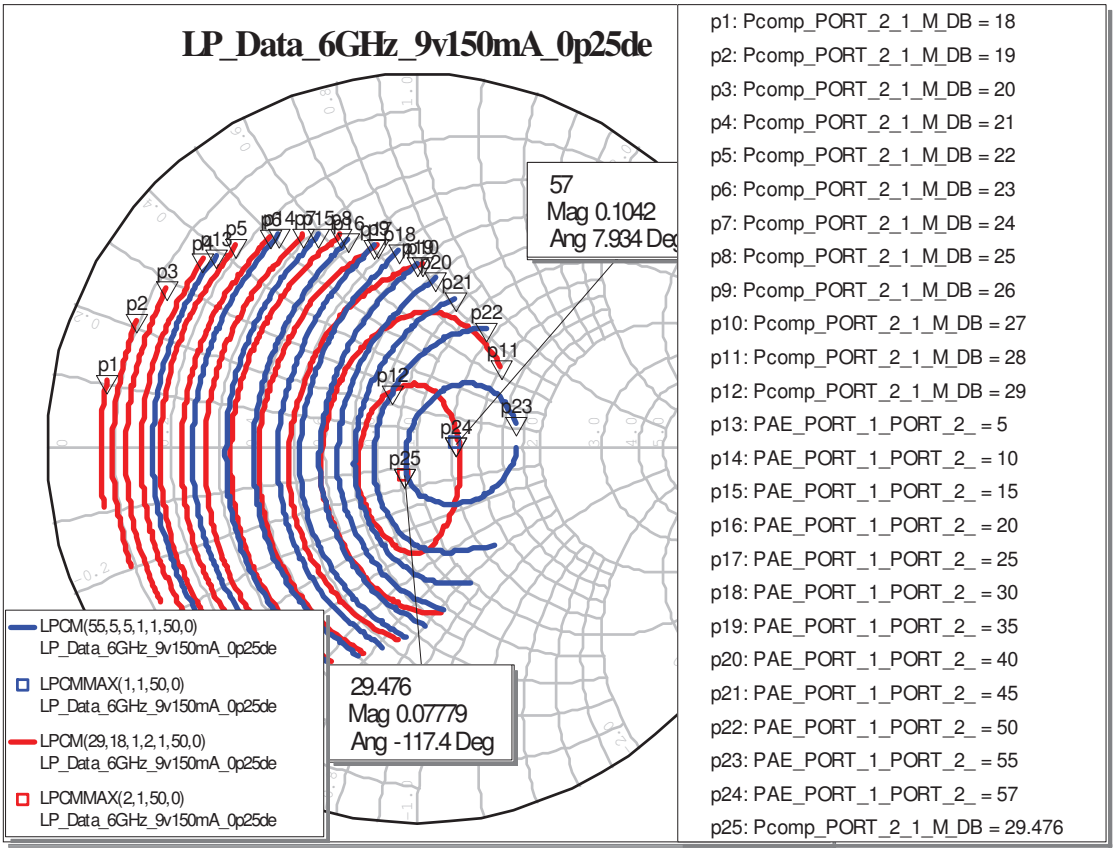


Figure 4-52, Simulated fundamental load pull conducted at the CG plane, by de-embedding output capacitance. Optimum PAE impedance de-embedding.

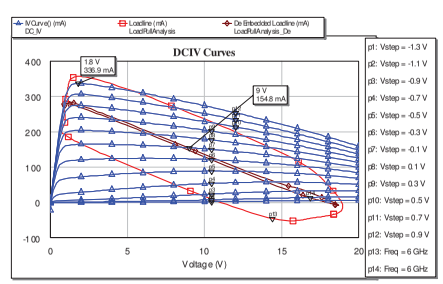
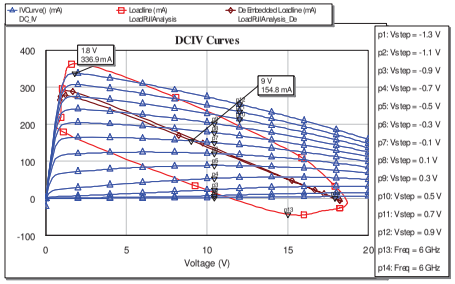
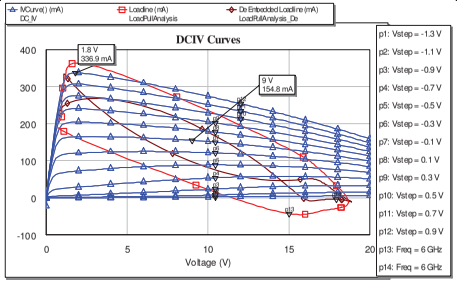
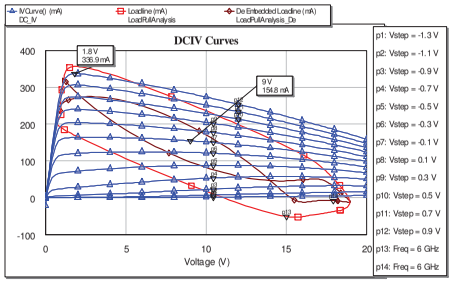
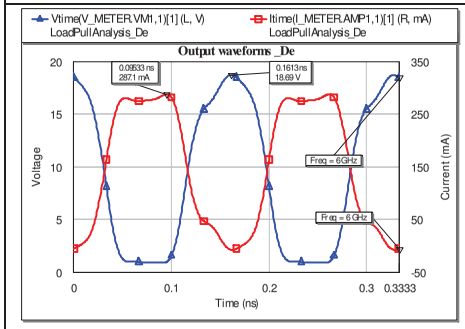
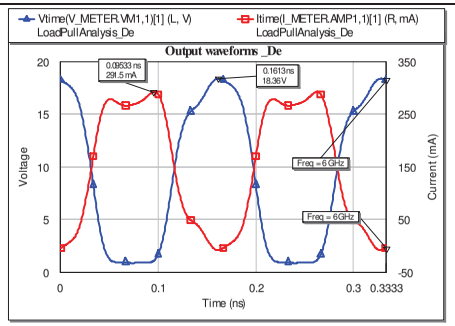
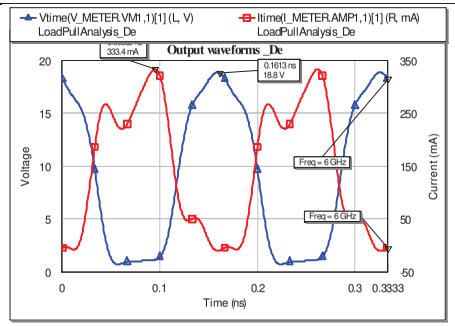
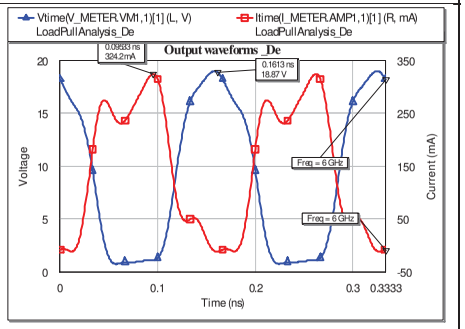
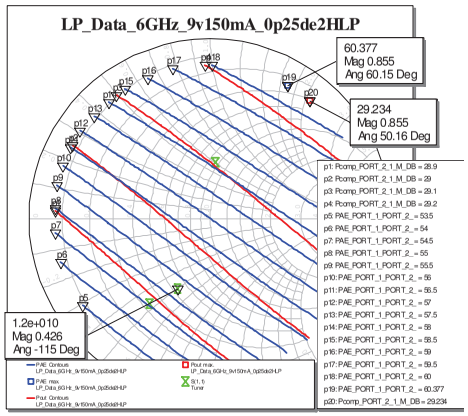
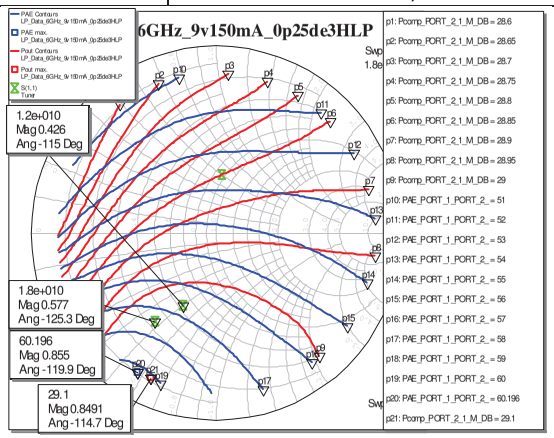
$\|\Gamma_{L2}\| = 0$ & $\|\Gamma_{L3}\| = 0$	$\Gamma_{L2} = 0.43/\_-115°$ & $\|\Gamma_{L3}\| = 0$	$\Gamma_{L2} = 0.43/\_-115°$ & $\Gamma_{L3} = 0.58/\_-125°$	$\|\Gamma_{L2}\| = 0$ & $\Gamma_{L3} = 0.58/\_-125°$
				
				
29.0dBm, 57.0%	28.9dBm, 55.3%	29.1dBm, 58.5%	29.1dBm, 59.4%	
Measurements are made at the CG plane with the Γ_{L1} terminated in the optimum PAE load, 0.10/_7.9°, from Figure 4-52. The effect of the harmonic load pull is shown (right) of the 2nd and 3rd harmonics, and as can be see the 2nd harmonic load is moving away from the optimum and the 3rd towards them, hence the improved performance is achieved with Γ_{L3} loading and Γ_{L2} = 50Ω. The deviation from the straight, resistive load line can be seen from the waveforms; the 3rd harmonic adds an asymmetry to the Voltage and Current whereas the 2nd doesn't.				

Table 4-6, Simulated effect of harmonic load on dynamic loadline, output waveforms and performance.

It should be noted that the simple equivalent models that have been used for the output of the device do not take into account all the 'parasitic' parameters, there is additional capacitance, the feedback capacitance, C_{gd} and both C_{gd} and C_{ds} are nonlinear and proportional to drain voltage [7]. In fact the relationship demonstrated over a very narrow range in these measurements has shown a proportionality of C_{ds} to V_{ds}. The next chapter is concerned with how these devices can be modelled, for the moment we are only observing the device behaviour and comparing that with the supplier's model and fundamental theory. This stage has mainly been to assess the effect of the fundamental load impedance; the next step was to consider the impact of the harmonic terminations.

4.5 Harmonic Impedance Effects

As has been shown the harmonic impedance seen by the device alters its performance. The advantages of manipulating the harmonic terminations in high efficiency amplifiers has been well documented, however much of the work has been conducted at communication bands below 4 GHz. Further, many of the applications are narrow band and thus readily addressable by harmonic manipulation. The key objective of this piece of work was to establish graphical display of harmonic impedance effects that could assist in the designing of output matching networks of wide bandwidth amplifiers. To this end a fundamental impedance at 6GHz was chosen for the initial investigation, primarily as this allowed for 3 harmonics (including fundamental) to be load-pulled (within the high power load pull amplifier bandwidth). The measurements up to this point have mainly been made with the harmonics in the region of 50Ω, largely determined by the circulator on the output.

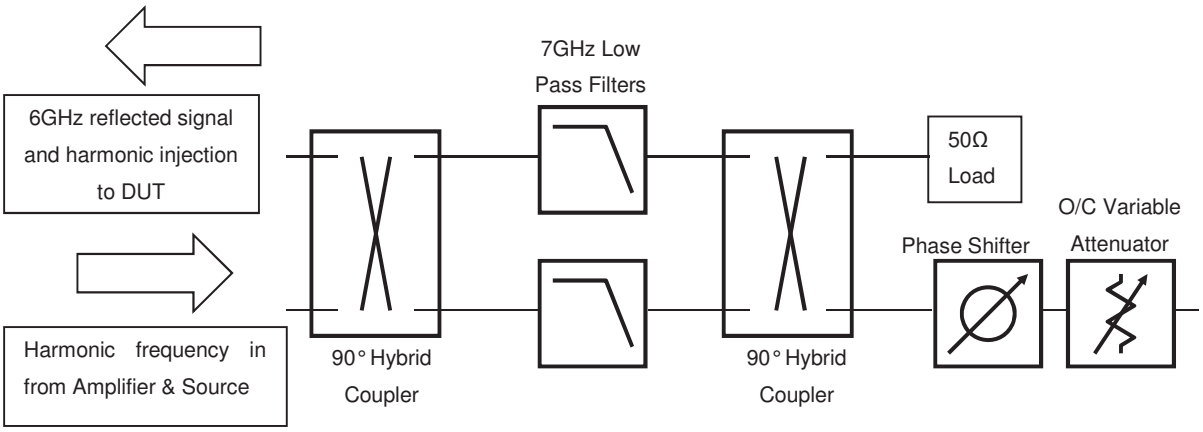


Figure 4-53, Diplexer Set-up for Passive Fundamental termination and Active Harmonic load-pull

With the diplexer connected, Figure 4-53, to the output and the fundamental passively tuned close to the optimum load impedance (Note: this could have been done actively however as the reflection coefficient was low it was easily realised passively and the measurement is much faster if only one impedance needs to be iterated), the harmonics were varied. The operation of the Triplexer is described in more detail in chapter 3, with active load pull at the fundamental as well as the harmonics.

The impedance grid can be seen to cover the majority of the real impedance plane of the Smith Chart in Figure 4-54. Note that the grid exceeds the real impedance plane at some

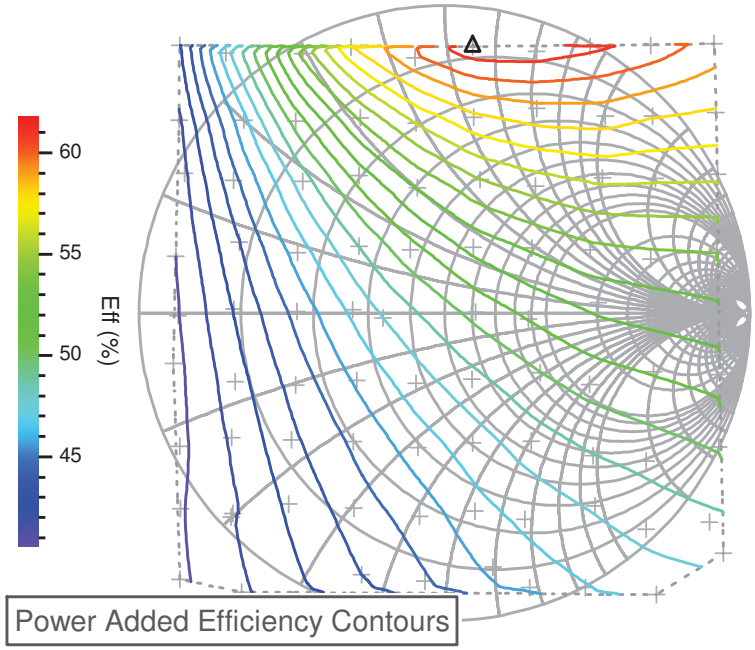


Figure 4-54, 6GHz measured 2nd harmonic load pull contours with the fundamental at 0.43/_63.2°

points, this shows the ability to produce |Γ|>1, i.e. harmonic injection. The contours of PAE at 6GHz, as a function of 2nd harmonic load, show that the efficiency varies from <41% to >61%. The results from the same measurement carried out using the PDK model show a similar trend with the value and location of the peak PAE being close, Figure 4-55, the impact of the minimum PAE is however less pronounced, ~55% as opposed to the measured <45%. It should be borne in mind however that the PDK model predicted a higher fundamental PAE (58%) thus the relative degradation across that section of the impedance plane is similar, however the enhancement across the higher efficiency portion is not as great.

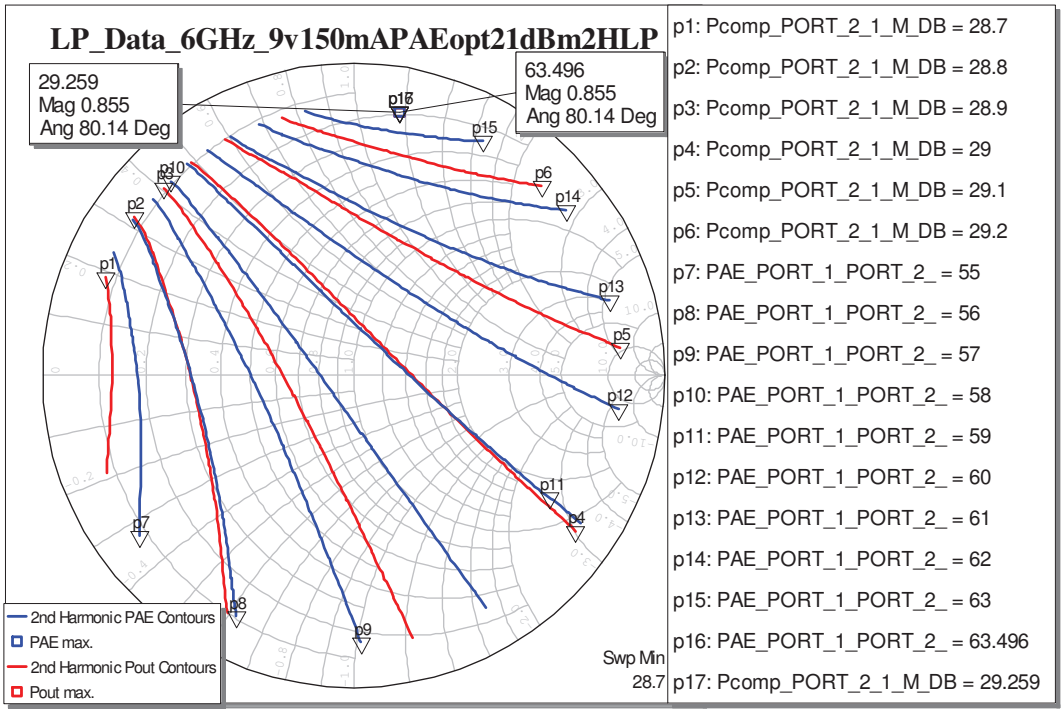


Figure 4-55, 6GHz simulated 2nd harmonic load pull contours with the fundamental load at 0.31/_84.6°

The measured 3rd harmonic load impedance behaviour is interesting from a different perspective; the fundamental is held at the previous point and the 2nd is near 50Ω (specifically a Γ_{L1} of 0.43/_63.2° and Γ_{L2} of 0.16/_143.1°). As the input power at the fundamental is increased the position of the optimum 3rd harmonic termination moves around the impedance plane. Referring back to Figure 4-40, we can see that there is a 'dip' in the 2nd harmonic level at the drive power level of Figure 4-56, which may be related to this dramatic change in the optimum 3rd harmonic load point. Figure 4-56 to Figure 4-59 show that the impact of the 3rd harmonic load impedance has a much less drastic effect than the 2nd. The total variation within the real impedance plane is of the order of 4% at the highest output power level. It is worth noting that the minimum PAE load point appears relatively constant in the top left hand corner of the Smith Chart. The change occurs as the DUT moves deeper (beyond -2dB) into the compression region and the load rotates almost 180° around the chart.

These measurements are conducted at the device plane as this is the point to which any output matching can be applied. Whilst it could be argued that for a better understanding of how and why harmonic terminations impact the device performance examination at the CG plane would be better, the objective of this part of the project was to

produce a tool that could be used to assist design engineers and prove the operation of the design approach at these frequencies.

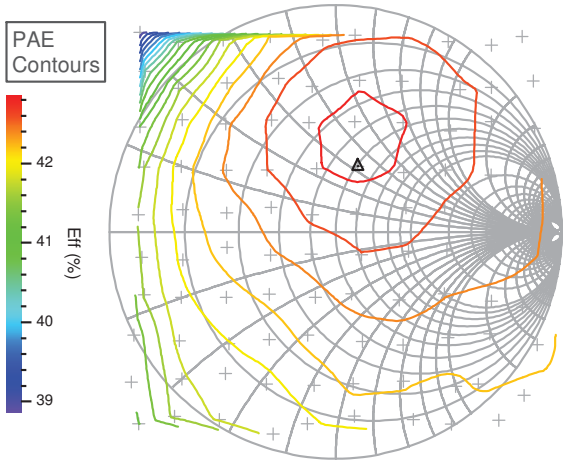


Figure 4-56, 3rd Harmonic Load pull; at optimum point (Δ) PAE 42.9%, Pout 26.6dBm, Gain 19.5dB

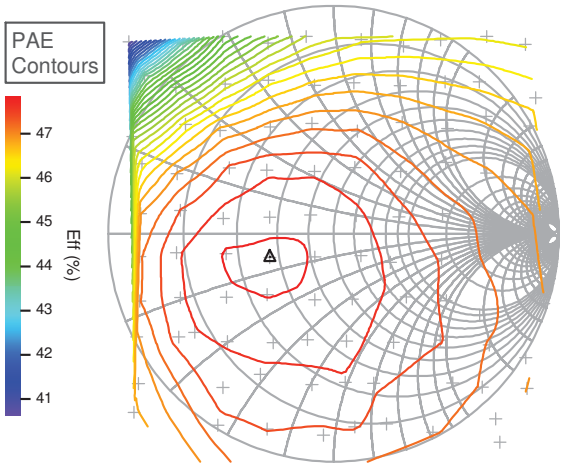


Figure 4-57, 3rd Harmonic Load pull; at optimum point (Δ) PAE 47.8%, Pout 26.8dBm, Gain 17.6dB

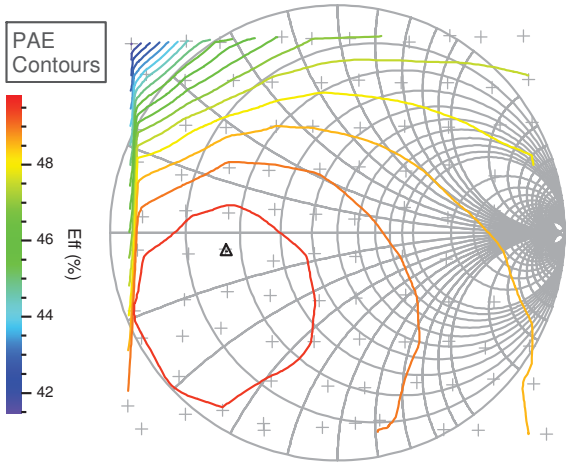


Figure 4-58, 3rd Harmonic Load pull; at optimum point (Δ) PAE 49.8%, Pout 27.0dBm, Gain 15.4dB

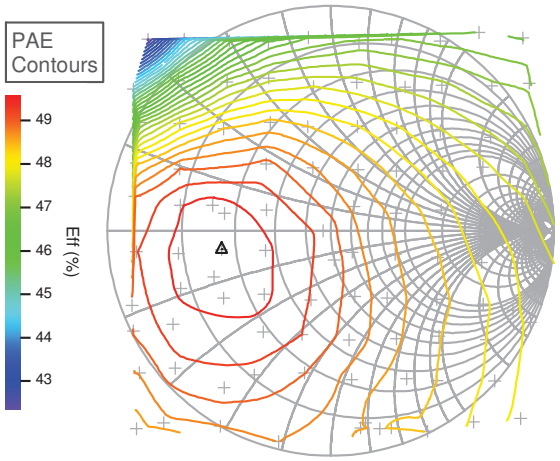


Figure 4-59, 3rd Harmonic Load pull; at optimum point (Δ) PAE 49.6%, Pout 27.2dBm, Gain 13.4dB

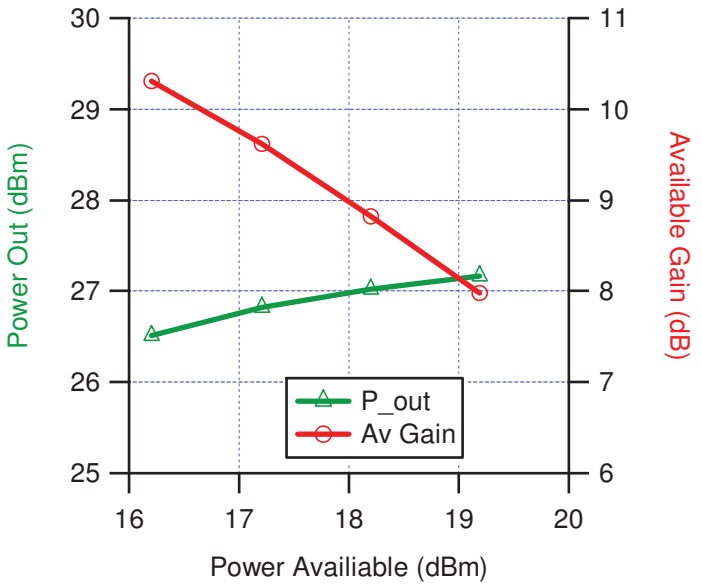


Figure 4-60, Fundamental Input Power sweep at the optimum load point of Figure 4-59.

As well as showing the regions where efficiency improvements can be obtained the plots show areas to be avoided. This is important information to a designer particularly when considering interstage matching. Typically when interstage matching is being developed, particularly over wide bandwidths, the driving parameter will be to optimise match into the load, with perhaps a frequency slope to improve gain flatness. This can be seen in Figure 4-61 which shows the load impedance presented by two different matching circuits compared with the ideal PAE load. In the conventional case (red trace) the load rotates around the centre of the Smith chart, with the top frequency having a lower $|\Gamma|$ than the bottom, hence improving gain. On the other hand with the optimum PAE load information a new matching solution (pink trace) stays in the same quadrant as the ideal load (it is difficult to exactly match the impedance trajectory of the ideal terminations over wide frequency ranges). There is also some gain equalisation as the upper frequency is closer to the centre of the impedance plane and is also significantly closer to the optimum load and hence will have improved performance (losses tend to be higher as frequency increases thus more emphasis tends to be placed on maximising power and efficiency at the top end). The rotation around the optimum PAE loads will mean that at some frequencies there will be a detrimental effect on PAE (whilst at others there may be a benefit), however as is clear from the plot the deviation from the optimum is considerably less than with the conventional approach.

Returning to the fundamental and harmonic measurements at 6 GHz, a combination of loads were created to examine the impact on performance, these range from all the

harmonics terminated in 50Ω, to the optimum harmonic terminations, and included an optimum fundamental and the worst case harmonic loads. The results are summarised in Table 4-7. The device plane dynamic loadline is shown in Figure 4-62 and the CG plane loadline in Figure 4-63.

Load	Γ_{L1}	Γ_{L2}	Γ_{L3}	PAE (%)	Pout (dBm)	Gain (dB)
1	0.09/82.6°	0.02/16.0°	0.02/58.8°	44.0	27.7	14.3
3	0.42/68.1°	0.01/171.0°	0.01/121.2°	51.9	27.1	15.3
4	0.42/68.1°	0.90/72.3°	0.00/22.5°	61.4	27.6	16.1
5	0.42/65.9°	0.90/89.3°	0.86/-132°	63.8	27.7	16.6
6	0.42/65.9°	0.87/-170°	0.50/-153°	44.4	26.6	14.8

Table 4-7, Summary of loads plotted in Figure 4-63 and associated performance.

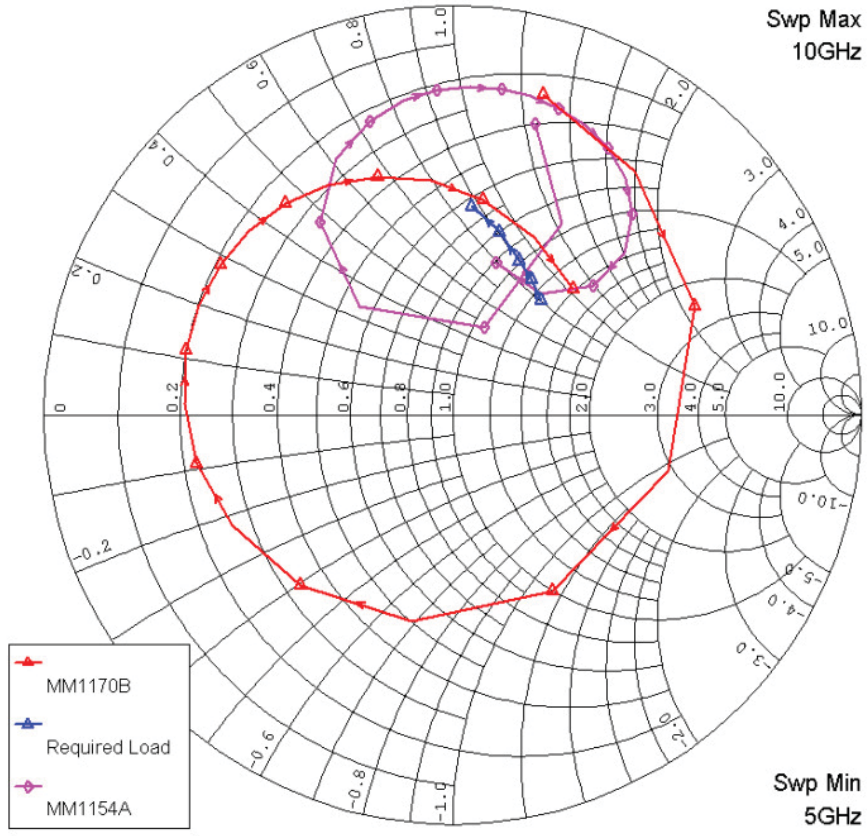


Figure 4-61, Example matching circuit solutions, blue - trace ideal PAE load, red - conventional wideband match, pink - attempt to realise optimum PAE load.

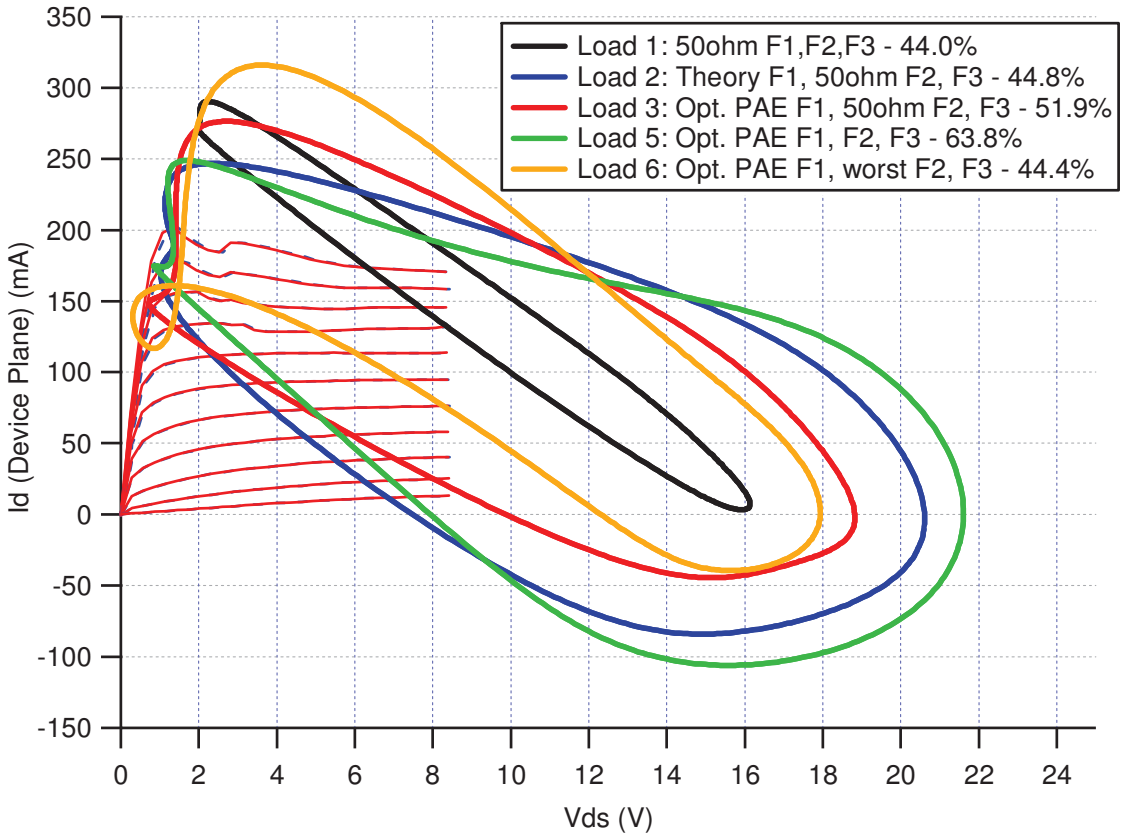


Figure 4-62, Measured RF Load lines for various combinations of fundamental, 2nd and 3rd harmonic loads.

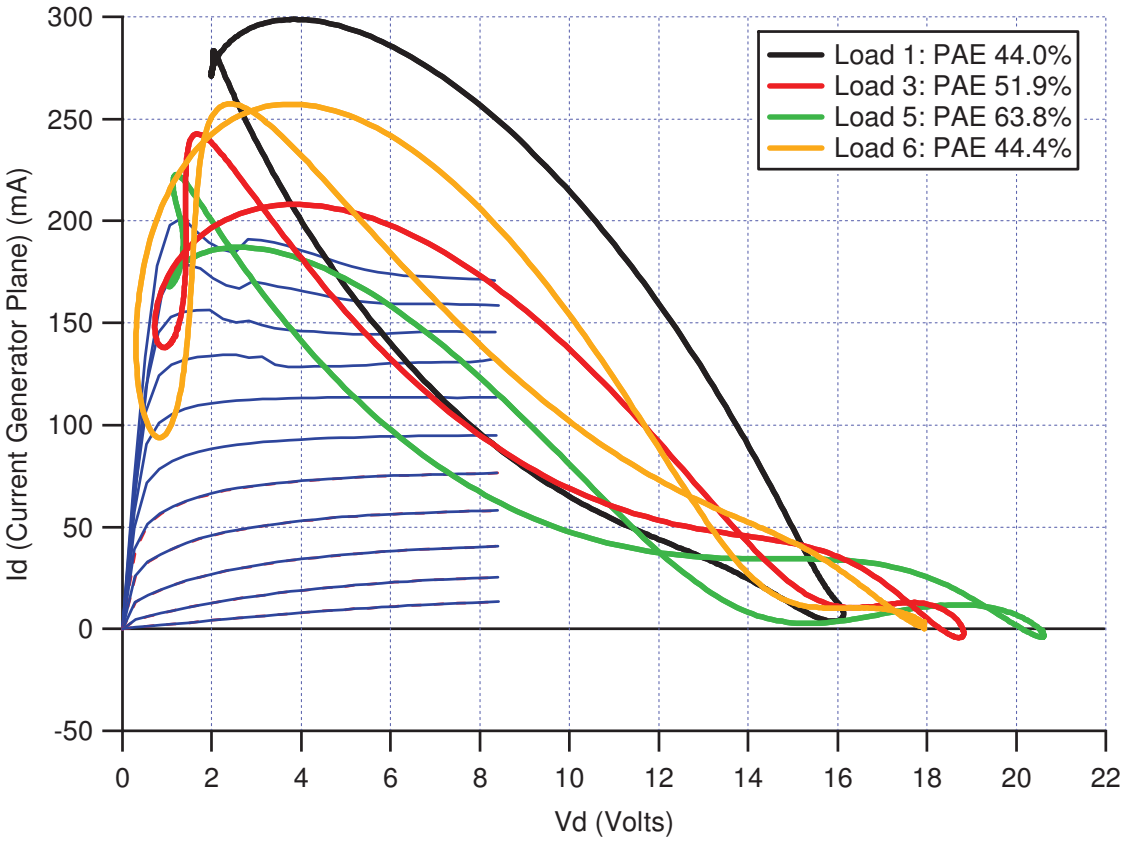


Figure 4-63, RF dynamic loadlines de-embedded to the Current Generator plane. Loads 2 & 4 removed for clarity.

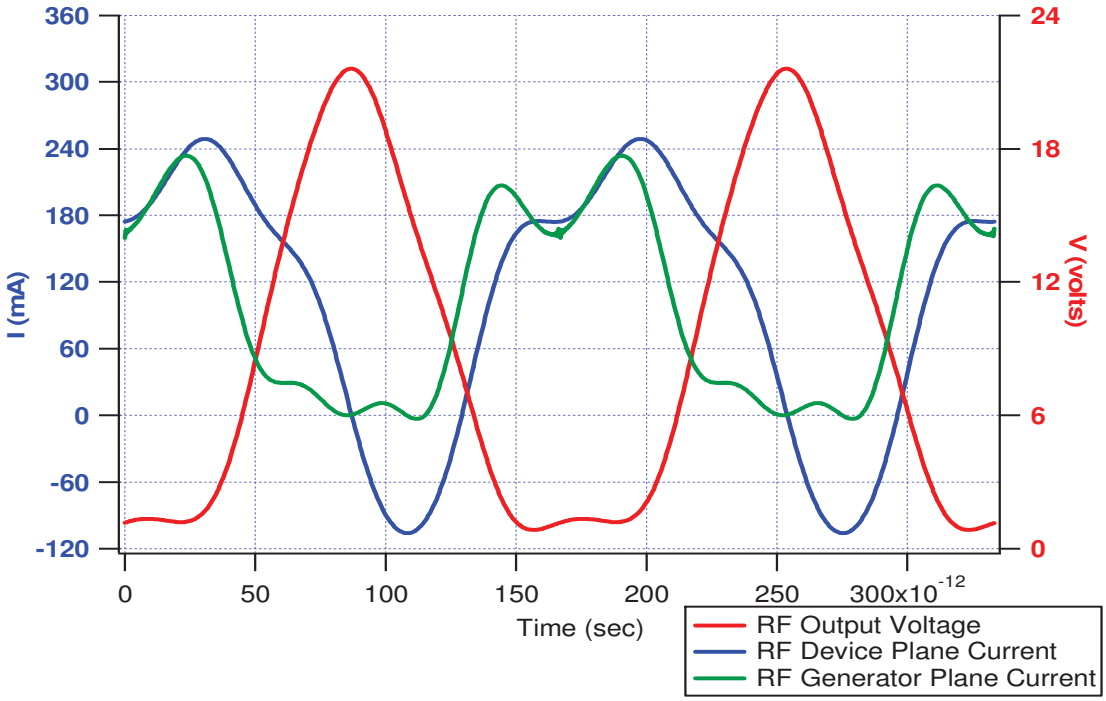


Figure 4-64, Voltage and current waveforms at the device and CG plane with the optimum loads.

Observing the dynamic loadline, Figure 4-63, it can be seen that at the optimum load:

- The drain voltage excursion is the greatest.
- The interaction with the knee region is less.
- The loadline moves 'away' from the class A bias, 45Ω nominal loadline, towards a class B bias loadline.
- The current and voltage waveforms, Figure 4-64, are close to being in anti-phase and the current waveform at the CG plane is 'squaring-up' (as much can be expected from only a few harmonic terms).

This last point can be emphasised by looking at the waveforms of the other loads measured.

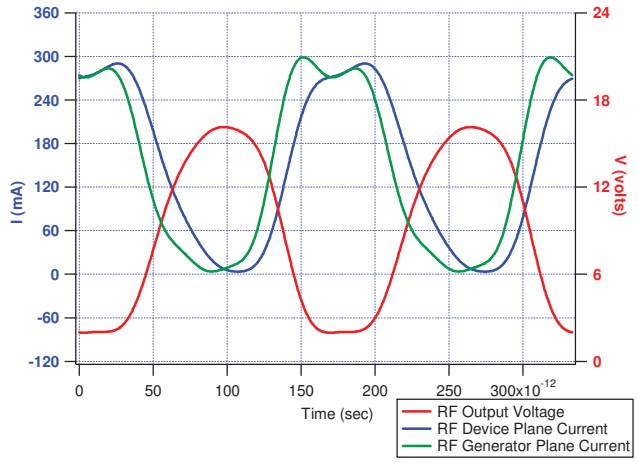


Figure 4-65, Current and Voltage Waveforms F1, F2, F3 in 50Ω Load #1 – PAE 44.0%

- Current peak, 300 mA
- Voltage peak, 16 V
- V & I waveforms are not in anti-phase resulting in poor efficiency (relative to theoretical class A maximum).

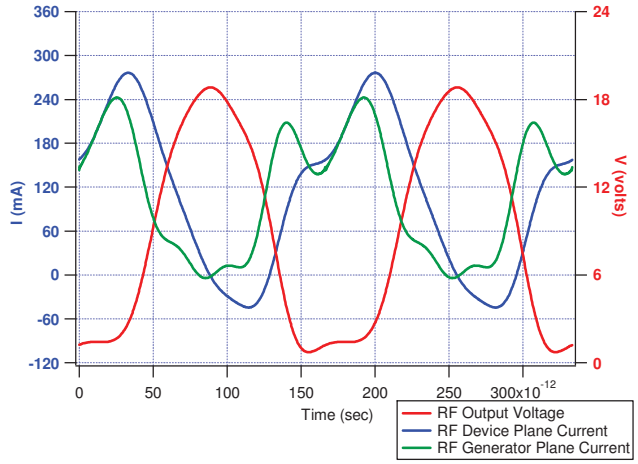


Figure 4-66, Current and Voltage Waveforms F1 Optimum, F2, F3 50Ω Load #3 – PAE 51.9%

- Current peak, 240 mA
- Voltage peak, 19 V
- V & I waveforms are closer to anti-phase resulting in an increase in efficiency.
- Voltage waveform greater excursion and significantly more 'peaked'.

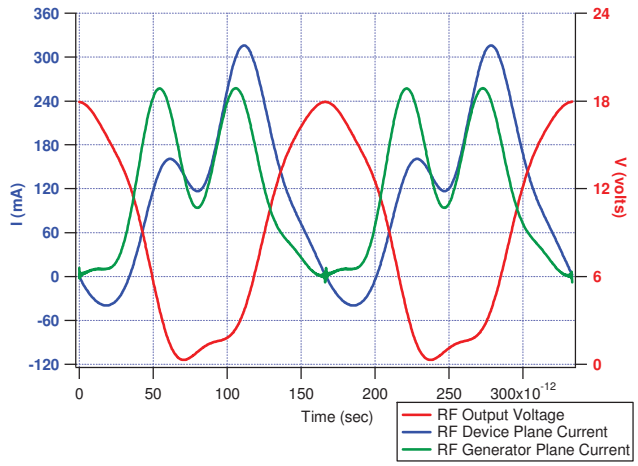


Figure 4-67, Current and Voltage Waveforms F1 Optimum, F2, F3 Worst case Load #6 – PAE 44.4%

- Current peak, 250 mA
- Voltage peak, 18 V
- V & I waveforms are closer to anti-phase however current is exceedingly 'rippled' resulting in a decrease in efficiency.
- Voltage waveform less flat at the trough.
- PAE degraded by 7.5% over 50Ω harmonics (Load 3).

A problem encountered during the 3 harmonic device measurements was the stability of the loads. The load pull signal generators were locked together using the 10MHz reference, however as described in chapter 3 the relative phase drift between the channels

made load convergence very difficult. The problem was partly alleviated by making the fundamental load passive, however simultaneously controlling both 2nd and 3rd harmonic during active load pull resulted in some cases poor load convergence. The system could be operated in phase coherent mode over limited frequency ranges (6.67-9.97GHz as described in chapter 3) which greatly improved the load convergence. Other improvements to the speed of the measurement system [8] and a move to broad band phase coherent sources have resulted in a reduction in this issue.

Three harmonic load pull was also conducted at 4GHz (at lower frequencies the load convergence was faster as the relative phase drift was less) and a similar performance to that observed at 6 GHz was seen, Figure 4-68 to Figure 4-75. These measurements were conducted with the 2nd and 3rd harmonics at a nominal 50Ω; however as can be seen from Figure 4-68 there was some drifting of these impedances. Also as can be seen from Figure 4-72 the device was not driven quite hard enough to reach the peak PAE. Figure 4-71, shows the variation in input reflection coefficient with drive level. The inflection point, 0.93/_-134° corresponds to the start of gain compression in Figure 4-75. This variation in Γ_{in} is technology dependant, the change is much less in GaN for example, and is important to consider when designing the input matching circuit to the device. For example, compression can be made much sharper (faster roll-off) by designing the input matching circuit to the low power Γ_{in} in, whereas matching to the higher power level can increase the gain expansion and improve the overall linearity (increase the power level of the 1dB compression point).

For the 2nd harmonic load pull the fundamental was held at 0.24/_40° and the 3rd at a nominal 50Ω. As was the case with the 6GHz measurements the optimum 2nd harmonic termination was at the edge of the impedance plane, 0.99/_63° (PAE of 61.2%). Note that this corresponds to the highest point on the measurement grid. By observing the shape of the load pull contours we may surmise that the actual optimum is at an angle greater than this, perhaps of the order of 85°. One would also expect the PAE to be a few % higher at this point. Note that compared to Figure 4-54, the 6 GHz 2nd harmonic load pull contours, although the optimum 2nd harmonic has moved little, the worst case PAE load point has rotated further round in an anti-clockwise direction. The input power was swept over only 3 power levels and as can be seen they continue on from those of the fundamental Figure 4-72 to Figure 4-75. The PAE can be seen to have still not peaked despite the device being 2.5dB compressed, hence these are pessimistic results in terms of what we can expect to

achieve in terms of PAE. Measuring the 3rd harmonic impedance behaviour with the fundamental at the optimum impedance and the 2nd harmonic termination near 50Ω we again see a much reduced impact on the PAE. Also the optimum load is power level dependant and so unless an amplifier was being designed for a particular output power level, the optimum load should be a compromise over the typical output power range. Compare this with the optimum 2nd harmonic as shown in Table 4-8; what variation there is more due to the approximation of the optimum at the edge of the grid.

Pin (dBm)	8.1	10.3	12.0
PAE (%)	55.7	58.8	61.2
Opt Load (Mag/_Ang)	0.88/_84°	0.86/_81°	0.99/_63°

Table 4-8, 4 GHz optimum 2nd harmonic load with increasing input power

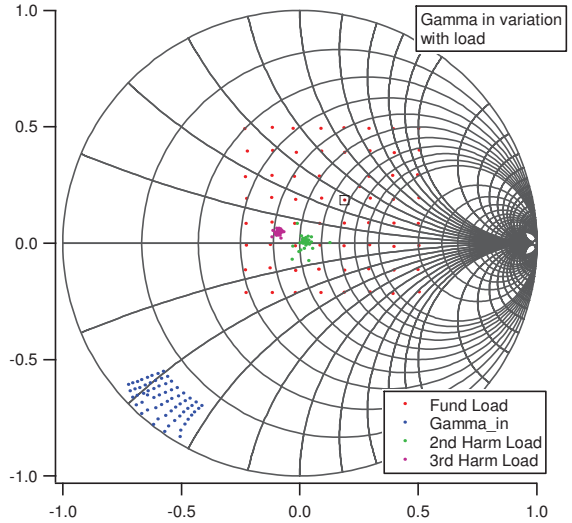


Figure 4-68, 4GHz Load pull grid and measured input and 2nd & 3rd harmonic loads achieved.

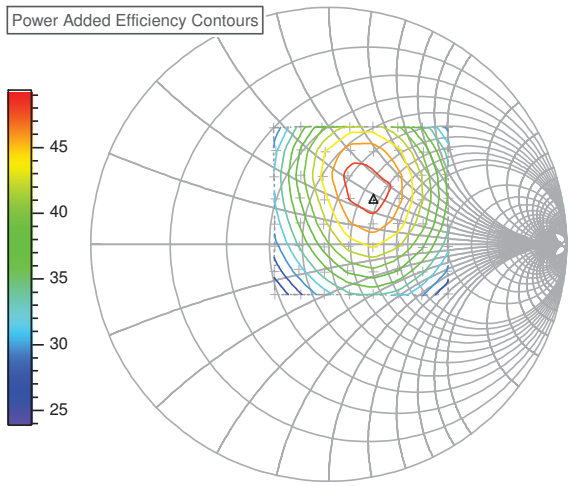


Figure 4-69, 4GHz PAE load pull contours at maximum input power

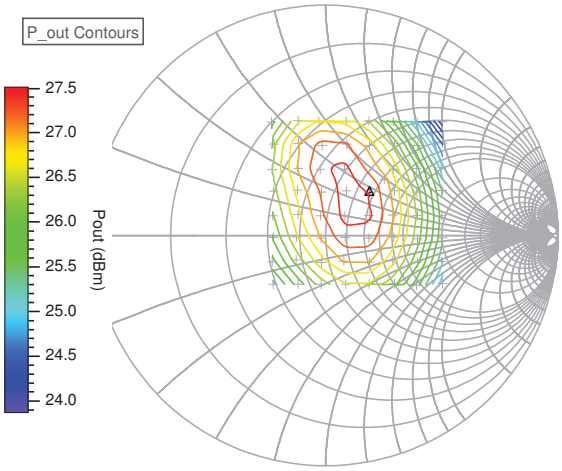


Figure 4-70, 4GHz Output power load pull contours at maximum input power

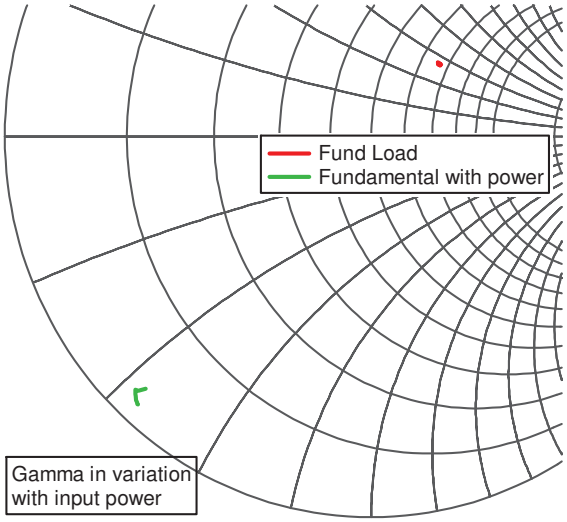


Figure 4-71, 4GHz Input reflection coefficient changes with drive level, from 0.95/_132° to 0.91 /_133°

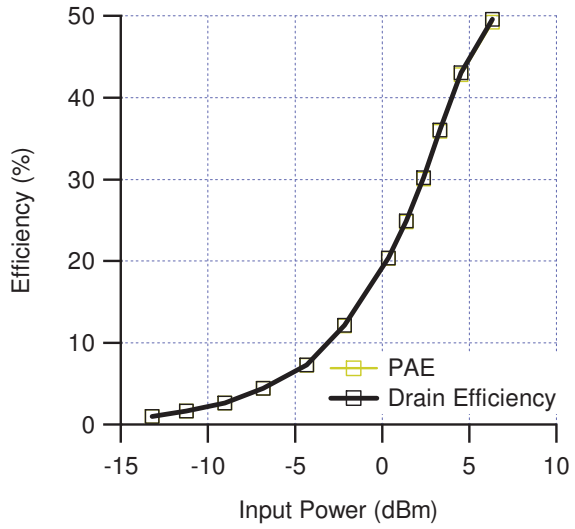


Figure 4-72, Efficiency measures at 4GHz with a fundamental load of 0.27/_45°

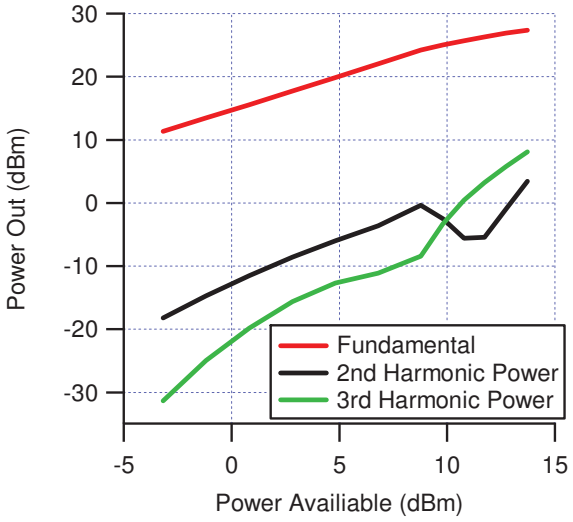


Figure 4-73, Output power at harmonics of 4GHz with a fundamental load of 0.27/_45°

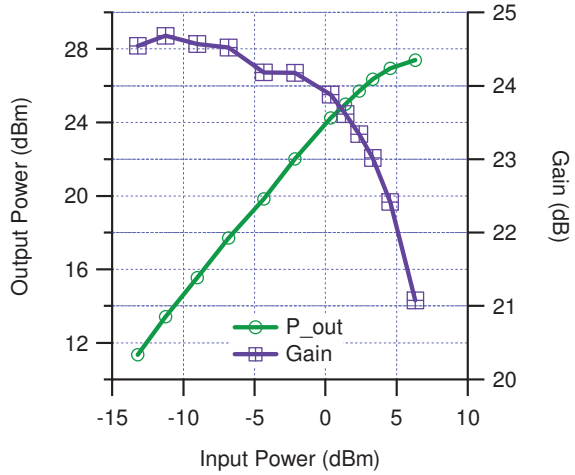


Figure 4-74, Gain and Output power against power into device at 4 GHz with a fundamental load of 0.27/_45°

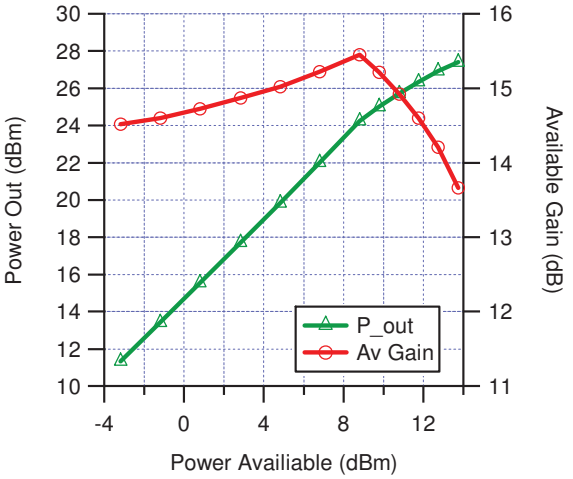


Figure 4-75, Available Gain and Output Power against available power at 4 GHz at a load of 0.27/_45°

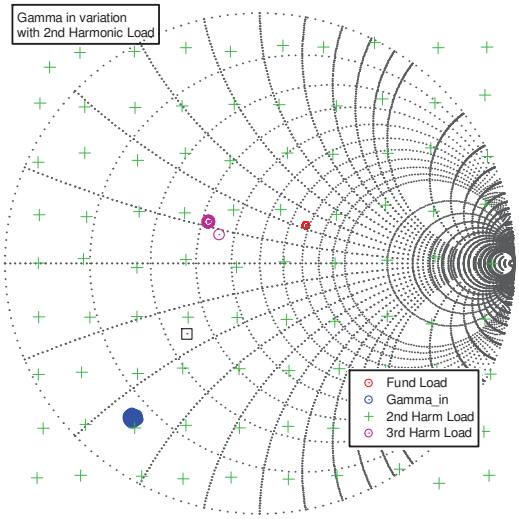


Figure 4-76, 4GHz 2nd harmonic load pull grid, fundamental input, load and 3rd harmonic load.

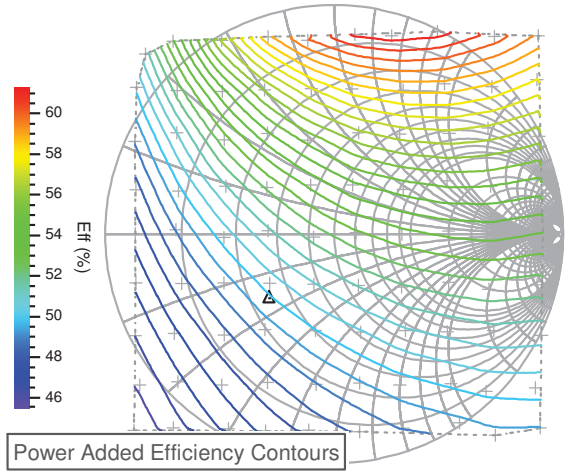


Figure 4-77, 4GHz fundamental 2nd harmonic PAE contours.

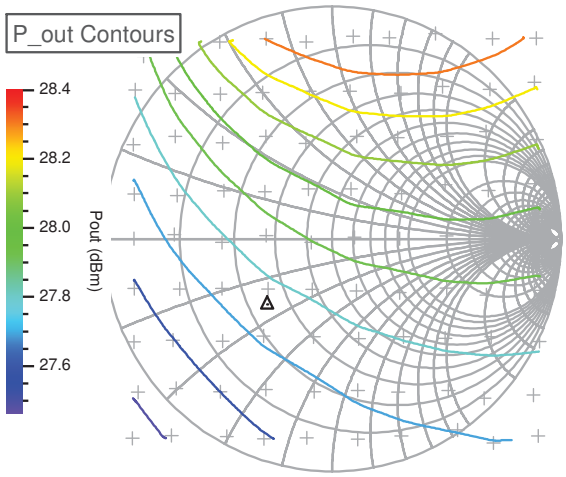


Figure 4-78, 4GHz 2nd harmonic Output power contours.

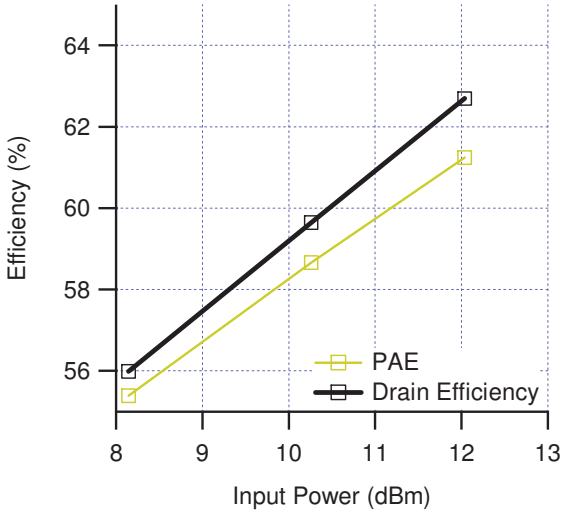


Figure 4-79, 4GHz Efficiency with swept input power at Γ_{L1} and Γ_{L2} of 0.24/_40° and 0.99/_63°.

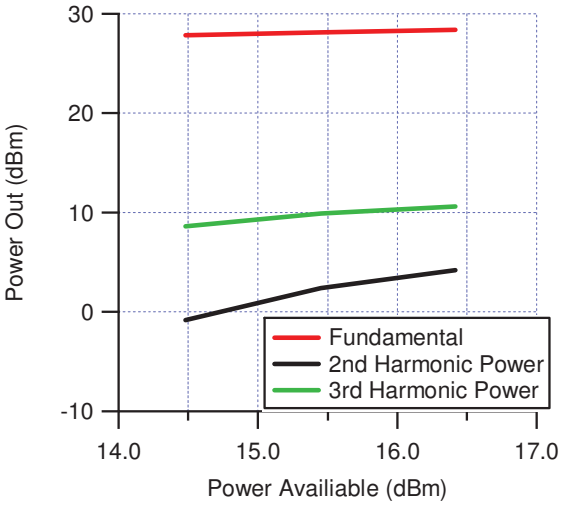


Figure 4-80, Harmonic output with swept input power at Γ_{L1} and Γ_{L2} loads of 0.24/_40° and 0.99/_63°.

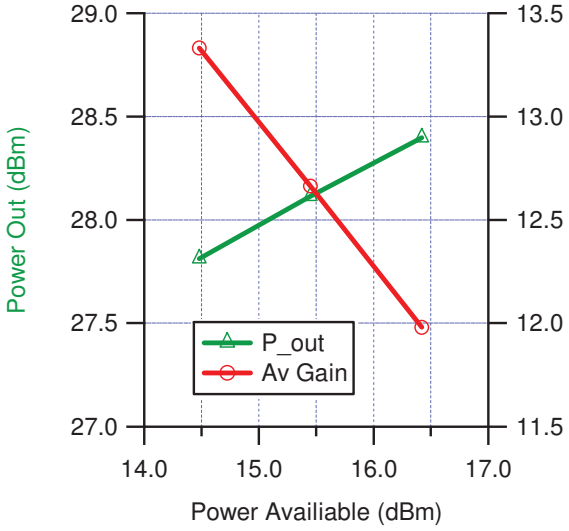


Figure 4-81, Gain & output power with swept input power at Γ_{L1} and Γ_{L2} loads of 0.24/_40° and 0.99/_63°.

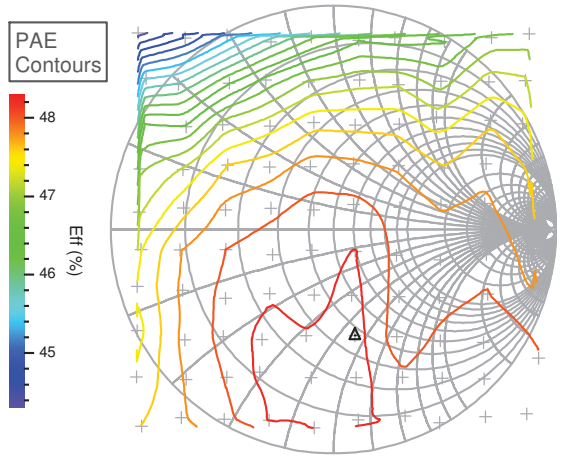


Figure 4-82, 4GHz 3rd harmonic PAE contours, maximum 48.3% 0.49/_-79°, Pin = 5.8dBm.

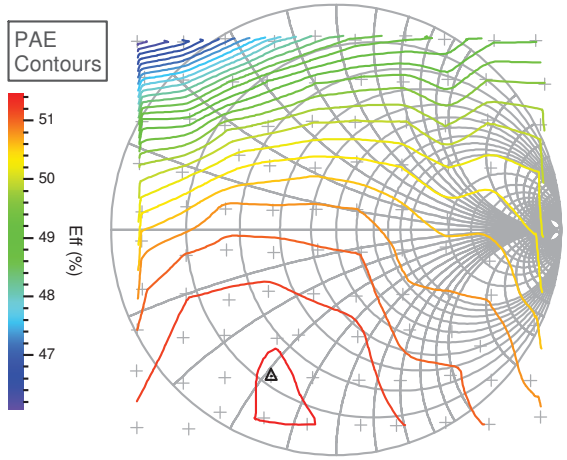


Figure 4-83, 4GHz 3rd harmonic PAE contours, maximum 51.4% 0.71/_-114°, Pin = 7.9dBm.

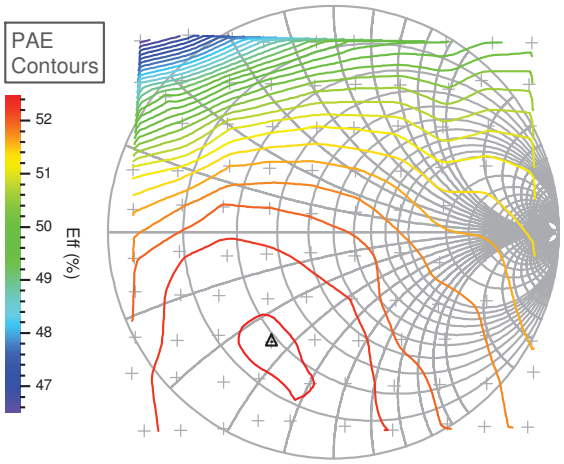


Figure 4-84, 4GHz 3rd harmonic PAE contours, maximum 52.5% 0.56/_-120°, Pin = 10.1dBm.

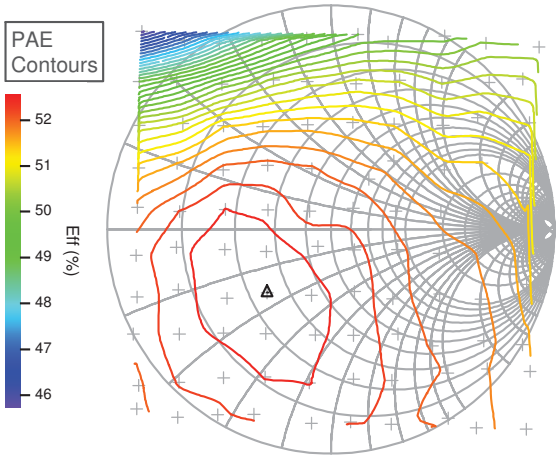


Figure 4-85, 4 GHz 3rd harmonic PAE contours, maximum 52.6 0.40/_-135°, Pin = 12.0dBm.

Having seen that the impact of the 3rd harmonic is significantly less than the 2nd on PAE and output power, further measurements concentrated on fundamental and 2nd harmonic output impedances only. This is practically more sensible as well; for the wide bandwidths of the amplifiers under consideration in this project, control of the 2nd harmonic impedance is enough of a challenge. Further, the performance of the manufactured circuit elements deviates significantly from the models above 26 GHz for the process under consideration. Finally the device process is a 0.3µm gate length and hence devices have f_ts of the order of 30GHz, thus the magnitude of any harmonics produced at frequencies approaching this limit are so small as to make little difference to the waveforms.

The device, still biased in class A (9V, 150mA), was measured at a fundamental frequency of 8 GHz. The load pull measurements are shown in Figure 4-86 to Figure 4-88, and the optimums in Table 4-9.

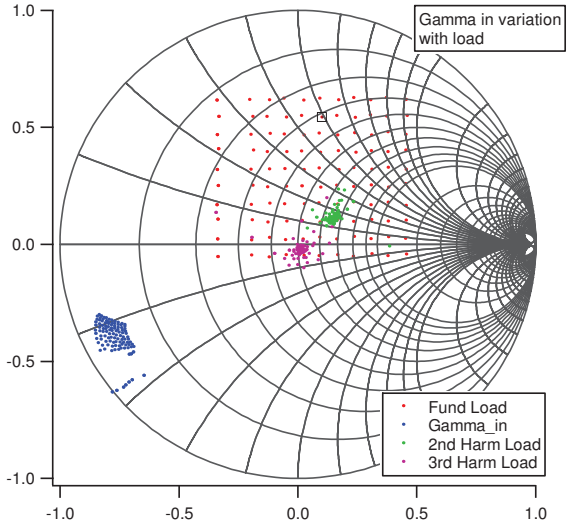


Figure 4-86, 8 GHz Fundamental load pull grid, measured input and 2nd & 3rd harmonic Γ's.

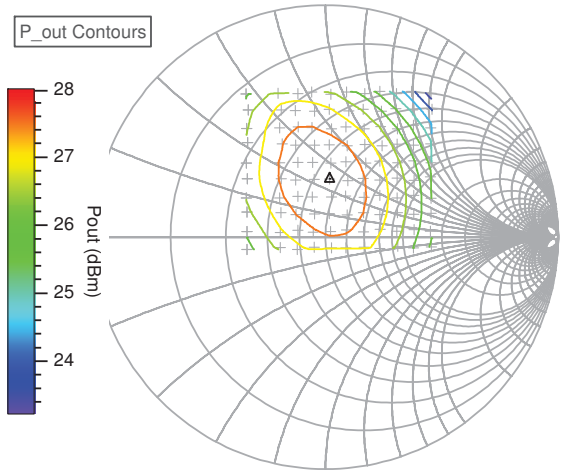


Figure 4-88, Measured 8 GHz Fundamental Output power contours

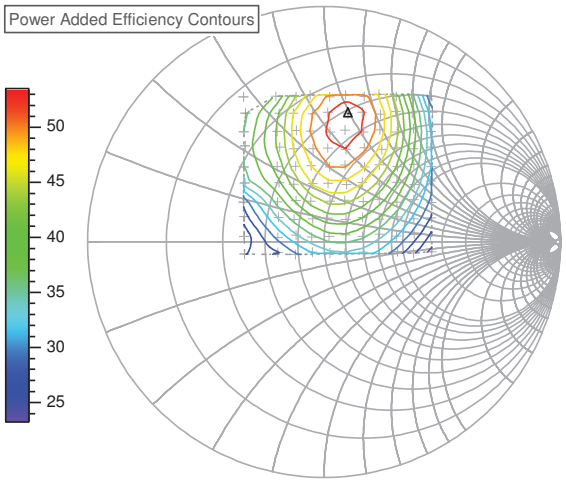


Figure 4-87, Measured 8 GHz Fundamental PAE contours.

	Load	PAE (%)	Pout (dBm)	Gain (dB)
PAE	0.55/_79.6°	53.3	26.8	14.2
Pout	0.25/_86.1°	47.6	28.0	14.2

Table 4-9, Measured 8 GHz optimum loads and associated performance.

Note the shift in the column of Γ_{in} in Figure 4-86, which corresponds to the left hand column of the load impedance grid. This may indicate an instability as referred to earlier.

Conducting a power sweep at a fundamental load of 0.47/_86.9° (a compromise between the PAE and Pout optimums), achieves a maximum PAE of 52.5% at an input power (optimum for PAE) of 13dBm and achieves an output of 27.4dBm. At this power level the 2nd harmonic is -22.6dBc. The results are shown in Figure 4-89 to Figure 4-92.

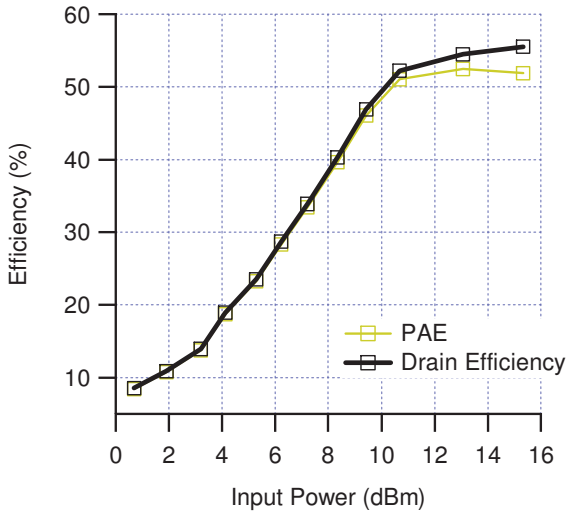


Figure 4-89, Efficiency at 8GHz with a fundamental load of 0.47/_86.9°

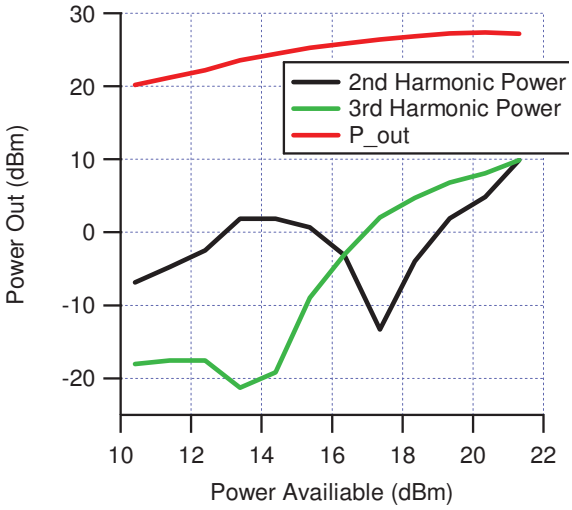


Figure 4-90, Harmonic output power with a fundamental load of 0.47/_86.9°

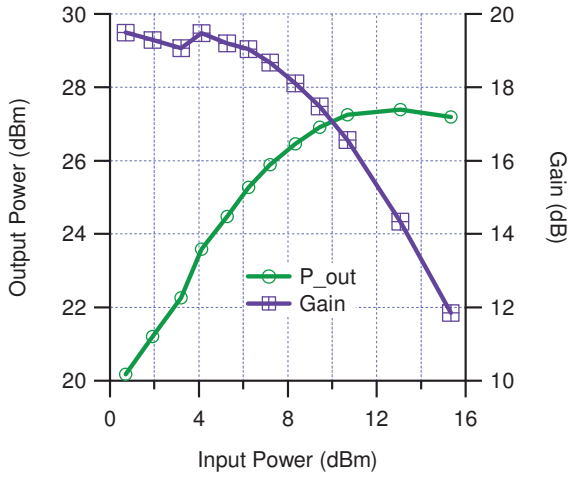


Figure 4-91, Gain and power out with a fundamental load of 0.47/_86.9°

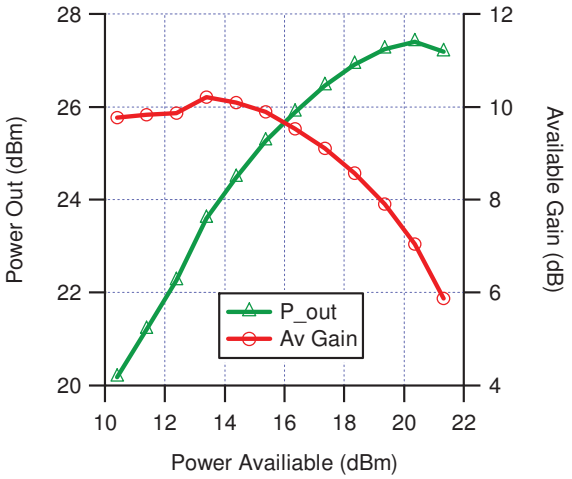


Figure 4-92, Available Gain and power out against available power with a fundamental load of 0.47/_86.9°

Having determined the device characteristics with the fundamental load whilst holding the harmonics in a nominal 50Ω; the impact of 2nd harmonic termination was measured. The fundamental impedance was passively tuned to 0.45/_86.1°, (the 3rd harmonic was consistently measured as 0.12/_65.7°) and the 2nd harmonic was stepped across the grid as set-up to cover the entire impedance plane, Figure 4-93. As can be seen the grid is less regular indicating the degree of difficulty in converging on the intended load points. Conducting a power sweep at the optimum 2nd harmonic termination shows that the optimum PAE input level has increased by at least 2dB, and that an improvement in performance is achieved as summarised in Table 4-10, and shows an increase in PAE of 6.5% and in increase in output power 28.0 dBm (mainly due to higher input drive level).

	Load	PAE (%)	Pout (dBm)	Gain (dB)
Fundamental Only PAE optimum	0.55/_79.6°	53.3	26.8	14.2
Fundamental Only PAE	0.45/_86.1°	51.5	27.6	14.6
Fundamental and 2nd Harmonic PAE optimum	0.45/_86.1° 1.0/_60°	58.0	28.0	12.8

Table 4-10, Measured 8 GHz performance with fundamrntal and harmonic loads.

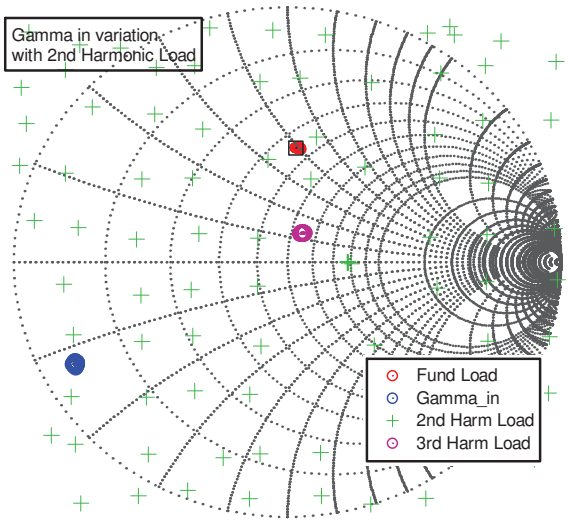


Figure 4-93, 2nd harmonic load pull grid with 8GHz fundamental and Γ_{in} & Γ_{L1} & Γ_{L3}

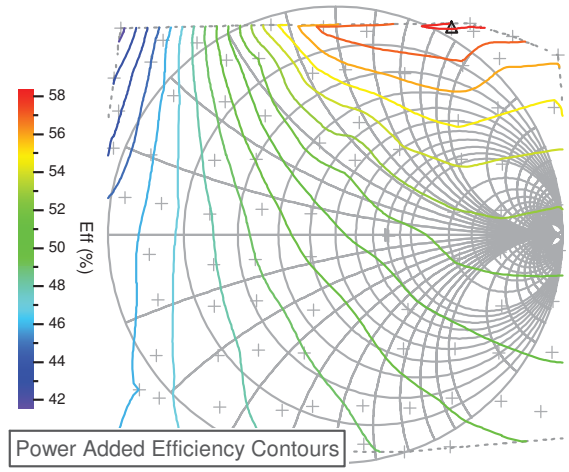


Figure 4-94, 2nd harmonic PAE contours with 8GHz fundamental load 0.45/_86.1°

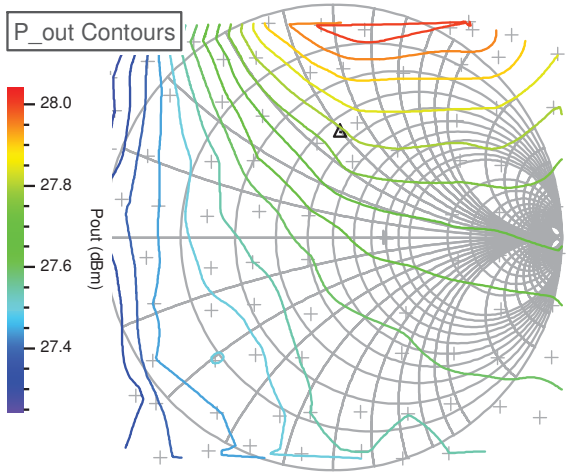


Figure 4-95, 2nd harmonic Pout contours with 8GHz fundamental load 0.45/_86.1°

Note:

- At the optimum 2nd harmonic termination of 1/_60° (best measured point) 58% PAE was achieved with 28dBm output power and 12.8 dB gain.
- When compared with Figure 4-54 and Figure 4-77 (4 & 6GHz 2nd harmonic PAE contours), the position of the optimum has not changed significantly, however the minimum has rotated clockwise round the load impedance plane, (compared to the anti-clockwise at 4GHz).

Thus we see a 7.5% improvement from the optimum 2nd harmonic termination. This may be an optimistic result as the magnitude of Γ_{L2} is 1.0, which may be impractical to achieve at 16GHz. Nonetheless it is clear that there is still both an improvement and

degradation in performance possible with selection of 2nd harmonic load. Further we can see that harmonic injection is a possible route to performance enhancement as taken up by [9], especially as the power level required at the harmonic is over 20dB below that of the fundamental, Figure 4-97. As was noted in the 4GHz measurements the position of the optimum is in the same vicinity as for 6GHz, however the minimum has rotated in this case clockwise. The opposite direction of rotation is not surprising as we are now considering a frequency above the reference (6GHz) as opposed to below in the 4 GHz case.

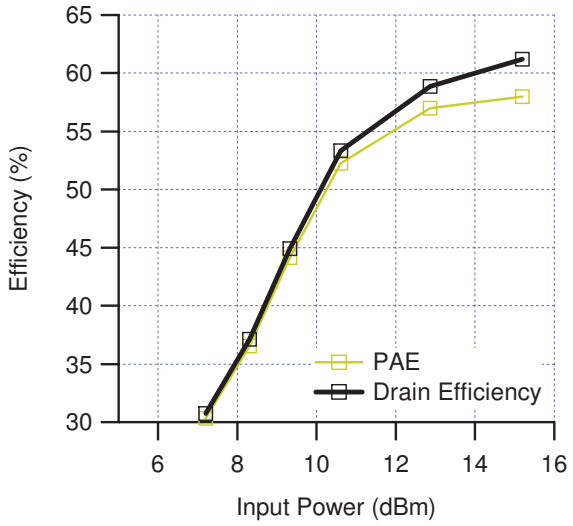


Figure 4-96, Efficiency at 8GHz with Γ_{L1} = 0.47/_86.9° and Γ_{L2} = 1.0/_60°.

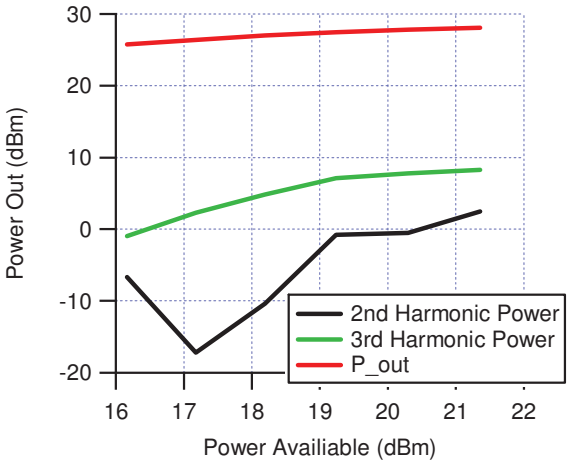


Figure 4-97, Harmonics at 8GHz with Γ_{L1} = 0.47/_86.9° and Γ_{L2} = 1.0/_60°.

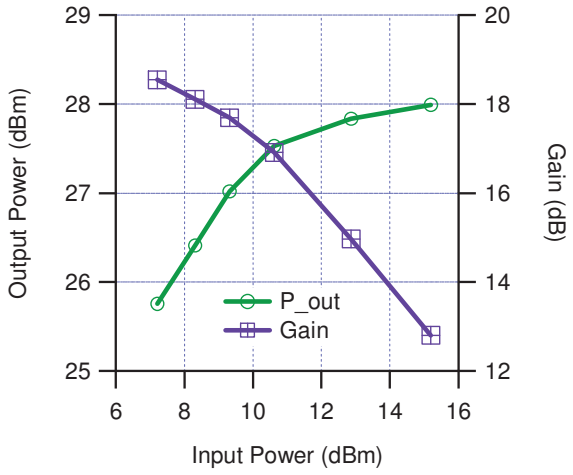


Figure 4-98, Power out and Gain at 8GHz with Γ_{L1} = 0.47/_86.9° and Γ_{L2} = 1.0/_60°.

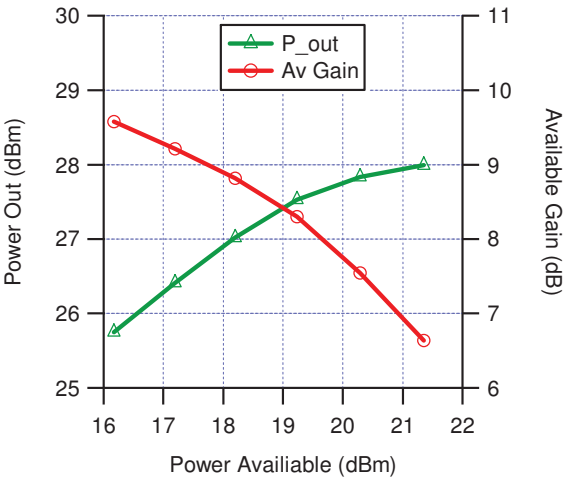


Figure 4-99, Power out and available Gain against available input power at 8GHz with Γ_{L1} = 0.47/_86.9° and Γ_{L2} = 1.0/_60°.

Device measurements at 10GHz; a fundamental load pull was conducted over the entire impedance plane. As can be seen from Figure 4-101, this confirmed that such

coverage is in fact very wasteful, with about 50% of the measurement points having PAE of ≤0%. The matching of the fundamental impedance will naturally tend to be centred in the regions of the optimum, whilst the same cannot be said of the harmonics, which without guidance could literally end up anywhere in the impedance plane.

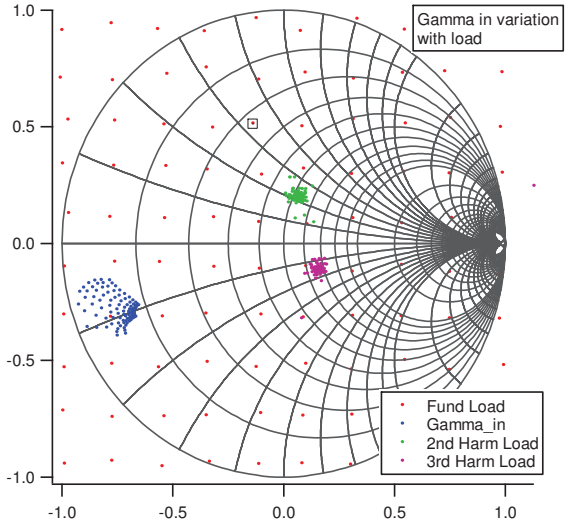


Figure 4-100, 10GHz Fundamental load pull grid, measured input and 2nd & 3rd harmonic Γ's

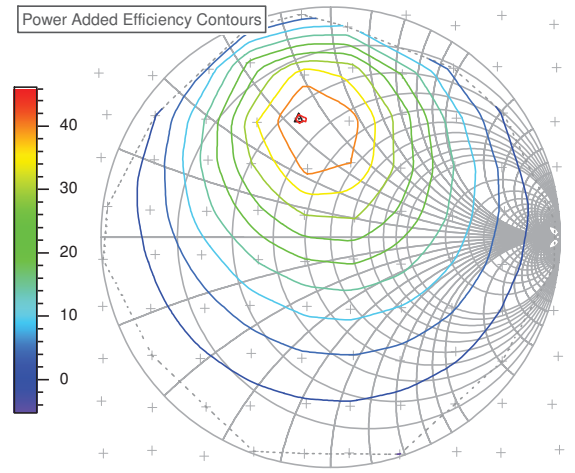


Figure 4-101, Measured 10GHz fundamental PAE contours

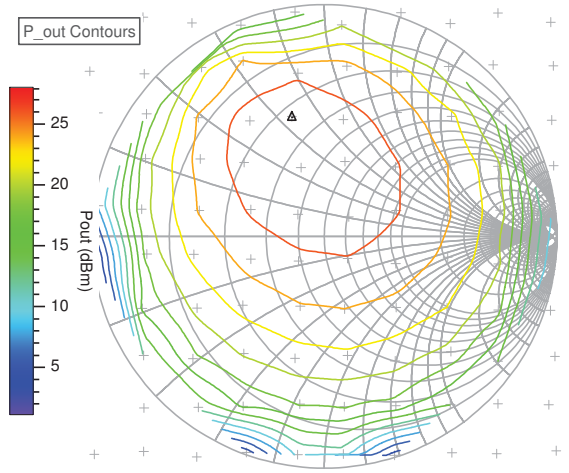


Figure 4-102, Measured 10GHz fundamental Output Power contours

	Load	PAE (%)	Pout (dBm)	Gain (dB)
PAE	0.54/_105°	45.8	27.4	12.7
Pout	0.31/_109°	41.5	27.9	11.9

Table 4-11, Measured 10GHz optimum loads and associated performance.

The power sweep was conducted at the optimum PAE fundamental load, 0.54/_105°.

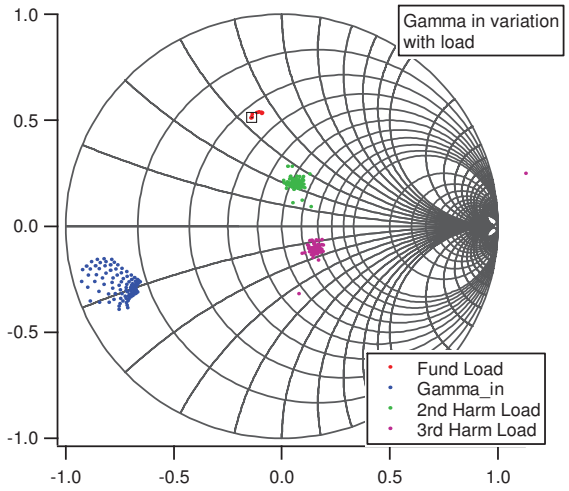


Figure 4-103 Fundamental load grid with 10 GHz fundamental and Γ_{in} & Γ_{L2} & Γ_{L3}

Note the slight movement in fundamental load during the power sweep, active load pull was utilised not passive as previously, as a passive load near enough to the optimum could not be produced with the current set-up. Hence not only does this slow down the measurement, but it also adds to measurement error.

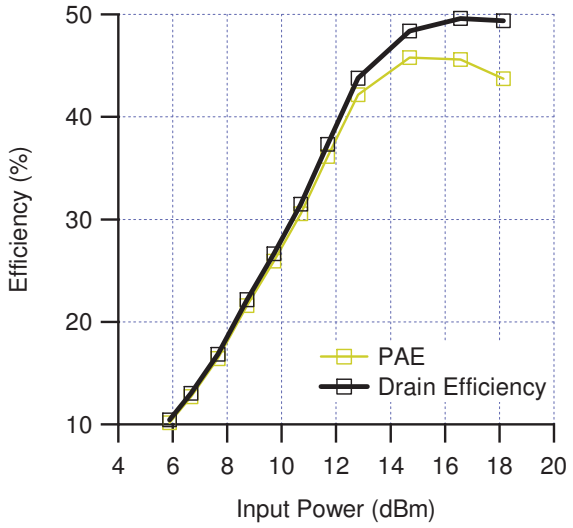


Figure 4-104, Efficiency at 10GHz with a fundamental load of 0.54/_105°

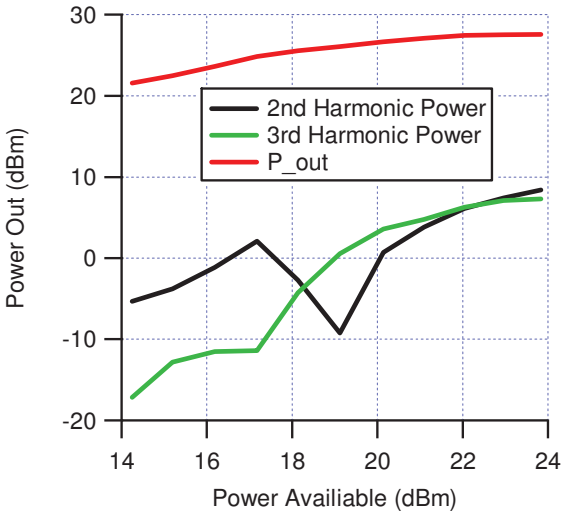


Figure 4-105, Harmonics at 10GHz with a fundamental load of 0.54/_105°

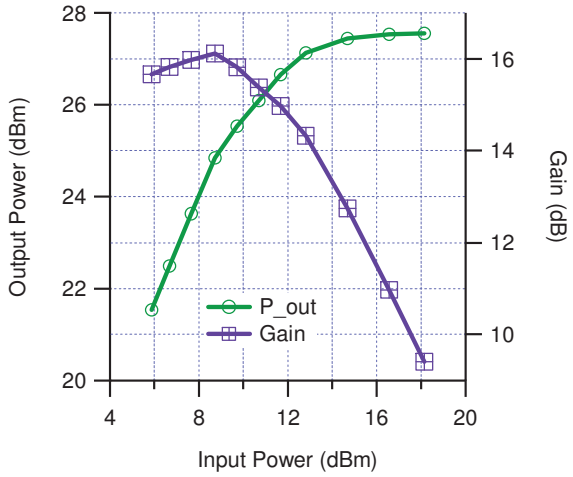


Figure 4-106, Output power and Gain at 10GHz with a fundamental load of 0.54/_105°

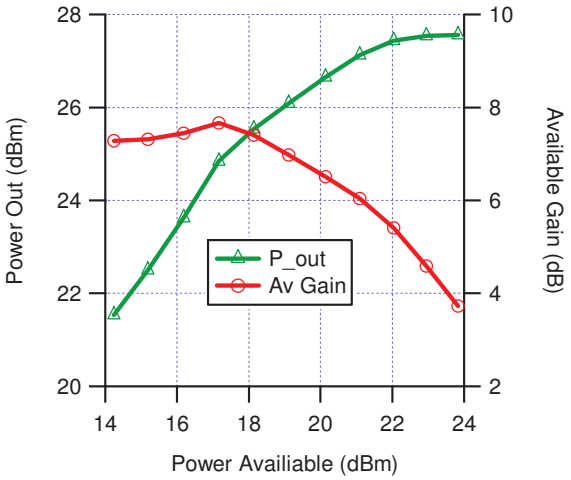


Figure 4-107, Output power and Available Gain at 10GHz with a fundamental load of 0.54/_105°

Conducting the 2nd harmonic load pull, again there is a slight variation in the fundamental load across the measurements.

	Load	PAE (%)	Pout (dBm)	Gain (dB)
Fundamental Only PAE optimum	0.54/_105°	45.8	27.4	12.7
Fundamental Only PAE[3]	0.50/_91.4°	-	-	-
Fundamental and 2nd Harmonic PAE optimum *(not driven hard enough to reach PAE max. hence lower PAE and Pout and higher gain than expected)*.	0.50/_91.4° 0.92/_53.6°	49.5	27.3	13.6

Table 4-12, Measured 10GHz performance with fundamrntal and 2nd harmonic loads.

[3] No results at this fundamental load as the grid was too widely spaced.

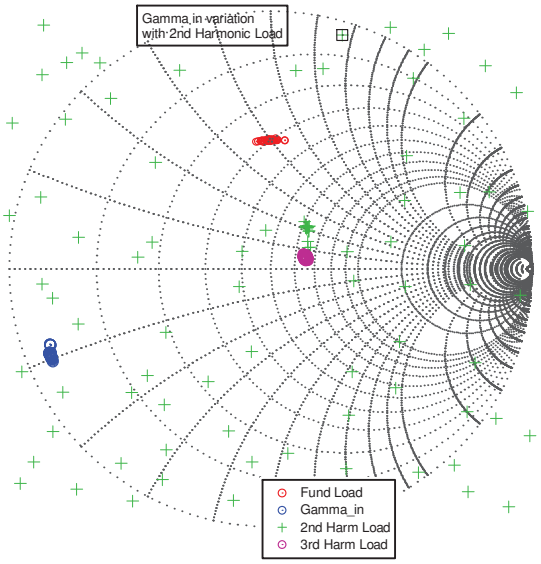


Figure 4-108 2nd harmonic load pull grid with 10GHz fundamental and Γ_{in} & Γ_{L1} & Γ_{L3}

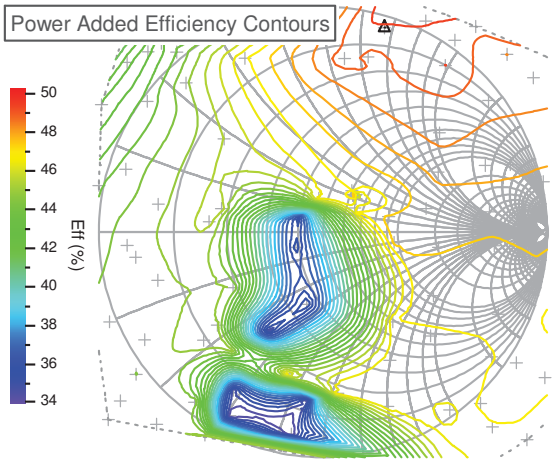


Figure 4-109, 2nd harmonic PAE contours with 10GHz fundamental and Γ_{L1} = 0.54/_105° (nominal).

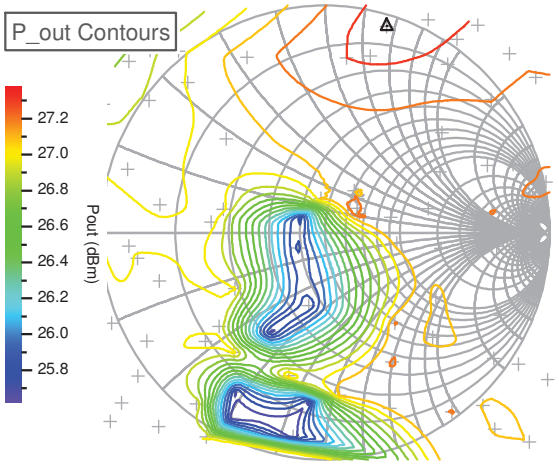


Figure 4-110, 2nd harmonic Output Power contours with 10GHz fundamental and Γ_{L1} = 0.54/_105° (nominal).

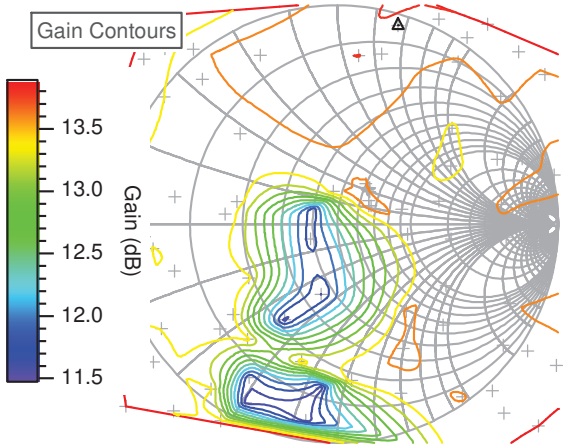


Figure 4-111, 2nd harmonic Gain contours with 10GHz fundamental and Γ_{L1} = 0.54/_105° (nominal).

There is a striking difference in the load pull contours, Figure 4-109 to Figure 4-111, in that there are clear regions of minima within the impedance plane, one of which encompasses the centre of the chart, the 50Ω load point. If in the fundamental measurement, the 2nd harmonic load had not been off this point, Figure 4-100, a dramatic reduction in PAE (from the 8GHz measurement) would have been seen, as well as exacerbating the difference in PAE with 2nd harmonic termination.

There appears to have been some problem with the measurements around the centre of the grid as a number load points have converged around the same area, Figure 4-112. Looking at the harmonic levels at the two points nearest the centre of the impedance plane, and therefore in the minima, there is no obvious difference in the power levels. Indeed looking at the 2nd harmonic at the optimum load point, Figure 4-116, it is about 35dB below the fundamental power level.

It is possible that the DUT was oscillating near the 50Ω 2nd harmonic impedance, which would explain the dip in efficiency and output power. This explanation was only considered during the write-up of the experiment and it was not possible to investigate further. A search on device instability due to harmonic terminations produced no results and is not a phenomenon that the author has heard of. It is interesting that the optimum is still in the same area as 4,6 and 8GHz, but that will not remain the case as seen at 12GHz.

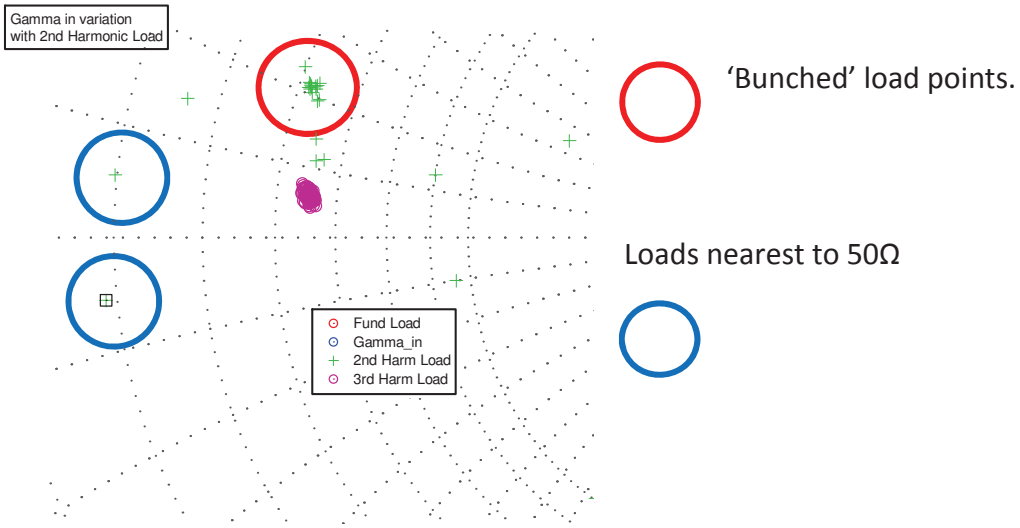


Figure 4-112, 10GHz 2nd Harmonic loads near centre of impedance plane.

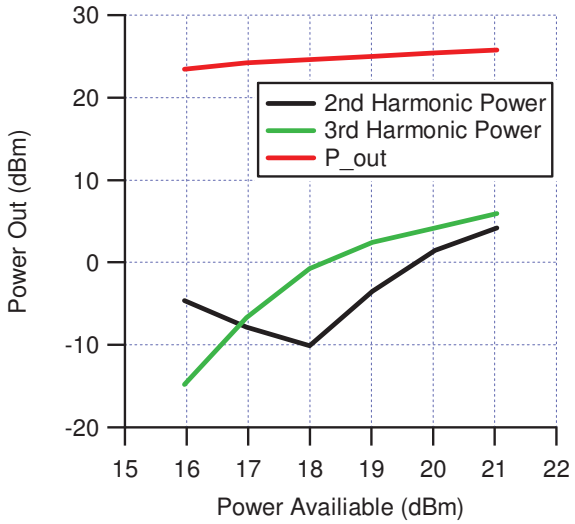


Figure 4-113, Harmonics characteristics of load #1 Figure 4-112

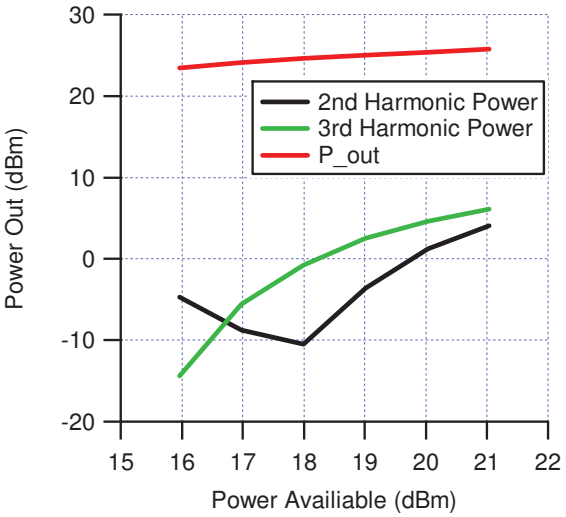


Figure 4-114, Harmonics characteristics of load #2 Figure 4-112

As before the optimum PAE 2nd harmonic load is in the top right hand corner, Figure 4-109. The gain contours of Figure 4-111 were included this time to see if they gave a clue as to why the minima occurred. The load pull was conducted at 6 power levels, and looking at the swept input power, with the optimum 2nd harmonic load point 0.92/_53.6° and the fundamental load at 0.50/_91.4°, it can be seen that the contours were not measured at the optimum PAE drive level. The input power of the 2nd harmonic 'dip' can be seen to be the same, but above this point the 2nd harmonic power level is about 10dB lower than with the fundamental only tuning, compare Figure 4-105 and Figure 4-116.

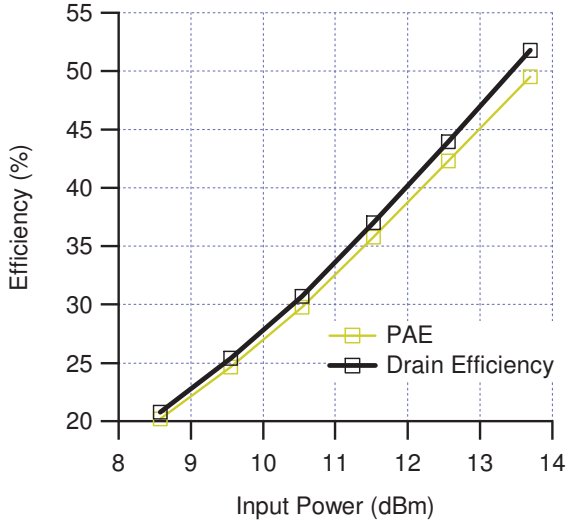


Figure 4-115, Efficiency at 10GHz with Γ_{L1} = 0.50/_91.4° and Γ_{L2} = 0.92/_54°.

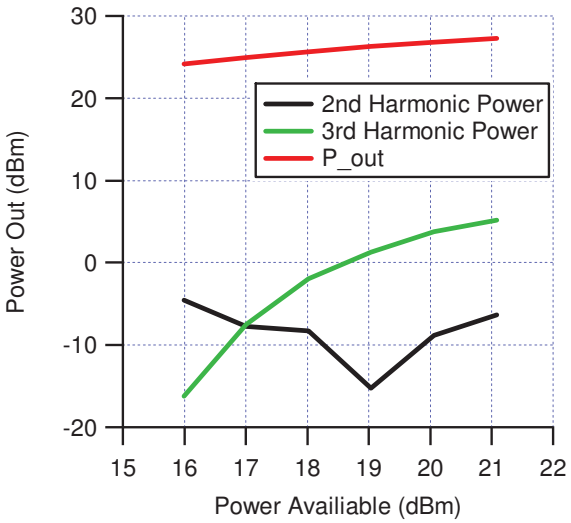


Figure 4-116, Harmonics at 10GHz with Γ_{L1} = 0.50/_91.4° and Γ_{L2} = 0.92/_54°.

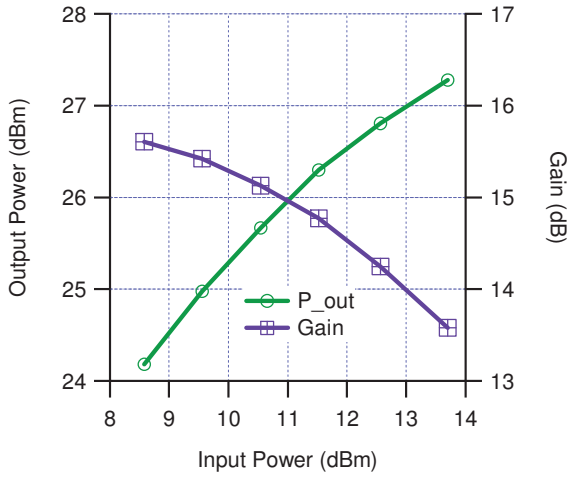


Figure 4-117, Output power and Gain at 10GHz with ΓL1 = 0.50/_91.4° and ΓL2 = 0.92/_54°.

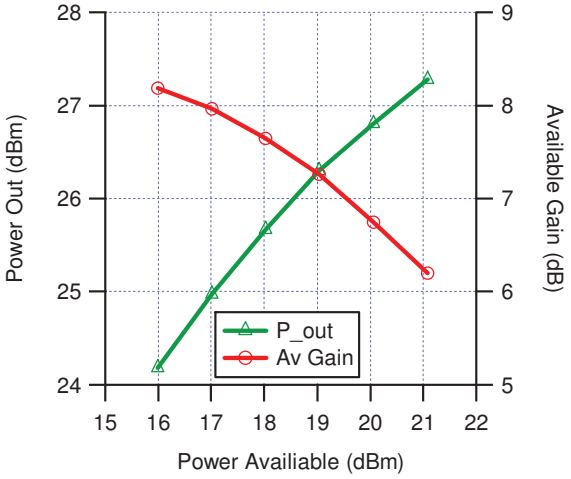


Figure 4-118, Output power and Available Gain at 10GHz with Γ_{L1} = 0.50/_91.4° and Γ_{L2} = 0.92/_54°.

Finally for this section, the performance of the device at 12GHz was measured.

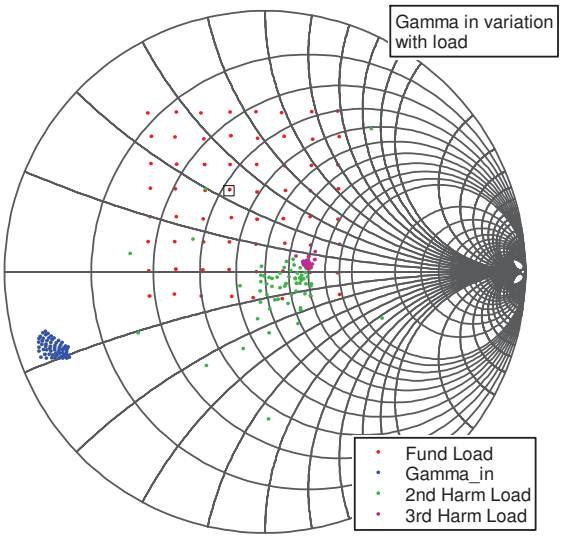


Figure 4-119, 12GHz Fundamental load pull grid, measured input and 2nd & 3rd harmonic Γ's

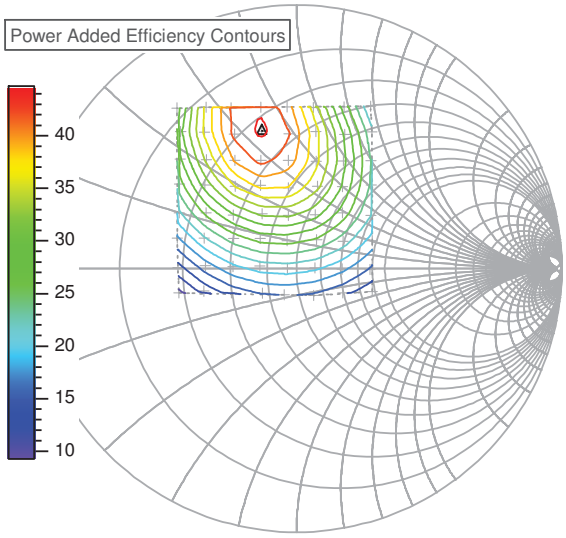


Figure 4-120, Measured 12GHz Fundamental PAE contours

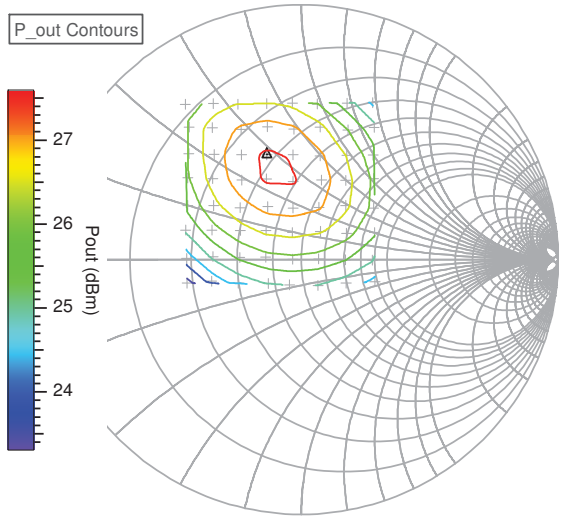


Figure 4-121, Measured 12GHz Fundamental Pout contours

	Load	PAE (%)	Pout (dBm)	Gain (dB)
PAE	0.54/_104°	44.5	27.2	10.8
Pout	0.43/_108°	42.7	27.6	10.7

Table 4-13, Measured 12GHz optimum loads and associated performance.

Conducting a power sweep at a fundamental load of 0.54/_104°, achieves an optimum PAE of 44.5% at an input power (optimum PAE drive) of 16.4dBm and an output of 27.2dBm. The peak PAE has been captured, Figure 4-122. At this power level the 2nd harmonic is -25.9dBc, having just increased from a dramatic dip. This suggests that it may be difficult to conduct a 2nd harmonic load pull, the levels varying markedly at these power levels. The results of the fundamental power sweep are shown in Figure 4-122 to Figure 4-125.

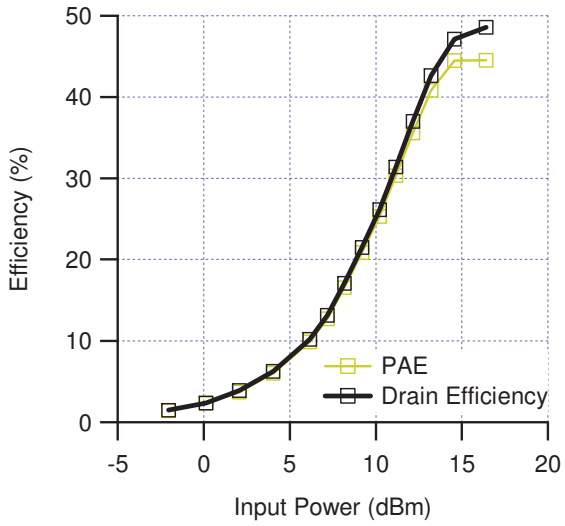


Figure 4-122, Efficiency at 12GHz with a fundamental load of 0.54/_104°

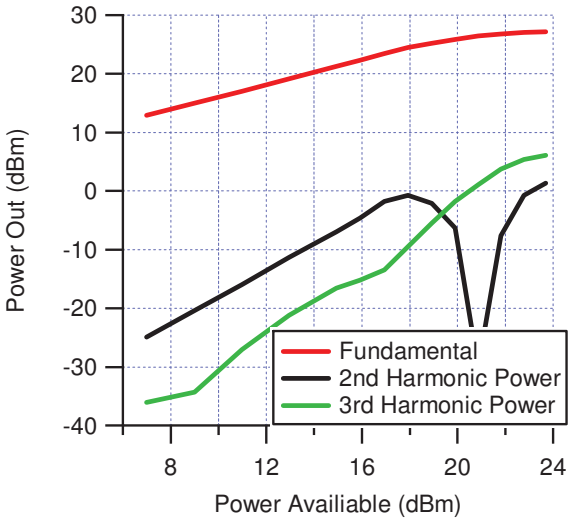


Figure 4-123, Harmonics at 12GHz with a fundamental load of 0.54/_104°

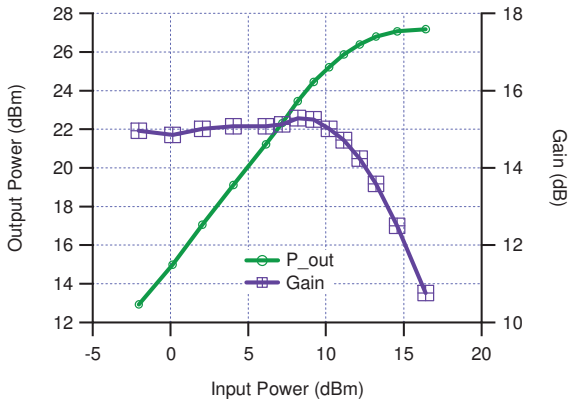


Figure 4-124, Output power and Gain at 12GHz with a fundamental load of 0.54/_104°

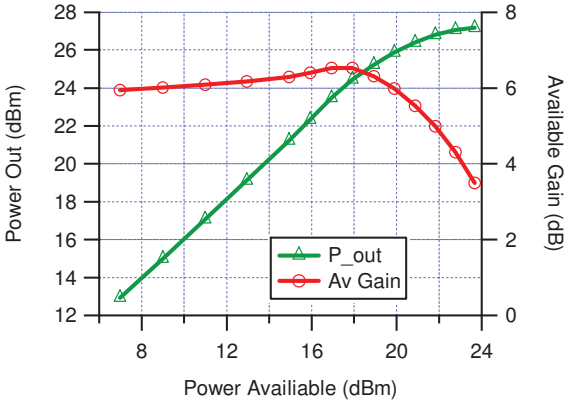


Figure 4-125, Output power and Available Gain at 12GHz with a fundamental load of 0.54/_104°

Following the now well-trodden path, the 2nd harmonic load pull was conducted. The fundamental load was passively tuned to a Γ of 0.36/_99°, which produces a PAE of 38.5% at an input power (ideal input match) of 20.3dBm and an output of 27.7dBm. This fundamental load is some way from the optimum found earlier (0.54/_104°), but a passive load was used to speed up the measurements, so that the system could concentrate on converging on the 2nd harmonic loads. Despite the same grid being set as in the 8GHz measurements as can be seen from Figure 4-127 the measurement points are no longer on the grid, this is mainly due to the fact that the fundamental and 2nd harmonic load sources cannot be locked phase coherently with this system, (12GHz fundamental derived from 6GHz multiplied by 2, whereas the 24GHz second is derived from 8GHz multiplied by 3). Another factor is that the low levels of harmonic change significantly with input drive level as was noted from Figure 4-123. The low levels of harmonic also result in very little variation in both PAE and Pout,

<3% and <0.3dB, across the impedance plane. As noted the position of the minimum and maximum PAE have both continued to rotate in opposite directions and thus have 'crossed over' by 12GHz. The unusual pattern of the minima within the impedance plane seen at 10 GHz which appears to be the frequency where the minimum and maximum would coincide, or perhaps more accurately clash, has been replaced by the more typical contours.

To check whether the optimum PAE load shifts with input drive level the 8GHz measurements are re-examined and the PAE contours at each drive level observed, Figure 4-126. From this a number of observations can be made, firstly the grid is better established at some power levels than others, and not necessarily corresponding to input drive level; the load points are clear at 7dBm on the left side of the Smith Chart and again at 15dBm. It is also seen that it is not necessary to have a consistent grid to see the trend in the performance with varying impedance, (it is however necessary to obtain a useful power sweep – in practice the device is more likely to see a constant load with drive power[4]).

[4] This actually depends on circuit topology and amplifier type; for example if the device is driving a class C transistor the load impedance will vary with drive.

If we observe the optimum PAE 2nd harmonic load, designated by a 'Δ' in Figure 4-126, with increasing input power this corresponds to the power sweeps of Figure 4-96 to Figure 4-99, and from these we can see that the regions of optimum PAE are clearly identifiable from the point where the gain response begins to deviate from linear, despite the fact that

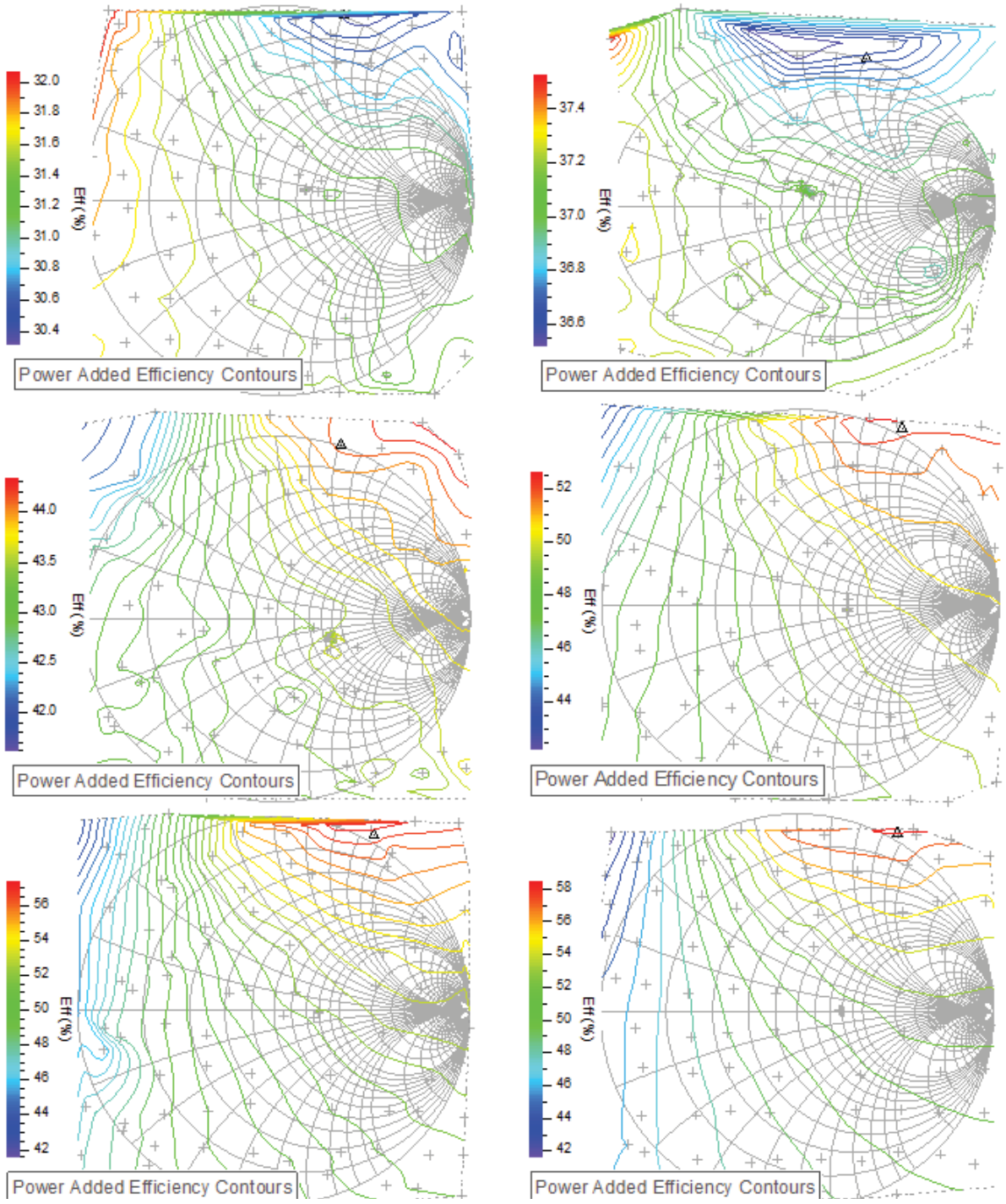


Figure 4-126, PAE 2nd harmonic Load pull contours at input power levels from 7 (top left) to 15 (bottom right) dBm.

at this point the 2nd harmonic level is in fact falling. Clearly, as described in chapter 1, when the class A biased transistor is in the linear regions the efficiency cannot be improved by harmonic addition (waveform engineering). One cannot read too much into the contours in the lower two power levels of Figure 4-126 as the variation across the whole impedance plane is 1%, whereas by the 4th power level there is >6% difference. Hence, provided we are operating in the nonlinear region we can predict the impedances for optimum performance and it is not imperative to be able to drive the device to the optimum PAE input power level. This is an important result as, due to the high input reflection coefficients, it can be difficult to provide sufficient input power to reach >2dB of compression, and this problem increases with frequency.

Returning to the measurements at 12GHz; having confirmed that being in the nonlinear region was sufficient for indicative results; harmonic load pull was conducted as before. As discussed earlier and in chapter 3 on the measurement system itself, 12GHz is outside of the phase coherent harmonic lock region (generated from twice 6GHz, whilst the 2nd harmonic, 24GHz, is from three times 8GHz), and so in order to speed the measurement acquisition time the fundamental load was produced passively. This meant that the maximum reflection coefficient that could be achieved was 0.36, which is significantly less than the 0.54 of the optimum load, and thus the efficiencies that we can hope to observe will be less than they could be. It was still possible to see that the PAE can be enhanced by altering the 2nd harmonic load from the nominal 50Ω load point. It should be recognised that these are compromise measurements, exploring how far the investigations can go with a non-optimum system.

As can be seen from the measurement results of Figure 4-127 to Figure 4-129, the grid is not well established, but also as has been said earlier there is enough spread across the impedance plane to be able to determine the impedances to target (and those to avoid). An unfortunate consequence of the lack of regularity to the impedance grid is that there are no load points that lie on the same impedance as the fundamental load grid (despite being set the same as a target). The nearest two fundamental loads are at 0.31/_96° and 0.42/_94°; whereas the fundamental load at the maximum PAE with 2nd harmonic tuning was 0.36/_99°. Plotting the performance of the two loads on ‘either side’ of the one used in the harmonic load pull measurement are shown in Figure 4-130 and Figure 4-131. The change in gain is surprising, which causes the offset to the right of the PAE curve, closer

investigation of Figure 4-127 shows that the input reflection coefficient was measured as 0.75/_172° as compared to the measurement at the fundamental, Figure 4-119, of 0.87/_-161°. It appears that the input de-embedding may not have been applied which has caused the rotation in the input gamma and the subsequent decrease in maximum gain; alternatively there may be some instability.

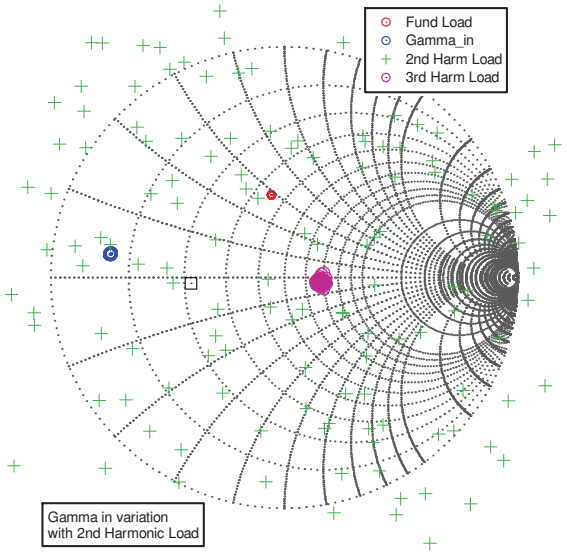


Figure 4-127, 2nd harmonic load pull grid with 12GHz fundamental and Γ_{in} & Γ_{L1} & Γ_{L3}

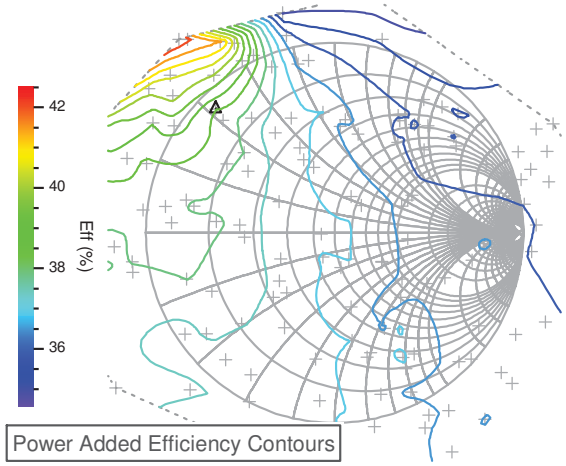


Figure 4-128, 2nd harmonic PAE contours with 12GHz fundamental load 0.36/_99°

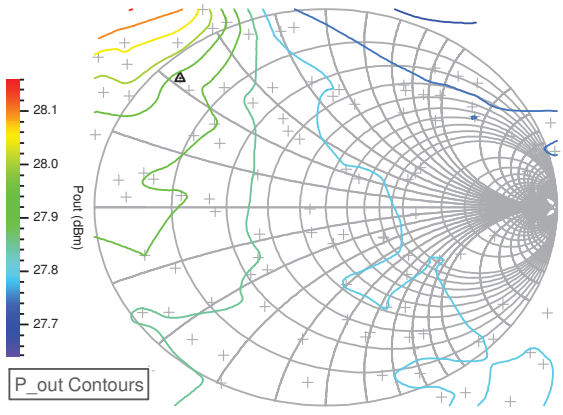


Figure 4-129, 2nd harmonic Pout contours with 12GHz fundamental load 0.36/_99°

Note:

- At the optimum 2nd harmonic termination of 0.91/_134° (best measured point) 38.5% PAE achieved with 27.9dBm output power and 6.8dB gain.
- When compared with Figure 4-54 and Figure 4-77 (6 & 4GHz 2nd harmonic PAE contours), the position of the optimum has now rotated anti-clockwise and the minimum has rotated further clockwise round the load impedance plane, so that they have effectively 'crossed over'.

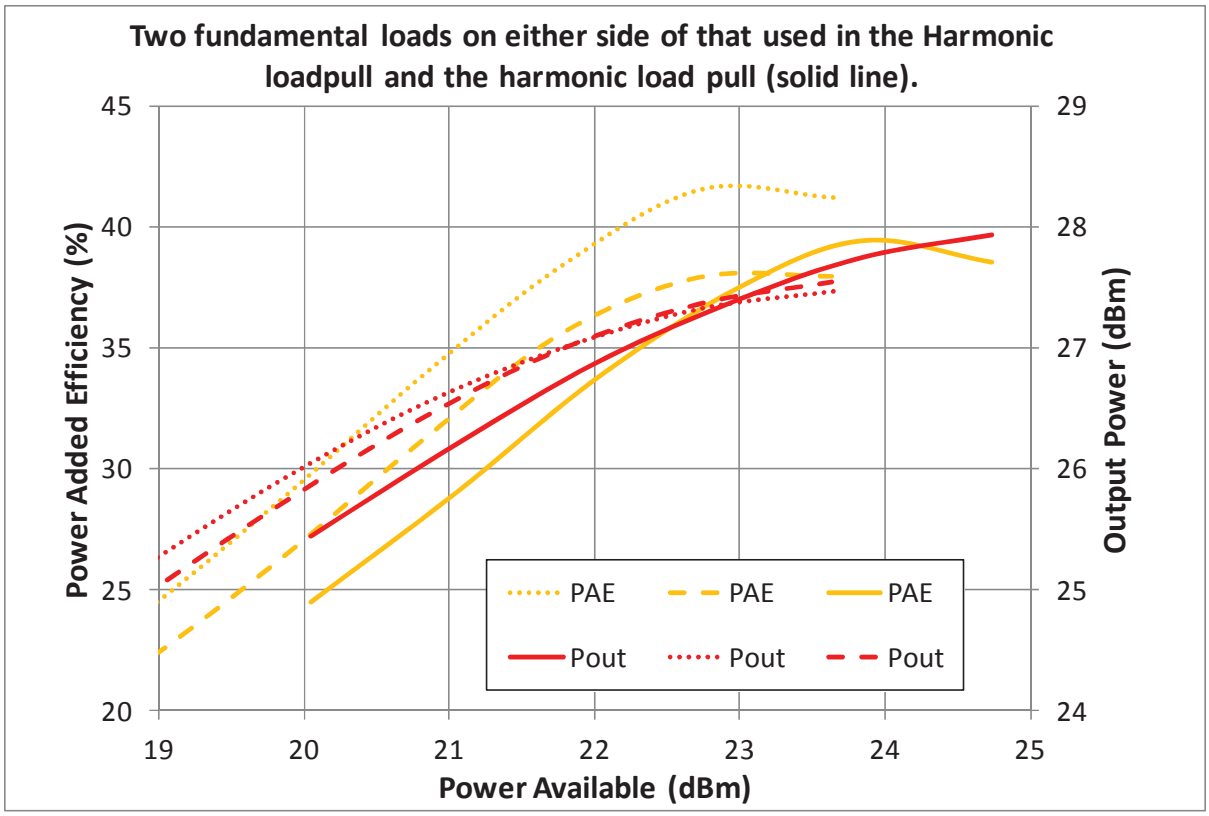


Figure 4-130, Efficiency at 12GHz with Γ_{L1} = 0.36/_99° and Γ_{L2} = 0.91/_134° (solid) and with Γ_{L1} = 0.31/_96° and 0.42/_94° and Γ_{L2} = 50Ω (dashed) – change in gain associated with input match change (see text).

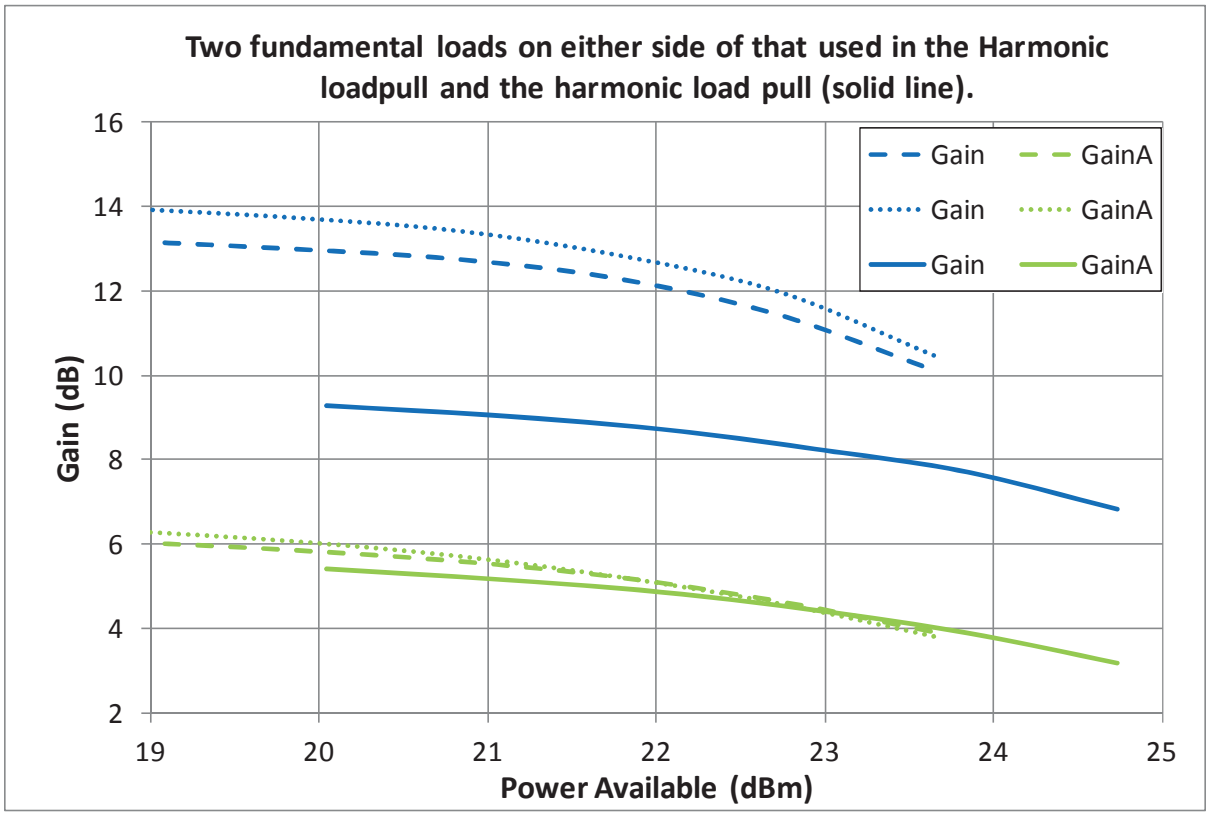


Figure 4-131, Gain at 12GHz with Γ_{L1} = 0.36/_99° and Γ_{L2} = 0.91/_134° (solid) and with Γ_{L1} = 0.31/_96° and 0.42/_94° and Γ_{L2} = 50Ω (dashed) – change in gain associated with input match change (see text).

A few points that are worth noting with regards to the measurement system as it operated during these measurements (improvements to the system are continually being made on both the hardware and the software).

The difference between the load convergence using a passively tuned load and an active can be demonstrated between the measurements made at 10 and 12GHz. As has been said, the passive load makes the measurements faster and as can be seen more repeatable, however especially as frequency increases it is not always possible to realise the load required passively due to the system losses. The actively tuned load depends on the convergence of power and phase of the injected signal to that required to match the b_{2n} wave from the device. This not only varies due to the load itself but also the noise in the system and as we have seen it is quite possible that there is a dip in the 2^{nd} harmonic power levels near compression hence reducing the 'headroom' above the noise floor. As can be seen in Figure 4-132 the passive load, as would be expected, provides excellent impedance stability. The active load on the other hand, Figure 4-133, shows significant phase variation (13.1°), although the magnitude error is <0.01. However at the lower power level the phase variation is of the same order, but the magnitude increases to ~0.05. When harmonic power levels are low they are close to the system noise level and hence it becomes difficult to match the reflected power level to these values, so as to produce the required reflection coefficient. This accounts for the scattering of the load grid points on the impedance plane, (we have even seen some points near to 50Ω move); it depends not only on the stability/phase coherence but also how close the harmonic level is to the noise floor of the system.

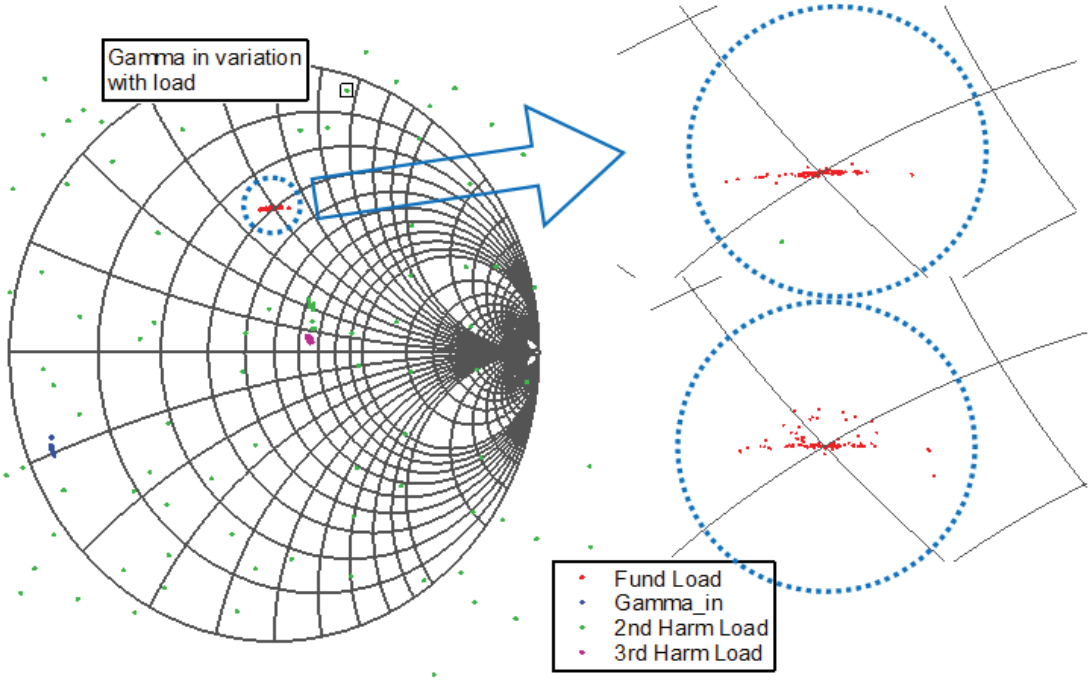


Figure 4-133, Fundamental active load variation with 2nd harmonic load pull at highest power, top right and lowest power, bottom right

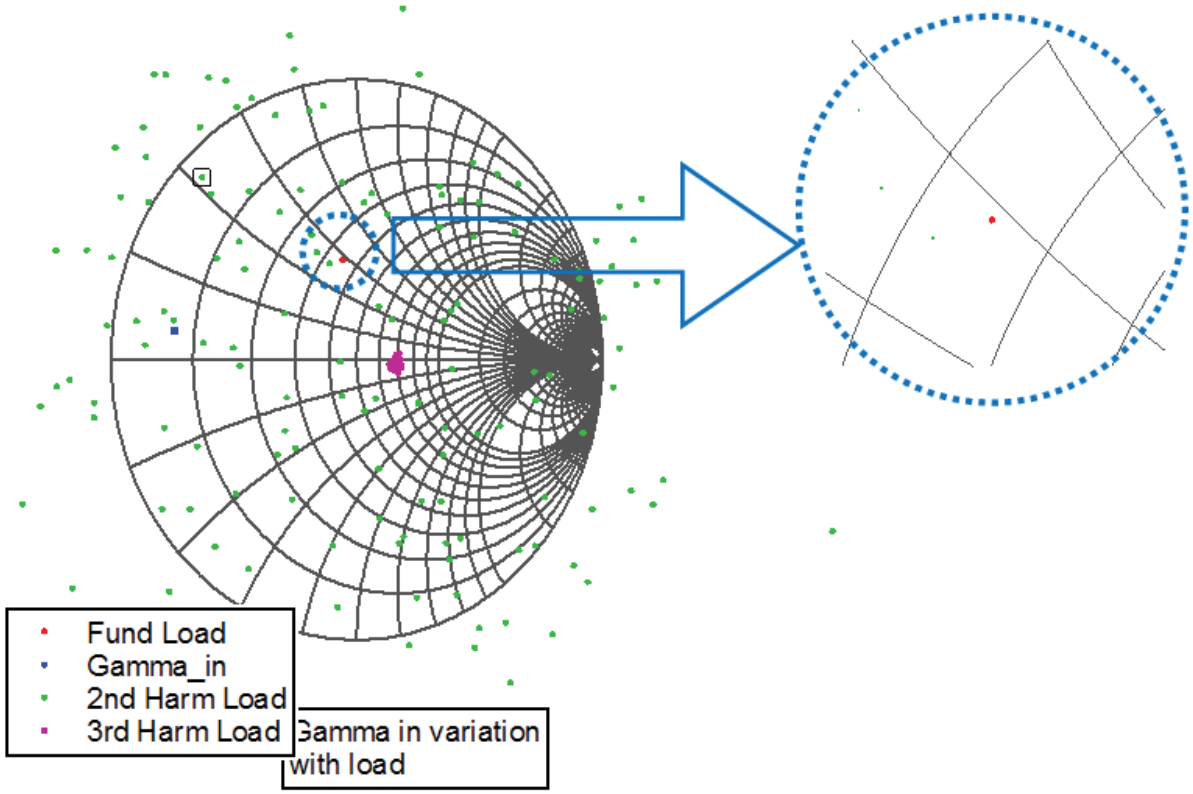


Figure 4-132, Fundamental passive load variation with 2nd harmonic load pull.

A summary of the harmonic measurements made is shown in Table 4-14 and the position of the minimum PAE in Table 4-15. Note that the measurements taken were not

ideal in that they did not all include the peak PAE that could be achieved, not having driven the device hard enough, also that the exact position of the optimum load and its associated performance could not always be determined as it lay outside the measurement grid. The objective of these experiments was to establish a process for acquiring the data and to observe the device characteristics over a large impedance plane and frequency range, giving graphical guidance to the design engineer.

PAE Maximum							
Freq (GHz)	Γ_{in}	Γ_{L1}	Γ_{L2}	PAE (%)	Pout (dBm)	Gain (dB)	Comment
4	0.80/-130°	0.24/39.8°	0.99/63.0°	61.2	28.4	16.4	Not opt Pin P2.5dB
6	0.86/-146°	0.43/63.4°	0.87/83.9°	61.8	27.8	14.4	Not opt Pin P3.5dB
8	0.87/-153°	0.45/84.7°	1.00/57.3°	58.0	28.0	12.8	Not quite opt Pin P7.8dB
10	0.91/-157°	0.50/91.4°	0.95/73.2°	49.4	27.4	13.7	1 dB short of opt.
12	0.90/-162°	0.54/105°	0.91/134°	44.5	27.2	10.8	Data from fund load pull

Table 4-14, Summary of Optimum PAE Loads and Performance

PAE Minimum							
Freq (GHz)	Γ_{in}	Γ_{L1}	Γ_{L2}	PAE (%)	Pout (dBm)	Gain (dB)	Comment
4	0.79/-128°	0.24/39.8°	0.94/-137°	46.9	27.6	15.3	Not opt Pin P2.5dB
6	0.84/-144°	0.43/63.4°	0.90/-170°	40.6	26.7	12.8	Not opt Pin P3.5dB
8	0.87/-152°	0.45/84.7°	0.94/170°	45.4	27.4	12.2	Not quite opt Pin P7.8dB
10	0.89/-161°	0.50/90.3°	0.92/-117°	34.1	25.6	11.5	Multiple minima
12	0.90/-162°	0.54/105°	0.91/44°	35.4	27.7	6.5	Data from fund load pull

Table 4-15, Summary of Minimum PAE Loads and Performance

The data from the summary measurements was used to create Figure 4-134 and Figure 4-135. The points plotted for the 2nd harmonic optimum loads are taken from the measured grids, and as has been noted the actual optimums sometimes lie outside of these areas, hence the values recorded here are not very accurate. Similarly, the minimum recorded at 10GHz was the lowest measured point, however it will be recalled from Figure 4-109, that there are in fact a number of low PAE points between this one plotted and the

50Ω load point, hence too much should not be read into this exact location. Plotted is also the conjugate match of the ideal equivalent circuit to the fundamental drain impedance.

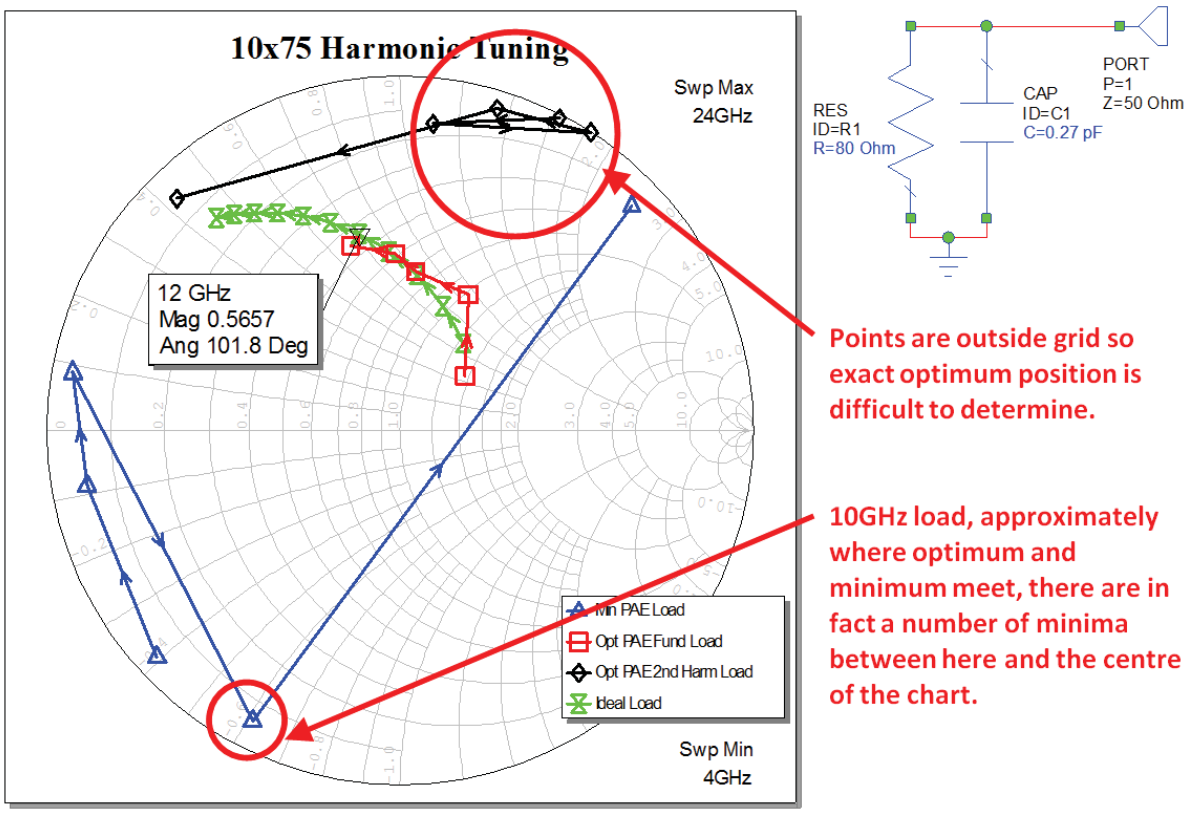


Figure 4-134, Trajectory of Optimum and Minimum PAE impedances with Frequency and the conjugate match of the equivalent circuit, top right

Investigating the trajectory of the 2nd harmonic loads the resistance of the equivalent circuit was increased to move the response to the edge of the Smith chart, without altering the estimated drain capacitance. If we consider the 2nd harmonic response we can see that using this equivalent circuit provides a good estimation of the optimum 2nd harmonic load. Clearly to confirm these results it will be necessary to conduct 2nd harmonic load pull at high reflection coefficients in a circular grid at the edge of the impedance plane.

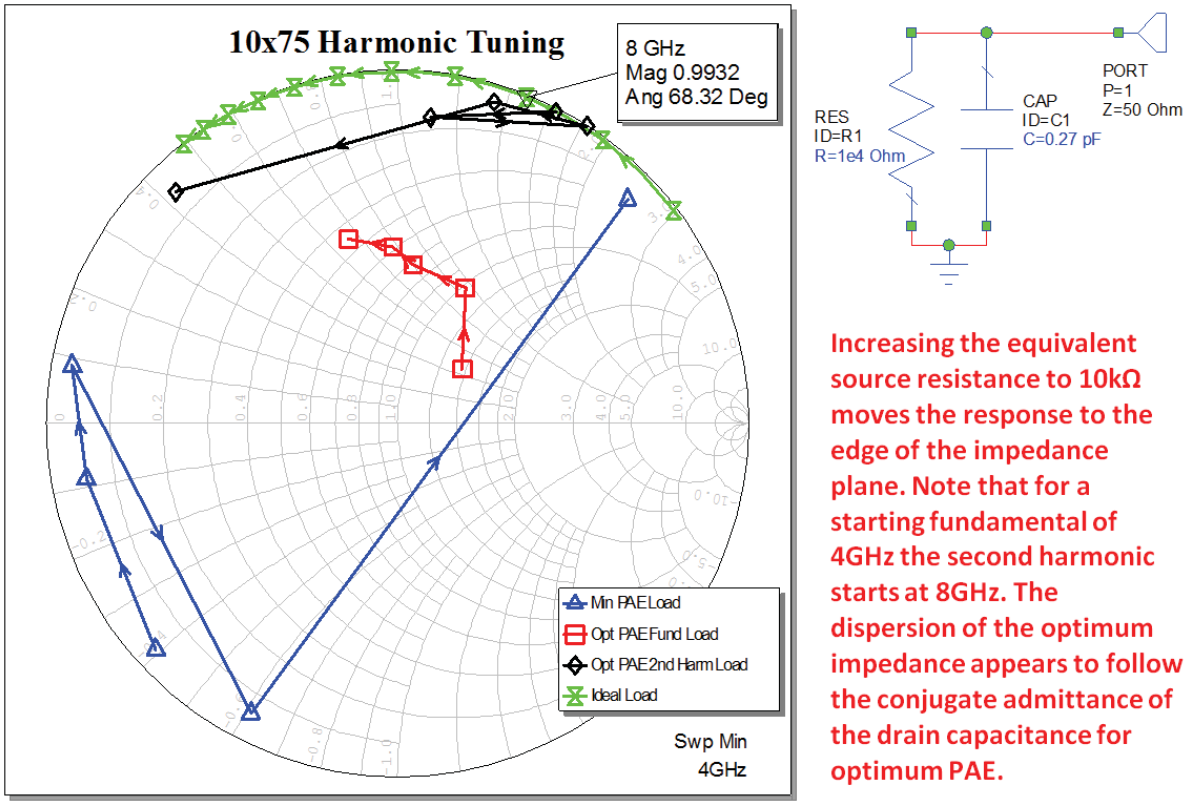


Figure 4-135, High resistance current source equivalent circuit response at 2nd harmonic frequencies

4.6 Reduced Conduction Angle Measurements

The measurements described so far have concentrated on devices biased in class A, mainly due to the fact that this was the standard approach used by the industrial sponsor in their broad band applications. This mode of operation produces good linearity, gain and power but at the expense of efficiency. We have seen however, Figure 4-49, that as devices are driven into saturation the load line moves towards that of a class B characteristic. Thus it may be that there are advantages to starting in this condition. The improvements in efficiency are well documented [10], but a further improvement may be the ability to operate from a higher voltage rail, which would not normally be used in class A due to the higher power dissipation. There are a number of mechanisms implicated in device breakdown and although in years past voltage was a key factor, process developments have been such that GaAs FET [11] (and even more so in GaN) devices are often thermally limited rather than by voltage breakdown. Further it has been observed [12] that the RF breakdown can be higher than the DC level standard device measurement data would suggest. The

literature is not extensive on this effect but the general feeling is that the duration of the RF voltage peaks is less than the period of the avalanche breakdown and so this failure mechanism cannot occur. As a result it is not always clear what the real voltage limits are. The device measured is from a process that is commercially described as being 10V; however as has been seen from the voltage waveforms the test transistors have seen drain source voltages in the region of 20V. By biasing the devices in class B the thermal effects will be reduced and thus it should be possible to safely operate the devices at a higher voltage. The impact of bias current and drain voltage were assessed and compared to that of the class A device and the PDK predictions.

Rather than follow a specific class of operation the effects of altering the supply rail and reducing the quiescent drain current were examined. This shows trends rather than specific performance which is naturally investigated in the light of a particular requirement. Here we are attempting to show the direction one should follow in order to achieve a specific improvement and what the consequences of such action would be.

Adjusting the gate bias such that the drain current with RF present was 40mA a fundamental load pull was conducted at 8GHz. The second harmonic was then load pulled with the fundamental load as close to the optimum that could be achieved with a passive tuned load (used for reasons explained earlier). The results are summarised in Table 4-16. Compared with the class A, fundamental only load tuning the change in bias appears to result in a >10% increase in PAE. The original 150mA measurement was not made at the peak PAE drive level, so both the PAE and the output power will be higher, whilst the gain will be lower. The results bear out the conventional wisdom, that as the bias is moved away from class A the output power and gain are compromised.

Similarly the measurements at 40 mA bias with harmonic tuning do not use the optimum fundamental load impedance, and hence it is difficult to make a direct comparison with the performance at the optimum. In an attempt to do this the two fundamental load points, above and below (with regards to the Smith Chart) the impedance of the fundamental used in the harmonic load pull would suggest a fundamental only PAE of ~60.6%, hence the improvement with the 2^{nd} harmonic tuning is not as significant as in the class A case. On the other hand the 2^{nd} harmonic tuning does appear to improve PAE, output power and gain, these latter two parameters usually having degraded somewhat compared to class A.

Freq (GHz)	Supply (V)	Current (mA)	Γ_{in} M/A	Γ_{L1} M/A	Γ_{L2} M/A	PAE (%)	Pout (dBm)	Gain (dB)	Comment
8	9	150	0.91 /-156°	0.55 /79.6°	0.20 /39.4°	53.3	26.8	14.2	Opt. PAE at 3.5dB Av. gain compressed.
8	9	150	0.87 /-153°	0.45 /84.7°	1.00 /57.3°	58.0	28.0	12.8	Pin just too low to reach PAE max.
8	9	40	0.93 /-149°	0.55 /80.3°	0.19 /29.0°	63.6	26.5	14.6	Fundamental only Opt PAE load
8	9	40	0.89 /-147°	0.48 /79.0°	0.19 /31.4°	61.5	27.0	12.3	Fundamental only Above 2HLP load
8	9	40	0.93 /-147°	0.45 /84.2°	0.97 /-51°	63.2	27.3	15.4	2HLP - not opt fund impedance
8	9	40	0.88 /-148°	0.41 /77.3°	0.14 /65.5°	59.7	27.1	13.1	Fundamental only Below 2HLP load

Table 4-16, 8GHz performance variation with bias.

Comparisons between the measured data and the PDK model are best done graphically, as rather than looking at the absolute difference between specific power levels it is more informative to look at performance characteristic shapes and trends. The choice of optimum points in the measured data is based on the maximum power grid load, the differences between the load points can be small, and so the actual peak may well be slightly different. Indeed if a graphical estimation based on the centre of the contour in Figure 4-137 were made the result would be much closer to that of the simulation. As seen earlier the PDK gives slightly more optimistic results.

Measurement and Simulation at 9V, 40mA

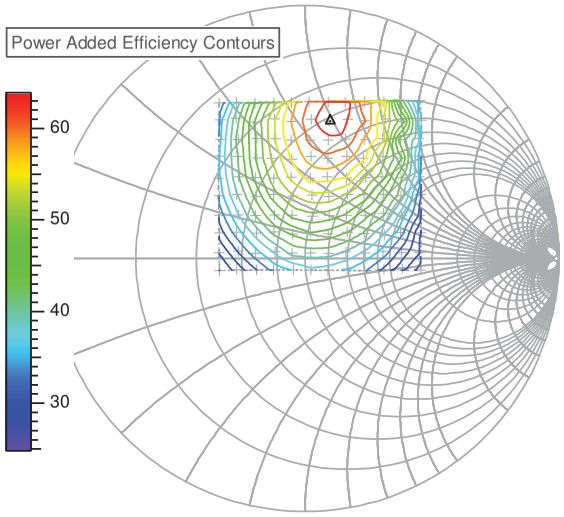


Figure 4-136, 8GHz PAE Contours, maximum PAE 63.6%, Γ_{L1} 0.55/_80.3°, 26.5dBm

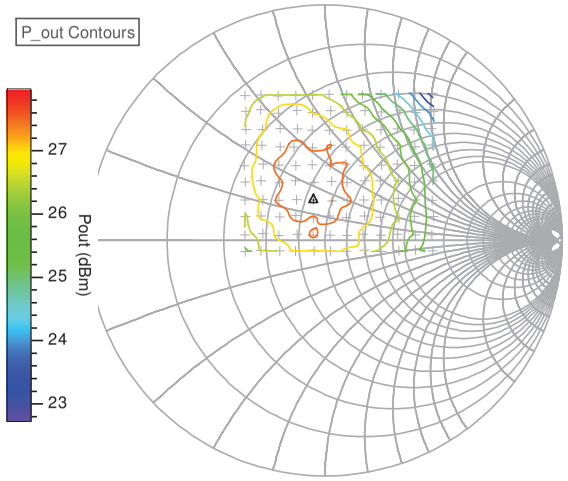


Figure 4-137, 8GHz Pout Contours, maximum Pout 28.0dBm with PAE 48.7%, Γ_{L1} 0.18/_106°.

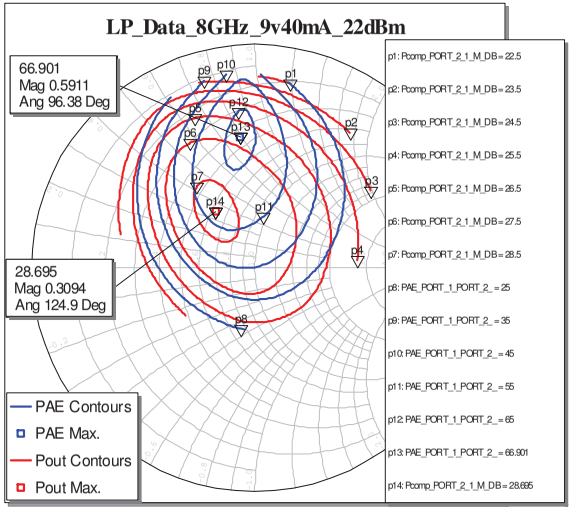


Figure 4-138, Simulated 8GHz Load pull Contours, max. PAE 66.9%, Γ_{L1} 0.59/_96.4°, max. Pout 28.7dBm, Γ_{L1} 0.31/_125°

Measurement and Simulation at 9V, 40mA

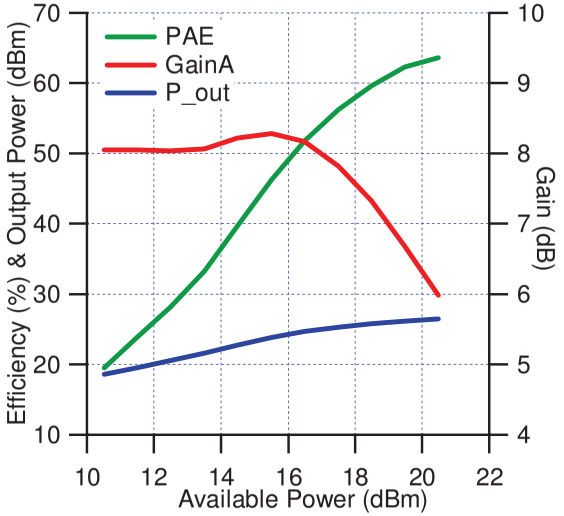


Figure 4-139, Measured swept input power performance at optimum PAE fundamental load Γ_{L1} 0.55/_80.3°

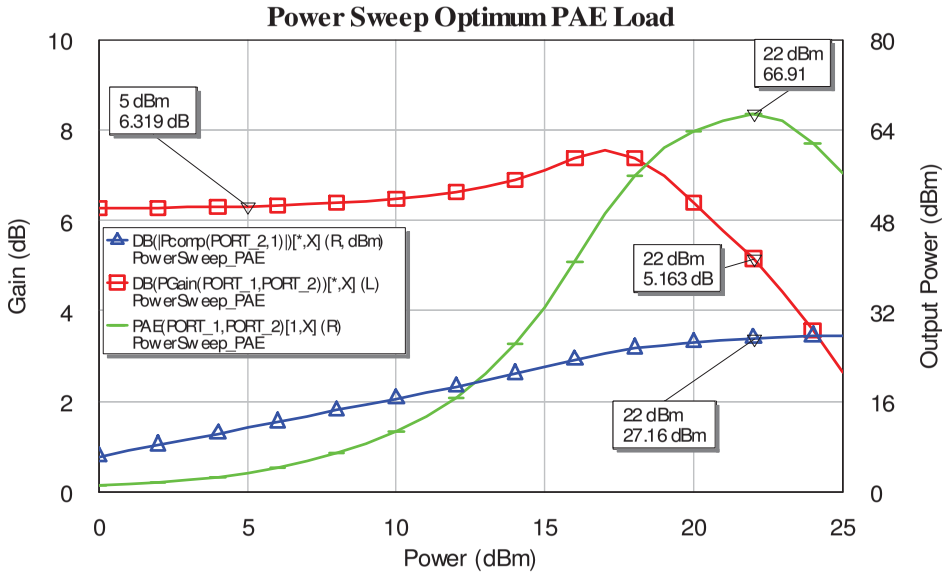


Figure 4-140, Simulated swept input power performance at optimum PAE fundamental load Γ_{L1} 0.59/_96.4°

Measurement and Simulation at 9V, 40mA

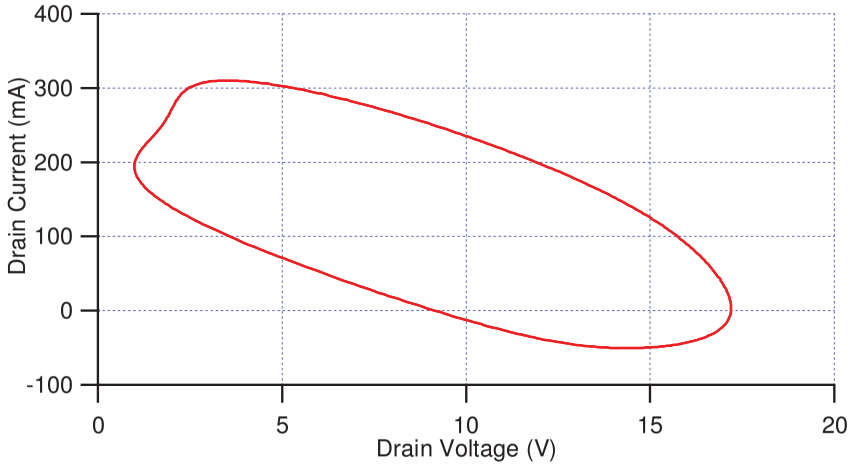


Figure 4-141, Measured 8GHz Dynamic loadline, at device plane, 9V 150mA bias, Γ_{L1} 0.55/_79.6°

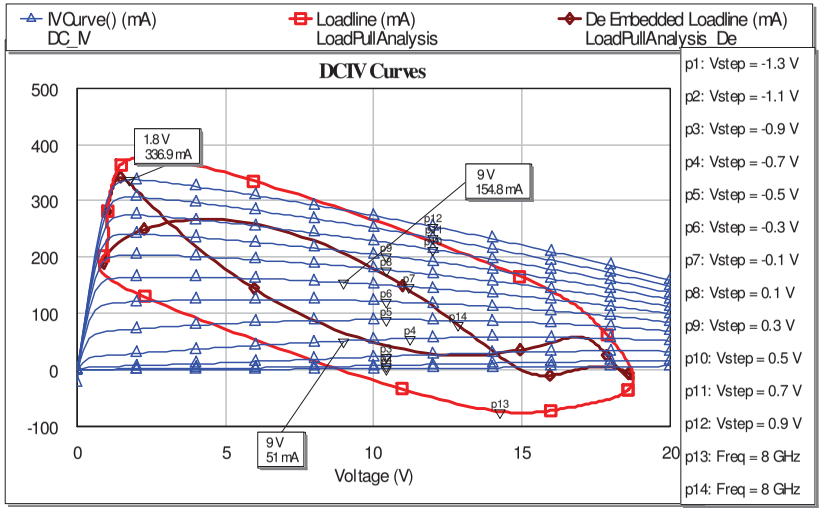


Figure 4-142, Simulated 8GHz Dynamic loadlines, at device (red) and CG (brown) planes, 9V 150mA bias, Γ_{L1} 0.39/_95.2°, de-embedding capacitance = 0.27pF.

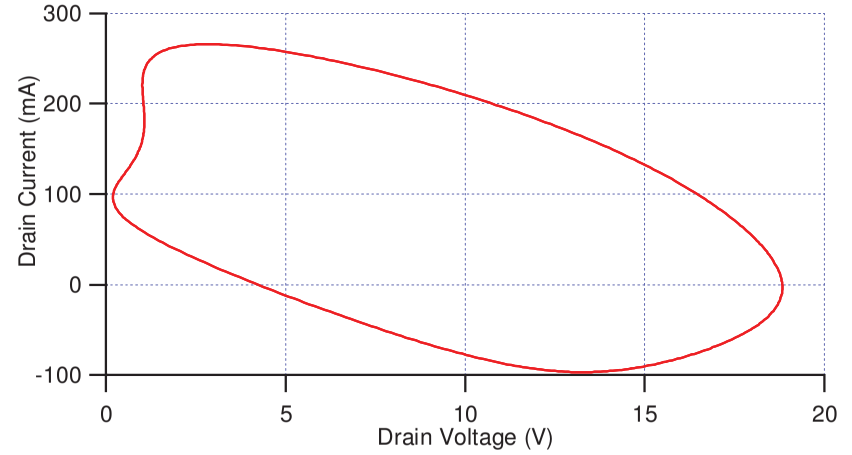


Figure 4-143, Measured 8GHz Dynamic loadlines at device plane, 8V 40mA bias, Γ_{L1} 0.55/_80.3°

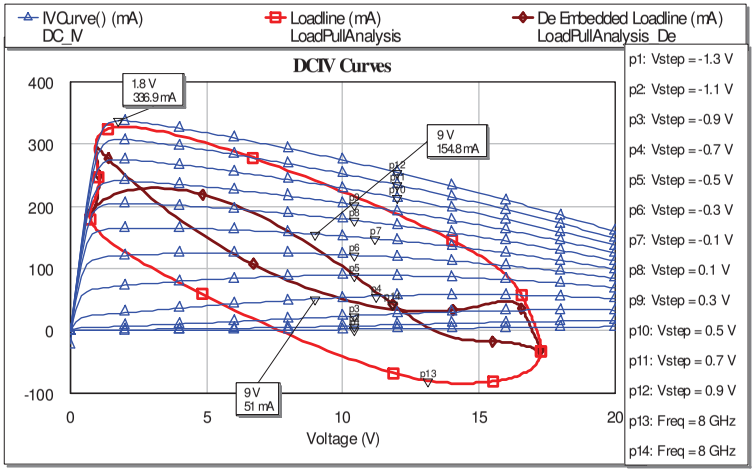


Figure 4-144, Simulated Dynamic loadlines, at device (red) and CG (brown) planes, 8V 40mA bias, Γ_{L1} 0.59/_96.4°, de-embedding capacitance = 0.27pF.

It appears that the measured data and the PDK model are closer to agreement with the lower bias current. It is interesting that in the case of the measured data the optimum PAE load has not changed significantly, whilst there has been a marked changed in the simulated model (from 0.39/_95.2° at 150mA to 0.59/_96.4° at 40mA). This may be indicative of one difference between the measurement and the simulation that has not been mentioned so far; the measured data is from a device mounted onto a metal carrier, whilst the PDK model is largely based on wafer measurements using pulsed measurements to avoid heating effects. Thus it may be that in the more efficient mode we get better agreement between the approaches as the temperature differences through self-heating are less. Looking at the dynamic load lines of Figure 4-141 to Figure 4-144, the agreement can be seen; in the case of Figure 4-141 the fact that the measurement had not reached the peak PAE, i.e. was not driven hard enough, can be seen by the fact that the top left corner of the 'loop' has not reached the edge of knee as in Figure 4-143. In the simulated responses the dynamic loadline has also been de-embedded to the Current Generator (CG) plane using a nominal capacitance of 0.27pF. In Figure 4-142 the de-embedded loadline just brushes the 0mA axis showing that the displacement current due to the drain capacitance has been accounted for, however in Figure 4-144 there is a negative current element to the loadline which indicates that the value of C_{ds} has changed. It has been well documented that the gate source capacitance changes significantly in the region of the gate voltage changes made (from -0.18v for 150mA to -0.75v for 40mA), however the effect of changes in drain capacitance have been less well recorded, but as the peak drain voltage has changed then this is a possible cause [6].

Looking at the 2^{nd} harmonic load pull with the fundamental load set nears its optimum, it can be seen that the difference between the measured contours and the simulated is larger than in the class A bias case Figure 4-55, also the simulation is now predicting a variation of 14% across the impedance plane whilst the measurement <10%. This is a reversal of the effects seen in class A. The 2^{nd} harmonic load pull contours were previously quite 'clean' whereas at 8GHz 40mA bias; the contours now display considerable deviation from the 'clean' curves. This occurs in both the PAE and output power graphs, Figure 4-145 and Figure 4-146, however as can be seen from Figure 4-150, the levels of 2^{nd} harmonic are very low, which makes load convergence difficult.

Measurement and Simulation at 9V, 40mA

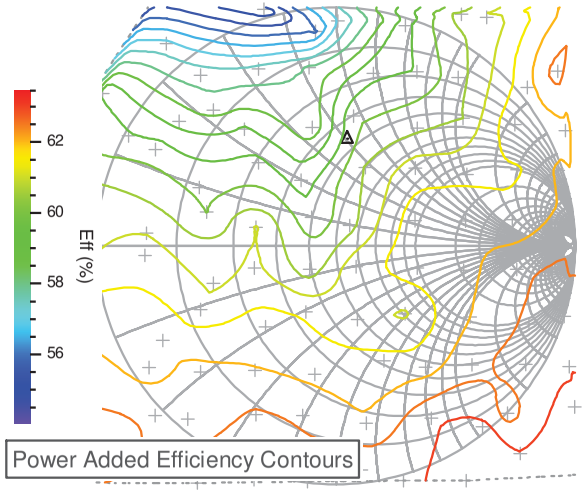


Figure 4-145, 8GHz measured 2nd harmonic PAE contours Γ_{L1}=0.45/_84.2°, optimum PAE load Γ_{L2}=0.97/_-51°, PAE 63.2%, 27.3dBm.

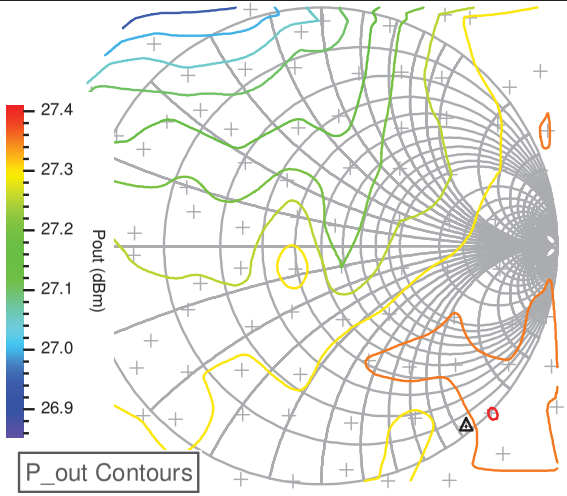


Figure 4-146, 8GHz measured 2nd harmonic Pout contours optimum Pout load Γ_{L2}=1.00/_-44°, PAE 62.3%, 27.4dBm.

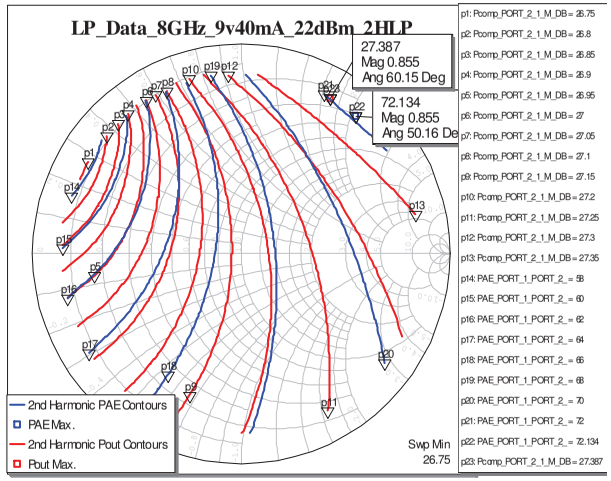


Figure 4-147, 8GHz simulated 2nd harmonic PAE and Pout contours Γ_{L1}=0.59/_96.4°, optimum PAE load Γ_{L2}=0.86/_60°, PAE 72.1%, 27.4dBm

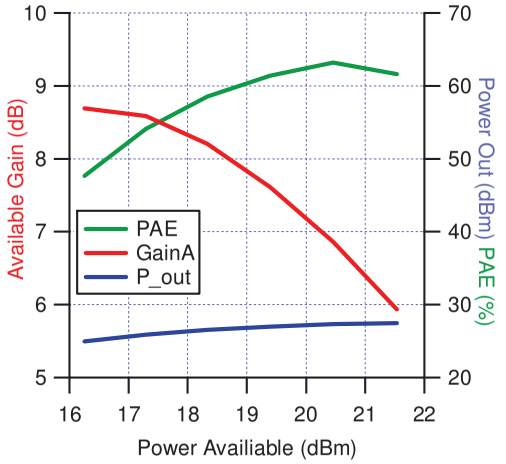


Figure 4-148, Measured power sweep at 8GHz, Γ_{L1}=0.45/_84.2° and Γ_{L2}=0.97/_-51°.

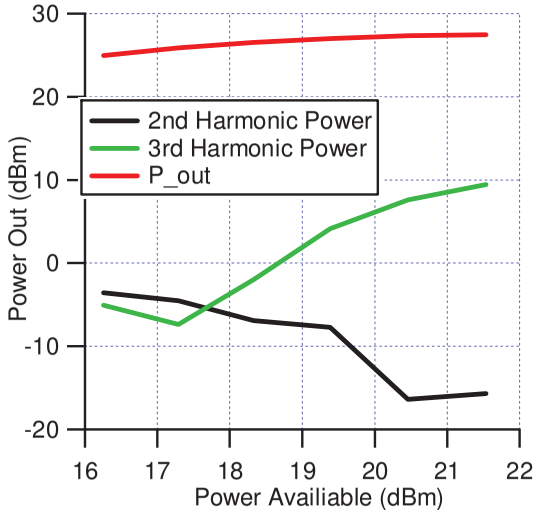


Figure 4-150, Measured harmonics at 8GHz, Γ_{L1}=0.45/_84.2° and Γ_{L2}=0.97/_-51°.

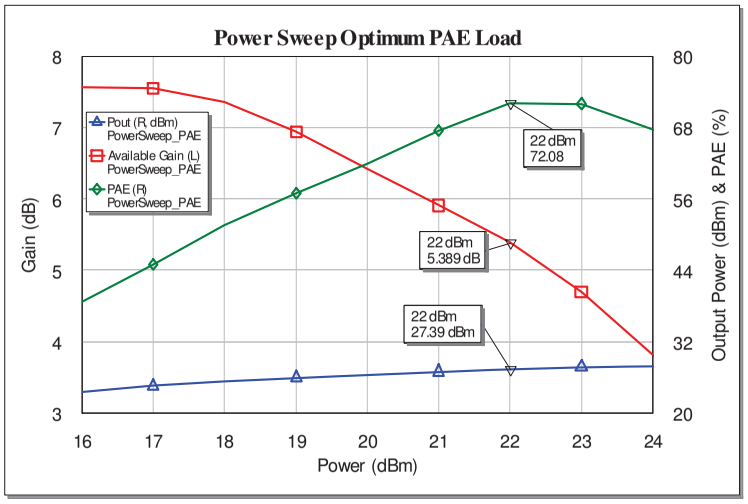


Figure 4-149, Simulated power sweep at 8GHz, Γ_{L1}=0.59/_96.4° and Γ_{L2}=0.86/_60°.

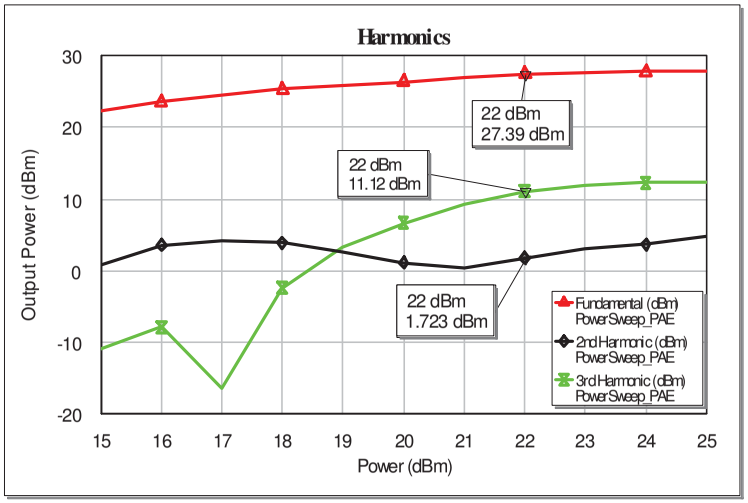


Figure 4-151, Simulated harmonics at 8GHz, Γ_{L1}=0.59/_96.4° and Γ_{L2}=0.86/_60°.

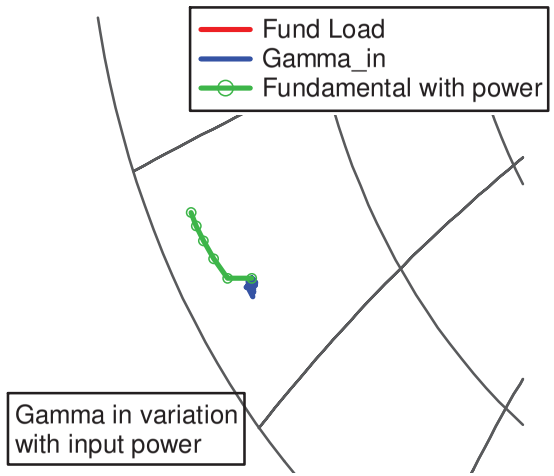


Figure 4-152, Measured input reflection coefficient with swept input power (green) at 8GHz, Γ_{L1}=0.45/_84.2° and Γ_{L2}= 0.97/_-51°.

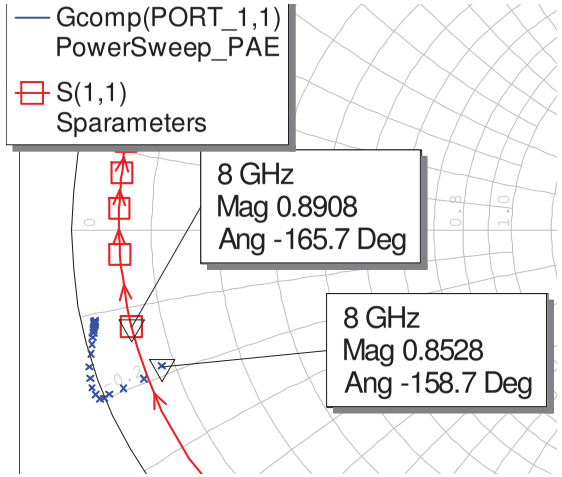


Figure 4-153, Simulated input reflection coefficient with swept input power at 8GHz, Γ_{L1}=0.59/_96.4° and Γ_{L2}=0.86/_60°.

In general the PDK seems to provide a good prediction of the measured performance in terms of typical trends; however the PAE is some 10% higher than measured, but note that the measurement was not at exactly the optimum fundamental impedance, hence the actual difference will be less. Further, the position of the optimum fundamental loads is different by 12° and the 2nd harmonic by ~100°. This difference in 2nd harmonic optimum load position may in part be due to the very low levels of 2nd harmonic predicted by the model ~20dB.

The frequency was changed to 10GHz and a new set of measurements were made as summarised in Table 4-17. Initially both the drain voltage and the drain current were reduced, to 8V and 40mA respectively. Note the much improved agreement in optimum load location between the measured and simulated results shown in Figure 4-154 to Figure 4-156. The PDK model is still predicting higher performance (PAE: 64.2:61.0% and Pout 27.6:26.4dBm). The device was then measured at 12V 80mA and 10V 40mA (with the 2nd harmonic shorted).

Freq (GHz)	Supply (V)	Current (mA)	Γ_{in}	Γ_{L1}	Γ_{L2}	PAE (%)	Pout (dBm)	Gain (dB)	Comment
10	9	150	0.91/-157°	0.50/91.4°	0.95/73.2°	49.4	27.4	13.7	1dB short of optimum PAE drive level.
10	8	40	0.90/-149°	0.54/97.5°	0.05/-7°	61.0	25.9	10.1	Significant decrease in output power and gain.
10	12	80	0.89/-160°	0.54/85.8°	0.25/91°	51.6	28.2	12.2	Slight increase in PAE over class A, significant increase in Pout.
10	10	40	0.86/-152°	0.55/82.6°	0.98/180°	62.7	26.4	10.3	2nd Harm S/C – High efficiency mode, poor Γ_{L2} control.

Table 4-17, 10GHz measurements summary with varying bias and harmonic loads.

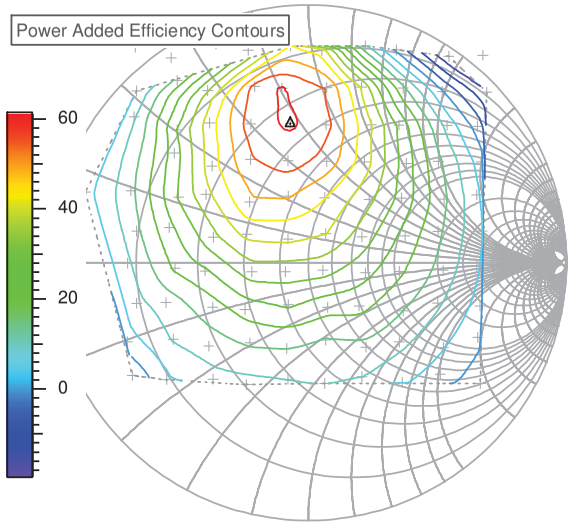


Figure 4-154, Measured 10GHz, 8V 40mA PAE contours, maximum PAE 61.0% 0.54/_97.5°, 25.9dBm.

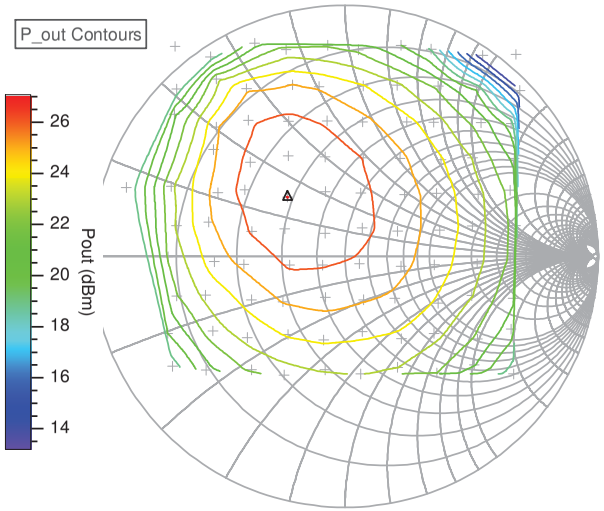


Figure 4-155, Measured 10GHz, 8V 40mA Pout contours, maximum Pout at 0.33/_135°, 27.0dBm, 49.4%.

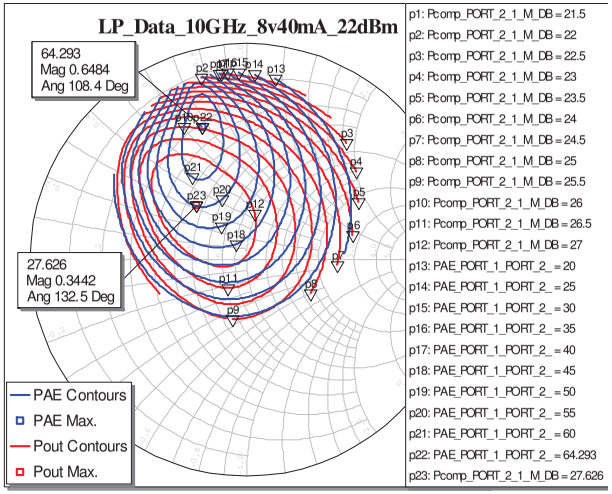


Figure 4-156, Simulated 10GHz, 8V 40mA PAE and Pout contours, maximum PAE 64.3% 0.65/_108°, 25.9dBm.

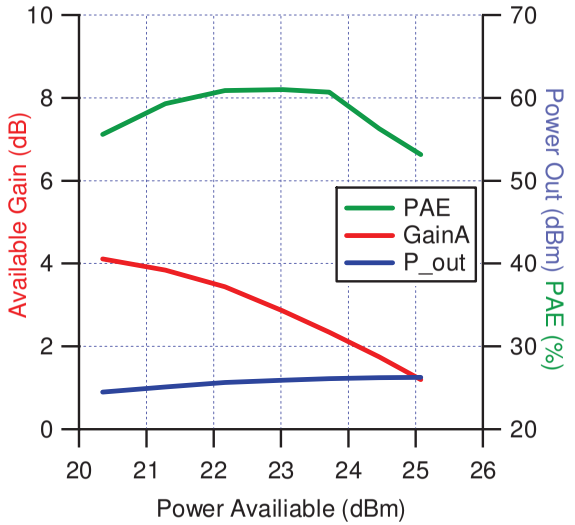


Figure 4-157, Measured 10GHz, 8V, 40mA power sweep with optimum PAE load.

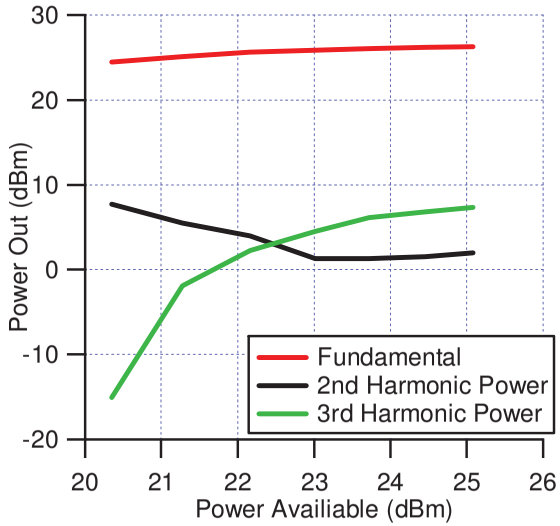


Figure 4-158, Measured harmonics at 10GHz, 8V, 40mA power sweep with optimum PAE load.

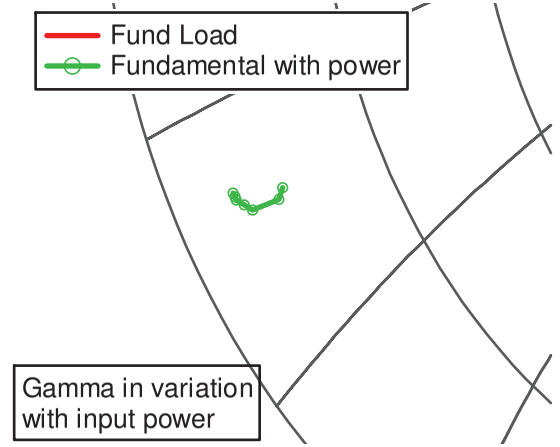


Figure 4-159, Measured input reflection coefficient with swept input power at 10GHz, 8V, 40mA power sweep with optimum PAE load. Γ_{In1}=0.91/_-151°, Γ_{In2}=0.84/_-148°, Γ_{L1}=0.54/_98.7.

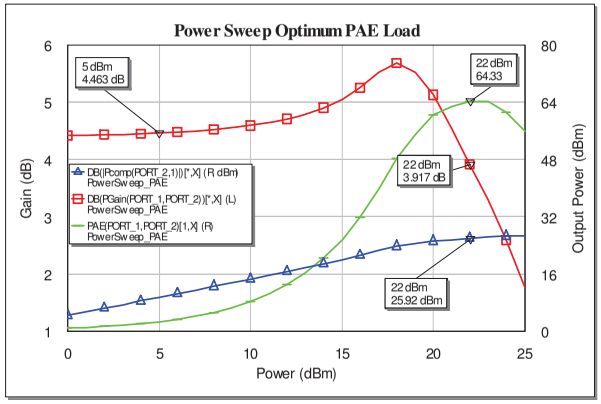


Figure 4-160, Simulated 10GHz, 8V, 40mA power sweep with optimum PAE load.

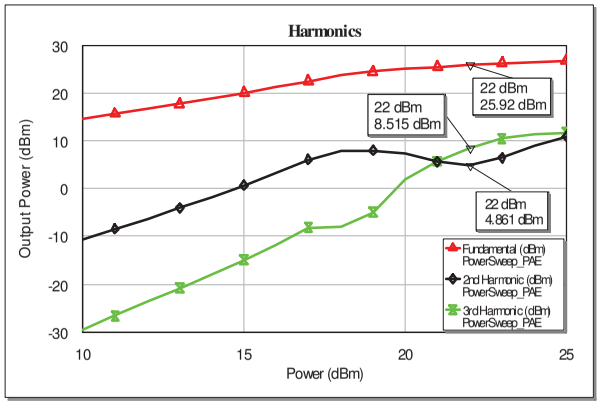


Figure 4-161, Simulated harmonics at 10GHz, 8V, 40mA power sweep with optimum PAE load.

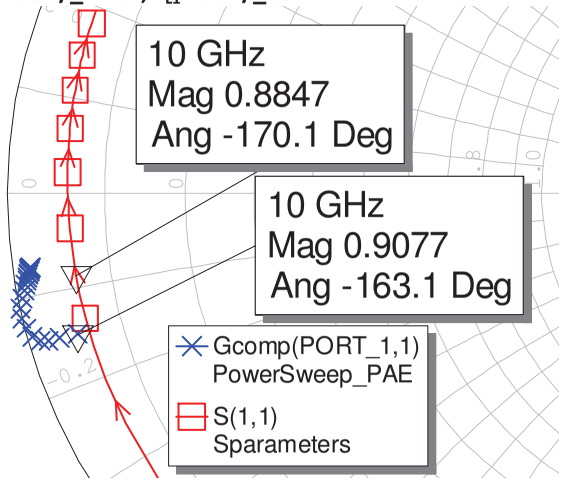


Figure 4-162, Simulated input reflection coefficient with swept input power at 10GHz, 8V, 40mA power sweep with optimum PAE load.

The results for the 8V 40mA bias show a significant output power and gain reduction, hence the next step was to increase the drain voltage to 12V and move the bias to a more A/B mode, by increasing the quiescent current to 80mA. The wafer process is specified as 10V, hence this is exceeding the recommended level however as has been said failure tends to be due to thermal issues and by operating in a more efficient class this should alleviate the reliability problem. This operating point is however at the top limit of the measured DC-IV curves from which the PDK model was in part generated so it will be interesting to see how this compares. As can be seen in Figure 4-163 to Figure 4-165, the positions of the optimum PAE and output power are fairly close, however the PDK model is optimistic by over 8% in PAE and 1dB in power.

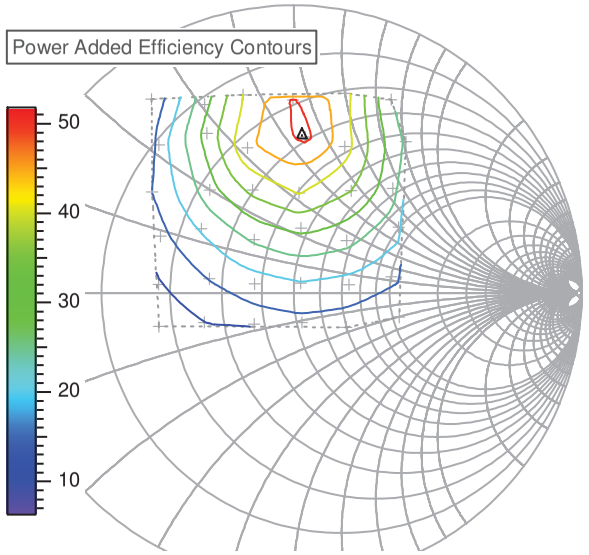


Figure 4-163, Measured 10GHz, 12V 80mA fundamental PAE contours. Optimum PAE 51.6% at 0.54/_85.8°, 28.2dBm

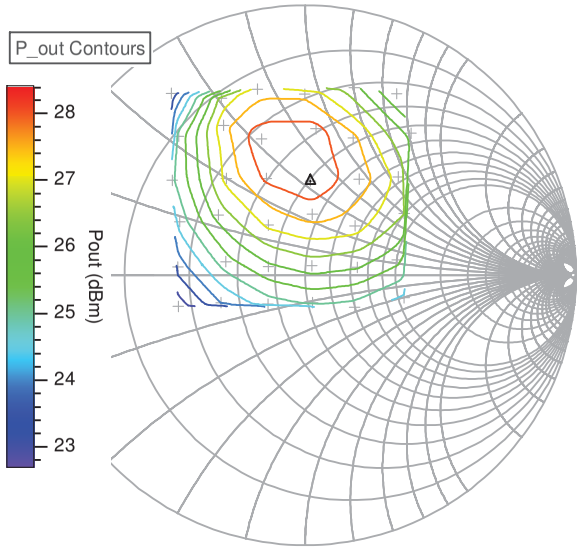


Figure 4-164, Measured 10GHz, 12V 80mA fundamental Pout contours. Optimum Pout 28.4dBm at 0.53/_108°, 40.3%.

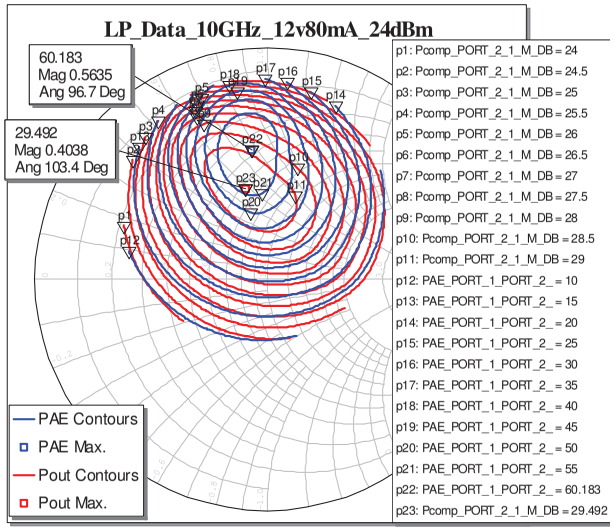


Figure 4-165, Simulated 10GHz, 12V 80mA fundamental PAE and Pout contours.

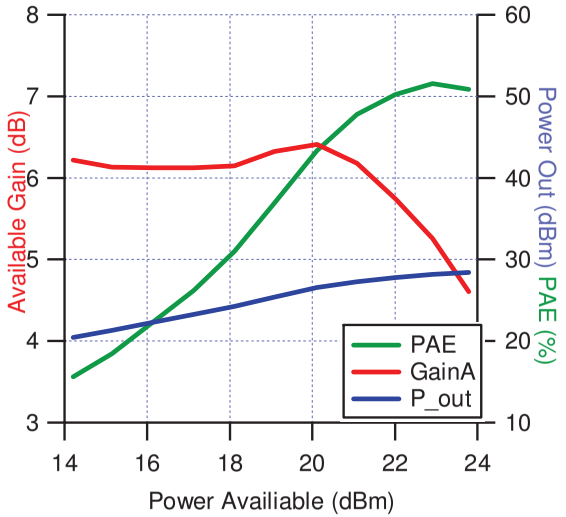


Figure 4-166, Measured 10GHz, 12V 80mA fundamental power sweep at optimum PAE load.

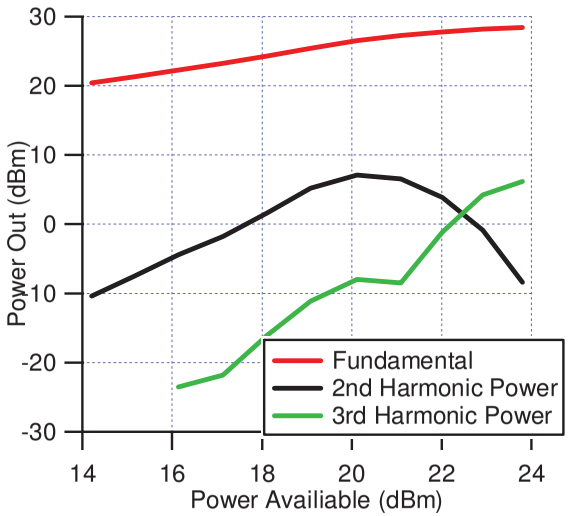


Figure 4-167, Measured 10GHz, 12V 80mA harmonics at optimum PAE load.

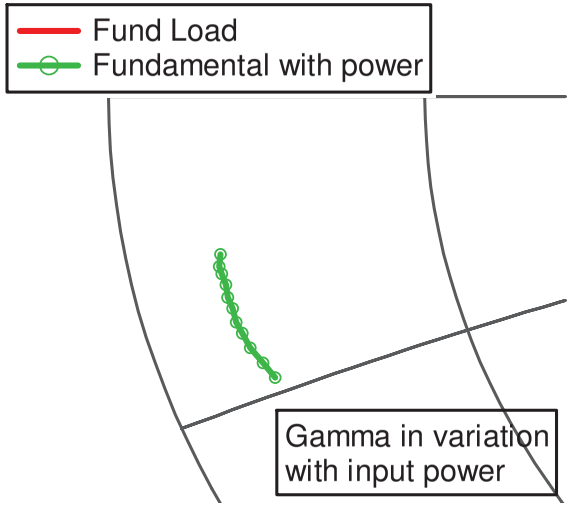


Figure 4-168, Measured 10GHz, 12V 80mA Γ_{in} at optimum PAE load. Γ_{In1}=0.90/_-168°, Γ_{In2}=0.89/_-158°, Γ_{L1}=0.54/_84.5°

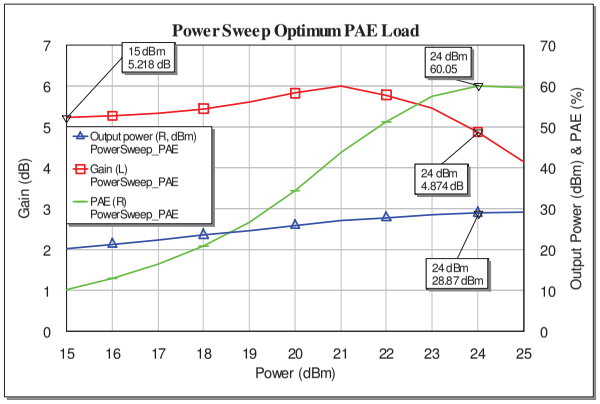


Figure 4-169, Simulated 10GHz, 12V 80mA fundamental power sweep at optimum PAE load.

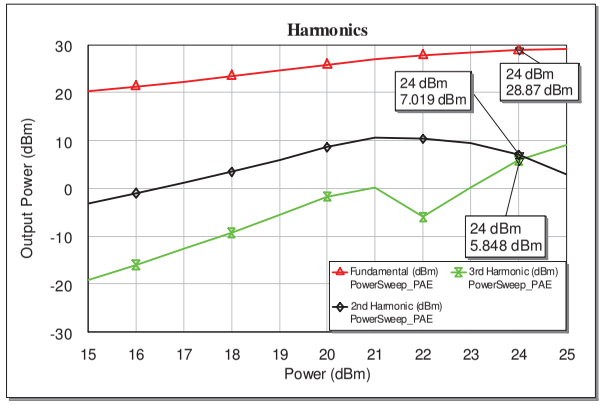


Figure 4-170, Simulated 10GHz, 12V 80mA harmonics at optimum PAE load.

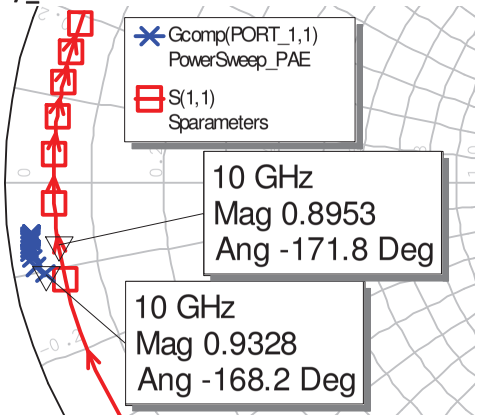


Figure 4-171, Simulated 10GHz, 12V 80mA Γ_{in} at optimum PAE load.

Looking at the swept power performance the PDK model is optimistic compared to this one device's efficiency and power performance; the 2nd harmonic agreement between simulation and measured is very good and the 3rd follows a similar pattern even if the absolute levels are not correct. The input reflection coefficient is higher than the measured, which makes the difference in Pout more surprising as more power is mismatched at the input.

In the final set of measurements the supply voltage was reduced to 10V and the quiescent current to 40mA. A common characteristic of high efficiency operation is the positioning of the 2nd harmonic at a short circuit in the CG plane. For class J operation (see chapter 1) the 2nd harmonic is moved around the edge of the Smith Chart, typically up to ±60° of the short circuit position. To demonstrate the capability of such a measurement at high frequency on the active load pull system, the 2nd harmonic was held at the short circuit position (at the device plane, which would correspond to a rotation of ~60°around the edge of the impedance plane) and a fundamental load pull was conducted. In effect this is the opposite of the earlier approach where we found the optimum fundamental load and then its corresponding optimum 2nd harmonic load; in this case we specify the 2nd harmonic and find the optimum fundamental.

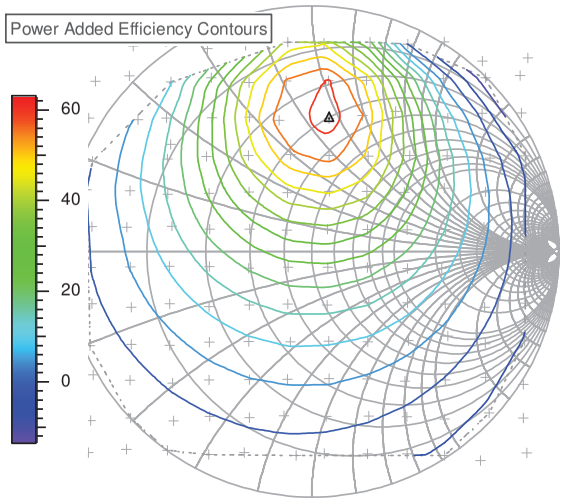


Figure 4-172, Measured 10GHz, 10V 40mA fundamental PAE contours, s/c 2nd harmonic. Optimum PAE 62.7% at 0.55/_82.6°, 26.4dBm.

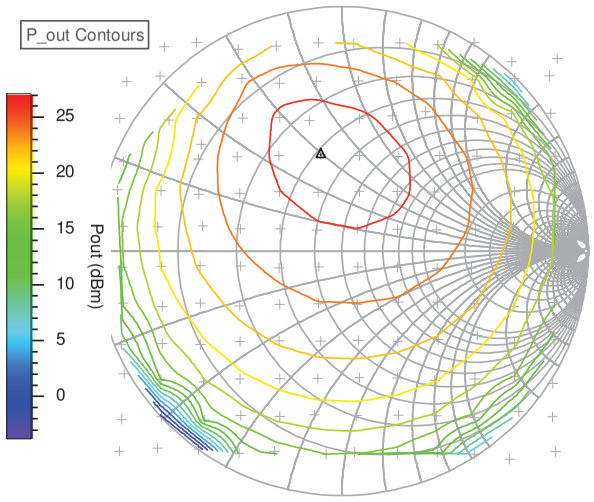


Figure 4-173, Measured 10GHz, 10V 40mA fundamental PAE contours, s/c 2nd harmonic. Optimum Pout 27.0dBm at 0.40/_102°, 54.3%.

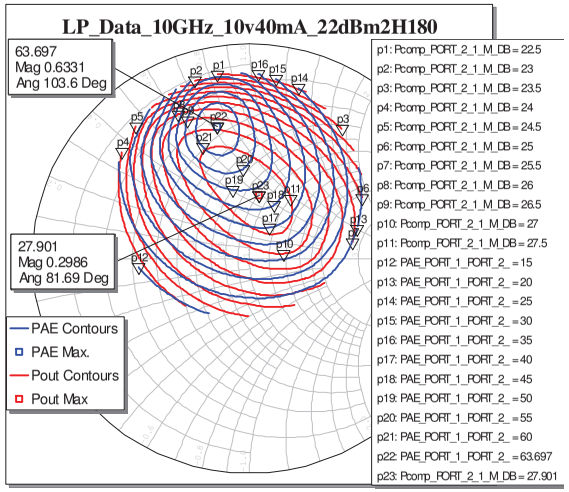


Figure 4-174, Simulated 10GHz, 10V 40mA fundamental PAE & Pout contours, s/c 2nd harmonic. Optimum PAE 63.7% at 0.63/_104 ° and optimum Pout 27.9dBm at 0.30/_81.7°.

Although the predicted PAE values are now similar, it is very noticeable that the phase of the optimum loads is different. In the PDK case Pout has the smaller angle, even allowing for the measured Pout optimum being between the two measured grid columns it would still have a greater angle than the PAE.

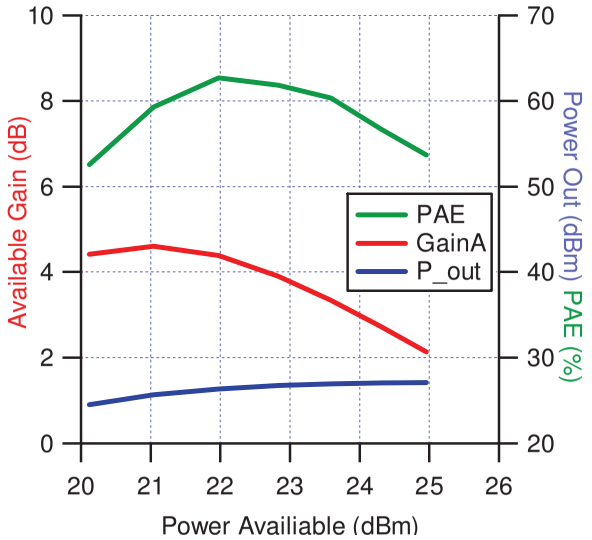


Figure 4-175, Measured 10GHz, 10V 40mA power sweep with 2nd harmonic s/c and fundamental at PAE optimum, Γ_{L1}=0.56/_84.1°.

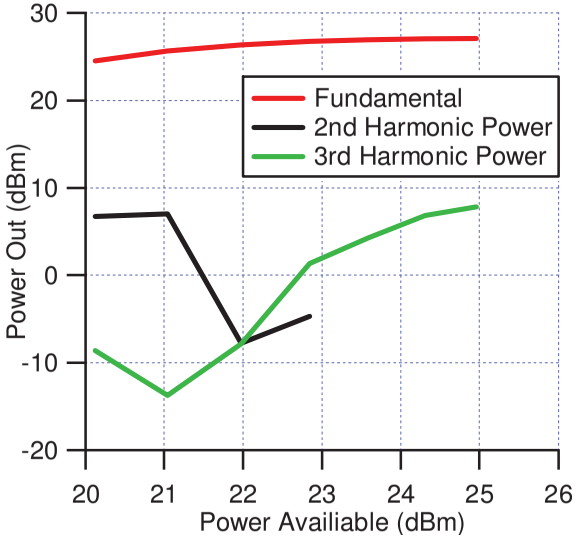


Figure 4-176, Measured 10GHz, 10V 40mA harmonics with 2nd harmonic s/c and fundamental at PAE optimum, Γ_{L1}=0.56/_84.1°.

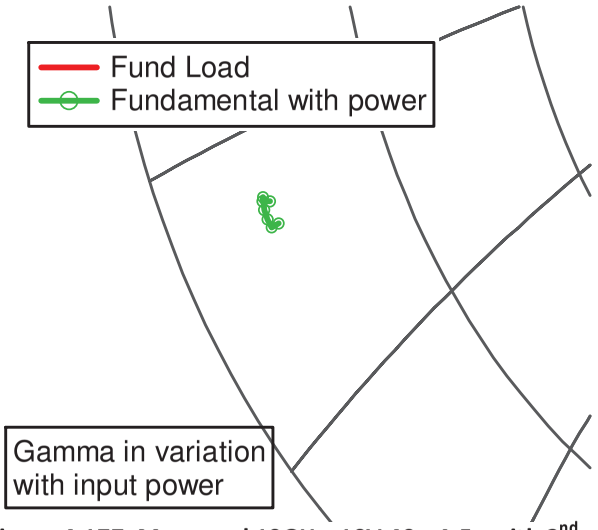


Figure 4-177, Measured 10GHz, 10V 40mA Γ_{In} with 2nd harmonic s/c and fundamental at PAE optimum, Γ_{In1}=0.87/_-152°, Γ_{In2}=0.86/_-150°, Γ_{L1}=0.56/_84.1°.

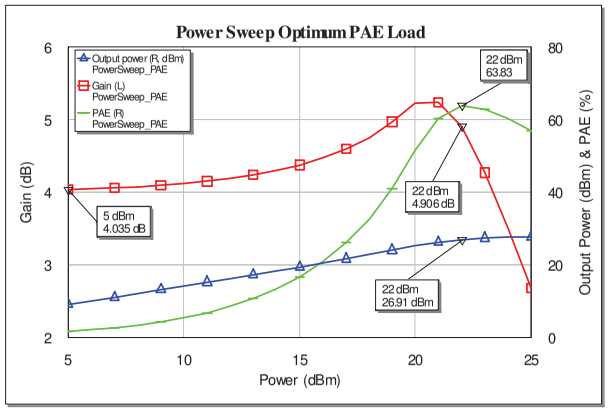


Figure 4-178, Simulated 10GHz, 10V 40mA power sweep with 2nd harmonic s/c and fundamental at PAE optimum, Γ_{L1}=0.63/_104°.

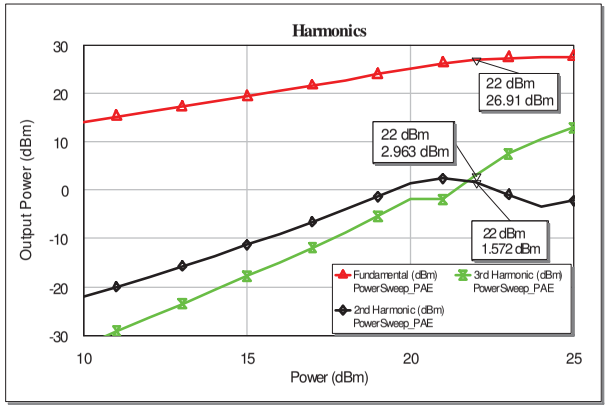


Figure 4-179, Simulated 10GHz, 10V 40mA harmonics with 2nd harmonic s/c and fundamental at PAE optimum, Γ_{L1}=0.63/_104°.

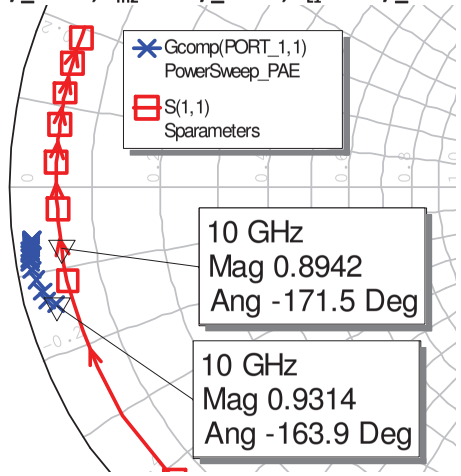


Figure 4-180, Simulated 10GHz, 10V 40mA Γ_{In} with 2nd harmonic s/c and fundamental at PAE optimum, Γ_{L1}=0.63/_104°.

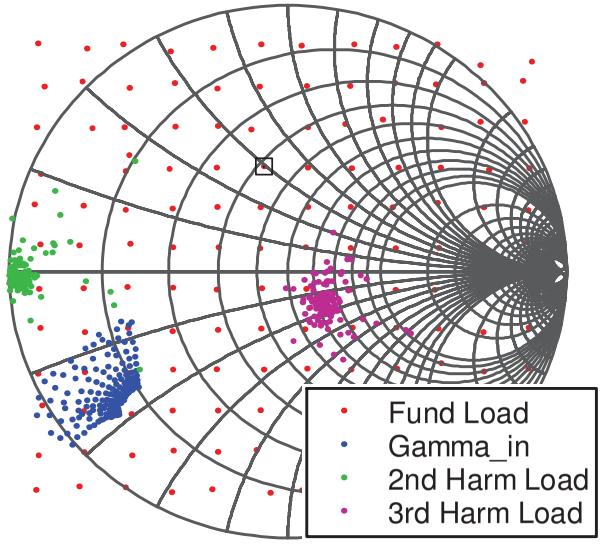


Figure 4-181, Measured Γ_{In} , Γ_{L1} , Γ_{L2} and Γ_{L3} at 10GHz

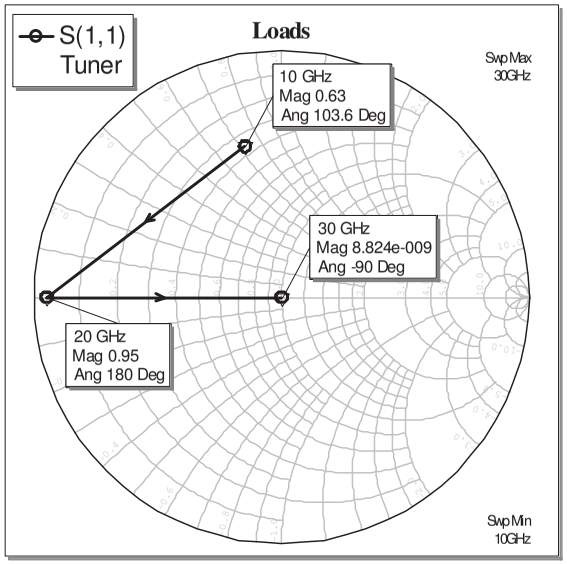


Figure 4-182, Simulation Γ_{L1} , Γ_{L2} and Γ_{L3} at 10GHz

The measured Γ's show a number of a number of relevant factors;

- The input reflection coefficients show the translation of the load impedance to the input.
- The fundamental load grid is reasonably regular.
- The 2nd harmonic grid although concentrated around the s/c point still has a number of occasions when it has failed to converge and hence a measurement has been recorded which must be checked at each fundamental load point.
- The 3rd harmonic is a reflection of the system impedance at 30 GHz and the noise level/dynamic range.

The effect of the 2nd harmonic short circuit (s/c) termination is to increase the optimum PAE by 1.7% and the associated output power by 0.5dB. The s/c termination had no impact on the optimum output power itself, remaining constant at 27 dBm, but the efficiency at this point has increased by 4.9%. The optimum PAE reflection coefficient has moved anti-clockwise by ~15°, whilst the optimum output power has rotated by ~30°. The behaviour (load rotation) was also replicated in the simulation; however the optimum PAE was in fact lower in the simulation with the s/c 2nd harmonic termination.

Although as can be seen in Figure 4-181, the majority of the 2nd harmonic terminations during the measurements were grouped around the s/c position there were a number of occasions when the 2nd harmonic was a long way off from the target load. Examining each point in the power sweeps recorded in Figure 4-175 to Figure 4-177 the specific 2nd harmonic load reflection coefficient was noted and recorded in Table

4-18. At the optimum PAE level the 2nd harmonic was very close to the short circuit, 0/_180°, however above this point Γ_{L2} varied significantly with both magnitude and phase. As there was no full impedance plane 2nd harmonic load pull conducted at 10GHz, it is difficult to estimate the impact of these deviations; the fact that the 2nd harmonic PAE improvement was not significant suggests that these will not cause a major change. The measurements made at 8GHz, Figure 4-145, indicate that the overall effect is likely to be < ±4%. The results show that although the system can be used, for example, for measurement of devices in high efficiency modes at X band, improvements need to be made to improve the harmonic load convergence and stability.

2nd Harmonic Load		Performance			
Mag	Ang	Power	Pin	Pout	PAE
0.72	162.9	0	20.1	24.5	52.6
0.70	167.8	1	21.0	25.6	59.3
0.98	-179.8	2	22.0	26.4	62.7
0.93	174.8	3	22.8	26.7	61.8
1.20	-172.6	4	23.6	26.9	60.3
0.84	94.0	5	24.3	27.0	56.7
1.08	162.7	6	25.0	27.1	53.7

Table 4-18, Actual 2nd harmonic termination at 10GHz, 10V, 40mA with fundamental at Γ_{L1}=0.56/_84.1°.

4.7 Input Pre-matching

As has been seen the input reflection coefficients of the type of devices measured at high microwave frequencies are large and hence significant amounts of power are wasted. This limits the ability of the system to drive transistors into compression and this problem is exacerbated as the frequency increases. Although input tuners can be used, for on wafer measurements these have to be mounted away from the device and the connecting cables add to the loss. If they are used between the device and measurement couplers the tuner needs extensive calibration, on the other hand if they are inserted before the couplers, i.e. outside of the calibration region, there is increased phase dispersion of the device input match with frequency. They are also expensive pieces of equipment that need to be controlled/optimised at each frequency. An alternative solution is the use of input pre-matching circuits on the wafer with the device itself. Obviously this cannot be implemented until the first design 'spin', however it does allow an increase in the frequency range when limited by the available drive power. Also if a broad band matching technique is used the actual accuracy of the device input match is less critical in the design of the pre-matching circuit. A small improvement in input can make a significant improvement in input power, Figure 4-29. Such an approach was included in the first design iteration.

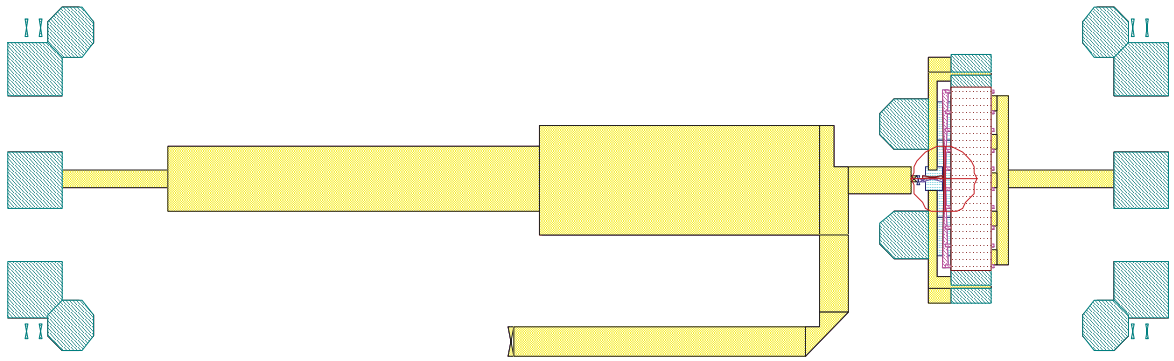


Figure 4-183, 20GHz Pre-matching for 10x75 device.

The input matching was based upon the measured small signal (+10dBm) S parameters of the 10x75 device. A relatively straight forward microstrip matching circuit was designed to improve the input match at 20GHz, as shown in Figure 4-183. The anticipated performance of this circuit is indicated by Figure 4-184 which shows the anticipated performance of the design. Note that although the 8dB input return loss corresponds to an increase in input power of about 7dB, Figure 4-29, this has to be offset against the insertion loss of the matching circuit of 0.9dB (from the linear circuit prediction).

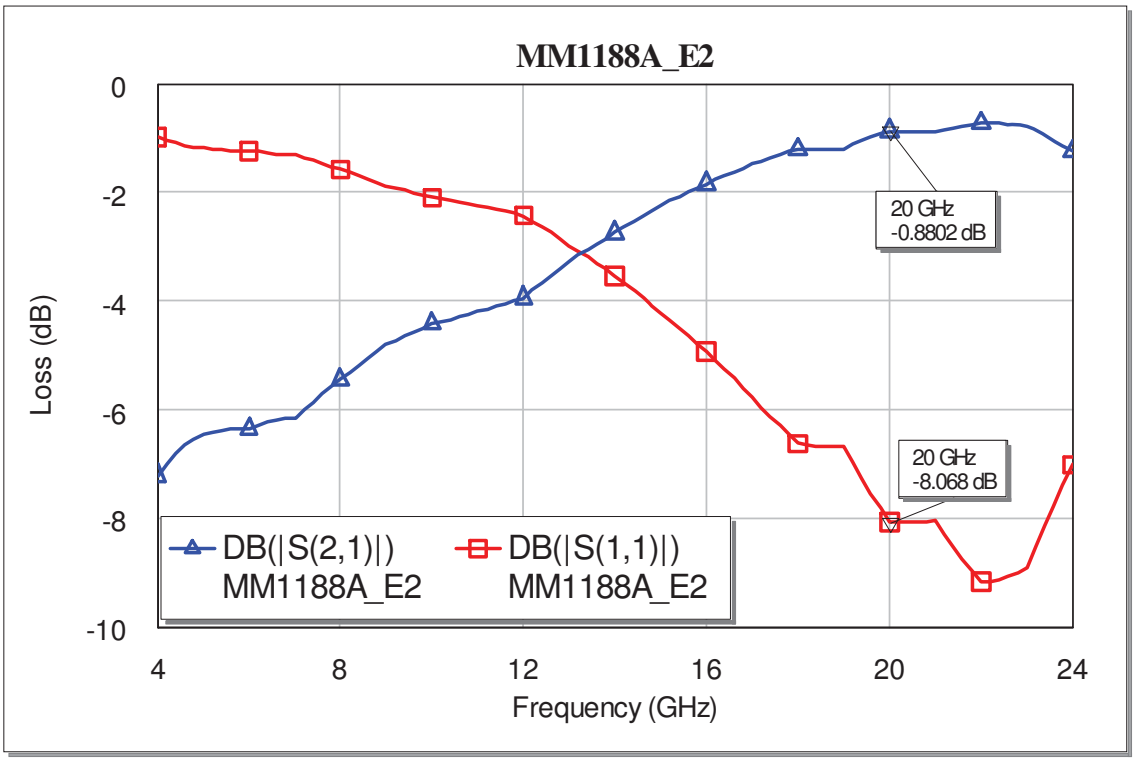


Figure 4-184, Simulated input match and insertion loss of pre-matching

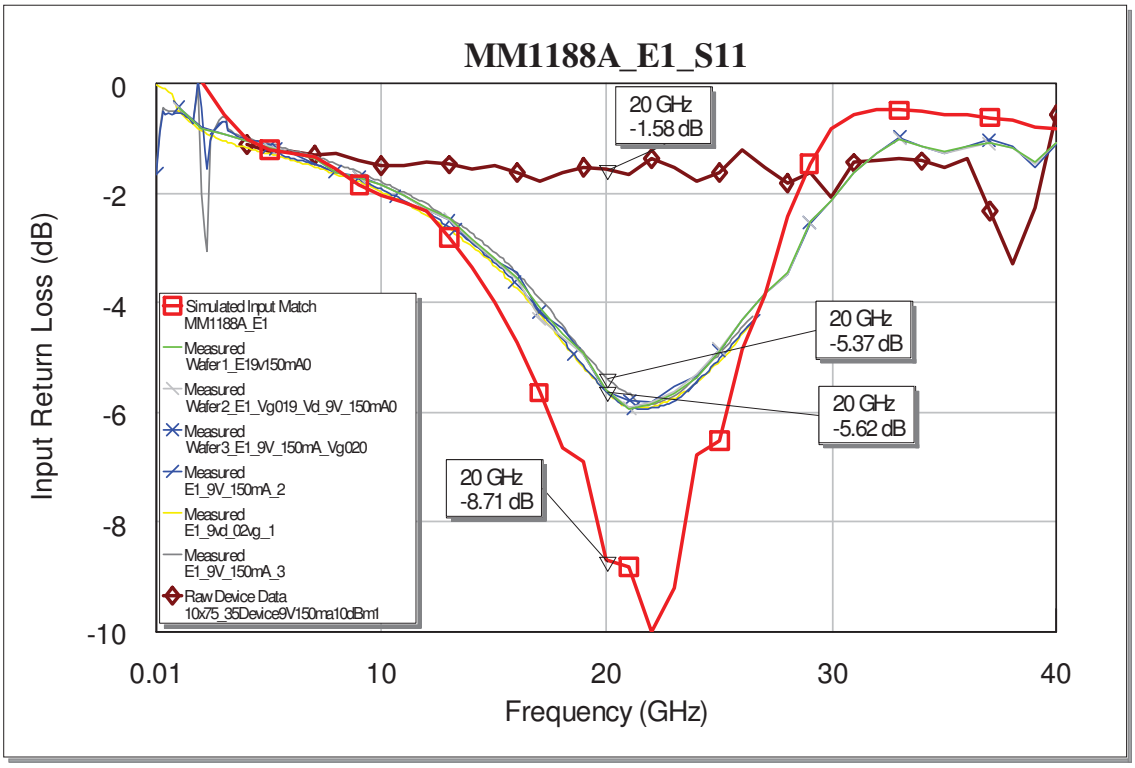


Figure 4-185, Response of input matching circuit, simulated IMAT with measured device S parameters (red), measured IMAT with biased device (multiple devices from wafer), biased device input match (brown).

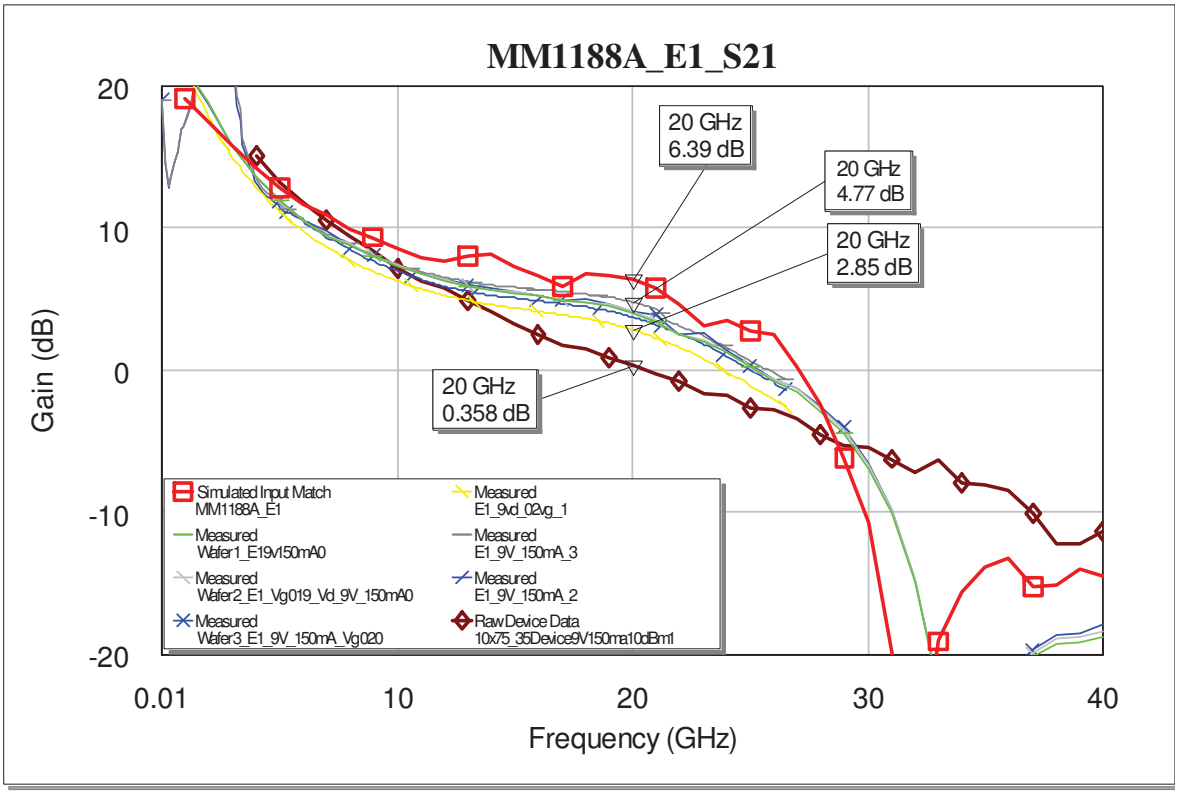


Figure 4-186, Measured gain response of input matching circuit, simulated IMAT (red) with measured device S parameters, measured IMAT with biased device, biased device S21 (brown).

The manufactured wafer was diced and tiles were tested on a vector network analyser. This included unmatched devices. A device from the new batch was measured and the S parameter data imported to the simulator and analysed with the predicted response of the input matching circuit, (red trace of Figure 4-185 and Figure 4-186). This showed the input match was very close to that of the original data. The response of the device on its own is shown in these figures (brown trace) and the Input Matching circuit (IMAT) with biased device was measured on 6 devices and the results compared. The input return loss at 20GHz ranged from 5.4 to 5.6dB as compared to 8.7dB for the simulated circuit. The actual shape and frequency centring of the response was very similar. The gain of the IMAT devices varied from 2.9dB to 4.8dB as compared to the simulated 6.4dB. These results would suggest that the insertion loss of the IMAT is higher than the predicted 0.9dB by of the order of 2dB. This seems high; alternatively the devices could be being compressed by the increased input power which is lowering the gain. Unfortunately insufficient data at 20GHz was available on the device to see if this was the case. However even if the loss was 2dB higher than expected, this still produces a worst case improvement of 2.5dB over the unmatched device which is a reasonable extension in the drive power range of the system. Load pull measurements could not be done on the devices at this stage as the un-backed tiles would not be able to handle the power dissipation. The pre-matched 10x75 device (connected between the wafer probe pads and feed lines) is shown in Figure 4-187.

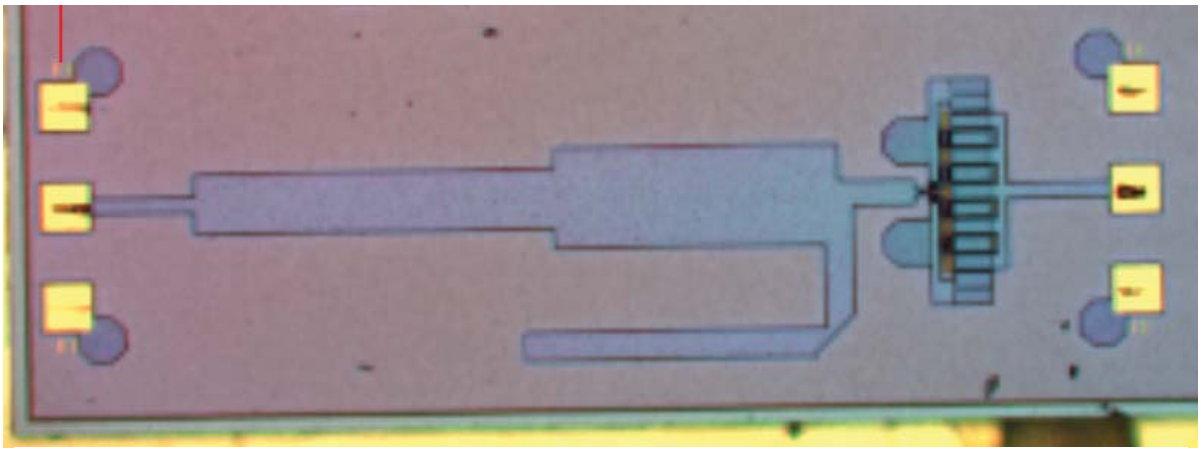


Figure 4-187, Input pre-matching circuit and 10x75 device.

4.8 Summary

This chapter has described the initial measurements undertaken using the high frequency active load pull system. The ability to conduct fundamental load pull up to 18GHz on a 0.5W on wafer device was shown. The results on the test device showed that although the PAE declines with frequency, maximum output power remains relatively constant, which is consistent with the 6dB/octave gain roll-off of this type of device. The ability to conduct up to 3 harmonic active load pull was demonstrated at 4 and 6GHz, whilst fundamental and 2nd harmonic measurements were carried out at frequencies up to 12GHz. The performance of the measurement system was limited by the lack of phase coherent sources for 3 harmonic load pull outside of the range 6.67-9.9GHz and 2 harmonic load pull outside of 6-10GHz. Nonetheless this performance marked a significant increase in the capability of the device characterisation systems available to researchers.

During the period of the research the measurement system control software underwent some significant changes (implemented by other researchers within the group), not least of which was the reduction in measurement file size from a measurement folder of an average size of 100MB and containing several thousand individual files for a 100 point grid and 10 power levels, to a single file of under 800kB for a 57 point grid and 10 power levels. These changes did of course introduce some issues in terms of data compatibility throughout the project, with analysis features present in earlier versions not always migrating to the later ones immediately.

The measurement system proved the usefulness of the design approach recommended by this research. Not only were the achievable performance levels observed, but confidence increased in the reliability and robustness of the chosen devices. It was also observed that after driving the device into heavy compression the gate voltage had to be increased (closer to 0V) in order to achieve the same I_{ds}. This indicates a certain amount of aging going on and might suggest that some burn in required pre-test (biased with enough RF to saturate the transistor) to avoid measuring devices which are still changing.

A behaviour that had not been observed (as far as this researcher is aware) is the opposite rotation of the 2nd harmonic maximum and minimum PAE load impedances and the PAE and output power contours that were observed at 10GHz, the approximate frequency where the contra rotating minima and maxima would coincide. This would benefit from further investigation; with improved phase coherent sources and the improved

measurement speeds both a finer frequency step and grid spacing could be implemented. It would also be useful to extend the investigation to devices of different sizes, different f_ts, and power levels (output capacitance) and see if the behaviour is consistent and upon what parameters does it depend so as to suggest the mechanism. Inevitably much work has been done elsewhere on determining how to achieve the maximum performance and it appears little has been found that investigates the detrimental factors to device behaviour, particularly PAE.

An area in need of more detailed research is to investigate the class A 2nd harmonic optimums more closely; having shown that the load impedances producing the highest PAE lie on the edge of the real impedance plane it would be useful to concentrate the impedance sweep along this edge to accurately determine the optimum point at each frequency. This could then be used to confirm the trajectory follows that of the output capacitance.

The final set of measurements worth more study would be that of the 2nd and 3rd harmonic load pull contours, but de-embedded to the CG plane. The simulation of the PDK model suggests that the 2nd harmonic load pull contours at this plane, shown in Table 4-6, are relatively straight lines across the impedance plane rotated from the horizontal, whilst the 3rd display a curved shape.

4.9 Bibliography

[1] I. Bahl, "0.7 - 2.7GHz 12W Power Amplifier MMIC developed using MLP technology," *IEEE Transactions on Microwave Theory and Techniques,* vol. 55, pp. 222-229, February 2007.

[2] P. Ladbrooke, N. Goddship, J. Bridge and D. Battison, "Dynamic I(V) measurement of contemporary semi-conductor devices," *Microwave Engineering Europe,* pp. 23-33, May 2000.

[3] P. Smith, W. Kopp, P. Ho and et.al., "Ku-Band high frequency high gain pseudomorphic HEMT," *Electronics Letters,* vol. 27, no. 3, pp. 270-271, 1991.

[4] B. Virdee, A. Virdee and B. Banyamin, Broadband Microwave Amplifiers, Artech House, 2004, pp. 133-135.

[5] R. Sakamoto, A. Kiyoyasu and I. Masataka, "Real-space transfer and Hot-electron transport properties in III-V semiconductor heterostructures," *IEEE transactions on Electron Devices,* vol. 36, no. 10, October 1989.

[6] J. Moon, J. Kim and B. Kim, "Investigation of a class-J power amplifier with a nonlinear Cout for optimised operation," *IEEE Tranactions on Microwave Theory & Techniques,* vol. 58, no. 11, pp. 2800-2811, November 2010.

[7] A. Grebennikov, RF and Microwave Power Amplifier Design, McGraw-Hill, 2004.

[8] R. Saini, S. Woodington, J. Lees, J. Benedikt and P. Tasker, "An intelligence driven active load pull system," in *75th Microwave Measurements Conference (ARFTAG)*, 2010.

[9] A. AlMuhaisen, P. Wright, J. Lees, P. Tasker, S. Cripps and J. Benedikt, "Novel Wide Band High-Efficiency Active Harmonic Injection Power Amplifier Concept," in *International Microwave Symposium*, Anahiem, 2010.

[10] S. Cripps, RF Power Amplifiers for Wireless Communications, 2nd Ed., Artech House, 2006.

[11] R. Davis, "GaAs FETs - physics, design and models," in *Handbook of RF and Microwave Power Amplifiers*, J. Walker, Ed., Cambridge, Cambridge University Press, 2012, pp. 57-60.

[12] J. Verspecht and D. Schreurs, "Measuring transistor dynamic loadlines and breakdown currents under large-signal high-frequency operating conditions," in *IEEE MTT-S International Symposium Digest*, 1998.

[13] J. Walker, The Handbook of RF and Microwave Power Amplifiers, Cambridge: Cambridge University Press, 2011.

5 Novel Device Modelling Techniques

"Before we can design a transistor circuit, we need to be able to model all important characteristics of the device we are using, and we need to know how the model works and precisely what it represents. Given a choice between an unmodelled device with super performance and a well-modelled one with more prosaic performance, I'll usually choose the latter because I'll almost always get better results with it. Perhaps I'm not very adventuresome, but I like to know what I'm doing. A device may have wonderful potential, but if I don't know how to realize it, I probably won't be able to."

Steve Maas, Why I hate base resistance, IEEE Microwave Magazine, June 2004

5.1 Introduction

The purpose of this chapter is to explain how the measurements made in chapter 4 and using the measurement system described in chapter 3 can be utilised in modern nonlinear CAD packages. As Maas says in the quote above, you can have the 'best' device, but it is the detailed knowledge of how it behaves and being able to replicate this that enables the practical application. Traditionally measured S parameters and pulsed DC I-V data are used to construct small signal and large signal equivalent circuit models. Small signal models are used when the transistor is operating in the linear region, i.e. for class A bias and at power levels well below (typically >10dB) the 1dB compression level. By contrast the large signal model is designed to represent the behaviour in the nonlinear region, such as in the higher efficiency classes of operation and at power levels where the gain has started to compress. Often simulators will incorporate both large and small signal models, the application of which being input power level dependant; this can lead to accuracy issues in the 'handover' region. Device modelling is a complicated process and takes considerable time and effort, inevitably leading to compromises, meanwhile the device processing and technology developments continue apace and as a result models are often left playing 'catch-up' or trying to replicate processes whose physics may not be fully determined or even understood.

The circuit models comprise of elements that are divided between "extrinsic" and "intrinsic". Extrinsic components are associated with device connections to the outside

world, for example source via holes and gate and drain feed lines and bonding pads. Intrinsic elements are representative of the internal features of the device such as the drain channel resistance or cross element capacitances. These elements can be inferred from S parameter measurements under different bias conditions, commonly referred to as 'hot' and 'cold'. Large signal models require 'best fitting' or optimisation of low frequency pulsed DC I-V measurements and the elements derived from the S parameter measurements. Load pull systems have used since the days of valve amplifiers and even active load pull has been around for 35 years [1], [2], but it has proved difficult to incorporate the load impedance information derived from such measurements directly into a CAD model. Instead these measurements have often been used in parallel with equivalent circuit based models [3] to 'fine tune' the desired loads information, but this then makes the nonlinear simulation results inconsistent. Part of the difficulty in the past in obtaining the raw data from which to generate large signal models has been the lack of measurement systems capable of acquiring simultaneous magnitude and phase data of the fundamental and harmonic signals. In the nonlinear region the sinusoidal input signal to a device is distorted by the generation of harmonics of the input frequency within the transistor due to either the voltage or the current clipping, and by the parametric variation of the intrinsic capacitances due to the changing voltages.

Both Large and Small Signal models are highly dependent upon the accuracy of the measurement systems used to acquire the data used to characterise the device. In the case of large signal high efficiency designs this includes the harmonic frequencies (some argue up to the 7^{th} although in practice controlling beyond the 3^{rd} in a design is the best that can be achieved) and thus the measurement system needs to have a bandwidth capability several times greater than the operating frequency of the device. Additionally, due to the practical limits on companies in creating device models and the power limitation on measurement systems, a route of scaling device models based on smaller geometries (such as number of gate fingers or their width) is followed. In this case any error in the smaller measured device is multiplied as the geometry is increased, and also some features like the via hole to finger ratio do not scale. Thus accuracy of the amplifier simulation is highly dependent on the device model and hence the final performance of the physical amplifier. In the early days of CAD approximations were made including using S parameters to get the device in the right 'ball-park', however this contained no harmonic data. In 1983 Cripps [4] produced the

classic paper on the prediction of load pull contours, which in a number of variants [5] and developments became the mainstay of power amplifier design; again, however this was not a model that could be integrated into CAD or deal with the increasingly complex harmonic terminations required in the drive to higher efficiency amplifiers. Nonlinear models required large amounts of processing time and were thus difficult to use in optimisation routines in simulators, hence a combination of small signal models (S parameters) to get a design in the approximate region and large signal models to estimate output power, were often used [6] in the design of power amplifiers.

The ability to construct device measurements, as described in chapter 4, replicating actual operating conditions; bias, impedance environment, drive level, etc. does not in itself lead to the solution to a particular design requirement. Arguably the best place to do this is within the CAD simulator where the designer can model the 'real world' performance of the components used to create the required matching (both linear and electro-magnetic) relatively quickly and with the ability to optimize values to meet particular requirements. The confidence in the simulator results is very much based upon the quality of the models used and the detail contained within the emulation of the circuit construction. The old adage, "rubbish in – rubbish out" is highly applicable. Thus the importance of good quality nonlinear models and the starting point for these must be measured data.

5.2 Measurements

The early models were based upon fitting equivalent circuit models to measured S parameter data, although similar in many aspects to the measurement system described in chapter 3 these measurements are made in a fixed 50Ω environment. The approach is still important for measuring passive structures such as device packages, impedance transformers and feed networks, and so from both a historical perspective and for completeness small signal S parameter measurements will be briefly described here.

Small signal models are almost exclusively based upon measurements made on Vector Network Analysers (VNAs). These instruments use couplers to sample the forward and reflected waves at the test ports. The sampled signals are down-converted using tuned receivers, whose local oscillators are phase locked to the fundamental frequency. Thus the measurements are relative in both magnitude and phase to the primary source. The errors

inherent in VNA measurements are usually categorised into three groups; systemic, random and drift.

Systemic errors are those which are consistent and therefore repeatable, which means that they can be identified, quantified and corrected. In a full two port S parameter measurement such errors are caused by coupler directivity, source and load mismatch, transmission and reflection tracking and finite port isolation. The accuracy of the measurements depend to a large extent on the quality of the calibration conducted to remove the unwanted and repeatable signals which are also measured by the tuned receivers, such as breakthrough between coupler ports, reflections from discontinuities, particularly at transitions, and non-ideal transmission lines in the connecting cables.

Random errors are by definition those whose occurrence cannot be predicted, such as noise and connector repeatability (although careful attention to connector care and torque, and the use of quality components can minimise this latter problem). The noise comes from two sources, the low level system noise floor of the receivers and the higher level phase noise of the local oscillators. To remove these effects narrower IF filtering and increased averaging are used, however both of these increase measurement time.

Drift errors are either a result of temperature changes or of the frequency sources. Temperature changes can cause errors in both phase and magnitude. Drift errors require recalibration and because this often involves disconnecting and reconnecting the unit under test, introduces their own difficult to quantify errors. Hence the best option is to maintain a constant temperature (including ensuring instruments have had a sufficient period to reach a thermal equilibrium) and to use instruments with good frequency stability and phase noise performance.

Calibration tends to be carried out at a point which is convenient for testing such as at coaxial connections or wafer probe tips; this may not however be the plane from which the measurement data needs to be extracted. Further, the devices themselves may be mounted in test fixtures, or even where the bare device is mounted such that it can be probed there is often a feed line in between the launch and the device, Figure 5-1. Thus a method is required which can remove the physical and parasitic effects surrounding the device.

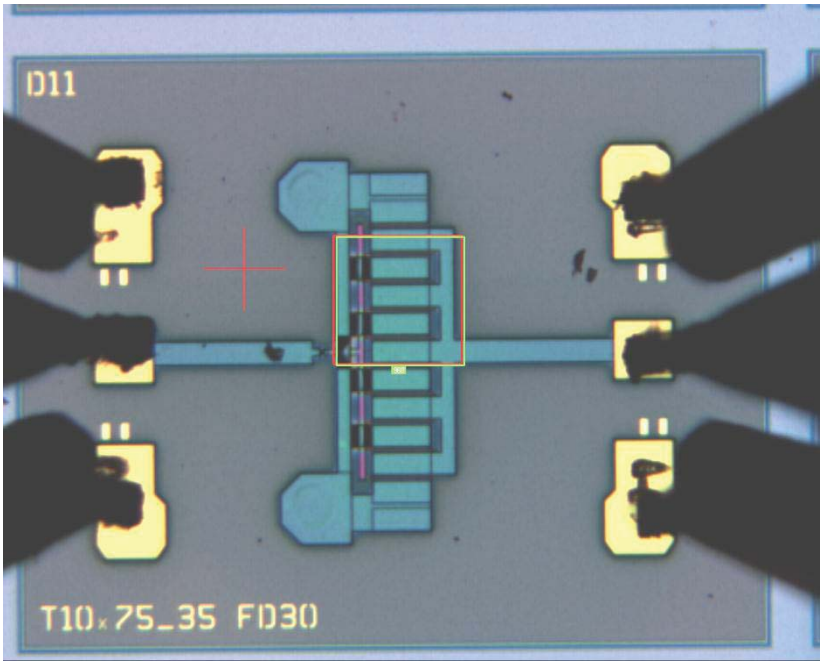


Figure 5-1, Test device with wafer probes and feed line.

In order to do this a process known as de-embedding is necessary. In its simplest form this may consist of using a negative circuit element in the simulation of the measured data, for example a negative shunt capacitance to remove bond pad capacitances. For more complicated de-embedding a number of circuit elements can be treated as a whole. This process is one by which the transmission and reflection properties of the structures between the calibration plane and the measurement plane are replicated and removed from the measurement. This may be done by measuring the structures involved which are specially laid out on a separate cell or by modelling them (for example by means of an Electro-Magnetic simulator as they are passive elements) and then removing their effect through matrix manipulation using Transmission or T parameters [7]. It should be remembered that it is not always necessary to remove all of the parasitic components around a device, for example in the device shown in Figure 5-2, in its practical application it will be necessary to bond to the chip, therefore the bond wires can stay in the measurement data, although as will be discussed later it may be useful to be able to analyse the currents and voltages at a plane within the device itself and for this it will be necessary to the 'remove' all of the external structures. The through lines between the probe and the bond wires are however purely for measurement purposes and thus need to be removed from the measured data.

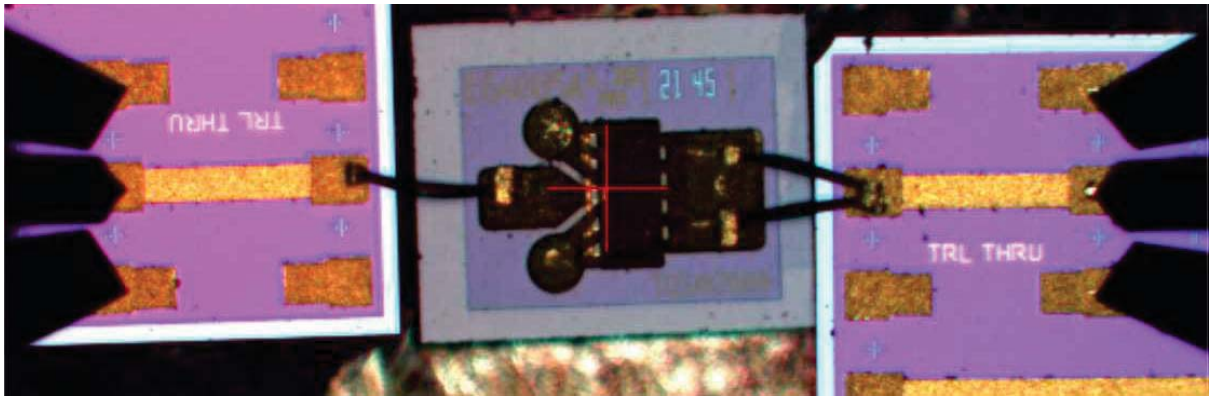


Figure 5-2, Device with bond wires and feed line. ***{Picture courtesy of Selex Galileo Ltd.}***

In the ideal world one would exactly replicate the way in which the device was physically used in practice, e.g. mounted on the same carrier with the solder or epoxy attach method, however this is not always practical, particularly from the device manufacturers perspective where they have no control over the implementation of the device by their customers, who are possibly using them in numerous different applications. It is also desirable to conduct as many measurements as possible at the wafer level as it is easier to handle larger substrates than individual chips and automated measurements can be conducted on many different devices with a single set-up. Wafer mapping builds up a picture of performance variation and yield. On-wafer measurement does however limit the power levels that can be handled, hence the tendency to carry out measurements in pulsed mode. This in itself brings out another possible source of error, the electrical characteristics of the device are temperature sensitive and in pulsed mode the device may have a different temperature profile to normal operation.

Returning to the issue of calibration itself, coaxial calibration itself is relatively simple; short, open and matched (typically 50Ω) loads of the same coaxial types are connected at the reference plane. In practice these standards are not perfect, but their deviation from the ideal is known and held in a calibration file on the VNA. Usually the standards are themselves calibrated to traceable national standards; this being done at regular intervals and the calibration file updated. This approach has been extended with the development of automatic calibration standards or electronic standards. In this case a single connection is made to the calibration box and various known (but not necessarily opens, shorts or 50Ω) loads are switched internally. For through connections on two port calibrations if the device is insertable (port connectors of opposite genders or genderless such as APC-7) then the two test ports are connected together and the through loss and

phase measured. If they are non-insertable a calibrated adapter must be used of known transmission properties. There are numerous calibration approaches used; for coaxial measurements one of the most common is Short, Open, Load, Through, (SOLT), however due to the difficultly in producing a consistent open circuit termination in wafer probe measurements a Through, Reflect, Match (TRM) approach, Figure 5-3, is more common. In coplanar wafer probe calibration the short and through lines are simple to fabricate, the load is commonly produced by using two 100Ω thin film resistors in parallel. This not only maintains the symmetry of the structure but halves the parasitic components. The open circuit is surprisingly more difficult, (one could expect that the probes could just be raised above the substrate), but at the discontinuity higer order modes are generated and these modes have different transmission characteristics. Good practice dictates [8] that where the raised probe approach is used the probes should be kept at least two wavelengths apart, which may be impractical (at 4 GHz λ_{air} = 7.5cm). Note that this applies to using an open

Figure 5-3, Wafer probe calibration standards, through, short and matched load.

circuit as a calibration standard, it is however advisable to use an open circuit as a rough check that a standard calibration has been implemented correctly, by observing the impedance of the open on a Smith Chart. Some calibration substrates do contain an element that they refer to as an open circuit, however this is more of a know mismatch and it needs to be clear and consistent how and where the other port (wafer probe) is terminated for consistency of calibration.

5.3 Equivalent Circuit Models

The earlist approximations of transistors for use in Computer Aided Design (CAD) software used standard circuit elements such as resistors and capacitors alongside current sources. The current sources generally being controlled by the voltage across another

element. To determine the element values devices would be tested under off or 'cold' conditions and on or 'hot' conditions. For FETs, the 'Cold' S parameters are measured with the gate voltage biased at pinch-off and 0 volts on the drain. 'Hot' S parameters are typically measured at a drain current 50% of Idss and the normal operating drain voltage. Care should be taken to ensure that the input power level to the device is at least 10dB below the 1dB compression point of the device, (but high enough for a good measurement signal to noise level). The measured data is optimised in a CAD simulator to fit one of the standard device equivalent circuit models. A problem can be the initial starting values; it is necessary to monitor and steer the optimiser to avoid the simulator finding a solution that is totally unrelated to the physical model of the device. For example, parasitic resistances that tend to 0Ω. Initial 'guesses' at some of the component values can be made based on the device physical properties, such as the inductance and resistance of source vias or gate bond wires.

A standard small signal model for a GaAs FET is shown in Figure 5-4, the intrinsic parameters are shown in the shaded box. When the device is biased in pinch-off there is effectively no current generator and a passive circuit model can be used to represent the device [9] as shown in Figure 5-5. The capacitors C_a, C_b and C_c represent the fringing capacitances due to the depletion layer extension at each side of the gate under high negative bias. The inductors and resistors represent the parasitic resistance and delay of the connections to the gate, drain and source channel. These structures can of course be considerably more complex, consider for example the source via connections; the number of these proportional to the number of gate fingers is not constant and as the number of fingers increases the phase length between each one and the vias is not constant. Also they can change considerably when devices are scaled up (the number and width of gate fingers increased), with device cells effectively being replicated in parallel. This is one reason why a more simplistic model of the these elements is used, a more complex model would require changes for different device sizes. By measuring the S parameters these can be transformed to Z parameters from which the element values can be calculated.

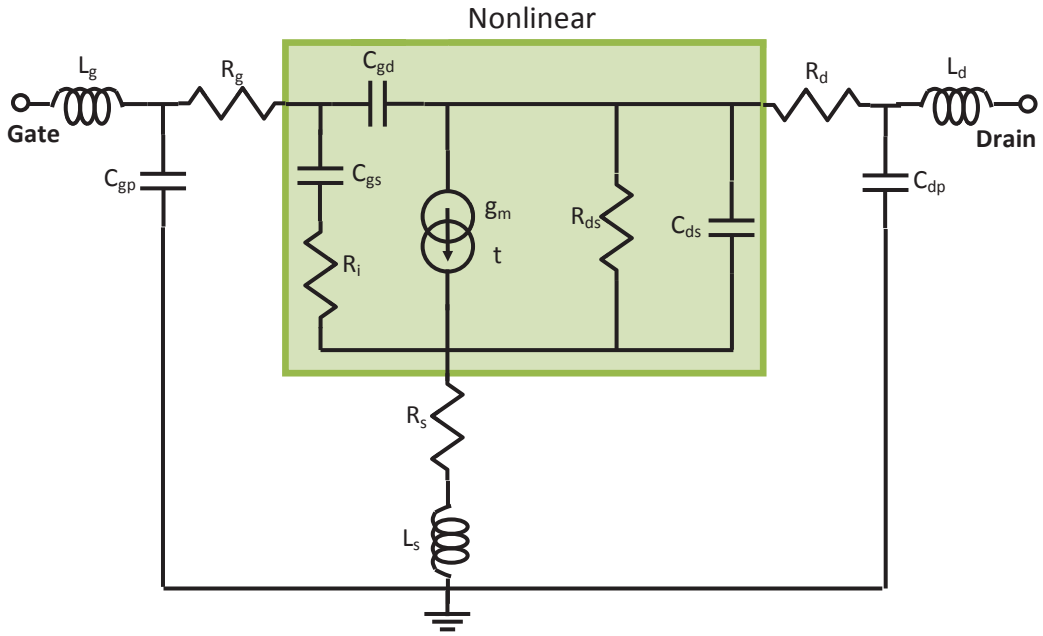


Figure 5-4, Small signal equivalent circuit model

Circuit Element	Description	Typical Value (0.35µm 0.5W pHEMT)
C_{gd}	Gate – Drain capacitance	0.05pF
C_{gs}	Gate – Source capacitance	1.03pF
R_i	Gate – Source resistance	0.97Ω
g_m	transconductance	200mS
t	transit time	2.5ps
C_{ds}	Drain – Source capacitance	0.13pF
R_{ds}	Drain – Source resistance	155Ω
C_{gp}	Gate parasitic capacitance	0.02pF
C_{dp}	Drain parasitic capacitance	0.06pF
L_g	Gate inductance	0.01nH
L_s	Source Inductance	0.04nH
L_d	Drain inductance	0.03nH
R_d	Drain resistance	0.65Ω
R_s	Source resistance	0.28Ω
R_g	Gate resistance	1.2Ω

Table 5-1, Calculated values for 10 x 75 µm pHEMT based on "Cold "and "Hot" S parameter measurements.

$$Z_{11} = R_g + R_s + j\left\{\omega(L_g + L_s) + \frac{1}{\omega C_{ab}}\right\} \quad \{5\text{-}1\}$$

$$Z_{12} = Z_{21} = R_s + j\left\{\omega L_s - \frac{1}{\omega C_b}\right\} \quad \{5\text{-}2\}$$

$$Z_{22} = R_d + R_s + j\left\{\omega(L_d + L_s) - \frac{1}{\omega C_{bc}}\right\} \quad \{5\text{-}3\}$$

where,

$$\frac{1}{C_{ab}} = \frac{1}{C_a} + \frac{1}{C_b} \quad \textit{and} \quad \frac{1}{C_{bc}} = \frac{1}{C_b} + \frac{1}{C_c} \quad \{5\text{-}4\}$$

hence,

$$R_g = Re(Z_{11} - Z_{12}) \quad \{5\text{-}5\}$$

$$R_s = Re(Z_{12}) = Re(Z_{21}) \quad \{5\text{-}6\}$$

$$R_d = Re(Z_{22} - Z_{21}) \quad \{5\text{-}7\}$$

$$\omega Im(Z_{11}) = \omega^2(L_g + L_s) - \frac{1}{C_{ab}} \quad \{5\text{-}8\}$$

$$\omega Im(Z_{12}) = \omega^2 L_s - \frac{1}{C_b} \quad \{5\text{-}9\}$$

$$\omega Im(Z_{22}) = \omega^2(L_d + L_s) - \frac{1}{C_{bc}} \quad \{5\text{-}10\}$$

The resistance values can be calculated directly from equations {5-5} – {5-7}. For the

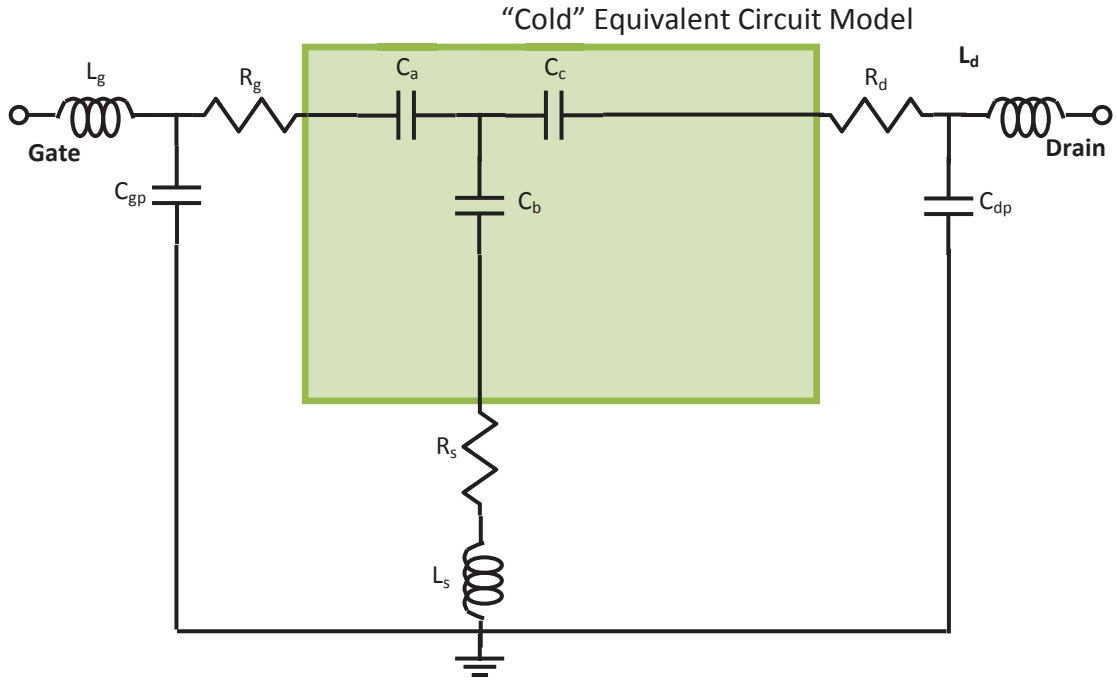


Figure 5-5, "Cold" equivalent circuit model, device biased at pinch off.

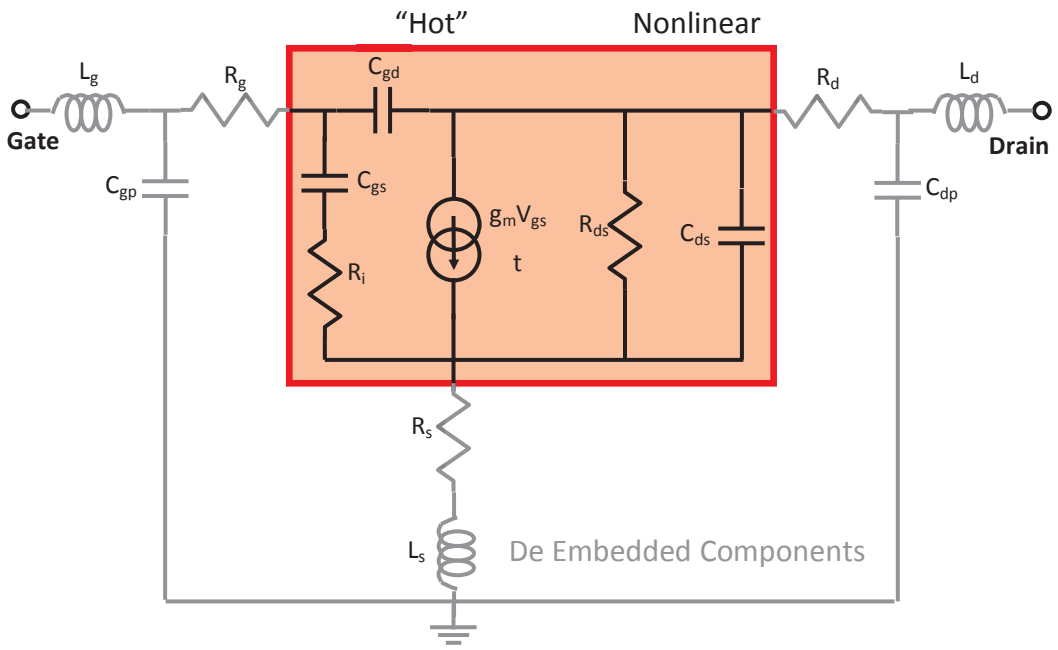


Figure 5-6, "Hot" equivalent small signal model, device biased at the normal operating point.

inductors it is necessary to plot *ωIm(Zxx)* against ω^2, the inductance being the gradient, and the capacitances C_a, C_b and C_c are determined from the constants. Once these values have been calculated they can then be used to de-embed to the intrinsic nonlinear elements of the "Hot" model, Figure 5-6.

The device S parameters are measured at the normal operating point and then de-embedded to remove the effects of the extrinsic components. The various circuit elements can then be calculated from the Y parameters as shown in the following equations:

$$Y_{11} = \frac{R_i {C_{gs}}^2 \omega^2}{D} + j\omega\left(\frac{C_{gs}}{D} + C_{gd}\right) \qquad \{5\text{-}11\}$$

where,

$$D = 1 + R_i {C_{gs}}^2 \omega^2 \qquad \{5\text{-}12\}$$

$$Y_{12} = -j\omega C_{gd} \qquad \{5\text{-}13\}$$

$$Y_{21} = \frac{g_m e^{-j\omega\tau}}{(1 + j\omega R_i C_{gs})} - j\omega C_{gd} \qquad \{5\text{-}14\}$$

$$Y_{22} = G_{ds} + j\omega(C_{ds} + C_{gd}) \qquad \{5\text{-}15\}$$

Note that the model is no longer bilateral as Y_{12} does not = Y_{21}. Separating the real and imaginary parts the component values can be calculated by:

$$C_{gd} = \frac{-Im(Y_{12})}{\omega} \quad \{5\text{-}16\}$$

$$C_{gs} = \frac{Im(Y_{11}) - \omega C_{gd}}{\omega}\left\{1 + \frac{[Re(Y_{11})]^2}{\left[Im(Y_{11}) - \omega C_{gd}\right]^2}\right\} \quad \{5\text{-}17\}$$

$$R_i = \frac{Re(Y_{11})}{\left\{\left[Im(Y_{11}) - \omega C_{gd}\right]^2 + [Re(Y_{11})]^2\right\}} \quad \{5\text{-}18\}$$

$$g_m = \sqrt{[Re(Y_{21})]^2 + \left[Im(Y_{11}) - \omega C_{gd}\right]^2\left[1 + \omega^2 {C_{gs}}^2 {R_i}^2\right]} \quad \{5\text{-}19\}$$

$$t = \left(\frac{1}{\omega}\right)\sin^{-1}\left\{\frac{-\omega C_{gd} - Im(Y_{21}) - Re(Y_{21})\omega C_{gs} R_i}{g_m}\right\} \quad \{5\text{-}20\}$$

$$C_{ds} = \frac{Im(Y_{22}) - \omega C_{gd}}{\omega} \quad \{5\text{-}21\}$$

$$R_{ds} = \frac{1}{Re(Y_{22})} \quad \{5\text{-}22\}$$

As can be seen, this process involves substantial effort in order to achieve a model that is only valid in the small signal region and for a specific bias. Arguably a quicker, simpler and as (if not more) accurate method is to measure the small signal S parameters. This process is relatively easy to automate and by including multiple S parameter data blocks within a single ".mdf" file (measurement data format), the data can easily be selected within a CAD file, Figure 5-7. The equivalent circuit model is useful in determining equivalent circuit component values, which aid in the design of matching circuits or understanding where device limitations come from, for example the effects on bandwidth of the output capacitance. However it should be remembered that these are approximations of the actual device physical structure, which by their nature are distributed rather than discrete. For example the transconductance for a GaAs MESFET can be described [8] in terms of the saturated carrier velocity, v_{sat}, gate width, w, semiconductor dielectric constant, ε_s, and the depletion layer depth, h_d, as shown in {5-23} or in terms of the extracted parameters as given by {5-19}.

$$g_m = \frac{v_{sat} w \varepsilon_s}{h_d} \quad \{5\text{-}23\}$$

True large signal S parameters pose a problem and therefore cannot be collected and utilised in the same way. As has been said devices are nonlinear and therefore contain frequencies in addition to the source, S parameters only contain fundamental frequency

information. There is a further practical problem in measuring the S_{22} of the device. As the characteristics are input drive level dependant it is necessary to drive the input whilst at the same time stimulating the output port with the signal to measure S_{22}. This therefore has to be different to the source frequency, which requires an offset Local Oscillator. The S_{22} measuring signal will mix with the main output causing reverse intermodulation products, necessitating careful selection of the receiver filter bandwidths. For these reasons significant effort has been put into the development of nonlinear analysis techniques and large signal modelling.

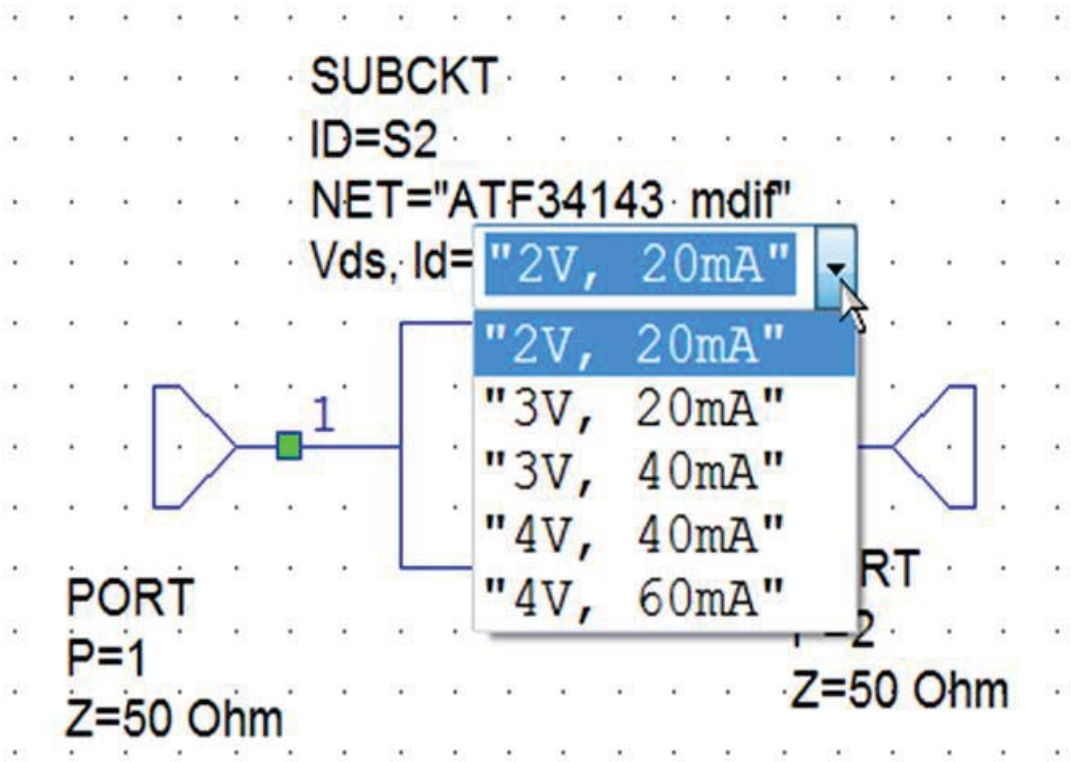


Figure 5-7, Using "mdif" format within a circuit simulator for a small signal device under different bias conditions

5.4 Large Signal Modelling

In order to utilise the model discussed earlier (Figure 5-4) in a large signal simulation the elements whose characteristics are bias and RF signal level dependant are commonly defined by equations rather than static values. These are largely related to the changing size of the drain channel which affects the capacitances C_{gd} (gate – drain), C_{gs} (gate – source) and C_{ds} (drain – source) and the drain source resistance R_{ds}, and also the breakdown characteristics, most commonly represented by diodes, Figure 5-8. As these features of the

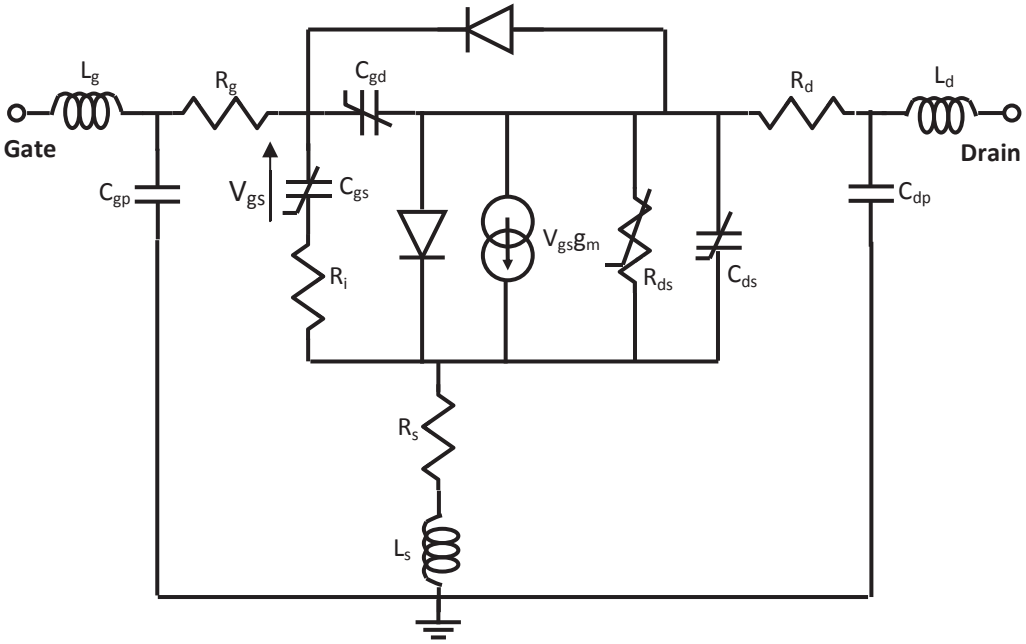


Figure 5-8, Large signal device model incorporating nonlinear and breakdown elements.

device are related to physical properties they can be derived from the geometry and materials of the parts, however such a theoretical approach needs to be proved by measurement and so, as the measurements need to be done anyway it is more typical to use an empirically based model. Before considering how the model is created it is worth reviewing the most common nonlinear analysis techniques in which they will be used.

5.4.1 Nonlinear Analysis Techniques

When the output from a network is not directly proportional to the input it is termed nonlinear. For RF and microwave design linear analysis is largely conducted in the frequency domain using data in S parameter format. This allows numerous blocks to be connected in series or parallel and an overall S parameter matrix to be constructed, thus allowing fast simulation. By contrast nonlinear analysis requires the solving of complicated multiple equations. There are four main approaches:

a) Time Domain
b) Harmonic Balance
c) Volterra Series
d) Describing Functions

a) Time Domain: This uses a series of nonlinear equations with respect to time. The relationship between the voltage and current in the time domain for each element can be

solved using Kirchhoff's Laws. Solutions are found in steps of Δt which is inversely proportional to solution time and accuracy. For higher frequency analysis smaller Δt increments are required and hence the higher the frequency, the more elements there are, the greater the circuit complexity and the more computation required to achieve a solution. One of the difficulties is that in order to reach a steady state solution all transients must be calculated, even though it may only be the steady state condition that is of interest. Some parts of circuits, such as bias circuits, have a much greater settling time than the RF networks, causing lengthy calculations. Also, sometimes for a given time step a solution to Kirchhoff's Laws cannot be found and the circuit fails to converge to a solution. For RF and microwave frequency design the major limitation is the ability to handle immitance parameters as the majority of elements within simulators such as SPICE are lumped elements or ideal transmission lines.

b) Harmonic Balance: This operates in both the time and frequency domains and is the method employed by virtually all commercial nonlinear simulators [10] and so will be described in more detail than the other approaches. It is based upon initially applying a hypothetical solution to a system and then measuring the error between this solution and the rules applicable to the system, for example Kirchoff's Current Law. The variables determining the result are adjusted and the resulting error function calculated, this is repeated until the error function reaches zero or an acceptably small value. The circuit is divided between its linear and nonlinear components, with the former being solved in the frequency domain up to the n^{th} harmonic (decided by the level of accuracy and simulation speed required), and the latter being solved in the time domain. The results are exchanged between the two using Fast Fourier Transforms (FFTs) and Inverse FFTs. Simulations usually run linear – nonlinear – linear and this cycle repeats until the error function between the time and frequency domains is less than a predetermined limit. Initial starting values are obtained using a simplified set of parameters ignoring some of the nonlinear effects, more complex simulators analyse circuit changes since the last run and determine whether or not to use previous values. The speed of the analysis depends upon the type of circuit being processed, the partitioning between the linear and nonlinear elements, the initial conditions and the number of harmonics used in the calculation. There are a number of approximations that are made that the designer needs to be aware of that impact the

accuracy of the simulation. The phasor equivalents of nonlinear waveforms consist of an infinite number of terms; however these are truncated to a finite number. The impact of higher order terms is circuit dependant; band limiting elements reduce the impact of higher order terms. The higher the nonlinearity of the waveform the more terms needed to approximate it. To solve the nonlinear currents the voltage phasors are converted to the time domain using Fourier transforms. Nonlinear elements are easier to describe in the time domain and so the nonlinear currents are calculated in this way, these currents are then turned into current phasors. These domain transformations introduce inaccuracies due to aliasing[1]. This can be reduced by increasing the number of frequencies, but this in turn increases simulation time. An alternative solution is by oversampling. The lower sampling limit is 2H (the Nyquist limit), where H is the number of significant frequencies. To reduce aliasing the number of samples can be increased above this to, for example, 2nH. After evaluating the nonlinear current a Fourier transform converts the current to the time domain it will consist of nH frequency components. The more nonlinear the waveforms the higher the oversampling required. In nonlinear simulations it is therefore necessary to decide how far to truncate the analysis frequency range and what level of oversampling to use. It is important to understand that the Harmonic Balance approach avoids the need to be able to directly analyse a nonlinear circuit, instead a hypothetical solution is proposed and as long as it is possible to measure the difference between this solution and the laws governing the circuit operation (the error function) then by successive adjustment of the hypothetical solution it may be possible to reach an acceptable solution, - convergence [11]. For accurate results it is necessary that the model of the device itself be an accurate reflection of its behaviour up to the n^{th} harmonic. Unfortunately the suppliers of the models rarely provide data showing how the model prediction performs against measured results or what the maximum frequency the model was optimised to was.

c) Volterra Series: This approach is good for analysing weakly nonlinear circuits and multiple input signals; however it does not handle higher harmonics well. Originally the Taylor Series was used to approximate a nonlinear response, however whilst this was able to deal with output responses to a single input at a particular time, it could not resolve behaviour that

[1] The operation of a Discrete Fourier Transform (a method of determining and FFT) has the potential to 'fold' the highest frequency content of a signal onto the lowest frequency, this is known as aliasing.

also depended upon the state of a system prior to the stimulus under consideration. This was a major drawback particularly when dealing with 'memory' effects seen in electronic circuits. The Volterra Series approach overcame this restriction, however the limitations imposed by the inclusion of high level, high order harmonics in mixers and power amplifiers has led to the Volterra approach being largely replaced by Harmonic Balance in commercial nonlinear CAD software.

d) Describing Functions: These are generally used where the nonlinearity is low. The nonlinear behaviour is converted into a number of linear systems – usually filters – and analysed in the frequency domain. The accuracy depends upon the filter error between the nonlinear and the equivalent linear systems. Thus the major disadvantage for more nonlinear networks is the increasingly large number of linear circuits that are necessary to describe their operation.

Similarly the methods of creating nonlinear models can be separated into four approaches;

i) Empirical: this approach relies upon fitting mathematic expressions to observed behaviour such as from measured DC I-V curves. It is not necessary for the equations used to have any relationship to the physical characteristics of the device [12]. There is also a question as to how far the measured DC performance is representative of the high frequency behaviour, and one set of measurements may not be sufficient to give a picture of the spread of performance, i.e. statistically where on the bell curve they lay. The new modelling technique using directly measured device voltages and currents originally called the Direct Look-Up Table Model (DLUT) and commonly now referred to as the Cardiff Model, can be considered as empirical as it is entirely based on measurements. The DLUT model is significantly different from the other modelling approaches as it does not attempt to fit the model to the data, but instead the data is the model. This approach is described in more detail later.

ii) Analytical: this uses equations to describe physical behaviour. However not all the physical mechanisms controlling charge transport are fully understood [8]. Also the accuracy of the model is reduced by any assumptions made in creating the equations.

iii) Semi Empirical: this uses a mixture of characterisation measurements and process parameters such as geometry and doping levels to determine the device equivalent circuit elements. This is a compromise between the empirical and analytical approaches and is probably the most common approach currently.

iv) Numerical: these are based on the actual device physics and are computationally very intense. These are most commonly used by foundries where the emphasis is on maximising parameters such as f_T or minimising parasitic capacitances rather than utilising the model within a circuit environment.

The latest development in modelling has been the introduction of X parameters and the more general Poly Harmonic Distortion (PHD) models. These are a mix of Empirical and Analytical as they are directly measurement based but rely on the formation of multi term equations where the coefficients are calculated from the measured data. These will be returned to later.

Attempting to create accurate, physically correct descriptions of features can introduce far more problems than using a simple empirically derived formula [13], hence the Semi Empirical approach has probably been the most common, and as a major part of the research in this thesis concerns comparisons with this approach it will be described in more detail. This approach has been developed for over 30 years, with the first empirical time domain model appearing in 1980, [14] and the first frequency domain model in 1981, [15]. Some of the circuit element characteristics are directly linked to process controlled parameters such as the pinch off voltage V_p, and the gate Schottky voltage, whilst others are arbitrary and determined by curve fitting. Examples often named after their lead developer(s) are the Curtice [14], Curtice-Ettenberg [16], Statz [17] and the TOM (Triquint Own Model) [18]. A model may provide a better approximation of a particular process or attempt to deal with factors such as frequency dispersion or temperature. Often there is a compromise between capturing a particular behaviour, overall accuracy or complexity (and hence simulation effort).

The element values of the models are optimised to fit the measured pulsed DC-IV curves. Pulsed measurements are used as they provide a better representation of the RF performance, mainly due to the self-heating effects if true DC measurements are made. A weakness of this approach is that harmonic effects are not a part of the model. Model

development has generally been a progressive refinement with each 'generation' attempting to tackle what were seen as deficiencies in previous models. For example, the Curtice-Ettenberg model introduced a diode between gate and drain which reproduced the drain gate avalanche current that occurs at high input drive levels. Similarly, the initial MESFET models used basic diode capacitance equations; these were further developed by Statz [17] and in the TOM3 model [19]. Typically they were developed in SPICE (Simulation Program for Integrated Circuit Emphasis), a time domain tool developed at the Electronics Research laboratory, University of California, Berkeley, which lead to the use of circuit elements as the natural building blocks of the model. Models that fit well in the mid region of the V_{gs} /I_d curve do not necessarily track in the region near pinch-off; hence some models are better for class A/B or B operation. Similarly, models tend to avoid some of the more complex charge related physics of the device which the manufacturers endeavour to reduce, such as surface charge and trapping[2]. There is also a discussion regarding how the model behaves at edges of the operating envelope. In amplifier design it is good practice to use a device within the safe operating area with a reasonable margin. Thus it can be argued that there is little need to have accurate modelling of the device behaviour in the breakdown regions, suffice it to indicate where there regions are. Researchers and device developers may however wish to understand not so much where breakdown occurs but what the mechanisms involved are.

The TOM model is to some extent a general purpose model and hence it has become popular, despite for particular processes the Curtice-Ettenberg model being a more accurate prediction of, for example, the gate voltage - drain current relationship. The main features of the TOM model are:

- Simple method of fitting g_m as a function of V_{gs} as the device changes from a square law dependence on V_{gs} to a linear relationship.
- Gradual pinch-off characteristic (i.e. the magnitude of V_p increases as V_{ds} increases).
- Ability to simply model R_{ds} dependence on V_{gs}, V_{ds} and the channel temperature.
- The ability to produce families of DC I-V curves that cover V_{gs} ranging from V_p to the forward bias voltage.

[2] Traps are due to imperfections in the device material lattice that show up as discrete energy levels in the material bandgap. They are thus difficult to quantify as they are unpredictable in location and density.

The latest version of the model is referred to as the TOM3 model [19]. This version improves the equations used to model the intrinsic capacitances by using quasi-static charge conservation in the implanted layer of the MESFET. It has not been possible to directly measure the intrinsic gate charges or capacitances; the model derives the gate charge from the drain current and the gate capacitances from the drain conductances. This reliance on the drain current therefore puts the emphasis on the accuracy of the equation for calculating I_{ds}.

An additional factor that needs to be considered is the dynamic behaviour of the device; this refers to how the previous or initial conditions the transistor experiences affect the behaviour and are thus sometimes referred to as memory effects. When a device is in a quiescent condition and it experiences a significant change, for example a significant increase in RF input power, the change does not result in a new steady state condition instantly. This can be clearly seen during long pulsed RF signals such as are generated in radar transmitters. Initially the rising edge of the pulse may 'overshoot' and be followed by ringing. The top of the pulse will rarely be flat, tending to 'droop' with pulse length; this is due to the falling supply voltage as charge is depleted from the local capacitors. However close examination of the rising edge of an optimally driven pulsed transistor [20] shows that before the 'droop' sets in there are regions of different output power gradients due to the different thermal time constants within the device, as shown in Figure 5-9.

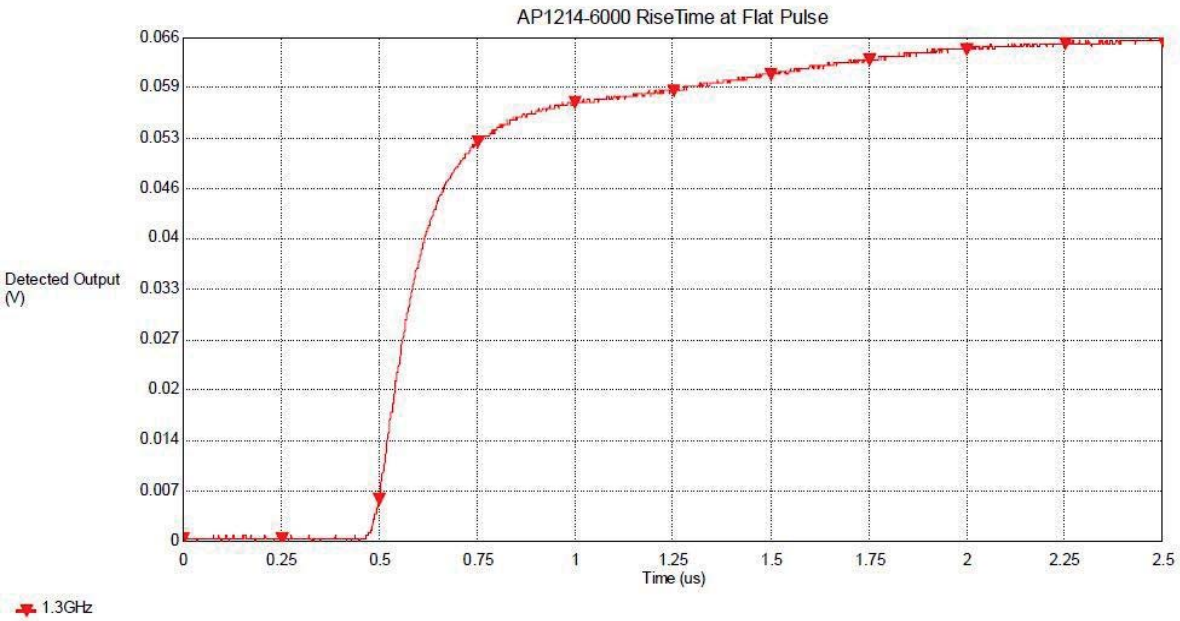


Figure 5-9 , Rising edge and initial pulse shape of an optimally driven solid state radar transmitter.

In these pulsed transistors a multitude of different effects come into play, which results in the overall pulse response. For example, with input drive level the device input impedance alters, the current drawn, gain and power out increase; as a result the junction temperature increases, due to the differing time constants, and as the efficiency changes during the rising pulse edge a non-uniform temperature profile results. In order to model these effects either equation based or electro thermal models must be created which have a significant dynamic range. For transistors biased in class A, as the device is driven in to compression it becomes more efficient, Figure 5-11[3], the junction cools which increases the gain and output power, whilst for a class B or C operating transistor the increasing drive turns the device on, so whilst efficiency increases so does the heating in the junction. At first sight it may appear that modelling these thermal behaviours would be a more understandable problem than intricate electrical models of device performance, as those that have looked into the issue have found [21], the more the problem is examined the more complicated it becomes!

[3] Measurements conducted by the author on a Milmega 1-2 GHz 30W module.

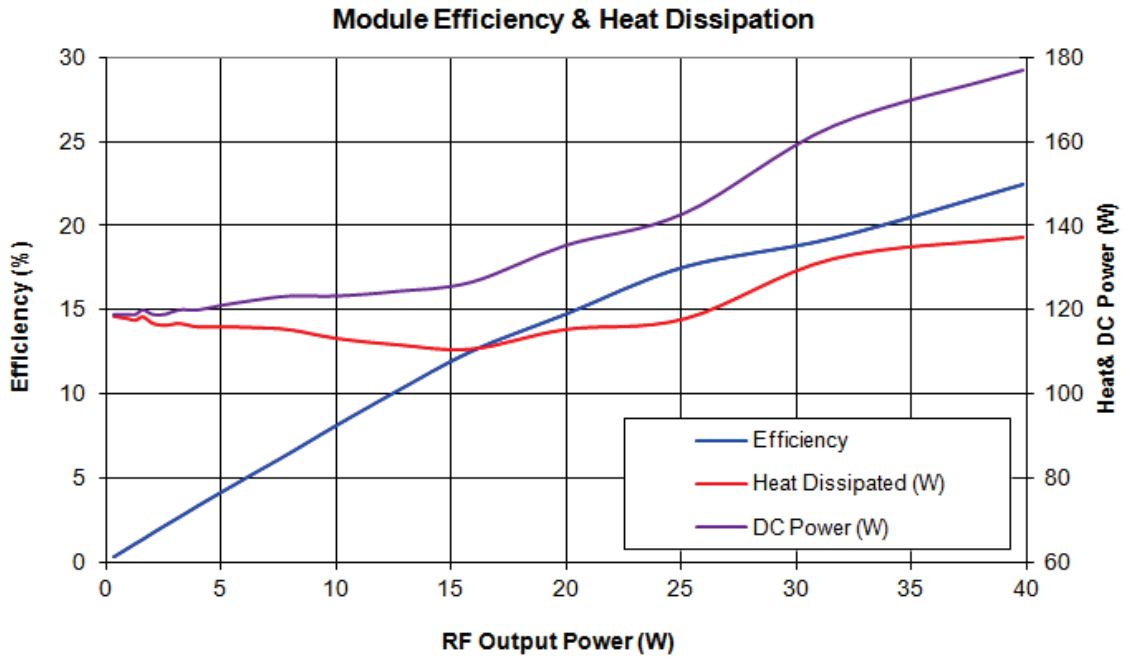


Figure 5-11, Efficiency and Heat Dissipation in a Class A 1-2 GHz 30W (nom.) Amplifier

As has been discussed earlier a key part of the standard approach to model development is the use of DC-IV measurements. It has been shown [22] that a better representation of the RF performance of the device is to use pulsed or dynamic DC-IV measurements. However the results obtained vary dramatically with the initial device bias as shown in Figure 5-10 [23]. These differences have been largely attributed to trapping of charge within the device structure and on the surface. The build-up of charge depends upon the electric fields present and temperature and can be due to either electrons or holes. Each of these has a different time constant; electron capture tends to be relatively fast, whilst electron emission is slower. These will therefore produce different effects with different signal modulation schemes. It is clear that models need to incorporate data taken at

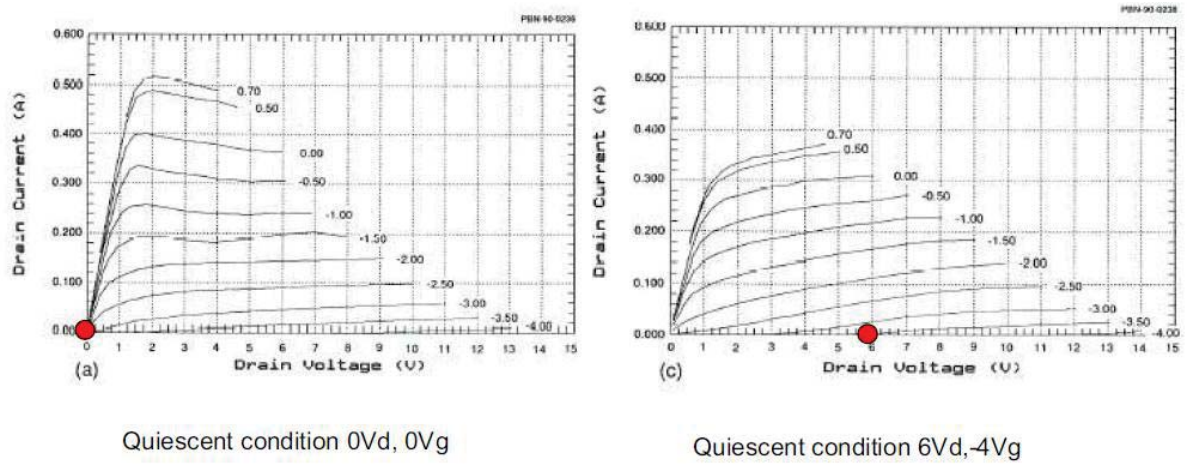


Figure 5-10, Pulsed IV data of pHEMT at different quiescent conditions

multiple bias points if they are to include the effects of dispersion.

It should also be noted that trapping is device technology dependant; for example they do not occur in Silicon LDMOSFET (Laterally Diffused) or VDMOSFET (Vertically Diffused) devices, which can therefore have simpler models.

Pulsed DC-IV measurements may also produce effects that are not seen in practice with an RF signal, such as shown in Figure 5-12. This effect is due to 'hot' electrons [24] being scattered from the InGaAs transport layer to the adjacent layers of AlGaAs where their mobility and velocity is lower. Although these effects can be seen when measured with a 100ns pulse, they are too slow to exist at frequencies in the GHz regions. This behaviour exemplifies the danger of basing a model purely on DC-IV measurements, where an observed behaviour is not replicated in the RF characteristics.

There is a balance to be struck between how accurately a model emulates actual behaviour and its complexity. Table 5-2 [25] summarises the accuracy in simulating I_{ds} as a function of V_{gs} and I_{ds} as a function of V_{ds} of a number of common modelling approaches and a rough measure of their complexity (number of fitting parameters).

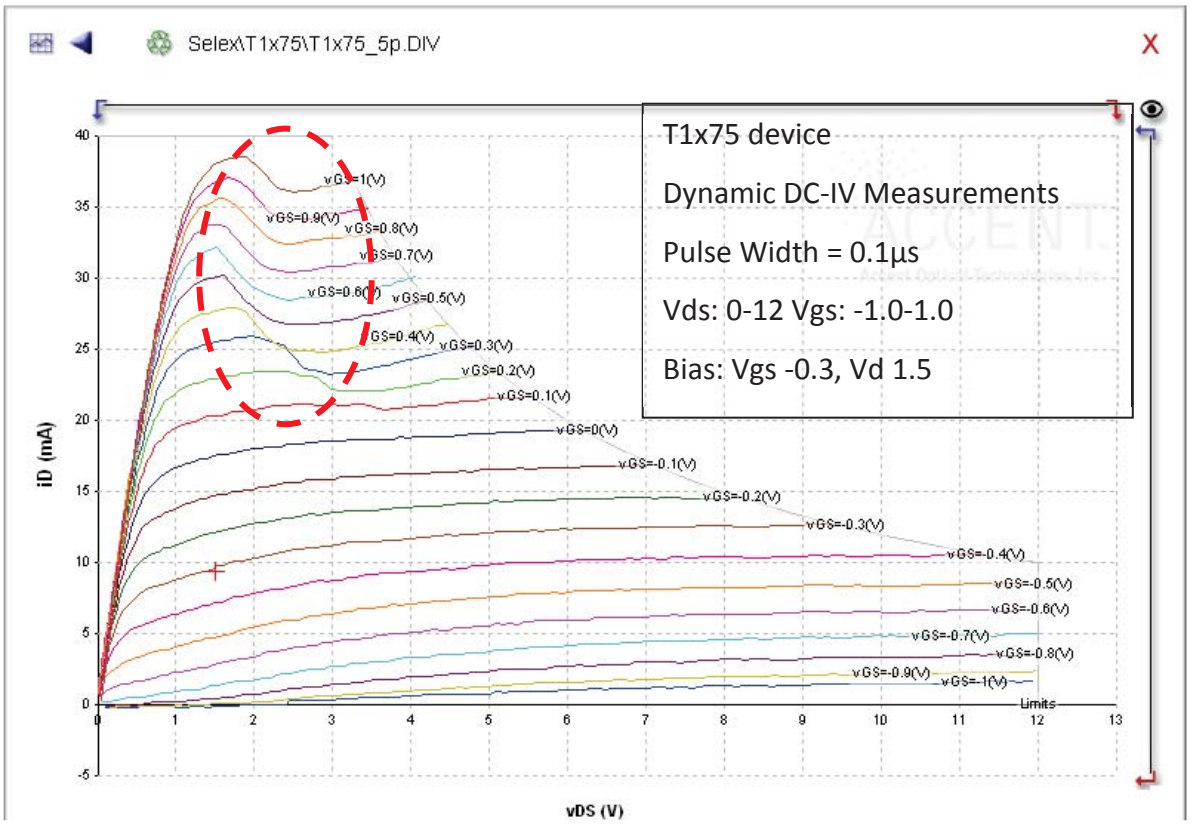


Figure 5-12, Hot electron tunnelling effect seen in DC-IV measurements of GaAs pHEMT

Model	$I_{ds} = I_{ds}(V_{gs})$		$I_{ds} = I_{ds}(V_{ds})$		Number of fitting parameters
	Max error (%)	RMS error (%)	Max error (%)	RMS error (%)	
Curtice	10.64	2.89	17.05	8.14	4
Curtice-Ettenberg	13.31	3.13	48.61	9.34	7
Statz	10.64	2.57	12.99	7.47	5
Materka	4.65	1.93	20.58	6.09	4
Triquint	6.21	2.88	7.14	2.30	6

Table 5-2, Accuracy of various Drain current Models and Complexity.

Such a comparison is crude and does not give the whole picture. The continued developing of models, with the striving for better and more comprehensive emulation of device behaviour, is indicative of both the need and the desire for the effort that has been applied. The question therefore must be asked whether the equivalent circuit model is the right approach, and as was referred to at the start of this chapter in an article by Maas [13], if the attempt to model even simple physical characteristics leads to greater complication and other problems to resolve.

If the number of measurements required to produce enough data to encompass the entire operating envelope is large then it is necessary (a) to have a stable process (to avoid skewing the results) and (b) a significant number of the devices before the model can be developed. Considerable effort must then go to processing and fitting the measured data to a suitable model, and even then the model may only be applicable to that particular device process. Thus models are only likely to be available sometime after the process becomes established, which in itself limits the take-up of the devices or produces circuits which do not fully exemplify the best possible device performance. Furthermore, the number of different types of measurement required in order to obtain the data, as shown in Figure 5-13 from Auriga [26] makes the whole process not only intensely specialist, but

prohibitively expensive for most organisations. It is with the dawning of this realisation that the use of measured data directly in models has begun to win over the amplifier design community; not that this is a new idea. Back in the early 90's Root published a paper describing the creation of models using automatic measurements, [27] showing how efforts were being made to try and reduce the time and effort required to produce accurate models for use in simulators. The key was to try and replicate the behaviour of the nonlinear elements by using data extracted from measurements. We can see here the basis for the next step in model generation; using measured data for the whole of the device rather than just for the nonlinear elements in the equivalent circuit model.

5.5 'Black Box' Modelling

The alternative to trying to develop either a physical or an equivalent circuit model is to treat the device under investigation as a 'Black Box', i.e. a component that has a transfer function, but with which we do not care how this transfer function actually occurs; we just

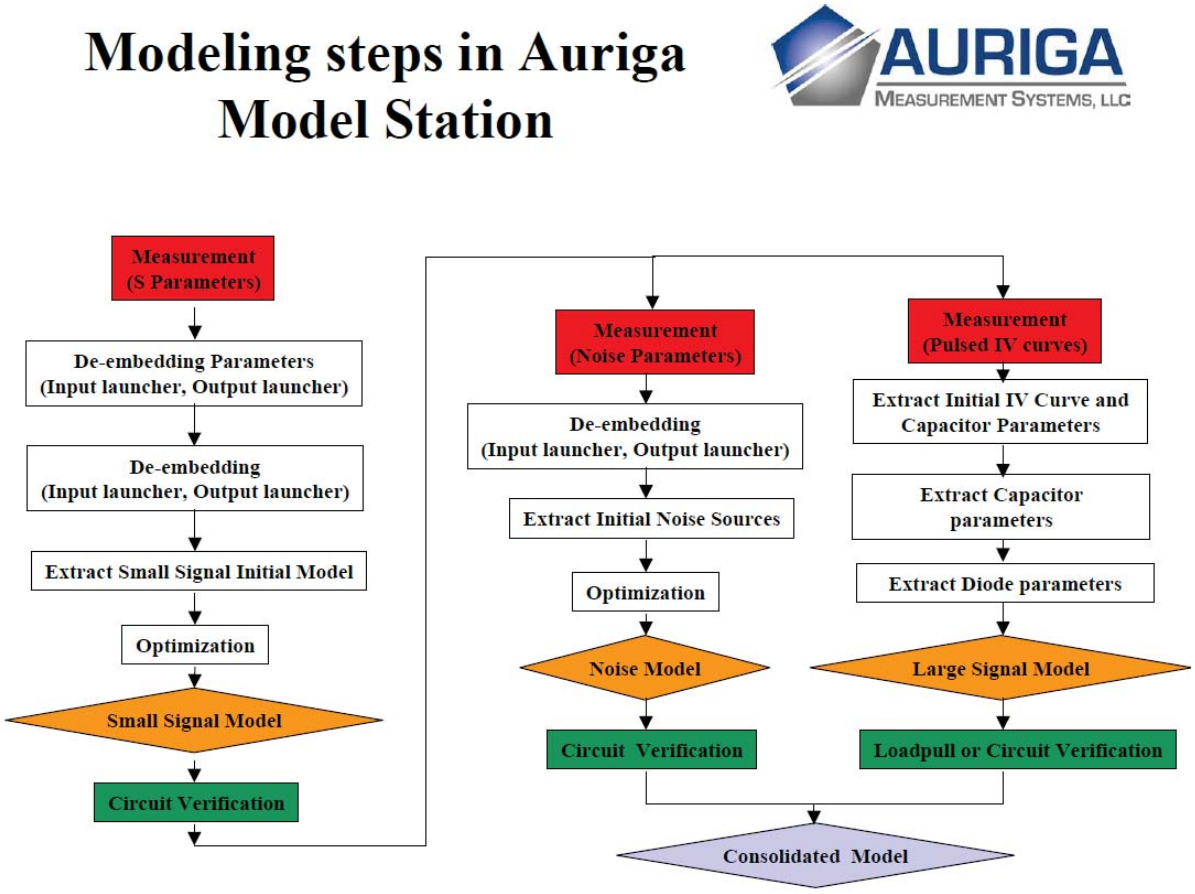


Figure 5-13, Model development process outlines by Auriga.

wish to be able to derive it from observing what happens to an input stimulus as it leaves the output.

There is an immediate problem with this proposition, as will be discussed later it may be important to know what happens at a specific point inside the 'Black Box', for example at the current generator plane, and so we do need to know enough information to be able to de-embed the measured data back to this point. But that aside, the basic principal holds. The most obvious and commonly used implementation of this approach, in the area of microwave circuit design, has already been mentioned; S parameters (S=Scattering). Conceptually the S parameters describe the linear behaviour of the component in terms of voltage wave ratios. The standard convention uses a_i to represent the wave incident upon port i and b_j to represent the wave resultant or scattered from port j. The ratio S_{ij} shows the relative scattered signal size from port i to the incident signal at port j assuming that the other ports are terminated in the system impedance, i.e. perfectly matched with no incident signals. Figure 5-14 shows the S parameter representation of a 2 port network. The scattered signals can thus be defined as shown in {5-24} and {5-25}. One of the key benefits of S parameters is that overall performance of multiple components can be assessed by combining their individual S parameter data blocks. The other major benefit, and certainly with the introduction of automated vector network analysers in the late 1980's, was the ease with which the S parameter data could be obtained and stored.

$$b_1 = S_{11}a_1 + S_{12}a_2 \qquad \{5\text{-}24\}$$

$$b_2 = S_{21}a_1 + S_{22}a_2 \qquad \{5\text{-}25\}$$

S parameters however contain only linear information, they assume that the

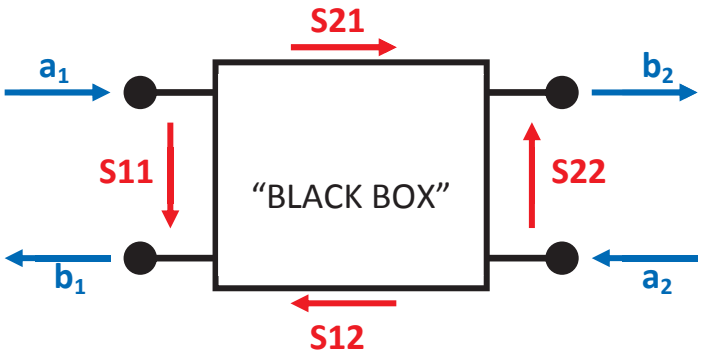


Figure 5-14, the Scattering of S parameter convention

relationships are valid for all magnitudes of the incident wave and no power is transferred to any other frequency (frequency is in fact not mentioned in {5-24} and {5-25}, and in S parameter data files there are different values of S_{ij} for each frequency). In order to overcome this limitation S-functions and X-parameters have been developed to capture the spectral response of a device to a stimulus. They are both in fact specific implementations of a more general Poly Harmonic Distortion (PHD) model [28], which seeks to be to the large signal nonlinear world what S parameters are to that of the small signal linear.

The PHD approach to creating models that can be integrated into high frequency simulators is based on frequency domain measurements of devices stimulated by a strong fundamental tone and a number of smaller harmonic tones. The approach relies upon the higher order harmonics being of a low level that does not affect the behaviour of the device and is a development of mixer theory where it is only the local oscillator signal that is strong enough to activate the mixer sufficiently so that the weaker signals can be analysed in a linear time dependent operating mode, i.e. frequency multiplication (mixing) occurs in a linearly predictable manner. This has been described graphically [29] and is reproduced in Figure 5-15. On the left hand side graph the input signal a1 is shown to consist of a fundamental signal (black) and a number of smaller harmonic signals (red, green and blue). On the right hand side the harmonic signals as a result of the fundamental in a1 are all shown in black. The input harmonic tones mix and using the superposition principle cause the net vector at each of the output b2 tones.

The measurement of the signals required for the PHD model depended on the development of Large Signal Network Analysers (LSNAs) with multiple sources so that vector measurements of the harmonic signals can be measured with known phase in addition to the fundamental (as in a conventional VNA). The PHD model itself uses the measured data to describe how the output B2 wave results from the input A1, in simple terms, {5-26}:

$$B_{pm} = F_{pm}(A_{11}, A_{12}, \dots, A_{21}, A_{22}, \dots) \quad \{5\text{-}26\}$$

In these equations the nomenclature used is that the first subscript refers to the port number and the second to the harmonic index. F_{pm} is referred to as a describing function and the overall relationship between B and A waves is called Spectral Mapping.

The PHD model is an approximation of {5-26} that involves linearization of the equation around the stimulus. It is important to note that $F_{pm}(.)$ describes a time invariant system, i.e. delaying the input signals (A waves) results in exactly the same delay for the output signals (B waves). In the frequency domain time change is translated as a phase change (that is proportional to frequency). Hence {5-26} becomes:

$$B_{pm}e^{jm\theta} = F_{pm}\left(A_{11}e^{j\theta}, A_{12}e^{j2\theta}, \dots, A_{21}e^{j\theta}, A_{22}e^{j2\theta}, \dots\right) \quad \{5\text{-}27\}$$

A point not mentioned in [28] should be commented on here, earlier in the discussion of equivalent circuit models it was pointed out that dispersion or memory effects were an added complexity to creating accurate models. This is equally true of the PHD model as the fundamental assumption just stated is that it is time invariant, which is not the case when the previous condition of a device has an effect on its behaviour. Although X parameters

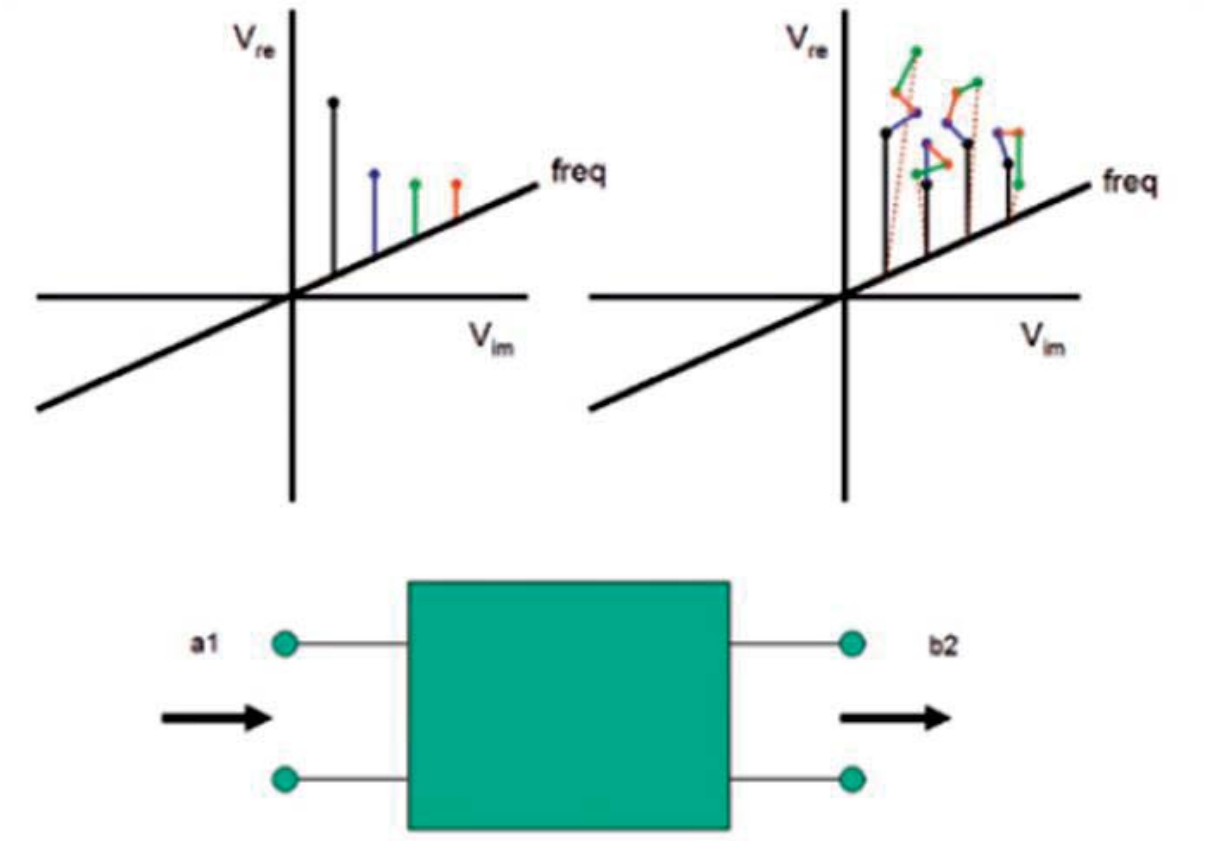


Figure 5-15, Harmonic superposition principle

have been modified to take into account some memory effect time variance [30], it is not currently clear that the implementation software running the models or measurement systems acquiring the data, incorporate the necessary functions to utilise these additions.

Continuing with the development of the PHD model, the next step is based upon the fact that the actual value of the phase of the primary stimulus, A_{11}, is not important; it is the phase relationship of the other signals in relation to it. Hence if θ is made equal to the inverted phase of A_{11} we are left with $|A_{11}|$ for the first term. Also at this stage if a phasor P is introduced such that $P=e^{j\phi(A11)}$, note therefore $e^{j\phi} = P^{-1}$. Hence we have normalised to the phase of the input signal.

$$B_{pm}P^{-m} = F_{pm}(|A_{11}|, A_{12}P^{-2}, A_{13}P^{-3} \dots, A_{21}P^{-1}, A_{22}P^{-2}, \dots) \quad \textbf{\{5-28\}}$$

$$B_{pm} = F_{pm}(|A_{11}|, A_{12}P^{-2}, A_{13}P^{-3} \dots, A_{21}P^{-1}, A_{22}P^{-2}, \dots)P^{+m} \quad \textbf{\{5-29\}}$$

Normally in strongly nonlinear systems the superposition theory is not valid, however it has been said for power amplifiers with one dominant input component (A11) it can be used[4]. This special case is referred to as the harmonic superposition principle [31]. Referring back to Figure 5-15, this principle holds provided the overall deviation of the output spectrum of B2 is the superposition of the individual deviations and has been verified experimentally in [31]. Note this is the key assumption leading to the PHD model. The final derivation of the PHD equation is given in {5-30}, the full details of the derivation are given in [28], but the key points are linearising {5-29}, dividing it into its real and imaginary components, and then substituting these with a combination of the input signals and their conjugates.

$$B_{pm} = \sum_{qn} S_{pq,mn}(|A_{11}|)P^{+m-n}A^{qn} + \sum_{qn} T_{pq,mn}(|A_{11}|)P^{+m+n}conj(A^{qn}) \quad \textbf{\{5-30\}}$$

The PHD model and its derivatives X-parameters and S-functions provide accurate representations of time invariant devices within the limitations of the model measurement detail, i.e. the number of harmonics measured and then used in the creation of the model.

[4] This proposition must surely be questioned in regards to multi-tone signals and may suggest that the applicability of the model is to certain specific modulation schemes.

For completeness the equations for X-parameters are given {5-31} and S-functions {5-32} as stated in [32], and can be seen to be very similar in approach.

$$B_{pm} = X_{pm}^{(F)}(|A_{11}|)P^{m} + \sum_{qn\neq 1,1} X_{pq,mn}^{(S)}(|A_{11}|)P^{+m-n}\, a_{qn} + \sum_{qn\neq 1,1} X_{pq,mn}^{(T)}(|A_{11}|)P^{+m+n} conj\left(a_{qn}^{*}\right) \quad \textbf{\{5-31\}}$$

$$b'_{pm} = Sf_{pm11}(|a_{11}|) + \sum_{qn\neq 1,1} \left(Sf_{pq,mn}\, a'_{qn} + Sfc_{pq,mn}\left(a'_{qn}\right)^{3}\right) \quad \textbf{\{5-32\}}$$

In the X-parameter expression {5-31}, the first term describes the fundamental output term resulting from the input wave A_{11}. The terms in the summation describe the mixing via the P term. They characterise the wave leaving port 'p' at harmonic 'm' due to the small harmonic signals at the input as shown in Figure 5-15. Each set of X-parameters is a function of the magnitude of A_{11}, thus allowing the model to accommodate the nonlinear behaviour. In the S-function equation {5-32} the first term describes the output signal resulting from the input stimulus, and the terms in the summation describe the changes from the smaller perturbations at the harmonics.

These approaches require that coefficients exist for each load impedance and that the simulator utilising the model interpolate between the points. An improvement on these techniques, [33] has shown that by considering higher order mixing terms a model can be developed that can describe load pull contours across the fundamental impedance plane, and then extended to include harmonic terminations, [34].

Many of the current examples used to demonstrate the accuracy of X-parameter are centred on 50Ω load impedances and the ability of such measurements to handle significantly reflective loads must be questioned. This is not to question the basis behind the X-parameter model itself, but to warn against extrapolating what is a measurement based model outside the of the measurement space.

Referring back to an earlier discussion about the number of harmonics required to be controlled in the design of high efficiency amplifiers, Table 5-3 shows the impact of increasing the number of harmonics on the number of coefficients needed in the various modelling approaches, [35] for a 2 port (n=2) network.

harmonic	mixing order	Maximum number of Coefficients		
		Cardiff PHD Model	General Formulation	Superposition Principle [28]
h	w	nh(w+1)	nh(w+1)^h	nh(1+2h)
3	1	12	48	42
3	2	18	162	42
3	3	24	384	42
3	4	30	750	42
3	5	36	1296	42
3	6	42	2058	42
3	7	48	3072	42
5	5	60	77760	110
5	6	70	168070	110
5	7	80	327680	110

Table 5-3, Coefficients required in behavioural model formation.

In the acquiring of the data for generating X-parameters the a and b waves are measured at each load point and for each input level of a_{11}. This is then repeated with a small extraction tone applied at each port to each harmonic, one at a time. Hence the complete data set is a matrix containing the response between the ports at the fundamental and all of the considered harmonic frequencies and combinations thereof. Note that in the limit when the amplifier is behaving linearly, within the dynamic range of the measurements, the harmonic terms tend to 0 and the data set becomes the S parameter set (not changing as a_{11} reduces).

Perhaps the strongest argument in favour of the measurement based modelling approach is that in order to ascertain the accuracy of other model types it is necessary to measure device performance and use this as the baseline for any comparison. Using the measured data itself for the model is therefore inherently the most accurate. This could bring into question the whole X-parameter basis as the models are based upon separate harmonic perturbations, whilst a 'proper' measurement would require the harmonics to be applied simultaneously, thus returning to the prime question that if you have to do this measurement anyway why not use it as the basis for the model?

The measured data model is however only valid within the measurement envelope, whereas the equivalent circuit model **may** be able to approximate performance at other bias settings or power levels. But to prove/confidence check this, measurements must be made, reinforcing the argument for the measured model approach. An apparent

contradiction to this argument is the tendency to compare the measured model performance to that of existing equivalent circuit models, [36] which if anything has confirmed how good some of these models are. The answer must lie in confidence levels. If the equivalent circuit model obtained from the device manufacturer is known to be accurate under the designers operating conditions, and if the model converges quickly and thus has reasonable simulation times, then these will be the obvious choice for the designer. However as nonlinear vector network analysers become more prevalent then it will become natural to use the verification data on a device in the simulator, cutting out the need for a modelling exercise. In the early stages this is likely to be taken up by the users of matched devices such as packaged MMICs and mixers operating in 50Ω, the equipment needed to conduct full multiple harmonic load pull being still prohibitively expensive and complicated. This will allow the nonlinear data to be passed to system simulators, a significant problem in the past. For power amplifier designers one of the key advantages of PHD models is that load pull data can now be captured and utilised directly in nonlinear simulators. Publications advocating the difference that X-parameters will make [37] are actually doing little more than describing the long accepted practice of fundamental load pull. As the PHD model and its 'cousins' are based on measurement data the question could rightly be asked "why not use this data directly itself?". This approach is in fact discussed in detail in the next section however in summary these behavioural models just discussed:

- Significantly reduce the size of the stored data set.
- Fewer measurements are required to produce the descriptive functions.
- The modifying the coefficients may be used to represent 'spread' in device performance.
- Measurement noise is filtered out by the Fourier transformation creating the models.
- Improved (but still risky) extrapolation on the edges of the measurement envelope is feasible.

One application where the modelling approach has made a marked and singular improvement to the measurement method itself is outlined in [38]. Load convergence in open loop active load pull is based upon:

$$A_{2h} - \Gamma_h B_{2h}(A_{21}, A_{22}, \ldots, A_{2h}) = 0 \qquad \{5\text{-}33\}$$

where h is the harmonic number.

This is an iterative process, complicated particularly at the fundamental frequency by the fact that A_{11} changes with Γ_{L1}, referring to Figure 5-16, {5-34} and {5-35} describe the relationships between the power waves and the loads reflection coefficients (the S parameters refer to the device); in practice convergence on a solution may take 5-10 iterations.

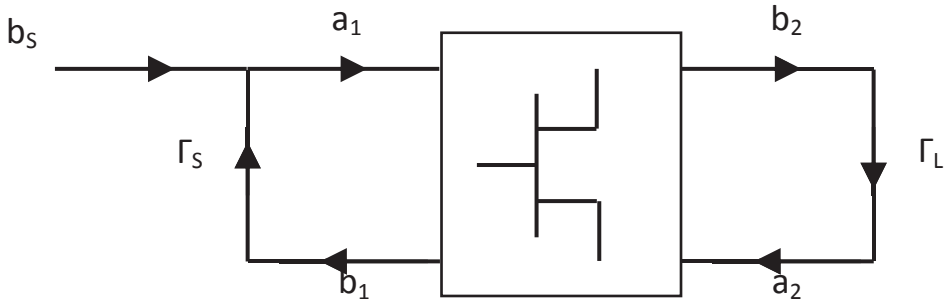


Figure 5-16, signal flow diagram around Device Under Test

$$a_1 = b_s + \Gamma_s\, b_1 \qquad \{5\text{-}34\}$$

$$\Gamma_S = S_{11} - \frac{S_{12}S_{21}\Gamma_L}{S_{22}\Gamma_L - 1} = \frac{b_1}{a_1} \qquad \{5\text{-}35\}$$

The relationship {5-33} has been utilised along with the Cardiff PHD model to create a local model which greatly enhances the open loop Active Harmonic Load Pull (AHLP). AHLP attempts to converge on specific loads – fundamental and harmonic - typically maintaining 1 or more of these whilst varying another. The problem being that as, for example A_{21} changes so do B_{22}, B_{23}, etc. and to maintain Γ_{2L} and Γ_{3L} constant A_{22} and A_{23} must also change. As Γ_{xL} is altered so Γ_S changes as predicted by {5-35} and shown in Figure 5-16. By knowing in advance how these changes will occur, the loads (Γ_{hL}) can be maintained as Γ_{xL} is varied. AHLP is non-ideal due to the behaviour of the system components (amplifiers, sources, etc.) so that adjustments in A_{2h} (to set Γ_h) are not exactly as the calculated values. This error, $T_{s,h}$ is frequency and power level dependent.

The impact of changes in load impedance on the input reflection coefficient and hence available source power from wafer probe measurements on a 10x75 µm GaAs pHEMT are shown in Figure 5-17. The central red grid shows the fundamental load impedances of the measurement. The blue 'curved' grid shows the fundamental input impedances as the load impedance changes. Figure 5-17a is at the lowest drive level and Figure 5-17b at the highest. Figure 5-17c shows how the input impedance changes at a single load point with drive power. From Figure 5-17e it can be seen that there is about 2 dB difference in the available input power as a result of the changes in the input impedance over the power range. This leads to gain expansion as seen in Figure 5-17d.

It should be noted that the behaviour of Γ_{in} is technology dependant. Similar measurements on GaN devices (10 & 25W Cree) have not exhibited this behaviour until much higher levels of compression, and hence a 'shallower' compression curve.

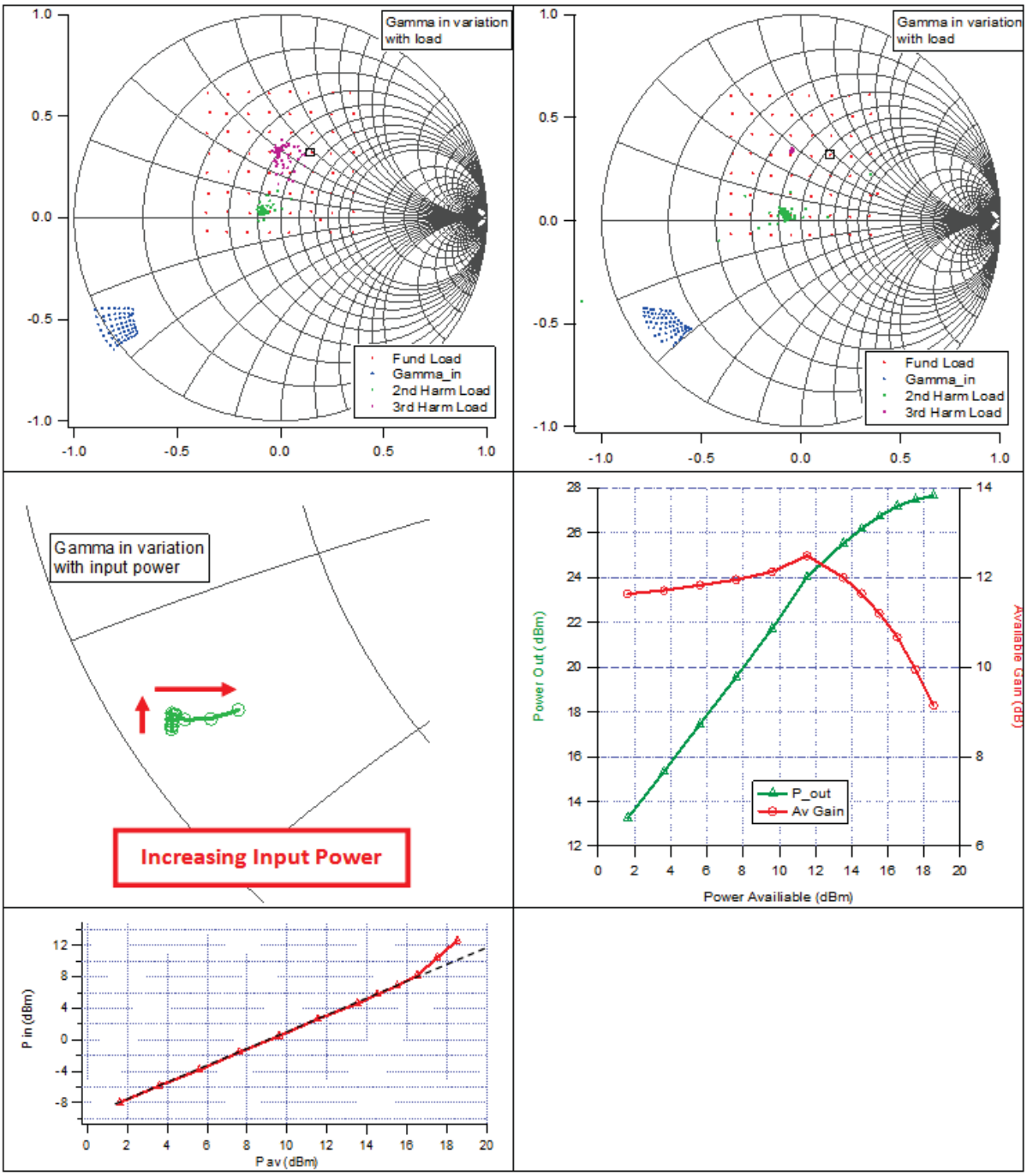


Figure 5-17, Changes to Γ_{in} with drive level, (a) to (e) clockwise from top left. (a) reflection coefficients at lowest drive level, (b) at highest, (c) variation with drive level at a fixed load point, (d) effects on gain and output power, and (e) impact of reducing Γ_{in} on power into device. Measurements made on 10x75 µm GaAs pHEMT biased in class A.

5.6 Direct Measurement Model Implementation

Arguably a variant of the 'Black Box' model described in the previous section there exists a form of model that basically stores a set (or sets) of measured data that can be read directly by a simulator. Although some would argue that these are not 'models' in the true sense of the word; in practice it is their use in CAD and simulation that determines their usefulness and applicability to the modelling scenario. Indeed this approach is not new; S parameters in the linear simulation domain are measurement based data sets. To be useful in a simulation environment it is necessary for the data to be stored in an accessible form which includes a method for choosing the appropriate data sets. Indeed one of the very early attempts at addressing modelling of the nonlinear behaviour in power amplifier design was to measure device S parameters at various input power levels and to store this data within "MDIF" (Measurement Data Interchange Format) files, the appropriate S parameter data set being accessed by referencing to the input power level. As described earlier S parameters are limited by only describing the fundamental frequency, however the measurement system described in the chapter 3 captures the fundamental and harmonic voltage waveforms and by extension the currents. These can also be stored in a MDIF file with the data accessed by reference to the appropriate load and input stimulus (plus other parameters such as temperature and bias if required) [39]. This is referred to as the Direct Look-Up Table (DLUT) Model and was the method used in the design process described in this thesis[5].

The basic model consists of a normalised (to the phase of the input voltage waveform) data table of the Fourier coefficients of the current components, indexed by frequency, bias, drive level (input voltage) and fundamental and harmonic load impedances. The measured data is stored in a large file, typically 70MB for a 10x10 impedance grid and a power sweep of 12 levels. By contrast the equivalent DLUT model is 168kB (to some extent this is due to the file structure used and the replication of data in a number of different formats). As mentioned the DLUT model is created in a "MDIF" file. This contains blocks of data identified by specific values of declared variables. The variables can be such parameters as gate and drain voltage, temperature, and load impedance. Within each data

[5] Nonlinear device modelling is still a fast evolving area despite having started over 30 years ago. Although now largely overtaken by other approaches DLUT models were the best available at the time of this research; the processes described are also applicable to the newer modelling methods.

block are a number of parameters describing the device as a function of the input voltage, v1. This model can be read in a nonlinear simulator, and by looking at the load impedance presented to the device and the input drive level, the input current and output voltage and current can be determined. The Cardiff DLUT is relatively quick to create (a 10x10 grid covering ~30% of the impedance plane, with 12 power levels takes about 1.5 hours/frequency of automated measurements) and the speed is constantly being enhanced. The simplicity of this model implementation allows all subcomponents and their complex nonlinear interactions to be fully accounted for. Therefore, the impact of a single design parameter on a complete RF system can be determined. This allows the investigation of design parameter sets to give the optimum performance of the complete RF system and not just its subcomponents.

In order to create a DLUT file a constant measurement grid must be established as the actual measured impedances will vary by some degree, depending on the success of the load convergence, as discussed in the chapter 3. A software program has been written which reads in all of the measurement data and creates a DLUT model, to an established 'best–fit' grid. As described earlier the model uses the MDIF format to store the data and uses the load impedances as one of the variable parameters to identify each data block. Thus there is a small source of error here in the model, as it is assumed the measured results come from a fixed regular load impedance grid. The extent of this error will depend upon how repeatability the AHLP system was able to converge upon the required load (at the different output power levels). An example of a part of the MDIF block is shown in Figure 5-19.

The model operates in MicroWave Office (MWO)[6] using a Voltage Controlled Current Source (VCCS) to create the input current waveform and a VCCS to create the output current, both determined by input voltage. An element called a Gyrator is used to produce the output voltage, these are dependent upon the input voltage and the load impedance presented to the model. Other variables exist such as gate and drain voltage; however within the existing model these remain fixed, as the measurement data acquired was at single bias sets.

[6] Commercial CAD package from AWR Corp., www.awrcorp.com

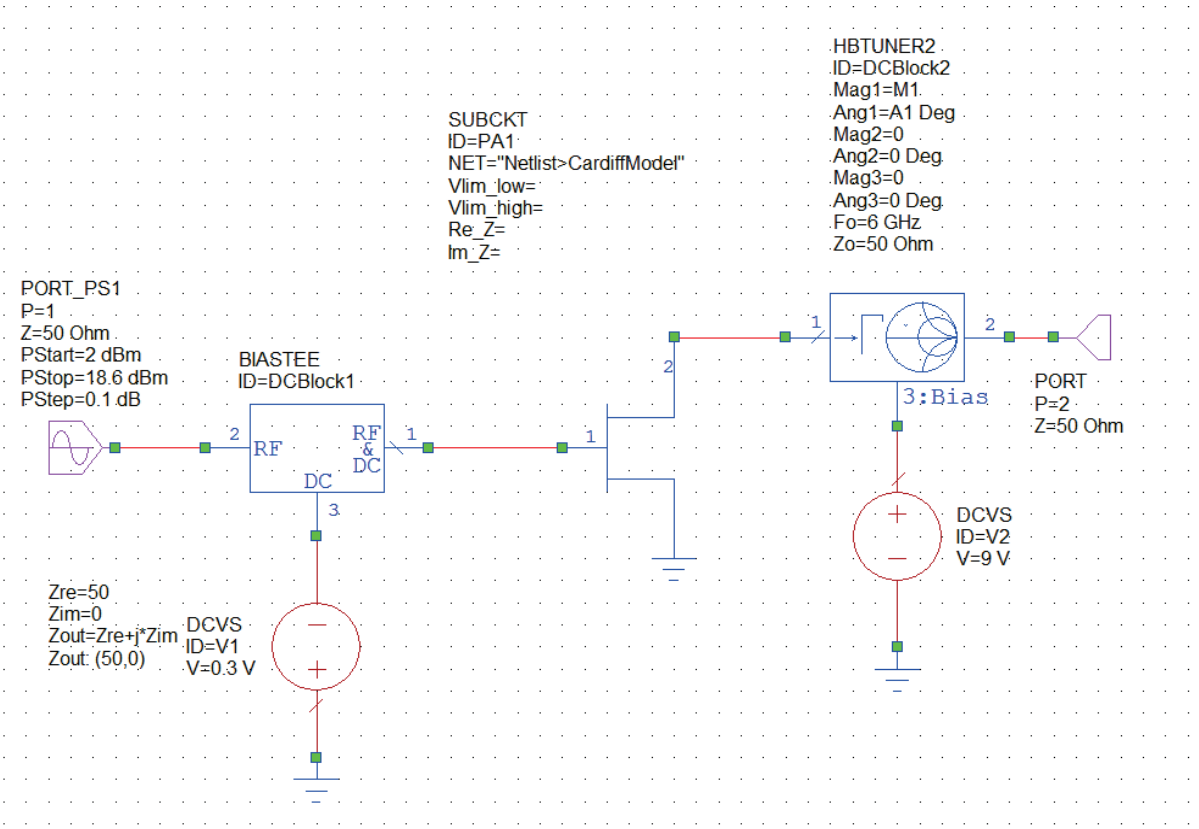


Figure 5-18, MWO circuit for swept power nonlinear simulation of Cardiff DLUT model

Within MWO the model is placed in a circuit such as the one shown in Figure 5-18. The fundamental load impedance is determined by two global variables M1 and A1 controlling a harmonic load pull tuner. The input power is swept across approximately the same range as the original measurements. The frequency and gate and drain voltages are also used as variables within the model file and are read from the values used in the circuit simulator.

The DLUT model was analysed with the load values corresponding to the measured maximums for PAE and output power (Pout), Table 5-4. For comparison the results of the measured and simulated data using the DLUT model were compared with the foundry or PDK (Process Design Kit) model.

Selex 10x75_35 at 9v 150 mA measured at 6 GHz at an input power level of 18.6 dBm.			
Maximums	Value	Γ mag	Γ pha (°)
PAE (%)	51.4	0.35	65.9
Pout (dBm)	28.0	0.22	78.1
Gain (dB)	16.4	0.61	94.3
Drain Eff. (%)	53.0	0.35	65.9

Table 5-4, Maximum measured performance values and corresponding load impedances.

The measurement used power steps of approximately 2 dB over the bottom of the power range and 1 dB at the top end. The simulation uses 0.1 dB steps across the range and thus shows the interpolation between measurement points. Similarly Figure 5-22 and Figure 5-27 show the load pull contours, which are using a circular grid (as opposed to the square 10x10 measurement grid), again the interpolation of the DLUT model when using points within the measurement grid is very good. Note however that where the simulation grid exceeds the measurement space the contours are perturbed, whereas the PDK model is consistent (if of the wrong value and centre). There is good agreement between measured and DLUT model data, as shown by Figure 5-20 to Figure 5-27. The swept input power results plotted are for two different fundamental load impedances, that for optimum PAE and for maximum output power. The PDK simulation is relatively close on its calculation of the optimum impedance points (from Figure 5-22 and Figure 5-27), however the calculation of both power output and PAE are optimistic (0.5 dB and 6%). In the PDK model the fundamental output power tracks closely at both loads simulated, however the harmonic performance and input reflection coefficient are markedly different.

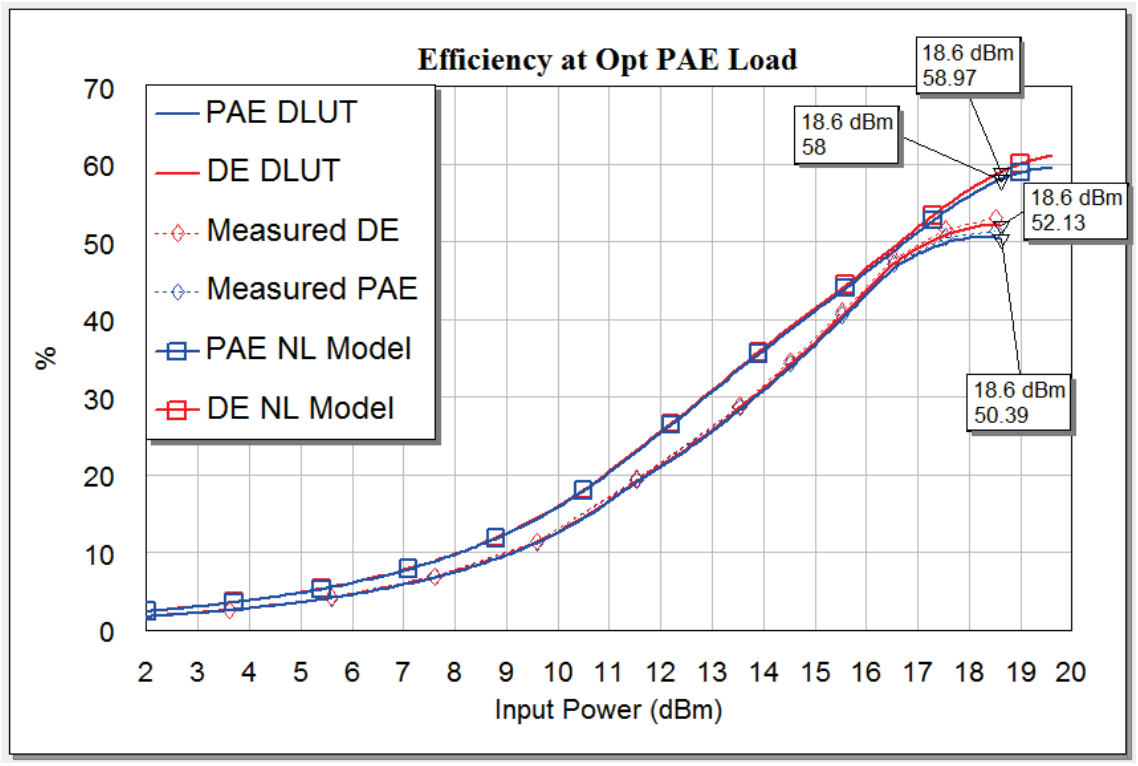


Figure 5-20, Comparison of Drain (DE) and Power Added Efficiency (PAE) performance of measured data, DLUT and PDK (NL) models at the Optimum PAE Load.

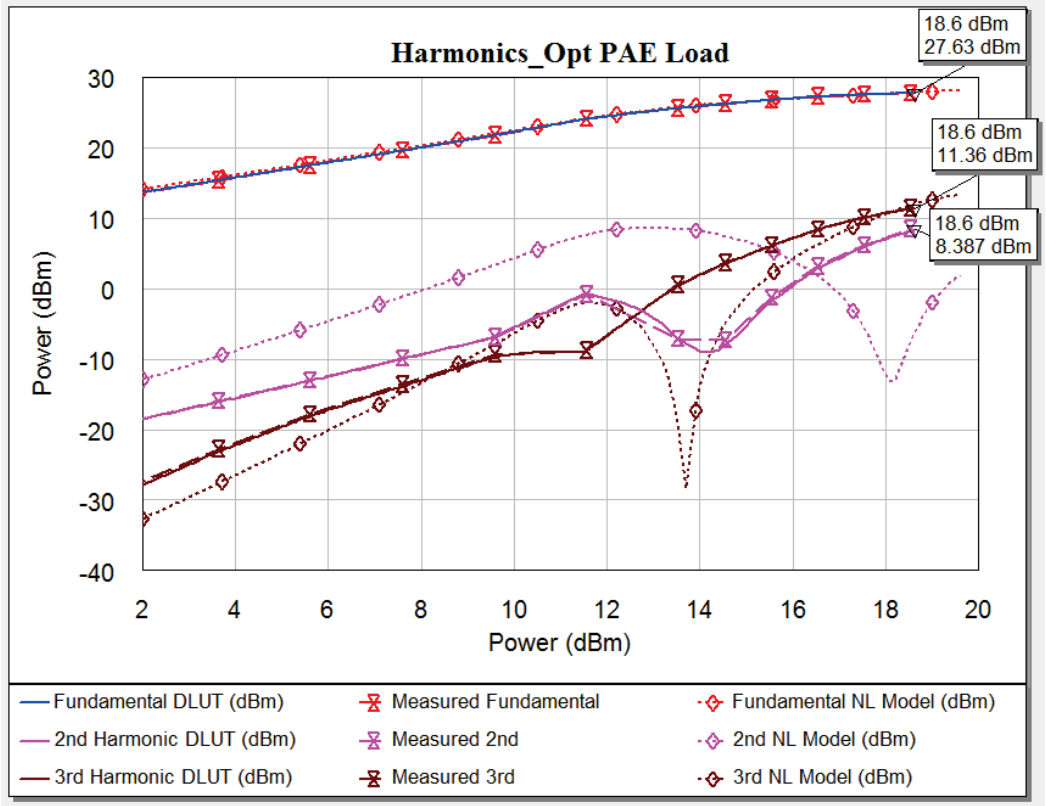


Figure 5-21, Comparison of Fundamental, 2nd and 3rd harmonic output power of measured data, DLUT and PDK (NL) models at the optimum PAE load.

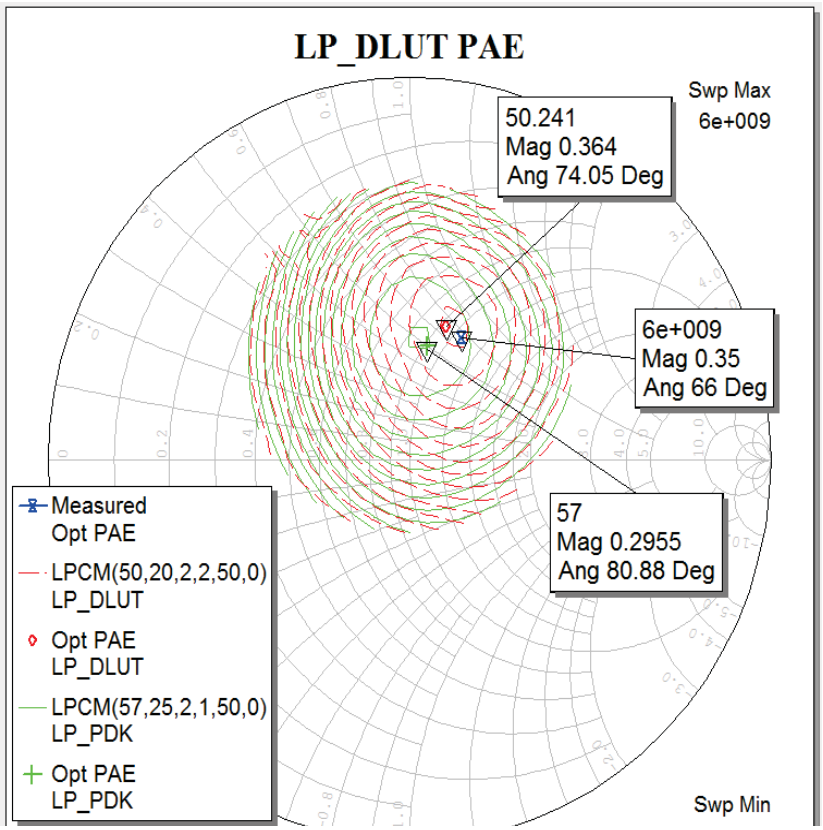


Figure 5-22, PAE Load pull contours for DLUT and PDK models, markers show optimum values and measured optimum (51.4% from table 5.4). Note perturbation in DLUT contours at the top left, where it meets the measurement grid edge.

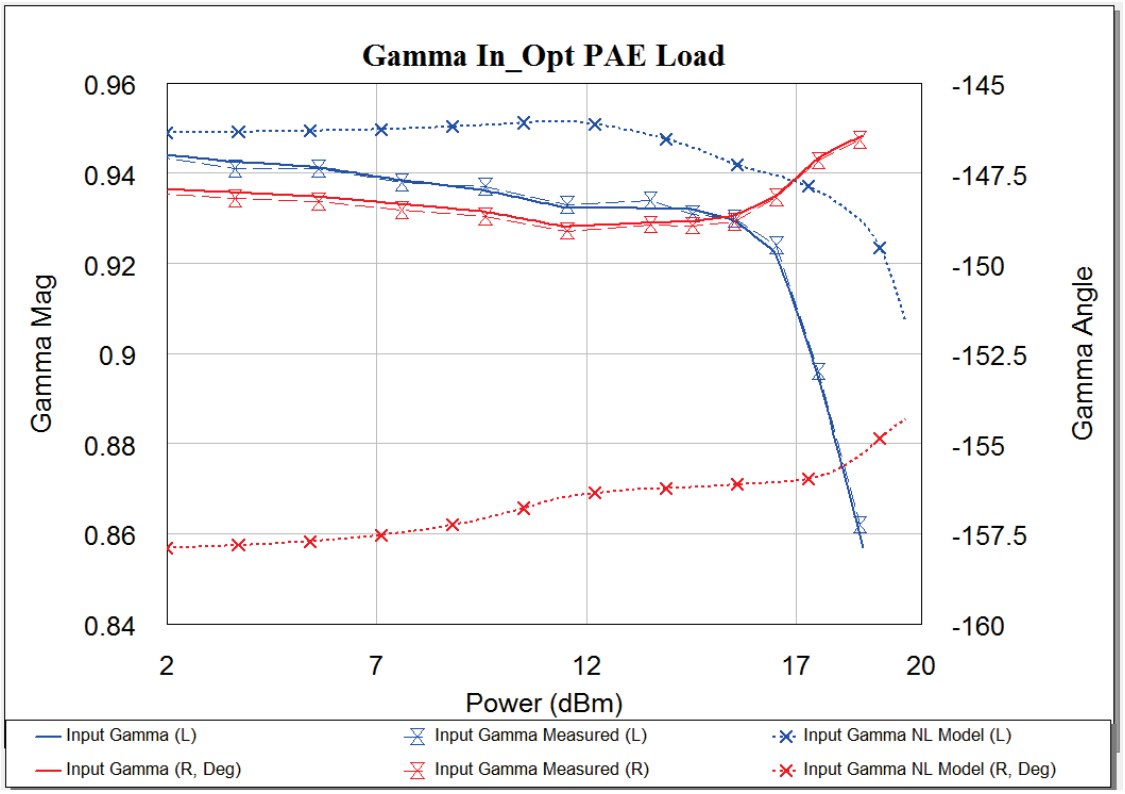


Figure 5-23, Comparison of Input Reflection Coefficient (Gamma) between measured data, DLUT and PDK (NL) models at Optimum PAE load.

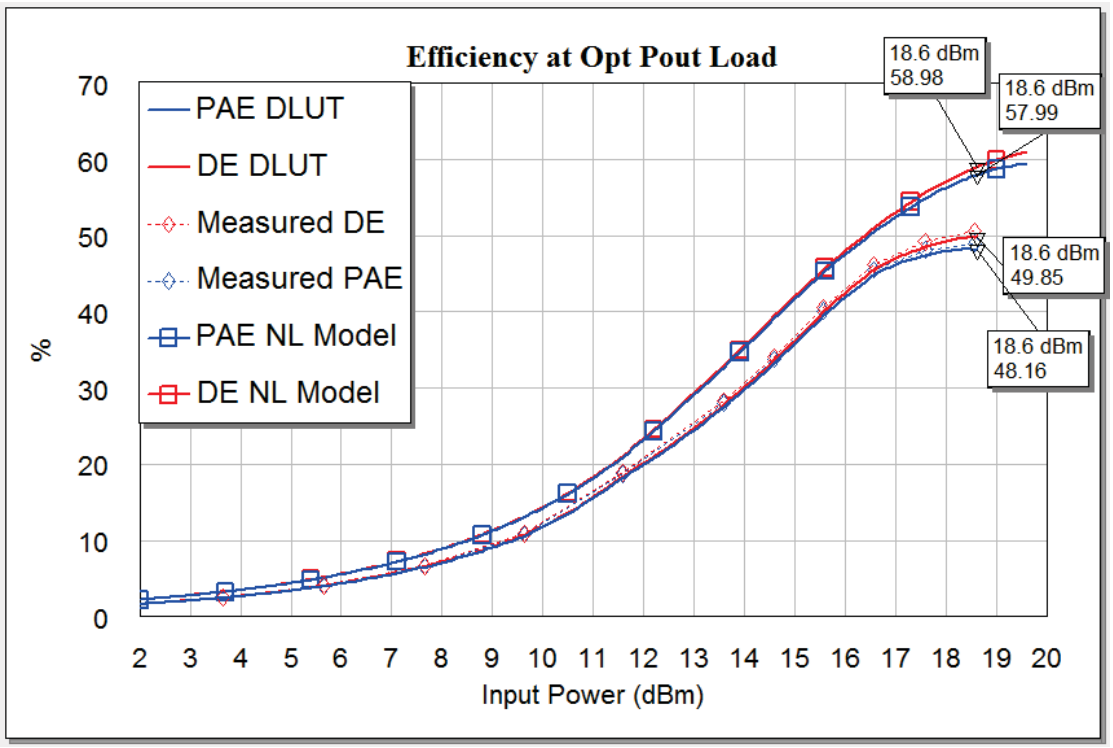


Figure 5-24, Comparison of Efficiency performance between measured data, DLUT and PDK Nonlinear (NL) models at Optimum Pout load.

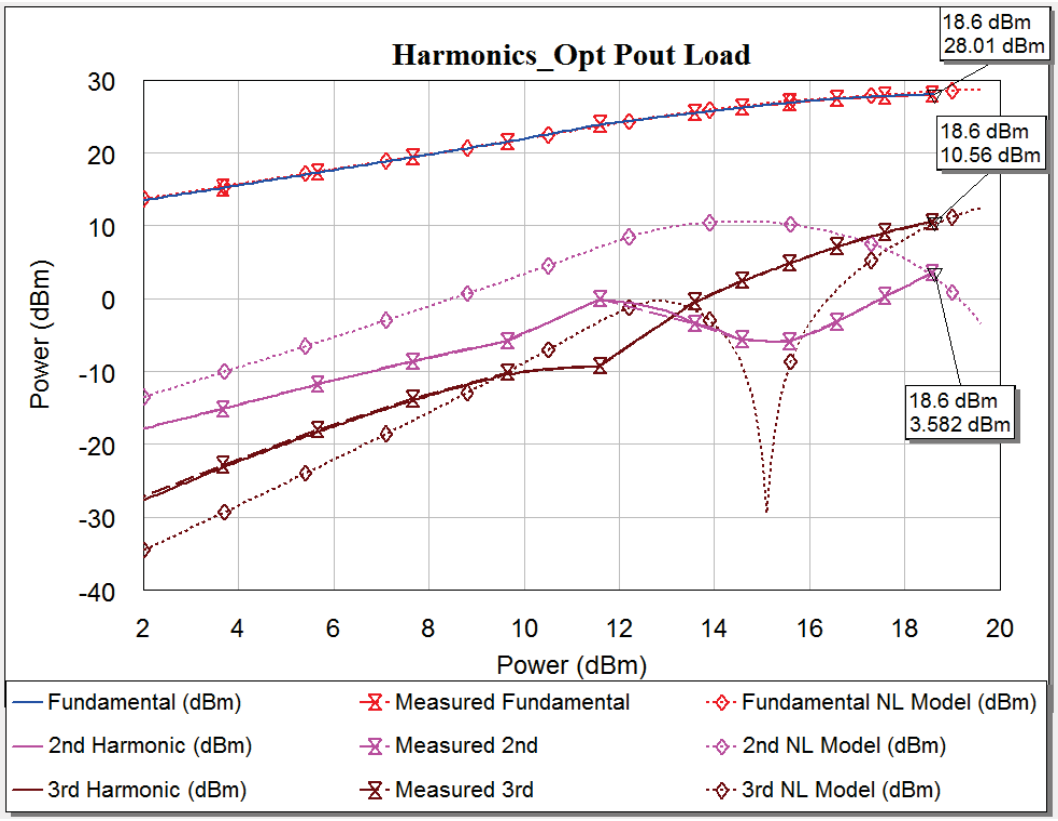


Figure 5-25, Comparison of Fundamental and Harmonic power levels between measured data, DLUT and PDK Nonlinear (NL) models at Optimum Pout load.

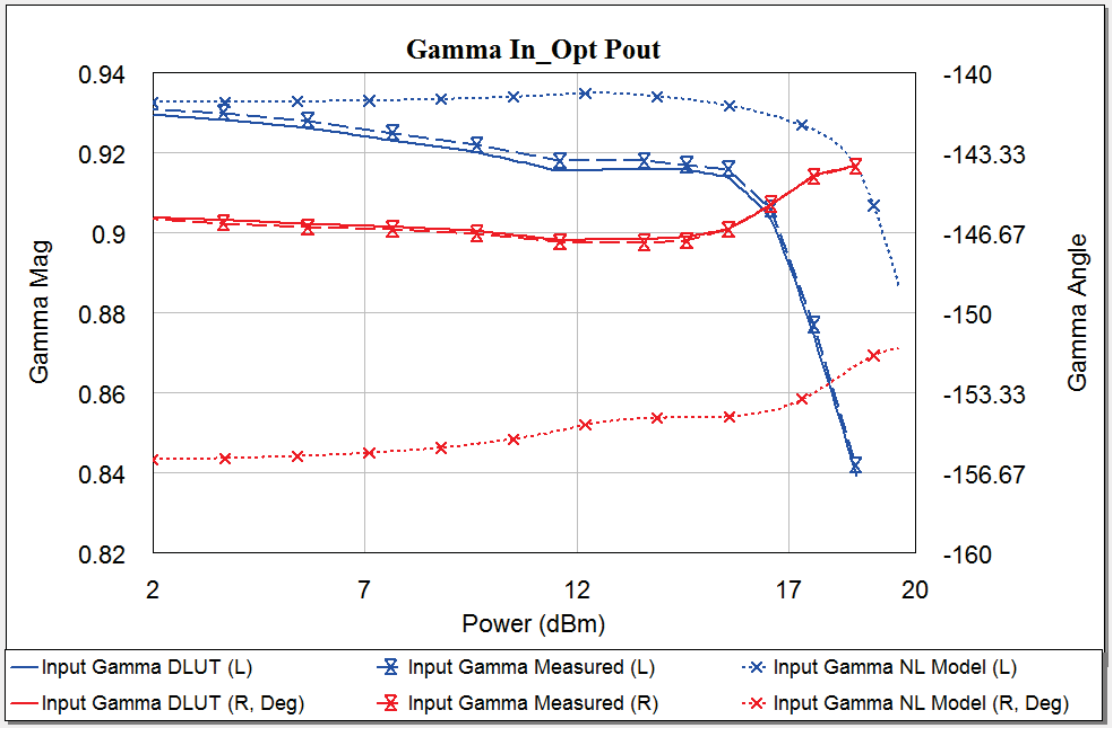


Figure 5-26, Comparison of Input Reflection Coefficient (Gamma) between measured data, DLUT and PDK Nonlinear (NL) models at Optimum Pout load.

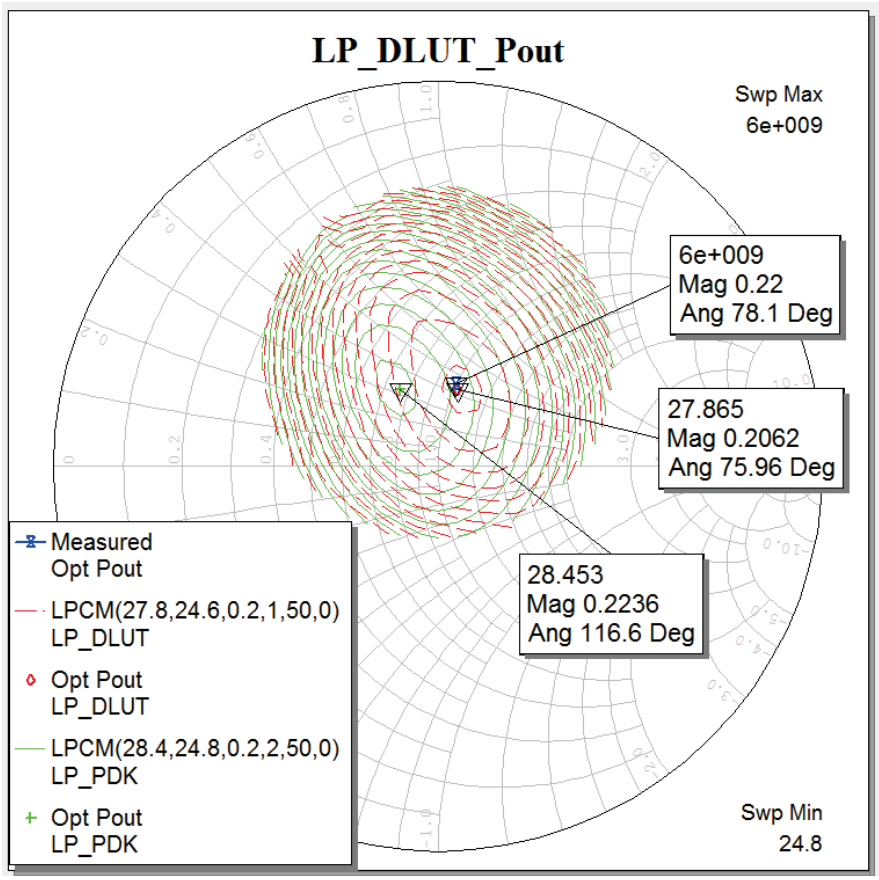


Figure 5-27, Pout Load pull contours for DLUT and PDK models, markers show optimum values and measured optimum (28.0 dBm from table 5.4). Note perturbation in DLUT contours at the top left, where it meets the measurement grid edge.

In order to maximise accuracy it is necessary to set the loads for the 2nd and 3rd harmonic within the simulator to those in the original measurement. The DLUT model used at this point did not contain the harmonic impedances as a look-up variable. Also as shown in Figure 5-28, during the measurements the 2nd and 3rd harmonic loads were not held fixed and so will be a source of error. In the ideal situation these would have been held constant at 50Ω, or at least at a fixed impedance. In this case nominal values of 2nd 0.1/_156.3° and 3rd 0.35/_98.0° where used in the simulation, (global variables M2, A2, M3 and A3). It is worth noting that in Figure 5-28 the input reflection coefficients are shown to go outside the Smith Chart (Γ>1) indicating a negative resistance. The nature of the measurement system is such that it can handle such potential instabilities.

As has been well documented [40], that for higher efficiency operation, the impedance of the harmonics are critical. Many models are limited by the original data used in their creation and hence may not include the region covered by harmonic frequencies. This was shown with the PDK model of the device measured. As shown in Figure 5-21 and Figure 5-25 the PDK model predictions for the 2nd and 3rd harmonic levels did not track the

measured performance. The AHLP system was used to map the effect of the 2nd harmonic load on the PAE (with the fundamental held at the optimum PAE load at 6 GHz). This showed that the optimum 2nd harmonic load was at a magnitude >0.95 and an angle of ~85°. The variation of PAE across the 2nd harmonic impedance plane was >16% and as Figure 5-29 clearly shows, the 2nd harmonic termination can have a detrimental as well as a positive influence on PAE.

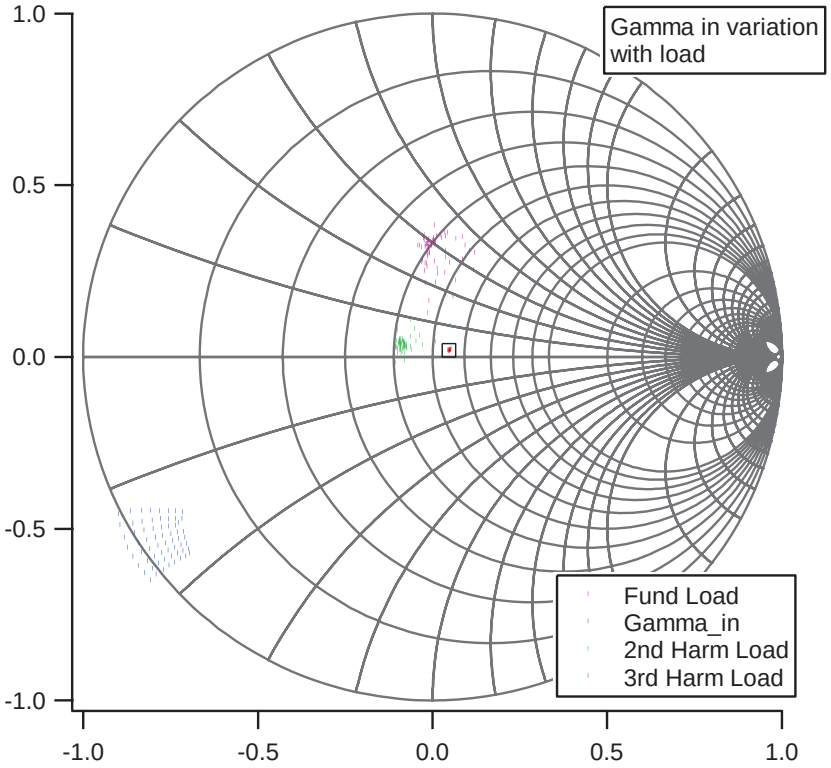


Figure 5-28, input reflection coefficients, fundamental, 2nd and 3rd harmonic load impedances.

By contrast a simulation of the PDK model shows not only a much reduced PAE variation over the 2nd harmonic impedance plane, <8%, but also the optimum is ~3° higher, as shown by Figure 5-30. The improvement in PAE from a 50Ω 2nd harmonic termination is only 5%, whereas in the measured case it is >9% and similarly the minimum PAE seen is 54% compared to the measured <43%. However the basic shape of the 2nd harmonic contours is very close. The reduced variation in PAE could lead a designer to underestimate the importance of the 2nd harmonic termination.

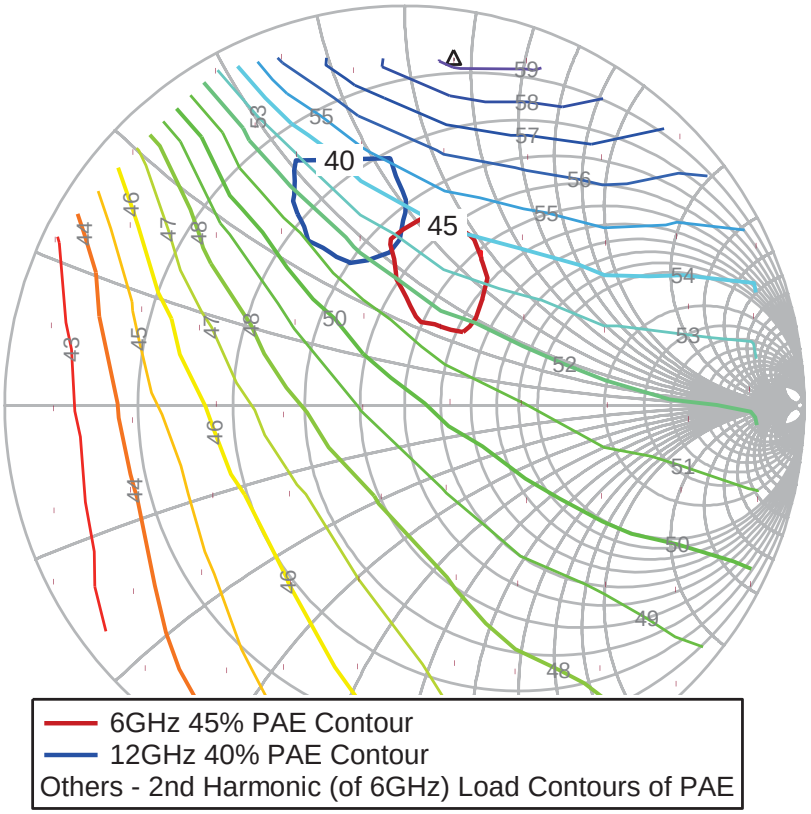


Figure 5-29, 2nd harmonic PAE load pull contours with the fundamental set to the optimum PAE load at 6 GHz. 45% circle is the 6 GHz PAE contour and the 40% circle is the 12 GHz PAE contour

Clearly the benefit of having a model directly based on measured data has been shown. Referring back to Figure 5-25, the harmonic growth shown by the PDK model does not accurately reflect that measured and replicated by the DLUT model. The DLUT model has been shown to accurately reproduce the measured data when the correct load impedances are used. From this we can see the importance of maintaining the load impedances at a constant, known, value during data acquisition; this should be greatly improved by the adoption of a fully phase coherent measurement system and the new model based load setting algorithm [38]. Interpolation between load impedances and power levels (input voltage waveforms) in the DLUT model is also good. However moving above the measured drive level or to the edges of the measured impedance plane (Figure 5-22 and Figure 5-27) the model accuracy breaks down, i.e. extrapolations are poor. This can be overcome by extending the measurement plane to cover the majority of the impedance plane, which should be greatly helped by improvements in the measurement speed.

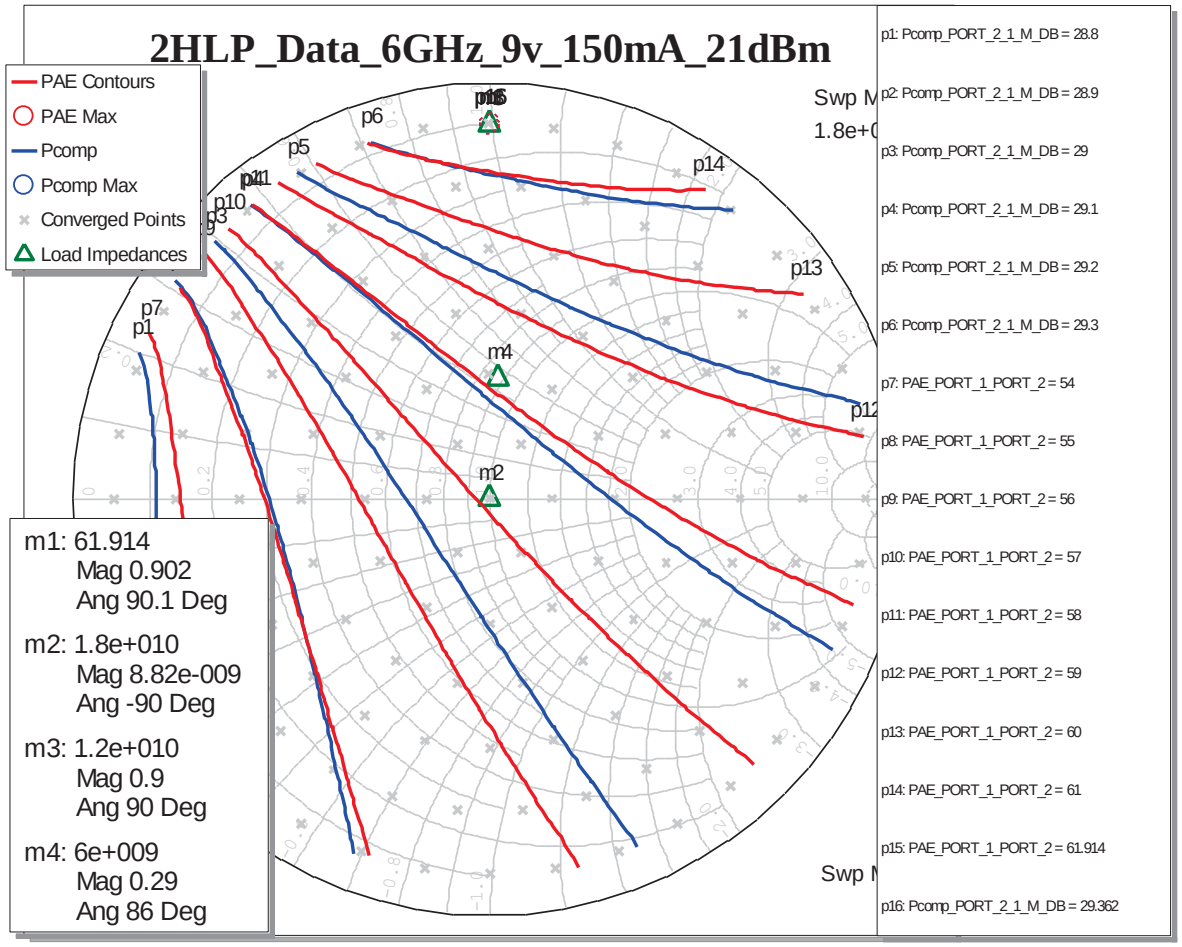


Figure 5-30, 2nd Harmonic Load Pull contours for the PDK model, with the fundamental impedance set to 0.35/_90° (the PDK models optimum PAE load).

In order to include the effects of the 2nd harmonic load impedance the model needs to include this data. However for the current DLUT fundamental only model there are 100 load points. For each load point there are 12 drive powers, thus there are 1200 data sets. If the same were repeated at each load point for the 2nd harmonic loads the data set would increase to 1,440,000 data sets. This would also be impractical as measuring the harmonic load pull contours for each fundamental load would correspond to 100 measurements. It is more practical to limit the measurements to a few impedances around the optimum load or at the specific load presented by a matching circuit (when used in conjunction with MMIC design for example). The usefulness of mapping the 2nd harmonic impedance plane as shown in Figure 5-29 can be seen from the indication of the 'direction' in which the matching circuit impedance should be driven and equally the areas to be avoided. A further advantage of the DLUT model is that it can be added to over time as other conditions are measured, increasing the comprehensiveness of the model. The wide operating bandwidth

of the measurement system makes it highly suitable for producing models for applications at C and X band where all the critical harmonic frequencies are covered.

As discussed earlier an alternative to the DLUT model, based on Poly Harmonic Distortion has been developed at Cardiff. As this uses coefficients to describe the load pull contours and because contours are assumed to be elliptical, fewer measurement points are required and typically only about 5 coefficients are necessary. This will reduce the model size and the measurement time. It may still however be impractical to measure every 2^{nd} harmonic termination impedance effect for every fundamental load impedance. This will be of great importance for further projects where high efficiency modes and the impact of inter-stage matching (where the load impedance can be further away from the optimum) are assessed. Rarely does a transistor sit in isolation. Measurements will need to be targeted at known or expected impedance areas. The models created can also be used in system simulators; however there is a problem with measurement based models used in individual stage design. In actual amplifier device line-ups the input stimulus will rarely be a single tone. Previous stages will also be nonlinear and therefore have a harmonic content. The extent of the problem this generates will depend upon the amplifier topology and band limiting components which will attenuate harmonics relative to the fundamental, and the extent of overdrive. For example, in base station applications driver stages will tend to be operated in the linear region and incorporate narrowband filters and circulators, thus reducing harmonics. They may, however have multiple tones closely spaced. A model generated from a single tone excitation will not necessarily accurately represent this behaviour. In contrast an electronic warfare (EW) jamming system may have many of the stages driven hard into compression to attain the maximum output power and flatness with frequency, hence having a high harmonic content. The measurement based models can include input harmonics, but this adds more dimensions to the data set (harmonic number and drive level). If we add to this measurements of more devices, so that any spread of performance across time and different wafer runs can be included, rather than depending upon the results from just one device, then one can see how the size of a truly comprehensive model can quickly ‘explode’. The behavioural models are able to handle this better than the DLUT as they are starting from a smaller base and if a relationship can be found between, for example, device variability and the Fourier coefficients, then these will be able to handle the distribution more economically. However more input tones

(harmonics or other signals) will require larger data sets even for these models, especially when relative phase is added to the scenario. Some degree of worst case and typical comparisons will probably be necessary.

Measurement based modelling has the additional advantage that characteristics causing performance changes away from the ideal can be incorporated without necessarily understanding the cause. It has been observed that the input reflection coefficient changes as the drive level increases, Figure 5-17(c); this is presumed to be due to changes of C_{gs} and C_{gd} [25], and as can be seen in Figure 5-23 and Figure 5-26 the PDK model does not replicate this behaviour very well. However the DLUT and other measurement based models accurately replicate the changes.

5.7 Summary

The history and progress in device modelling has been described. It can be seen as an iterative process; especially in the beginning where the complexity of the models was increased to account for what was seen as deficiencies in earlier models. Recent developments have seen changes in a more fundamental nature with the PHD and X-parameter models seeking to describe measured behaviour with parameterised equations. These have been enabled by the increase in numerical processing and the developments in nonlinear measurement capability, especially phase coherent harmonic measurements and harmonic load pull. The developments at Cardiff have been as a result of a change to the perceived wisdom of frequency domain measurement to one of the time domain. This return to a more fundamental analysis of the voltage and current waveforms in device operation (which RF and microwave engineers have tended to lose mainly due to the way that measurement equipment operated) came as the understanding of higher efficiency operation in terms of waveforms became more common, championed by the likes of Cripps [40]. This has seen a burst of activity in the industrial arena with new companies such as Mesuro and NMDG entering the test market alongside an increase in the capabilities offered by the existing players Agilent, Maury, Focus and Rohde & Schwarz.

Similarly, in the CAD area developments in nonlinear simulation approaches has led to the refinement of Harmonic Balance and increases in processing speed have made even optimisation of nonlinear circuits possible. The problem has been with believing the results

of such simulations, primarily due to a lack of faith in the nonlinear models. Manufacturers are very reticent to release the conditions under which models were developed or the extent to which they have been proved. Designers often do not even know over which frequency ranges the models operate. As has been shown models can accurately predict the fundamental performance but do less well with harmonic behaviour. This should not however come as any great surprise; those creating the models do so to a budget which limits the capabilities of the measuring equipment and the time available. Also as models improve so customers expect more. The author recalls saying to a nonlinear CAD supplier in the early '90s that he would be happy with a model that accurately predicted the fundamental output power. Now we are looking for not only the power of the harmonics but accuracy in the phase as well. Although the DLUT method described in this chapter could be used to acquire massive data sets, it is unlikely that they would ever be all-encompassing (including all permutations of frequency, load, bias, driver, etc.). Further, such a deluge of data does not in itself point towards the best design solution. Instead targeted measurements which are design driven are a more attractive and efficient approach.

Device manufacturers will find these models of benefit because they can be produced quickly; however they are unlikely to become so responsive to customer requests that they will provide custom models to each application. With the current test equipment costs the advantage of these models are most applicable to design processes where design iterations are expensive and processing takes many weeks (such as MMIC design); here being able to accurately predict the performance of active devices will make major strides in reducing development time and cost. However, in the future as test system costs are reduced the approach will become accessible to more amplifier development teams, which will enable the engineers to create their own accurate models under the operating conditions relevant to their designs, which is the fundamental proposal of this work, that measurement, simulation and design be integrated.

The design process that is advocated by this thesis moves the responsibility for modelling away from a traditionally remote and independent function firmly into the amplifier design arena. The DLUT model has been shown to be very accurate within the operating parameter boundary; and there should be no need to simulate outside this boundary, instead the envelope itself should be expanded to encompass the required

conditions. Extrapolation is by its very nature risky, moving out of the area of the known into the unknown, whereas the model has been shown to handle interpolation very well, given a sufficient primary data set.

Access to this approach will lead to the movement of the modelling segment away from the start of a linear design process and into an integral part of the design cycle. This is the subject of the next chapter.

5.8 Bibliography

[1] Y. Takayama, "A new load-pull characterization method for microwave power transistors," in *International Microwave Symposium*, 1976.

[2] R. Soares, M. Goudelis and E. d. L. Reyes Davo, "Novel measurement technique allows full two port characterisation of GaAs power MESFETs," in *10th European Microwave Conference*, 1980.

[3] I. Bahl, "0.7 - 2.7GHz 12W Power Amplifier MMIC developed using MLP technology," *IEEE Transactions on Microwave Theory and Techniques,* vol. 55, pp. 222-229, February 2007.

[4] S. Cripps, "A Method for the Prediction of Load-Pull Power Contours in GaAs MESFETs," in *International Microwave Symposium*, 1983.

[5] E. Griffin, "Application of Loadline Simulation to Microwave High Power Amplifiers," *IEEE Microwave Magazine,* vol. 1, pp. 58-66, June 2000.

[6] I. J. Bahl, "Ku-band MMIC Power Amplifiers developed using MSAG MESFET Technology," *Microwave Journal,* vol. 49, no. 2, February 2006.

[7] L. Besser, "Skillful De-embedding," *Applied Microwave & Wireless,* pp. 86-90, Fall 1994.

[8] B. Virdee, A. Virdee and B. Banyamin, Broadband Microwave Amplifiers, Artech House, 2004, pp. 133-135.

[9] S. Yanagawa, H. Ishihara and M. Ohtomo, "Analytical Method for Determining Equivalent Circuit Parameters of GaAs FET's," *IEEE Trans MTT,* vol. 44, no. 10, pp. 1637-1641, October 1996.

[10] M. Edwards, “AWR's Support of Polyharmonic Distortion and Nonlinear Behavioural Models,” [Online]. Available: http://web.awrcorp.com/content/Downloads/AWR-Nonlinear-Modeling-White-Paper.pdf. [Accessed 2011].

[11] S. Maas, “Computer-aided design of power amplifiers,” in *Handbook of RF and Microwave Power Amplifiers*, Cambridge, Cambridge University Press, 2011.

[12] D. Root, “A measurement-based FET model improves CAE accuracy,” *Microwave Journal,* pp. 126-139, September 1991.

[13] S. Maas, “Why I hate base resistance,” *Microwave Magazine,* June 2004.

[14] W. R. Curtice, “A MESFET model for use in the design of GaAs integrated circuits,” *IEEE Trans. on Microwave Theory and Techniques,* Vols. MTT-28, pp. 448-456, 1980.

[15] Y. Tajima, B. Wrona and K. Mishima, “GaAs FET Large Signal Model and Its Application to Circuit Designs,” *IEEE Trans Electronic Devices,* Vols. ED-28, February 1981.

[16] W. R. Curtice and E. M., “A nonlinear GaAs FET model for use in the design of output circuits for power amplifiers,” *IEEE Trans. on Micorwave Theory and Techniques,* Vols. MTT-33, pp. 1383-1394, 1985.

[17] H. P. Statz, “GaAs FET device and circuit simulation in SPICE,” *IEEE Trans. on Electron Devices,* Vols. ED-34, pp. 160-169, 1987.

[18] A. J. McCamant and G. D. S. D. H. McCormack, “An improved GaAs MESFET model for SPICE,” *IEEE Trans. on Microwave Theory and Techniques,* Vols. MTT-38, pp. 822-824, 1990.

[19] R. B. Hallgren and L. P. H., “TOM3 Capacitance Model: Linking large- and small-signal MESFET Models in SPICE,” *IEEE Trans. on Microwave Theory and Techniques,* vol. 47, no. 5, pp. 556-561, May 1999.

[20] FitzPatrick, “Design of an Optimally Flat Pulsed Microwave Power Amplifier,” in *Military Electronics Show*, Baltimore, 2003.

[21] S. Cripps, “Taking Heat,” *Microwave Magazine,* vol. 12, no. 2, pp. 24-31, April 2011.

[22] P. Ladbrooke, N. Goddship, J. Bridge and D. Battison, “Dynamic I(V) measurement of contemporary semi-conductor devices,” *Microwave Engineering Europe,* pp. 23-33, May 2000.

[23] L. Dunleavy, “Transistor and Amplifier Modeling Methods of Microwave Design,” in *IMS Workshop*, 2008.

[24] R. Sakamoto, A. Kiyoyasu and I. Masataka, “Real-space transfer and Hot-electron transport properties in III-V semiconductor heterostructures,” *IEEE transactions on Electron Devices,* vol. 36, no. 10, October 1989.

[25] A. Grebennikov, RF and Microwave Power Amplifier Design, McGraw-Hill, 2004.

[26] Y. Tajima, “Progress in RF device modelling: from MESFETs to GaN PHEMTs,” in *Advnaced Device Modeling Symposium*, National Nano Device Lab, Taiwan, 2006.

[27] D. Root, S. Fan and J. Meyer, “Technology Independent Non Quasi-Static FET Modles by Direct Construction from Automatically Characterised Device Data,” in *21st European Microwave Conference*, Stuttgart, 1991.

[28] J. Verspecht and D. Root, “Polyharmonic Distortion Modeling,” *IEEE Microwave Magazine,* pp. 44-57, June 2006.

[29] M. Edwards, “Nonlinear behavioural models, measurement systems and their use in circuit simulators,” in *ARMMS Conference November*, Corby, 2009.

[30] J. Verspecht, J. Horn, L. Betts and D. Root, “Extension of X parameters to include long-term dynamic memory effects,” in *International Microwave Symposium*, Boston, 2009.

[31] J. Verspecht and P. Van Esch, “Accurately characterising hard nonlinear behaviour of microwave components with nonlinear network measurement system:Introducing 'nonlinear scattering functions',” in *International Workshop Integrated Nonlinear Microwave Millimeterwave Circuits*, Germany, 1998.

[32] C. Baylis, R. Marks, J. Martin, H. Miller and M. Moldovan, “Going Nonlinear,” *Microwave Magazine,* vol. 12, no. 2, pp. 55-64, April 2011.

[33] S. Woodington, T. Willimas, H. Qi, D. Williams, L. Pattison, A. Patterson, J. Lees, J. Benedikt and P. Tasker, “A novel measurement based method enabling rapid extraction of a RF waveform look-up table based behavioural model,” in *International Microwave Symposium Digest IEEE MTT-S*, 2008.

[34] S. Woodington, R. Saini, D. Williams, J. Lees, J. Benedikt and P. Tasker, “Behavioural model analysis of active harmonic load-pull measurements,” in *International Microwave Symposium Digest IEEE MTT-S*, 2010.

[35] P. Tasker and J. Benedikt, “Waveform inspired models and the harmonic balance emulator,” *Microwave Magazine,* vol. 12, no. 2, pp. 38-54, April 2011.

[36] D. Bespalko and S. Boumaiza, “X-parameter measurement challenges for unmatched device characterisation,” in *75th Microwave measurements Conference (ARFTAG)*, 2010.

[37] G. Simpson, “High power load pul with X parameters - a new paradigm for modelling and design,” in *IEEE Wireless and Microwave Technology Conference*, Melbourne, Fl., 2010.

[38] R. Saini, S. Woodington, J. Lees, J. Benedikt and P. Tasker, “An intelligence driven active load pull system,” in *75th Microwave Measurements Conference (ARFTAG)*, 2010.

[39] H. Qi, J. Benedikt and P. Tasker, “Nonlinear data utilization: from direct data lookup to behavioural modeling,” in *IEEE Trans Microwave Theory & Techniques*, 2009.

[40] S. Cripps, RF Power Amplifiers for Wireless Communications, 2nd Ed., Artech House, 2006.

6 MMIC Design Using Active Harmonic Load Pull

"There is never enough time to do it right the first time, but there is always enough time to go back and do it again"

Truism, originator unknown, but probably Adam!

6.1 Introduction

One of the most revealing conversations held during the period of this research was during the early days with one of the sponsoring industrial partners. This company designed their own MMICs (Monolithic Microwave Integrated Circuits) and had a considerable reputation in the field. During the discussion it was explained that:

- They would normally expect to make several versions of a design and see which one came nearest to the requirements.
- They would accept that there would be several iterations (wafer runs) in order to refine the design.
- The manufacturer's models had 'short-comings', they would therefore pay a 3rd party to create specific models, but budget would limit how many transistors they could do this for and the range of the environment (bias, temperature and drive level) that would be covered.
- That often models were software version specific, i.e. if the software were upgraded they quite expected to go through a period of de-bugging the issues.

The result was that there was a reticence to change beyond minor 'tweaks' to existing designs. Any advances were of an iterative nature. Further the company tended to use those devices for which they had tried and tested experience rather than the optimum ones indicated by theoretical frequency and power performance. The company were very well aware of the short-falls of this approach and bearing in mind that a wafer run would cost in excess of £50,000 and rarely be shorter than 6 weeks (just to get the manufactured wafers back, let alone tested), it is hardly surprising that they were interested in the possibilities the measurement system offered. Furthermore, many of the devices are used in RF systems for emerging ultra-wideband radar products (6-18GHz) where any improvement in output

power and efficiency of the final RF transistor stages will effectively minimise the power consumption, thermal dissipation, and the size of RF system as a whole (for a given performance), resulting in benefits to a wide range of applications with a particular relevance to mobile or field-deployable units, which are effected by significant size and power consumption constraints. The standard approach is summarised in Figure 6-1. Note that it is typically not possible to alter the foundry PDK models and so often any refinement consists of adding external components or altering the bias to fit the behaviour observed in practice.

To overcome the incremental improvements in the design methodology of these military RF systems, the research sought to extend the techniques developed for the optimisation of mobile communications technologies [1], to the higher frequencies and wider bandwidths of interest. These techniques, based on the novel measurement concept, allow complete access to, and manipulation of, the information contained in the RF waveforms. The approach involves the measurement and engineering (including dynamic load line control) of current and voltage waveforms and represents a major departure from currently established non-linear design and analysis methods. It is important to note that despite its novelty, the information provided still allows direct reference and comparison with past and present design techniques and thus allows the inclusion of the significant know-how which has been developed in the last few decades.

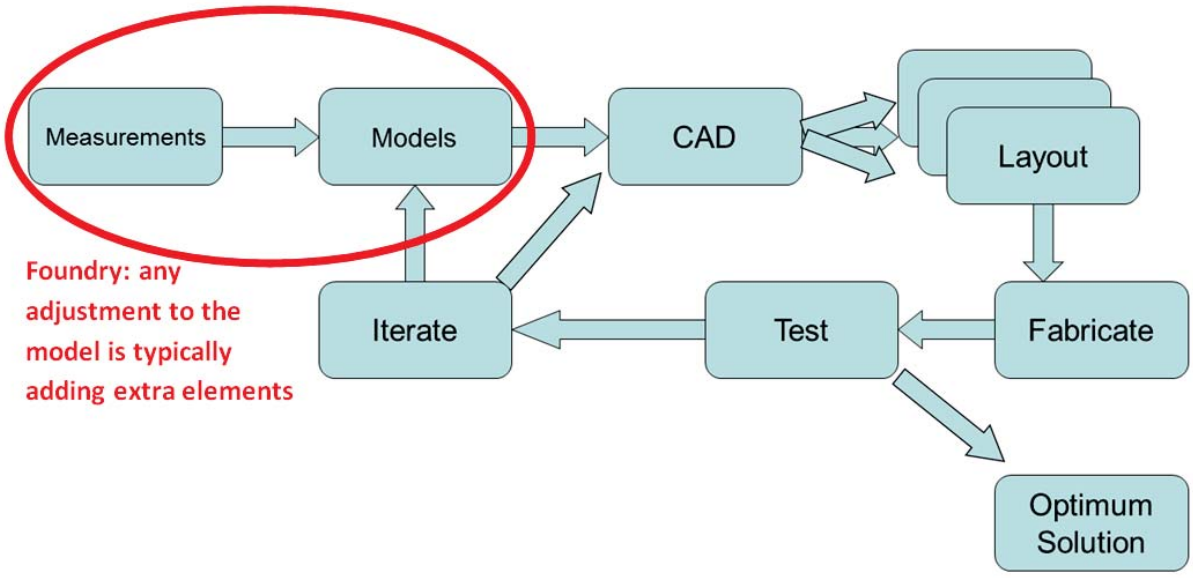


Figure 6-1, Typical Process Flow for the MMIC Design.

To fully utilise the additional information, which is directly obtained from the measurement system, a new development methodology is required. This integrates the measurement, modelling and CAD domains, previously seen as separate, into a cohesive systematic approach, resulting in not only optimum performance but also a high confidence in success.

As a first step the measurement system can be used to produce accurate models under the actual conditions required by the design. This will of itself allow the optimum device to be used (rather than for which there is known good model) and, provided the foundry devices are repeatable, a better solution. This is shown in Figure 6-2; by the replacement of the "Models" block from Figure 6-1, by the orange "Look Up Based Model". It is worth noting that it would be necessary to conduct measurement in the first place to verify foundry nonlinear models so this is not introducing any additional tasks. The benefit in this case is that the information acquired can be used in the design simulation itself. The obvious next step is to incorporate the measurement system within the design process – moving it "on-line" – so that as output matching solutions are developed the impedances they create are applied to the device and the actual performance assessed. The ultimate objective is of course to achieve a first pass design success. We are not so naive as to suggest that simply by adopting the process described here will MMIC design become a straightforward and trivial process. There are many other factors that come into play such

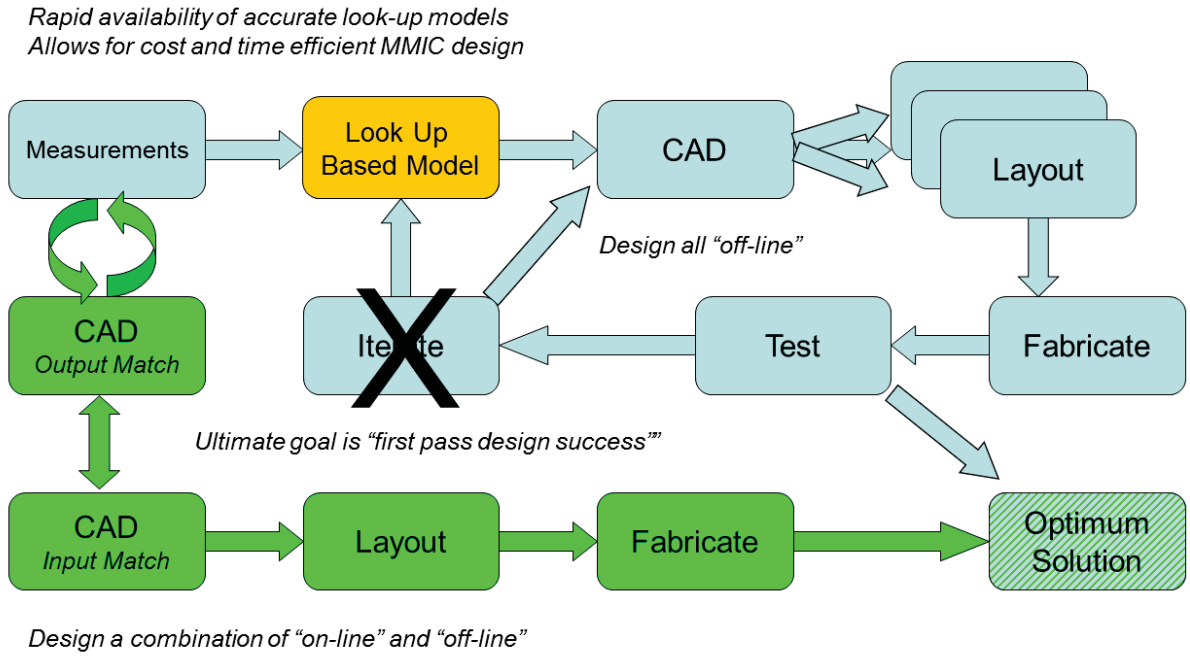


Figure 6-2, Improved process based on measurement system, initially through better models, then through incorporating the measurement system in the design process.

as odd mode stability, parasitic element coupling and thermal behaviour; however as CAD tools become more comprehensive (we now have 3 and 2.5D E-M simulators to tackle the short falls in distributed element models and the implications of coupling) so the accuracy of our simulations will improve. The method developed in this research is another weapon in the MMIC designer's armoury, taking this analogy further; it is equivalent to moving from bows and arrows to sniper rifles – faster and more accurate. Also the approach allows a subject that hitherto has been largely ignored in high frequency broadband MMIC design, the impact of harmonic terminations on device performance, to be incorporated.

Returning to the quote at the start of this chapter, one of the obstacles that inevitably will be put in the way of the adoption of the methodology described here is that there isn't sufficient time or resources to invest more in the upfront design stages. The virtually inevitable consequence of which is that more time and expense will be incurred at the backend of the design process, or that a poorer performance is accepted.

Through the use of the engineering of waveforms the impact of changes to harmonic terminations can be clearly seen. This is important to wideband designs both in the inter-stage match between transistors and the impedance presented by the following components in the system, such as antennas. As will be seen the impact of 2^{nd} harmonic terminations can alter Power Added Efficiency (PAE) by ±10%, and this could explain why in previous amplifier designs PAE performance has been lower than expected [2], [3], etc. (when only considering the fundamental frequency).

6.2 Design Approach

The starting point for any design is the requirement or system specification. This should outline the minimum performance and physical characteristics, output power, efficiency, gain, size, weight, etc. as well as targets or 'nice to have' details. It is not necessary to go into system specifications in detail here, however these must form the backbone of the design and all decisions should conform to achieving the requirements defined in the system specification.

The drivers of the amplifier architecture are not only technical; commercial and risk factors also come into play. To some extent these are judgement calls, hence different design teams will take different approaches. Take for example a requirement for a 6-18 GHz

10W MMIC amplifier; a purely technical approach may suggest opting for a GaN solution, a commercial consideration may push for far-eastern or European GaAs foundry to avoid ITAR[1] restrictions, whilst a risk reduction consideration may push for a particular GaAs foundry with a tried and tested process.

The outcomes of the architecture design stage should be:

- A selected process and foundry.
- A power-budget; the number of stages and the distribution of gain and power.
- Size, including whether the design can best be met with a single chip or multiple devices.
- Cost, this is largely related to wafer real-estate and technology (GaAs or GaN).
- Test strategy, especially if things don't work first time how will the problem be identified.
- Risk assessment, and what options are necessary to minimise critical risk factors.
- Time scales, probably with a number of options depending upon the risk assessment.

For MMIC Power Amplifiers (PAs) the choice of process from a technical point of view can be boiled down to maximum frequency (hence gate length) and power (W/mm). However factors such as export restrictions can rule out many foundries. Management caught out in the past by export restrictions put on products and processes, which had been readily available are extremely cautious on this front. It is not unknown for a specification to explicitly prohibit the use of ITAR components; however the technical advantages offered by these processes make them extremely tempting!

The general approach is to list the foundries from whom one is prepared to purchase and tabulate the gate length of their process and the W/mm. This latter parameter may be harder to compare as some degree of 'specmanship' goes on and it may not be clear that comparisons are identical. With regards to gate length, also the difference between processes needs to be kept in mind. The drain capacitance of GaN is significantly lower than that of GaAs for an equivalent output power (higher W/mm = less periphery required = lower output capacitance) capability and hence its frequency performance will generally be

[1] ITAR: International Traffic in Arms Regulations, U.S.A. government restrictions on the trade in components that could be used in military products. Besides the obvious prohibitions the paperwork involved in getting approval for using components even in clearly no-military applications has led many companies to specifically prohibit the use of such U.S. components in their products.

better, however other parasitic capacitances such as those of the field plate may cause degradation as frequency increases. The quality and reliability of the process models will be another factor, as Steve Maas said "*Given a choice between an unmodelled device with super performance and a well-modelled one with more prosaic performance, I'll usually choose the latter because I'll almost always get better results with it. Perhaps I'm not very adventuresome, but I like to know what I'm doing. A device may have wonderful potential, but if I don't know how to realize it, I probably won't be able to.*" [4]. A problem with this though is that it tends to rule out new processes as it takes time and data to create a good model set, (this is speaking in a current industry sense rather than applying the method described in chapter 5).

Of course in the end the determining factor may be cost, the price of a particular foundry run may just be too prohibitive to be used (especially as there are no guarantees of success). This area is commercially sensitive and also rapidly changing and so will not be discussed further here.

The output power of the amplifier determines the gate periphery required, however it is also necessary to take into account the loss of the output matching circuit which is made up of the insertion loss of the structure and the mismatch loss due to circuit reflection. From Fano's theory [5] it is not possible to perfectly match a reactive circuit over an infinite bandwidth. Generally speaking the operating bandwidth is defined by the Quality Factor, Q, of the transistors output admittance (Y_{out}) which is mostly the result of the equivalent drain resistance, R_{ds}, and the output capacitance, C_{out}, (a combination of C_{ds} and C_{gd}). Assuming a required upper frequency of F_U and a lower frequency of F_L the following equations describe the output circuit reflection coefficient, Γ, from which the decrease in output power can be determined.

From {6-4} we see that as the bandwidth-Q product decreases the magnitude of Γ also decreases; which makes sense; over a narrower bandwidth or with a lower Q output load we should be able to achieve a better match.

$$Y_{out} = G_{out} + j\omega C_{out} \qquad \textbf{\{6-1\}}$$

$$BW = \frac{F_U - F_L}{\sqrt{F_U \times F_L}} \qquad \textbf{\{6-2\}}$$

$$Q = \frac{2\pi C_{out}\sqrt{F_U \times F_L}}{G_{out}} \quad \{6\text{-}3\}$$

$$|\Gamma_{max}| = e^{\left(\frac{-\pi}{Q \times BW}\right)} \quad \{6\text{-}4\}$$

$$Insertion\ Loss\ (dB) = 10\log(1 - \Gamma^2) \quad \{6\text{-}5\}$$

Plotting {6-5} in Figure 6-3 the insertion (mismatch) loss resulting from the reflection coefficient (and hence the equivalent return loss) can be seen. Note that this does not include the physical circuit loss, but purely that resulting from reflections. Also it should be

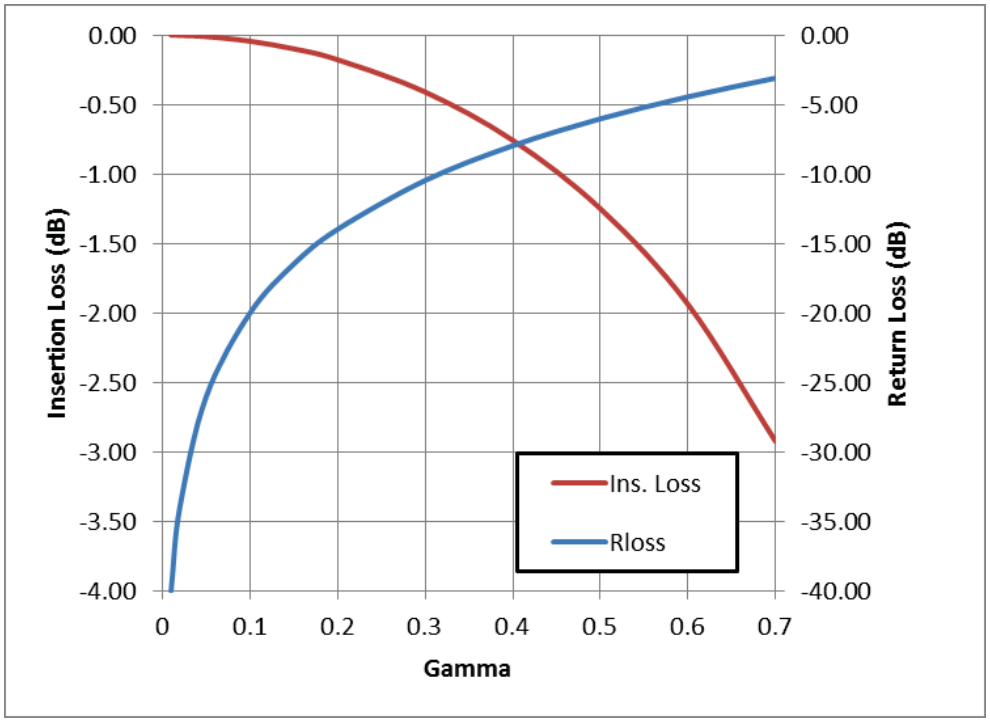


Figure 6-3, Relationship between Insertion & Return Loss and Reflection Coefficient (Γ).

remembered that this is the worst case, there will be points when matching across a wide bandwidth where a better match can be achieved and hence less loss.

Some interesting results come from these relationships, if we consider the 1.25mm GaN device load-pulled on the measurement system we find that the equivalent output circuit for the optimum output power corresponds to an equivalent circuit of a shunt resistor and capacitor (51Ω and 0.46pF), whilst for the optimum PAE the equivalent circuit is 85Ω and 0.47pF. Using the equations {6-2} to {6-5} and applying them to a 6 - 18 GHz amplifier, we see that if we design to the optimum output power load the mismatch loss is very low,

<0.15dB, whilst designing to the optimum PAE load we inherently get over 0.4dB mismatch loss, Table 6-1, and as the increase in PAE was only ~6% it has pretty much all been lost in the mismatch. Furthermore, generally the more complex the matching circuit the higher the associated circuit losses, so in this case it is likely that designing to the optimum PAE load will result in a lower PAE than if the optimum output power load were targeted.

Parameter	Opt. Pout	Opt. PAE	Unit
Rds	51	85	Ω
Cout	0.46	0.47	pF
Upper frequency	18	18	GHz
Lower frequency	6	6	GHz
Q	1.53	2.61	
Rel. BW	1.15	1.15	
Gamma Max	0.17	0.35	
Return loss	15.43	9.06	dB
Mismatch Loss	-0.13	-0.58	dB

Table 6-1: Effect of Rd and Cout on Mismatch loss,

Further if we consider {6-3}, and replace F_L as a fraction of F_U and also note that the ratio of C_{out} to G_{out} is constant with gate periphery we get {6-6} where x is the lower frequency fraction of the upper (in the case of 6-18GHz, x = 3). Hence we can see that Q and so Γ depend upon F_U only (with a particular process technology, obviously if the ratio of C_{out} to G_{out} changes Q will change – this leads to an useful comparison between processes, the ratio of C_{out} to G_{out}).

$$Q = \frac{2\pi F_U}{\sqrt{x}} \times constant \qquad \{6\text{-}6\}$$

As mentioned, in addition to the mismatch loss an allowance for circuit losses is also necessary. Simulations will suggest that the loss of a λ/4 50Ω line at 20 GHz on a 100µm GaAs substrate is of the order of 0.1dB, as the line impedance increases the loss increases so that a 75Ω line is about 50% higher and conversely a 25Ω line is about 10% lower. Series elements such as decoupling capacitors also have loss, about 0.05dB. Although these amounts are small and the mismatch loss seems to dominate, as the number of elements increase not only do their contributions add up, but discontinuities are introduced which not only increases loss but narrows bandwidth, hence they should not be underestimated,

further experience has shown that these losses are on the optimistic side (see later in this chapter for output matching circuit losses) by a factor of up to two times. In summary the output power and the bandwidth determine the Power Periphery of the output stage(s).

So the question should arise, if Q is not related to periphery why don't we just build cells as large as we need for the power? The problem is that as device size increases gain decreases, which means that a higher drive level is required and hence more stages. There are a number of factors which cause this decrease in gain:

i. Increased parasitics as size increases, particularly capacitances.
ii. Varying distances from source terminals to ground via, or complicated track routing to accommodate 'internal' vias. Increasing source inductance results in negative feedback – the gain 'killer'.
iii. Increased phase variation between the combined signals due to varying path lengths through each gate finger. Similarly when there is significant phase variation along the width of a gate finger.
iv. Increased heating, higher temperature increases resistance (more carrier collisions with increased vibrations in the lattice). $G_m \alpha 1/T_j$.

It is important to note that these effects are frequency dependent. At low frequencies the increasing parasitics and phase lengths are less important than the thermal effects. At the higher (mm wavelengths) frequencies it is the parasitics and phase imbalance that are the limiting features and the thermal considerations are less of a problem. As has been discussed above, selecting the device with the right output impedance results in less power being lost both in mismatch loss but also in the loss of matching circuits (they may be able to be largely avoided). However, the device needs to have sufficient current capacity and an advantage of lower impedances is that they have a higher current capacity and wider line widths. Thus it may not be best to target a device with an output impedance at 50Ω. In some configurations where balanced amplifiers are used it may be possible to include an impedance transformation within the combiner, thus allowing for both higher current and minimal matching. Such combiners are however difficult to produce in MMICs with significant bandwidths.

The optimum device size can be summarised as that which produces the maximum output power whilst still having useable gain. A rule of thumb suggests that for practical

applications G_{max} >10dB, (always remember that rules are meant to be broken!). Increased gain increases the power added efficiency by reducing the input drive level, {6-7}, thus besides less power being required from the driver fewer overall stages may be needed in the RF amplifier line-up. This not only increases efficiency itself but reduces wafer area and therefore cost.

$$\eta = \frac{P_{out} - P_{in}}{P_{dc}} \qquad \{6\text{-}7\}$$

It is important that the driver stage shall not be the limiting factor in the output power. For linear amplifiers all of the compression should ideally occur in the output stage. For saturated amplifiers, although this is less critical it is still important that the output power performance is determined by the output and not the driving devices, although compression in the earlier stages can be used to improve output gain flatness and prevent the output device(s) from being overdriven. The clearest evidence of earlier stages compressing is a very 'soft' compression curve resulting from gain being lost through nonlinear input to output power transfer at more than one stage. This is partly where the 10dB gain rule comes from. If a device has a gain of 10dB then obviously the input power required is 10dB less than the output. Assuming that the saturated power compression is 3dB then the saturated drive level is 7dB below the output power. The significance of this is that theoretically the same device used as a driver could drive 4 output devices in parallel. In practice the problem comes from not being able to match into a perfect 50Ω load. The Γ of {6-4} assumes that the match is into a perfect load. In many systems, such as where the load is a wideband antenna, this is not the case. Similarly the driver stage is not seeing the same load that the output stage is (it is seeing the output stage's input through the splitter) which will be far from perfect. Thus mismatch losses are likely to be higher. Nonetheless if the gain is 10dB minimum across the frequency band then there will also be points where it is higher and to some extent the law of averages helps to balance things out. Thus the increased gain of GaN can provide more margin than GaAs designs of a similar architecture.

In the case of this research the selection of which foundry and process to use was determined by the project sponsors; being one that was preferred by them for wideband radar and Electronic Warfare (EW) applications. It was a GaAs DpHEMT technology having a maximum DC power density of 1.9 W/mm and a gate length of 0.3µm. The maximum V_{ds} DC

operation is 12V. The normal gate finger pitch is 25µm, however the project sponsor has access to a non-standard 35µm variant, which has slightly better thermal performance; however the RF model used is that of the 25µm pitch which is therefore an obvious source of error.

The device size is selected based on the output power required and maximum operating frequency. Two designs were undertaken as part of this work, the first 5-10GHz used a 6x100µm stage and the second 5-10GHz with harmonic enhancement, used a 10x75µm, the shorter gate width facilitating higher frequency operation.

6.3 5 – 10 GHz 26dBm Driver Stage Design

As a first step in proving the applicability of the novel design process, a device was chosen that was typically used in the industrial partner's products. The selection was also based upon the measurement capabilities of the system at the time. A major consideration is the drive power available; to evaluate the optimum PAE it is necessary to be able to supply a sufficient amount of drive power to the device, which is ideally high enough for the device to be several dBs into compression at the highest level. The signal generators incorporated into the system had a maximum output power in range of 21 to 25dBm depending upon frequency. Between these and the device input are cables, switches, coupler, bias tee and wafer probe. Despite efforts to keep losses to a minimum the system insertion loss of this line-up amounted to just over 4dB at 18GHz. On the input side the other significant factor is the reflection coefficient of the transistors gate. Figure 6-5 shows the effect input reflection coefficient has on mismatch loss, and the typical Γ_{IN} presented by microwave devices. This shows that this is often the dominant feature affecting drive level requirements.

No yield or variability information was available as would be the case for the introduction of a new technology. Nonlinear models for the devices do exist within the foundry Process Design Kit (PDK), however these were not used in the design. Initially a section of a previously manufactured wafer containing 6x100 µm device cells was measured. The resulting measurement data is interesting from a number of perspectives, firstly it allows interrogation of the device performance, secondly it allows optimisation of matching circuit designs and finally as GaAs is a mature technology, the measurements can

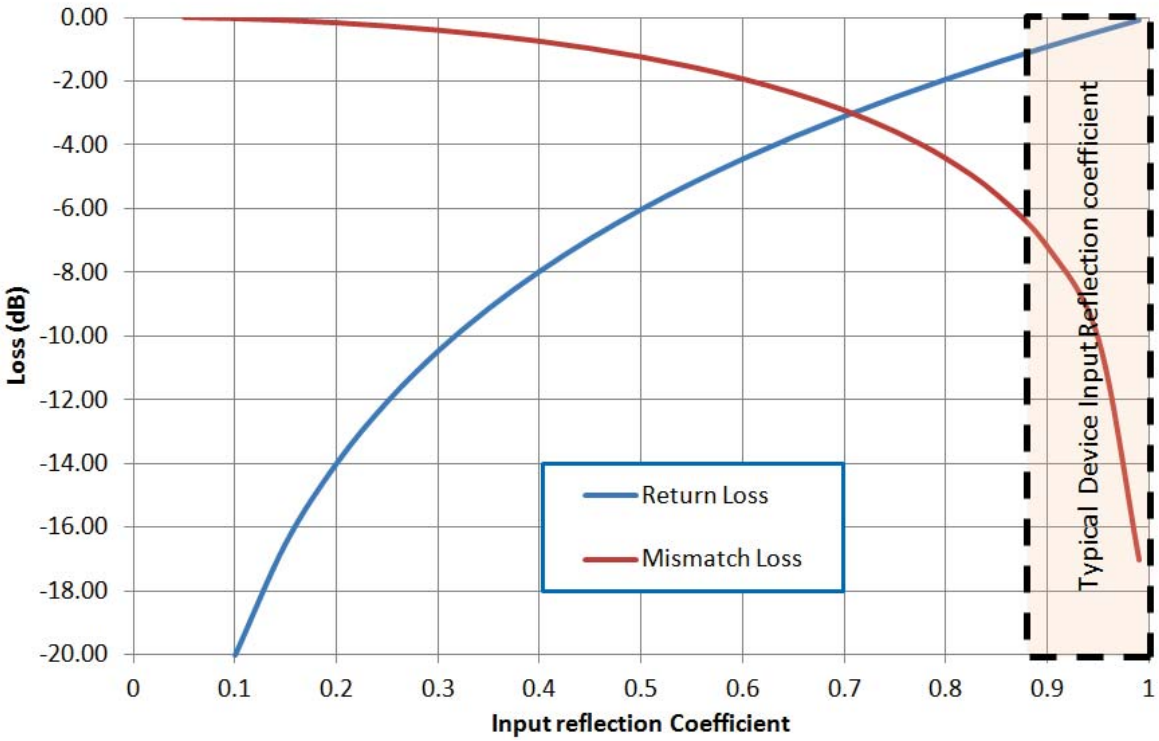


Figure 6-5, Impact of Input Reflection Coefficient on Mismatch Loss.

be used to validate the measurement system under large signal operation as we would expect the results observed to be in good agreement with theory.

The first measurement was performed to observe the dynamic RF load-line, with 50Ω fundamental and harmonic impedances, achieved by plotting output voltage versus output current and to compare it to the DC boundary conditions as the device is driven into compression. The result is shown in Figure 6-4, and allows analysis of any dispersion caused by the device parasitics. In this case little dispersion is observed, with virtually no difference

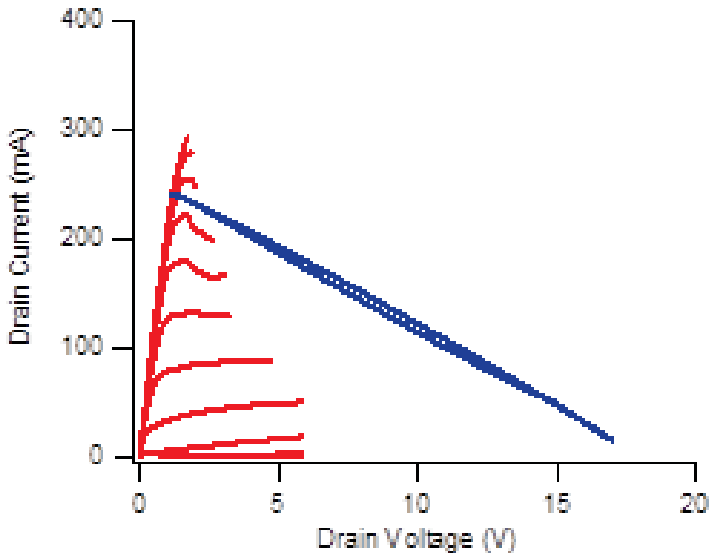


Figure 6-4, Dynamic Load-line analysis using measured DC and RF performance.

between DC and RF performance, this gives a good verification of large signal measurement system and calibration whilst providing positive feedback to the device manufacturer.

This type of analysis becomes particularly useful when characterising new device technologies such as GaN transistors, where dispersion mechanisms become far more prevalent. It has been shown that using such analysis enables separation (removal) of the dispersion mechanisms allowing observation of both knee walkout and soft pinch-off [6] - [7]. The measurement system can also be employed to perform automated power sweeps along with load-pull sweeps. Figure 6-6 shows the waveforms with the device biased in class B with 6 GHz stimuli, as an input power sweep is conducted into a 50Ω load impedance. The corresponding Pout versus Pin plot is shown in Figure 6-7. Further validation of the measurement system performance is achieved through observation of the waveform shape with the expected half rectified output current waveform confirming class B operation.

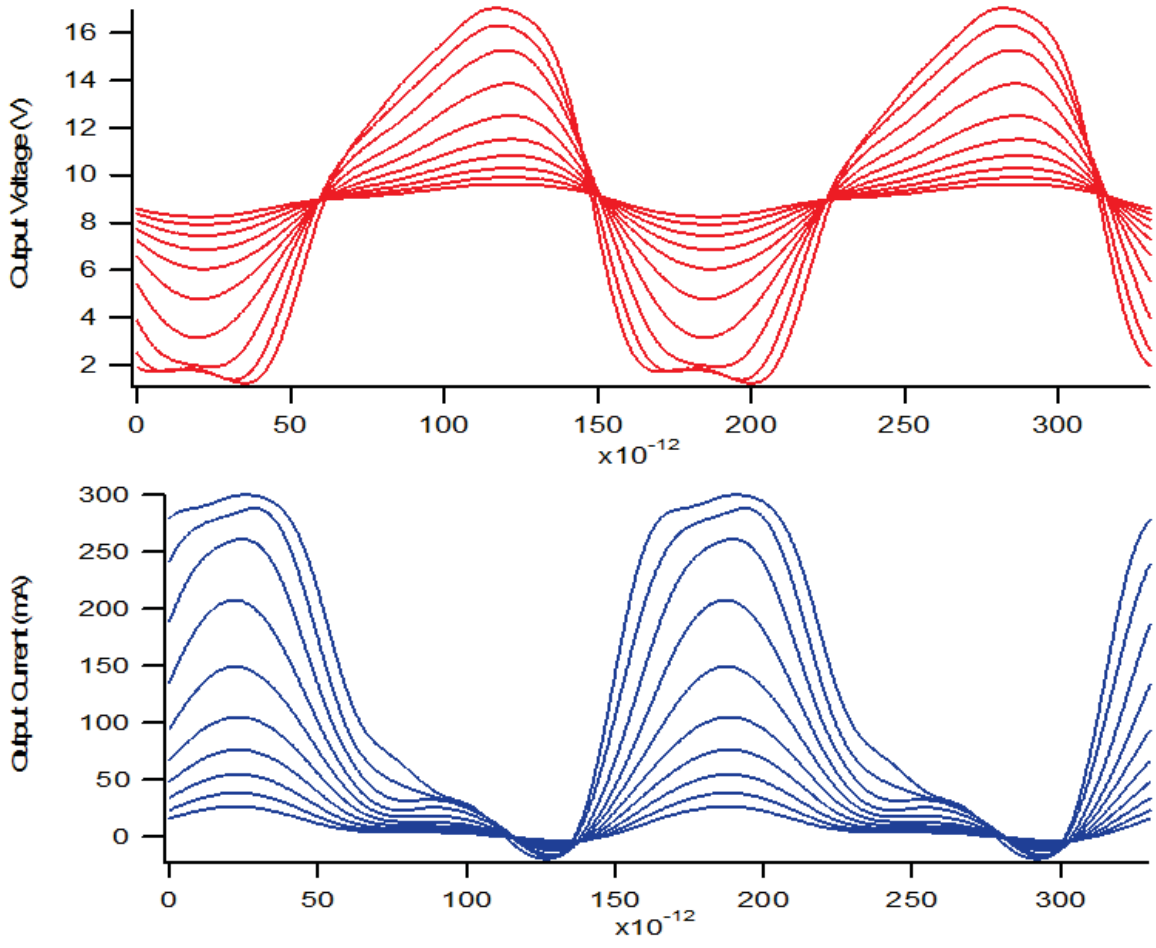


Figure 6-6, Measured class B output waveforms as device is driven into compression.

By combing automated power sweeps with load-pull it is possible to create power dependent contour plots showing the gain, maximum output power and efficiency. Figure 6-8 shows the measured contours of Power Added Efficiency (PAE) at 3dB into compression across a square grid on the Smith chart, in this case the device was biased in class A, (9V 120mA) as this was the standard mode used in the sponsors systems. It should be noted that these contour plots show the device efficiency for impedance presented at the probe tips. As discussed earlier the device is embedded between two line structures, it is therefore necessary to de-embed back to the device plane to see the performance of the device itself. Figure 6-10 shows the optimum device plane impedance required for maximum PAE at 5, 7.5 and 10GHz before and after de-embedding. These results indicate clean device performance with measurements agreeing excellently with theory for ideal device operation. It can be seen that the points lie almost perfectly on a circle of constant G of

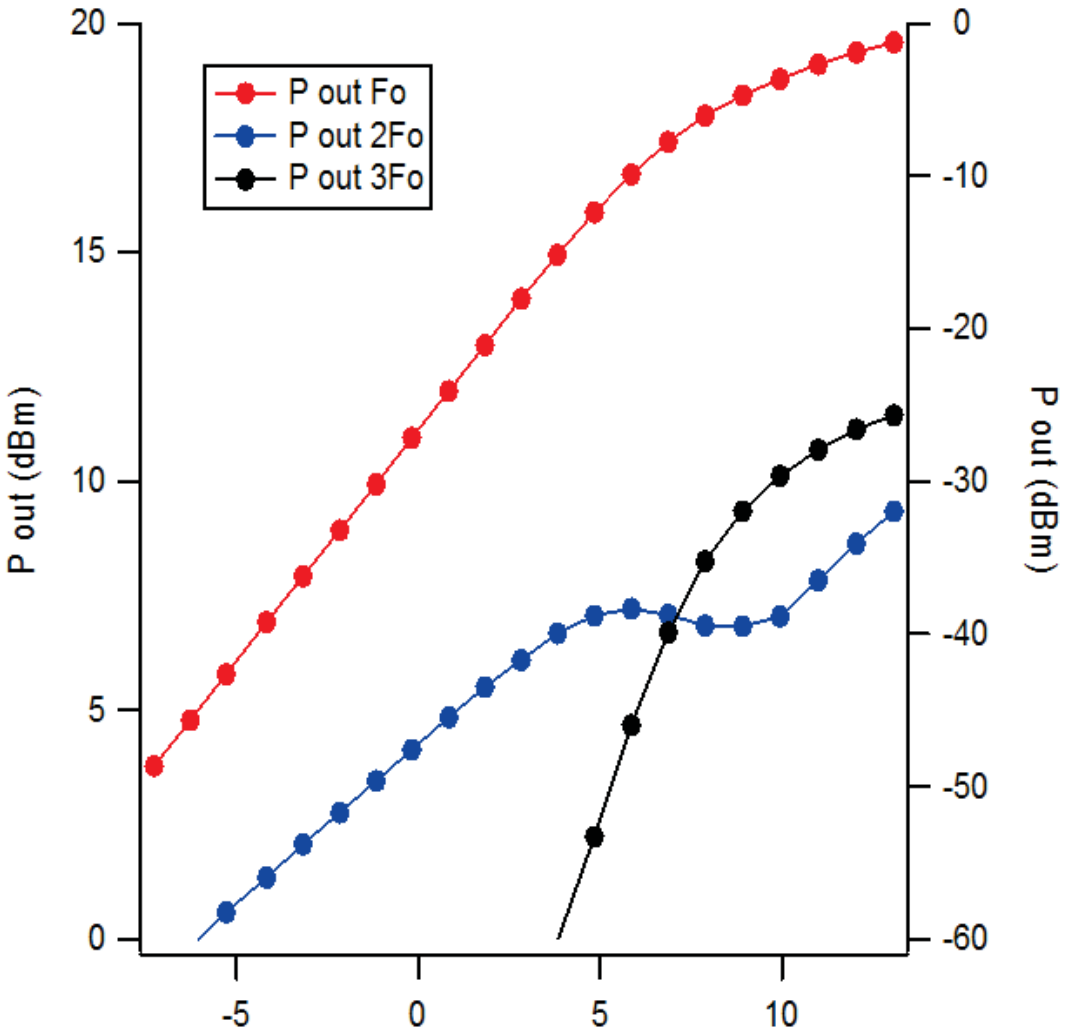


Figure 6-7, Measured fundamental (L) and harmonic (R) output power with swept input in class B at 6 GHz

0.55, and that the value of susceptance (B) doubles as expected with frequency. Such good agreement with theory also offers further validation of the large signal measurements. This data can also clearly be used to design an optimum output matching circuit to maximise efficiency. Using the values of G and B at each frequency in a standard formulation shows that the matching circuit for optimum efficiency is required to resonate out an output capacitance of about 0.3pF and provide a real load of 91Ω. Figure 6-9 shows the measured optimum load impedances, the de-embedded load impedances and the conjugate match of the simple equivalent circuit (shunt R and C) for the output of the device. The agreement between the measured device loads and the generally accepted device output equivalent circuit was in itself re-assuring and gave confidence in the system operation.

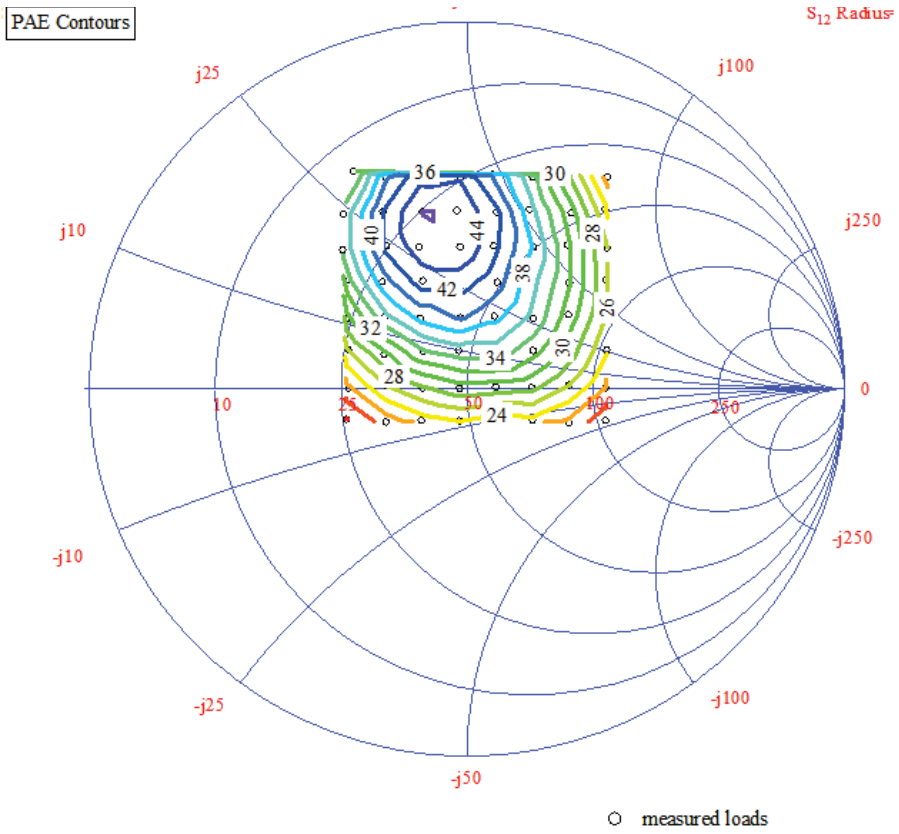


Figure 6-8, Measured Load pull contours at 6GHz at the measurement plane.

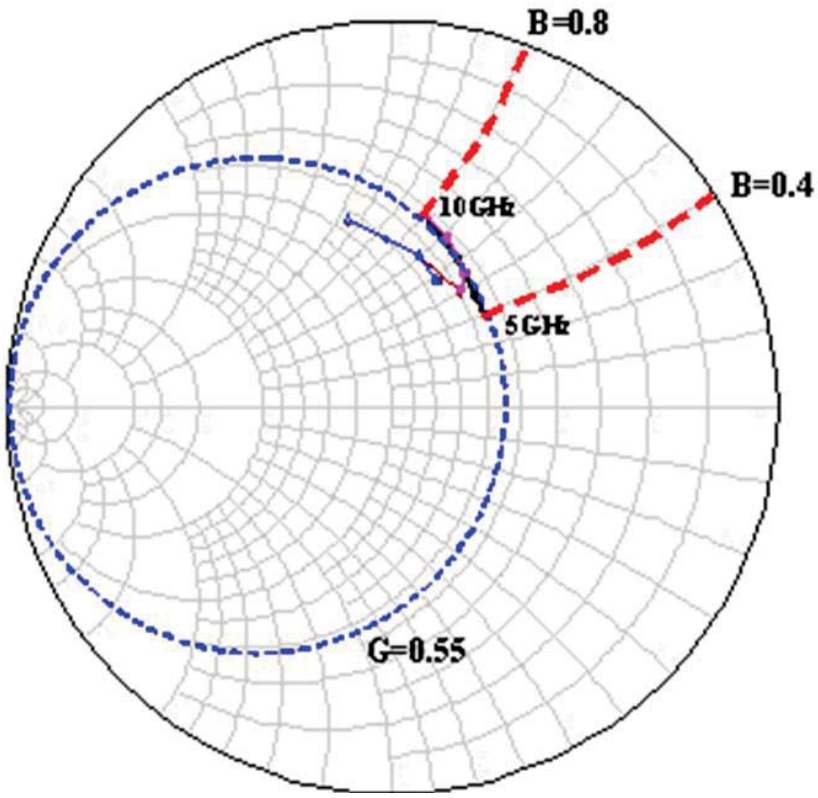


Figure 6-10, Optimum PAE impedances measurement and device plane at 3dB compression

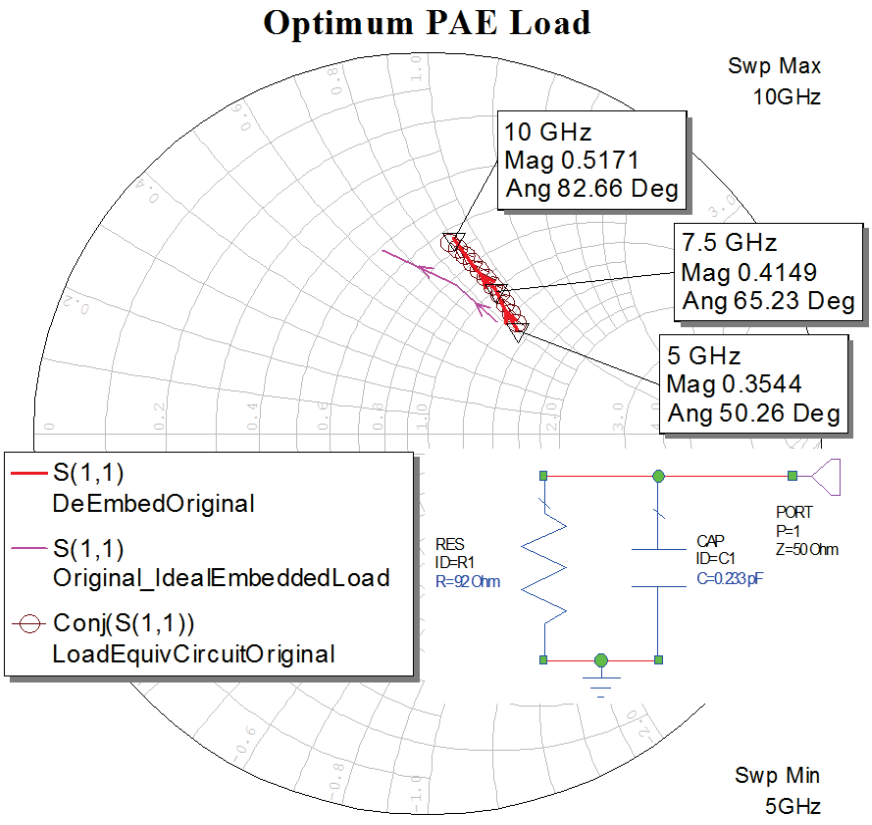


Figure 6-9, Optimum PAE Load impedance for 6x100 at 9v 120mA bias, (pink) and these loads de-embedded to the device plane (red). The conjugate of a simple parallel R-C network (Brown circles) proves in good agreement with the measured device load.

The measured results at the optimum PAE loads are summarised in Table 6-2. Based on the performance of previous designs a target across the operating bandwidth of 26dBm saturated output power and 35% PAE were set (allowing 0.5dB matching circuit loss). Also by way of comparison the measured data was supplied to the sponsor who then implemented two designs, the first using the standard S parameter based matching technique (extracted from the PDK model) and the second based on the impedance and performance data measured. The resulting load matches are shown in Figure 6-11, this is the load presented to the stage under consideration and also the input match of the following stage; the standard approach only seeks to achieve a minimum input return loss from the following stage across the operating frequency band (this is for a driver circuit, not a power stage) but consequently presents impedances which are significantly far from the optimum PAE. Although the driver stage has a secondary effect on the overall efficiency, depending on the gain of the output stage, it will still contribute and as can be seen by using the information available the best performance through the system can be achieved. The question arises as to why was the PAE predictions from the PDK model are not used in standard design; the answer given was that there was little confidence in the nonlinear model but experience had shown that the S parameter data (small signal) was reasonably accurate.

DpHEMT 6x100 at 9v 120mA				
Frequency (GHz)	Γ (mag/ang)	PAE (%)	Pout (dBm)	Gain (dB)
5	0.35/50.3°	50.6	27.1	14.8
7.5	0.41/65.2°	49.8	26.9	12.3
10	0.50/82.7°	46.1	26.5	10.4

Table 6-2, Measured performance at Optimum PAE load impedances.

Wide bandwidth impedance matching requires some compromise as it is a very difficult (if not impossible) task to achieve the exact desired impedances at each frequency. More matching elements generally increase loss, which in turn degrade output power and efficiency. Further, there are limits on the achievable transmission line impedances, as well as available circuit area.

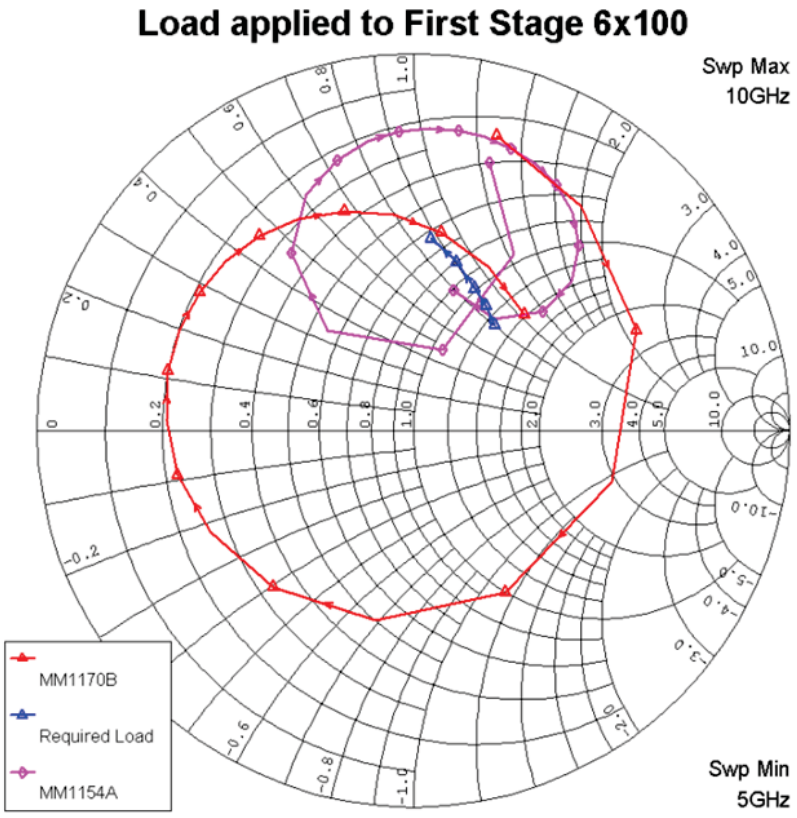


Figure 6-11, Load impedances presented by standard (red) and measured data (pink) with target loads (blue).

One circuit solution derived has an impedance trajectory as shown in Figure 6-12. In this case the compromise is that the performance at the band edges is closer to the optimum than at the centre. Measuring the device with the impedances of the output matching circuit predicted by the circuit simulator, the performance was found to be within target and is summarised in Table 6-3. The gain quoted is maximum gain and accounts for power lost due to device input mismatch. Also the PAE at this stage is optimistic as it considers only the load affects and ignores circuit losses, as well as assuming that the device can be perfectly input matched.

DpHEMT 6x100 at 9v 120mA				
Frequency (GHz)	Γ (mag/ang)	PAE (%)	Pout (dBm)	Gain (dB)
5	0.31/85.2°	47.8	27.1	15.4
7.5	0.25/102.3°	44.0	27.2	11.8
10	0.37/99.3°	44.7	26.9	10.2

Table 6-3, Measured performance at fundamental load impedances from simulated matching circuit.

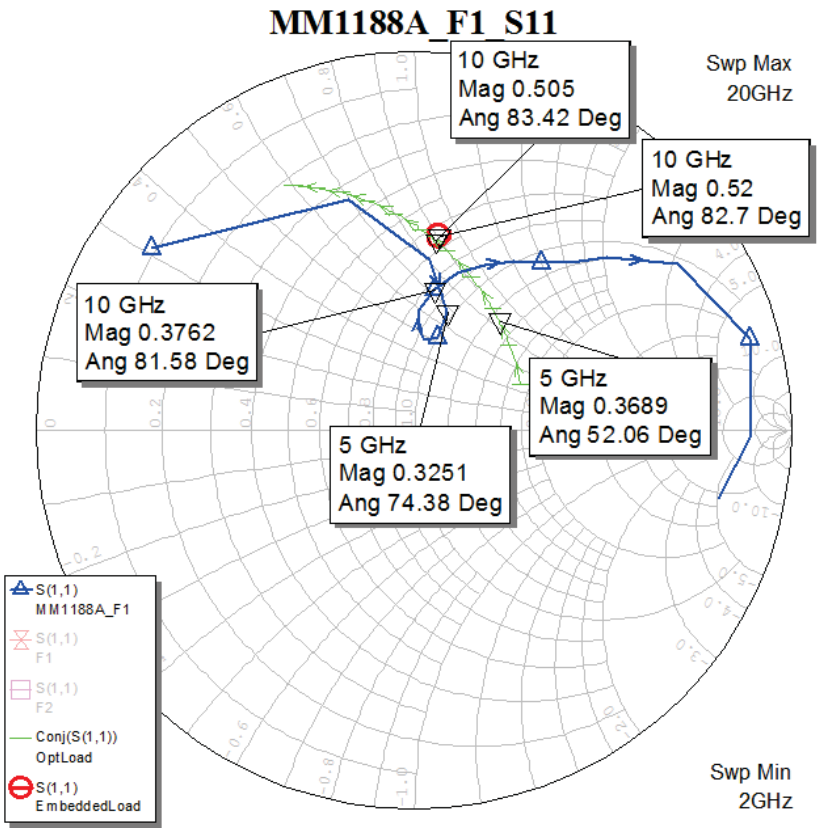


Figure 6-12, Optimum PAE load impedances extrapolated over a wider frequency range, 2-20GHz, (green) and the predicted (linear analysis) impedance presented by the matching circuit (blue).

Input matching and stabilisation circuitry were designed to produce an acceptable compressed stable gain over the frequency band. The complete test cell is shown in Figure 6-13. This comprises of two output matching solutions to the design problem, C1 and F1, their complete circuits, L1 and M1, with input matching networks, 6x100 and 10x75 devices on their own, 10x75 with pre-matching (to reduce the drive level required for 14-18GHz testing) and de-embedding line. This typifies the practice of trying to get as many designs as possible into the available wafer cell area, as the opportunity to process circuits is usually limited. It is usual to try and anticipate problems hence including the design components so that they can be independently verified. The input matching was based upon the measured large signal input gamma. The input and output matching circuits were included as separate entities so that they could be measured and compared with the PDK simulations as at this stage in the project we had no measure of the veracity of even the passive PDK models. The 10x75 devices were included for future research. Due to time constraints an E-M simulation

of the matching circuits was not carried out before manufacture[2]. At this point in the research the DLUT model had not been implemented in the design software and so a nonlinear simulation was not possible.

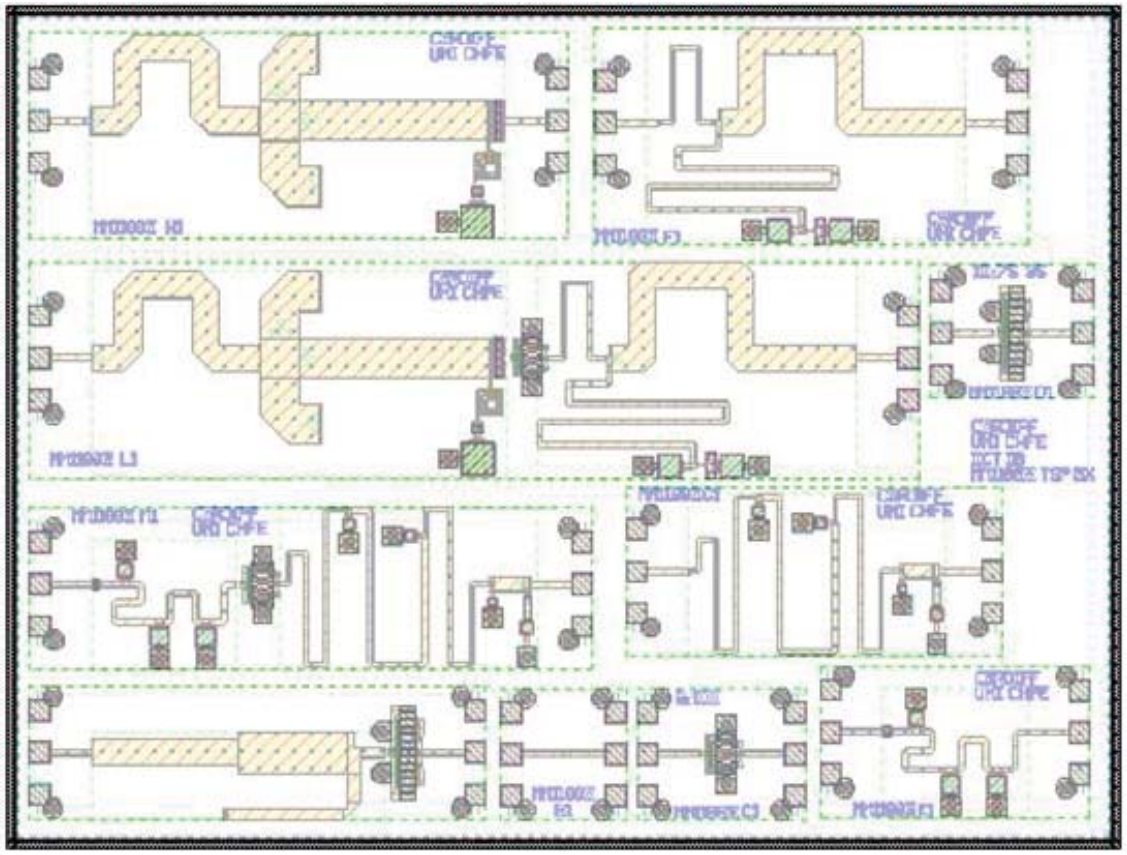


Figure 6-13, Layout of test cells for manufacture. Two different matching approaches where tried (L1 and M1). Also included are 6x100 device cell, de-embedding line, matching circuits on their own, and another device (10x75_35) also under evaluation.

The manufactured circuits, Figure 6-14, were initially tested small signal on both the Automated Load Pull System (ALPS) and a PNA-X vector network analyser. The small signal gain response is shown in Figure 6-15 and Figure 6-16. These results show reasonable agreement over the pass band between simulation and measurement, bearing in mind the designs were based on large signal not small signal data, although there appears to be some frequency shifting. Of particular note is the high level of consistency between the ALPS (frequency range 1-40 GHz) and the PNA-X (frequency range 0.01-26.5 GHz); this was the first wide band comparison between active circuits on the two systems. Although the small signal response of circuit M1 appeared to closely match the simulated, an examination of

[2] A problem experienced throughout the period of the research was that space on wafer runs for experimental circuits was unpredictable and thus opportunities had to be grasped when available.

the load presented by the output matching circuit revealed that the load impedance was actually some way off as shown in Figure 6-17. Circuit "Wafer1_G10" was cracked hence it's dramatically different response. The other circuits are consistent but not in line with the simulated performance. Examination of the manufactured structure showed this to be due to a missing metal bridge to the first shunt capacitor. A comparison was made of the PAE and output power performance between the device on wafer loaded with the matching circuit F1 impedances and the amplifier stage itself, Table 6-4. A reminder here, that the measurements on a device on its own do not include the loss of the output matching circuit and hence are optimistic. Similarly the drive power level on the device itself can be accurately measured; however in the MMIC circuit the losses of the input matching circuit (which vary significantly across the frequency band) make an accurate comparison difficult. As a result of the implications of this point for the next design the input matching circuit is omitted. The ability of the measurement system as a diagnostic tool now comes in to play; to analyse and help understand what is actually causing the differences in performance between the predicted and the measured.

DpHEMT 6x100 at 9v 120mA				MMIC L1 Stage in 50Ω	
Frequency (GHz)	Γ (mag/ang)	PAE (%)	Pout (dBm)	PAE (%)	Pout (dBm)
5	0.46/99.0°	30.8	25.5	25.1	25.3
7.5	0.25/98.1°	35.7	26.1	38.3	26.9
10	0.28/111.8°	30.7	25.8	33.0	26.8

Table 6-4, Measured performance of device with manufactured circuit load impedances and performance of MMIC circuit L1 in a 50Ω

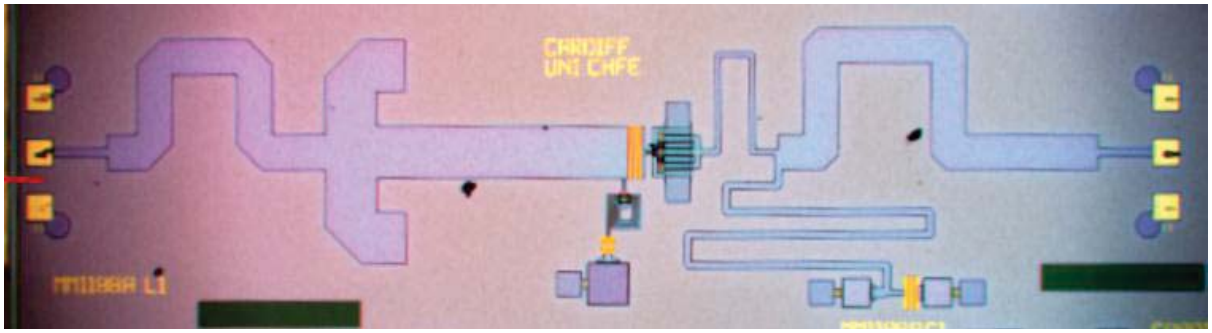


Figure 6-14, Manufactured MMIC L1

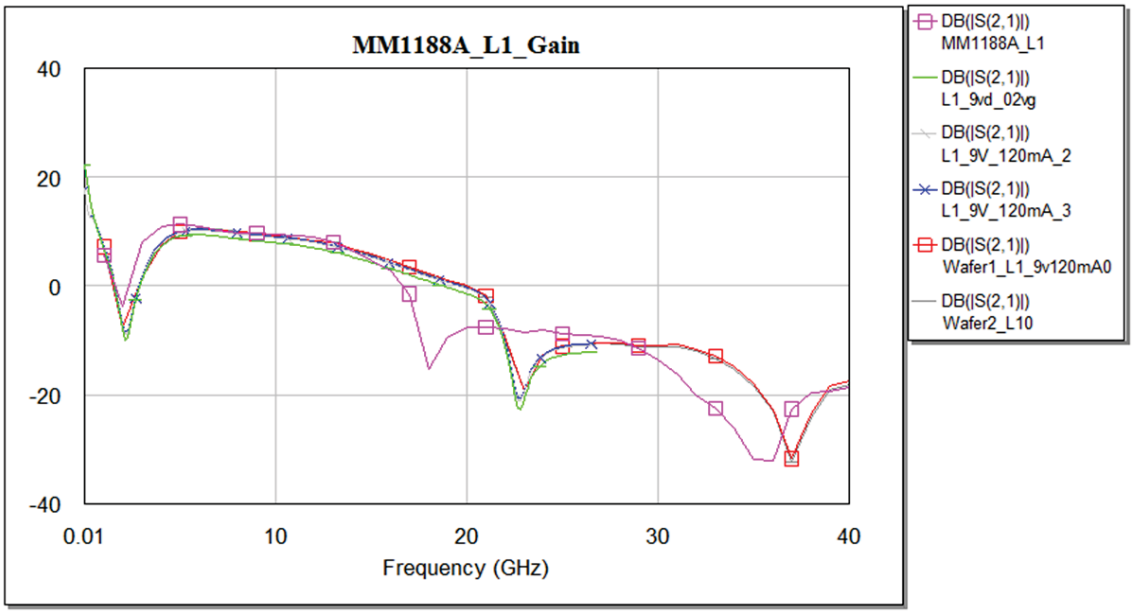


Figure 6-15, Small signal performance of MMIC stage L1, simulated (Pink), PNA-X (26GHz) ALPS (40GHz)

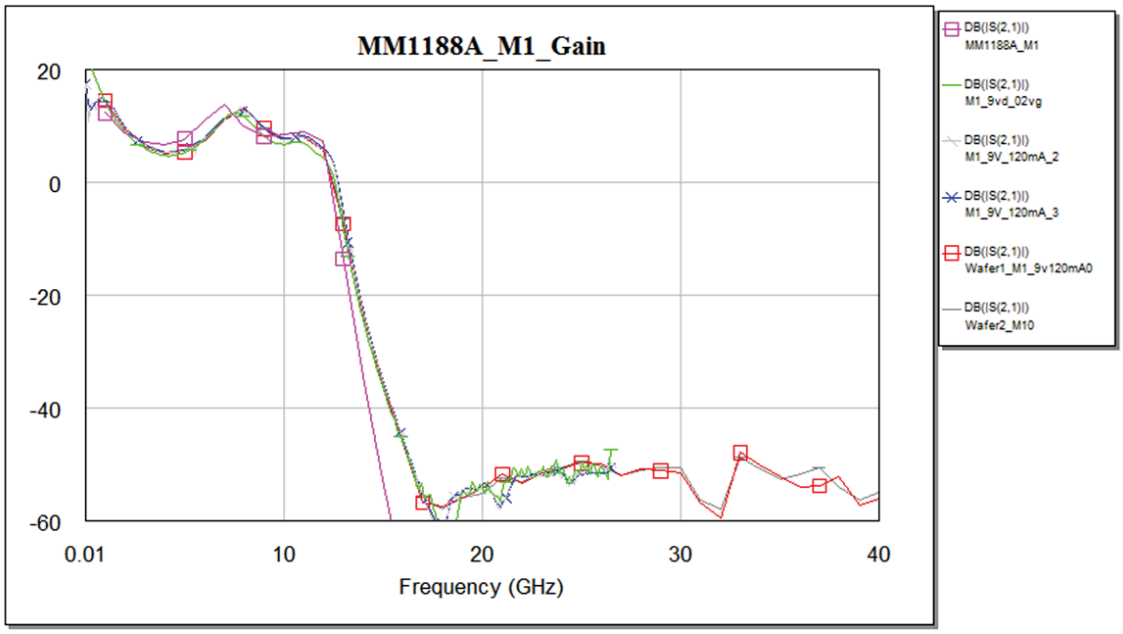


Figure 6-16, Small signal performance of MMIC design M1, simulated (pink), PNA-X (26GHz), ALPS (40GHz)

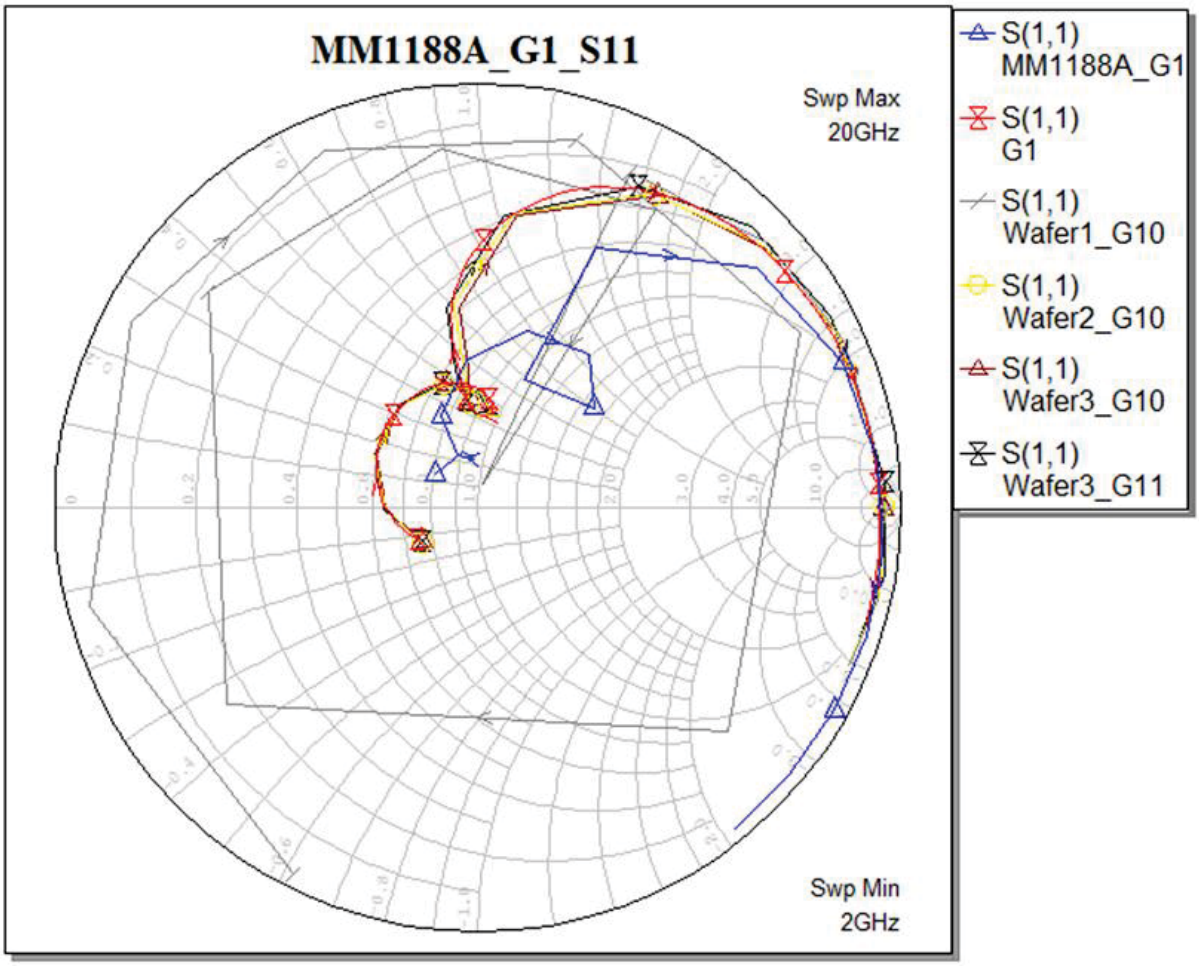


Figure 6-17, Output matching network for M1, wafer 1 circuit damaged, simulation shown in blue.

The performance of the load presented by the output match of design L1 followed the predicted performance much more closely, Figure 6-18, although with a slight frequency shift and less dispersion at the top end of the frequency band. Note that these load impedance measurements include the feed lines and therefore require de-embedding for comparison with the desired loads; but as the simulation includes the feed lines as well, comparisons between simulated and measured are valid. No further work was carried out on MMIC M1 (containing G1) due to the capacitor connection error. To examing MMIC L1 in detail we start by looking at the load impedances presented by the output matching circuit and circuit losses.

A detailed plot showing the differences between measured and modelled de-embedded load impedances is shown in Figure 6-19. It can clearly be seen that the centre frequency, 7.5GHz is very close to the predicted however the impedances at 5 and 10GHz are considerably further away from the target loads. To examine the impact of these changes and to include effect of circuit losses, Figure 6-20, we return to the original

measurement data and assess the performance with the designed and manufactured load impedances.

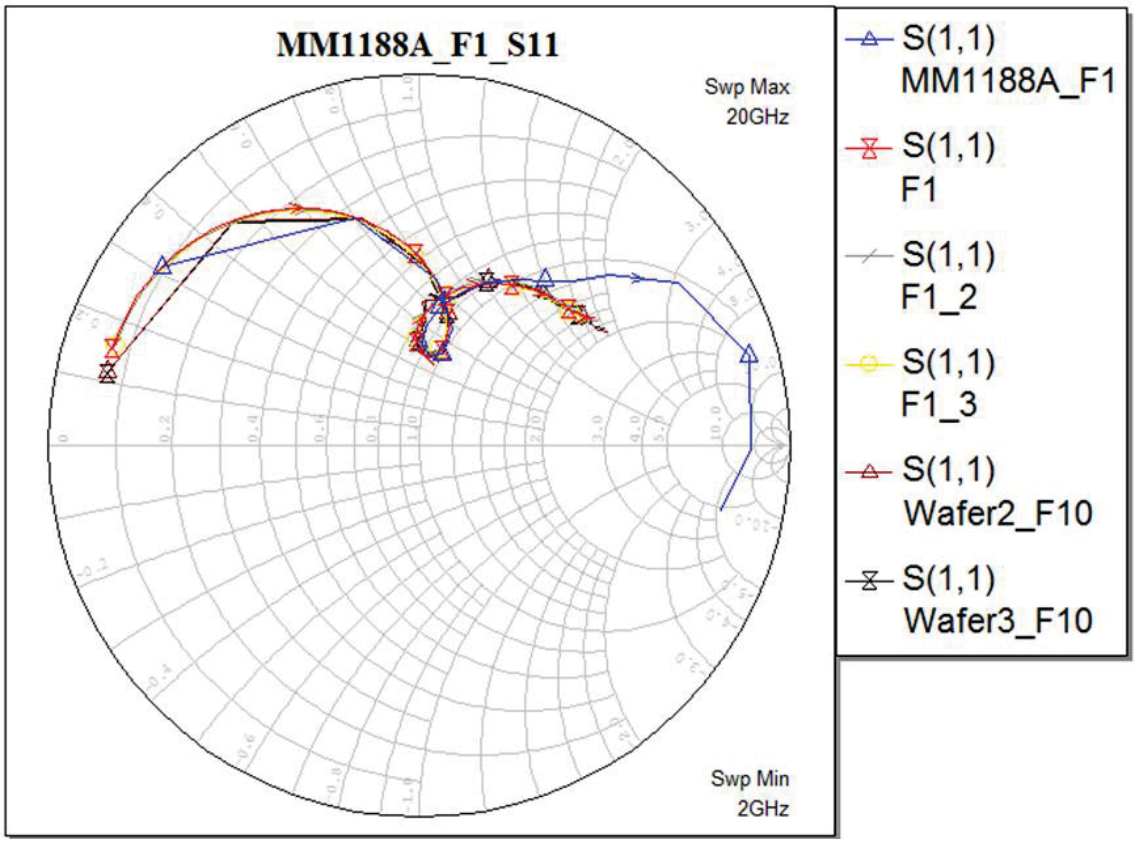


Figure 6-18, Output matching circuit for design L1, simulation in blue.

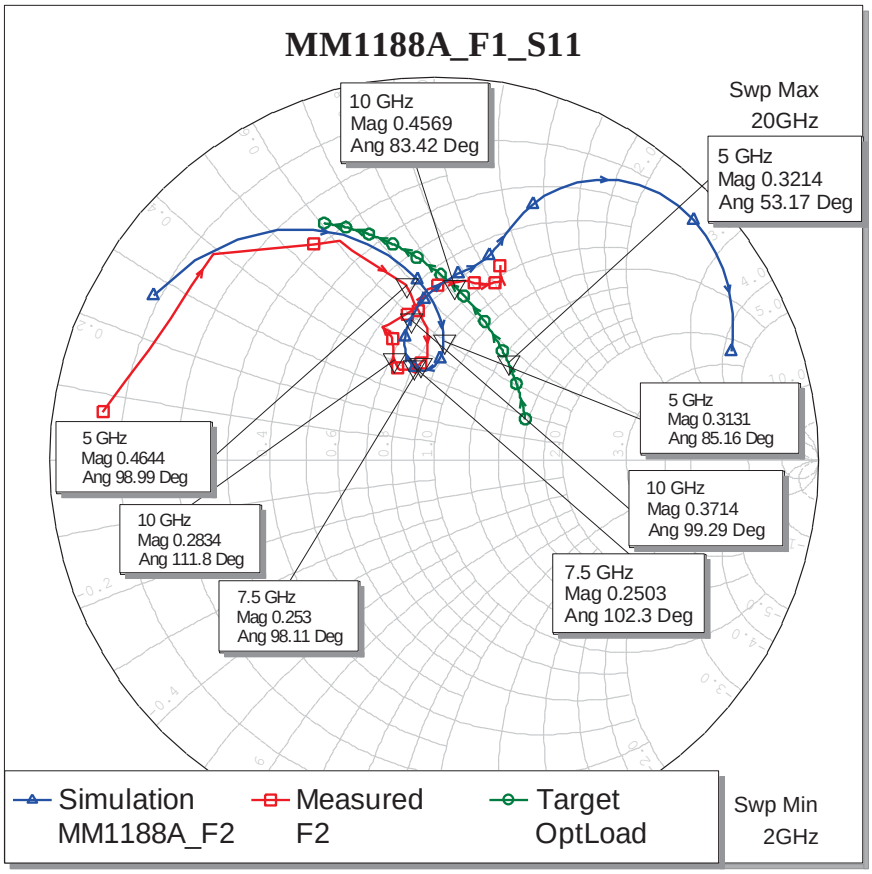


Figure 6-19, De-embedded measured and simulated load impedances and the optimum PAE impedance trajectory (green).

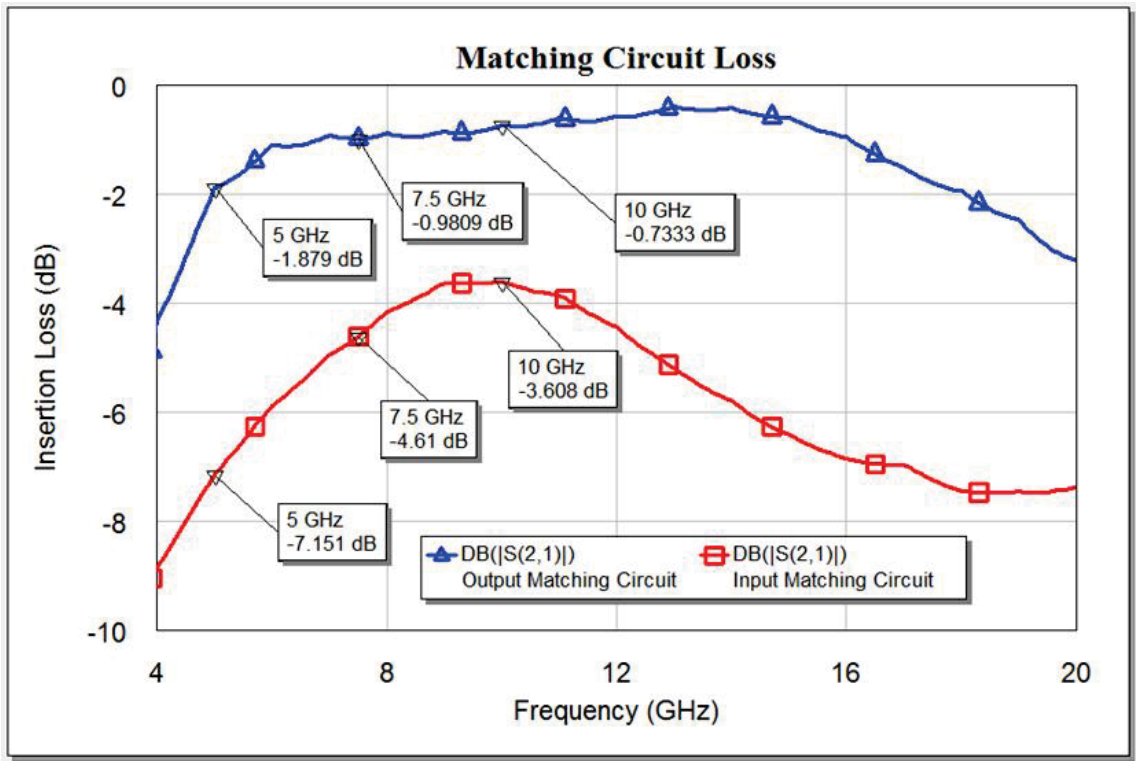


Figure 6-20, Measured matching circuit losses when terminated with the device impedances.

A further factor to take into account when looking at performance deviation from expected is the behaviour of the devices themselves. There are not only differences in behaviour across wafers but often more significant differences between wafer runs. The extent of the variations depends upon the tightness of the process control in the foundry and the quality of the materials used. Foundries will vary between those that offer a very few processes and produce high volumes in each (for example for mobile communications components) to those that offer more processes and tend to operate on a batch basis. A consequence of this latter mode is that there tends to be increased variability between wafer runs. An individual 6x100 device was included in the test cell so that it could be measured and compared with the original device the design was based on. Although there is no guarantee that this will have exactly the same performance as the one within the matching circuits themselves, it is likely to be closer than the one from the earlier wafer run. Conducting an automatic load pull measurement on this device revealed an 11.9° shift in the optimum PAE load impedance at 10GHz. This result is plotted in Figure 6-21, which also shows the optimum PAE impedance of the previous device measurement and the impedance actually presented by the manufactured circuit load. The actual optimum PAE values of the devices are within 0.6% although the power out at the optimums are different by 0.5dB. A comparison of the optimum values of the key parameters is summarised in Table 6-5. It is worth noting that the devices were biased at exactly the same drain current. This therefore takes no account of differences in transconductance between the devices. When attempting to match device performance a common approach is to measure the I_{DSS} and adjust the I_D to 50-60% of that value. However in typical EW applications it is more normal to have the devices run at a fixed drain current; relying on a good repeatability between transistors. Loading the device on the same wafer as the MMIC with the output circuit load the PAE is 30.7% and output power 25.8dBm. This is lower than measured from the MMIC itself (33.0% and 26.8dBm), which considering the matching circuit losses is a surprising result. There are two further factors that may account for the difference, one is the impact of harmonic terminations (which will be examined in detail in the next design) and the other is the drive levels which directly effects the compression of the device. This latter point is complicated by the input matching circuit. In the original device measurements the power incident upon the device is accurately known. In the MMIC circuit there is an input matching circuit which attempts not only to improve the input match to

the device, but also to stabilise the circuit, which is achieved in part with a lossy resistive series element. In the solution used the matching networks have insertion losses as shown in Figure 6-20. A problem with such matching circuits when used with device performance evaluation is that the transistor input and output impedances are drive level dependant and thus the matching circuits are only optimum at a single input power level. In practical amplifier design this can be utilised to, for example modify the compression curve: a device where the input match is tailored to the highest drive power will have a less steep compression curve, caeteris paribus. Thus with a varying and high Γ_{IN}, and a relatively high insertion loss between the device and the calibration plane, the measurement of the actual input power to the device becomes subject to a wider tolerance.

DpHEMT 6x100 at 9v 120mA: **Original Device**, **2nd batch**				
Frequency (GHz)	Γ_{LOAD} (mag/ang)	Opt. PAE (%)	Assoc. Pout (dBm)	Assoc. Gain (dB)
5	**0.35/50.3°** **0.40/39.4°**	**50.6** **52.0**	**27.1** **26.7**	**14.8** **14.9**
7.5	**0.41/65.2°** **0.45/45.7°**	**49.8** **49.4**	**26.9** **26.4**	**12.3** **12.7**
10	**0.50/82.7°** **0.50/70.8°**	**46.1** **45.5**	**26.5** **26.0**	**10.4** **10.4**
Frequency (GHz)	Γ_{LOAD} (mag/ang)	Assoc. PAE (%)	Opt. Pout (dBm)	Assoc. Gain (dB)
5	**0.22/57.1°** **0.30/37.6°**	**49.4** **50.9**	**27.5** **27.1**	**15.3** **15.0**
7.5	**0.22/52.9°**[3] **0.24/45.5°**	**45.7** **45.7**	**27.3** **27.0**	**12.3** **12.7**
10	**0.21/57.1°** **0.31/56.8°**	**40.2** **41.4**	**27.2** **26.8**	**10.1** **9.2**

Table 6-5, Measured performance at Optimum PAE load impedances.

[3] The reflection coefficient of the load is very similar to that at 5GHz and one would expect the angle to be larger, however the maximums are calculated based on the nearest load point and thus there is an overlapping area of the impedance plane for the optimum values of output power at the two frequencies. As the resistance of the output load is closer to 50Ω there is less spread in the output load impedance over the range 5 – 10GHz.

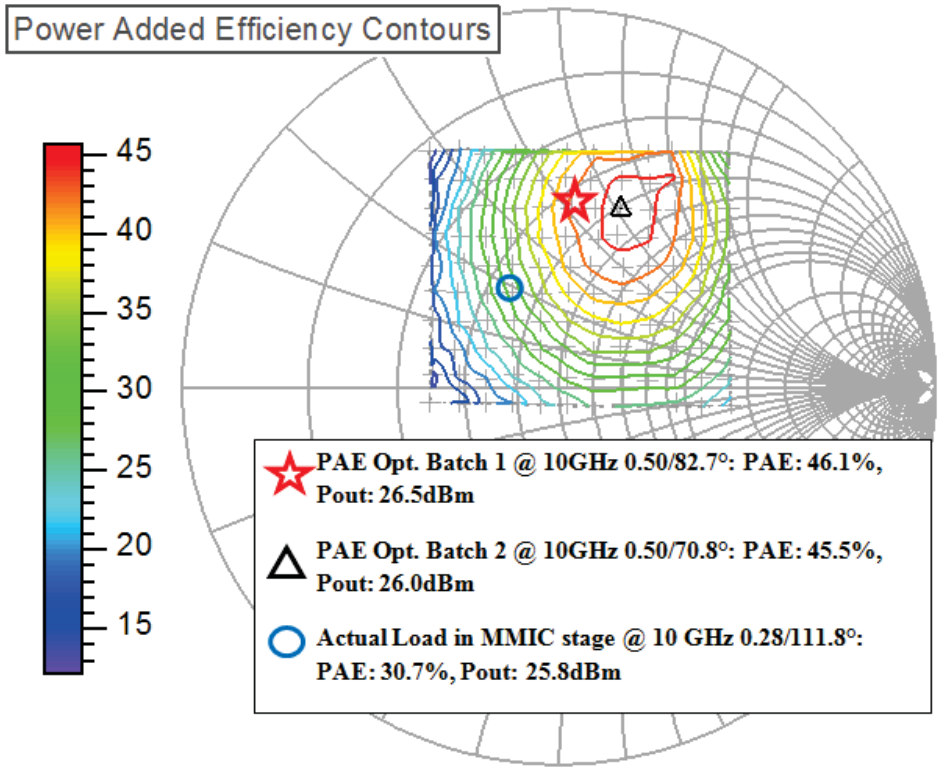


Figure 6-21, PAE Load pull contours at 10GHz for the device on the wafer with the MMIC, with the optimum load from the original device and the measured matching circuit load.

It thus now becomes much more difficult to make a direct comparison between the measured results and the simulation data, consider also that there may be some difference in the input reflection coefficient as well as the gain and output of the device. At the time of the original design the DLUT model was not operational within MWO (software in which the PDK operates), and thus it was necessary to make do with a small signal analysis using measured S parameters. A measurement of the MMIC performance small signal, Figure 6-22, made on the VNA gives only limited information regarding the device performance. Using the newly measured data (from the same wafer as the MMIC), a model was created and the circuit analysed retrospectively. Figure 6-23 shows the simulation schematic for the analysis of the MMIC performance. The passive input and output matching circuits are loaded as S parameter data blocks. Bias "Tees" are incorporated so that the model can select the appropriate data set if multiple bias measurements have been made.

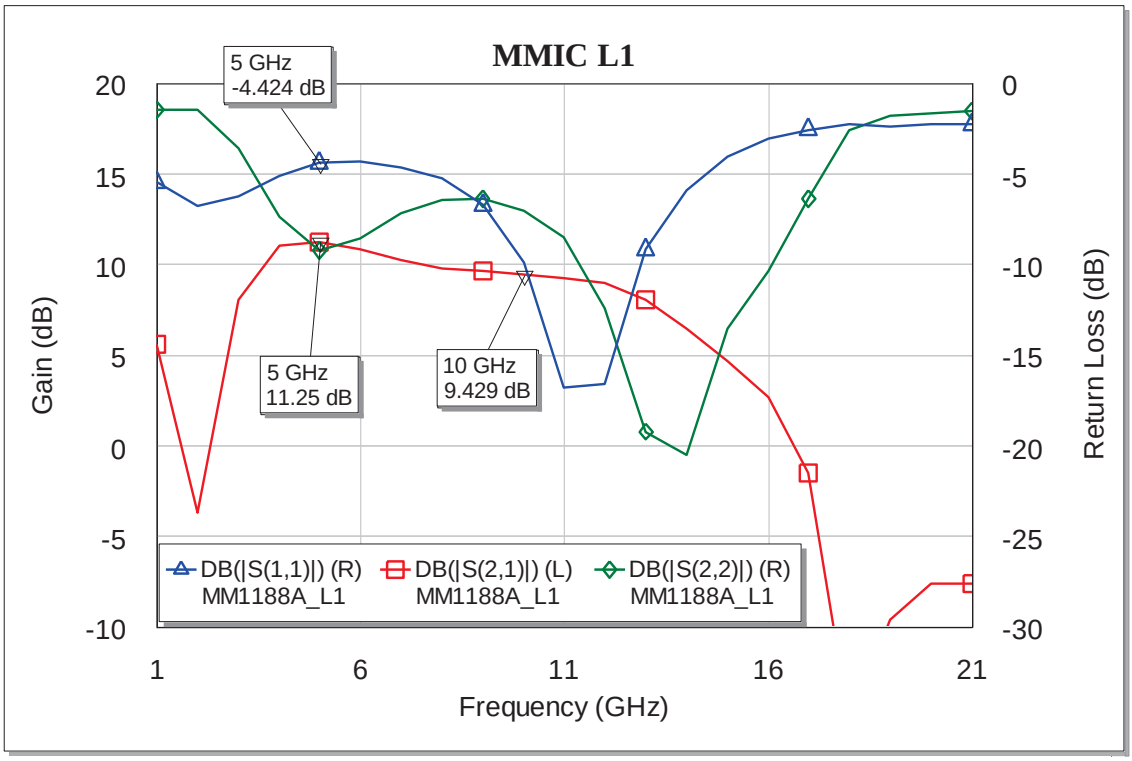


Figure 6-22, Small signal Gain and Match of MMIC L1

The first step was to compare the device data from the measurement system with the results from the nonlinear simulation. Figure 6-24 and Figure 6-25 are contour plots from the measurements of the 6x100 device at 5GHz. The load presented by the output matching circuits is indicated by the black triangle (nearest measured grid point) and with this impedance the PAE is 30.6% and the output power 25.6dBm.

The results of a fundamental load pull simulation are shown in Figure 6-26. Note that

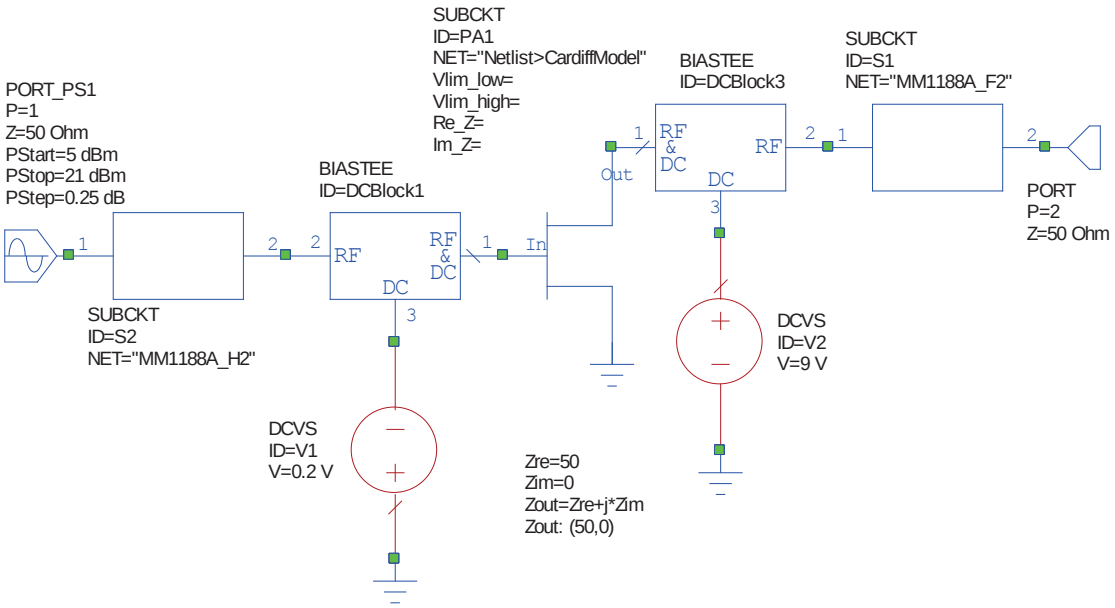


Figure 6-23, Nonlinear simulation schematic for MM1188A in MWO using DLUT model based on original design measurements.

the measurement grid points are arbitrarily set within a nominal area and not the same as those used in the original data acquisition, this demonstrates the ability of the model to interpolate satisfactorily between measured data points. Both the shape of the curves and the maximum levels show good consistency with the measured data. Also shown on the graph is the load applied to the DLUT model for the assessment of the device on its own, which is that presented by the output matching circuit in the MMIC. An advantage of the DLUT model over the measured data is that there is an algorithm in the simulation software which predicts where the optimum load impedance will be, whereas in the measurement software it looks for the grid point with the best performance, which can lead to some confusion as discussed in Table 6-5.

Knowing that we have a good replication of the device in the DLUT model we can now assess the impact of the matching circuits. Figure 6-27 shows three power sweeps. The red trace is the simulation of the DLUT model on its own, terminated in the load impedance of 0.46/_99°, i.e. that presented by the measured matching circuit. Note that the red output power trace shows an atypical increase at the top of the range. This is because the drive power level in the simulation has at this point exceeded that of the measurement data and the model is extrapolating – showing the inaccuracy resultant from going outside the measurement envelope. The blue trace shows the performance of the DLUT model in between the measured input and output matching circuits; in this case the marked change in slope above marker 2 is at a higher power level due to the circuit losses of the input matching. The gain has decreased, the curve shifted to the right, as would be expected due to the input and output matching circuit losses. This is by just under 4dB in the linear part of the curve, and to nearly 6dB at the top end as the input reflection coefficient changes with drive power, as seen in the blue curve of Figure 6-28. In the case of the actual measured data the maximum output power has fallen by 0.8dB which is less than we would expect based on the insertion loss of the output matching circuit (~1.5dB at 5GHz – note Figure 6-20 is measured in 50Ω and hence needs to be terminated in the output impedance of the device for a correct measurement of insertion loss). The simulation (green trace) and the measured power sweep for the MMIC are close, particularly at the lower power levels; however the actual device performs about 0.6dB better than the prediction. Again this is postulated to be due to the actual harmonic terminations as will be investigated further in the next section. In the simulation of the DLUT model on its own the harmonics are

terminated in 50Ω as this complies with the measurement conditions used to create the model. The DLUT model was created with data from specific harmonic device impedances, the difference between these and the actual impedances of the output matching circuit are not accounted for in this simulation.

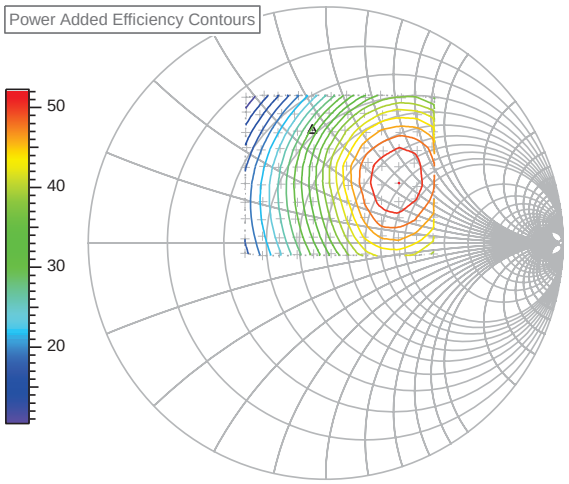


Figure 6-24, Measured PAE contours at 5GHz, original device.

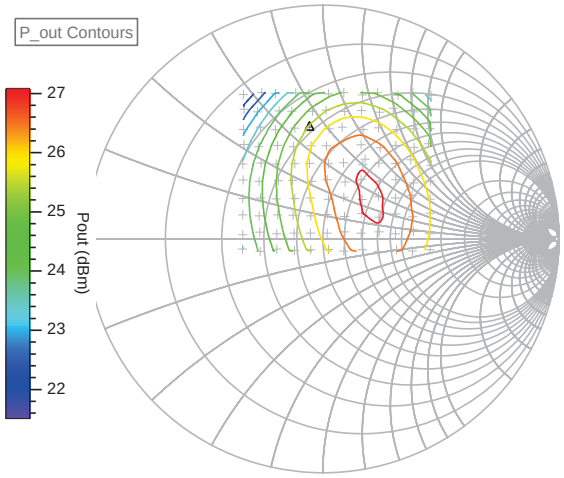


Figure 6-25, Measured Output Power contours at 5GHz, original device.

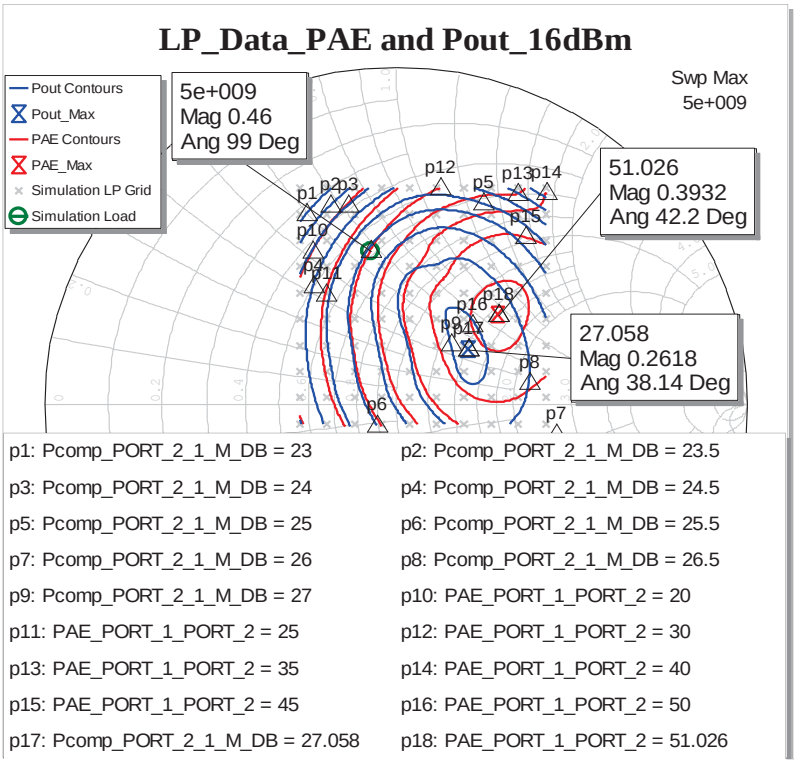


Figure 6-26, Simulated PAE and Output Power contours from DLUT model of original device and load applied to device.

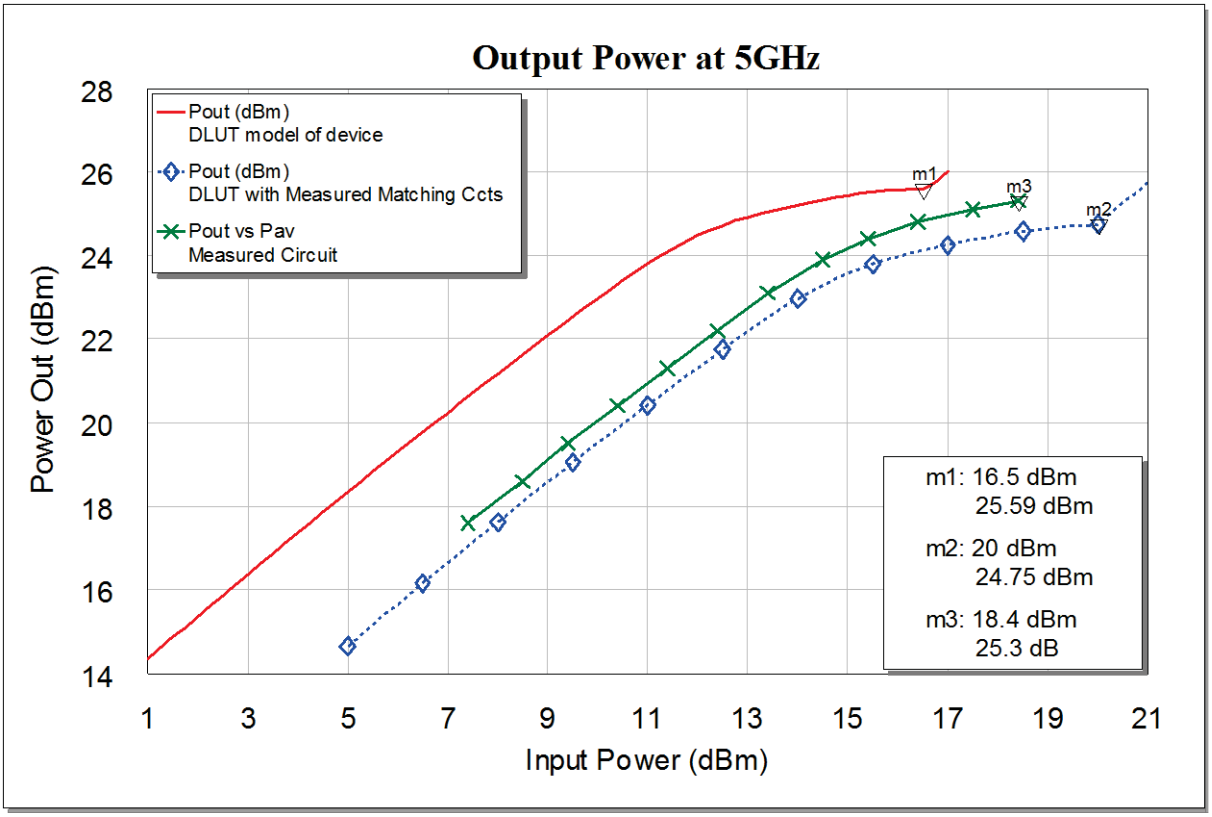


Figure 6-27, NL Simulation of device, device with matching and measured MMIC – Output Power.

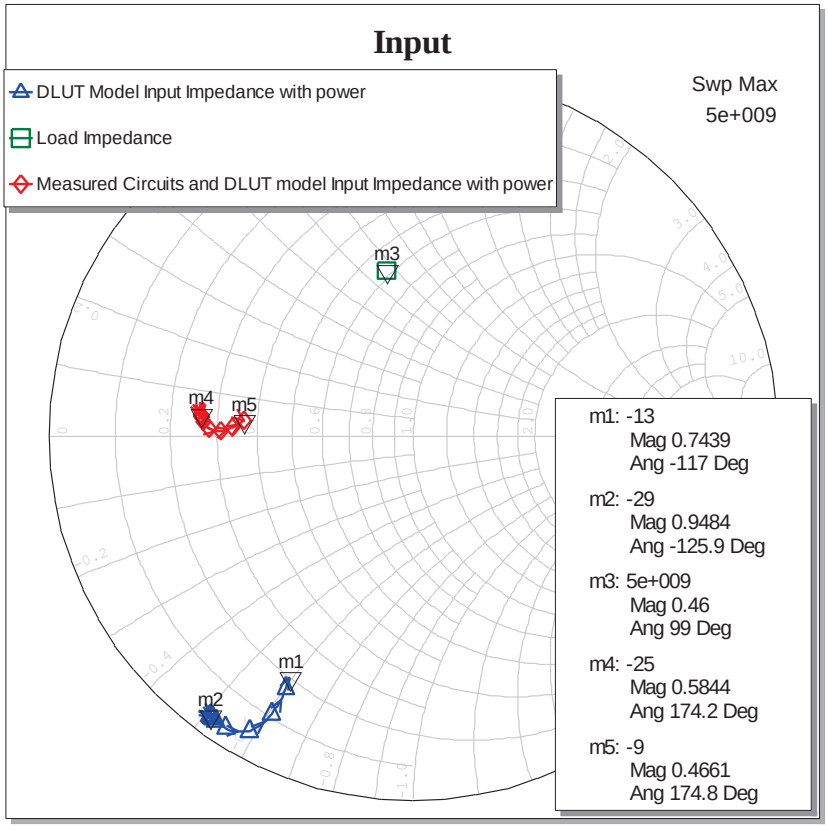


Figure 6-28, Variation of input impedance with drive level.

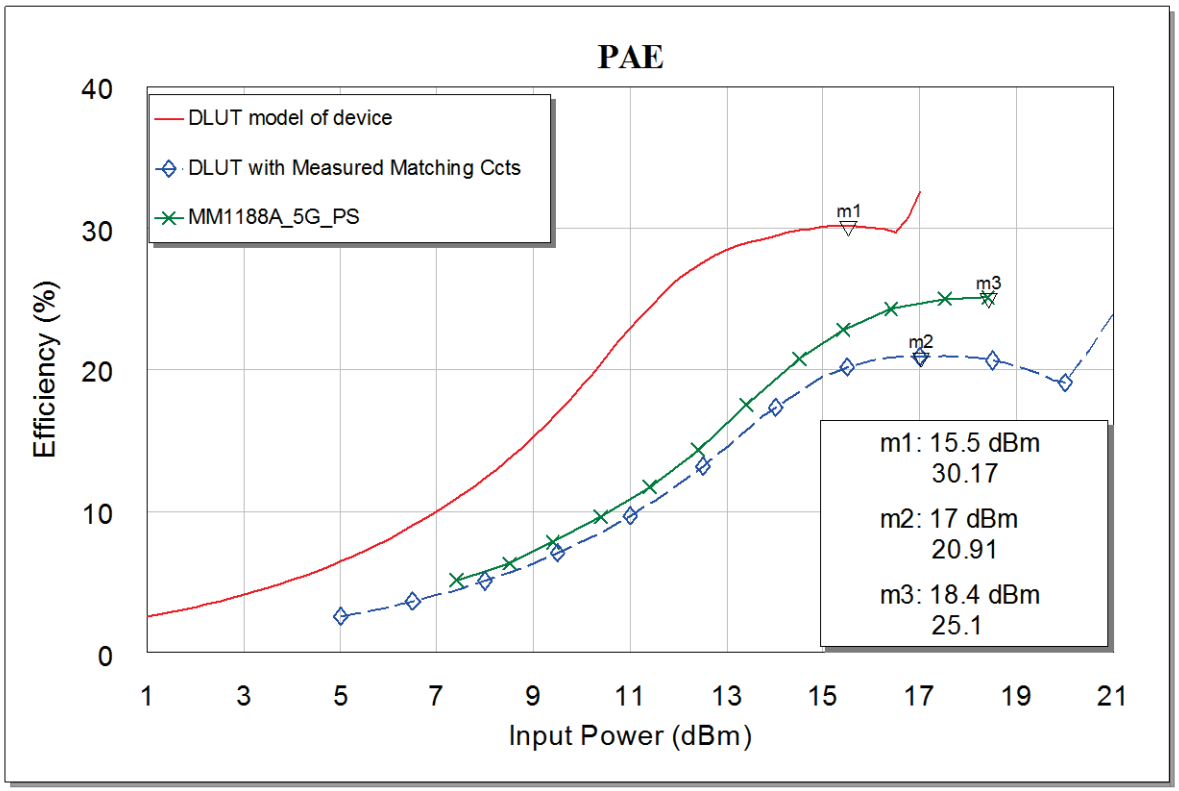


Figure 6-29, Nonlinear simulation of PAE; DLUT model with measured loads (red), DLUT model with measured output matching circuit (blue) and measured MMIC data.

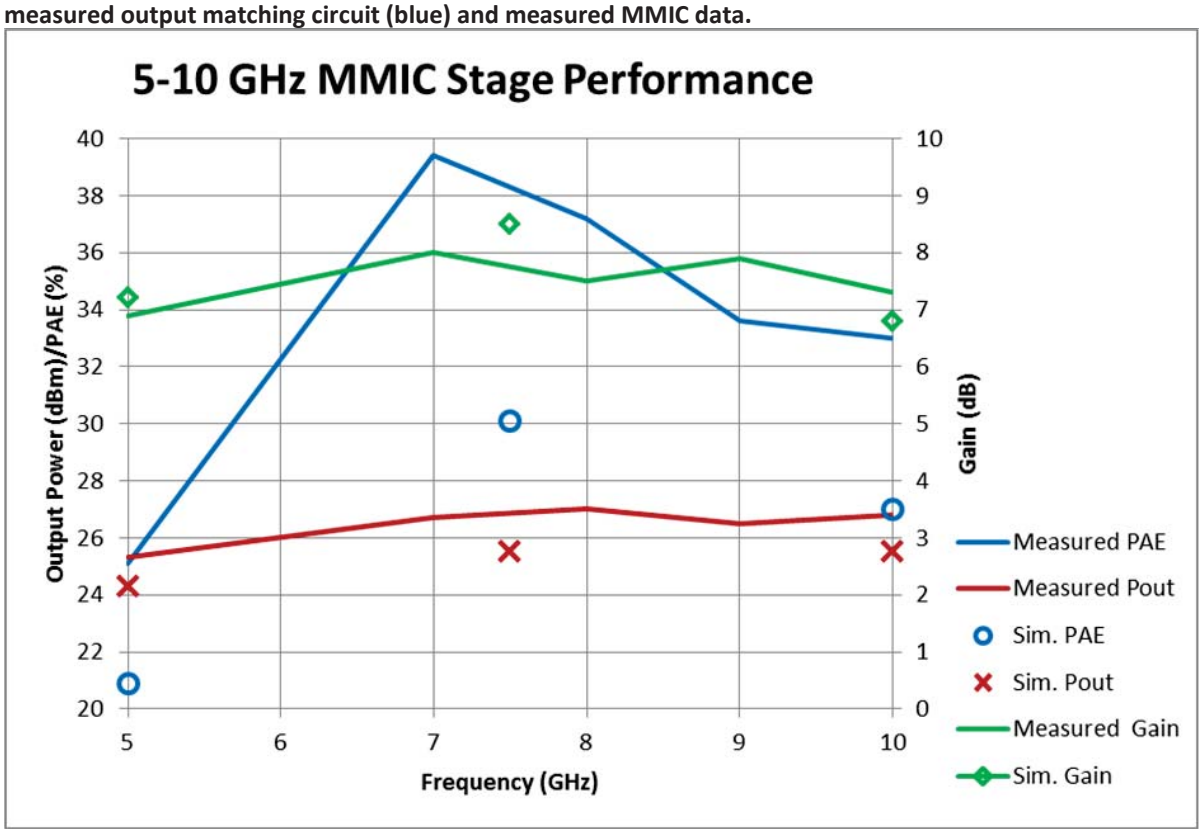


Figure 6-30, Summary of simulated data (points at 5, 7.5 and 10GHz) and measured performance of manufactured MMIC (traces).

Looking at the PAE performance again the manufactured MMIC is better. The peak PAE is 25.1 as opposed to 20.9%, Figure 6-29. The MMIC was measured at 5, 6, 7, 8, 9 and

10GHz at the maximum PAE drive level. The PAE and associated output power and gain are recorded along with the simulated performance from the measured matching circuits and the DLUT models from the MMIC wafer devices at the original measurement frequencies of 5, 7.5 and 10GHz, Figure 6-30. Cleary the MMIC has performed better than we would expect from the fundamental frequency analysis. An appropriate course of action at this stage would be to investigate the hypothesis that this increased performance is as a result of the harmonic load impedances. However time pressure from the objectives of the funding contract made it imperative to implement a harmonically tuned MMIC design and the results of this research and other work undertaken by the industrial partner suggested that a larger periphery device was required to meet the 27dBm (0.5W) output power requirement. Thus an analysis of the impact of the harmonic terminations was not made on the 6x100 DpHEMT but instead investigations were carried out on the 10x75 device and it is assumed that the behaviour would be similar.

A reasonable explanation for the behaviour of the device was established based on the load-pull and matching circuit measurements. The nonlinear model was successfully implemented within MWO, demonstrating excellent agreement with the measured device data.

6.4 Design of 0.5W Harmonically Enhanced MMIC stage

The results of the previous design had given confidence in the measurement system and experience of both the DpHEMT performance and the foundry PDK. A strategy was therefore proposed for designing a gain stage with increased efficiency as a result of the harmonic terminating impedance; an issue with ≥octave bandwidths is that for the lowest frequency the 2^{nd} harmonic will fall in band and thus the impedance will be determined by that required for the fundamental at this frequency. The problem is displayed graphically in Figure 6-31; in the case of an above octave amplifier the 2^{nd} harmonic of the lowest in band frequency, $2F_L$, falls within the operating bandwidth. Therefore either higher order terminations must be utilised, or a higher in band frequency (such that $2F_X > F_H$) selected. It was therefore decided that an octave band amplifier, 5-10 GHz, would be designed with harmonic enhancement at 6 GHz. This frequency was chosen as not only were both the 2^{nd} and 3^{rd} harmonics within the range of the measurement system load pull driver amplifiers,

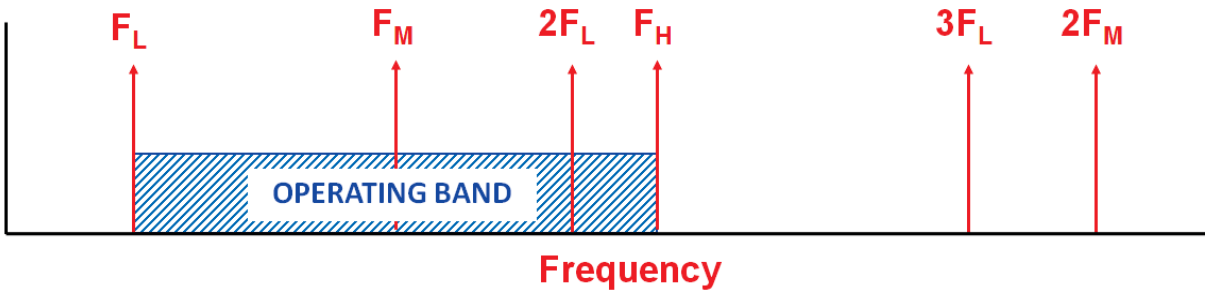


Figure 6-31, Harmonic positions in >octave amplifiers.

but also the accuracy of the PDK models for the matching circuits looked reasonable up to these frequencies, based upon the previous design results. The plan would be to design a fundamental matching circuit that gave close to optimum performance at the top end of the frequency band and then implement harmonic tuning to enhance the performance at the lower part of the band. The objective would be to achieve a minimum of 27 dBm at a PAE of >40% in a class A bias.

In line with the described process the device was initially biased in class A and load pull measurements were conducted across the frequency and input power range to determine the optimum load points at each frequency and the performance that could be achieved. These are summarised in Table 6-6; in order to achieve higher power a larger periphery device was selected, 0.75mm, which therefore required a proportionate increase in bias current for continued class A operation. The measurements showed that the simple RC equivalent circuit model for the output of the device still held broadly true at least up to 18 GHz.

Freq. (GHz)	Γ_L (Mag/Ang)	PAE (%)	Pout (dBm)	Gain (dB)
4	0.26/44.6°	49.2	27.4	21.1
6	0.35/65.9°	51.4	27.7	15.0
8	0.52/80.8°	51.2	27.0	13.5
12	0.54/104.1°	44.5	27.2	10.8
16	0.70/118.0°	36.6	26.2	9.7
18	0.70/118.6°	35.1	26.0	8.1

Table 6-6, Measured optimum PAE loads and associated performance for 10x75_35 DpHEMT operated at 9V and 150mA

As mentioned earlier there is still an issue with the calculation of the optimum load within the measurement software, which is typified in Figure 6-32 where it can be seen that the optimum PAE load will lie at the geometric centre of the four grid points (faded grey crosses) on the innermost PAE contour. The measurement software however calculates the optimum based upon the highest PAE grid point. It can also be seen on this graph that one of the load points failed to converge to the desired grid impedance. The data at this point is still valid; it just offsets the regular grid and misses the information at that point. An analysis of the small signal stability circles suggests that this may be due to the load entering the output instability region for the device, causing the convergence failure. A further issue with the measurements is the ability to compare the load pull contours at the same level of compression, e.g. 2dB or 3dB, for each frequency. For convenience a power sweep with 1dB steps is frequently employed and it is only afterwards when the data is analysed that the amount of gain compression is established. It is often convenient to define a step size of 1 dB, which is adequate for the linear region of the compression curve, however above P1dB the slope of the curve changes quickly and it would be advantageous to have a finer power increment at this stage. As the speed of the measurement system increases the practicality of more and finer power steps improves, however as a result there is some deviation from the ideal impedance trajectory of the optimum PAE and Output Power loads as shown in

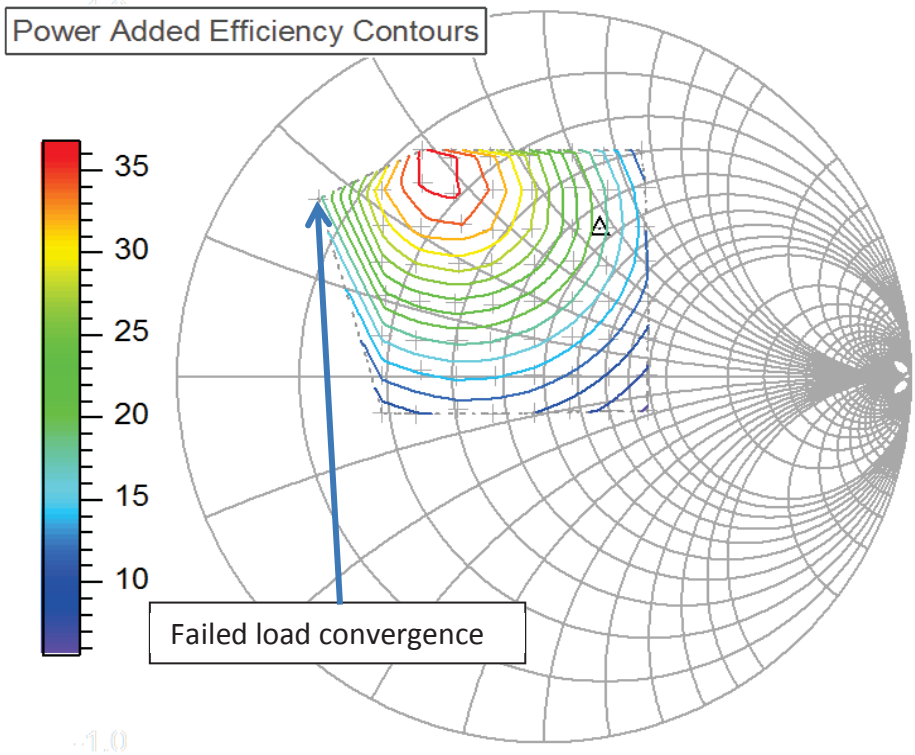


Figure 6-32, 16GHz Load pull contours at the optimum drive level

Figure 6-33, also remember that the optimum impedance points are approximate as they are the highest value measurement grid point. Although the error bars about these measurement points are not insignificant for the reasons outlined the general trend is still clear as is the reasonable approximation offered by the output equivalent circuit also plotted on the admittance chart, the value of the equivalent drain source resistance being chosen to match the appropriate level of compression in the operating environment of the circuit. In summary to improve the accuracy there are two steps that could be taken:

i. Use a finer step size in the drive power in the region of compression.
ii. Either incorporate a better peak level contour algorithm in the measurement software or determine it in the analysis software.

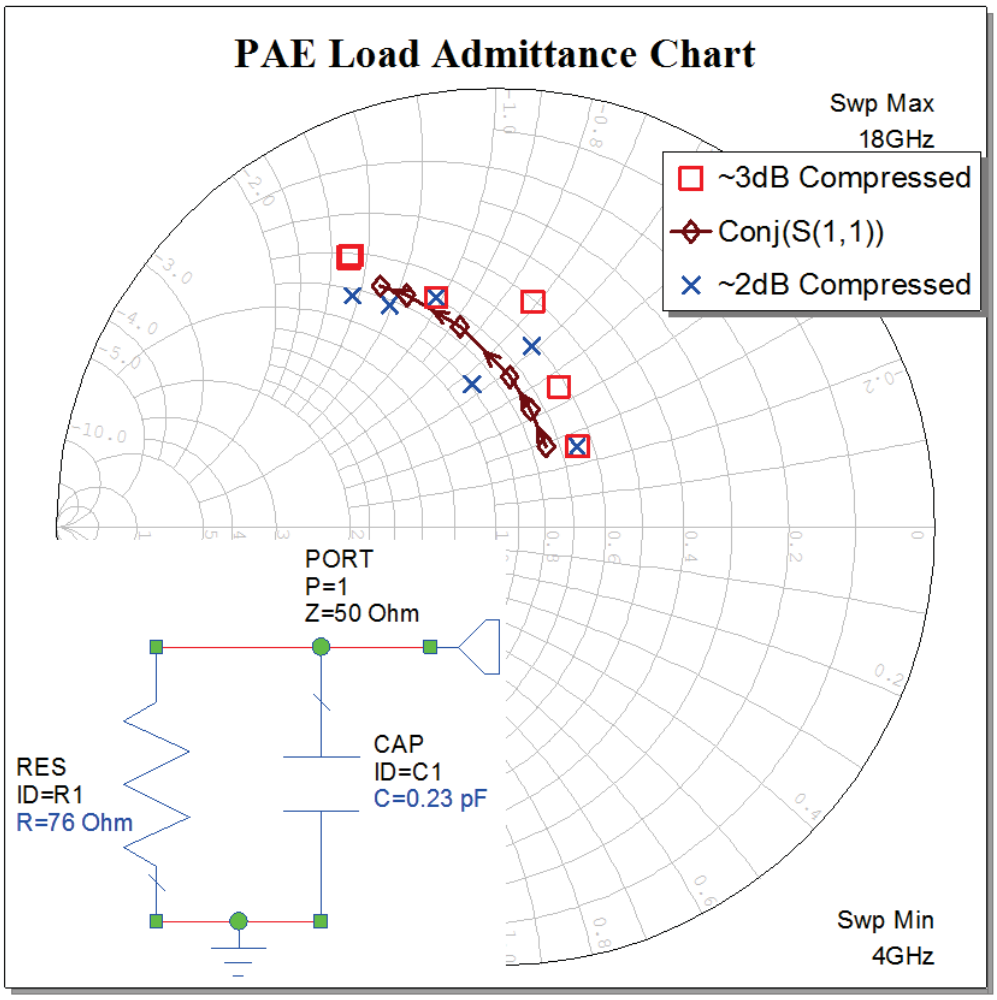


Figure 6-33, Measured PAE loads at 4, 6, 8, 12, 16 and 18GHz and the load equivalent circuit.

It has been established [8] that harmonic terminations can be used to improve the PAE of RF power amplifiers. The theory behind these modes is based upon class B operation of the device. It was suggested that similar conditions existed for class A. Terminating the device with the optimum load impedance at 6GHz a 2nd harmonic load pull across the majority of the real impedance plane was implemented. This showed that the impact of the 2nd harmonic load impedance could be ~±10% on the PAE as shown in Figure 6-34. This graphical representation is useful to the designer as they can visualise the efficiency benefits of 'steering' the out of band impedance. Note that the measurement grid extends beyond the real impedance plane and is equivalent to harmonic injection, a technique that has been explored elsewhere [9]. Examining the power sweep at the grid point where the PAE is highest it is clear that the PAE has not in fact peaked, Figure 6-35. However the positive benefits of the 2nd harmonic termination are equally clear as summarised in Table 6-7. Clearly the 2nd harmonic impedance can improve all of the main measures of device performance and it is noted that the graphical representation of Figure 6-34 suggests that there is further improvement to be had by increasing the magnitude of the 2nd harmonic

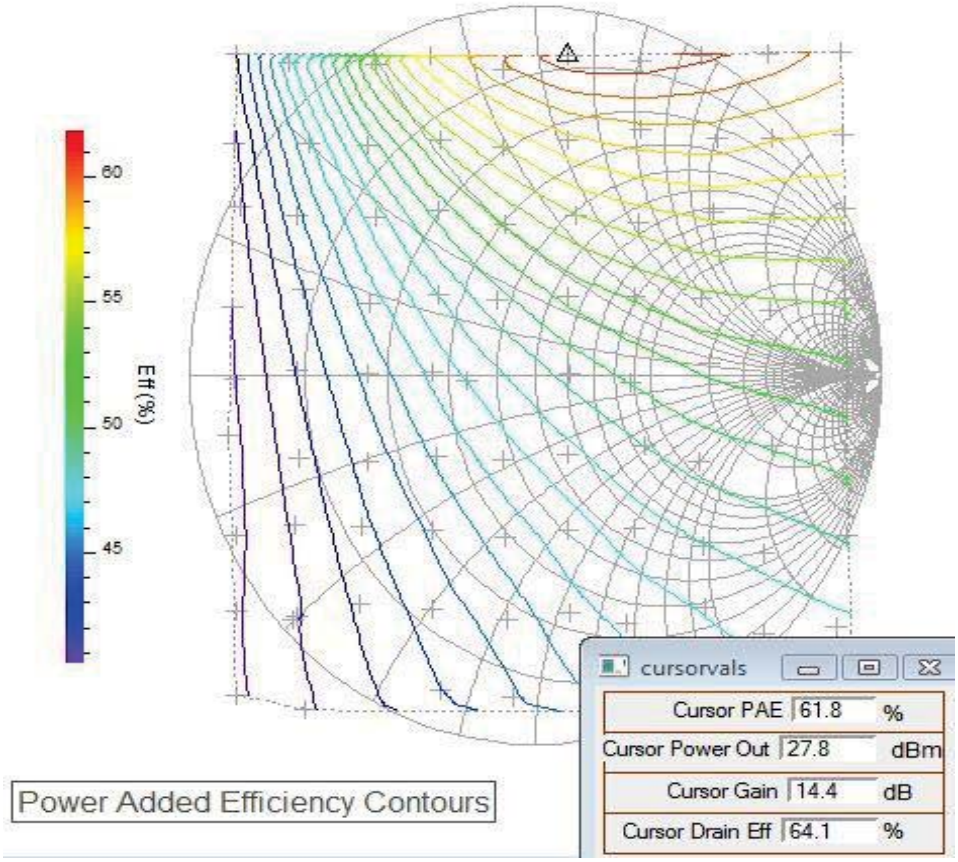


Figure 6-34, 2nd harmonic load impedance contours at 6GHz with the fundamental at optimum PAE load 0.35/_65.6°.

reflection coefficient. Looking at Figure 6-37 we can see the self-limiting action of $\Gamma_{2\ LOAD}$, moving away from the 50Ω point to the perimeter of the Smith Chart in the direction of the optimum, the PAE increases but also the magnitude of 2nd harmonic decreases. Thus in order to get the maximum PAE we have to move outside of the Smith Chart and inject harmonic power, as the device itself under these load conditions does not produce a high enough level to construct a waveform that gives higher efficiency (as will be investigated later).

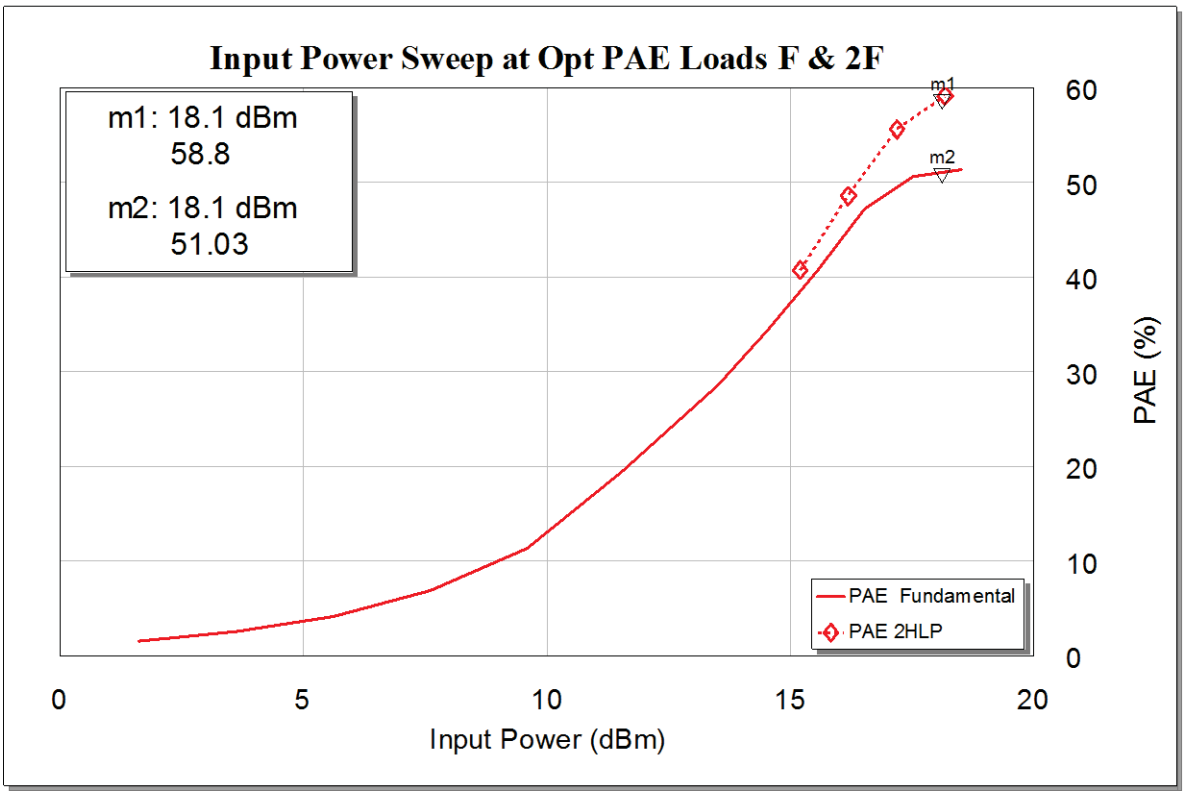


Figure 6-35, PAE with swept input power at the optimum PAE fundamental load with the 2nd harmonic in 50Ω (solid) and then at a load of 0.89/_79° (dashed).

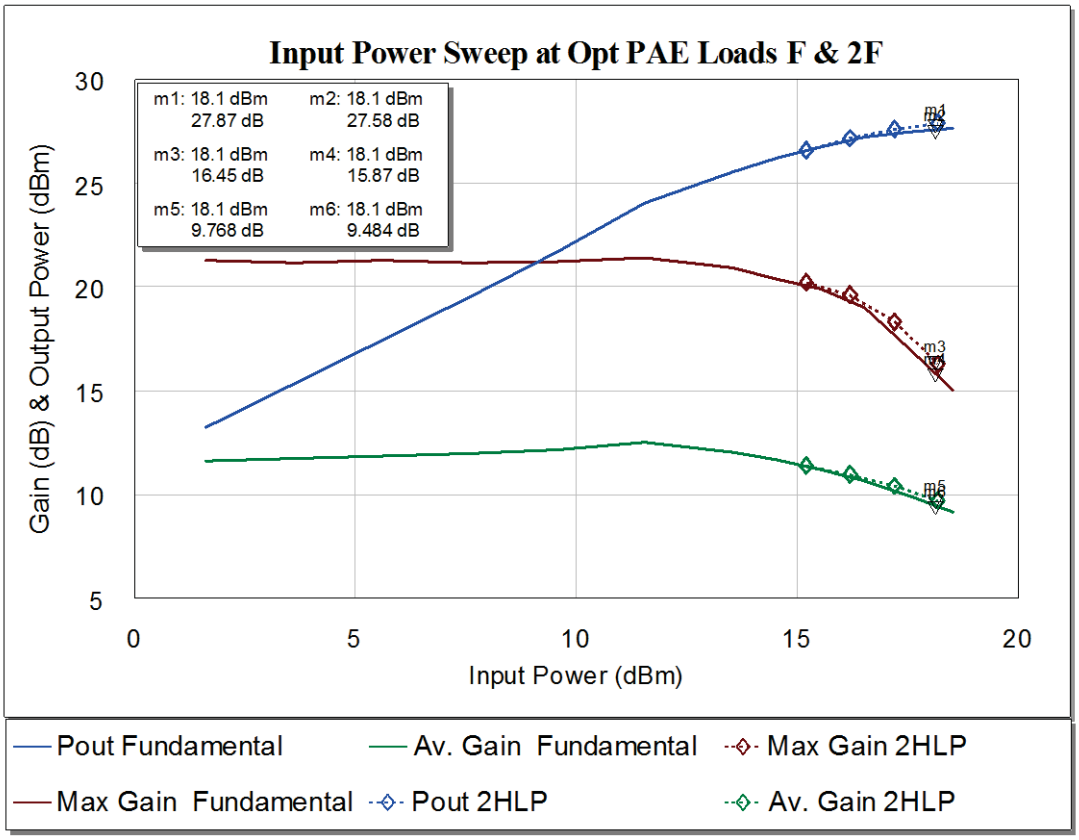


Figure 6-36, Power sweep at the optimum PAE fundamental load with the 2nd harmonic in 50Ω (solid) and then at a load of 0.89/_79° (dashed).

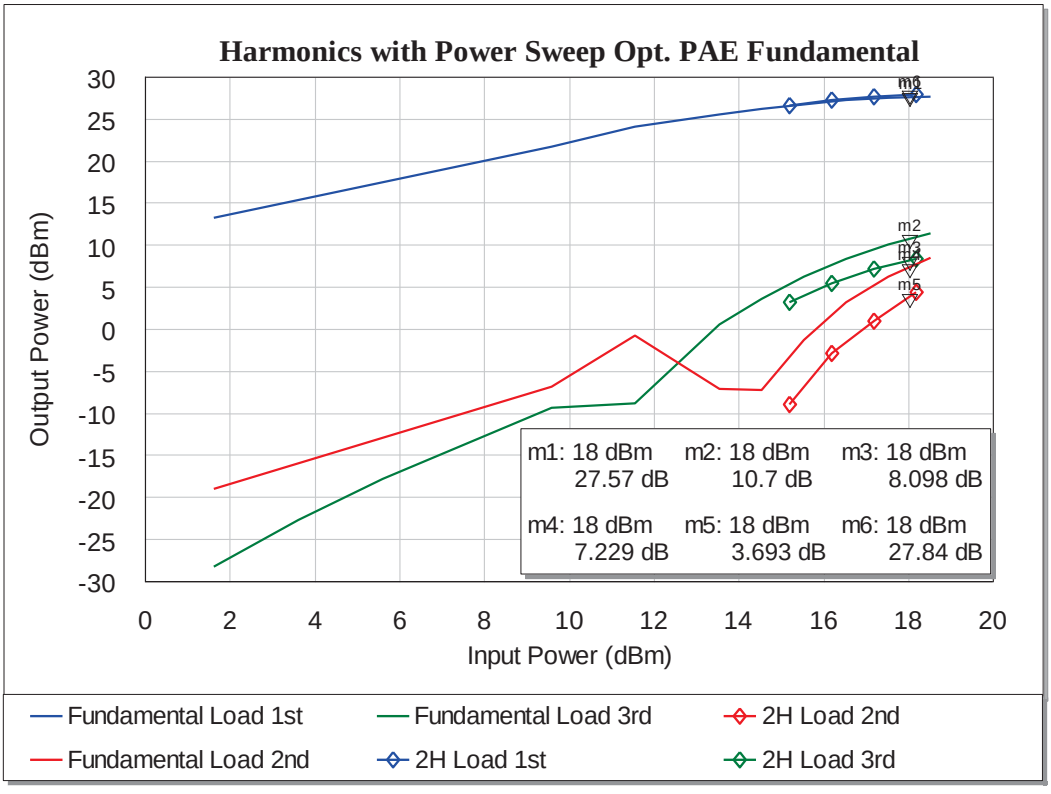


Figure 6-37, Harmonic levels with swept input power at the optimal fundamental load with the 2nd terminated in 50Ω (solid line) and 0.89/_79° (dashed).

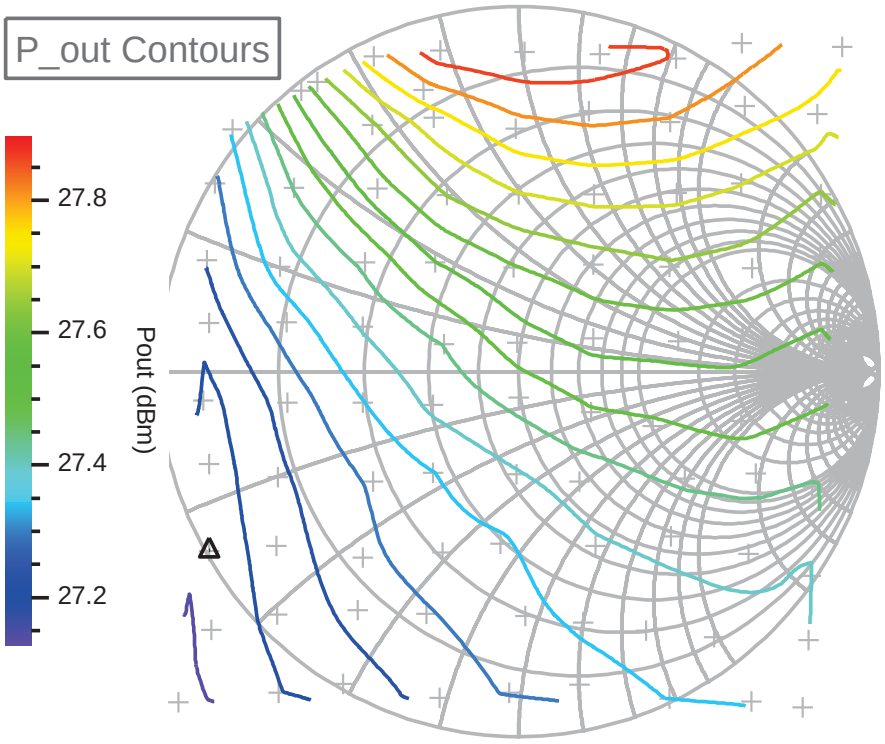


Figure 6-38, 2nd harmonic output power load pull contours with the fundamental (6GHz) at the optimum impedance.

Γ_2 Load (Mag/Ang)	P_{IN} – Avail. (dBm)	P_{OUT} (dBm)	PAE (%)	Max. Gain (dB)	Av. Gain (dB)	P_{OUT} 2nd H (dBm)	P_{OUT} 3rd H (dBm)
0.0/_0°	18.1	27.6	51.0	15.9	9.5	7.5	10.8
0.89/_79°	18.1	27.9	58.8	16.5	9.8	4.0	6.2
	Improvement:	0.3	7.8	0.6	0.3	3.5	2.6

Table 6-7, Performance improvements from 2nd harmonic tuning at a fundamental frequency of 6GHz.

Before moving on to look at the impact of the higher harmonics it should be noted that over about 50% of the impedance plane there is actually a negative impact on PAE and the other key parameters. There is almost a straight line that can be drawn from NW to SE on Figure 6-34 with positive improvement in the upper right half and degrading performance in the lower left. Similarly, although curving upwards in the bottom right, the output power contours, Figure 6-38, roughly divide the impedance plane into a positive contributing half and a degrading half. The power variation across the real impedance plane is ±0.5dB. This clearly shows that whether or not harmonic enhancement is used the impact should not be ignored.

The performance changes with 3rd harmonic load variation are significantly different. The variation across the impedance plane is less, <4% in PAE, but also the

optimum point moves across the impedance plane depending upon the drive power level, Figure 6-39. This obviously makes defining a matching circuit only possible over a very narrow power range. Examining the primary components of PAE, Output power and Gain, Figure 6-40 and Figure 6-41, we can see that the variation across the plane reduces as the peak power encompasses a greater area and whilst there are clear circular contours for the power this is much less obvious for gain. In order to speed up the measurement time a passive load[4] was used for the fundamental impedance. This was because the phase variation due to signal drift was making load convergence take longer. This was discussed in detail in chapter 3 regarding the phase coherence of the load sources. The system is able to phase coherently lock 6 and 12GHz signals together, fundamental and 2nd harmonic, however 18GHz is produced by 2x 9GHz and hence is not phase coherent. The passive reflection coefficient realised, 0.43/_63.6° although close to the optimum was far enough away so as to reduce the output power level and PAE. A comparison is made between the 50Ω and optimum PAE Γ_{3LOAD} in Table 6-8. Although the broad conclusion is that only the 2nd harmonic makes a significant contribution the effect on the dynamic RF loadline was examined to gain some insight as to why the particular load impedances produced the improvements seen.

Γ_3 Load (Mag/Ang)	P_{IN} –Avail. (dBm)	P_{OUT} (dBm)	PAE (%)	Max. Gain (dB)	Av. Gain (dB)	P_{OUT} 2nd H (dBm)	P_{OUT} 3rd H (dBm)
0.0/_0°	18.2	27.1	49.3	15.4	8.8	10.8	10.6
0.55/_-152°	18.2	27.0	49.8	15.4	8.8	11.9	9.4
Improvement:		-0.1	0.5	0.0	0.0	-1.1	1.2

Table 6-8, Performance changes from 3rd harmonic tuning at a fundamental frequency of 6GHz. (Negative improvement means decrease in performance)

[4] By placing a variable phase shifter and an appropriate amount of attenuation on the fundamental port of the Triplexer.

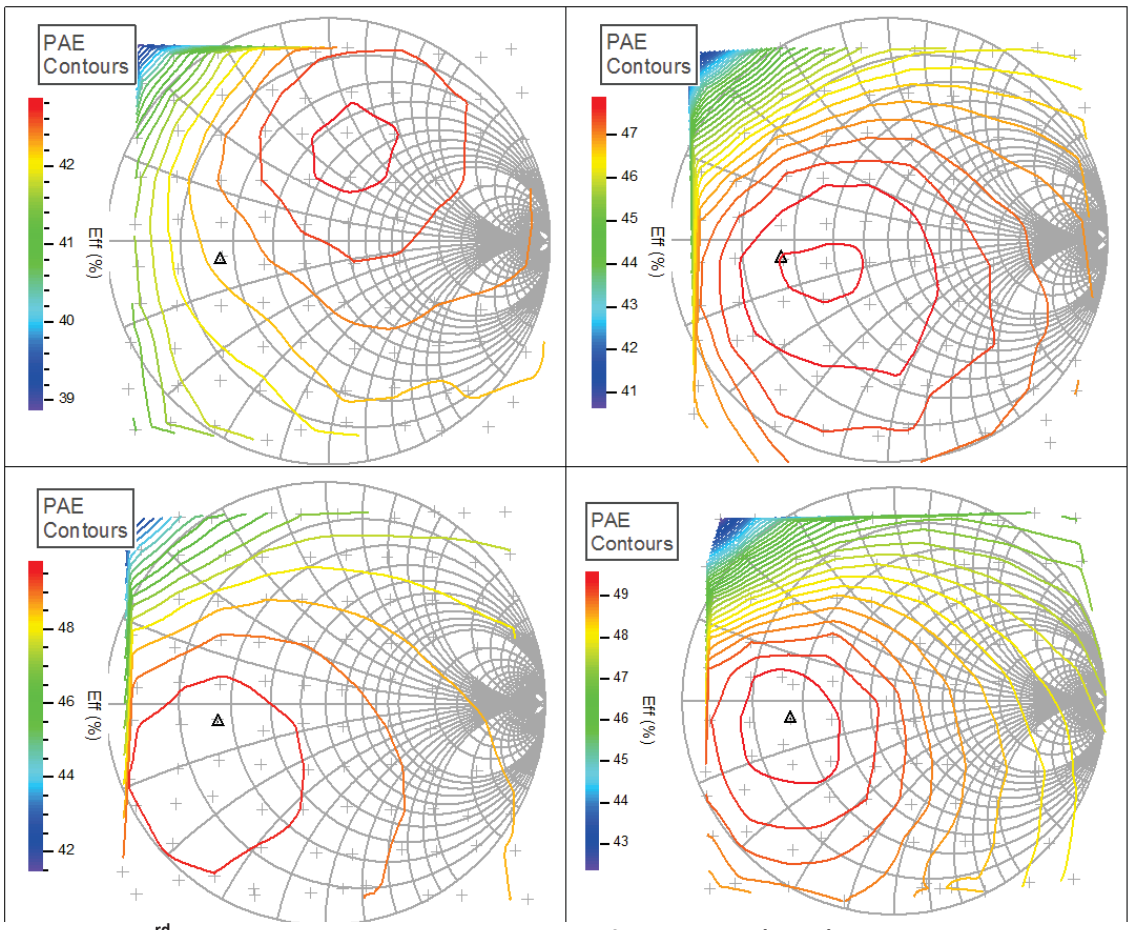


Figure 6-39, 3rd Harmonic load pull contours with the fundamental (6GHz) near the optimum load at 0.43/_63.6°. Drive power levels (clockwise from top left) of 16.2, 17.2, 18.2 and 19.2dBm.

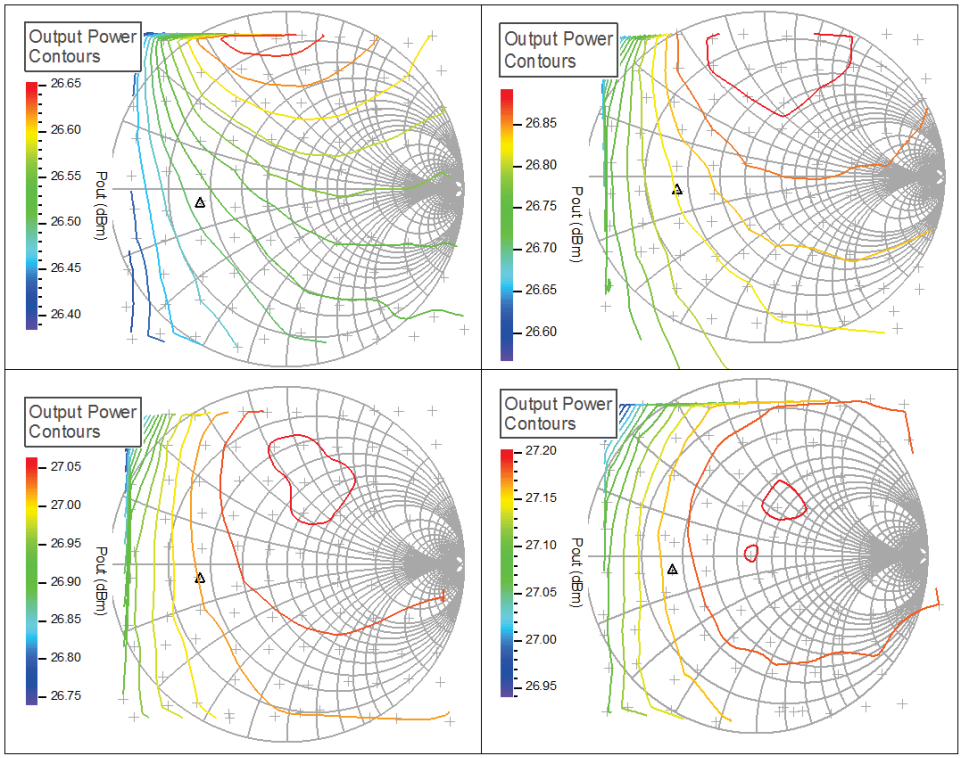


Figure 6-40, 3rd Harmonic Output Power load pull contours with the fundamental (6GHz) near the optimum load at 0.43/_63.6°. Drive power levels (clockwise from top left) of 16.2, 17.2, 18.2 and 19.2 dBm.

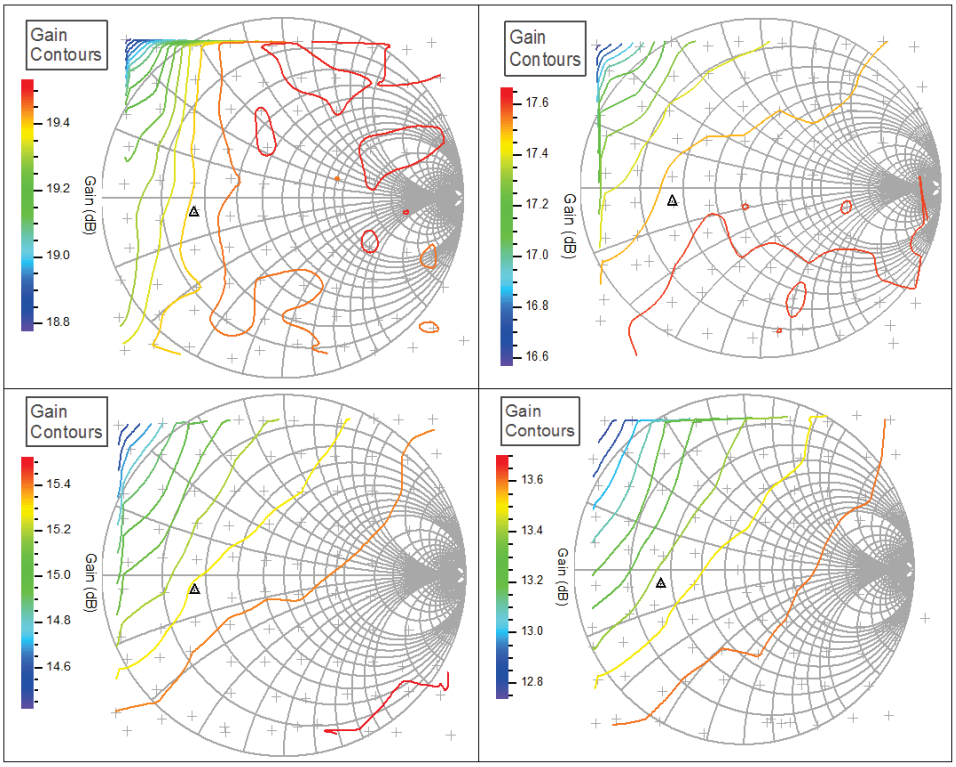


Figure 6-41, 3rd Harmonic Maximum Gain load pull contours with the fundamental (6GHz) near the optimum load at 0.43/_63.6°. Drive power levels (clockwise from top left) of 16.2, 17.2, 18.2 and 19.2 dBm.

Superimposing the measured RF load line (output current and voltage curve) on the measured DC-IV traces of a device allow an understanding of the device behaviour in relation to its operating envelope. Figure 6-42 includes the load lines for 5 combinations of harmonic terminations with a 6 GHz fundamental frequency, displayed in the legend is the associated PAE. It is difficult to see why Load 5 should produce the highest PAE due to the

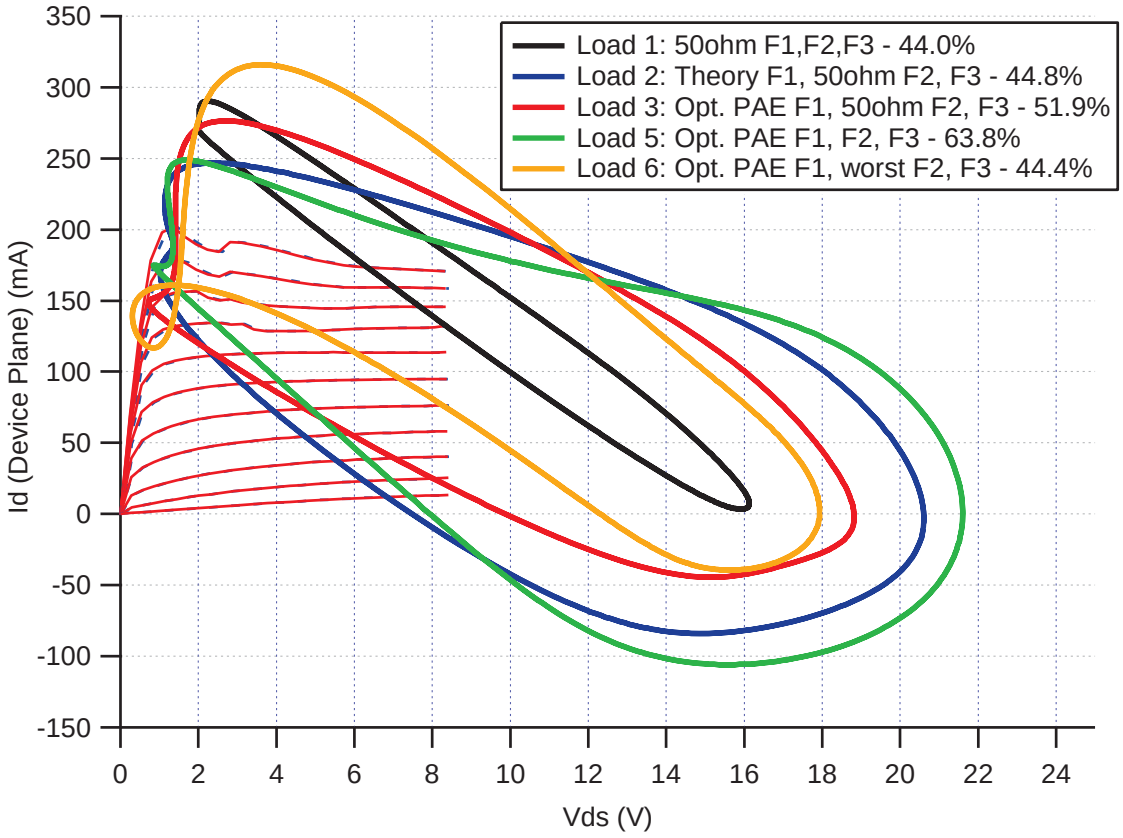


Figure 6-42, Measured RF Load lines for various combinations of fundamental, 2nd and 3rd harmonic loads.

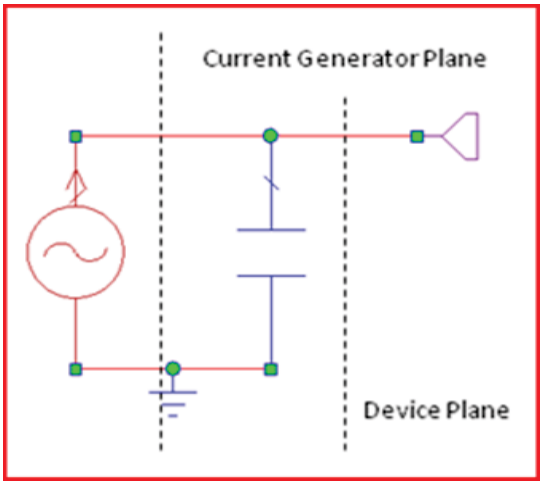


Figure 6-43, Removing the output capacitance to move from the Device to the Current Generator Plane

effect of the displacement current from the output capacitance. Having de-embedded the

calibration plane up to the device plane by removing the effects of the probe 'lands' and the feed line, it is now necessary to remove the output capacitance Figure 6-43, so that the waveforms at the Current Generator (CG) plane can be observed, Figure 6-44; the loads and associated perfomrance are summarised in Table 6-9.

Name	Description	Γ_1 (Mag/Pha)	Γ_2 (Mag/Pha)	Γ_3 (Mag/Pha)	PAE (%)	Pout (dBm)	Gain (dB)
Load 1	50Ω F1 F2 F3	0.09/ 82.6°	0.02/ -116.0°	0.02/ 158.8°	44.0	27.7	14.3
Load 2	Theory F1 50Ω F2 F3	0.23/ 108.9°	0.01/ -179.4°	0.01/ 38.8°	44.8	27.6	14.7
Load 3	Opt F1 50Ω F2 F3	0.42/ 68.1°	0.01/ 171.0°	0.01/ 121.2°	51.9	27.1	15.3
Load 4	Opt F1 F2 50Ω F3	0.42/ 68.1°	0.90/ 72.3°	0.00/ -22.5°	61.4	27.6	16.1
Load 5	Opt F1 F2 F3	0.42/ 65.9°	0.90/ 89.3°	0.86/ -132.1°	63.8	27.7	16.6
Load 6	Opt F1 Worst F2 F3	0.42/ 65.9°	0.87/ -170.4°	0.50/- 152.6°	44.4	26.6	14.8

Table 6-9, Summary of harmonic loads and associated performance on 10x75 device at 9V 150mA bias. Note Loads 2 & 4 are not plotted for clarity.

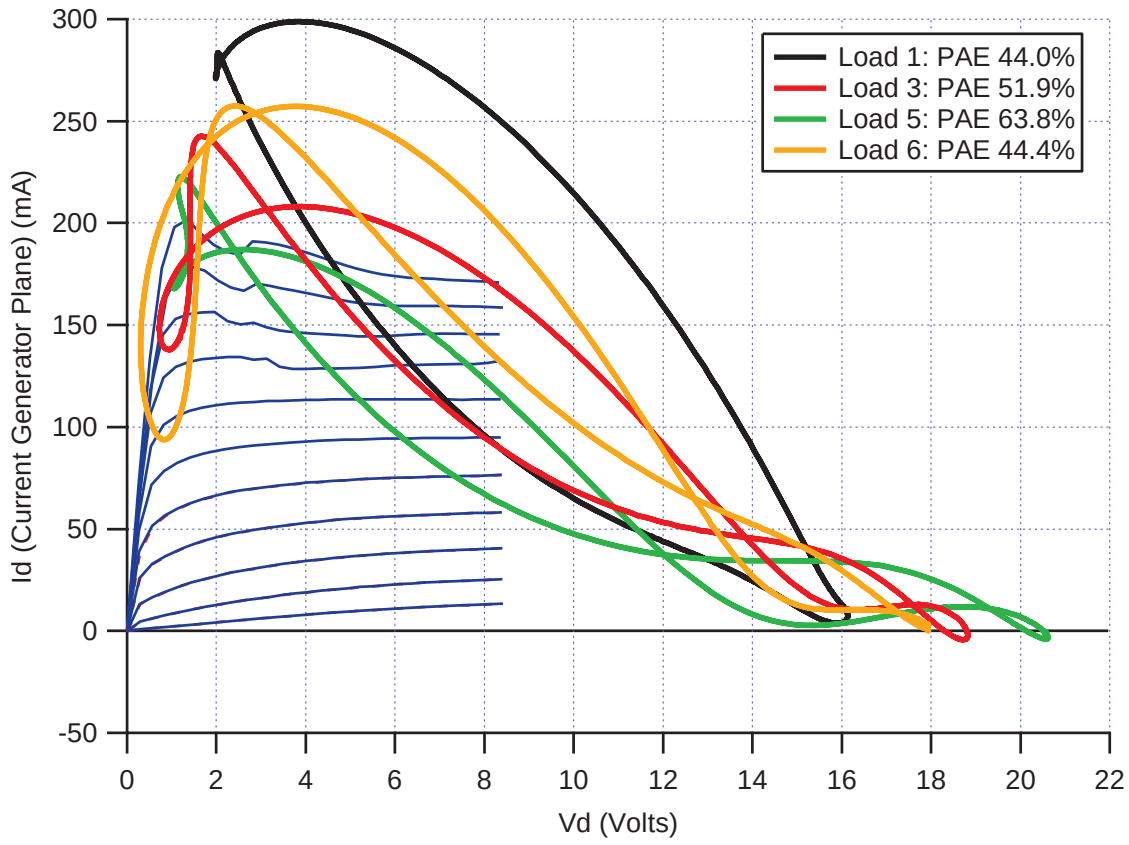


Figure 6-44, RF dynamic loadlines de-embedded to the Current Generator plane.

It is clear that as the harmonic tuning improves the PAE the dynamic load lines moves from being approximately centred on the bias point (9V, 150mA) and moves towards what would be expected from a more class B loadline. It is also seen that in the case of the worst case harmonic loads (Load 6), the effect of the harmonic waveform engineering is to shift the load line to the left such that it is limited by the knee voltage of the device, whereas with the optimum terminations the loadline is shifted to right and less is incident upon the knee and higher peak voltages are achieved. Consistent with load line theory [10] more efficient performance equates to a higher resistance load (reduced slope on the load line). This is perhaps shown more clearly in Figure 6-45; only Load 5 has been included, but at both the Device and C.G. planes. Also included is the loadline of a 45Ω load, (which travels between the I_{DSS} current of 350mA and the bias point) and that for an 82Ω load, determined by the equivalent circuit for the optimum PAE loads measured at 3dB compression. It is important to remember that the fundamental load has been kept constant except for Load 1 where all the harmonics were terminated in 50Ω. The change in the loadline is purely as a result in the change in harmonic impedances. It is also worth noting that a good check of whether the right value of the output capacitance has been chosen is to see whether all of the negative displacement current has been removed by the

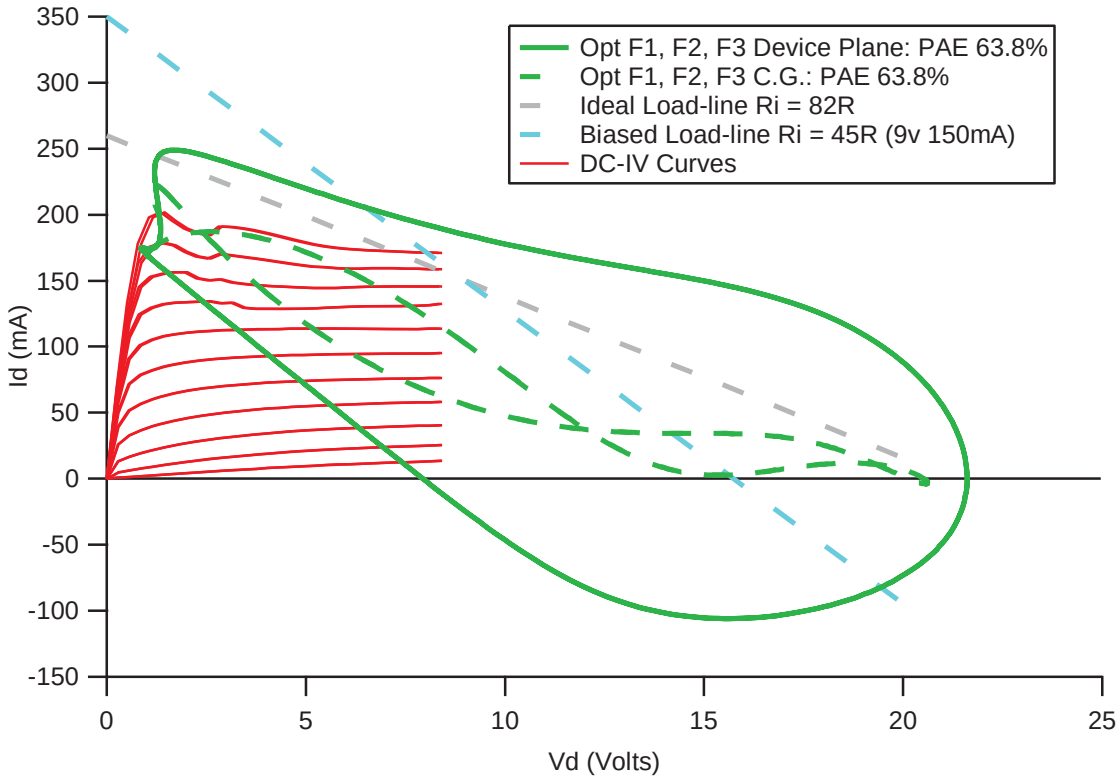


Figure 6-45, Optimum PAE Load lines, Device and C.G. Plane, with resistive load lines.

de-embedding.

Also useful in understanding the behaviour of the device is to view the output current and voltage waveforms. With reference to Figure 6-46, Load 1, with purely resistive terminations (1^{st} – 3^{rd} harmonics) there is only a slight shift in the output current between the two planes and there is only a slight voltage clipping at the knee, with a peak voltage of about 16V. At Load 2 (optimum fundamental load, harmonics in 50Ω) the voltage minimum is now flatter and the peak is >18V. The CG plane current has two, roughly equal, peaks that are more centred on the voltage minima. By adding in the 2^{nd} and 3^{rd} harmonic terminations the reflected harmonic waveforms of the correct phase and magnitude, the waveforms of Load 5 are constructed in such a manner as achieve voltage peaks of >21 volts, but with wide relatively flattened minima (determined by the knee voltage) and CG plane current waveform which is largely in anti-phase with the voltage and tending towards a 'squarish' shape; in the ideal PAE case the voltage and phase are in anti-phase and the current waveform would be square. Wave theory says that an ideal square wave consists of purely odd harmonics, however the nonlinearities in the device generate both even and odd harmonics and so the load must be constructed such that the even harmonics are cancelled and the odd set to the appropriate phase and magnitude. Using only 3 harmonics we will be unable to create an approximation of a square current waveform. The plot of Load 6 shows what happens when the harmonics add destructively. The voltage peak is reduced to 18V and the trough of the voltage waveform has a clear slope rather than being a flat minimum. The CG plane current waveform now has a very pronounced dip in the centre, despite being still in a reasonable anti phase with voltage waveform.

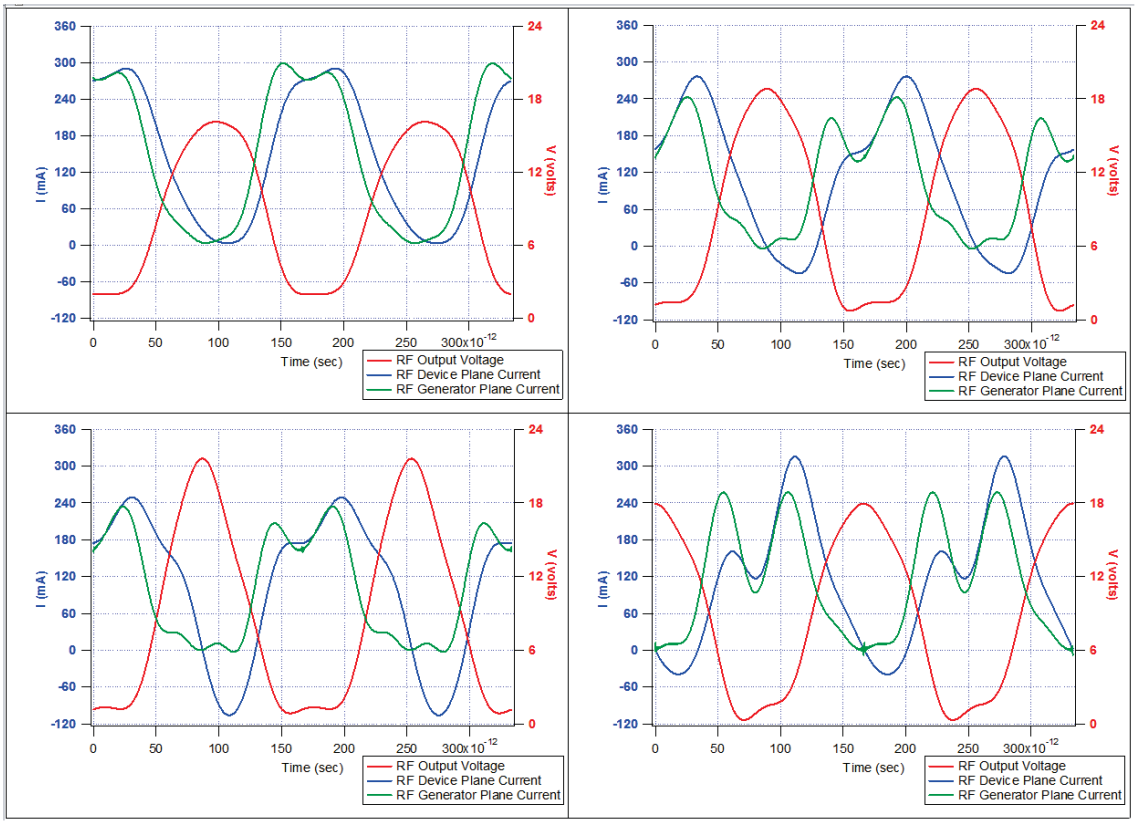


Figure 6-46, Output Voltage and Current Waveforms at (clockwise from top left) Load 1, 3, 6 and 5.

At its simplest the information gathered can be used graphically to guide the designer in the creation of the load matching circuits. For example in the case of an octave band 6 – 12GHz amplifier the fundamental efficiency contours at 6 and 12GHz can be over

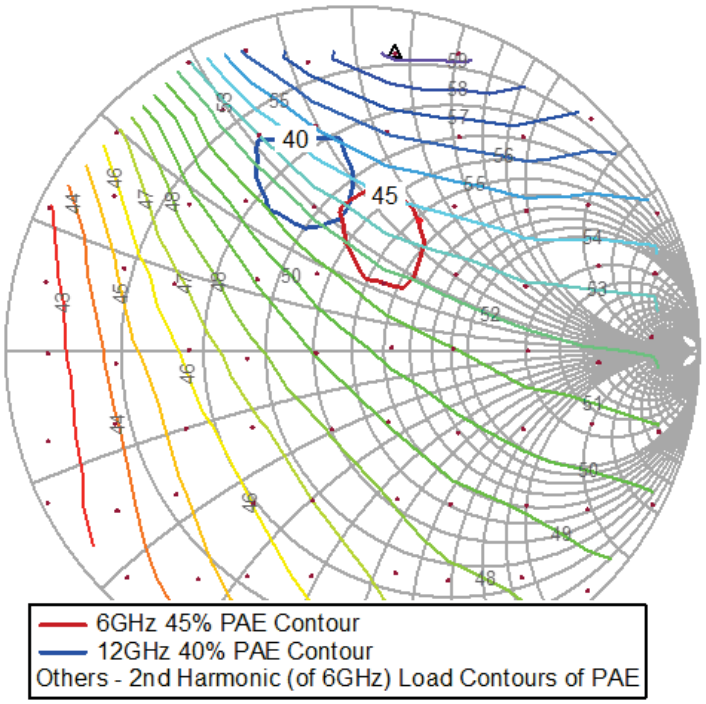


Figure 6-47, Over plotting on 6GHz 2nd harmonic PAE contours of fundamental 6 and 12GHz 45 and 40% PAE contours.

plotted on the 6GHz 2nd harmonic contours as shown in Figure 6-47. Although it was said earlier that harmonic enhancement cannot be implemented for the lowest frequency in an octave band circuit, we can see here how it can be used as guide, informing the designer that a load at 12GHz at the highest point on the 40% contour will give an additional 4% to the PAE at 6GHz compared to the lowest point on the 12GHz contour. Conversely for a two octave, 6-18GHz design, looking at the 3rd harmonic contours, Figure 6-48, shows that the lower portion of the 18GHz 30% PAE contour gives a better result at 6GHz. Even where it is decided that the difference in PAE doesn't warrant the extra effort this graphical technique gives the designer the information necessary to **know** that they don't need to specifically target a harmonic impedance. Without the mapping of the harmonic behaviour designs are implemented with a 'blind spot'.

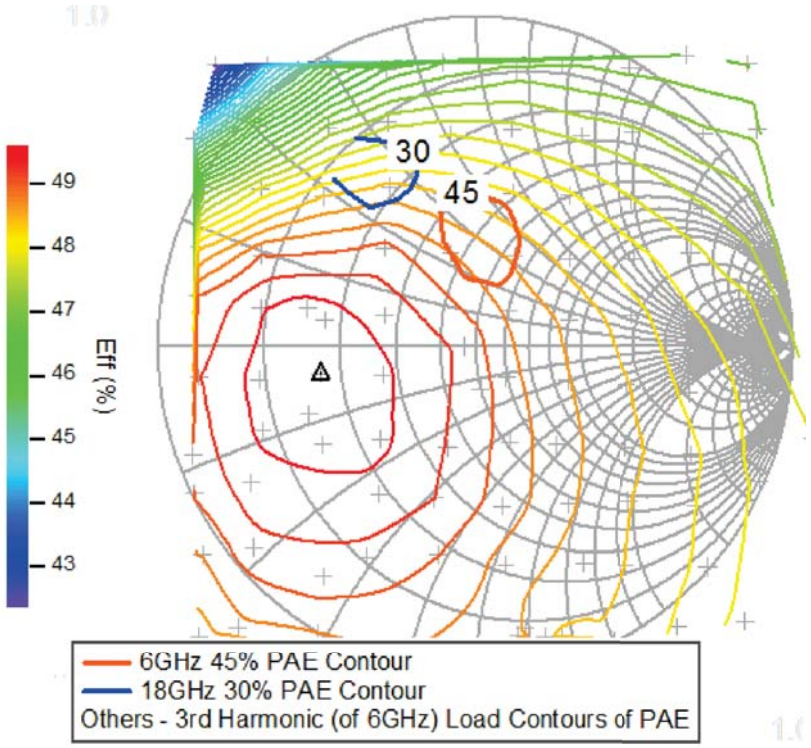


Figure 6-48, Overlay of 6 and 18GHz 45 and 30% PAE contours on 6GHz 2nd harmonic PAE load pull contours.

Using the harmonic information and fundamental optimum load impedances obtained, the next step was to produce a matching circuit that would target the 2nd harmonic load impedance in the region of highest efficiency improvement. Note that the power and efficiency levels recorded are at the device plane and hence an allowance needs to be made for the insertion loss of the output matching circuit.

Broad bandwidth design is about balancing conflicting objectives. In order to come as close as possible to the ideal impedances multiple matching sections are required, however this increases insertion loss and hence sacrifices efficiency and power. A compromise is necessary which provides adequate overall performance. To achieve this it is extremely helpful to work within the CAD environment, firstly with a simple equivalent circuit model of the device output for speed and then with the DLUT model, where the device performance and circuit losses can be examined together. At the time of this research a limitation of this approach was that to produce a model that contained the information on fundamental, 2nd and 3rd harmonics was not practical. For each fundamental load impedance point it would be necessary to conduct a 2nd harmonic load pull, and remembering that for each fundamental load point there are 10-15 power points, that there would need to be >5 power points on the 2nd harmonic sweep, and finally that these would not be the same for each fundamental load point, thus the fundamental measurement, data

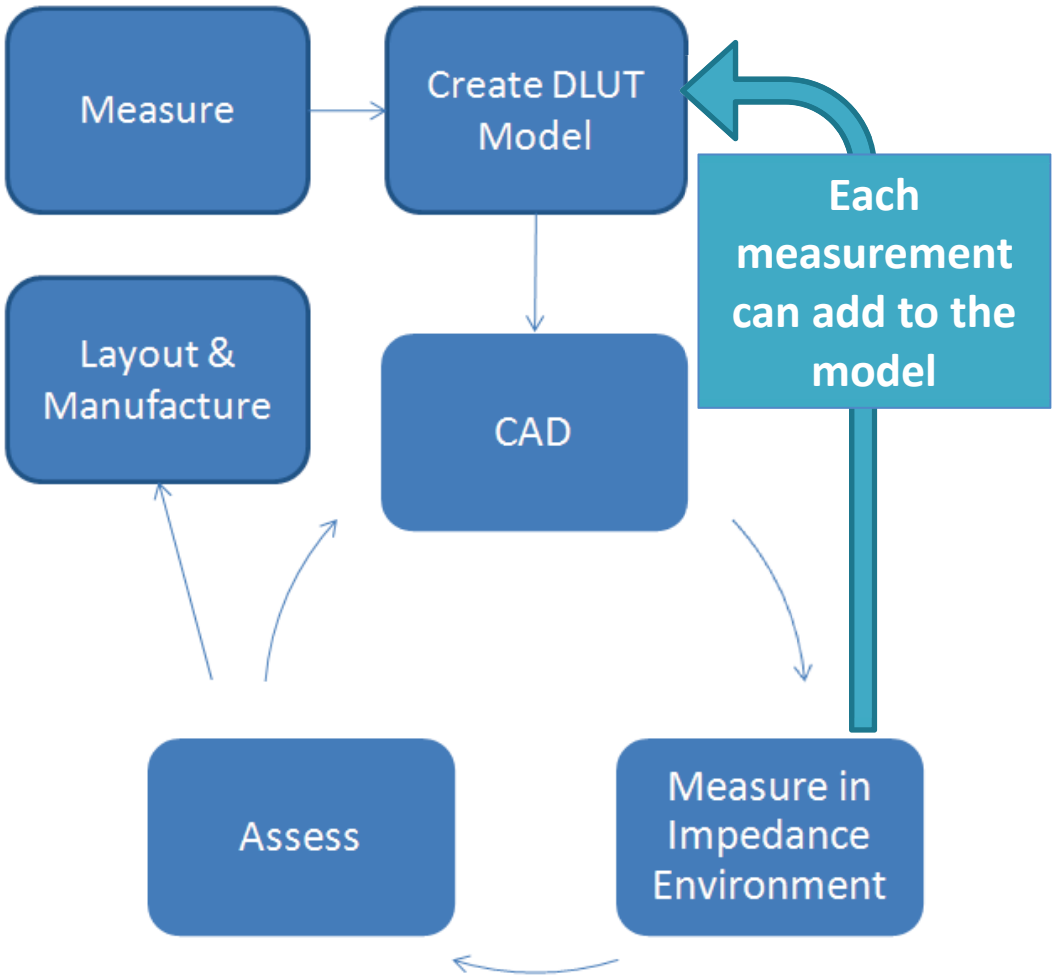


Figure 6-49, Design process for the harmonically enhanced output matching.

analysis, 2nd harmonic measurement and acquired data file become very difficult to manage, (the fundamental only data set for the 6GHz load pull, 64 grid point and 12 power levels took >130MB of disk space – by contrast the DLUT model was 163kB). Future improvements in measurement speed and the move to Poly Harmonic Distortion or Behavioural models, where instead of large data banks multi term equations are used to describe the contours, will make this process more feasible and the full harmonic based models more practical.

In this case therefore the process has started with measuring the optimum fundamental loads and then measuring the harmonic performance at these optimum impedances; creating a DLUT model from the measured data. Within the CAD environment the load equivalent circuit and simple transmission line models are used for the initial

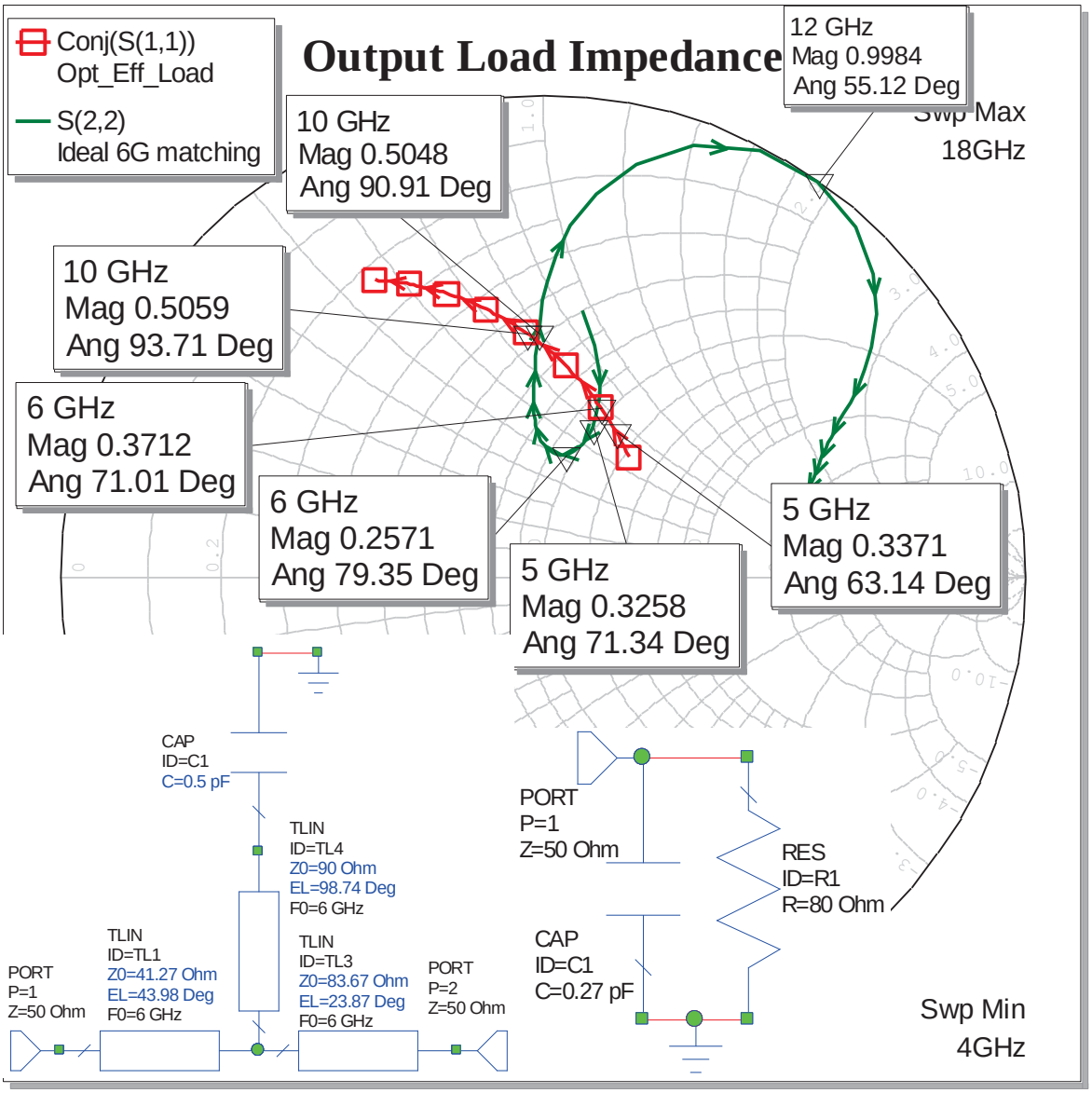


Figure 6-50, Initial output matching circuit trajectory, tuned to match at 5 and 10 GHz and harmonic enhancement at 6GHz.

matching circuit design as a first approximation of the solution. These are optimised such that the load matches the equivalent circuit at the band edges and the harmonic termination at 6 GHz is close to the measured optimum at near 1/_80°. With each set of measurements data is added to the model, thus its comprehensiveness increases with each data set acquired, Figure 6-49. The first pass solution is shown in Figure 6-50. Although the angle of the harmonic is not exactly right until the improved circuit models are included it is a satisfactory start. An advantage of this matching topology is that drain bias could be incorporated at the end of the capacitor coupled stub.

The next iteration of the matching circuit includes discontinuities for the 'Tee' junction and the microstrip line models for the GaAs substrate. This alters the impedance

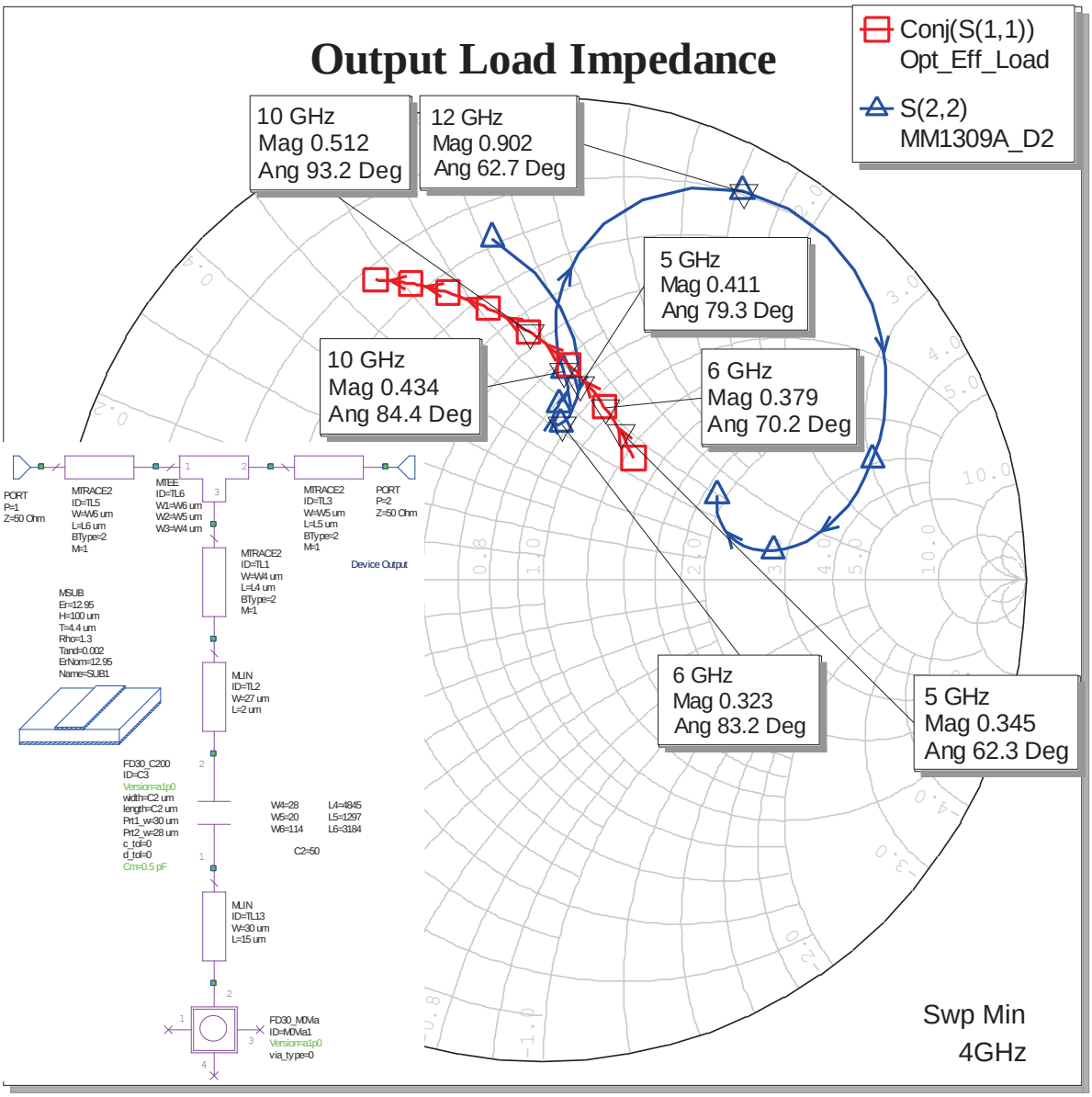


Figure 6-51, Linear simulation of output circuit and target load impedances with matching circuit inset.

response as shown in Figure 6-51. To assess the impact of these loads on the device a load pull was conducted at 10GHz, this was not a frequency measured earlier and thus the additional information can be used to increase the accuracy of the DLUT model, as depicted in Figure 6-49. From the load pull measurements at 10GHz it is seen that with a load reflection coefficient of 0.43/_84° the device will deliver an output power of 27.4 at a PAE of 43% and a maximum gain of 11.5dB. The loss of the output matching circuit needs to be calculated to see whether this will meet the target of 40% PAE and 27dBm output power. The insertion loss simulation is shown in Figure 6-52 and effect of insertion loss on PAE can be determined from the graph of Figure 6-53. This suggests that the output power would be reduced to 26.9dBm and the PAE to 33%, which will not therefore meet the requirements.

At this point in the research the accuracy of the PDK models was not known and so an Electro-Magnetic (E-M) simulation of the circuit was performed. This required separating the simulation into two parts, an E-M section for the transmission lines and the PDK model for the capacitor (there was insufficient information on the capacitor dielectric material to include this in the E-M). This gave a more favourable result both in terms of insertion loss and the magnitude and angle of the 2nd harmonic termination, Figure 6-55. Reading these loads across to the measurement data the estimated PAE at 10GHz is 44.3% at an output power of 27.6dBm. Thus allowing for the output circuit loss of -0.21dB, Figure 6-54; the simulation gives an estimated performance for the device with output matching of 40% PAE and 27.4dBm.

At 6GHz without any harmonic enhancement the performance at a load of 0.32/_98° is a PAE of 46% and an output power of 27.8dBm, thus allowing for circuit losses we would expect a PAE of 42% and an output power of 27.2dBm. This would suggest that we could afford to shift the load closer to the optimum at 10GHz, however the performance at 5GHz looks to have less margin and hence the decision was taken to proceed with this matching circuit. Measurements were not made at 5GHz due to a shortage of time and the restriction on 2nd harmonic phase locking at 5GHz (10GHz was a fundamental frequency and hence the two would not be phase coherent).

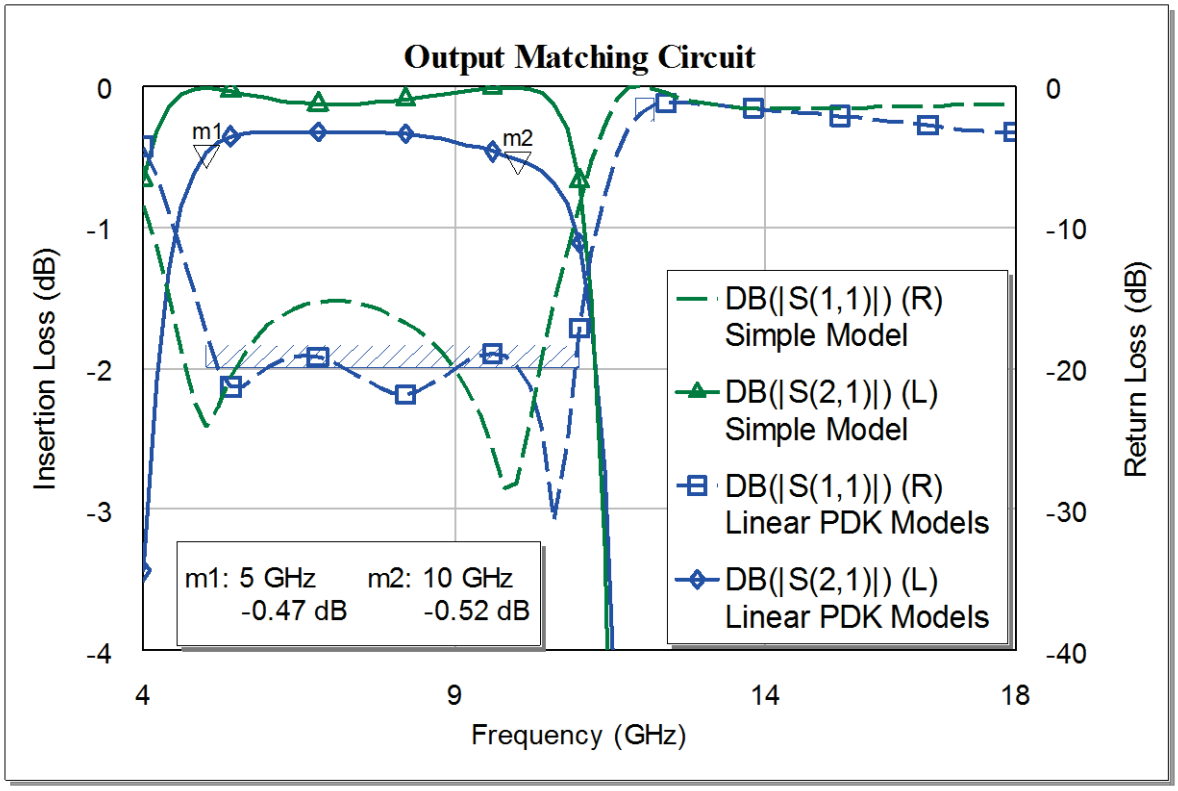


Figure 6-52, Insertion and Return Loss of output matching circuits, simple and PDK models.

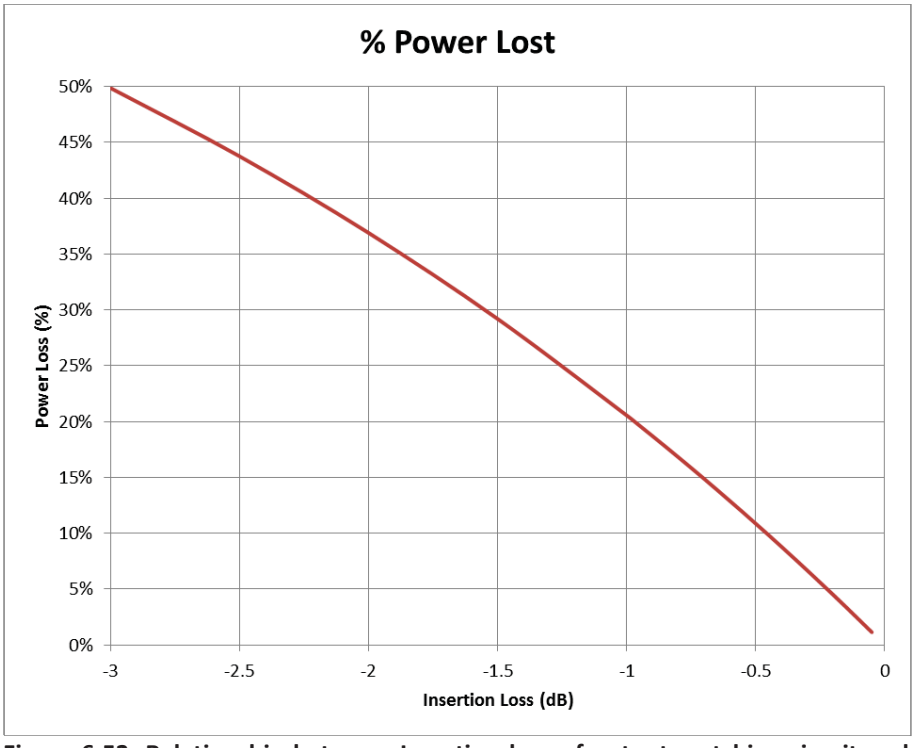


Figure 6-53, Relationship between Insertion loss of output matching circuit and PAE.

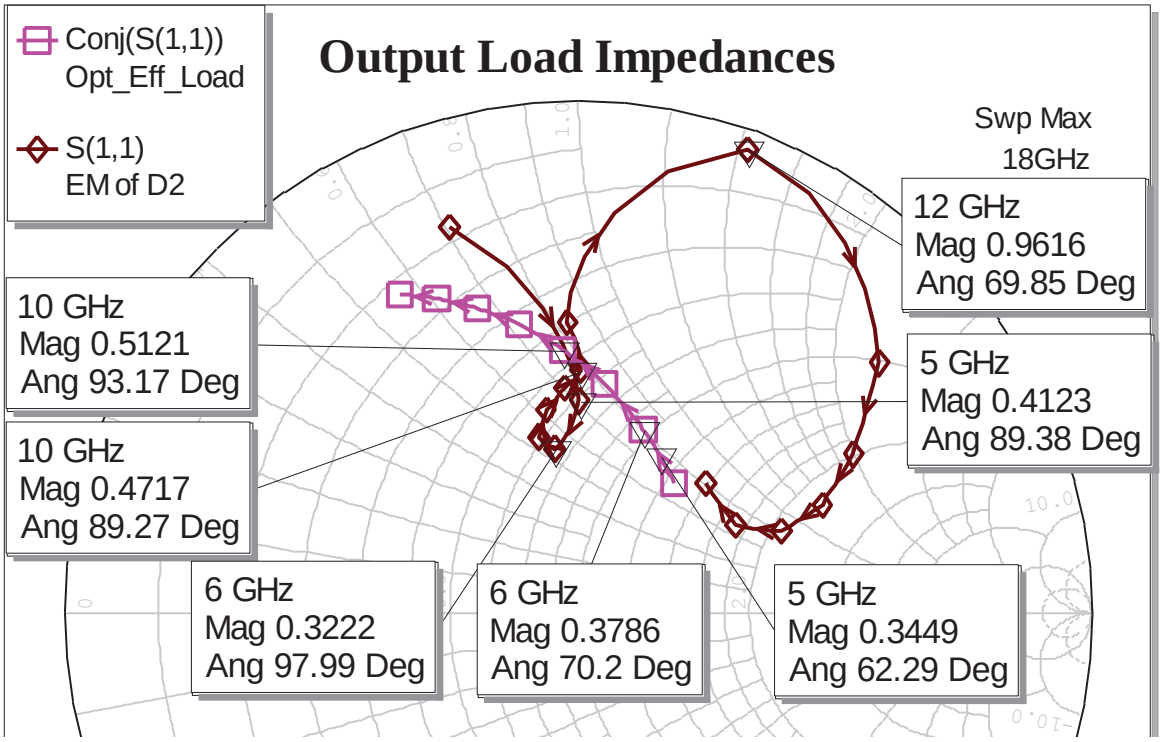


Figure 6-55, E-M Simulation of Output Matching circuit load impedances.

These estimations are based on the performance of the current device, as before a device is included on the MMIC to evaluate any changes in the device performance between

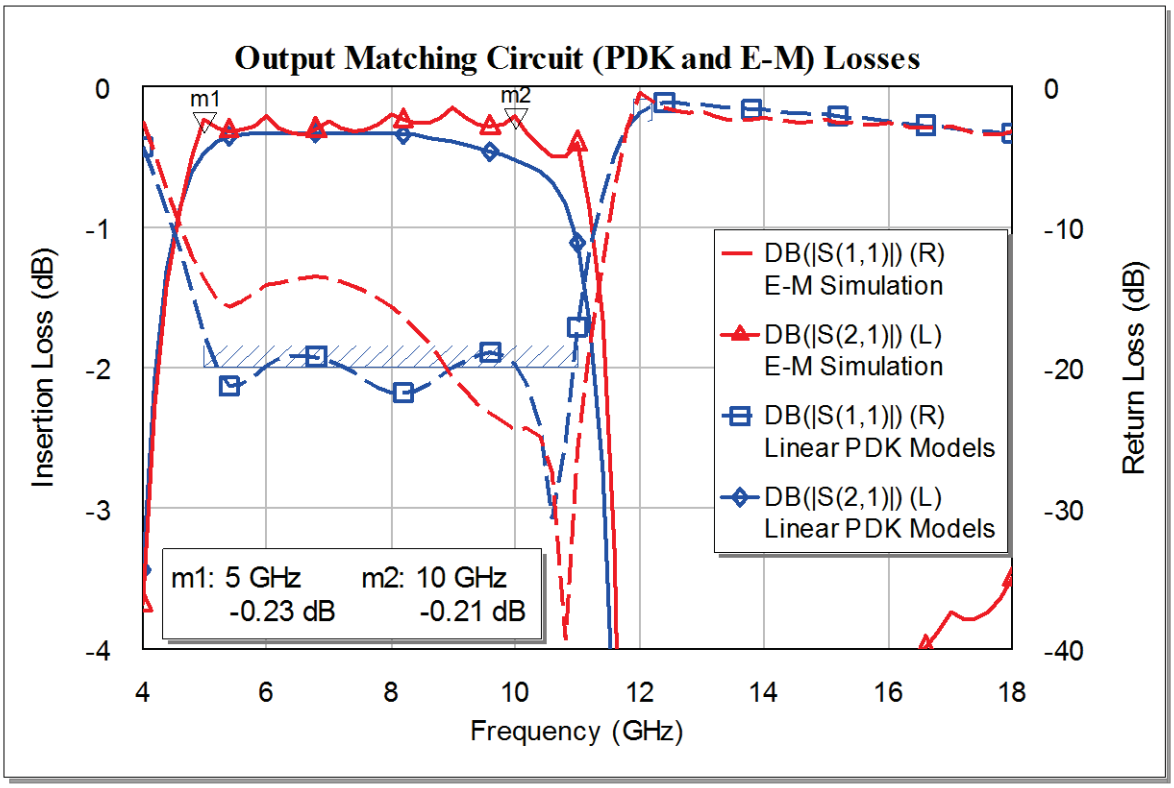


Figure 6-54, Circuit losses from the PDK linear and E-M simulations of the output matching circuit.

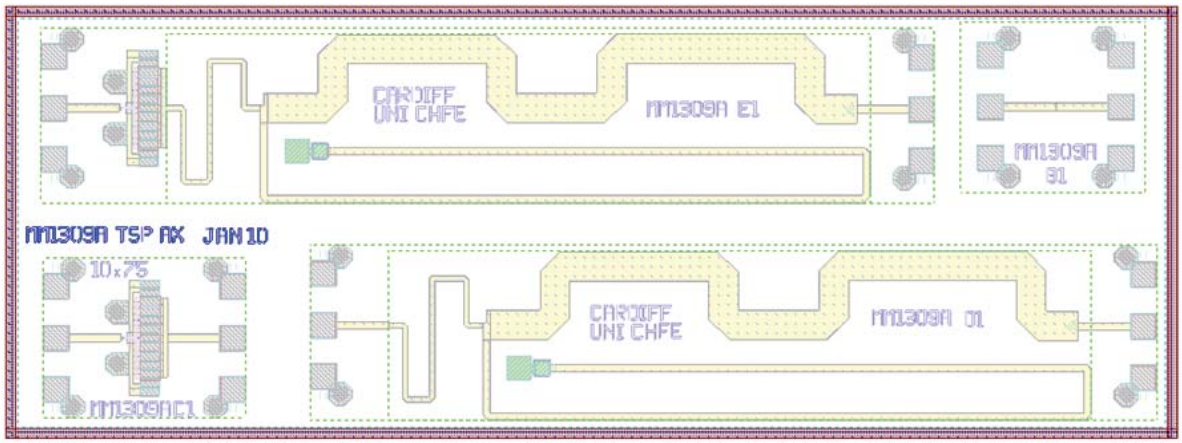


Figure 6-56, Wafer cell layout for harmonically enhanced MMIC design

wafer runs. It was decided to leave off the input matching so that a more accurate measure of the input power to the device could be made and thus a more accurate comparison between transistor on its own and one with output matching circuit. It was found in the first MMIC design iteration that the input matching by nature of the fact that it attempts to do gain equalisation and improve stability is lossy, this and the fact that a perfect match is not achieved results in a significant increase in uncertainty as to the actual power into the device. Input matching is by its nature a less significant challenge than high efficiency harmonic output matching and so deemed less import at this stage of the research. The designed wafer cell is shown in Figure 6-56. As before the output matching circuit on its own has been includes as has a de-embedding cell and the device on its own.

It is worth noting that analysing the circuit performance with the foundry PDK nonlinear model suggests that the harmonic termination will not improve the performance compared with terminating the harmonics in 50Ω, Figure 6-57.The model also predicts a higher PAE than the measured results (of the device with harmonics in 50Ω). To re-iterate a point made earlier in this work, this must not be taken as a criticism of the PDK model, but an observation of the fact that the PDK model cannot be valid under all conditions. It is a compromise between many measurements. A valid criticism of the DLUT model is that it is historical and based upon a single set of measurements (currently this is being further developed [11] [12]) and therefore depends upon the consistency and repeatability of the foundry process.

The manufactured circuit, Figure 6-58, was received and the performance measured. Across the design bandwidth of 5-10GHz a minimum PAE of 43% and associated output

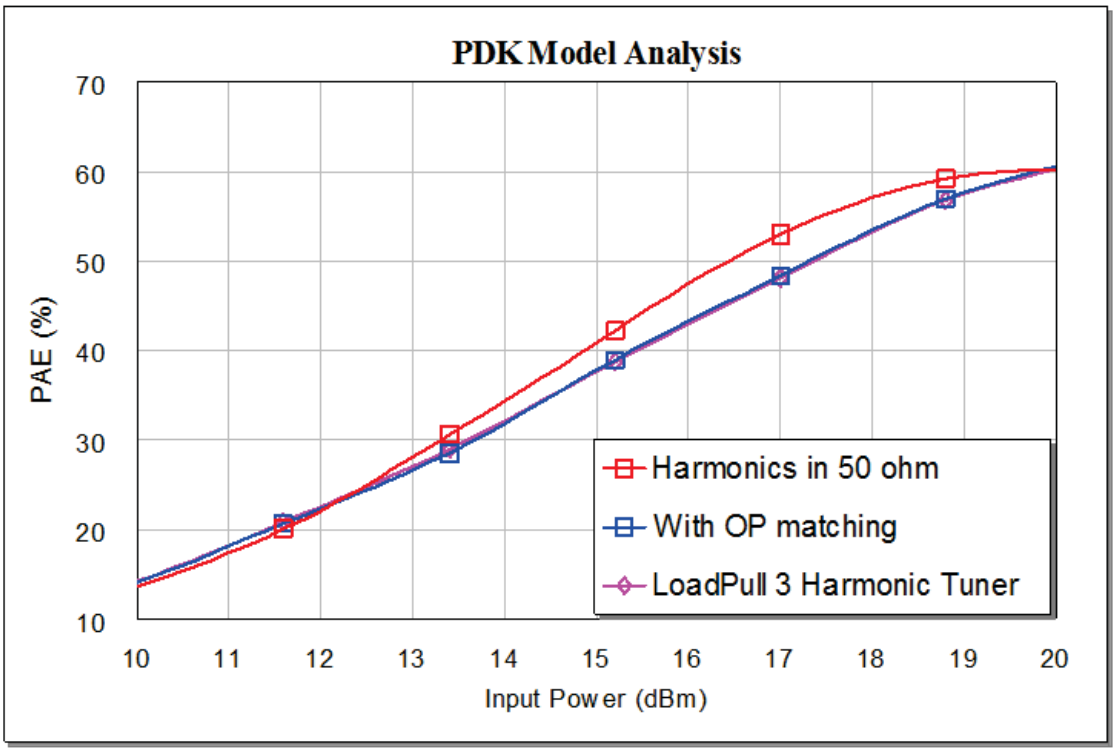


Figure 6-57, Simulation of PDK NL Model with LP Tuner with (a) harmonics terminated in 50Ω, (b) output matching circuit and (c) tuner using matching circuit impedances.

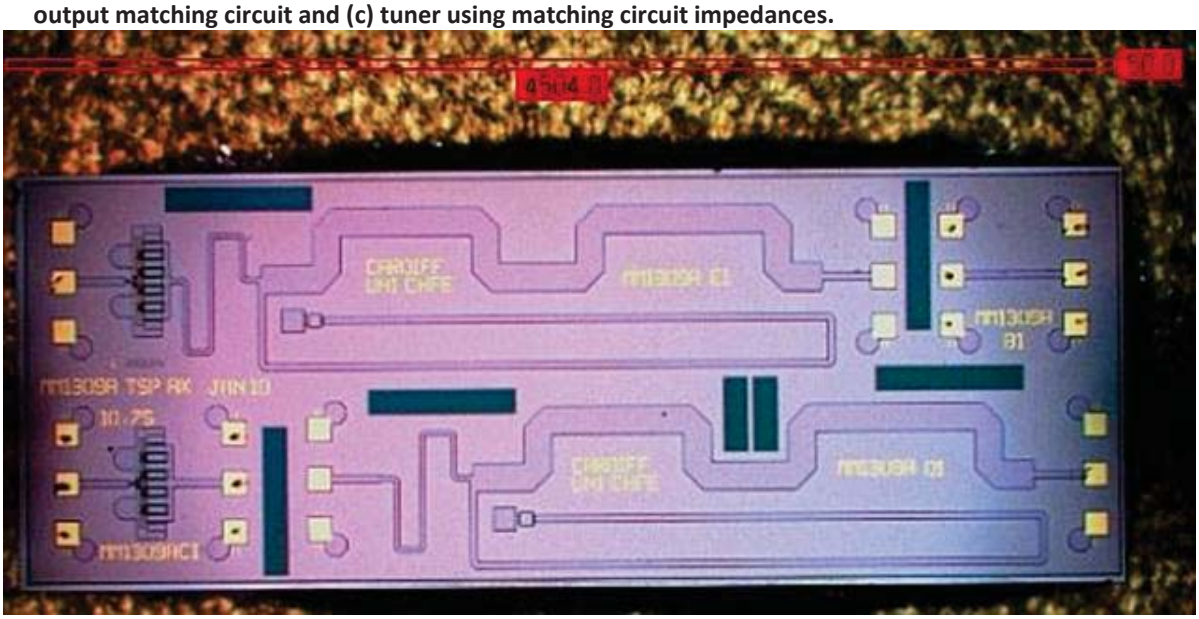


Figure 6-58, Manufactured wafer cell with scale bar in microns.

power of 27.4dBm was achieved, Figure 6-60. The results are in fact better than those estimated from the initial device measurements.

In order to investigate why the manufactured circuit behaves as it does the output matching circuit and device on wafer were measured. The first step was to examine the loads presented by the output matching circuit. Figure 6-59, shows the measured results from 2 wafer cells and compares them with the simulation.

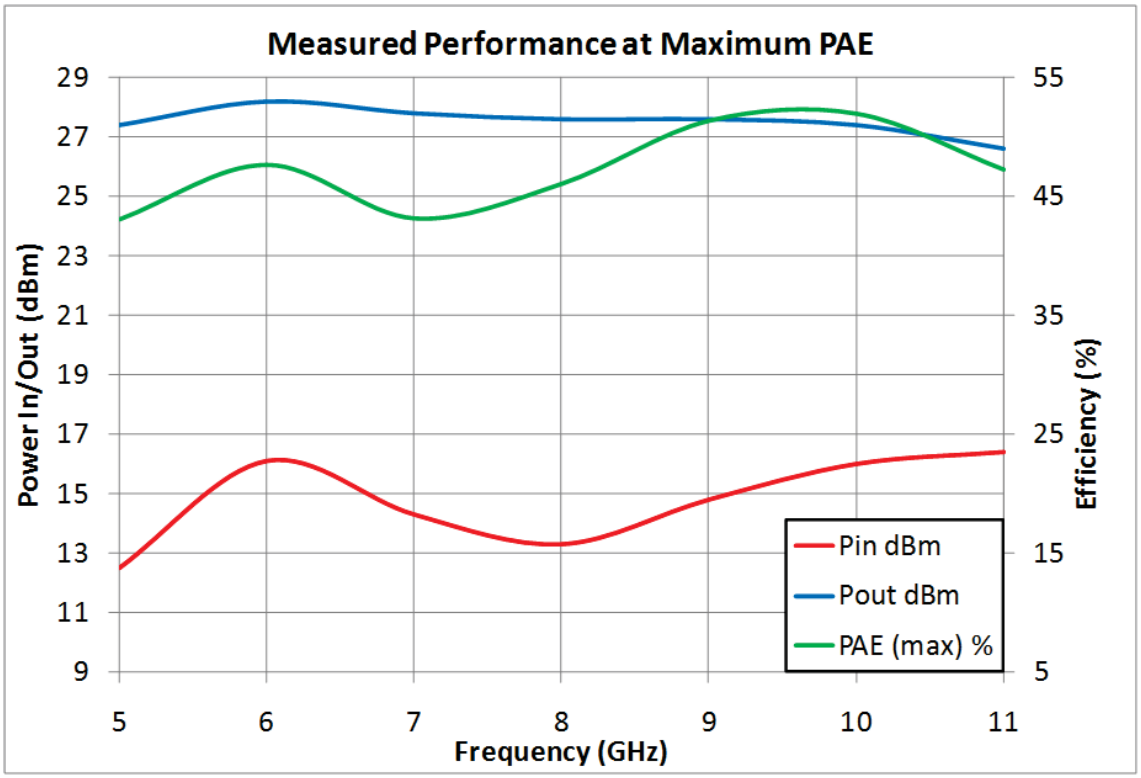


Figure 6-60, Measured performance of MMIC MM1309A including optimum drive level.

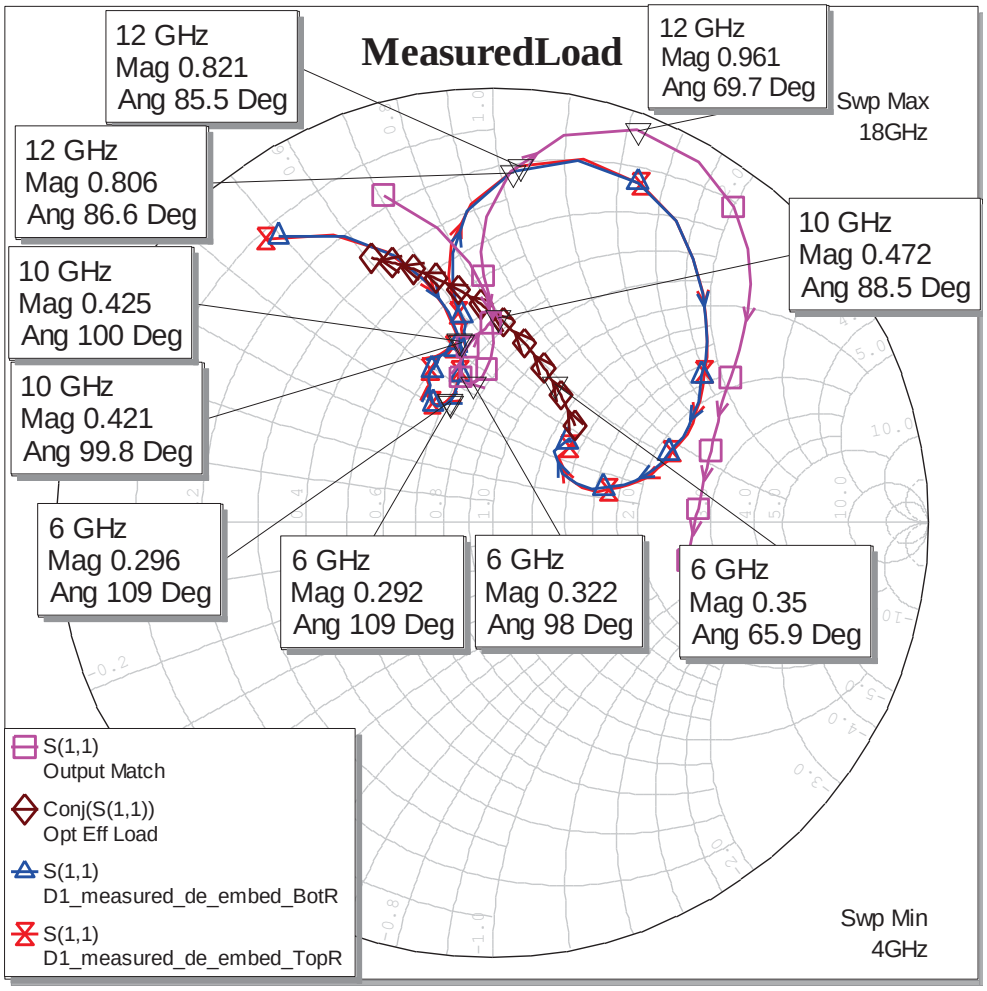


Figure 6-59, Target, design and measured load impedance trajectories for MMIC 1309A.

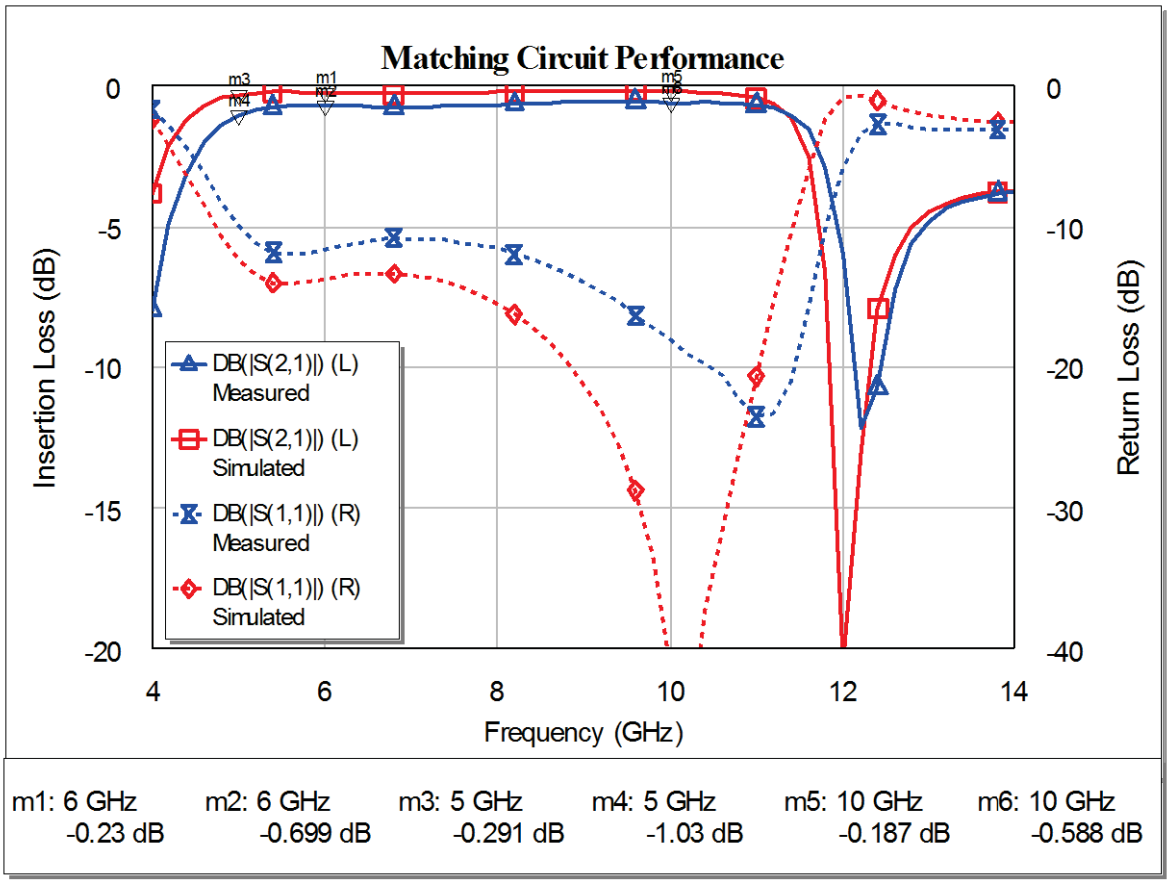


Figure 6-61, Comparison between output matching circuit simulated and measured losses.

The original device at 6GHz when loaded with the fundamental impedance of 0.29/_109° (the manufactured circuit load reflection coefficient at 6GHz) would have had a PAE of 42.5% and an output power of 27.6dBm. The 2nd harmonic load pull that was conducted on the original device was with a fundamental load set to 0.43/_63°; it is not known what the exact effect of harmonic enhancement will be at a different fundamental load, but if we assume a similar improvement, a 2nd harmonic load of 0.82/_86° should enhance PAE by 10% and the power by 0.1dB. This is of course at the device plane; the insertion loss of the output circuit must be accounted for in order to compare with the MMIC performance. As shown in Figure 6-61 the insertion loss is about 0.5dB higher than expected. This 0.5dB insertion loss would be expected to decrease the PAE by 10% negating the impact of the harmonic tuning and the associated output power would be expected to be 27.2dBm. As can be seen from Figure 6-62, 48% PAE and 28.0dBm output power was actually achieved at the matched circuit output. Adjusting the data in this graph to account for the output circuit losses Figure 6-63 is produced, showing a peak PAE of 58% and an output power of 28.8dBm. The next stage was thus to measure the device in the wafer cell, this data is summarised in Figure 6-64 and shows that not only is the optimum PAE 3.7%

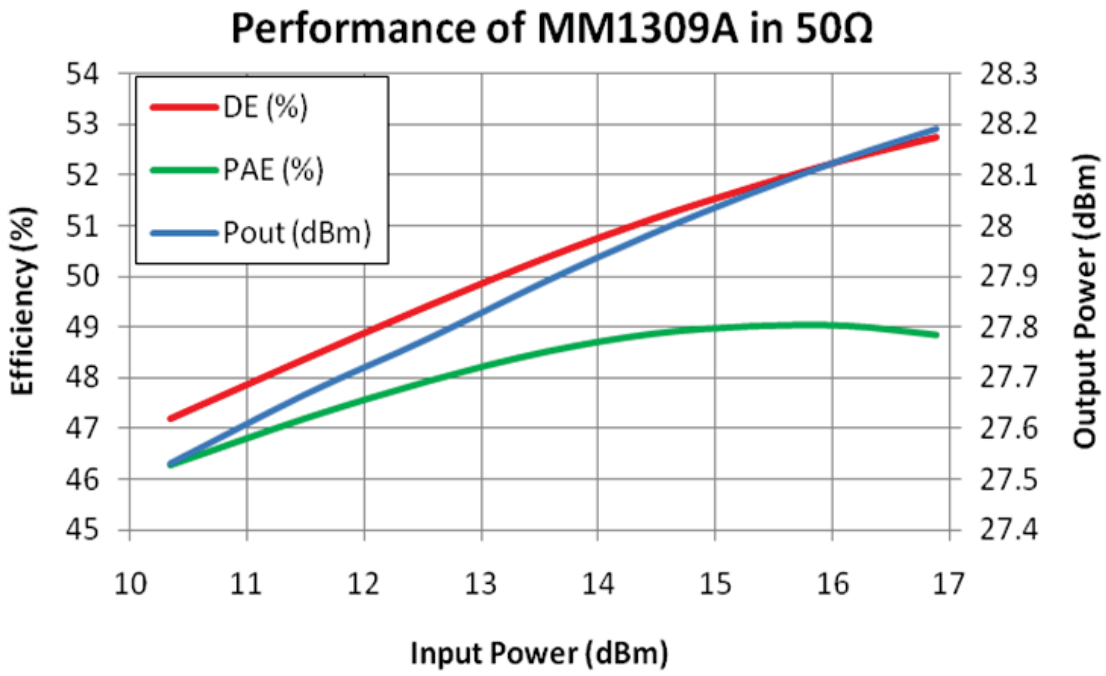


Figure 6-62, Performance of MMIC MM1309A at 6GHz with swept input power.

higher, but that the 50% PAE contour now covers a significantly larger proportion of the impedance plane. This accounts for the higher PAE achieved with the MMIC. Loading the device with the closest fundamental load that could be achieved with the passive load pull tuner, 0.33/_90° a 2nd harmonic load pull was conducted.

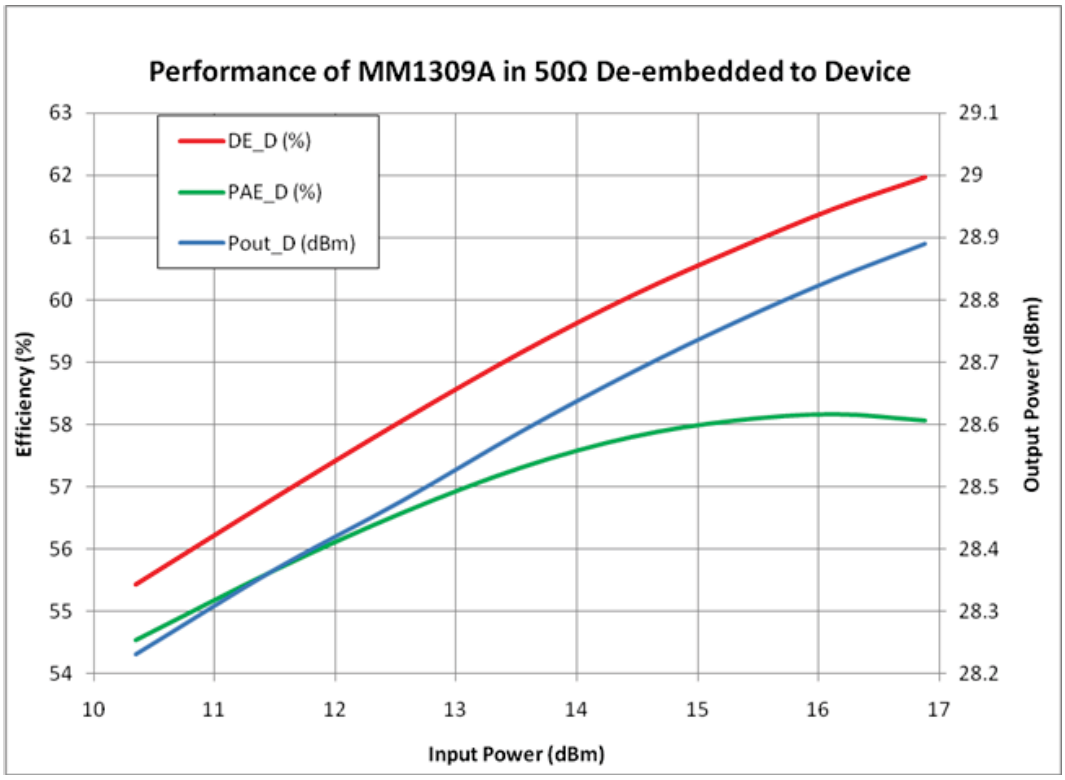


Figure 6-63, Performance at the output of device by de-embedding the output circuit losses.

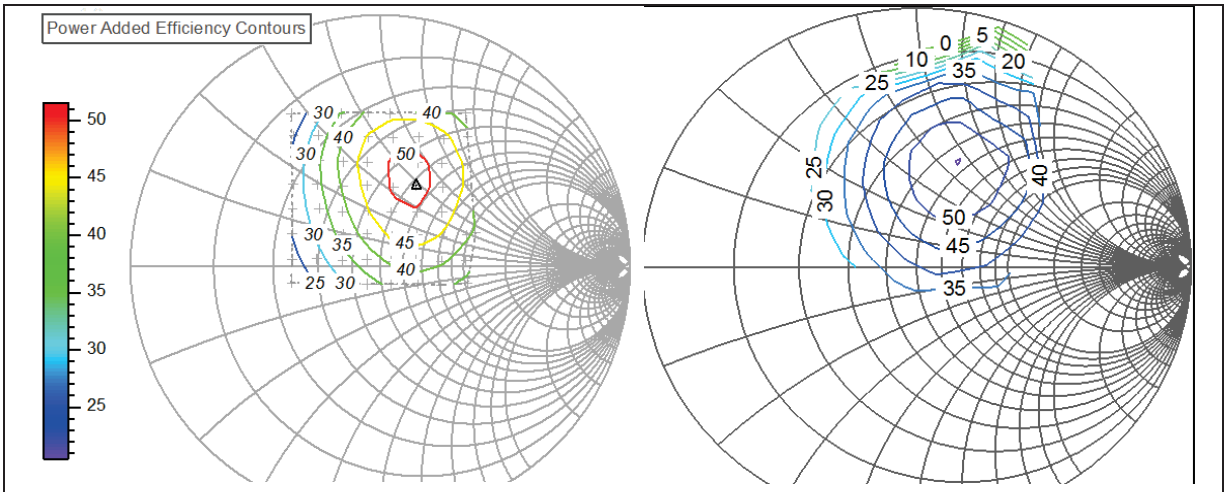


Figure 6-64, Original 10x75 device and device on MMIC wafer PAE LP contours at 6GHz.

Original Device on which design was based: 6GHz PAE opt. Load: 0.35/_66° Performance at optimum load: PAE = 51.4% Pout = 27.7dBm Gain = 15.0dB	Device on MM1309A Wafer Run: 6GHz PAE opt. Load: 0.43/_70° Performance at optimum load: PAE = 55.1% Pout = 27.5dBm Gain = 14.6dB

The 2nd harmonic load pull measurement grid, output power and PAE contours are shown in Figure 6-65 to Figure 6-67, although the fundamental load is slightly different from the measured fundamental output circuit load at 6GHz, 0.33/_90° as opposed to 0.29/_109°, the 2nd harmonic enhancement clearly follows the pattern measured on the original device and the power levels and efficiency are consistent with those measured on the MMIC when corrected for the 0.4dB output matching circuit loss.

The efficiency enhancement provided by the harmonic tuning at 6GHz is clear from the PAE 'hump' seen in Figure 6-60; as the actual load has moved further from the optimum PAE point it has moved closed to the optimum output power and hence there is an increase in this level as well. It can also be seen from Figure 6-60 how the input drive level requirements vary across the frequency band for optimum PAE. In typical amplifier designs this can be used to guide the design of the input matching circuitry to provide the required drive power profile. This will however invariably require lossy matching elements, which although useful for stability are not ideal for gain and efficiency. It is conceivable in 'smart' systems that the drive level can be set depending on the operating frequency to maximise efficiency according to this data.

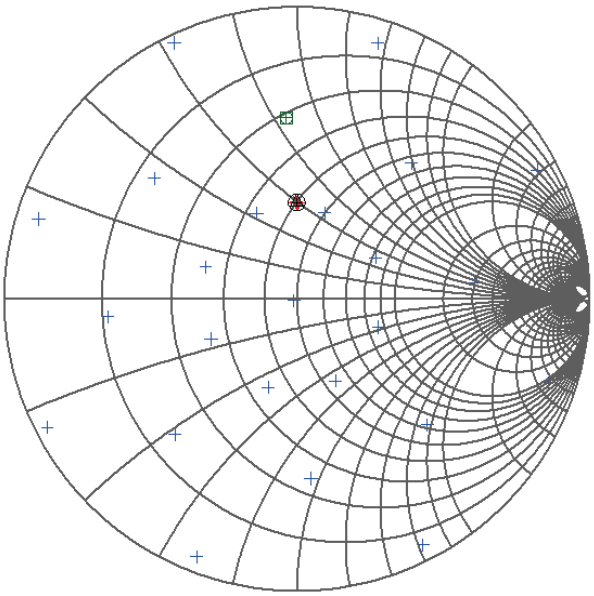


Figure 6-65, 10x75 device from MMIC 1309A wafer 2nd harmonic load pull grid (blue crosses) and fundamental load (red crosses) ~0.33/_90°.

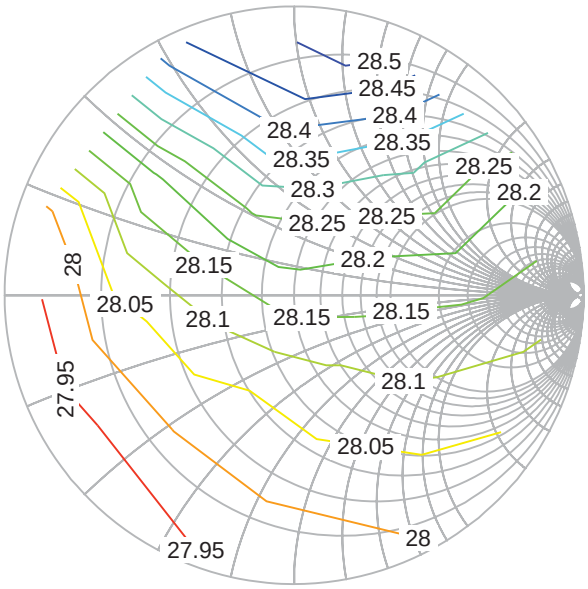


Figure 6-66, 10x75 device 9V, 150mA 2nd harmonic load pull output power contours.

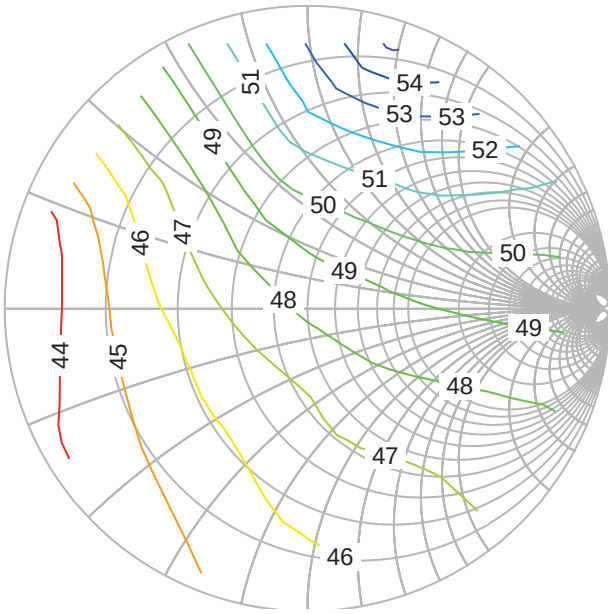


Figure 6-67, 10x75 device 9V, 150mA 2nd harmonic load pull PAE contours.

Measured input power for these contours was ~16.8dBm, which referring to Figure 6-63, suggests is slightly above the peak PAE level, which would account for a slight increase in output power and reduction in PAE from the expected levels.

We have noted from the measurement of the output matching circuit that the ideal output impedances have not been achieved for optimum performance. This can also be demonstrated by conducting a load-pull on the MMIC itself, an interesting additional application of the measurement system. Conducting such a load pull about the centre of the MMIC shows that the output matching has not been optimised into 50Ω and that increased efficiency (Figure 6-68) and power (Figure 6-69) can be obtained by further tuning.

Furthermore as the offset is not far from the 50Ω system impedance the amount of tuning is not great. Using the measured data from the device on the wafer and the output matching circuit a nonlinear simulation of the MMIC (based on the DLUT model) was created and load pulled. Although the results were pessimistic in terms of performance – because 2nd harmonic enhancement had not been included in the model (for the purposes of demonstration it was a fundamental model only) they replicate the load pull contours in shape and position, Figure 6-70.

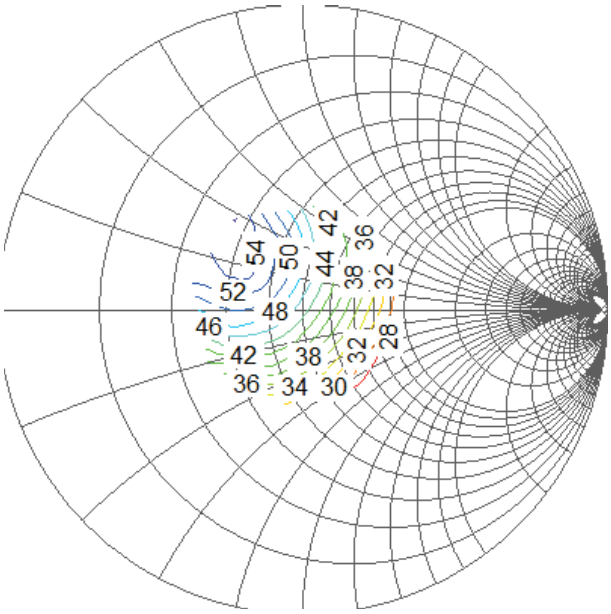


Figure 6-68, PAE load pull contours measured on MMIC at ~16dBm

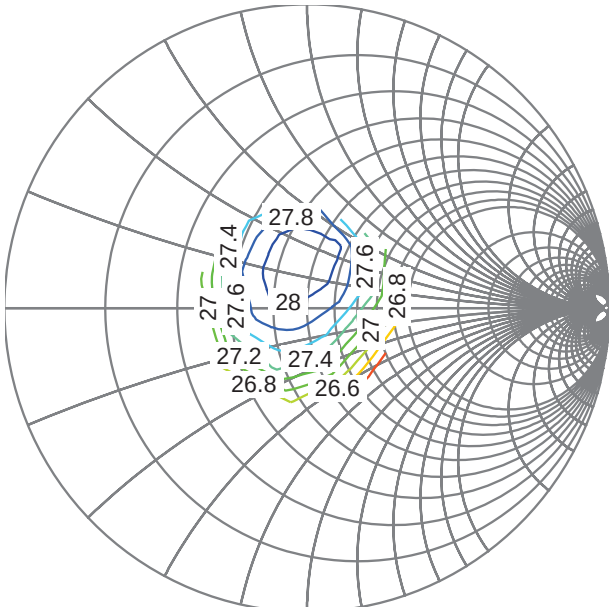


Figure 6-69, Output power load pull contours measured on MMIC at ~16dBm.

A simple tuning circuit that could be included 'off-chip' is shown inset to Figure 6-70 which presents the MMIC with the optimum PAE load impedance. A swept power simulation shows the relative improvement (remembering the offset due to the missing harmonic enhancement) that can be made. As this could potentially be incorporate in the interconnecting substrate it would be a simple and low cost modification (especially when compared to another design iteration). Obviously it would be necessary to assess the impact across the full frequency band.

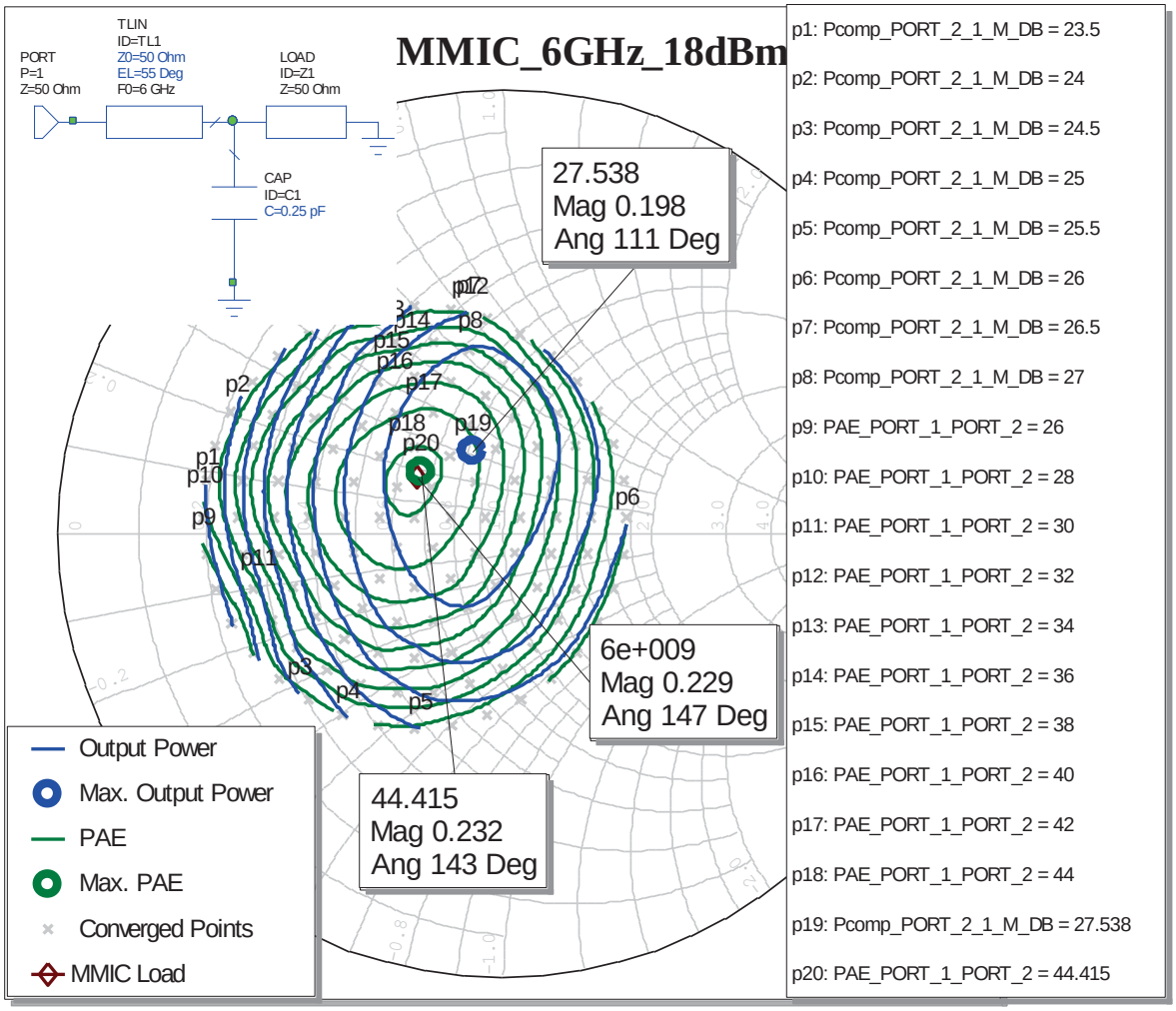


Figure 6-70, Simulated MMIC load pull contours based on DLUT device model and measured output matching circuit. Inset is a possible simple matching solution to 'pull' the MMIC to the optimum.

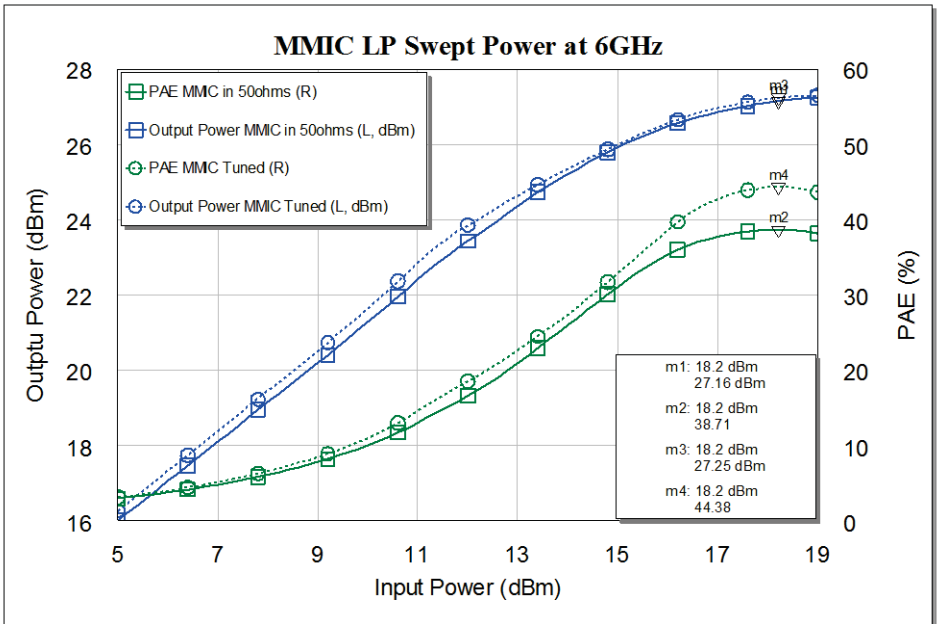


Figure 6-71, Comparison of swept input power performance on MMIC in 50Ω (solid) and with post chip tuning (dashed).

6.5 Summary of MMIC Design Process Implementation

The advantages of the novel design approach described in this chapter can be summarised as:

1. The ability to create, through a semi-autonomous process, nonlinear device models efficiently with a very high degree of accuracy.
2. The ability to construct, through waveform engineering specific current and voltage waveforms appertaining to particular theoretical high efficiency mode of operation.
3. The ability to understand why particular optimum conditions occur, especially when considering non-standard modes of operation (i.e. do not fit any of the 'classical' definitions), such as overdriven class A.
4. The measurement system gives designers the ability to explore 'What – if' scenarios; this is particularly important when investigating areas of operation near the device operating envelope. In particular such areas as device plane voltage swings which inform the designer of how close they getting to the process limits[5].
5. The effect of negative resistance impedance point, i.e. signal injection has been shown. This indicates the scale of performance improvements that may be achieved through harmonic injection, a technique that would not be limited in the way that greater than octave bandwidth amplifiers are by harmonics of lower frequencies falling in band.
6. It has been shown that the process also has a part to play in the evaluation and analysis of manufactured parts. This can help in the rapid development of designs. It has also been shown that the system can load pull the MMICs themselves which can both indicate if the optimum matching has been achieved and what needs to be done if not.
7. A key advantage is that this technique can be used with new processes before CAD device models have been developed to prove the viability of the new devices produced. This has clear 'time to market' advantages for foundries trying to recoup the massive capital investment in new processes.

[5] Indeed in this work, RF voltages in the region of 22 volts have been observed in constant use with no measureable harm to the device (gate current increase or output power decrease), on what is specified as a 12V process. There has been much debate on how DC voltage limits are related to RF voltage breakdown.

8. Finally and most importantly, the method and process used has proved to be accurate and implementable within industry standard CAD environments. This is essential for the acceptance of this technique by the industry; it must offer designers an improvement on existing methodologies.

There are still a number of issues that need to be addressed, that have become clearer as this work has progressed:

A. Sources need to be phase coherent. This has not been an issue for fundamental work as the primary signal within both source and load signal generators are the same. In effect the signal source is disconnected in one of the units (the slave) and the output circuitry is fed from the source of the other (master) unit. This does require that amplitude and phase adjustment are done in the output circuitry. The major problem has occurred due to the way that frequency multiplication is implemented within the unit, restricting frequency bands of operation. Where it was necessary to operate outside these limits then the oscillators had to be run from the same 10 MHz reference which resulted in phase drift with time, which, as the frequency increased, made load convergence more difficult; the amount of phase drift being more significant within the given measurement time. To attempt to overcome this, passive loads were used where possible, but this tended to be at the fundamental frequency where low reflection coefficients were required. Even here the passive load did not always replicate the optimum impedance exactly and in some instances 'muddied the waters' in comparisons with other data.
B. It appears that such systems will always struggle with the need for more drive power. In some instances the peaks in efficiency were tantalisingly close but not quite achieved. The expense of wide bandwidth power amplifiers puts them in great demand, especially when a number of measurement systems are on the go concurrently. It should be stated that an advantage of the approach is that the amplifiers are outside of the measurement couplers and so can be removed and swapped without 'upsetting' the calibration. It should also be noted that on a number of occasions it wasn't the system drive power that was the limiting factor but the way the measurement was configured. It was not always possible to access the measurement system for long periods of time and so a measurement would be

defined and set running and the results analysed afterwards. If it was found that a high enough power level had not be used, or the power step size was too large (to give adequate definition of an optimum for example), it was not always possible to repeat the measurement and still keep to the time scales dictated by wafer processing deadlines.

C. A persistent problem with the devices used is that they have high input reflection coefficients and errors in the measurement of Γ_{IN} have a significant impact on the measurement of gain and PAE. High reflection coefficients also obviously require higher drive levels (see B. above). An attempt to overcome this problem was tried where input matching was included on the manufactured circuit. Simple input matching proved very effective, however it was found that more complicated designs including gain equalisation and stability improvement made the analysis more difficult due to the increased uncertainties and number of variables. For a detailed accounting for performance differences it was much easier to compare devices on their own with those with output matching circuitry only. Input matching was successfully implemented which allowed a device to be driven into compression in the range 14-18GHz which would have been beyond the source amplifiers available at the time.

D. Finally the models used, DLUT, depend upon the consistency of the foundry processing. The version of the model used in this research does not take into account yield or tolerancing. It would be possible to introduce more variables into the model to refer, for example, to data set number, however this is not an ideal solution and does not directly relate to process variables. As mentioned the move to behavioural models offers a potential solution to this problem.

The consequence of poor/unreliable models is that multiple design iterations may need to be included which results in wasted wafer real estate, additional design time to produce the variants and test time to evaluate them when manufactured. Even when a variant is selected as being the most suitable to take forward (nearest fit to the design goals) there may still need to be further multiple iterations to ‘hone’ the design, assessing the variants has not improved the original nonlinear model (just shown it’s short comings). It may be possible to add external components to ‘tweak’ the model but there is rarely the

opportunity to do this 'internally' where it is most effective and device modelling is beyond the capabilities and time allowed for most engineers (if those expert in modelling can't get it right then what chance those who only work in the area occasionally?). The technique described in this work has been proved to accurately replicate the measured data under the conditions defined by the operator; ideally matched to those likely to be experienced in the application space. Indeed an advantage of the approach is that the parasitics associated by actual device mounting, such as the thermal implications of the device attachment method and the effects of bond wires can be incorporated into the model.

Although the approach described could be implemented with a fully passive tuner set-up, the active load pull has a number of advantages, a key one is the control of the harmonic impedances. Even where higher harmonics are not actively load pulled wide band 50Ω termination can be implemented. To do this in passive tuner based systems involves more loss and this is a limiting factor. Of particular importance is the need to present very high reflection coefficients at the harmonic frequencies, which is nearly impossible in a passive on wafer system due to the losses between the probe and tuner and becomes increasingly difficult at higher frequencies. In contrast active systems just require higher power amplifiers, and as the harmonic levels are typically 15-20dB below the fundamental the amplifiers needed are not excessively high powered.

By integrating the measurement into the design flow the design risk is significantly reduced, giving a greater degree of confidence in the performance of the circuits manufactured. Thus more wafer real estate can be devoted to other circuits. The goal of 'Right First Time' is still not guaranteed; as has been seen there is still a question as to the accuracy of the passive component models, (which can be addressed by more device measurements – more viable with increased wafer space available). The repeatability of device performance wafer to wafer is a problem that cannot be overcome by better modelling and is indicative that more effort needs to go into process control in the foundry. This highlights another advantage of our approach, effort and resource does not need to go into producing models at the early stages of the process development, thus allowing more attention to be paid to getting the most out of the process and bringing it on line as fast as possible (important considering the huge investment involved).

A final point that should be made about the accuracy of the DLUT modelling approach as described here is that the model is only as good as the measurements made

and this includes the range of impedances, voltages and drive levels encompassed. As has been shown the models interpolate between measurement points fairly well, however extrapolation invariably leads to wildly inaccurate results. The model size can be increased to accommodate the extra data so it is better to extend the measurement time and cover a wider range of the impedance plane (there usually needs to be sufficient drive power to reach the edges of the real impedance plane so this is not a limiting factor) and input power levels. The device bias should be set to that actually used as this is to the key to the operation of the RF load line.

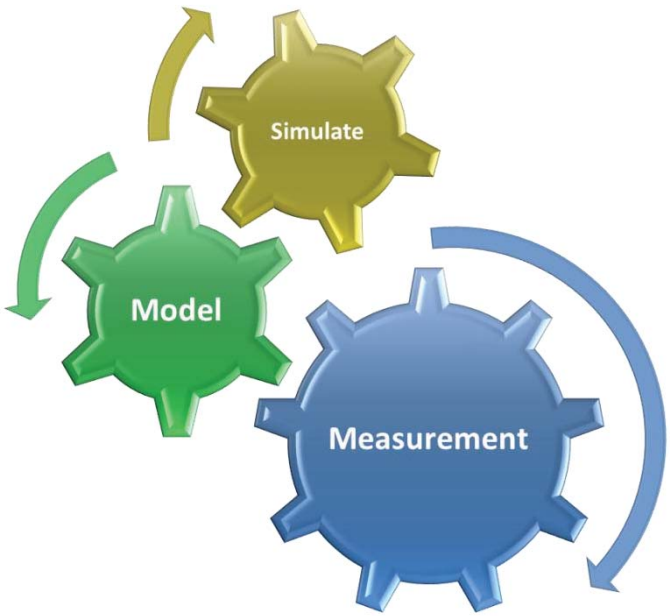


Figure 6-72, Waveform engineering based design process integrates measurement, modelling and CAD.

6.6 Bibliography

[1] J. Benedikt, T. P.J and e. al., "High power time domain measurement system with active harmonic load-pull for high efficieny base station amplifier design," *IEEE Transactions Microwave Theory and Techniques,* vol. 48, pp. 2617-2624, 2000.

[2] I. Bahl, "0.7 - 2.7GHz 12W Power Amplifier MMIC developed using MLP technology," *IEEE Transactions on Microwave Theory and Techniques,* vol. 55, pp. 222-229, February 2007.

[3] W. Bosch, J. Mayock, M. O'Keefe and J. McMonagle, “Low cost X band power amplifierMMIC fabricated on a 0.25um GaAs pHEMT process.,” in *Proceedings of the IEEE International Radar Conference*, 2005.

[4] S. Maas, “Why I hate base resistance,” *Microwave Magazine,* June 2004.

[5] R. Fano, “Theoretical Limitations on Broadband Matching of Arbitrary Impedances,” *Journal of Franklin Institute,* vol. 249, pp. 57-84 and 139-154, January-February 1950.

[6] P. McGovern, J. Benedikt, P. Tasker, J. Powell and e. al., “Analysis of DC-RF dispersion in AlGaN/GaN HFETs using pulsed I-V and time-domain waveform measurements,” in *International Microwave Sympsium*, 2005.

[7] C. Roff, P. McGovern, J. Benedikt, T. P.J, R. Balmer and et.al., “Detailed analysis of DC-RF dispersion in AlGaN/GaN HFETs using waveform measurements,” in *1st European Microwave Integrated Circuits Conference*, 2006.

[8] S. Cripps, RF Power Amplifiers for Wireless Communications, 2nd Ed., Artech House, 2006.

[9] A. AlMuhaisen, P. Wirght, J. Lees, P. Tasker, S. Cripps and J. Benedikt, “Novel wide band high-efficiency active harmonic injection power amplifier concept,” in *International Microwave Symposium*, Anahiem, 2010.

[10] S. Cripps, “A Method for the Prediction of Load-Pull Power Contours in GaAs MESFETs,” in *International Microwave Symposium*, 1983.

[11] R. Saini, S. Woodington, J. Lees, J. Benedikt and P. Tasker, “An intelligence driven active load pull system,” in *75th Microwave Measurements Conference (ARFTAG)*, 2010.

[12] R. Saini, J. Bell, T. Canning, S. Woodington, D. FitzPatrick, J. Lees, J. Benedikt and P. Tasker, “High speed non-linear device characterization and uniformity investigations at X-band frequencies exploiting behavioural models.,” in *77th Microwave Measurement Conference 2011 (ARFTAG)*, 2011.

7 Summary and Conclusion of Research into a Novel MMIC Design Process Using Waveform Engineering

"If I have seen further it is by standing on ye sholders of Giants"

Letter from Sir Isaac Newton to Robert Hooke, 5th February 1676.

7.1 Summary

Perhaps it is rather ironical that the quote above from Sir Isaac Newton was in fact a re-phrasing of quote attributed to Bernard of Chartres, a 12th century scholar, "we are like dwarves on the shoulders of giants, so that we can see more clearly than they, and things at a greater distance, not by virtue of any sharpness of sight on our part, or any physical distinction, but because we are carried high and raised up by their giant size", [1]. Going even further back in Greek mythology the blind giant Orion carried his servant Cedalion on his shoulders. We are all able to carry out our work in our specialist chosen fields because of the foundation and insights of what has been done before us. It is hoped that in some small way that this work too may contribute to the development of an understanding of the operation and behaviour of RF and microwave power amplifiers and by following the process described have a path to better design new ones.

The importance of harmonic terminations in improving the efficiency of solid state RF and microwave power amplifiers, even in class A, has long been understood [2] and useful approximations based on simplifying either the operation of the device or the nature of the load have been able to show how improvements can be made. As the boundaries of performance are pushed, in terms of bandwidth and efficiency more accuracy is however needed. To this end the measurement system described in this research has been shown to operate to a fundamental frequency of 18GHz with a harmonic frequency of 36GHz. Further, the method has been shown to be limited only by the constituent components, the power available from the load pull amplifiers, the bandwidth of the couplers, the power handling of the wafer probes, etc. In MMIC design this approach gives greater confidence that the

manufactured product will behave as predicted, i.e. the devices will have already been tested with the load impedances presented by the actual circuit. If the opportunity is taken, the mechanical mounting can also be replicated, and of course the bias and drive powers will be the same; thus the most significant variables will have been controlled. The results of the measurements are captured and available for interpretation within commercial nonlinear microwave design packages. The general aim of this project is not new, the problems with generic models have long been identified [3] but the typical approach has been to re-optimise a particular nonlinear model to a set of measured data. These measurements are not without problems, both in the need to conduct large and small signal measurements and the difficulty of embedding the measured data within the model; the data itself is lost - the model tries to approximate it. This approach has had success over bandwidths up to 30% but optimising the model over a greater frequency range leads to compromises in specific frequency behaviour. Despite these drawbacks it has been difficult to escape from the ease with which nonlinear simulations can be carried out using such models.

Although within the project timescales available and the file size of the early models, the model creation and implementation was limited to single frequency multi harmonic (up to 3rd) versions, recent developments in model creation [4], [5], greatly reduce not only the size of the data files but the number of points required to describe a load pull contour. These coupled with the increases in the speed of the measurement system [6], [7] makes the ability to map large impedance planes with their associated harmonic impedances possible.

The ultimate goal would be to integrate the measurement and CAD environments into one, such that after the main data collection, a first load approximation of the device output impedance and the assigned bias can be ‘tweaked’ “on-line” with the aide of the nonlinear analysis software, and this repeated with increasing complexity of the designed matching circuit and comprehensiveness of the model, until the ultimate performance is achieved. Not only does this remove the time lag between data acquisition and use of the standard approach, but also the separation of the designer from the device that they are working with - a sculptor would not fashion a statue without first spending time with and ‘understanding’ the marble they are to work with. Similarly the designer should get to know

the device; and yet how often in the past has the transistor only existed as a data sheet or a S parameter data file until the first prototype was built?

If this approach were adopted data would now be targeted in a more logical fashion, being acquired where it was needed and not just in a random, scalar fashion. This approach may be necessary at the start of the process, but as in a game of Battleships[1] once there is a hint of where the target is you hone in on a solution.

The attractions offered by MMIC amplifiers are numerous. In general they are smaller, lighter and have higher gain (due to less loss). Compared to discrete hybrid amplifiers they offer wider bandwidths and higher reliability (less connections and component parts), lower production parts cost, less unit to unit variation and better phase tracking; thus they are more repeatable. They are particularly suited to broadband power amplifiers due to the reduction of parasitics (through the integration of stages), of interconnects and discrete components. There are however many challenges, once built they are difficult to modify or tune, unlike hybrid amplifiers, where track widths can be adjusted and components values changed. Hence the manufactured design has to be right. The low PAEs traditionally achieved (<40% see chapter 2) means that there is relatively higher heat dissipation and hence lower gain and output power. A main objective is to reduce size and hence power density. The biggest drawback is the set-up cost, with foundry runs costing in excess of £50,000 a time (for GaAs, GaN is significantly higher and for smaller wafers). The unit cost is lower because of the large number of circuits than can be obtained from a wafer, but that is provided that they meet specification, repeatedly. Clearly the financial penalty for failure is extremely high. Contrast this with hybrid designs where a prototype design may cost only a few thousand pounds.

The decision on the optimum size of a MMIC is a trade-off between different parameters as illustrated in Figure 7-1. From the list of advantages stated above it would appear sensible to include as many parallel stages as possible on the MMIC; however the MMIC size cannot increase indefinitely. Thus there is a need to maximise the performance, particularly of the output stage, within the size available and with a properly centred design yield will also increase.

[1] Popular game, where naval warships are 'hidden' on a grid. Players take turns trying to locate and destroy them by calling out the co-ordinates.

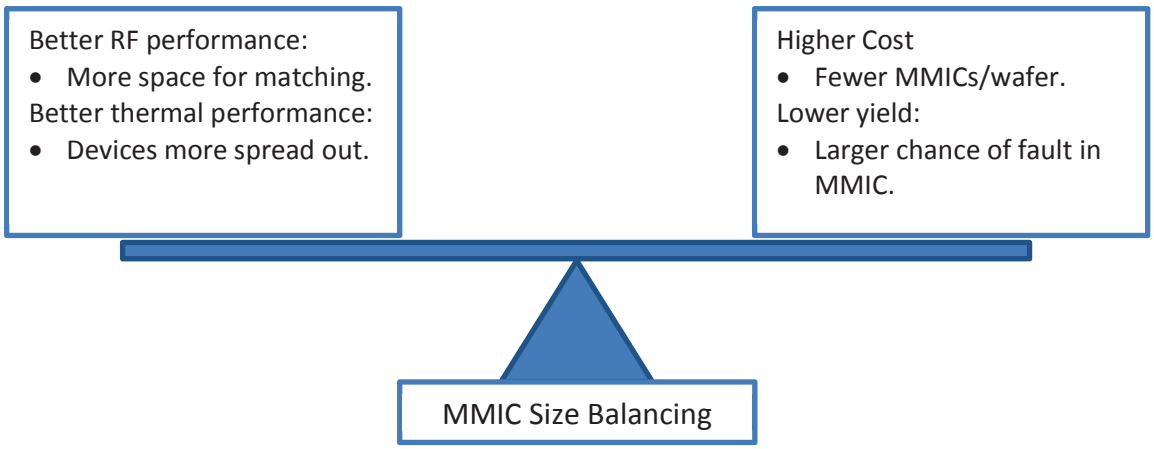


Figure 7-1, Some of the trade-offs involved in deciding on the MMIC physical size.

The results achieved in the first pass of the design technique proposed in this work produced a >42% PAE at >27dBm output power over an octave bandwidth (5-10GHz) which compares favourably with state of the art results reported for GaAs [8] (5W, 40% PAE 8-14GHz). Reviewing the papers referenced in the summary of at the end of chapter 2, key points raised are the critical importance in getting the output stage load impedance right and yet that the general approach is still largely based on linear models and the Cripps loadline technique [9]. It has been recognised that there is a need to integrate measurement and modelling including harmonic data [10], however this approach has so far been limited to optimising and models such as the EEHEMT1 to match measured load pull data. This therefore limits the optimisation to that indicated by the model. Of particular note in this regard is the measurements seen in chapter 4, where it was seen that the 3rd harmonic optimum did not lie on the perimeter of the Smith Chart, but at an optimum load within the real impedance plane. Encouragingly, the orders of magnitude of increase in PAE as a result of harmonic tuning agreed with those found in [10].

The need for better modelling is emphasised by Bahl, [8], one of the leaders in broadband amplifier design work, who highlighted the need for, “Accurate nonlinear models, predicting the output power and harmonics, are needed to improve PAE over multi octave bandwidths”. In this paper he also suggests that currently the most suitable method for assessing yield is to use 4 S parameter data sets, low gain, high gain, low current, and high current, and assess the circuit performance with these. This is not unlike the method proposed in this work to tackle the important design issue of yield; however this only gives the spread and not variation information. It has been proposed in this work that the

coefficient of the new behavioural models may be able to be allocated a distribution such that this replicates the actual yield seen in a MMIC process. To develop this will however require significantly more data.

There is perhaps an intractable argument between those who advocate the nonlinear model and those who defend the measurement based modelling approach, in fact some of those in the modelling camp do not recognise the DLUT model as a model at all. An imaginary argument between Tom the nonlinear modeller and Luke the DLUT using engineer might go something like this:

Tom: "Look how close my model is to the measured data".

Luke: "But the DLUT model almost exactly fits the measured data"[2].

Tom: "It is not a model, it is just looking up the measured data".

Luke: "It works in the simulator, it looks like a model, and it operates like a model, as far as I'm concerned…"

Tom: "But if you change the environment, the bias for example, it is no longer valid."

Luke: "If you change the environment then you need to re-measure to check your nonlinear model is still valid, add those measurements to the DLUT model and it is spot on again."

Tom: "But it depends upon the measured data".

Luke: "So does yours".

Tom: "No the nonlinear model will give you results without any measurement data".

Luke: "But you don't know if they're right unless you conduct a measurement to prove them".

And so the argument continued. Paul Dirac the Swiss/British fundamental theoretical physicist, author of the Dirac Equation describing the behaviour of fermions, and predictor of anti-matter, is famously quoted as saying "It is more important to have beauty in one's equations than to have them fit experiments". It is undeniable the insight that equations and theoretical models gave him into the operation of the universe, however the Holy Grail of the all-encompassing nonlinear model still appears to be a pipe-dream[3]. Especially as rather than the fixed universe that Dirac was tackling the structures of devices and the materials used are constantly changing.

[2] See chapter 5, nonlinear modelling.

[3] Victorian expression relating to the dreams experienced by those partaking of opium pipes.

The understanding of the impact of harmonics has developed from the early theoretical approach in 1967, [2] to an experimental approach in 1983, [11] to a CAD/nonlinear model based approach in 1988, [12]. Indeed the earliest reference to harmonic tuning that has been found in this research goes back to 1958, [13]. Early waveform analysis work [14], was based upon very idealised DC-IV curves, which as has been shown [15] and in chapter 4 are not valid at RF. In the early papers most transistors were connected directly to a tuned load and only the case of shorted even harmonics were considered. In practice the theory of device operation was waiting for a number of areas to catch up, the measurement capabilities to be able to accurately measure device impedances and waveforms (and hence power levels) to multiple harmonics, and the simulation side to have the 'number crunching' capabilities to develop the nonlinear models. As shown in this research we now have the ability to control the fundamental and harmonic load impedances, even so far as to extend outside of the real resistance plane, [16], to capture this data and analyse it within a nonlinear simulation environment. By de-embedding the output capacitance the voltage and current waveforms can be analysed at the current generator plane and the full operation of the device understood and if necessary improved. Reliability can be improved or safe operation confirmed, by looking at the peak voltage excursions. Correct phasing of the harmonic voltages can reduce the peak drain voltages to within safe operating limits. Equally the technique allows where these limits are to be explored, it has long been observed that the DC and RF limits are not the same.

Contour mapping of performance, giving engineers the ability to visualise the impact of impedance changes on device performance was in use before the end of the 60's, as shown in Figure 7-2. The "mapping" of the impedance plane was used not only to identify the areas of best performance and efficiency but also those areas of high collector dissipation so that they could be avoided (to improve reliability). This is still one of the best ways for engineers to steer their designs. Creating matching circuits to the 3rd harmonic over a bandwidth are impractical in terms of hitting each optimum impedance, however using the performance map derived from the measured data; show the regions to be aimed at and equally importantly those to be avoided. By viewing the contours of power and efficiency simultaneously the best trade-offs can be achieved.

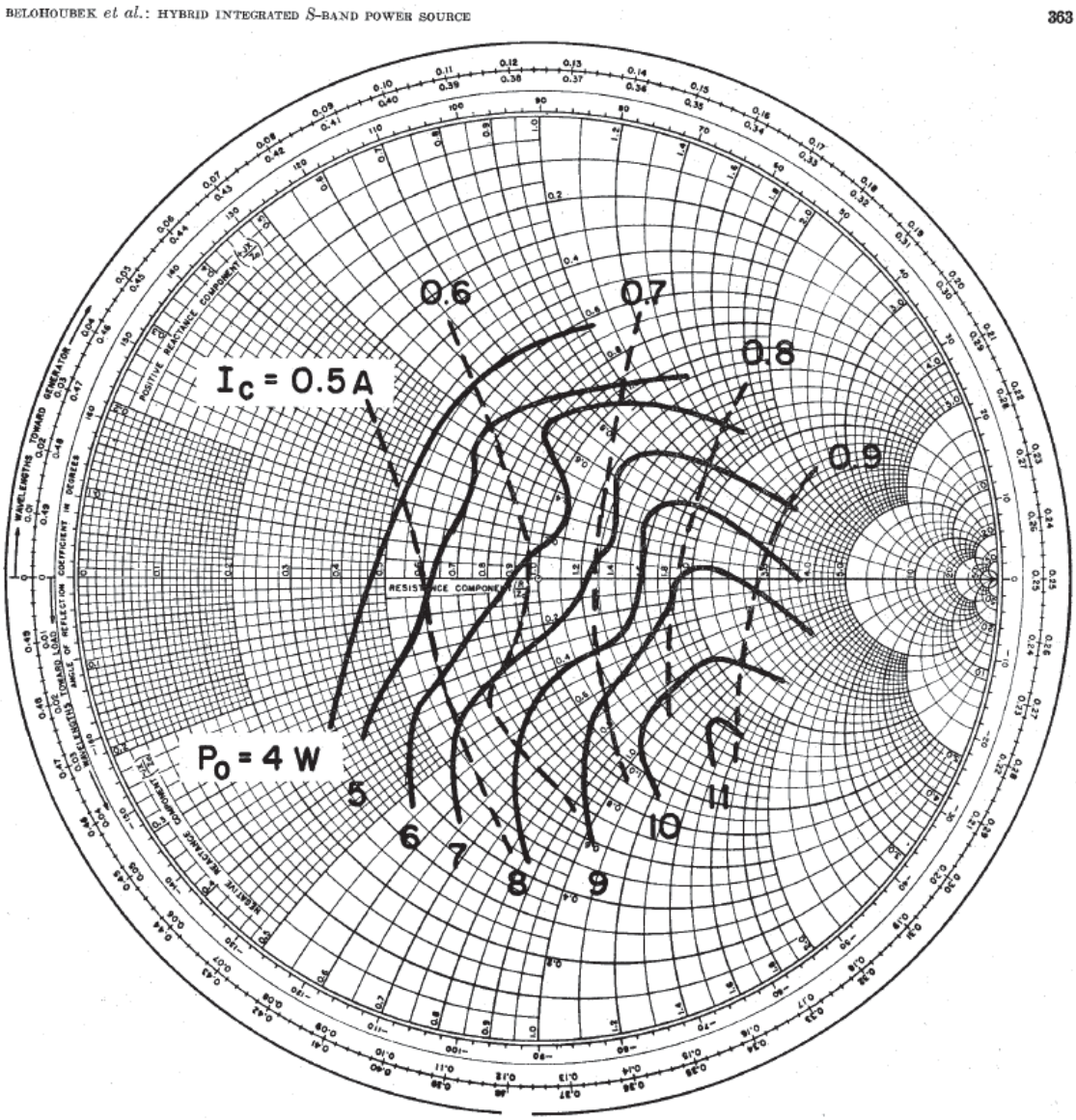


Fig. 7. Curves of constant output power and collector current on chart of impedance presented to the power amplifier at doubler input plane; F = 1600 MHz.

Figure 7-2, Contours of constant output power and collector current displayed on a Smith Chart from [22].

Of course it is not only harmonic optimisation that is important. Too often designers have been constrained at the initial design stage to specific bias conditions because those are the ones for which data is available, and sometimes it is questionable as to whether these were chosen for the convenience of the measurement system rather than their usefulness to the device application. As has been said design is about balancing competing needs, thus the optimum bias solution may lie between class A and B operating modes. But what is class AB? Is it 10% I_{DSS} or 25%? The slope of the G_m, transconductance curve, increases as the DC gate bias increases above pinch-off, the downside is the increasing DC drain current and hence decreasing efficiency. Surely the bias should be a parameter that

the designer can control, not be limited because of a possibly arbitrary decision by a measurement engineer, weeks, months or even years earlier.

The research conducted also highlighted a common problem in driver stage design. The standard approach is to use the interstage to flatten the gain by mismatching where the gain is highest and presenting a near conjugate match at the frequency where the gain is lowest. However as was shown in chapter 4 on Device Measurements, this can result in harmonic impedances being presented to the device with a detrimental performance on PAE. No indication has been found of the effects of this being taken into account in current design approaches. Interstage matching is typically a combination of dissipative and mismatch loss; by understanding the impacts of the harmonics the balance can between these two can be adjusted to maximise performance.

The issues surrounding cost and yield have led to innovative solutions with GaN MMICs. In part this is due to the immaturity of the process, which causes more variation in device performance and also as smaller wafers are being used the real-estate cost of GaN

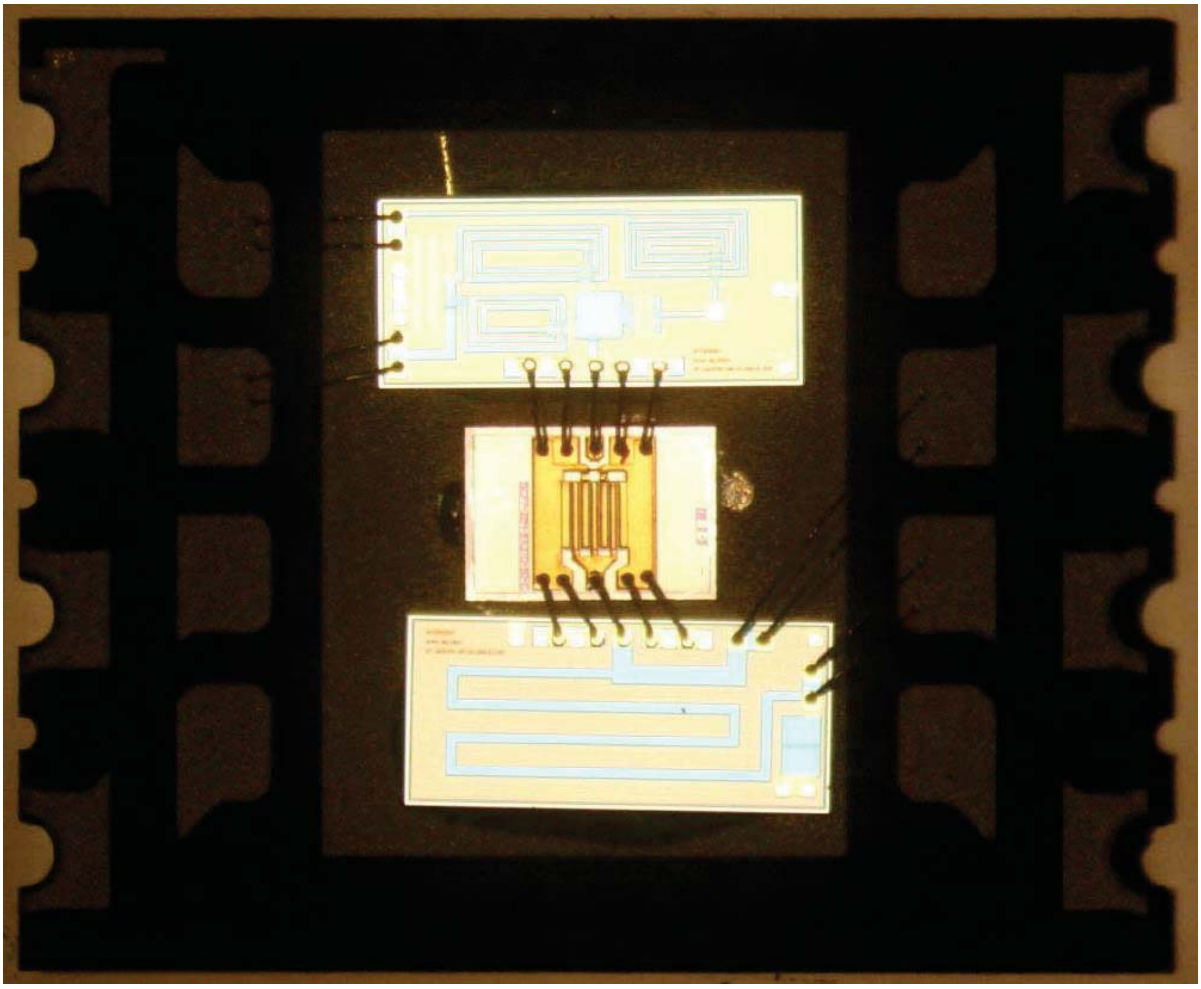


Figure 7-3, 8W GaN HEMT with GaAs input and output matching circuits in a SO8 AlN package.

circuits is very high. Although in part the example cited in chapter 2, [17], used off chip matching so that lower loss capacitors could be used, it is also true that this meant the chip size was smaller and more units could be produced per wafer and that a certain amount of tuning could be incorporated, for example by altering the bond wires (length or number) to account for device variation. It also made measuring each device much easier (output matching did not need to be de-embedded). In another paper by the same author, [18] he describes how GaAs circuits were used for the input and output matching for an 8W GaN device, Figure 7-3. The logic behind the choice of GaAs over for example Si is not clear, other than that RFMD are the world's large GaAs supplier for the RF industry. However the attraction of using an alternative substrate for the matching circuits is clear (until the cost of GaN on SiC comes down significantly and device repeatability improves). Perhaps a less obvious advantage may be that the passive circuit models would be well established for the more mature GaAs process and thus such solutions could be adopted whilst still in the early development stages of the new technology (before the circuit models for the passive structures have been constructed and tested). Coupled with the design process described in this work such prototype amplifiers could be being designed before nonlinear model creators have finished analysing the measurement data.

In industry, managers and project leaders want to know what can be really achieved in practice, not what the device supplier's claim on their datasheets. Experience has those responsible for delivering projects on time and on budget to be wary of such claims. The design engineer's answer is for them to be allowed to build a circuit to find out; there are too many variables and uncertainties to be able to know the performance of a new device or process with any degree of uncertainty. It's not what we know that limits us; it's what we don't know! To commit to the cost and time of designing a MMIC to find out whether or not a particular process will meet the system specifications is expensive and risky, and if there is more than one possible supplier, is the project supposed to support 2 or 3 trials to decide between them? What if the first run fails; do you have another go or discount possibly the best choice? The integrated measurement, modelling and CAD process described in this works allows a cost effective and thorough evaluation of device and process performance to be conducted directly, giving both engineers and management confidence in proceeding with a particular development path.

7.2 Conclusion

The design process described in this research is based upon a tight integration between the measurement, modelling and simulation technologies which enables devices to be tested in a manner that accurately replicates the operating environment for which they are intended. Designers are no longer tied to 'known' devices or operating conditions, but are free to adopt the conditions which produce the optimum performance for their requirement. At a very early stage in the development process confidence is gained in a particular process or device, radical strategies can be explored and their consequences shown. The project is de-risked and target specifications can be firmed up into requirement specifications with confidence. Although never risk free (a problem can occur during wafer processing) there is a much greater likelihood that the manufactured MMIC will meet speciation and hence there can be reduced time in the development program due to reiterations of the design.

7.3 Further Work

More work is required in a number of areas, primarily in the integration of the measurement data directly into a model that can be read by the simulation software. In this way all of the data display and manipulation capabilities of the commercial programs can be utilised avoiding duplication and benefiting from a standardised approach. This will also allow any problems with the data to be identified immediately and therefore prompt action to be taken to correct them. This will also improve the longevity of the data. During the research the fundamental data set format changed at least 3 times, which meant amongst other things that new data viewer software had to be written for each version. Inevitably display or search routines which were in one version were not available in others.

There are a number of areas where adding intelligence to the automatic load pull control will improve the quality of the data and assist in improving accuracy. When comparing performance at different frequencies it was found that data did not always exist for the same amounts of compression as the step size in the drive power sweep was too large. This included a number of occasions when the peak in the PAE had not quite been reached. The nature of device compression is such that less data is required in the linear region and more as the gain characteristic is changing most rapidly in compression. In the

current control software the power sweep is determined before the measurement, the start and stop power and the step increment. It should be possible instead to request a start power and for the system to step intelligently, for example increase the power by 1dB, if the output also increases by 1dB then continue with 1dB steps. When the relationship is no longer linear the step size should be decreased (including back tracking one step) so that more detail is acquired in this region. The changes in step size should be adjusted according to the rate of change of gain. Displaying the final step size would give an indication of whether the device had successfully been driven into compression. In some cases for example a device in class B the gain will not be linear and a finer measurement resolution will be required across the power sweep than for a device in class A.

Work is already on going on improving both the measurement speed and reducing the number of points required to describe a load pull contour, the associated reduction in data file size will be a significant advantage, particularly as multiple frequency point and harmonic data files are created. Describing contours in the impedance plane will also extend the model validity beyond the measurement plane, although with the obvious caveats on accuracy, but unlike the current implementation where the model behaviour breaks down as soon as the edge of the envelope is reached.

This raises another point which is applicable to all models. It would be very useful to designers to know where the model measurement envelope boundaries lie. For the standard nonlinear model this is much less clear, particularly as they are often scaled up from smaller devices. There may not be validity data for a particular device size or bias. The DLUT and Cardiff behavioural model on the other hand has the measurement information within it, thus it should be possible to display the measurement impedance points, bias and power range information. The measurement grid points could be displayed on a Smith Chart, whilst the measured bias ranges on the schematic.

The purpose of this work was not to advocate a particular Automatic Load Pull measurement system, although obviously the benefits of that used have been highlighted. A number of commercial companies are now offering solutions in varying degrees of similarity to the system used, such as Mesuro (www.mesuro.com), NMDG (www.nmdg.be) and Maury (www.maurymw.co). For the frequency ranges under consideration in this work passive tuner based systems are not applicable as they cannot overcome the increasing system losses with frequency and thereby achieve high reflection coefficients. Arguably a passive

fundamental tuner on the input would be an advantage as this would decrease the power requirements of the input driver. For the devices measured in this work the ~0.5W drive levels available were just about sufficient however with a move to higher power devices the cost of the additional input power would probably justify the investment in the tuner. From the measurements taken it is clear that any system should have a bandwidth that includes the 2nd harmonic as a minimum. For designing narrower band higher efficiency amplifiers the 3rd harmonic will also be required for the ultimate performance, e.g. for X band ≥10W MMICs. Although using a similar approach [10], achieved PAEs >7% above similar amplifiers (as listed in [19]), its performance still fell short of the 60% PAE predicted from the simulation. Based on an input power of 21dBm and taking the nominal values for I_D, Pout and PAE from the graphs plotted the performance at the output of the MMIC and at the output the final stages (taken as a sum total and assuming an output circuit loss of 0.7dB as predicted) the PAE at the stage output is ~48.8% as shown in Table 7-1. This falls some way short of the 60% predicted from the model. There are a number of possibilities for this which won't be explored here, but as shown in the final example in chapter 6 and described in [20]; the measurement system can also be used to understand why performance is different to that expected. It should (as was shown in chapter 6 and [20]) be possible to demonstrate consistency between the device behaviour, output circuit losses and MMIC performance.

MMIC Output				O/P Circuit Loss	Device Plane	
Pin	Pout	I_D	PAE		Pout	PAE
(dBm)	(dBm)	(A)	(%)	(dB)	(dBm)	(%)
21	40.2	3.1	41.5	0.7	40.9	48.8

Table 7-1, Average performance at 10GHz taken from [10], and calculated performance at the output of the devices based on the predicted insertion loss from reference.

7.4 **Bibliography**

[1] J. Salisbury, “Bishop of Chartes,” in *The Metalogicon of John Salisbury*, University of California Press, 12th Century, p. 167.

[2] D. Snider, “A theoretical analysis and experimental confirmation of the optimally loaded and overdriven RF power amplifier,” *IEEE Transactions on Electron Devices,* vol. 14, no. 12, pp. 851-857, 1967.

[3] F. van den Bogaart and A. de Hek, “First-Pass design strategy for high power amplifiers at X band,” in *IEE Tutorial Colloquim on Design of RFICs and MMICs*, 1997.

[4] S. Woodington, T. Willimas, H. Qi, D. Williams, L. Pattison, A. Patterson, J. Lees, J. Benedikt and P. Tasker, “A novel measurement based method enabling rapid extraction of a RF waveform look-up table based behavioural model,” in *International Microwave Symposium Digest IEEE MTT-S*, 2008.

[5] S. Woodington, R. Saini, D. Williams, J. Lees, J. Benedikt and P. Tasker, “Behavioural model analysis of active harmonic load-pull measurements,” in *International Microwave Symposium Digest IEEE MTT-S*, 2010.

[6] R. Saini, S. Woodington, J. Lees, J. Benedikt and P. Tasker, “An intelligence driven active load pull system,” in *75th Microwave Measurements Conference (ARFTAG)*, 2010.

[7] R. Saini, J. Bell, T. Canning, S. Woodington, D. FitzPatrick, J. Lees, J. Benedikt and P. Tasker, “High speed non-linear device characterization and uniformity investigations at X-band frequencies exploiting behavioural models.,” in *77th Microwave Measurement Conference 2011 (ARFTAG)*, 2011.

[8] I. Bahl, “0.7 - 2.7GHz 12W Power Amplifier MMIC developed using MLP technology,” *IEEE Transactions on Microwave Theory and Techniques,* vol. 55, pp. 222-229, February 2007.

[9] S. Cripps, “A Method for the Prediction of Load-Pull Power Contours in GaAs MESFETs,” in *International Microwave Symposium*, 1983.

[10] G. van der Bent, A. de Hek, A. Bessemoulin and F. van Vliet, “Low cost high efficient 10 watt X band high power amplifier,” in *IEEE International conference on Microwaves, Communications, Antennas and Electronics Systems*, 2009.

[11] K. Chiba and N. Kanmuri, “GaAs FET power amplifier module with high efficiency,” *Electronics Letters,* vol. 19, no. 24, pp. 1025-1026, 1983.

[12] B. Kopp and D. Heston, “High-efficiency 5-watt power amplifier with harmonic tuning,” in *IEEE MTT-S International Microwave Symposium Digest*, 1988.

[13] V. Tyler, “A new high efficiency high power amplifier,” *Marconi Review,* vol. 21, pp. 96-109, 1958.

[14] L. Kushner, “Output performance of idealized microwave power amplifiers,” *Microwave Journal,* pp. 103-116, October 1989.

[15] P. Ladbrooke, N. Goddship, J. Bridge and D. Battison, “Dynamic I(V) measurement of contemporary semi-conductor devices,” *Microwave Engineering Europe,* pp. 23-33, May 2000.

[16] A. AlMuhaisen, P. Wirght, J. Lees, P. Tasker, S. Cripps and J. Benedikt, “Novel wide band high-efficiency active harmonic injection power amplifier concept,” in *International Microwave Symposium*, Anahiem, 2010.

[17] K. Krishnamurthy, D. Green, R. Vetury, M. Poulton and J. Martin, “0.5-2.5GHz, 10W MMIC power amplifier in GaN HEMT technology,” in *IEEE Compound Semiconductor Integrated Circuits Symposium Digest*, 2009.

[18] K. Krishnamurthy, D. Wang, B. Landberg and M. J., “RLC matched GaN HEMT Power Amplifier with 2 GHz bandwidth,” in *Compound Semiconductor Integrated Circuits Symposium*, 2008.

[19] C.-K. Chu, H.-K. Huang, H.-Z. Liu and et.al., “A 9.1-10.7 GHz 10-W, 40-dB gain four-stage pHEMT MMIC power amplifier,” *IEEE Microwave and Wireless Components Letters,* vol. 17, no. 2, pp. 151-153, 2007.

[20] D. FitzPatrick, R. Saini, J. Lees, J. Benedikt, S. Cripps and P. Tasker, “A Waveform Engineering Approach to the Design of Improved Efficiency Wideband MMIC Amplifiers,” in *WAMICON 2011*, Clearwater, 2011.

[21] K. Krishnamurthy, D. Lieu, R. Vetury and J. Martin, “A 0.1-1.8 GHz, 100W GaN HEMT Power Amplifier Module,” in *Compound Semiconductor Integrated Circuit Symposium*, 2010.

[22] E. Belohoubek, A. Rosen, D. Stevenson and A. Presser, “Hybrid Integrated 10-Watt CW

broadband power source at S band," *IEEE Journal of Solid-State Circuits,* Vols. SC-4, no. 6, pp. 360-366, 1968.