Modular Nonlinear Characterization System and Large-Signal Behavioral Modelling of Unmatched Transistors for Streamlined Power Amplifier Design

by

Dylan Bespalko

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

Waterloo, Ontario, Canada, 2016

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

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

Abstract

This thesis provides a comprehensive approach to the characterization and modelling of large-signal nonlinear RF/microwave devices, circuits and systems. This research is motivated by the increased linearity and power-efficiency requirements of modern power amplifier technology for wireless communications. For instance, maximizing the power amplifier's efficiency can only be achieved by operating RF transistors under strong nonlinear conditions, however this is contradictory to maximizing PA linearity. Simultaneously designing for efficiency and linearity is a challenging trade-off in today's fragmented design process, therefore the advancement of computer-aided design (CAD) tools is essential for achieving an optimal solution. The successful and effective CAD tool based PA design relies on the availability of accurate nonlinear models to mimic the electro-thermal behaviour of RF transistors. The accuracy of these models depends on three factors:

1. The formulation of the model.
2. The model extraction procedure.
3. The accuracy of the measurement data.

While prior work focuses separately on the improved model formulations or improving characterization accuracy, this thesis provides a comprehensive analysis of all three factors. This thesis proposes a modular large-signal RF device characterization system, and a nonlinear behavioral model capable of handling strongly nonlinear unmatched RF transistors, each necessary to streamline the design process and achieve a first-pass PA design.

As a first step, a large-signal characterization system has been developed to measure the multi-harmonic frequency response of RF transistors and has the ability to i) Perform high-power measurements, ii) Characterize unmatched transistors, iii) Operate the DUT under any possible operating condition, iv) Synthesize any multi-harmonic stimulus, and v) Reconstruct the time-domain I/V waveforms at the ports of the DUT. The proposed characterization system eliminates fragmentation between measurement and simulation environments by providing seamless integration with Harmonic Balance simulations. This provides a common framework that integrates all steps of the PA design process from device-level characterization, to circuit-level measurement and validation. This system is implemented using modular instruments consisting of mixer-based receivers, arbitrary waveform generators, impedance tuners, and a multi-harmonic phase-coherent reference source. It also integrates sequential calibration routines to provide receiver, port match, and source-power corrections to the DUT measurement plane and measurement routines for automated data collection.

The second part of the thesis researches black-box frequency-domain behavioral models that can approximate strongly nonlinear, unmatched devices. Our investigation yielded two complimentary solutions to ensure the targeted modelling accuracy. First, improving the accuracy of a first-order expansion-based Poly-Harmonic Distortion (PHD) model by $5dB$, in terms of Normalized Mean-Squared Error (NMSE), by minimizing multi-harmonic reflections that artificially increase the order of the nonlinear system. While this addresses the fictitious need for higher-order models due to the deficiencies in the model extraction procedure, strongly nonlinear devices will require high-order models to achieve the targeted accuracy over a larger measurement distribution. Hence, a variable order Multi-Harmonic Volterra (MHV) model is proposed to extend the PHD model formulation to strong nonlinear devices. This model is extracted by utilizing the proposed characterization system to extract higher-order multi-variate model coefficients not included in the PHD model. The resulting model improves DC drain current prediction by $5dB$ and improves fundamental output-power prediction by $2dB$. The MHV model improves the vector power-gain prediction by $3.4dB$ in realistic PA design applications, thereby providing better emulation of linearization techniques within a simulation environment.

Finally, a concurrent dual-band PA design is studied as an example of how the pro-

posed nonlinear characterization system and behavioural modelling approach can be used to enable complex PA designs. First, a $10W$ Class-AB PA is designed using dual-band matching-network theory, however it is difficult to implement because the design technique does not control the matching fractional bandwidth as a design parameter. Therefore, an alternative Class-J $45W$ dual-band PA was designed using a low-impedance matching network, combined with a trans-impedance dual-band filter. Although the dual-band PA can achieve comparable performance to an equivalent single-band PA at each separate frequency, further development of characterization, modeling, and circuit design techniques is needed to achieve high-efficiency during concurrent operation.

Acknowledgements

It's hard to reflect on my achievements as my own doing, because it's hard to imagine past generations of my family having the same opportunity. My grandparents were molded by their ability to survive harsh winters and the great depression, relying solely on their own abilities to solve problems. By my age, my Grandfather had built a house twice to pursue his career as an electrician, while supporting a wife and two kids. On my Mother's side, my Uncles left home early to become welders in order to support the family. My Father was the first in the family to receive a college education. He worked long hours, many weekends, and rarely took vacation time so that I would have the means to pursue a vocation of my choosing. Therefore, it was my Mother that guided me through each milestone of my adolescence, helping me with my homework, science fair projects, and ensuring my advancement to the highest tier of education. She raised three boys to become a medical doctor, a lawyer, and a stock broker – unfortunately they all fell off course during their awkward teenage years and pursued Ph.D degrees in Engineering. Sorry Mom! I couldn't possibly compare the problems that I have faced to the problems my predecessors solved in order for me to be here.

While pursuing a graduate degree it's easy to get lost in the details of your work. I've missed every one of my friends' bachelor parties and weddings. I'm thankful for the friends that have accepted me in my dormant state and I hope that we can make up for the lost times. I am very fortunate to have a cherished friend in Scott Branton since Junior High School, as he has been a great source of competition and enduring friendship.

I would like to thank my professor, Dr. Slim Boumaiza, for introducing me to this industry through his tireless mentorship, ambitious research facilities and determination

to support meaningful research. He has instilled a determination in me to be the best at what I do and use that knowledge to solve meaningful problems.

Finally, I would like to thank Radhika Phatak for putting me back together when I fell apart. One night, I remember talking to you on the phone while the sun was rising; I couldn't understand how such a beautiful woman, with a vastly superior intellect and personality, could find me that interesting. You brightened each day and kept me going.

Dedication

For Radhika

Table of Contents

List of Tables

List of Figures

Nomenclature

ABCD 2-Port ABCD Network Parameters

AM Amplitude Modulation

ARF Multi-Port LSOP RF Incident Power-Wave

AVS Available Source Power

AWG Arbitrary Waveform Generator

$\mathbf{A_p}$ Large-Signal Incident Power-Wave

$\mathbf{A_{RF}}$ Incident Travelling-Wave at the Calibration Plane

$\mathbf{A_{jl}}$ Incident Travelling-Wave at port j (and harmonic l)

$\mathbf{B_p}$ Large-Signal Reflected Power-Wave

$\mathbf{B_{ik}}$ Reflected Travelling-Wave at Port i (and Harmonic k)

CAD Computer Aided Design

COM Commmunication Object Model

CW Continuous-Wave

$\mathbf{C_P}$ Parallel Capacitor

$\mathbf{C_S}$ Series Capacitor

$\mathbf{C_{ds}}$ Drain-Source Capacitance

$\mathbf{C_{gd}}$, Gate-Drain Capacitance

$\mathbf{C_{gs}}$, Gate-Source Capacitance

DAC Digital-to-Analog Converter

DPD Digital Pre-Distortion

DUT Device-Under-Test

ET Envelope-Tracking

FET Field Effect Transistor

GSOLT General Short Opend Load Thru Calibration

GUI Graphical User Interface

$\mathbf{G_m}$ Large-Signal Trans-conductance

$\mathbf{G_{SRC}}$ Source Vector Calibration Gain Coefficent

$\mathbf{G_{ik,nr}}$ A two-dimensional $n^{th} \times r^{th}$ degree Fourier Series that models the response on port i, harmonic k.

IF Intermediate Frequency

IFFT Inverse Fast Fourier Transform

IVI Interchangeable Instruments Virtual

$\mathbf{I_D}$ DC Drain Current

K Impedance Inverter Factor

$\mathbf{K_l}$ Absolute Vector Calibration Factor at harmonic l

LO Local Oscillator

LPE Low-Pass Equivalent Filter

LRM Line Reflect Match calibration

LSE Least Squared Error

LSNA Large-Signal Network Analyzer

LSOP Large-Signal Operating Point

LTE Long-Term Evolution

LUT Look-Up Table

$\mathbf{L_P}$ Parallel Inductor

$\mathbf{L_S}$ Series Inductor

MEMS Microelectromechanical Systems

MHV Multi-Harmonic Volterra Series

MIMO Multiple-Input, Multiple-Output System

MN Matching Network

MTA Microwave Transition Analyzer

MaxInputs Maximum Kernel Inputs used during MHV model extraction

MaxOrder Maximum Kernel Order used during MHV model extraction

N Number of Measurements in a sweep plan

NMSE Normalized Mean Squared Error (at harmonic k)

NVNA Nonlinear Vector Network Analyzer

NumPorts Number of Ports of MHV model extraction

PAE Power-Added Efficiency

PAPR Peak-to-Average Power Ratio

PHD Poly-Harmonic Distortion model

PM Phase Modulation

$\mathbf{P^k}$ Phase of A_{11} providing phase normalization at harmonic k

$\mathbf{P_L}$ Delivered Load-Power

$\mathbf{P_{AVS}}$ Available Source-Power

$\mathbf{P_{IN}}$ Delivered Input-Power

QR QR Matrix Decomposition

QSOLT Quick Short Open Load Thru calibration

RF Radio Frequency

RMSE Root Mean Squared Error

$\mathbf{R_{ik,n}}$ An n^{th} degree Fourier Series that models the response on port i, harmonic k.

SISO Single-Input, Single-Output System

SRF Self-Resonance Frequency

SS Small-Signal

SSE Sum of Squared Error

SVD Singular Value Matrix Decomposition

TRL Thru, Reflect, Line Calibration

$\mathbf{T_A}$ Amplitude Reference Transmission Coefficient

$\mathbf{T_F}$ Transmission Parameters of the De-Embedded Fixture

$\mathbf{T_{RF}}$ Transmission Parameters of the Receiver Error Matrix

$\mathbf{T_{p,F}}$ Generalized Transmission Parameters of the De-Embedded Fixture

VISA Virtual Instrument Software Architecture

VNA Vector Network Analyzer

VSA Vector Signal Analayzer

VSG Vector Signal Generator

$\mathbf{V_{DS}}$ Drain-Source Voltage

$\mathbf{V_{GS}}$ Gate-Source Voltage

$\mathbf{V_{i0,1}}$ The Multi-Harmonic Volterra kernel that predicts the response on port i, harmonic 0

$\mathbf{V_{ik,m}}$ The m^th Multi-Harmonic Volterra kernel that predicts the response on port i, harmonic k

$\mathbf{V_{j0}}$ DC voltage on port j

$\mathbf{X^F_{ik}}$ The X-Parameter model LSOP kernel that determines $b_{ik} = f\left(A_{11}\right)$

$\mathbf{X^S_{ik,jl}}$ The X-Parameter model kernel that determines $b_{ik} = f\left(a_{jl}\right)$

$\mathbf{X^T_{ik,jl}}$ The X-Parameter model kernel that determines $b_{ik} = f\left(a^*_{jl}\right)$

$\mathbf{Y_O}$ Characteristic Admittance of Open-Circuit Transmission Stub

$\mathbf{Y_S}$ Characteristic Admittance of Short-Circuit Transmission Stub

ZRF Multi-Port, Multi-Harmonic impedance during MHV extraction

$\mathbf{Z_0}$ System Characteristic Impedance

$\mathbf{Z_a}$ Characteristic Impedance of Impedance Inverter Transmission Line

$\mathbf{Z_b}$ Characteristic Impedance of Impedance Inverter Transmission Stub

$\mathbf{Z_{IN}}$ Input Impedance

$\mathbf{Z_{jl}}$ Impedance on port j at harmonic l

$\mathbf{\Delta_f}$ The fractional frequency separation of Dual-Band design frequencies f_1 and f_2

$\mathbf{\Gamma_L}$ Load Reflection Coefficient

$\mathbf{\Gamma_P}$ Phase Reference Reflection Coefficient

$\mathbf{\Gamma_{jl}}$ Reflection Coefficient on port j at harmonic l

$\hat{\mathbf{x}}$ Predicted Value of the LSE Input

$\hat{\mathbf{y}}$ Predicted Value of the LSE Output

κ Condition Number of the Problem Matrix

$\mu_{\mathbf{Source}}$ Maximum Magnitude of Source Reflection Coefficient before Potential Instability

$\sigma_{\mathbf{min}}$ Minimimum Singular Value

$\theta_{\mathbf{a}}$ Electical Delay of Impedance Inverter Transmission Line

$\theta_{\mathbf{b}}$ Electical Delay of Impedance Inverter Transmission Stub

aHarmonics Number of Incident Power-Waves used to extract MHV model

aRFStep Maximum Small-Signal Variation used to extract MHV model

$\mathbf{a_P}$ Phase Reference Incident Travelling Wave

$\mathbf{a_p}$ Small-Signal Incident Power-Wave

$\mathbf{a_{RF}}$ Incident Travelling-Wave at the Receiver Reference Plane (on port j)

$\mathbf{a_{jF}}$ Incident Travelling-Wave at the Calibration Reference Plane (on port j)

$\mathbf{a_{jl}}$ Small-Signal Incident-Wave on port j, harmonic l

$\mathbf{a_{p,F}}$ Incident Power-Wave at the Calibration Reference Plane

$\mathbf{a_{p,SRC}}$ Incident Power-Wave at the Signal-Source Reference Plane

$\mathbf{a_{pjl}}$ Small-Signal Incident Power-Wave on port j, at harmonic l

bHarmonics Number of Reflected Power-Waves predicted by the MHV model

$\mathbf{b_A}$ Amplitude Reference Reflected Travelling Wave

$\mathbf{b_p}$ Small-Signal Reflected Power-Wave

$\mathbf{b_{RF}}$ Reflected Travelling-Wave at the Receiver Reference Plane (on port i)

$\mathbf{b_{ik}}$ Small-Signal Reflected-Wave on port i, harmonic k

$\mathbf{b_{jF}}$ Reflected Travelling-Wave at the Calibration Reference Plane (on port i)

$\mathbf{b_{pik}}$ Small-Signal Reflected Power-Wave on port i, at harmonic k

$\mathbf{error_k}$ Percent Error at harmonic k

$\mathbf{f_0}$ Fundamental Frequency

$\mathbf{f_1}$ Low-Band Fundamental Frequency

$\mathbf{f_2}$ High-Band Fundamental Frequency

$\mathbf{f_s}$ Sample Frequency

$\mathbf{f_t}$ Threshold Frequency

$\mathbf{f_{IF}}$ Intermediate Frequency

$\mathbf{f_{LO}}$ Local Oscillator Frequency

$\mathbf{f_{RF}}$ Radio Frequency

$\mathbf{fund_j}$ Fundamental Frequency on port j

i Reflected-Wave Port Index

$\mathbf{i_d}$ Drain-Source Small-Signal Current

j Incident-Wave Port Index

k Reflected-Wave Harmonic Index

l Incident-Wave Harmonic Index

$\mathbf{l_O}$ Electical Length of Open-Circuit Transmission Stub

$\mathbf{l_S}$ Electical Length of Short-Circuit Transmission Stub

n Order of a Nonliner Problem

r Order of a Nonliner Problem

ssfreq Small-Signal Input Frequency Sweep

ssphase Small-Signal Input Phase Sweep

ssport Small-Signal Input Port Sweep

tstart Start Time

tstop Stop Time

$\mathbf{v_{DS}}$ Drain-Source Small-Signal Voltage

$\mathbf{v_{GS}}$ Gate-Source Small-Signal Voltage

x Actual Value of the LSE Input

y Actual Value of the LSE Output

Chapter 1

Introduction

The recent trend in wireless communications towards a continuous increase in data throughput has motivated the deployment of advanced modulation and access technologies to maximize spectrum efficiency. However, the characteristics of the resulting signals brought stringent design requirements on the radio system on both sides, namely mobile and base-station infrastructure. This was particularly the case for the most expensive and power-hungry building block in wireless radio systems, the power amplifier (PA). In fact, PAs for wireless communications must operate linearly to meet signal quality requirements (Error Vector Magnitude (EVM) and Adjacent Channel Power Ratio (ACPR)), while providing competitive power-efficiency performance to minimize operating costs across a large network.

Trends in wireless infrastructure towards high peak-to-average power (PAPR) communications signals require advanced PA circuit topologies to maximize average power-efficiency rather than peak efficiency. Whereas, Class-F and Class-E PAs maximize peak efficiency, they do so at the detriment of average efficiency, linearity, and RF bandwidth. Hence, average efficiency enhancement techniques, such as Doherty, Envelope Tracking (ET), Linear Amplification Using Non-linear Components (LINC), and Envelope Elimination and Restoration (EER), are applied to maximize average efficiency[4][5][6]. Furthermore, the diverse allocation of radio spectrum and the need to support multiple wireless protocols must be achieved while reducing capital costs through hardware minimization. As the

PA dominates the physical size and cost of the RF front-end, it is critical to develop PAs that operate over multiple frequencies *concurrently*, while not compromising efficiency and linearity performance. Reconfigurable PAs such as switch-based single-band PAs or Micro-Electro-Mechanical Systems (MEMS) devices, can only provide a single service at any given time, and tend to produce loss between the PA and the antenna that greatly deteriorates RF Front-End performance. Alternatively, Multi-Band and Wideband PAs provide complimentary support for wide or narrow RF bandwidth separation, enabling concurrent ("always-on')' operation that maximizes data-throughput, while reducing the physical hardware footprint. Fig. 1.1 summarizes the PA design goals as a maximal compromise between power-efficiency, linearity and RF bandwidth.

The challenging multiple-goals of modern PA design requires very sophisticated design techniques and approaches. More particularity, the conventional empirical design approach that relies on Load-Pull measurements and/or major post-fabrication tuning cannot achieve a global optimum given the high level complexity of the problem. This places emphasis on the importance of an integrated and comprehensive computer aided design (CAD) methodology, relying heavily on the accuracy of models used to emulate both passive and active components in the design. Given that emerging PA topologies operate the transistor in strongly nonlinear modes, accurate nonlinear models of the RF transistor is paramount. However, the accuracy of the transistor model is controlled by the model formulation, the model extraction procedure, and the accuracy of the measurement data. While compact circuit models aim to improve model formulation, it relies heavily on conventional linear measurement techniques. Advanced large-signal measurement techniques, such as multi-harmonic load-pull, focus heavily on replicating specific design conditions, however few characterization solutions are tailored towards extracting measurement data to construct a model. Furthermore, it is unlikely that a nonlinear characterization system can replicate an exact electro-thermal operating condition while using test-bench instrumentation that differs vastly from the bandwidth and power-handling specifications of the actual design. While the literature continues to focus on improving the accuracy of the measurement data and the model formulation, there is little work that explores the impact of the model extraction procedure. Successful CAD-based design requires tight integration between *all* design processes, specifically device characterization, model extraction, and advanced

circuit design, as illustrated in Fig. 1.2.

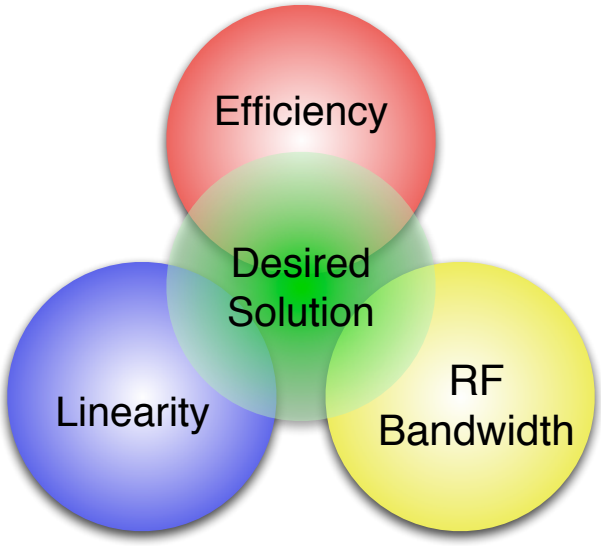


Figure 1.1: Power Amplifier Performance Goals

The core of the PA, the transistor, represents the source of nonlinearity and the main source of inaccuracy in Computer Aided Design (CAD) simulations. Although a linear device can be described using Scattering Parameters (S-Parameters), the performance of nonlinear devices is communicated using many high-level figures-of-merit, such as DC I/V characteristics, Hot S-Parameters, Load-Pull Contours, Intermodulation Distortion, AM-AM, and AM-PM[7]. While each figure-of-merit provides qualitative information that looks good on a specification sheet, few provide root-cause analysis that can be used to improve a nonlinear design. Hence, prior work has focused on the development of nonlinear measurement systems that capture sparse frequency spectrum and convert the data into time-domain I/V waveforms[8][9]. The extraction of accurate time-domain waveforms at the intrinsic transistor reference plane, provides unambiguous information that a designer can use to achieve optimal behaviour, and it can also be used to calculate all traditional design figures-of-merit.

As PA design processes become increasingly dependent on CAD simulation, the accuracy and availability of nonlinear compact transistor models is crucial to achieving the final

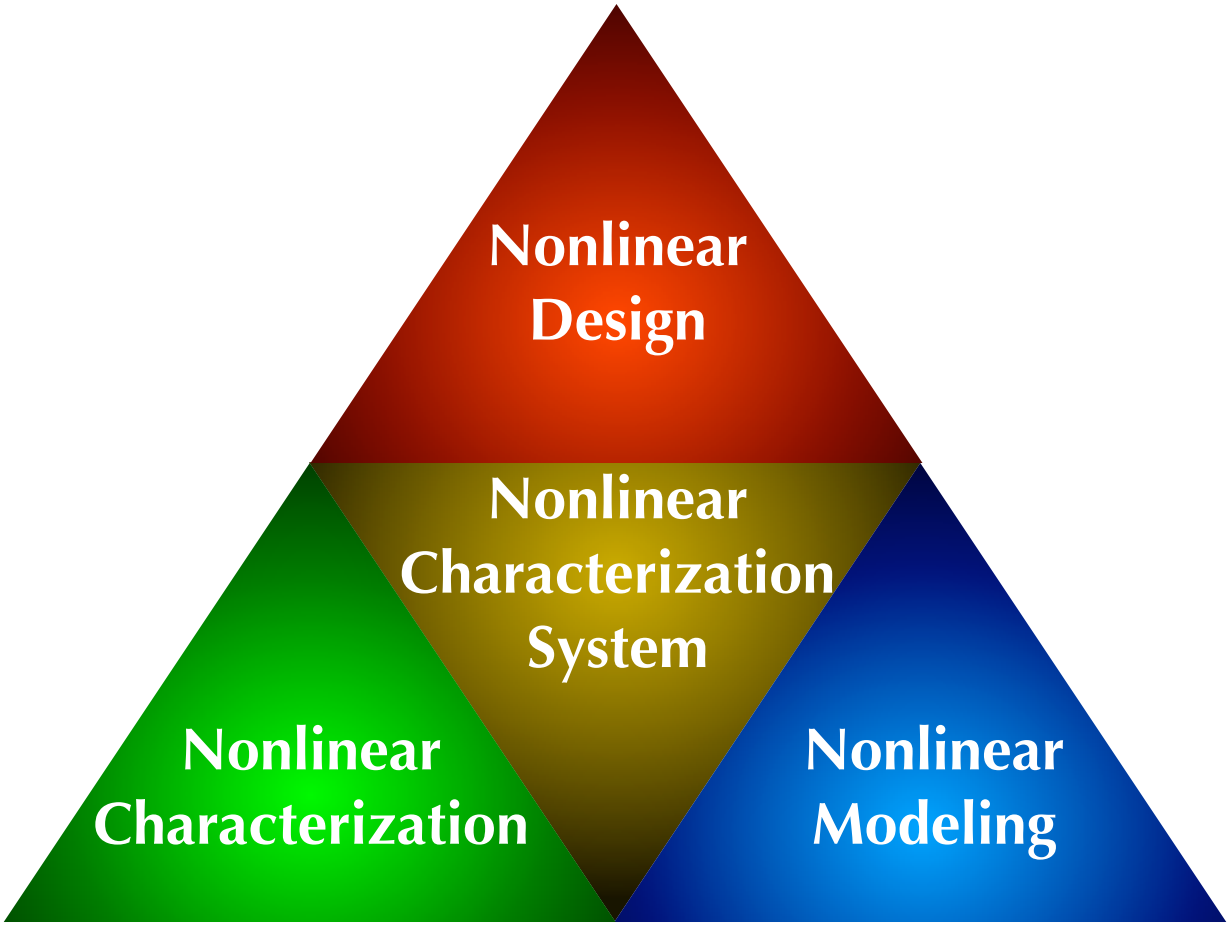


Figure 1.2: Cornerstones of Nonlinear Electronic Design Automation

design objectives. Unlike linear circuit models, whose equations often transcend fabrication processes and mechanical substrates, compact circuit transistor models are technology specific and have a greater dependency on numerical methods[10][11]. Furthermore, high-power devices rely heavily on exotic substrates and semiconductor doping profiles to maximize electro-thermal performance, hence they cannot be described using a single model formulation. This proliferation of semiconductor fabrication processes requires significant effort by device manufactures to produce a proprietary compact-circuit model to approximate the behaviour of nonlinear devices. Alternatively, behavioural modelling utilizes analytical methods to predict the behaviour of nonlinear devices based on information collected from many nonlinear measurements. This analytical model can represent a transistor behaviour using a mathematical expression that does not require any proprietary information about how the device was manufactured. While compact-circuit models are created using a combination of linear and nonlinear measurements, collected under DC or small-signal RF stimuli, the behavioural model can guarantee nonlinear "measurement-based" accuracy by extracting the model from "real-life" large-signal measurements.

The research in this thesis explores the multi-faceted approach of Nonlinear Electronic Design Automation that is needed to design and validate next generation nonlinear circuits. The proposed solution is based around a nonlinear characterization system that enables the design process flow outlined in Fig. 1.3. This process begins with characterization of a transistor, followed by compression of this measurement data into a comprehensive mathematical model, which is validated using independent measurements. The resulting transistor model is used to design a PA in a CAD environment, and then the fabricated PA is validated using the same characterization system. This process is successful when the elements of Fig. 1.2 are tightly integrated, thereby curating accurate design data in a regenerative information pipeline. Unfortunately there is a large fragmentation between the measurement and simulation environments that stems from a lack of standardized measurement equipment for characterizing arbitrary nonlinear systems. Hence the scientific contributions of this thesis are summarized as follows:

1. An automated nonlinear characterization system, based on the process flow in Fig. 1.3 extracts calibrated multi-harmonic measurement data in a controlled environment.

2. A method for enhancing the accuracy of a Poly-Harmonic Distortion (PHD) model is proposed without modifying the existing model formulation. For stronger nonlinear devices, a higher-order Multi-Harmonic Volterra (MHV) model extraction is investigated to further improve model accuracy.

3. A PA design example, a concurrent multi-band PA is derived from matching network theory and filter theory, and is used to illustrate the complexity of modern nonlinear circuit design.

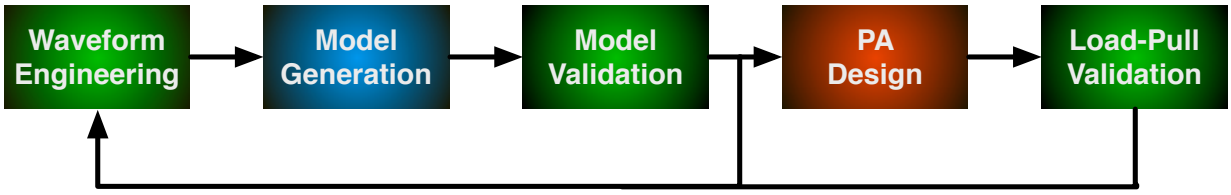


Figure 1.3: Feedback Design Process

Chapter 2 reviews the prior work in nonlinear measurement science and nonlinear behavioural modelling. It analyses existing nonlinear characterization techniques, impedance modulation techniques, and behavioural model synthesis.

Chapter 3 presents the proposed Nonlinear Characterization System (NCS), a modular hardware abstraction that isolates the functional elements of a nonlinear measurement test-bench. By standardizing the hardware interface, rapid development of generalized high-level calibration and measurement routines can be implemented without specific knowledge of the measurement test-bench. Seamless integration between the NCS and the Harmonic Balance simulator provides a 1:1 measurement and simulation comparison with the most common nonlinear circuit design tool. This enables early diagnosis of model inaccuracy down to the transistor-level, isolating the root cause of simulation versus measurement discrepancies, inside a comprehensive analytical RF test backbone that can be used throughout all stages of the design process. It also provides an ideal test environment where advanced test algorithms can be developed without performing tedious test-bench calibrations or incurring the cost of measurement system down-time. Finally, the NCS is used to control a multi-harmonic RF test-bench and several test-bench design consider-

ations are presented to achieve a measurement system that is optimized for behavioural model extraction.

Chapter 4 presents an analytical study of the PHD, Cardiff and MHV models that compares the implications of each model formulation and extraction procedure. An improved PHD model extraction methodology uses multi-harmonic impedance tuners to minimize reflections within the measurement system. This reduces the order of the nonlinear system, thereby permitting perceived stronger nonlinear device operating conditions to be described sufficiently by the PHD model. For strongly nonlinear systems, an extraction methodology of a higher-order MHV model is proposed to approximate high-order multivariate systems. In contrast to the fixed formulation of the PHD model, the proposed MHV model extraction uses a variable-order synthesis that can be practically scaled with respect to the number of harmonic system inputs.

Finally, Chapter 5 proposes two methods of designing a dual-band PA; using matching network theory, and using trans-impedance filter theory. It suggests ways that concurrent dual-band operation can be optimized by utilizing the research presented in Chapter 3 and Chapter 4.

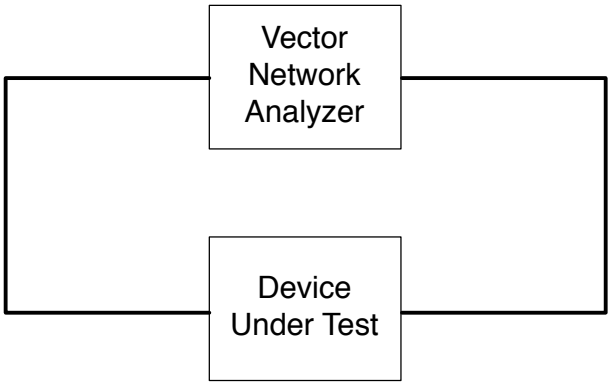
Chapter 2

Literature Review

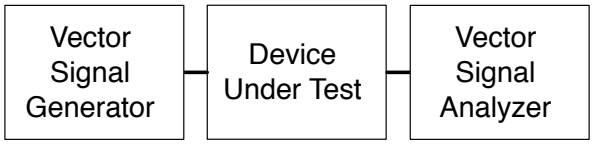
2.1 Nonlinear Characterization Techniques

Traditionally, nonlinear devices have been characterized using measurement architectures based on frequency-domain vector network analysis, shown in Fig. 2.1a[12], or a time-domain envelope measurement system, shown in Fig. 2.1b[13]. The four phase coherent receivers inside a Vector Network Analyzer (VNA) can be utilized to measure the 2-port incident and reflected vector power, useful for characterizing unmatched devices, however it can only characterize the device under test (DUT) using Continuous Wave (CW) signals. Alternatively, a time-domain envelope measurement, implemented using a Vector Signal Generator (VSG) and Vector Signal Analyzer(VSA) architecture, supports modulated signals, however it only contains one receiver and is only suitable for measuring the output response or transfer function of a matched system. While both architectures stimulate and measure the device under test (DUT) around the carrier frequency (f_0), they do not measure the DUT at DC ($0f_0$) or harmonic frequencies ($nf0$), hence these are linear characterization methods that can only be applied to weakly nonlinear systems. Neither architecture captures the full picture, and there is no ubiquitously accepted test and measurement solution that can fully characterize the large-signal response of a nonlinear device.

To provide a clear explanation, we can look at simulation architecture of computer-



(a) Vector Network Analyzer



(b) Vector Signal Generator and Vector Signal Analyzer

Figure 2.1: Traditional Nonlinear Characterization Techniques

aided design (CAD) tools such as the Keysight Advanced Design System (ADS). While the nonlinear steady-state frequency response can be extracted using a Harmonic Balance simulation[14], the time-domain nonlinear response around the carrier is determined using the Transient simulator[15]. Again, neither solution encapsulates the entire nonlinear response of the circuit, because the carrier frequency (f_0) is much greater than the frequency of the modulated envelope (f_m), the information of the signal (2.1).

$$f_0 >> f_m \tag{2.1}$$

If we tried to measure the composite carrier and envelope signal in the time-domain, it would demand a sample rate of at least $2f_0$ to satisfy the Nyquist rate, and we would need an extensive playback memory to record enough measurements to include the modulated signal envelope. For example, an LTE signal modulated around a carrier frequency $f_0 = 2.7GHz$ would require a sample period $t_s < 0.37ns$ and would require 27 million samples to record a $10ms$ playback of the signal. In order to measure a *multi-harmonic* time-domain response of the DUT, the minimum sample rate and memory size must be multiplied by the number of harmonics to be sampled. Alternatively, we could characterize the nonlinear swept frequency response, however much of the spectrum between f_m and f_0 is empty and this would be very inefficient. To efficiently characterize a nonlinear DUT, we must combine the frequency-domain multi-harmonic carrier response (at nf_0), with the time-domain response of the modulated envelope signal ($A(t)$) around each harmonic carrier, as shown in Fig. 2.2. Therefore, the objective of nonlinear characterization is to measure input and output signals in the form of (2.2) by sequentially tuning the receiver to each harmonic carrier frequency ($^{\pm jn2\pi f_0 t}$) and measuring the complex envelope signal ($A_n(t)$).

$$X(t,f) = \sum_{n=-\infty}^{\infty} A_n(t)\, e^{\pm jn2\pi f_0 t} \tag{2.2}$$

Where,

n is the harmonic index ($DC = 0$).

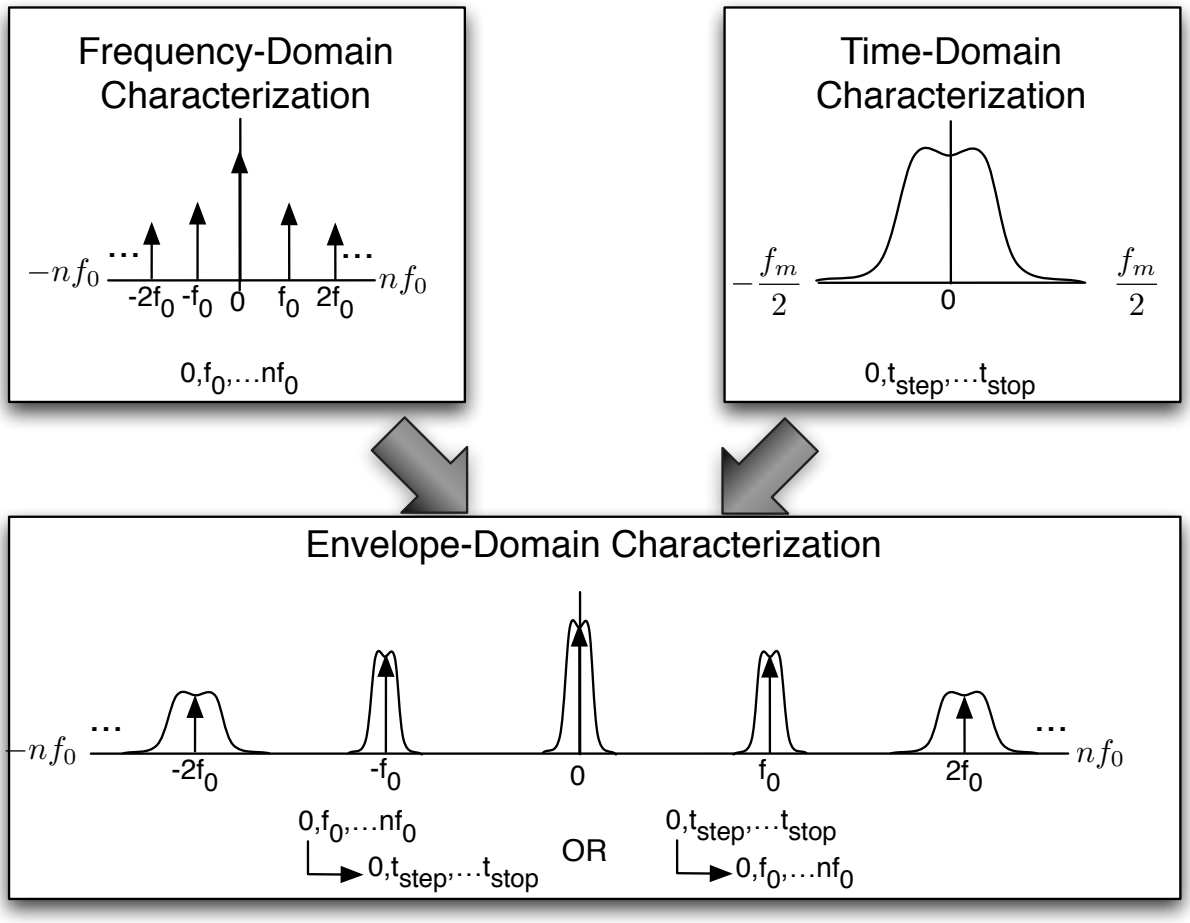


Figure 2.2: The generalized response of the DUT

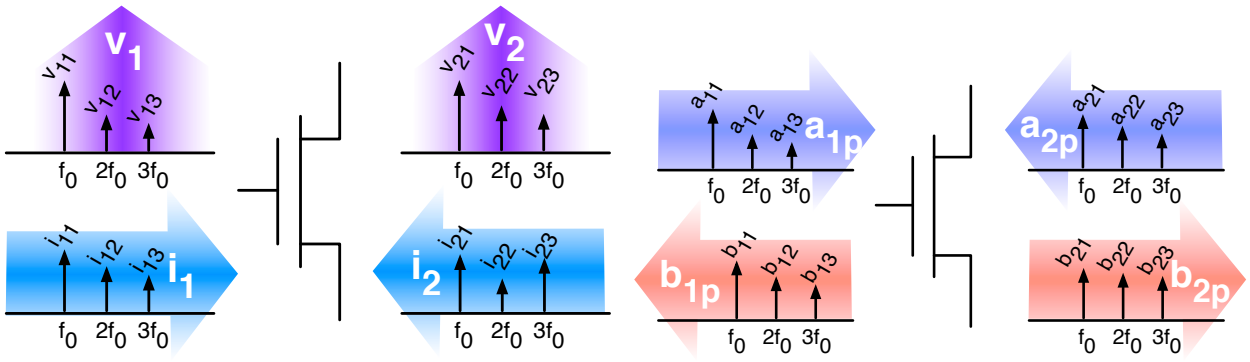
$e^{\pm jn2\pi f_0 t}$ is the n^{th} harmonic frequency carrier signal.

$A_n(t)$ is the n^{th} harmonic complex IQ-modulation envelope.

$X(t, f)$ is the measured multi-rate time-domain/frequency-domain signal.

As RF measurement constraints often demand that we measure incident and reflected power, rather than voltages and currents, the multi-harmonic measurement of nonlinear devices can be mapped between equivalent representations shown in Fig. 2.3a and Fig. 2.3b . While an envelope simulator might represent the data in voltages and currents, a

measurement test-bench would measure incident (a) and reflected (b) travelling-waves and then calculate voltage and current using (2.3 - 2.6). Since the RF stimulus signal-source is typically defined as a power source (a voltage source connected to an output impedance), it represents a Thevenin equivalent circuit, hence the incident power-wave can be controlled on each port independently (assuming the voltage source is linear). A controlled voltage source in an RF measurement requires source impedances that are matched to the DUT in order to maximize the amount of power transferred to the DUT, as shown in Fig. 2.4a. Since the current on *each* port is a function of the voltages on *both* ports, the voltages at the port of the DUT is uncontrolled. Alternatively, the power-sources defined in Fig. 2.4b, are defined independently of each other, therefore the available source power P_{AVS} can be controlled under the assumption that the reflected power-waves are fully absorbed by the source impedances[14].



(a) Voltage/Current Measurement (b) Incident/Reflected Power-Waves Measurement

Figure 2.3: Nonlinear Multi-Harmonic Measurement Data

$$v = \frac{1}{\sqrt{\Re\left(Z_0\right)}}\left(a + Z_0^* b\right) \tag{2.3}$$

$$i = \frac{1}{\sqrt{\Re\left(Z_0\right)}}\left(a - b\right) \tag{2.4}$$

$$a_p = \frac{1}{2\sqrt{\Re\left(Z_p\right)}}\left(v + Z_p i\right) \tag{2.5}$$

$$b_p = \frac{1}{2\sqrt{\Re\left(Z_p\right)}}\left(v - Z_p^* i\right) \tag{2.6}$$

Where,

Z_0 is the system characteristic impedance (typically 50Ω).

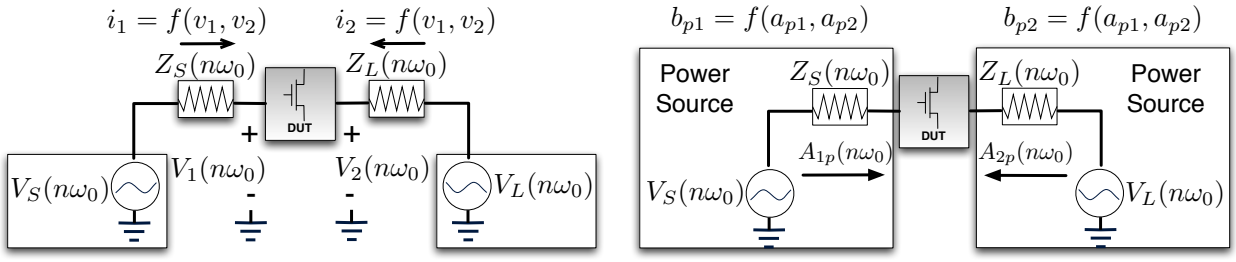
a is the incident *travelling-wave* (normalized by Z_0).

b is the reflected *travelling-wave* (normalized by Z_0).

Z_p is the port impedance.

a_p is the incident *power-wave* (normalized by Z_p).

b_p is the reflected *power-wave* (normalized by Z_p).



(a) Controlled Voltage Sources

(b) Controlled Power Sources

Figure 2.4: Controlled Voltage Sources vs. Controlled Power-Sources

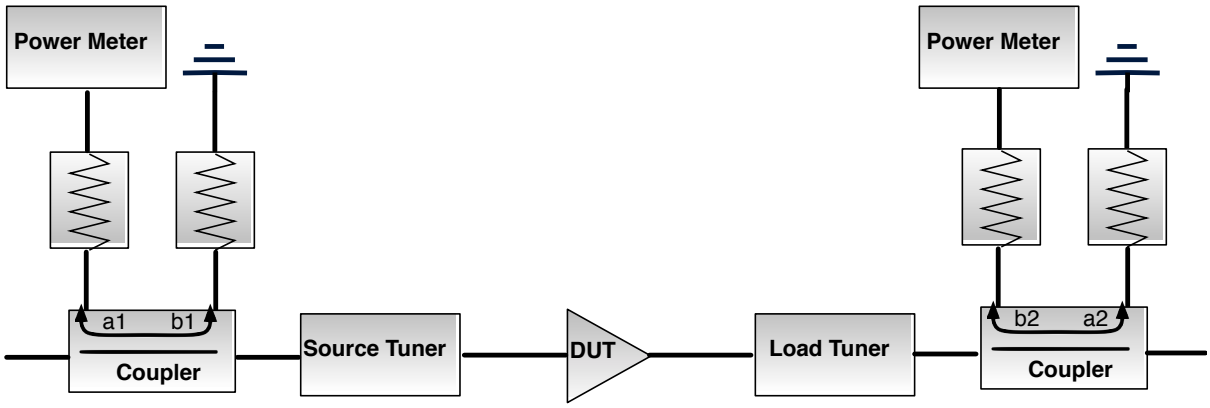
2.1.1 Nonlinear Measurement Receivers

Traditional scalar receivers, shown in Fig. 2.5a, used power meters to measure the scalar incident travelling-wave on the input port ($|a_1|$) and the scalar reflected travelling-wave on the output port ($|b_2|$). For unmatched devices, additional power meters can be added to measure $|b_1|$ and $|a_2|$, however without measuring the relative phase between these travelling-waves, we cannot evaluate (2.3 - 2.6), thus we cannot perform the following operations:

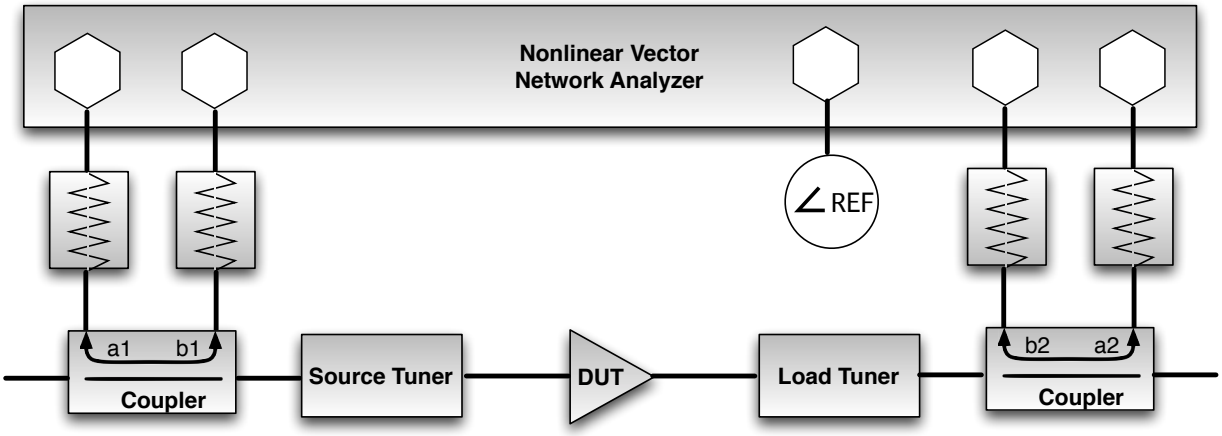
- De-embed the measurement plane from the power-meters to the DUT reference plane.
- Measure the vector reflection coefficient such as Γ_{IN} and Γ_L
- Compute power-waves, voltages and currents.
- Cannot compute time-domain waveforms.

Another reason why time-domain waveforms cannot be computed is because the power meters use broadband sensors that do not distinguish between the spectral content of the measurement. While this configuration in rudimentary, it provides a sufficient empirical method for finding the optimal output power and efficiency by sweeping the load impedance and the source power.

By using vector receivers, such as calibrated VNA receivers, we can de-embed the measurement to the DUT reference plane and compute power-waves and voltage/currents at the frequency of measurement. If we measure the DUT with an Nonlinear Vector Network Analyzer (NVNA), as shown in Fig. 2.5b, we can simultaneously measure a multi-harmonic frequency response and we can use an Inverse Fast Fourier Transform (IFFT) to compute time-domain signals at the ports of the DUT. Hence, by converting to a vector based receiver architecture, it is possible to achieve the same comprehensive measurement that an oscilloscope would provide, but at microwave frequencies.



(a) Scalar-Receiver Based Measurement System



(b) Vector-Receiver Based Measurement System

Figure 2.5: Scalar vs. Vector Receiver Based Nonlinear Measurement Systems

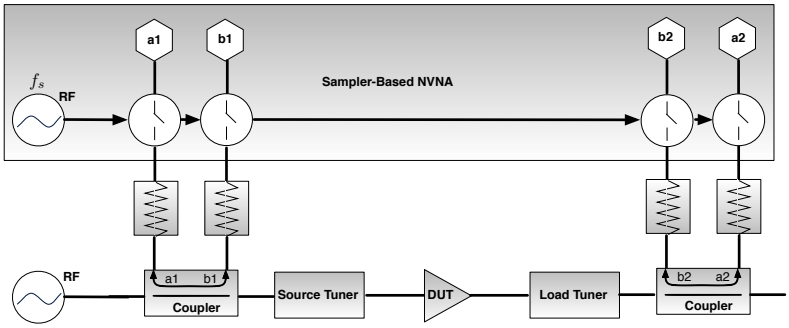
Nonlinear Vector Network Analyzer

A Nonlinear Vector Network Analyzer represents a conceptual test and measurement instrument that is described in [8][9]. There are multiple possible NVNA architectures, each with its own technical advantages and challenges:

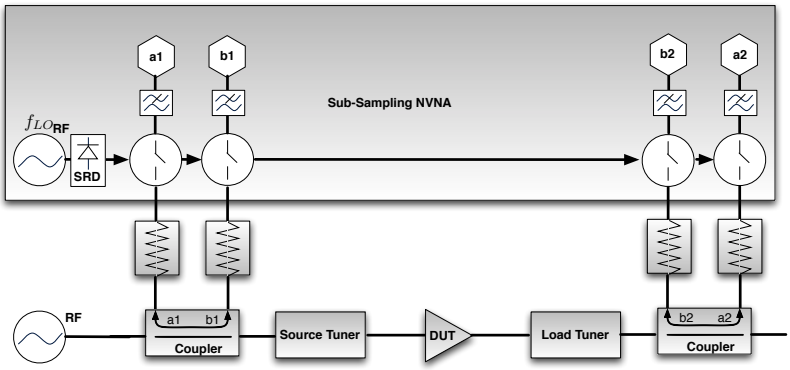
- Sampling-Based (Fig. 2.6a)
- Sub-Sampling Based (Fig. 2.6b)
- Mixer-Based (Fig. 2.6c)

The sampling-based method, shown in Fig. 2.6a, uses a high-frequency sampling oscilloscope [16] to sample the RF signal without down conversion. Given the fundamental frequency (f_0) and the number of harmonics (n) a sufficiently high sampling frequency $fs > 2nf_0$ must be chosen to measure the signal with integrity. A common sampling clock is shared between all receivers and all harmonics are measured simultaneously, therefore the measured signal does not need to be periodic and an additional phase synchronization source is not required. The analog-to-digital converters (ADCs) inside a high-speed oscilloscope typically have no greater than 8-bits resolution ($48dBc$ dynamic range), therefore for practical reasons, a periodic signal is often used so that the dynamic range can be extended using coherent measurement averaging. An example of this hardware architecture has been implemented by Cardiff University (Mesuro) [17].

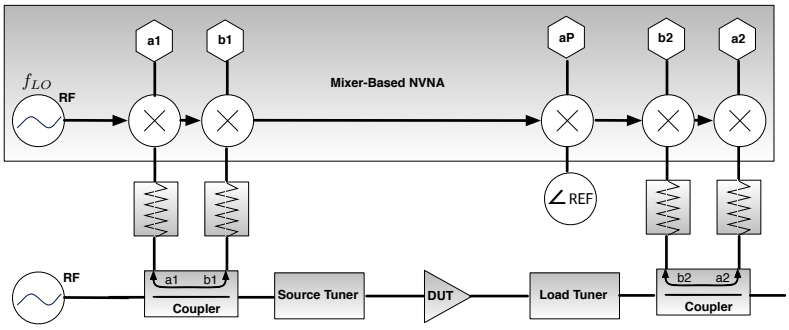
As most of the relevant spectral bandwidth is located around each harmonic of the carrier frequency, most of the sampling bandwidth captured by the sampling-based method is not utilized. The sub-sampling technique, shown in Fig. 2.6b, uses a Step Recovery Diode (SRD) to step through the RF signal, distributing samples over multiple repetitions of a periodic signal, effectively reducing the sampling rate. Various sampling algorithms down-convert all harmonic signals from RF to IF by carefully selecting the sampling frequency, f_s, to avoid aliasing between harmonic spectral content. By compressing the RF measurement data into a smaller IF bandwidth using techniques described in [18], the ADC only measures the modulated bandwidth around each harmonic carrier. These systems provide better dynamic range than the sampling-based approach, but are limited by harmonic aliasing and



(a) Sampling Based NVNA



(b) Sub-sampling Based NVNA



(c) Mixer Based NVNA

Figure 2.6: Nonlinear Vector Network Analyzer Architectures

timing sensitivity. While the Hewlett Packard Microwave Transition Analyser (MTA) [19] was the first commercialized application of this measurement technique, similar hardware architectures have been implemented in the NMDG LSNA[20] and the VTD SWAP[21].

The mixer-based method, shown in Fig. 2.6c, down converts the signal using mixers on each receiver, driven by a shared local oscillator (LO). While all travelling-waves can be measured simultaneously, the LO frequency ($f_{LO} = f_{RF} + f_{IF}$) can only be tuned to measure a single harmonic at a time, thus the input signal is assumed to be periodic, and an independent phase reference signal is required to provide cross-measurement synchronization. This solution provides superior dynamic range (typically $80 - 90dB$) by using narrow-band receivers and by measuring each signal relative to a constant phase reference. Unfortunately, the mixer-based solution requires an additional phase coherent RF source, dedicated to generating a phase reference signal, and it also requires a fifth receiver that is dedicated to measuring the phase reference. While other methods capture harmonics simultaneously, the acquisition time of the mixer-based solution is multiplied by the number of harmonics. The mixer-based solution has been implemented by modifying existing 4-port VNAs, such as the Keysight PNA-X [22] and the Rhode and Schwarz ZVx (by NMDG)[23].

Table 2.1: Nonlinear Vector Network Analyzer Architecture Comparison

Architecture	**Sampling**	**Sub-Sampling**	**Mixer**
Input Signal	Arbitrary	Periodic	Periodic
Frequency Range	Moderate	High	High
Dynamic Range	Low	Moderate	High
Bandwidth	High	Moderate	Low

A comparison of the NVNA hardware architectures is presented in Table 2.1. Although the mixer-based solution requires added source and receiver complexity, high-power circuits require extensive dynamic range to account for the attenuation (thermal noise) needed to reduce the power of the signal before it enters the receiver. When characterizing nonlinear

devices at harmonic frequencies, the signal integrity of the harmonic signals is degraded when signals are well below the $0.1dB$ compression point of the receiver, hence the effective dynamic range for this application is $10 - 20dB$ below the dynamic range of the receiver itself.

2.1.2 Nonlinear Operating Point Conditioning

The previous section discussed the different possible types of receivers needed to capture the input and output signals of a nonlinear device, but further research is needed to determine how to synthesise the nonlinear operating conditions. This section will specifically analyse multi-harmonic source and load impedance modulation, assuming the DC bias condition and frequency of operation are already fixed. Source/Load impedance modulation, otherwise known as load-pull, is the variation of the source/load impedance seen by the DUT, and has been traditionally achieved using a mechanical circuit consisting of a 50Ω slotted coaxial transmission line, whose impedance is altered by a sliding capacitively coupled transmission stub. The impedance seen looking into this system is controlled by changing the *x-coordinate* of the sliding stub along the transmission line, and *y-coordinate* capacitive coupling between the stub and the transmission line to achieve a reconfigurable single-stub matching network. Automated mechanical impedance tuners use electrical motors to control the X/Y position of the stub so that the DUT performance can be evaluated over many load impedances in a controlled, repeatable environment that accurately generates a reflection coefficient to within $-40dB$ precision. Due to conductance loss in the impedance tuner, it is impossible to synthesize a perfect open/short circuit, thus all mechanical impedance tuners can only generate impedances over a portion of the Smith Chart (typically $\Gamma < 0.95$). As the DUT is rarely connectorized, an additional DUT fixture is implemented using lossy transmission lines or wafer probes, therefore it is unlikely that the maximum tuner reflection coefficient can be presented to the DUT in many measurement applications. In many applications the tuning range provided by the impedance tuners is acceptable, however high-power transistors are constructed by connecting many small transistor fingers in parallel, thus the cumulative input impedance is equal to the input impedance of a single finger divided by the number of fingers ($Z_{IN} = Z_{finger}/\#fingers$).

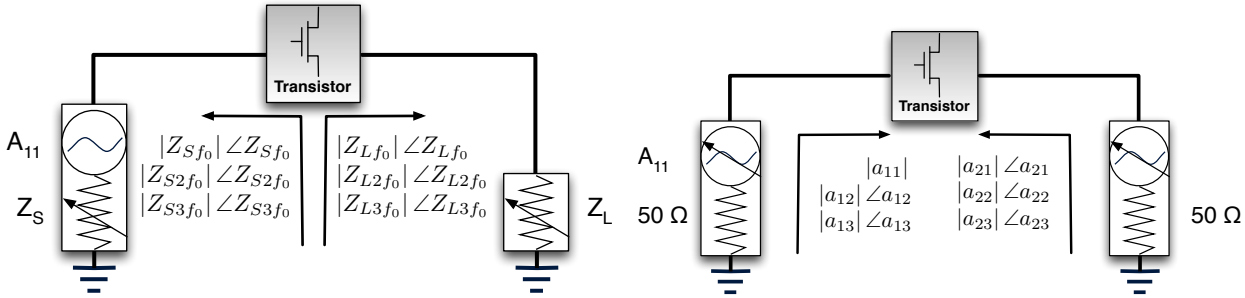
For transistors that exceed $100W$ output power, impedance tuners may not be able to synthesize a conjugate matching impedance. To extend the impedance tuner range, the DUT fixture can be constructed using Klopfenstein tapered lines that provide a broadband impedance transformation from the characteristic impedance of the input/output transmission line of the DUT to the $Z_0 = 50\Omega$ of the impedance tuners[24][17]. The reconfiguration speed and vibration of mechanical impedance tuners represent two complimentary drawbacks that can never be completely removed. While increasing the speed of the motors would allow for faster measurements, it would likely create larger vibrations that could potentially destroy wafer probes that use rigid connections. Hence, complex impedance tuner mountings on air tables have been designed to minimize vibration and insertion loss between the wafer-probe and the impedance tuner.

It is conceptually difficult to generate a multi-harmonic mechanical impedance tuner that provides a controllable impedance at an arbitrary number of harmonics. Three common techniques are currently used:

1. Multiple impedance tuners combined with a multiplexor
2. A fundamental frequency tuner combined with harmonic resonators
3. Multiple cascaded impedance tuners

The first solution is the simplest and works with existing hardware, however the multiplexer has an unacceptable amount of insertion loss that will greatly limit the tuning range of each tuner. The second solution uses narrowband resonance circuits at the harmonics to provide any impedance on the edge of the Smith Chart, followed by a fundamental tuner that provides a moderate tuning range. Since the resonators must reside between the DUT and fundamental tuner, the fundamental impedance range will be limited by the insertion loss of the resonators. The third solution provides simultaneous tuning capability for all harmonics over most of the Smith Chart. This multi-harmonic tuner requires 2-degrees freedom (X and Y coordinates) for each harmonic impedance that is controlled. To reduce the physical dimensions of the multi-harmonic impedance tuner, it is advantageous to integrate these tuning stubs into a single tuner body [25]. These multi-harmonic tuners

compute the stub positions using search algorithms, hence they often have larger memory requirements and use pre-calculated positions to generate the desired multi-harmonic impedance. Another disadvantage of the cascaded multi-harmonic impedance tuner is that it provides a narrower bandpass impedance bandwidth, thereby limiting the matching bandwidth for increasing reflection coefficients and potentially violating DUT stability requirements at lower frequencies.



(a) Passive Load-Pull

(b) Active Load-Pull

Figure 2.7: Passive Load-Pull vs. Active Load-Pull

While passive load-pull directly controls the reflection coefficient seen by the DUT, the reflection coefficient is a ratio of the travelling-waves described in (2.7).

$$\Gamma_{Port} = \frac{a_{Port}}{b_{Port}} \tag{2.7}$$

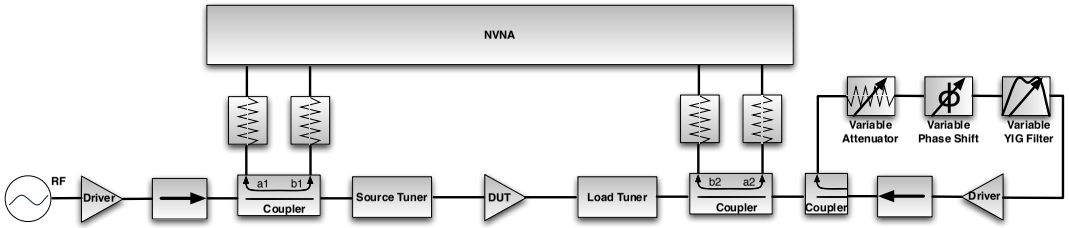
Instead of directly controlling the reflection coefficient (Γ_{Port}) , we could indirectly control the reflection coefficient by injecting a signal into the port using a power source, thereby directly controlling a_{Port}. Although we no longer directly synthesize Γ_{Port}, active load-pull can create any reflection coefficient and can even generate $\Gamma_{Port} > 1$ when $a_{Port} > b_{Port}$. Although this appears to be straight forward, traditional large-signal Single-Input, Single-Output (SISO) measurement systems (2.1b) are built on the assumption that a unitary source is applied at the input port and that the response is measured at the output port. Hence, active load-pull demands that we move to a Multiple Input, Multiple Output (MIMO) characterization system, where each port that is connected to the DUT serves as both an input and an output. Active Load-Pull systems can also synthetically equalize

the matching impedance over bandwidth, and can perform rapid impedance modulation by injecting modulated signals whose PAPR indirectly controls the amount of impedance modulation. In some applications, such as behavioural modelling, it is computationally advantageous to sweep the injected signal into the output, rather than sweep the output impedance.

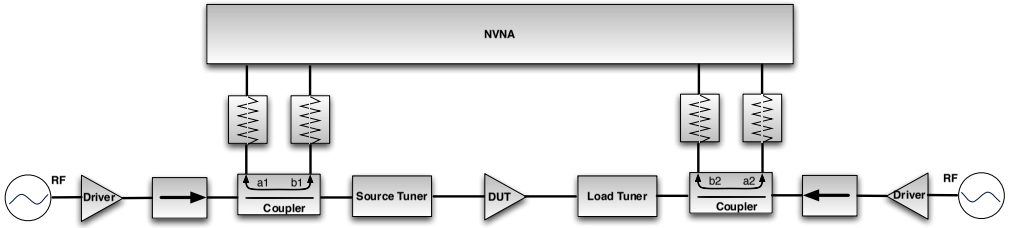
The work of Takayama et al in [26] provided the first active impedance modulation [27]. A single signal-source is split between the input and the output, while a tunable attenuator and tunable phase shifter is used to vary the magnitude and phase of the output source relative to the input source. This approach provides an elegant solution that only uses one signal source and requires minimal hardware resources. Another solution, presented in [28], utilizes an output feedback loop to condition the magnitude and phase of the b_2 reflected-wave before injecting the signal back into the system as the a_2 incident-wave. This solution is known as closed-loop active load-pull because it utilizes a feedback loop and does not require a separate signal-source to control incident power on the output port. One advantage of this solution is that it utilizes the amplified output-power of the DUT to synthesize the a_2 wave whose power requirements are typically similar to the output power of the DUT. Another advantage of this solution is that the magnitude of a_2 is implicitly tied to the magnitude of a_1, thus the synthesized load reflection coefficient will not change dramatically when the input power is varied. The disadvantages of this solution is that the feedback loop increases the likelihood of oscillation. It has been historically difficult to generate phase stable RF signal sources, hence the open-loop load-pull technique shown in Fig. 2.7b would traditionally result in unreliable measurements due to drift errors. A solution presented in [29] eliminates this problem by synthesizing both input and output signals using a high-frequency Arbitrary Waveform Generator (AWG) that contains multiple phase coherent digital-to-analog converters (DACs). While the AWG can synthesize multiple phase coherent tones, its sampling frequency must obey the Nyquist criteria ($f_s > 2nf_0$), and this solution can contribute broadband noise to the characterization system. Another solution [30] eliminates this restriction by using multiple phase coherent AWGs for each port/harmonic combination that is fed by a multi-harmonic local oscillator (LO). While the open-loop active load-pull concept is more complicated and expensive, it maximizes the signal generation bandwidth. It also allows the signal-sources to become

fully decoupled, however the sources can always become coupled using iterative software defined algorithms.

Extending active load-pull to support multi-harmonic load-pull is much easier than supporting multi-harmonic passive load-pull. Multiple active load-pull sources can be combined using a multiplexer without reducing the tuning range, assuming that extra gain is added to the specification of each source. Unlike the broadband nonlinear receiver and the cascaded multi-harmonic passive impedance tuner, almost all solutions presented require additional hardware sources for each harmonic on each port. The only exception is the Cardiff solution that only requires one source per port, however the maximum number of harmonics is limited by the sampling frequency of the AWG [29]. As the number of required sources can vary depending on the application, active load-pull can be seen as cost prohibitive because it implicitly demands a modular hardware solution.



(a) Closed-Loop Active Load-Pull



(b) Open-Loop Active Load-Pull

Figure 2.8: Closed-Loop vs. Open-Loop Load-Pull Architecture

Active load impedance modulation has many benefits compared to passive load-pull. It provides superior reconfiguration speed, a larger theoretical tuning range, and is better suited for multi-harmonic load-pull. It can also synthesize load impedances that would otherwise cause the DUT to oscillate, exceed current limitations, or voltage breakdown ratings by temporarily synthesizing *instantaneous* virtual impedances for small periods of time. In high-power applications, cost prohibitive PA drivers are needed to synthesize the full-range of impedances, therefore passive load-pull impedance tuners are used to pre-match the DUT, thereby pre-matching the active load-pull source, and reducing the power requirements of the system.

2.2 Nonlinear Modelling Techniques

Nonlinear modelling is required to compact large amounts of empirical data into a small, but representative, mathematical expression. Different model types can be classified under three categories: Physics-based, Compact Circuit, and Behavioural models as shown in Table 2.2. Physics-based models use knowledge of the semiconductor doping profiles and electromagnetic fields to model outputs such as electron mobility[31]. They are highly accurate, but become too cumbersome for circuit-level designs that consist of discrete and distributed components implemented on multiple substrates. Compact-circuit models approximate device behaviour by extracting a large-signal and small-signal equivalent circuit that can be implemented in spice-based circuit simulators[10][11]. While they provide a good overall prediction of the device behaviour using inexpensive measurement techniques, their formulation is technology dependent, they provide poor accuracy under nonlinear operation, and they require significant computation resources when evaluated in complex system-level simulations. The behavioural model is extracted from large measurement datasets of empirical measurements that are extracted under large-signal conditions that approximate the probability distribution function (PDF) of the final intended stimulus condition. They represent a purely empirical modelling solution with measurement-based accuracy, however they only encapsulate the behaviour of the DUT where they were measured, thus they give little insight into the internal circuitry and provide poor extrapolated results.

Table 2.2: Nonlinear Model Architecture Comparison

	Physics-Based	Compact Circuit	Behavioural
Concept	Derived from semiconductor physics	Power series and linear circuit elements combine to model I/V and Q/V relationships	Volterra Series black-box representation of the nonlinear transfer function
Application	Device Level	Device/ Circuit Level	Circuit/ System Level
Creator	Semiconductor Manufacturer	Manufacturer/ ling company	Manufacturer or Circuit Designer
Strengths	• Synthesized directly from the device physics. • Evaluation of fabrication process	• Global prediction of device behaviour • Amenable to circuit simulators • Mature concept	• Extracted under actual design conditions • Encapsulates Memory Effects • Fast Simulation
Drawbacks	• Requires knowledge of proprietary fabrication process • Too complex/slow to model a larger circuit layout	• Extracted under unrealistic stimulus • Process dependent, empirical curve-fitting process • Slow/inaccurate for system-level simulation	• Little analytical interpretation of results • Complex calibration and measurement apparatus • Simulation convergence problems due to poor extrapolation

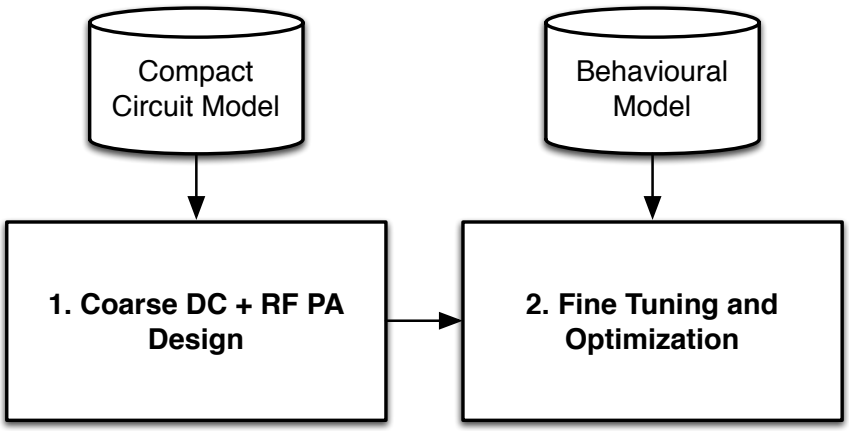


Figure 2.9: Complimentary Transistor Model Solutions for Power Amplifier Design

As the Physics-based model is too complex for most design applications, PA design has traditionally been reliant on two parallel solutions shown in Fig. 2.9; 1. a compact-circuit model extracted using puled DC and S-parameter measurements, and 2. a behavioural model (usually a look-up table (LUT)) extracted using load-pull measurements. Together, the compact-circuit and behavioural models represent complementary "course" and "fine" solutions, that provide global and localized accuracy respectively. For example, a compact-circuit model can be used to determine the proper DC bias and initial MN design, however the behavioural model can be extracted using application specific modulation schemes. This two-model design philosophy is vital to maximizing the performance of complex nonlinear systems such as efficiency-enhancement techniques, or Monolithic Microwave Integrated Circuits (MMICs).

Modern technologies such as GaN and GaAs have a large available bandwidth that enables designers to maximize efficiency using waveform engineering to design switching-mode amplifiers. This desire to accurately control the multi-harmonic signal behaviour of the transistor requires additional model accuracy to predict output signals at harmonic frequencies that can be atleast an order of magnitude below the fundamental frequency output. Therefore, an overview of existing modelling techniques will analyse how the compact circuit model and behavioural modelling techniques can be advanced to provide a more accurate modelling solution.

2.2.1 Compact Circuit-Based Modelling

The compact circuit model is a divide-and-conquer approach to modelling that seeks to extract elements of an equivalent circuit element-by-element using a sequence of measurements. The extrinsic model is extracted using the Cold FET measurement technique [32], followed by the intrinsic model that is extracted using information from isothermal DC large-signal measurements and small-signal RF measurements.

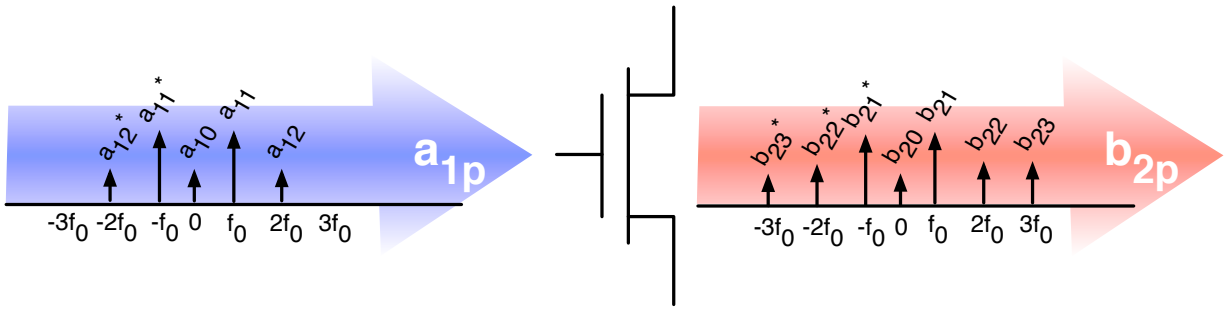
By decomposing the problem into several equivalent circuit components, the compact model achieves stable coefficient extraction. This approach assumes that the behaviour of the system is separable, implying that the large-signal RF behaviour can be fully predicted using a superposition of coefficients. In high-power applications, scaling techniques are used to predict the combined performance of multiple transistor fingers sub-circuits[33]. Unfortunately, when the extracted compact model is compared to large-signal load-pull measurements, it often contains discrepancies when the transistor is operating under non-linear conditions. Therefore, load-pull driven optimization is often performed around specific bias conditions to curve-fit measurement data to suite an intended application.

To increase the accuracy of the compact model there is an increasing dependency on non-analytical solutions ranging from data optimization to more sophisticated neural networks[34]. Due to the sequential nature of this model extraction, measurement or numerical error during each procedure is propagated through the rest of the model extraction. Hence, while significant effort could achieve slight improvements in compact circuit model accuracy, they would likely be technology or device specific ideas that would have a limited impact.

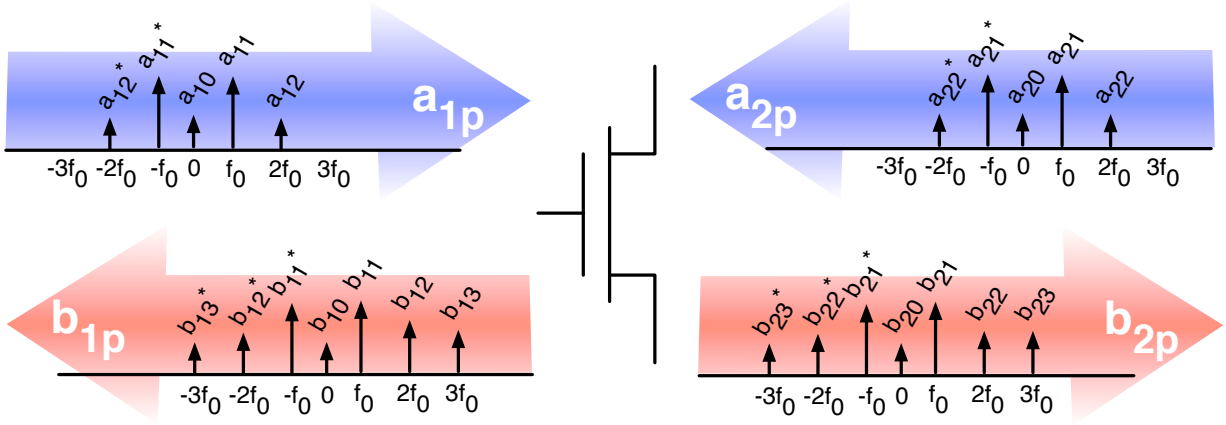
2.2.2 Behavioural Modelling

While the compact model achieves moderate accuracy using present characterization technology, it will never provide the inherent accuracy of measuring the large-signal behaviour of the DUT on the test-bench. Instead of trying to improve a mature methodology, we could alternatively focus on improving the intelligence of load-pull measurements by moving from load-pull LUTs [35] to more sophisticated numerical models. Extensive work

in behavioural modelling already exists and is commonly used in signal pre-distortion techniques[36][37][38], however it's application has been traditionally limited to Single-Input, Single-Output (SISO) systems as shown in 2.10a. The goal of multi-harmonic behavioural modelling is to extend existing behavioural modelling concepts to Multiple-Input, Multiple-Output (MIMO) systems that are characterized using multi-harmonic active load-pull.



(a) Single-Input, Single-Output System



(b) Multiple-Input, Multiple-Output System

Figure 2.10: Single-Input, Single-Output vs. Multiple-Input, Multiple-Output Systems

The Volterra Series

The Volterra-Series[39] [40] is a black-box generalization of the power-series that is applied to nonlinear problems that cannot be decomposed into separate linear and nonlinear components[41][42]. The frequency-domain formulation in Fig. 2.8 demonstrates that the system response cannot be accurately predicted by the individual transfer functions ($H(\omega_{q1}, \omega_{q2}, ...\omega_{qn}) \neq k_n H(\omega_{q1})H(\omega_{q2})...H(\omega_{qn})$). Equivalently, the time-domain formulation, presented in (2.9), demonstrates how the Volterra Series predicts the response of dynamic memory effects. The time-step used in (2.9) is proportional to the envelope frequency ($\tau \propto 1/f_m$), while the static impulse response ($h(t)$) encapsulates information about the carrier frequency. The assumptions made by the Volterra Series are that the system is weakly nonlinear, consisting of non-commensurate inputs signals that do not dramatically affect the DC response of the system.

$$b\,(t) = \sum_{n=1}^{N} \frac{1}{2^n} \sum_{q1=-Q}^{Q} \sum_{q2=-Q}^{Q} ... \sum_{qn=-Q}^{Q} H(\omega_{q1}, \omega_{q2}...\omega_{qn}) \prod_{l=1}^{n} A_{ql} e^{j\omega_{ql}t} \tag{2.8}$$

$$b\,(t) = \sum_{n=1}^{N} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} h_n(\tau_1, \ldots \tau_n) \prod_{l=1}^{n} a(t-\tau_l) d\tau_l \tag{2.9}$$

Where,

Q represents the number of frequency inputs (number of tones).

l is an frequency input indexation term.

n represents the nonlinear degree of the equation.

ω_{q1}, ω_{q2}, ...ω_{qn} are multi-tone frequencies.

τ_1, τ_2, ...τ_n are time-delays.

A_{ql} is the frequency-domain input (incident travelling-wave).

$a(t)$ is the time-domain input (incident travelling-wave).

$b(t)$ is the time-domain output (reflected travelling-wave).

The assumptions made in the Volterra-Series are no longer valid for a Multi-Harmonic model because the input signals are commensurate and many of their mixing products will occur at DC. Consequently, the time step in (2.9) is now proportional to the frequency of the carrier ($\tau \propto 1/(nf_0)$), thus the "memory" that the Volterra series would predict is due to a variation of the multi-harmonic matching network. This variation in the multi-harmonic behaviour produces a variation in the DC response, therefore the synthesis of a Multi-Harmonic Volterra (MHV) model can be seen as a Volterra-Series where each kernel has "memory" that is controlled by the magnitude and phase of the multi-harmonic input stimuli [43].

The derivation of the MHV model in [43] takes the constant kernels ($H(\omega_{q1}, \omega_{q2}, ...\omega_{qn})$) from (2.8) and converts them to functions of the mixing terms that produce mixing products at DC ($V_{i0,1}(\vec{x}_n(0))$), as shown in (2.10-2.12). A similar approach is taken in the VIOMAP algorithm developed in [44].

$$B_{i0} = V_{i0,1}(\vec{x}_n(0)) \tag{2.10}$$

$$B_{ik} = \sum_{m=1}^{M_{n,k}} \vec{x}_{n,m}(k) V_{ik,m}(\vec{x}_n(0)), \text{ for } k \neq 0 \tag{2.11}$$

$$\vec{x}_{n,m}(k) = \prod_{j=1}^{p} \prod_{l=-n}^{n} A_{jl}, \text{ such that } \sum l = k \tag{2.12}$$

Where,

n is the nonlinear order of the system.

A_{jl} is the large-signal incident travelling-wave applied at input j, harmonic l.

B_{ik} is the reflected travelling-wave measured at output i, harmonic k.

m is an indexation term.

$M_{n,k}$ is the number of indexation terms for output harmonic k of an n^{th} order formulation.

$\vec{x}_{n,m}(k)$ is the mixing term combination of A_{jl} inputs that produce a mixing product at harmonic k.

$V_{ik,m}(\vec{x}_n(0))$ is a MHV kernel function of the DC mixing terms that represents the weighting of mixing terms that produce a mixing product at harmonic k.

An sample formulation of a MHV model for a 1-port, 2-harmonic system is shown in (2.14-2.16). Substituting (2.17) into (2.14-2.16) will multiply the number of mixing products of a Volterra Series by the number of mixing terms in (2.17), therefore a significant number of kernels must be solved to extract this model.

$$B_{20} = V_{0,1}(\vec{x_2}(0)) \tag{2.13}$$
$$B_{21} = A_{11}V_{1,1}(\vec{x_2}(0)) + A_{11}^{*}A_{1,2}V_{1,2}(\vec{x_2}(0)) \tag{2.14}$$
$$B_{22} = A_{1,2}V_{2,1}(\vec{x_2}(0)) + A_{11}^{2}V_{2,2}(\vec{x_2}(0)) \tag{2.15}$$
$$B_{23} = A_{11}^{3}V_{3,1}(\vec{x_2}(0)) + A_{11}A_{12}V_{3,2}(\vec{x_2}(0)) + A_{11}^{*}A_{12}^{2}V_{3,3}(\vec{x_2}(0)) \tag{2.16}$$

where,

$$\vec{x}_2(0) = A_{10},\ A_{11}^{*}A_{11},\ A_{12}^{*}A_{12},\ A_{12}^{*}A_{11}^{2},\ A_{11}^{*2}A_{12} \tag{2.17}$$

Load-Pull Look-Up Table

Traditional load-pull measurements sweep the load impedance while recording the large-signal response in a LUT, and then display the trade-offs between performance goals as contours on a Smith Chart. This brute-force method of modelling a DUT provides a

simple, fast, and accurate way of modelling the performance of any unknown system. Unfortunately, as the available bandwidth of the DUT increases, the problem becomes multi-dimensional, essentially requiring two-dimensions of analysis for each harmonic frequency, multiplied by the number of ports on the DUT. Table 2.3 and 2.4 illustrate the large number of load-pull measurements for a 2-port transistor that is sensitive to 3-harmonics. To constrain the maximum distance between interpolated points in a higher dimensional LUT, the density of measurements within each dimension must be multiplied by $\sqrt{D}$ based on the formulation of a multi-dimension distance equation, where D is the number of sweep dimensions.

Table 2.3: Increase in Measurements for Multi-Harmonic Load-Pull

Sweep ID	Description	Meas. per Sweep
V_{DC1}	DC Gate Voltage	1
V_{DC2}	DC Drain Voltage	1
f_0	Fundamental Frequency	1
A_{11}	Port 1 Input Power at f_0	N^1
Γ_{21}	Port 2 Reflection Coefficient at f_0	N^2
Γ_{12}	Port 1 Reflection Coefficient at $2f_0$	N^2
Γ_{22}	Port 2 Reflection Coefficient at $2f_0$	N^2
Γ_{13}	Port 1 Reflection Coefficient at $3f_0$	N^2
Γ_{23}	Port 2 Reflection Coefficient at $3f_0$	N^2
Total		N^{11}

Table 2.4: Number of Measurements as a Function of Measurements per Sweep (N)

N	2	4	10
# of Measurements	2048	4.194×10^6	1×10^{11}

Although a load-pull LUT provides measurement-based accuracy, extending this logic to multi-harmonic load-pull problems will dramatically increase measurement time and/or decrease interpolation accuracy. Therefore, a more sophisticated load-pull based model

needs to provide enhanced predictive capability in order to minimize the number of multi-harmonic impedance tuner positions that are needed to extract the model.

Poly Harmonic Distortion Model

The derivation of the PHD model can be seen as the desire to modify a linear S-parameter model such that it predicts the characteristics of a nonlinear device. Swept S-parameters versus input power (2.18) can be generated, however $A_1 = 0$ when extracting S_{12} and S_{22}, therefore these terms don't change as a function of input power. To resolve this problem, "Hot S_{22}" measurements extract S_{12} and S_{22} when the input power is turned on (2.19)[45], thereby providing information about how the output impedance of the DUT is changing with respect to input power. This solution creates another problem by converting a SISO model into a MIMO model and now the performance of the DUT is simultaneously dependent on A_1, A_2 and the relative phase difference between the two inputs, $\phi(A_2) - \phi(A_1)$. To explain this another way, when a sinusoidal A_1 or A_2 signal is applied to a DUT at frequency f_0, there is also a conjugate A_1^* and A_2^* that is simultaneously applied at the negative frequency $-f_0$. When only one of these signals is applied, a phase normalization ($P^-1 = \phi(A_1)$) can be applied to each input signal such that A_1 is always real, however when a second input is applied, it is not necessarily phase coherent. Therefore, the formulation in (2.20-2.21) adds the A_2^* input term term and characterizes its response as R_{22}. One important observation is that the response resulting from A_2 was assumed to be much lower than A_1, therefore A_2 is replaced with a_2 in (2.22) to signify that $f(a_2) << f(A_1)$.

$$B_2P^{-1} = S_{2,1}\left(|A_1|\right)A_1P^{-1} + S_{22}A_2P^{-1} \tag{2.18}$$

$$B_2P^{-1} = S_{2,1}\left(|A_1|\right)A_1P^{-1} + S_{22}\left(|A_1|\right)A_2P^{-1} \tag{2.19}$$

$$B_2P^{-1} = S_{2,1}\left(|A_1|\right)A_1P^{-1} + S_{22}\left(|A_1|\right)A_2P^{-1} + R_{22}\left(|A_1|\right)conj(A_2P^{-1}) \tag{2.20}$$

$$B_2 = S_{2,1}\left(|A_1|\right)A_1 + S_{22}\left(|A_1|\right)A_2 + R_{22}\left(|A_1|\right)A_2^*P^2 \tag{2.21}$$

$$B_2 = S_{2,1}\left(|A_1|\right)A_1 + S_{22}\left(|A_1|\right)a_2 + R_{22}\left(|A_1|\right)a_2^*P^2 \tag{2.22}$$

Where,

A_1 is the large-signal incident travelling-wave applied at port 1.

A_2 is the large-signal incident travelling-wave applied at port 2.

A_1 is the large-signal reflected travelling-wave applied at port 1.

A_2 is the large-signal reflected travelling-wave applied at port 2.

P represents the phase of the input signal $\phi(A_1)$.

$S_{2,1}$ is the output response at port 2 (at $+f_0$) when an input is applied at port 1 (at $+f_0$).

$S_{2,2}$ is the output response at port 2 (at $+f_0$) when an input is applied at port 2 (at $+f_0$).

$R_{2,2}$ is the output response at port 2 (at $+f_0$) when an input is applied at port 2 (at $-f_0$). This coefficient is zero in linear systems, because nonlinear frequency mixing products do not exist.

As the nonlinear response to an input is often represented by a power-series, the response of a sinusoidal input consists of spectral tones at multiples harmonics of the fundamental frequency. As the DUT produces an output at harmonic frequencies, these output signals can also be reflected back into the system, thus acting as additional inputs to the MIMO system. Thus, the PHD model, shown in (2.23), is a generalization of (2.22) that includes additional inputs over an arbitrary number of harmonic frequencies. As the response to A_{11} is typically much greater than the rest of the inputs in most real-life applications, the validity of using the principle of superposition to describe a nonlinear system is justified by the assumption that all other inputs (signified by a lower-case a_{jl}) are acting as a small-signal "linear perturbation" of the nonlinear system around a large-signal operating point (LSOP), determined by $|A_{11}|$. This approximation implies that the model can be extracted by sequentially applying a signal at each harmonic, on each port, such that a miniature active load-pull is used to extract a complex first-order Taylor series around the LSOP. In (2.23), the LSOP response to $|A_{11}|$ is encapsulated in the X^F term, while the small-signal linearized response to each other signal is encapsulated in the X^S terms

for the positive frequency inputs, and the X^T terms for the negative frequency inputs. As a positive frequency input cannot be synthesized without also synthesizing a conjugated negative frequency input, the X^S and X^T terms must be extracted by solving atleast two simultaneous equations, hence multiple measurements are generated by sweeping the phase of a_{jl} to satisfy the number of degrees of freedom[46].

$$B_{ik} = X^F_{ik}(|A_{11}|, V_{10}, V_{20})P^k + \sum_{(jl)\neq(11)} X^S_{ik,jl}a_{jl}P^{k-l} + X^T_{ik,jl}a^*_{jl}P^{k+l} \quad (2.23)$$

$$B_{ik} = X^F_{ik}(|A_{11}|, A_{21}, V_{10}, V_{20})P^k + \sum_{(jl)\neq(11),(21)} X^S_{ik,jl}a_{jl}P^{k-l} + X^T_{ik,jl}a^*_{jl}P^{k+l} \quad (2.24)$$

$$B_{ik} = X^F_{ik}(LSOP_{mn})P^k + \sum_{(jl)\neq(mn)} X^S_{ik,jl}a_{jl}P^{k-l} + X^T_{ik,jl}a^*_{jl}P^{k+l} \quad (2.25)$$

Where,

a_{jl} is the small-signal incident travelling-wave applied at input j, harmonic l.

A_{jl} is the large-signal incident travelling-wave applied at input j, harmonic l.

B_{ik} is the large-signal reflected travelling-wave measured at output i, harmonic k.

P represents the phase of the input fundamental harmonic $\phi(A_{11})$.

X^F_{ik} is the model kernel that determines $B_{ik} = f(A_{11})$.

$X^S_{ik,jl}$ is the model kernel that determines $b_{ik} = f(a_{jl})$.

$X^T_{ik,jl}$ is the model kernel that determines $b_{ik} = f\left(a^*_{jl}\right)$.

The PHD model can be described as a complex first-order multi-dimensional (multi-harmonic) Taylor series approximation of the nonlinear behaviour around a given point (LSOP). For strongly nonlinear systems, it is not surprising that the PHD model may no longer provide a good approximation of the DUT behaviour, therefore a LUT-based extension known as the Load-Dependent X-Parameter model (2.24) provides an extension

to the PHD model in the event that A_{21} is no longer assumed to be a small-signal model input[47][48]. As the PHD model is already a LUT with respect to $|A_{11}|$, the Load-Dependent X-Parameter model merely generates a three-dimensional LUT with respect to $|A_{11}|$, $|A_{21}|$, and $\phi(A_{21})$. A generalization of this concept (2.25), suggests that any perceived "large-signal" input can be added to a multi-dimensional LSOP LUT of PHD models that are extracted for all inputs that are assumed to be "small-signal" inputs. Although there is no commonly accepted figure of merit that would distinguish a large-signal input from a small-signal input, empirical analysis has suggested that the response to small-signal inputs must be atleast an order of magnitude below the response to the large-signal input ($f(SS_{Input}) << f(LSOP_{Input})$).

While the PHD model is easiest to explain as an extension to the S-Parameter model, it is mathematically derived as a simplification of the MHV model. The PHD model formulation approximates the MHV model by only extracting kernels that consist of A_{11} and one other input, as all other high input combinations are treated as hidden variables. This dramatically decreases the number of model kernels, and simplifies the instrumentation that is necessary to extract the model coefficients.

Cardiff Model

As implied above, the PHD model may not be able to predict the behaviour of a strongly nonlinear device, hence the Cardiff Model first tried to model the response to the input phase variation ($\phi(A_{jl})$) of any input signal by using a higher-order Fourier Series. A graphical comparison of how the S-Parameter, PHD, and Cardiff model would model the nonlinear response of a stimulus signal with constant amplitude and varying phase is depicted in Fig. 2.11. The results shown in Fig. 2.12 show the Cardiff models extracted versus $\phi(A_{21})$ for different$|A_{21}|$ LUT values as described in (2.26)[1]. Unlike the PHD model, the Cardiff model can be scaled to predict any nonlinear order by specifying the n^{th} degree of the Fourier Series in (2.26). In Fig. 2.12, it can be seen that a low-order Fourier series could predict the DUT behaviour at low values of $|A_{21}|$, while higher-order equations are required to predict the nonlinear response at higher values. It has also been suggested that inter-model interpolation could be improved by using a top-level polynomial function

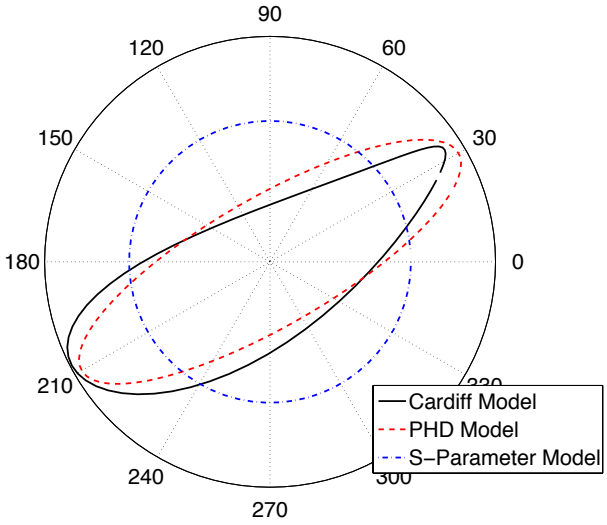


Figure 2.11: Comparison of S-Parameters, the PHD Model, and the Cardiff Model vs. $\angle a_{j,l}$

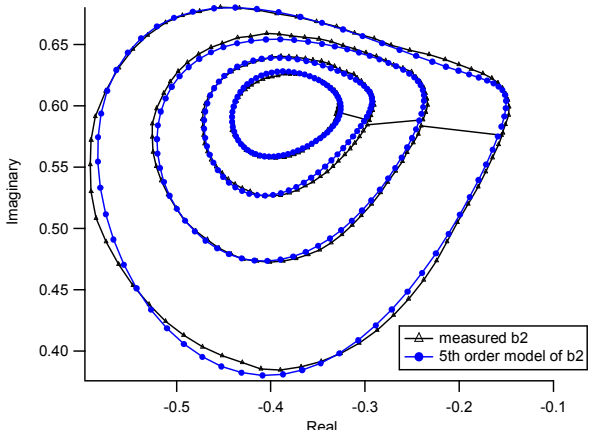


Figure 2.12: Fundamental Harmonic Cardiff Model as a function of A_{21}[1]

with respect to $|A_{21}|$ in 2.28[2].

$$B_{ik} = P^k \sum_m R_{ik,m} \left[|A_{11}|, |A_{21}| \right] \left(\frac{Q_1}{P_1} \right)^m \tag{2.26}$$

$$B_{ik} = P^k \sum_m \sum_r G_{ik,mr} \left[|A_{11}|, |A_{21}|, |A_{22}| \right] \left(\frac{Q_1}{P_1} \right)^m \left(\frac{Q_2}{P_1^2} \right)^r \tag{2.27}$$

Where,

A_{jl} is the large-signal incident travelling-wave applied at input j, harmonic l.

B_{ik} is the large-signal reflected travelling-wave measured at output i, harmonic k.

P_1 represents the phase of the input fundamental harmonic $\phi(A_{11})$.

Q_1 represents the phase of the output fundamental harmonic $\phi(A_{21})$.

Q_2 represents the phase of the output 2^{nd} harmonic $\phi(A_{22})$.

$R_{ik,m}$ is an m^{th} degree Fourier Series model that determines $B_{ik} = f(A_{21})$ as a function of the relative phase shift Q_1/P_1 at a constant $|A_{11}|$ and $|A_{21}|$ LUT entry.

$G_{ik,mr}$ is a two-dimensional $m^{th}Xr^{th}$ degree Fourier Series that determines $B_{ik} = f(\phi(A_{21}), \phi(A_{22}))$ as a function of the relative phase shift Q_1/P_1 and Q_2/P_1^2 at a constant $|A_{11}|$, $|A_{21}|$, and $|A_{12}|$ LUT entry.

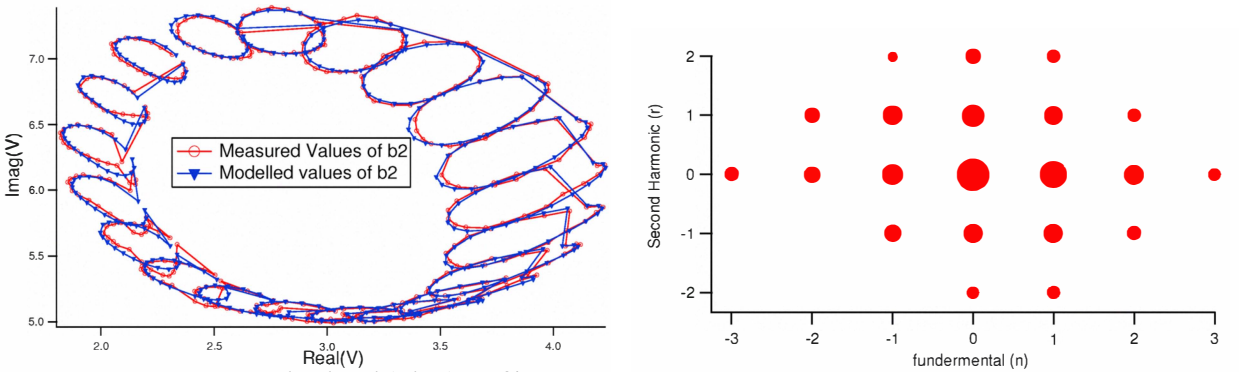
Each $R_{ik,m}$ in the LUT with respect to $|A_{11}|$ and $|A_{21}|$ (2.26) can additionally be described as a magnitude dependent polynomial represented by (2.28).

$$R_{ik,m} = \sum_{x=0}^{X} \alpha_{x-1} |A_{21}|^{x-1} \tag{2.28}$$

In order to extend the functionality of the Cardiff model to multi-harmonic, multi-port systems, a nested modelling formulation in (2.27) was proposed in [2]. In Fig. 2.2.2 the nonlinear response to simultaneous inputs, $\phi(A_{21})$ and $\phi(A_{22})$ is accurately predicted at constant input power magnitudes. Further analysis, shown in Fig. 2.13b, suggests

that cross-harmonic behavioural modelling is required to accurately model the response to the combined harmonic stimulus. This result indicates that the small-signal harmonic superposition assumption in the PHD model will provide limited accuracy for higher-order nonlinear systems, and it ultimately implies that behavioural models must be extracted when all harmonic input signals are varied simultaneously. While only a finite number of harmonics have been modelled using this formulation, it is suggested that this technique could be extended to model the response to an arbitrary number of input harmonics[49].

Although the Cardiff model extracts more coefficients in the MHV model than the PHD model, it separates the complex inputs into magnitude and phase inputs. This assumption is not perfect because the necessary order of the phase model in 2.26 increases with the magnitude of the input signal.



(a) Cardiff Model vs. A_{21} and A_{22}

(b) Harmonic Cross-Coefficient Sensitivity Analysis

Figure 2.13: Multi-Harmonic Cardiff Model as a function of A_{21} and A_{22}[2]

Although there are several existing nonlinear characterization solutions in the form of nonlinear measurement receivers and load-pull applications, they lack the simplicity and generality of nonlinear CAD simulation tools. A general-purpose nonlinear characterization system should contain flexible hardware support and it should conform to harmonic balance measurement science theory. To model the multi-harmonic behaviour of the DUT, we need to treat the problem as a MIMO system, where inputs are simultaneously applied at each harmonic, on each port. The synthesis of the MHV model is a generalization of the Volterra series when multiple input signals are presented a commensurate frequencies. Although the

Multi-Harmonic Volterra Series presents a complete description of a nonlinear system, the load-pull LUT, PHD model, and Cardiff model are progressively accurate approximations achieved through advancements in characterization systems and model synthesis. Although the PHD model has been commercialized as the Keysight X-Parameter model for system-level behavioural modelling, it is questionable whether a first-order expansion is a valid approximation of unmatched broadband transistors. Although the Cardiff model provides a higher-order expansion, it requires more complex instrumentation and may demand a much larger distribution of measurements. Therefore, Chapter 3 will investigate the intrinsic requirements of a multi-harmonic nonlinear characterization system and proposes an implementation that improves behavioural model extraction. Chapter 4 examines the limitations of the PHD, Cardiff and MHV models and it proposes a solution for improving the accuracy of the PHD model for unmatched devices. It also presents a higher-order model suitable for strongly nonlinear operating conditions based on the MHV model.

Chapter 3

Harmonic Balance Inspired Nonlinear Characterization System

This chapter proposes a modular Nonlinear Characterization System (NCS) and the required calibration routines. This system is carefully developed to enable seamless integration with CAD tools used for design simulation of PAs. It also enables the synthesis of nonlinear operating conditions to replicate the electro-thermal properties of the DUT. This chapter also studies the characterization requirements of high-power, unmatched, broadband nonlinear transistors and describes the building blocks, namely a mixer-based receiver augmented by multiple phase-coherent signal sources and multi-harmonic impedance tuners. Furthermore, a sequential calibration procedure is developed to ensure accurate nonlinear measurements and stimuli at the ports of the DUT, thus permitting a 1:1 mapping between measurement and simulation data. Several traditional and advanced measurement techniques are performed, to demonstrate that the NCS can be used to perform a comprehensive study of nonlinear devices.

Traditionally, nonlinear characterization systems come packaged with fixed measurement routines that perform a specific measurement task, such as load-pull. These turnkey solutions lack the general-purpose usability of the instruments they are built on. Conversely, Electronic Design Automation (EDA) tools are essential to the design of nonlinear circuits, however unlike linear circuits, the scalability of nonlinear circuit design is strongly

dependent on the accuracy of active device models. To achieve the desired accuracy of information, characterization, modelling and design must become tightly integrated to minimize the propagation of error. Hence, it is imperative that the NCS perform the following tasks:

1. Calibration at the non-connectorized ports of the DUT.
2. Synthesis of an exact operating condition.
3. Automated characterization based on parametric sweeps.
4. Model synthesis from datasets of measurement data.
5. Integration of the model into existing EDA simulation tools.

The NCS can achieve seamless integration between measurement and simulation by organizing the system architecture based on the Harmonic Balance simulator. Fig. 3.1 illustrates the system architecture of the Keysight Advanced Design System (ADS) that will be used in the design of the NCS.

The process flow diagram in Fig. 3.1 is built to match the generic structure of CAD circuit simulation environments. During any given measurement, a sequencer processes a list of sweep plans and sets the instruments to the next stimulus condition. Each consecutive measurement is saved in a dataset and the process is repeated until all measurements have been taken. Once the measurements are completed, the current and saved datasets can be post-processed using measurement equations, as well as custom plot functions.

Nonlinear characterization differs from linear characterization in that the principle of source superposition does not apply, hence all measurement ports theoretically must contain a dedicated signal source, both DC and RF. Similarly, nonlinear systems are impedance termination dependent, hence they must contain an impedance control circuit on each port. As a nonlinear system produces spectral content at multiple harmonics of the input signal frequency, all instruments must provide stimuli and measurements at multiple harmonic frequencies. A high-level equivalent block diagram of a nonlinear characterization system is described in Fig. 3.2, and is distinguished from a linear characterization system in the following regards:

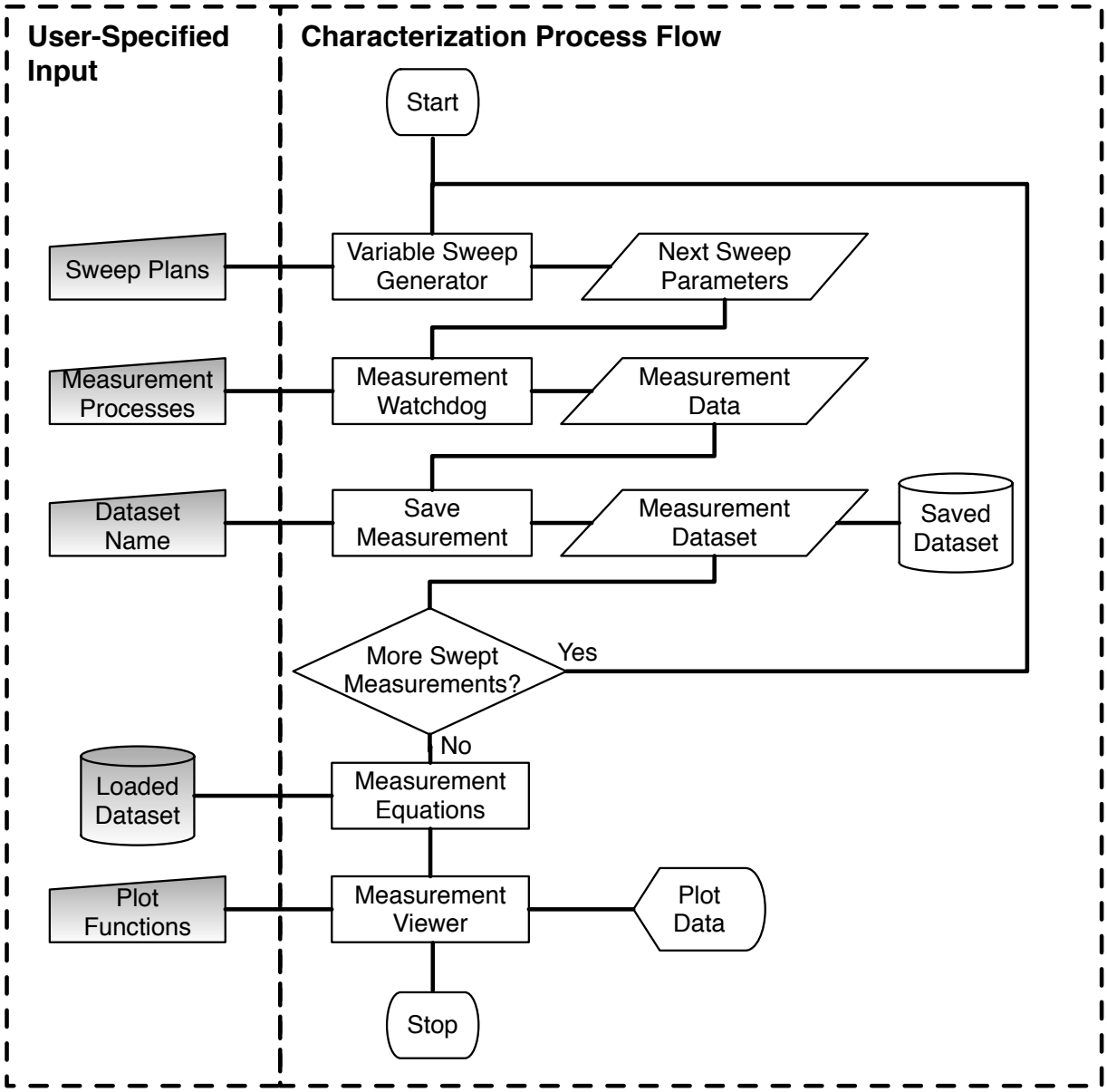


Figure 3.1: NCS Process Flow Diagram

1. A DC Receiver is added to measure the DC bias and the response to baseband intermodulation products.

2. A DC Source must be added to control the DC bias of the transistor.

3. A Multi-Harmonic RF Receiver must characterize the DUT at n harmonics of the fundamental frequency.

4. A Multi-Harmonic RF Source must be injected on each port, at every harmonic. As multiple signals are being applied during a vector measurement, each source must be phase coherent.

5. A Multi-Harmonic RF Impedance Tuner is needed to control the port impedance, because nonlinear circuits are match dependent.

Due to it's inherit hardware complexity, a nonlinear characterization system must be implemented in a modular fashion so that a system can be implemented based on resource availability and affordability. Regardless of which receiver architecture (Section 2.1.1), or which load-pull technique (Section 2.1.2) is chosen, a simplified description of the system is described in Fig. 3.2. Hence, a hardware dependent characterization system is avoided through the means of hardware abstraction, and hardware specific routines are avoided by organizing the test framework in a similar fashion to CAD simulation tools.

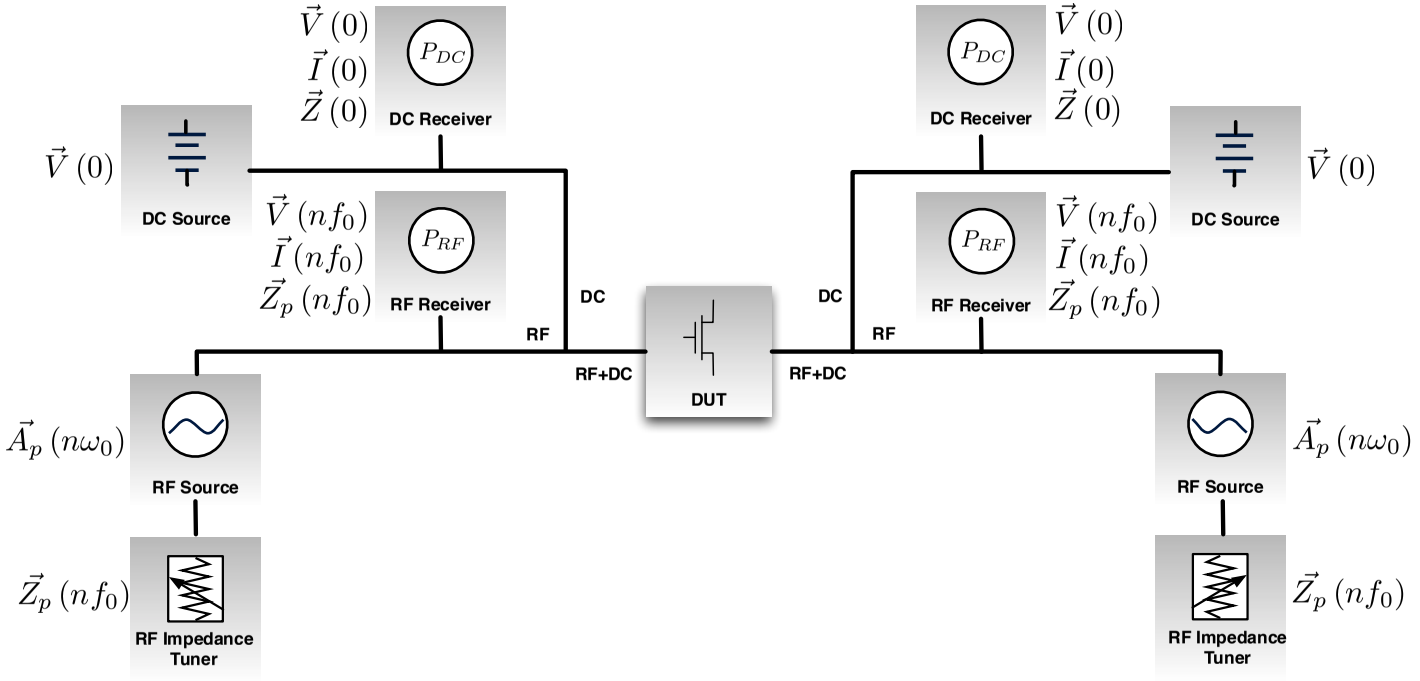


Figure 3.2: High-Level NCS Hardware Abstraction

3.1 Proposed Test-bench Configuration

Based on the system architecture outlined above, the goal of this research was to construct a test-bench that could be used to study high-power, unmatched, and broadband transistors, operating under large-signal RF stimulus conditions. The measurement objectives of the test-bench are listed below and are presented in the following subsections.

- High-power measurements to support high-power transistors and PAs that are the target of this research.
- Unmatched measurements to optimize performance of a PA when designing the matching networks.
- Multi-harmonic stimulus and receiver to enable multi-harmonic efficiency enhancement techniques of broadband devices.
- Stability enhancement to minimize or eliminate low-frequency oscillation during measurement extraction..

Unmatched devices produce output power-waves at both the gate and drain of the transistor (multiple ports), while multi-harmonic matching networks variably reflect the outputs back into the system at multiple harmonics. Therefore, the DUT must be treated as a MIMO system, and a mixer-based vector measurement architecture was modified to meet the aforementioned measurement objectives.

3.1.1 High-Power Devices

A VNA constructs small-signal S-parameters by injecting a low-power RF signal sequentially at each port of the DUT, while measuring the electrical response on each receiver. Although S-parameters fully describe the performance of a linear device, the measurement response of nonlinear devices is dependent on input power, therefore modifications must be made to support high-power measurements. This is accomplished by adding PA drivers to amplify the signal source, while adding attenuation to reduce the output power entering the receivers in accordance to the following considerations:

1. The measured power must remain within the linear power range of the receivers.
2. The signal-source and PA driver amplifiers must operate in their linear power range.
3. In general, high-power circuits restrict the bandwidth of the test-bench and may significantly alter the 50Ω measurement port impedance outside of the supported bandwidth. When connecting an impedance tuner or impedance transformer, this could result in an undesirable impedance transformation.
4. An isolator (or attenuation) is often connected to the port to provide an ideal match to the DUT when poorly matched PA driver amplifiers are used in the system

All test and measurement applications demand a controlled experimental environment where any spurious behaviour of the test-bench is prevented or corrected. While a high-power linear measurement limits the maximum power to the receiver 0.1dB compression point, a nonlinear characterization must operate far below this specification. The actual power-limit is empirically determined by ensuring the linearity of the receiver provides a sufficient harmonic measurement dynamic range. When performing nonlinear characterization it is important to insert PA drivers between the signal-source and the receiver couplers because linear measurement de-embedding techniques will produce inaccurate measurements if the drivers operate in power-compression. This optionally enables the receiver to be calibrated at lower power without any driver amplifiers connected. Although Isolators provide the DUT with a good match at the design frequency, they are narrow bandpass devices that may cause oscillation at low-frequencies, and they present a non-50Ω impedance at harmonic frequencies. The mismatch of the port at multiple harmonics may cause the measurements to disagree with simulations that used broadband 50Ω terminations, hence the match of the port is determined during the receiver vector calibration (and port mismatch calibration), and should be stored in the measurement dataset. Further information regarding high-power modifications can be found in [50].

3.1.2 Unmatched Impedance Termination

While the characterization of linear devices is independent of port impedance Z_p, the performance of nonlinear devices depends on the port impedance at the fundamental frequency as well as the harmonic frequencies. For matched devices, the multi-harmonic port impedance is fixed to $Z_p = 50\Omega$ because this is also the port impedance of the final design. When characterizing a transistor, the final port impedance is unknown, therefore it is desirable to measure the transistor under different harmonic port impedance conditions using multi-harmonic source/load impedance tuners. As travelling waves (a and b) are proportional to the characteristic impedance (Z_0) of the port, it is ideal to use a more general definition called power-waves (3.3-3.4) [51] where $Z_0 = Z_p = Z_{Harmonic,Port}$, a variable characteristic impedance with respect to the measurement port and harmonic frequency. Only under these conditions, will the incident power-wave (a_p) and reflected power-wave (b_p) represent the incident and reflect power at the DUT, however post correction is necessary to calculate the power-waves using (3.1-3.4) based on explicit knowledge of the port impedance(Z_p). When modifying the test-bench to support unmatched DUTs, it is traditionally optimal to minimize the insertion loss between the impedance tuner and the DUT, however recent developments in Active Load-Pull technology have made this decision complicated and two alternative configurations are presented in Table 3.1 [52][53].

$$v = \frac{1}{\sqrt{\Re\left(Z_0\right)}}\left(a + Z_0^* b\right) \tag{3.1}$$

$$i = \frac{1}{\sqrt{\Re\left(Z_0\right)}}\left(a - b\right) \tag{3.2}$$

$$a_p = \frac{1}{2\sqrt{\Re\left(Z_p\right)}}\left(v + Z_p i\right) \tag{3.3}$$

$$b_p = \frac{1}{2\sqrt{\Re\left(Z_p\right)}}\left(v - Z_p^* i\right) \tag{3.4}$$

As high-power transistors often have small input/output impedances, the first solution would appear to be highly desirable, however by supplementing the second solution with active load-pull and a tapered-line $\lambda/4$ impedance transformer it is possible to have the accuracy of the second solution without any drawbacks. While the second solution is the

Table 3.1: Alternative Passive Load-Pull Configurations

Configuration	Coupler ⇔ Tuner ⇔ DUT	Tuner ⇔ Coupler ⇔ DUT
Advantages	• Maximizes the achievable range of Γ_{Port} using Passive Load-Pull • Nearly eliminates the need for hybrid Passive/Active Load-Pull to generate high reflection coefficients	• Maximizes the measurement accuracy of a_p and b_p • May not require an impedance tuner calibration
Disadvantages	• Error in a_p and b_p is proportional to the error/repeatability of the tuner calibration • De-Embedding coefficients must be updated when the tuner impedance is repositioned • Test-bench reconfiguration during calibration procedure	• Smaller achievable range of Γ_{Port} using Passive Load-Pull • Requires hybrid Passive/Active Load-Pull to achieve high reflection coefficients for high-power devices

best measurement solution, it relies heavily on iterative Active Load-Pull techniques that were beyond the scope of this research. The first solution is much simpler to implement and provides an approximate measurement accuracy of $-30dB$ to $-40dB$.

3.1.3 Multi-Harmonic Signal Injection

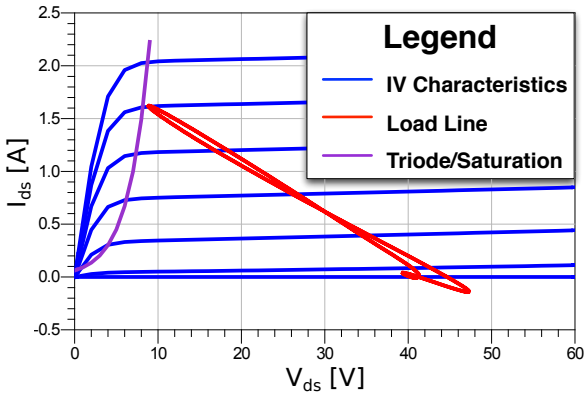
As described in section 2.1.2, Active Load-Pull provides faster electrical measurement reconfiguration compared to the mechanical Passive Load-Pull. While most characterization applications do not benefit greatly from this additional speed, the formulation and extraction of behavioural models is severely impacted as they require significantly more information. For this application, signal-sources must be applied to each harmonic and port of the DUT and this could be very expensive. Depending on the design frequency, it may be possible to use high-frequency arbitrary waveform generators (AWGs) to simultaneously generate multiple harmonic sources connected to each measurement port. Some of the important considerations when building a multi-harmonic Active Load-Pull system are listed as follows:

- The baseband, fundamental RF, and harmonic RF frequency bands represent three bandpass systems distinguished by a method of signal generation and maximum power requirements. The power requirements of the harmonic PA driver is significantly lower than the fundamental PA driver.
- Conjugately matching the DUT over all harmonics using an impedance tuner or static matching network will minimize the power requirements of PA drivers by ensuring the maximum power is delivered to the DUT. It will also ensure that the incident power is equal to the delivered power to the DUT, which is useful for modelling purposes.
- Isolators, while typically narrowband, increase the power-range of PA drivers by minimizing the load impedance modulation seen by the PA drivers. They also provide a consistent 50Ω matching condition to the DUT at a specific frequency.

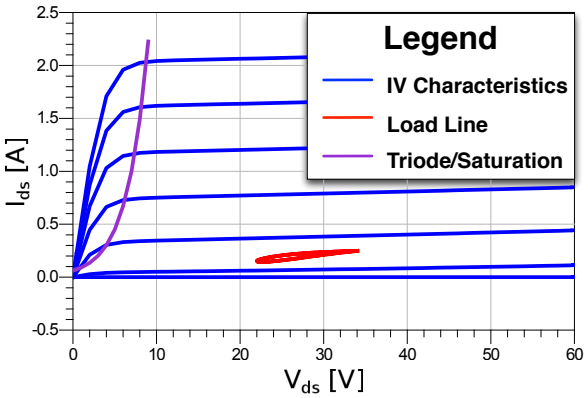
It may seem unintuitive to use narrowband isolators on a multi-harmonic signal source, however Fig. 3.3 demonstrates the load modulation experienced by a driver PA whose drain is connected to the drain of the DUT. Fig. 3.3a shows the load line induced by a signal injected into the input of the driver, while Fig. 3.3b shows the load line induced by a signal injected into the output. If the input and output injected signals are uncorrelated, the composite load line (Fig. 3.3c) will cover an extensive range of the IV Characteristic plot, thus increasing the chances of operation in cut-off or triode where the driver becomes nonlinear. Just like the receiver, it is vital to ensure the signal-source is linear so that it does not contaminate the nonlinear measurement of the DUT. It is better to multiplex isolated frequency bands than it is to operate a single driver at back-off, because the cost of broadband PAs increases dramatically with power, and the power requirements of the harmonic frequencies are typically $> 10dBc$ relative to the fundamental frequency.

3.1.4 Stability

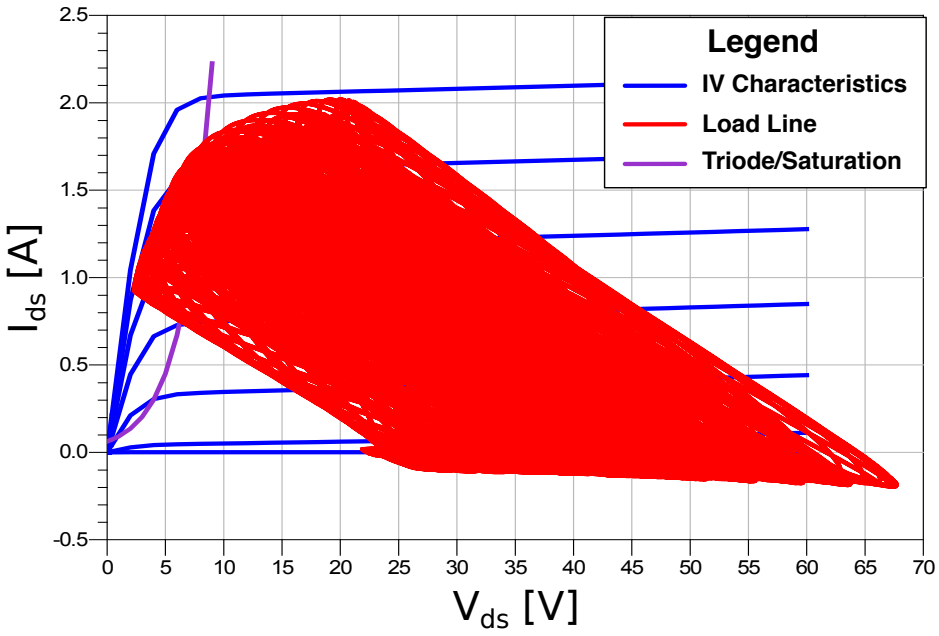
Transistors are susceptible to oscillation over the entire frequency spectrum, most commonly lower-frequencies due to an increased available gain. As a mixer-based harmonic receiver only measures a small subset of the frequency spectrum, it is easier to detect an oscillation using a spectrum analyzer. Due to the complexity of the test-bench apparatus, many considerations are made during component selection to ensure that the DUT remains stable during characterization. Providing a conjugate match at all frequencies would eliminate any chance of oscillation, however multi-harmonic impedance tuners provide narrowband tuning to several harmonic frequencies, and their physical dimensions prevent them from operating at such low frequencies. Although very low frequency $< 100MHz$ oscillations must be mitigated in the design of the DC bias network, moderately low frequency oscillations must be mitigated in the RF matching network. While narrowband multi-harmonic impedance tuners have unpredictable impedance configurations outside their operating frequencies, they must intrinsically degenerate to a 50Ω transmission line at low frequencies, whereas narrowband isolators degenerate to $50\Omega//50\Omega = 25\Omega$ ($\Gamma_S = -0.333$) at low-frequencies. Therefore, most oscillation problems can be attributed to the isolator when $\mu_{Source} > 0$[54]. Although stability can be guaranteed by inserting a stability net-



(a) Driver Response to Input Stimulus



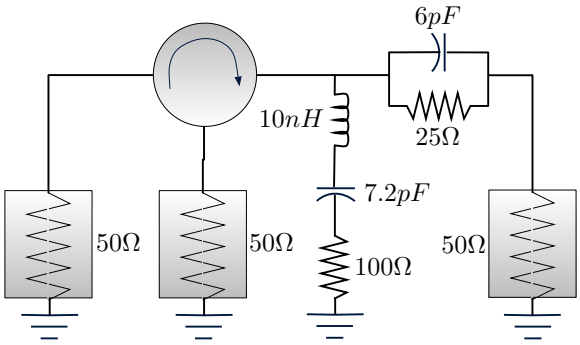
(b) Driver Response to Output Stimulus



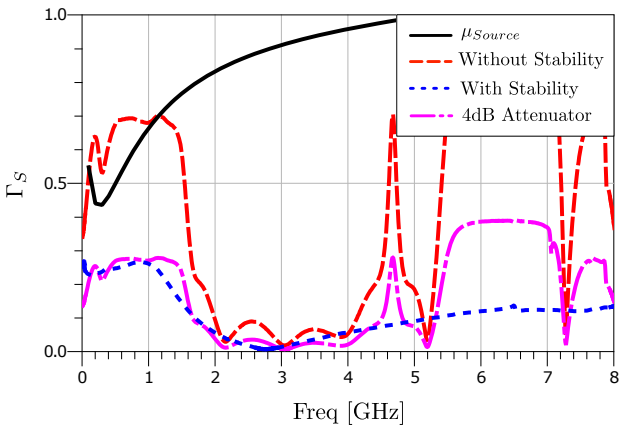
(c) Combined Driver Response to Uncorrelated Input and Output Stimuli

Figure 3.3: Driver PA Dynamic Load-Line vs. Input and Output Stimuli

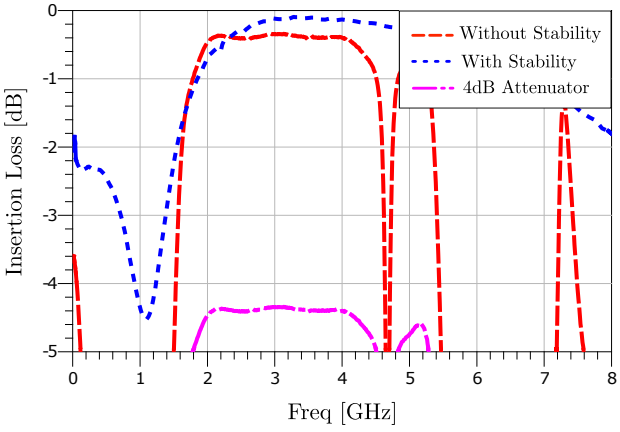
work directly at the input of the transistor, this would dramatically increase the insertion loss between the source impedance tuner and the DUT. Alternatively, if the oscillation is purely attributed to the isolator, the stability network can be connected to the output of the isolator, in the $Z_0 = 50\Omega$ transmission path, as demonstrated in Fig. 3.4a. Fig. 3.4b overlays the magnitude of the source reflection coefficient that could induce oscillation (μ_{Source}) with the reflection coefficient of the circuit in Fig. 3.4a. The stability network is designed to match the isolator to 50Ω at low-frequencies (ensuring that $|\Gamma_S| < \mu_{Source}$), while minimizing insertion loss at higher frequencies.



(a) Stability Circuit Schematic



(b) Source Reflection Coefficient

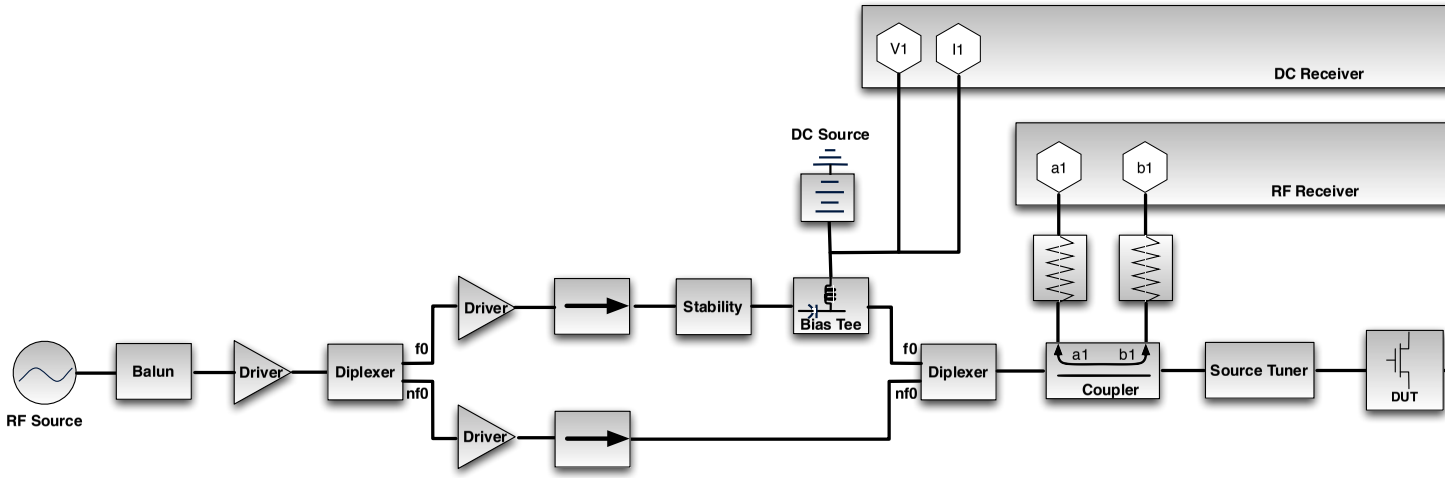


(c) Insertion Loss of Stability Network

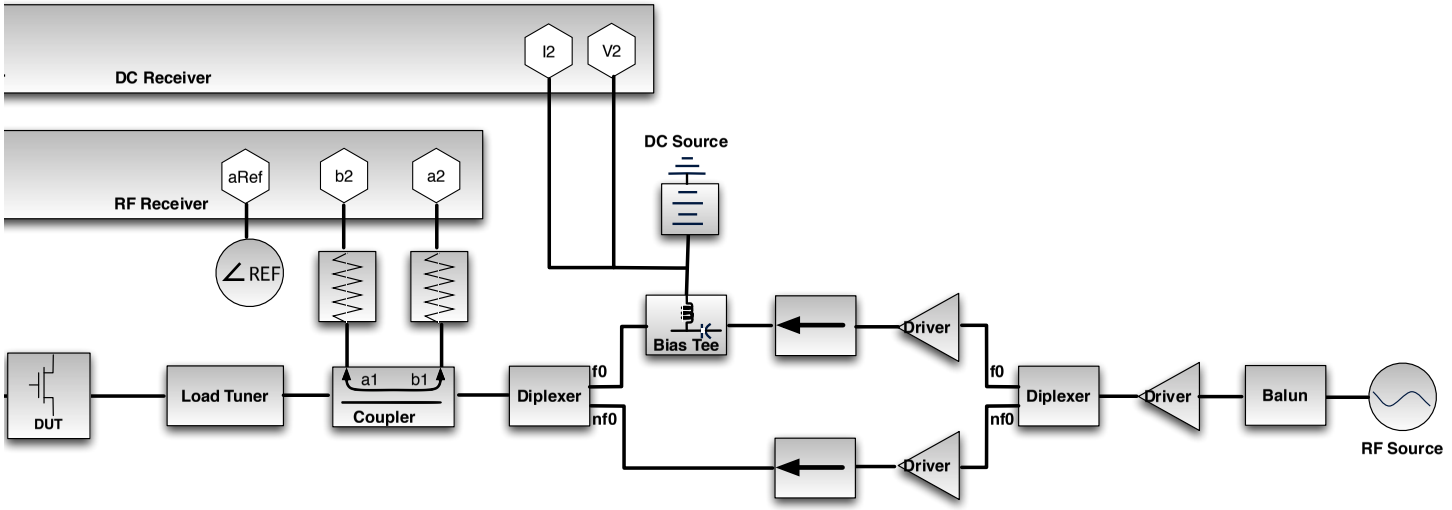
Figure 3.4: Improvement in Low-Frequency Matching with Stability Network

Based on the requirements proposed in this section, the 2-port nonlinear test-bench is

constructed as illustrated in Fig. 3.5. The assembled test-bench, shown in Fig. 3.6 is built on the considerations described in this section that are needed to measure high-power, unmatched, multi-harmonic nonlinear devices.



(a) Port 1



(b) Port 2

Figure 3.5: Proposed High-Power Unmatched Multi-Harmonic NCS Test-bench

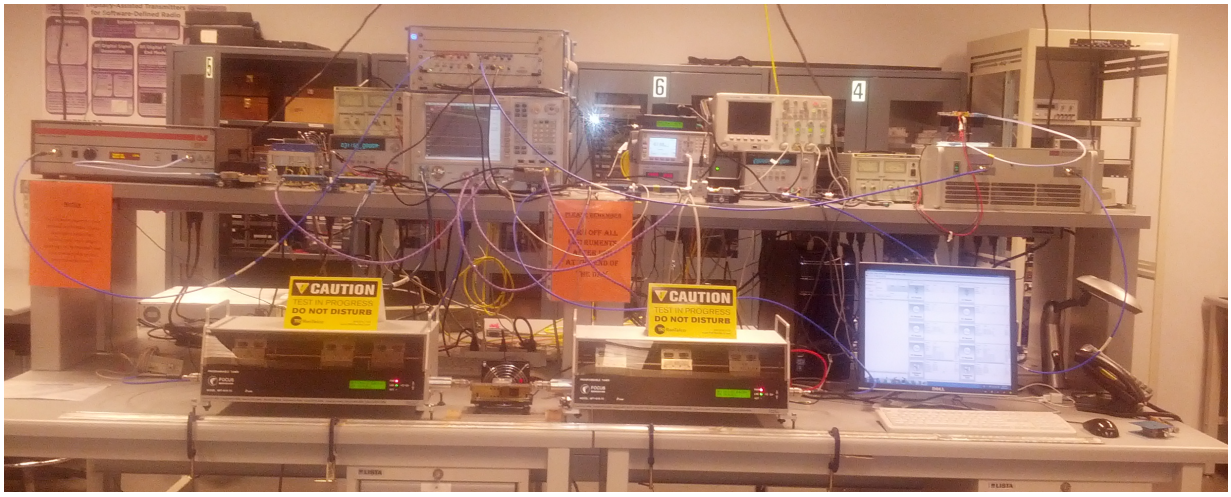


Figure 3.6: Final Nonlinear Characterization Test-Bench

3.2 Calibration and Measurement De-Embedding

The relatively large electrical length of the test-bench components and the relative phase coherency requirement needed to characterize MIMO systems, demands a sequence of calibration routines that are designed to de-embed an equivalent RF Source, RF Receiver and RF Impedance precisely to the DUT measurement plane. A set of calibration routines is described using the error model, shown in Fig. 3.7. The error-model primarily consists of:

- T_{RF} a multi-harmonic cascaded Transmission Matrix which defines the 2-port error between the RF Receiver and the calibration measurement plane.
- T_F a multi-harmonic cascaded Transmission Matrix which defines the 2-port error between the calibration plane and the DUT measurement plane (de-embedding).
- Γ_{SRC} a multi-harmonic reflection coefficient looking into the RF Source from the calibration measurement plane.

- G_{SRC} a multi-harmonic vector gain between the RF Source and the calibration measurement plane.

Since the error model contains *linear* correction coefficients, it is vital that the test-bench operates in a highly linear state during calibration and during measurement. During each of the calibration routines described below, two-port networks can be de-embedded, and the calibration routine will automatically remove these networks from the correction data. Prior to calibration, the user should calibrate the impedance tuners using the impedance tuner software. Some of the calibration routines rely on previous correction data, therefore it is advised that the calibrations be performed in the order of which they are introduced.

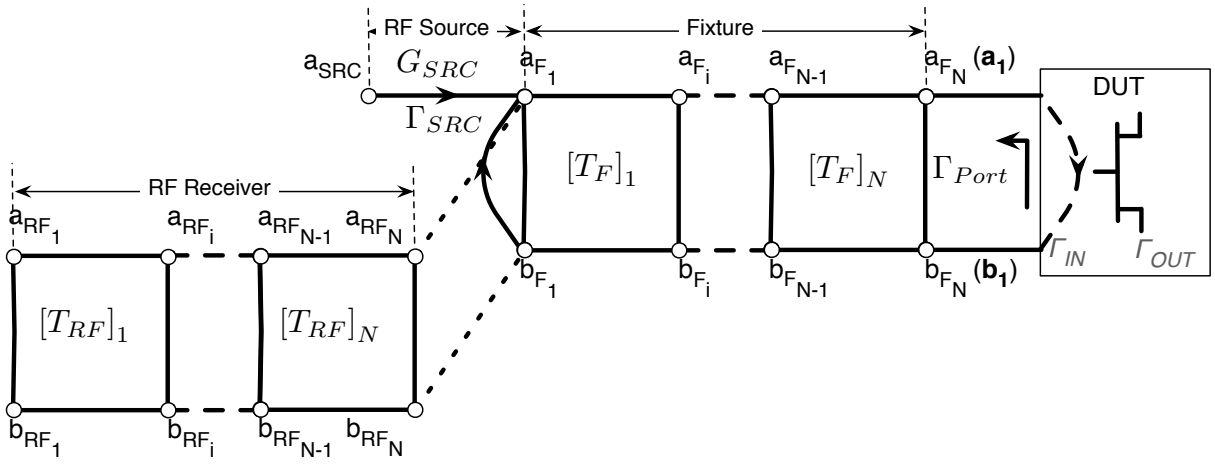


Figure 3.7: Error Model of NCS Correction Coefficients

3.2.1 RF Receiver Vector Calibration

The receiver calibration is a three step process described in [55] that consists of: 1. RF Receiver Vector Calibration, 2. RF Receiver Phase Calibration, and 3. RF Receiver Amplitude Calibration. The calibration terms collected during the Receiver calibration are stored in the T_{RF}, a 2x2 matrix that is swept versus port j, harmonic l. All receiver measurements are corrected using equation (3.5).

$$\left[\begin{array}{c} \boldsymbol{a} \\ \boldsymbol{b} \end{array} \right]_{jl} = \left[\left[T_{RF} \right]_{jl} \left[T_F \right]_{jl} \right]^{-1} \left[\begin{array}{c} a_{RF} \\ b_{RF} \end{array} \right]_{jl} \tag{3.5}$$

Where,

j is the port index.

l is the harmonic index.

a_{RF} is the incident travelling-wave at the receiver measurement plane.

b_{RF} is the reflected travelling-wave at the receiver measurement plane.

$\boldsymbol{a}$ is the incident travelling-wave at the DUT measurement plane.

$\boldsymbol{b}$ is the reflected travelling-wave at the DUT measurement plane.

T_F is the Transmission Parameters between the calibration measurement plane and the DUT measurement plane.

T_{RF} is the Transmission Parameters between the RF Receiver and the calibration measurement plane. These transfer parameters are defined as separate coefficients as follows:

$$\left[T_{RF} \right]_{jl} = \left[\begin{array}{cc} \alpha & \beta \\ \gamma & \delta \end{array} \right]_{jl}^{-1} \tag{3.6}$$

The RF Receiver Vector Calibration is the same method used to calibrate a VNA and multiple techniques have been developed[56], most of which are implemented in METAS VNA Tools II[57] or StatisiCAL[58]. Implementing this calibration using external software circumvented the S-Parameter friendly port designations, allowing direct connection to the five measurement receivers of the PNA-X without mechanical switches. As a result, the measurement accuracy and speed was improved, and the RF Impedance Mismatch

calibration was easier to implement. The goal of the vector calibration routines is to calculate the components of T_{RF} as shown in (3.6) using S-Parameter measurements of known reflection and transmission standards. If we assume that we are performing a 2-port calibration, then a 16-term 2-port error model describes all of the possible error terms in (3.7), where the entire matrix is normalized by α_1 (K_l).

$$\begin{bmatrix} a_{1F} \\ b_{1F} \\ a_{2F} \\ b_{2F} \end{bmatrix}_l = K_l \begin{bmatrix} 1 & \beta_1 & 0 & 0 \\ \gamma_1 & \delta_1 & 0 & 0 \\ 0 & 0 & \alpha_2 & \beta_2 \\ 0 & 0 & \gamma_2 & \delta_2 \end{bmatrix}_l \begin{bmatrix} a_{1RF} \\ b_{1RF} \\ a_{2RF} \\ b_{2RF} \end{bmatrix}_l \tag{3.7}$$

The zeros in (3.7) are an assumption that that there is no cross-port leakage between the multi-port receivers. The terms inside the matrix are mathematically sufficient for measuring the relative difference between a and b travelling waves at the calibration reference plane, however the K_l term is required to measure the absolute value of the travelling-waves independently. The inside matrix terms (β, γ, and δ) are solved by measuring three reflection standards ($\Gamma^{(1)}$, $\Gamma^{(2)}$, and $\Gamma^{(3)}$) on each port, equating $\Gamma a_{jF} = b_{jF}$ and by substituting in terms of the raw measurements, a_{RF} and b_{RF}[52].

$$\begin{bmatrix} \beta \\ \gamma \\ \delta \end{bmatrix}_{jl} = \begin{bmatrix} \Gamma^{(1)}b^{(1)}_{RF} & -a^{(1)}_{RF} & -b^{(1)}_{RF} \\ \Gamma^{(2)}b^{(2)}_{RF} & -a^{(2)}_{RF} & -b^{(2)}_{RF} \\ \Gamma^{(3)}b^{(3)}_{RF} & -a^{(3)}_{RF} & -b^{(3)}_{RF} \end{bmatrix}_{jl} \begin{bmatrix} -\Gamma^{(1)}a^{(1)}_{RF} \\ -\Gamma^{(2)}a^{(2)}_{RF} \\ -\Gamma^{(3)}a^{(3)}_{RF} \end{bmatrix}_{jl} \tag{3.8}$$

A transmission measurement is then used to calculate α_2 by equating $a_{1F} = b_{2F}$ and by substituting in terms of the raw measurements, a_{1RF} and b_{2RF}, as shown in

$$\alpha_{2l} = \left[\frac{a_{1RF} + \beta_1 b_{1RF}}{\gamma_2 a_{2RF} + \delta_2 b_{2RF}}\right]_l \tag{3.9}$$

The Receiver Vector Calibration provides enough correction to measure S-Parameters, hence the calibration can be verified by measuring the S-Parameters of a known passive component. Often it is easiest to measure the S-Parameters of a flush thru, as shown in Fig. 3.8, however a proper verification should be done by performing a measurement that was not used during the calibration process.

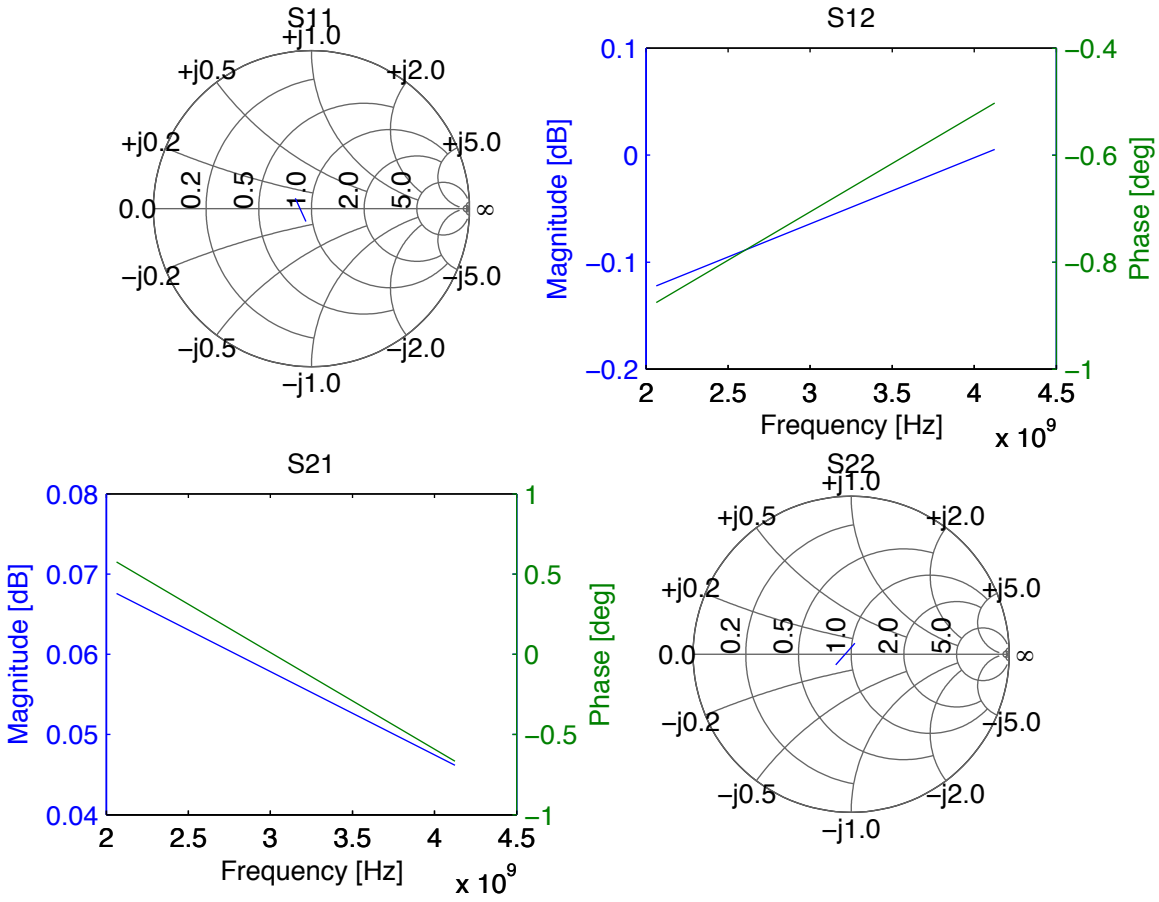


Figure 3.8: Receiver Vector Calibration Thru Measurement Verification

3.2.2 RF Receiver Phase Calibration

The receiver phase calibration solves for the phase of the K_l term, an arbitrary frequency-dependent linear phase shift whose time-domain equivalent is a linear time delay that aligns the multi-harmonic measurements at a single time reference. A multi-harmonic phase reference is connected to one of the vector calibrated ports and the phase of K_l is solved using the signal flow diagram in Fig. 3.9, by substituting in the terms of the raw measurements (a_{1RF} and b_{1RF}) using (3.10) (assuming calibration on port 1)[55]. This calibration routine needed to be implemented because the RF Receiver Vector calibration was implemented externally.

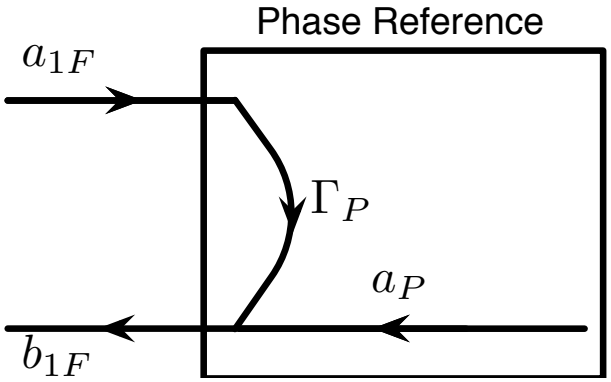


Figure 3.9: Phase Calibration Signal-Flow Diagram

$$\angle K_l = \left[\phi\left(\frac{a_P}{(\gamma - \Gamma_P)\, a_{1RF} + (\delta - \Gamma_P \beta)\, b_{1RF}}\right)\right]_l \tag{3.10}$$

3.2.3 RF Receiver Amplitude Calibration

The receiver amplitude calibration solves for the amplitude of the K_l term that is required to measure the absolute values of the a and b waves separately. An amplitude reference (power meter) is connected to one of the vector calibrated ports and the amplitude of K_l is solved using the signal flow diagram in Fig. 3.10 by substituting in the terms of the

raw measurements (a_{1RF} and b_{1RF}) using (3.11) (assuming calibration on port 1)[55]. The value of Γ_A can be obtained by measuring the reflection coefficient looking into the power meter, while T_A is taken from the power meter insertion loss table. This calibration routine needed to be implemented because the RF Receiver Vector calibration was implemented externally.

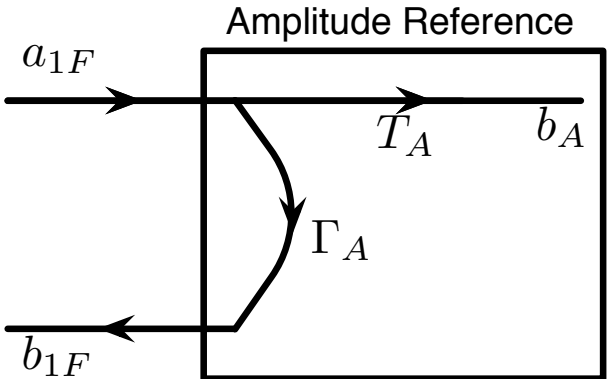


Figure 3.10: Amplitude Calibration Signal-Flow Diagram

$$|K_l| = \left[\left|\frac{b_A}{T_A\left(a_{1RF} + \beta b_{1RF}\right)}\right|\right]_l \tag{3.11}$$

If the receiver phase and amplitude calibration is performed on one of the measurement ports (usually the port with the lowest attenuation and highest dynamic range), than the correction can be applied to all ports by multiplying the T_{RF} matrices by K_l. Once both the amplitude and phase of K_l has been determined, the receiver calibration can be verified by measuring a flush thru, looking at the unratioed incident and reflected travelling-waves where both ports are normalized to the same reference impedance ($Z_0 = 50\Omega$). This verification requires that a phase-coherent multi-harmonic source is applied to one or more ports simultaneously to ensure that the relative phase/amplitude can be measured between multiple harmonics as shown in Fig. 3.11. The user should verify that $a_{1F} = b_{2F}$ and $a_{2F} = b_{1F}$, and that the phase difference between all harmonics remains constant over multiple measurements. The measurement traces in Fig. 3.11 do not overlap because neither port is terminated in exactly 50Ω and we do not have the means to generate a

conjugate match until the impedance tuners have been inserted in to the system. The variation in response overlap in Fig. 3.11 was caused by the non-50Ω output impedance of the driver PAs ($\Gamma_{Port} < 0.15$ at f_0 and $\Gamma_{Port} < 0.25$ at $2f_0$).

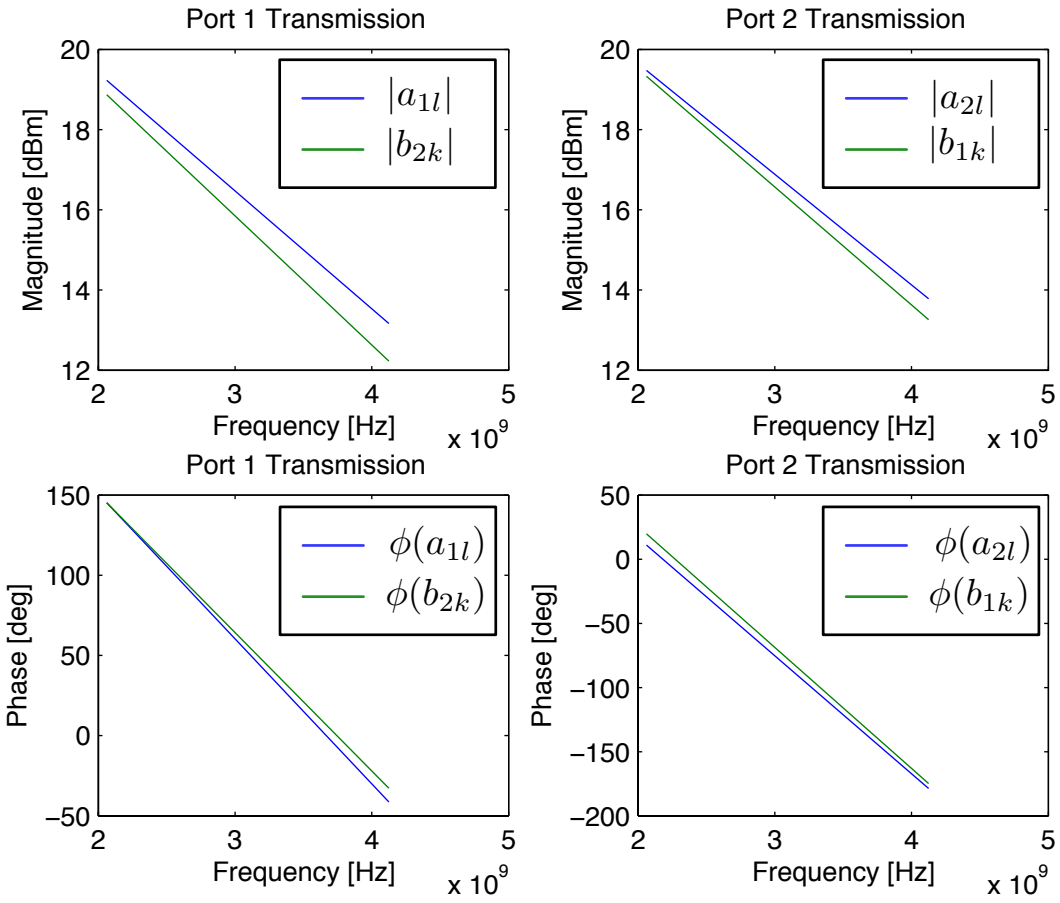


Figure 3.11: Receiver Calibration Verification

3.2.4 RF Port Mismatch Calibration

The RF Port Mismatch Calibration is included in the RF Receiver Vector Calibration, however it is typical to add high-power drivers to boost the signal after the receiver has

been calibrated. Adding high-power drivers will often result in poorly matched devices that do not present the same port impedance that was presented during the receiver calibration. This data is used by the Impedance Tuner to generate the specified port impedance and it is used by the RF Receiver to normalize the power-waves correctly. This calibration determines the Γ_{SRC} complex correction coefficient using (3.12), and the measurements are corrected using (3.13). This calibration routine needed to be implemented because the accurate calculation of both RF Source and RF Receiver power-waves demanded highly accurate knowledge of the port impedance, and de-embedding of the RF impedance tuner is very sensitive to port match error when synthesizing high reflection coefficients.

$$[\Gamma_{SRC}]_{jl} = \left[\begin{array}{cc} T_F(1,1) + \frac{T_F(1,2)}{T_F(2,1)} & T_F(2,2) \end{array} \right]_{jl} \left[\begin{array}{c} \Gamma_{Port} \\ 1 \end{array} \right]_{jl} \tag{3.12}$$

$$[\Gamma_{Port}]_{jl} = \left[\frac{T_F(1,2) - T_F(2,2)\Gamma_{SRC}}{T_F(2,1)\Gamma_{SRC} - T_F(1,1)} \right]_{jl} \tag{3.13}$$

Where,

j is the port index.

l is the harmonic index.

Γ_{Port} is the reflection coefficient looking into the DUT measurement port.

Γ_{SRC} is the reflection coefficient looking into the RF Source

T_F is the Transmission Parameters between the calibration measurement plane and the DUT measurement plane.

The port mismatch calibration is verified by specifying a port impedance and then measuring the reflection coefficient looking into the port with an S-parameter measurement as shown in Fig. 3.12.

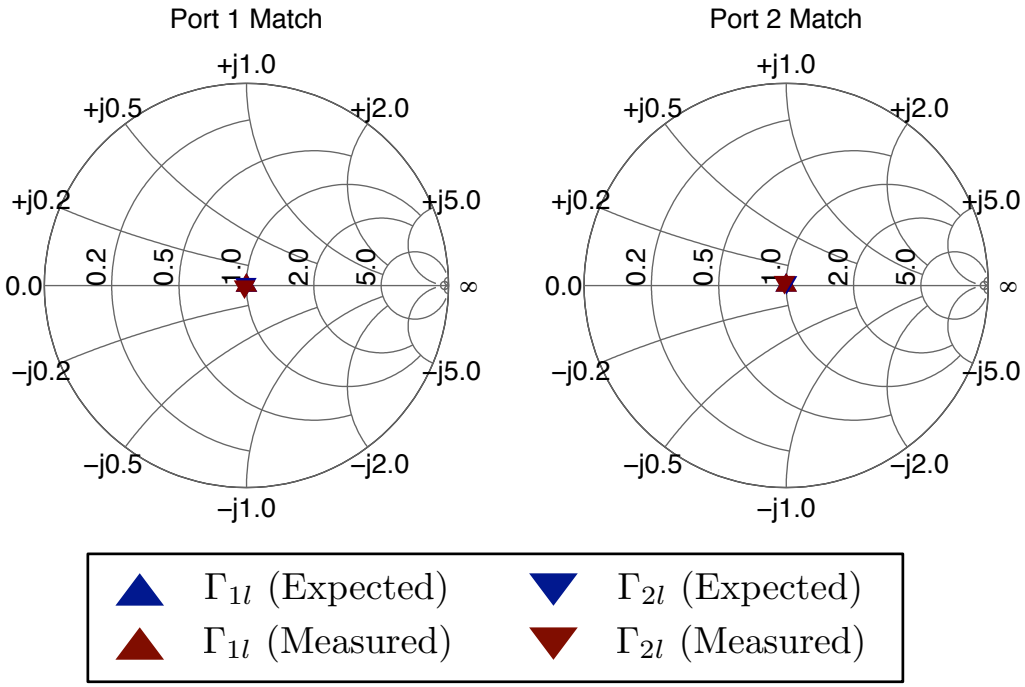


Figure 3.12: Port Mismatch Calibration Verification

3.2.5 RF Source Vector Calibration

The RF Source Vector Calibration requires an RF Receiver Calibration and RF Port Mismatch Calibration, and it is used for applications that require a specific RF Source power stimulus at the DUT reference plane. Since the RF Receiver is already calibrated to accurately measure the incident power-wave, it may seem unimportant to also correct the RF Source gain/loss, however this is essential because active load-pull demands multi-harmonic and multi-port phase coherent RF sources. Even when randomizing the RF Source power sweep, it is important to ensure that the source power conditions do not deviate from the chosen modelling region. This calibration determines the G_{SRC} complex correction coefficient using (3.14), and the source power is corrected using (3.15) by the RF Source. This calibration needed to be implemented because a scientific comparison of behavioural models demands accurate and repeatable synthesis of large-signal operating conditions.

$$[G_{SRC}]_{jl} = \left[T_{p,F}(1,1)\frac{a_{p,F}}{a_{p,SRC}}\right]_{jl} \tag{3.14}$$

$$[a_{p,SRC}]_{jl} = \left[\frac{T_{p,F}(1,1)}{G_{SRC}}a_{p,F}\right]_{jl} \tag{3.15}$$

Where,

j is the port index.

l is the harmonic index.

$a_{p,SRC}$ is the corrected incident power-wave at the DUT measurement port.

$a_{p,F}$ is the uncorrected incident power-wave at the RF Source port.

$T_{p,F}$ is the two-port *generalized* Transmission Parameters between the calibration measurement plane and the DUT measurement plane.

G_{SRC} is the vector gain between the RF Source port and the calibration measurement plane.

If the DUT fixture includes an impedance tuner, the value of $\Gamma_{Port} \neq 0$. Therefore, T_F must be converted to generalized transmission parameters ($T_{p,F}$) to calculate the source-power correction each time the port impedance is modified. This calculation is described in (3.16-3.18).

$$S_F = s2t\left(T_F\right) \tag{3.16}$$

$$S_{p,F} = s2s_p\left(S_F, \Gamma_S = \Gamma_{SRC}, \Gamma_L = \Gamma^*_{Port}\right) \tag{3.17}$$

$$T_{p,F} = s2t\left(S_{p,F}\right) \tag{3.18}$$

The source vector calibration is verified by specifying a complex incident power-wave and then ensuring that the same power is measured by the corrected incident power-wave receiver as shown in Fig. 3.13. Due to the fact that the ports are conjugately matched using the impedance tuners and the fact that the source amplifiers amplify the signal so that they use the maximum dynamic range of the receivers, the verification results in Fig. 3.13 show a significant improvement to the results in Fig. 3.11.

3.2.6 Additional Calibrations

The calibration menu in the proposed characterization system is meant to contain an ordered list of required and optional calibration routines. The NCS treats a "calibration routine" as any automated measurement routine that collects information about the system or the DUT and pre-conditions the test-bench for an automated measurement. A calibration routine has been added to empirically measure the receiver noise floor in the presence of noise that is generated by the RF sources and the DUT itself. Another calibration was added to measure the nonlinear compression of the test-bench when the input power levels are swept (AM-AM, AM-PM). Although this calibration procedure cannot differentiate between the compression of the RF sources or the RF receivers, it can be used to generate warnings when the received power-level exceeds a "system-wide" compression threshold. This would help ensure the integrity of each measurement, because the test-bench cannot differentiate between the nonlinearity of the test-bench and the nonlinearity of the DUT.

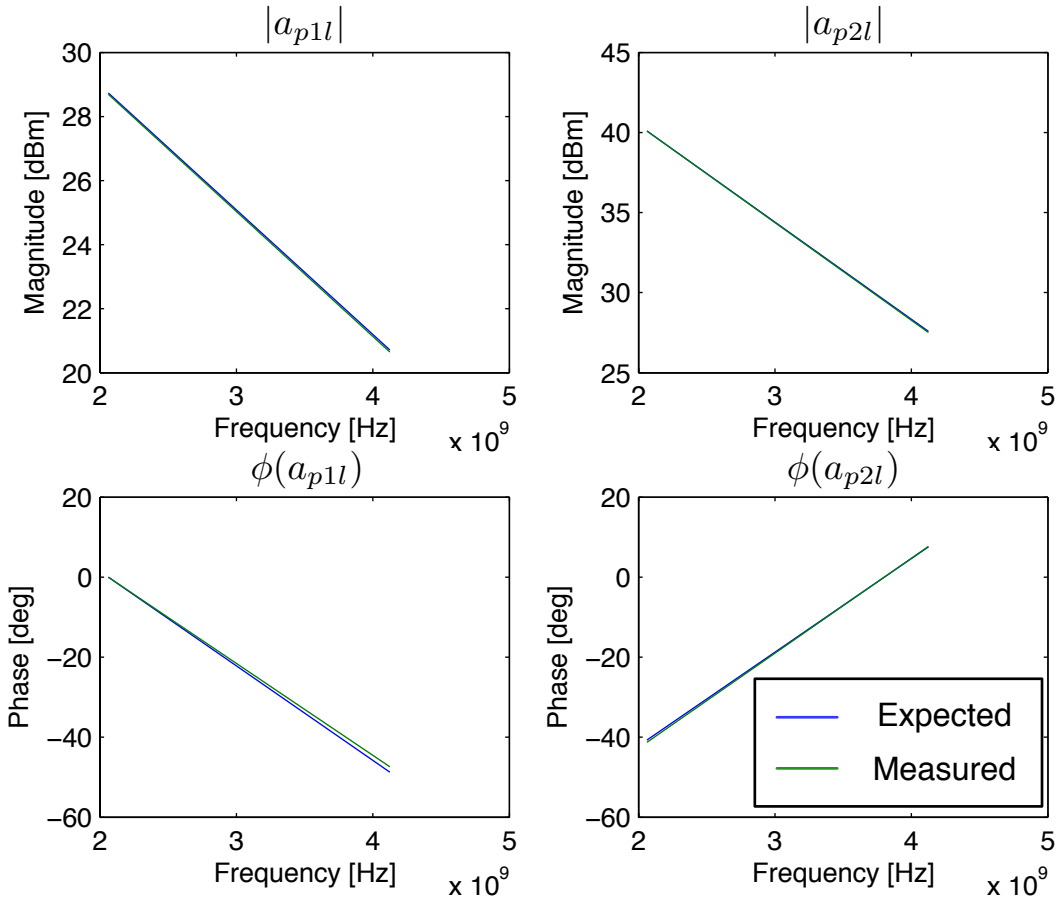


Figure 3.13: Source Vector Calibration Verification

3.3 Example NCS Measurement Applications

This section describes some of the measurement applications inherently supported by the NCS. Each characterization technique presented in this section describes the necessary test-bench configuration, and the resulting information obtained from the measurement data.

3.3.1 Pulsed Stimulus Based Characterization

By controlling high-power pulse generators, it is possible to characterize the DC IV Characteristics with and without thermal power dissipation, using the test-bench configuration in Fig. 3.14. This can be used to extract large-signal compact model coefficients, such as transconductance (G_M), knee voltage (V_k), breakdown voltage (V_{BR}) and the thermal resistance (R_{TH}). From a design perspective, this information is also useful for selecting a DC bias point, and performing load-line PA design techniques, as shown in Fig. 3.15.

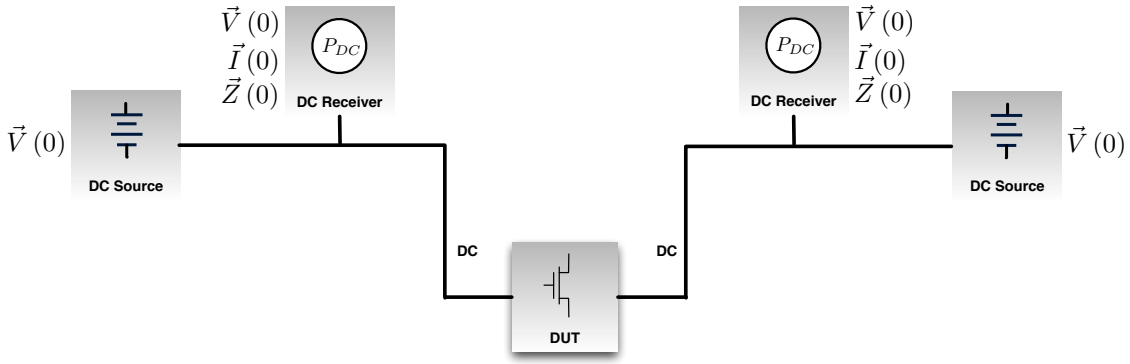


Figure 3.14: Pulsed-DC Test Configuration

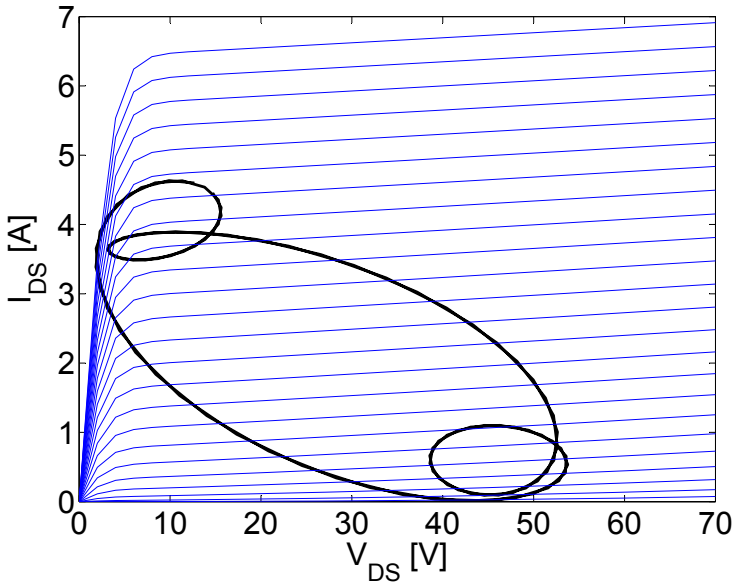
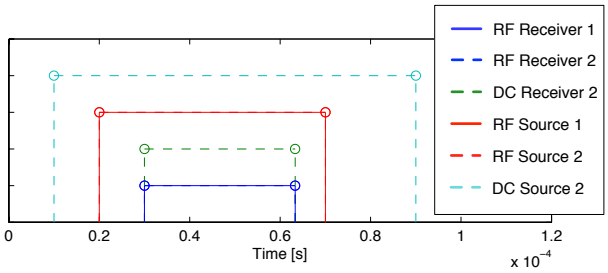


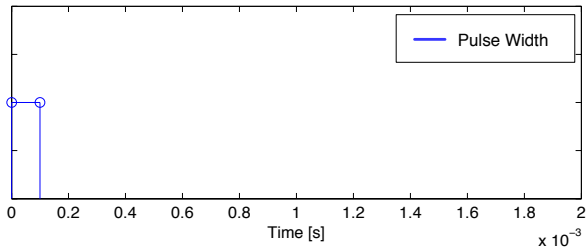
Figure 3.15: DC I/V Characteristic Plot and RF Dynamic Load-Line Plot

Performing CW RF measurements of high-power devices under large-signal conditions can result in thermal power dissipation, that may result in deteriorated device performance or device failure. When performing large-signal measurements, the RF stimulus is on the same order of magnitude as the DC stimulus, therefore the DUT is no longer assumed to remain "off" when the DC stimulus is turned off. In order to characterize the DUT under operating conditions that more accurately reflect the PAPR of the final application, it is important to characterize the DUT under a pulsed RF stimuli. It can also be used to test the DUT under operating conditions that would otherwise exceed the maximum power dissipation of the device.

The minimum/maximum timing limits for each instrument have been embedded into the instrument drivers and the final timing diagram that is presented to the user is shown in Fig. 3.17. A hardware-specific timing diagram for the Focus Microwaves Modular Pulsed IV (Pulsed-DC)[59] and the Keysight PNA-X pulse generators (Pulsed-RF)[60] is described in Fig. 3.16.



(a) Pulse-Width Profile



(b) Pulse-Period Profile

Figure 3.16: Pulsed-DC and Pulsed-RF Timing Diagram

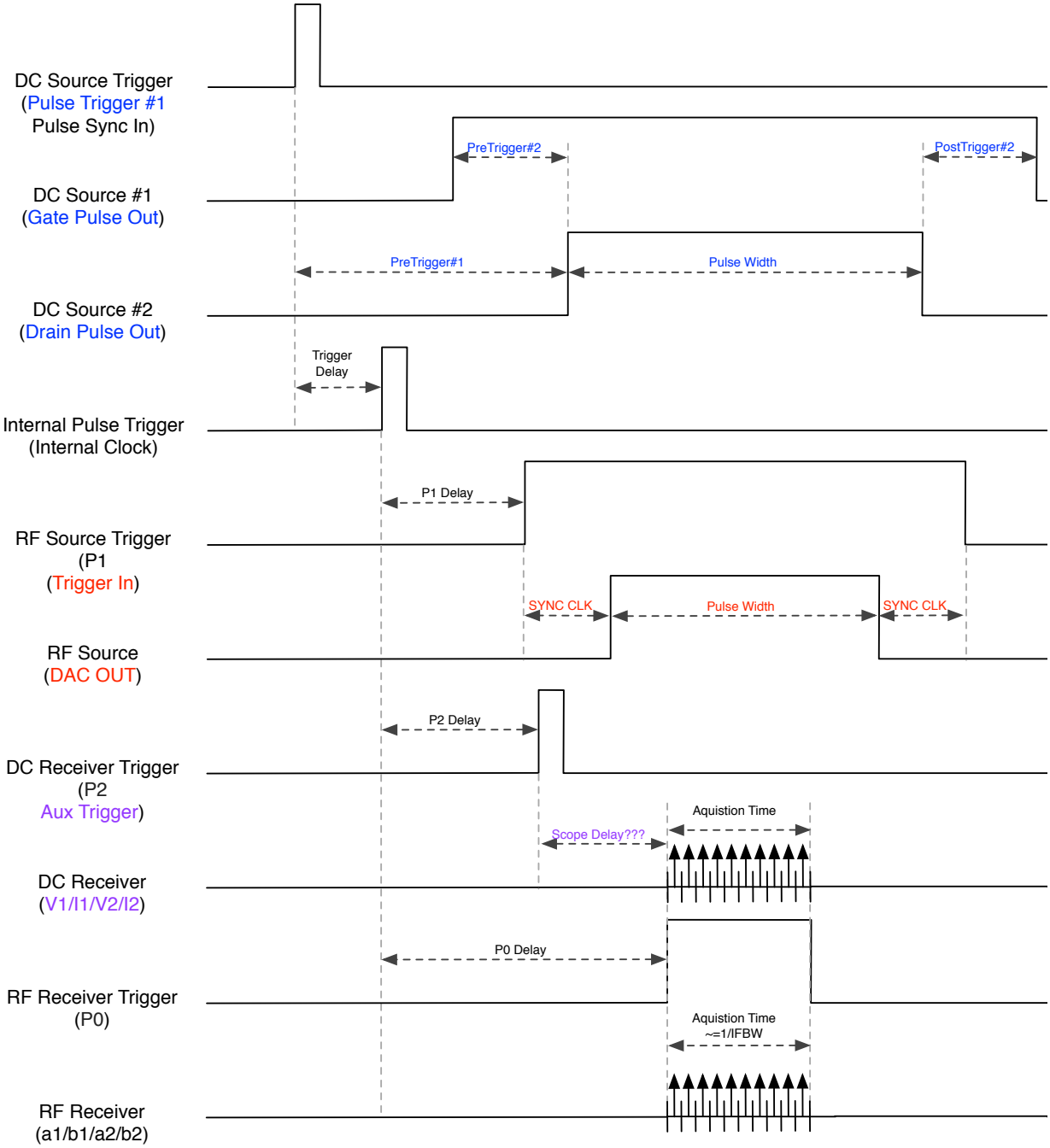


Figure 3.17: Pulse Profile Timing Information for Focus Microwaves Modular Pulsed IV and Agilent N5242A

3.3.2 Multi-Harmonic Active/Passive Hybrid Load-Pull

The measurement application can be configured to separately perform passive and active load-pull measurements at each port and harmonic frequency, as demonstrated in Fig. 3.18 and Fig. 3.20. Whereas pure active load-pull systems implement complex iterative measurement techniques to approximate equivalent passive load terminations, this non-linear characterization combines the advantages of active and passive load-pull techniques. Passive impedance tuners match the DUT, reducing power reflection of the DUT and reducing the power requirements of the test-bench, while active load-pull measurements provide a fast, pulsed, localized impedance sweep, as shown in Fig. 3.21. This implementation is computationally simpler than other active load-pull solutions and is preferable for the purposes of behavioural modelling. Additional programmable search algorithms can be implemented to automatically tune the system in order to maximize the DUT performance.

All signals injected into the system are synthesized using the same LO, hence they are phase coherent. This ensures that active load-tuning can be implemented in a deterministic manor because the relative phase between the signals does not drift. In order to perform active load-pull of high-power devices, passive impedance tuners are necessary to improve the power transfer of the active signal-sources. The multi-harmonic impedance tuners ensure that signals can be injected at all harmonic frequencies while avoiding costly high-power, narrow-band PAs. As a result, the impedance tuners are directly connected to DUT fixture so that low impedance matching of high-power transistors can be achieved.

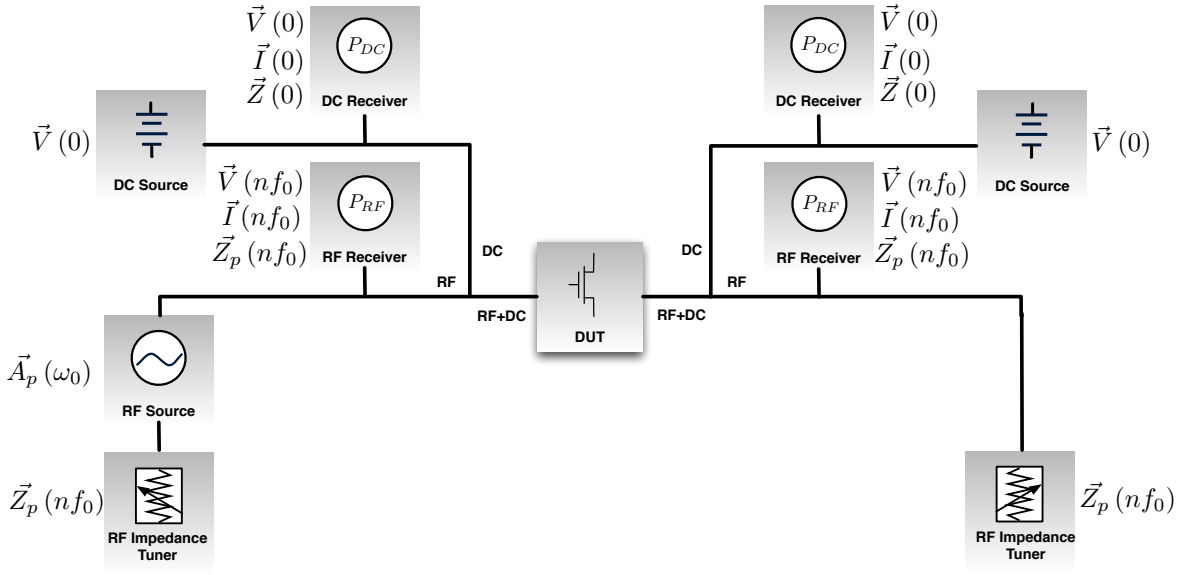
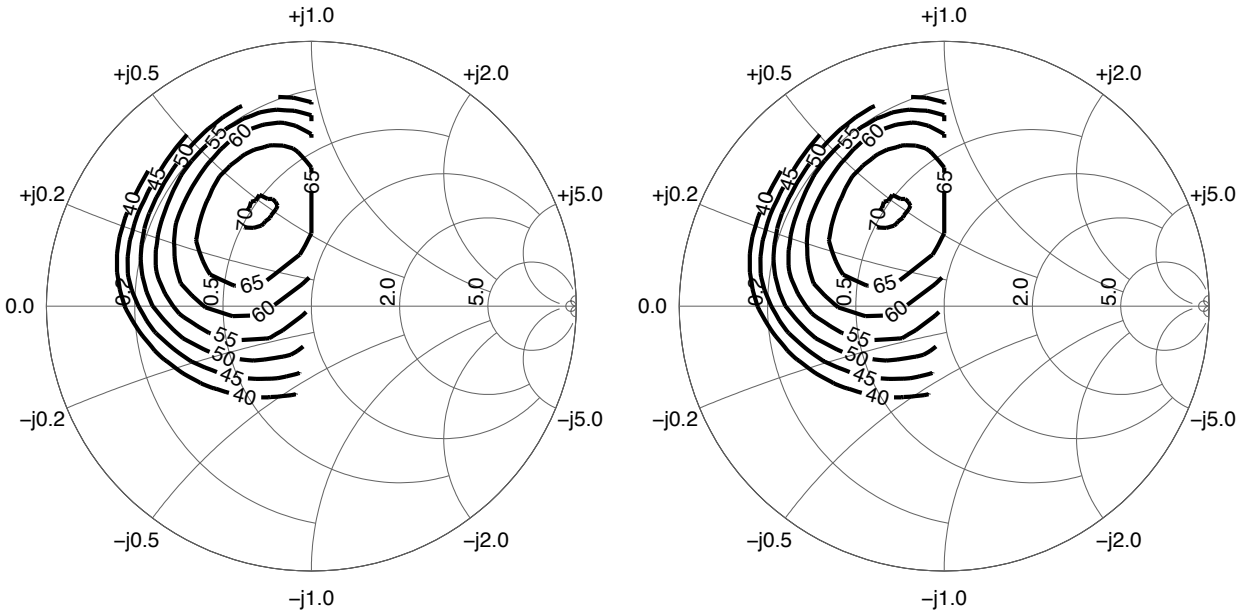


Figure 3.18: Passive Load-Pull Test Configuration



(a) PAE vs. Γ_{21} (b) P_L vs. Γ_{21}

Figure 3.19: Fundamental Harmonic Passive Load-Pull Contours

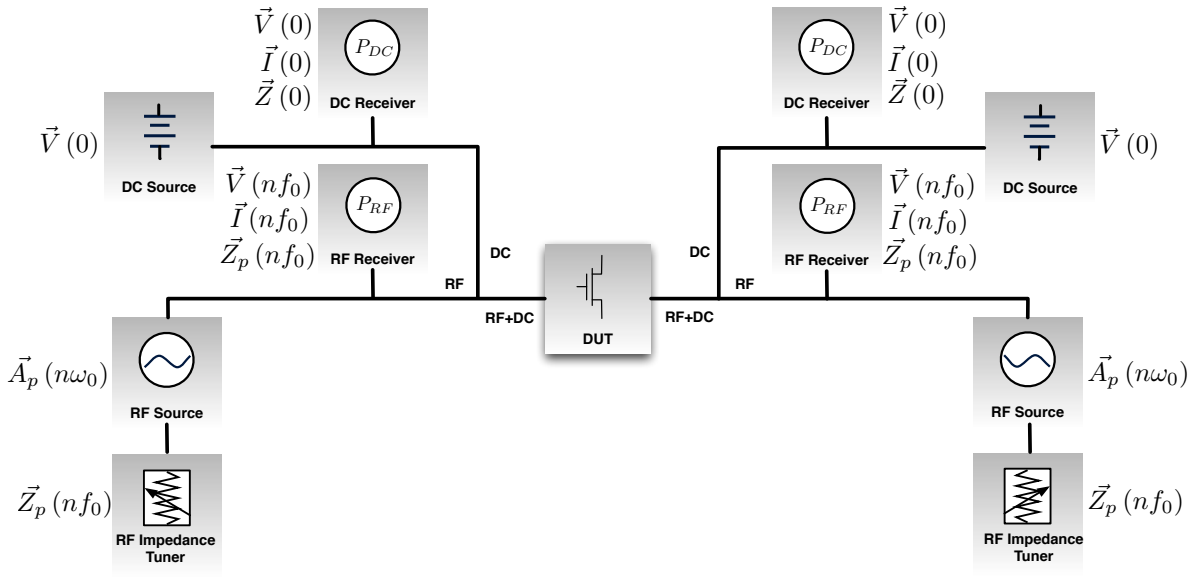
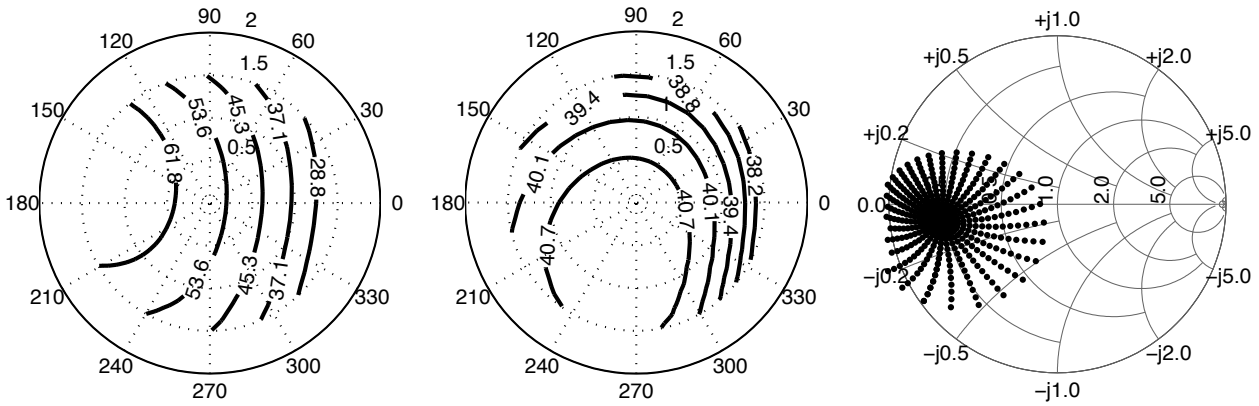


Figure 3.20: Active Load-Pull Test Configuration



(a) PAE vs. a_{p21}

(b) P_L vs. a_{p21}

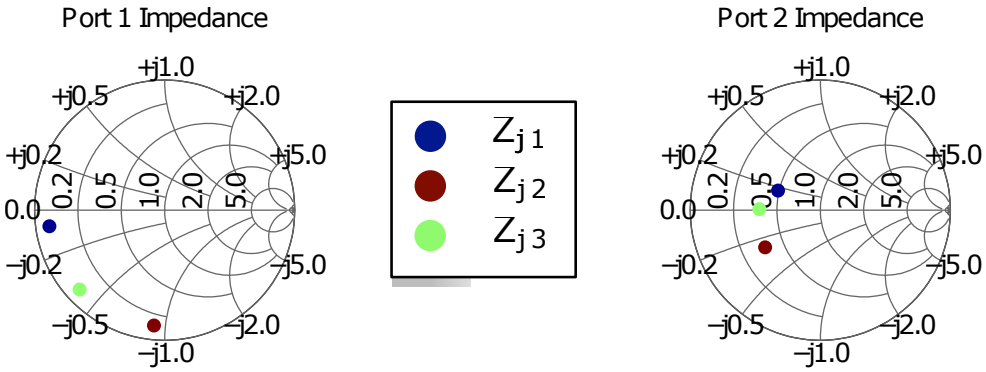
(c) Equivalent Passive Load-Pull Coverage

Figure 3.21: Fundamental Harmonic Active Load-Pull Contours

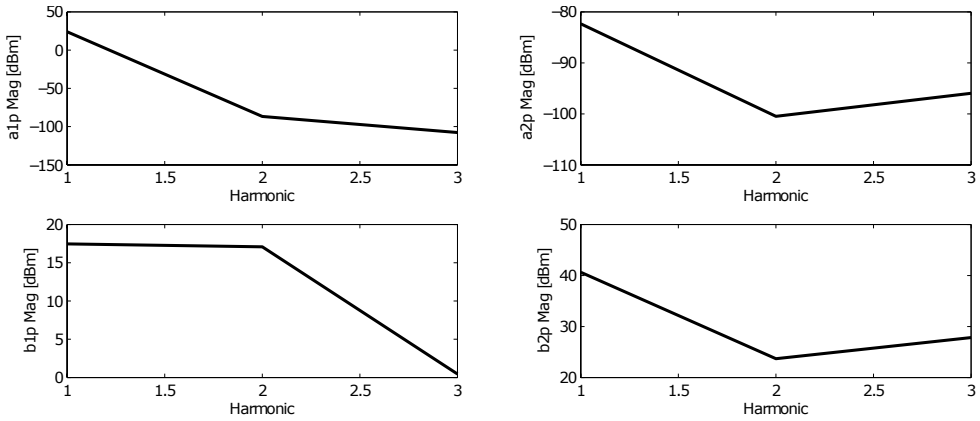
3.3.3 Multi-Harmonic Time-Domain Measurements

When harmonic efficiency design techniques are utilized, traditional load-pull search algorithms are replaced with Waveform Engineering techniques that enable global optimization of switching mode amplifiers. The proposed characterization system provides time-domain reconstruction of power-waves and voltage/current waveforms at the DUT measurement plane, as shown in Fig. 3.23 using the frequency-domain measurement shown in Fig. 3.22. In order to better support Waveform Engineering techniques, a nonlinear large-signal de-embedding technique must be implemented to synthesize the voltage and current waveforms at the DUT intrinsic reference plane.

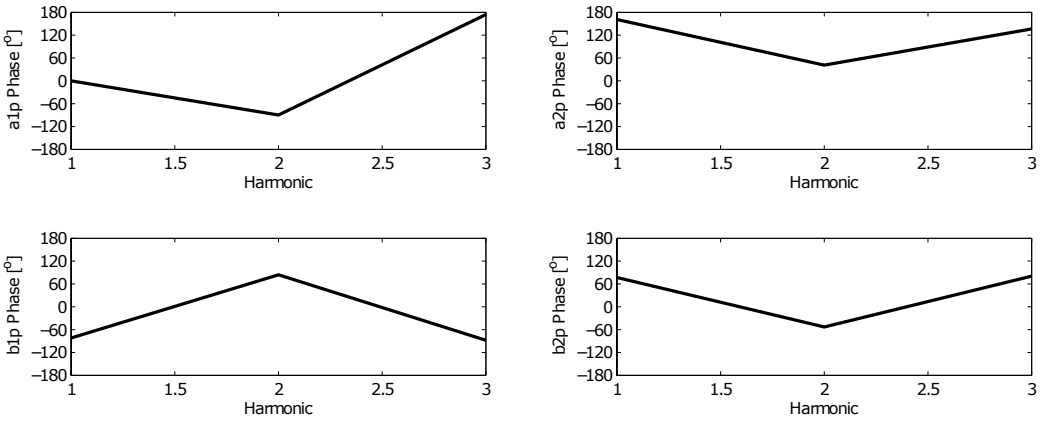
To measure multiple harmonics coherently, the measurement system must normalize each receiver measurement by a periodic multi-harmonic reference signal that is generated using the same LO as the RF Sources. In low-power measurement applications the phase reference is generated by passing a $10MHz$ sinusoid through a nonlinear diode, thereby producing a rich output spectrum of tones separated by $10MHz$. High-power applications require large attenuators on each receiver, hence the $10MHz$ input will generate signals below the noise floor. Therefore, a $687MHz$ phase coherent synchronization signal was taken from the same AWG that generated the RF Source signals to produce a clean phase reference for the RF Receiver measurements.



(a) Frequency-Domain Port Impedances



(b) Frequency-Domain Magnitude Response



(c) Frequency-Domain Phase Response

Figure 3.22: Single Measurement Frequency-Domain Response

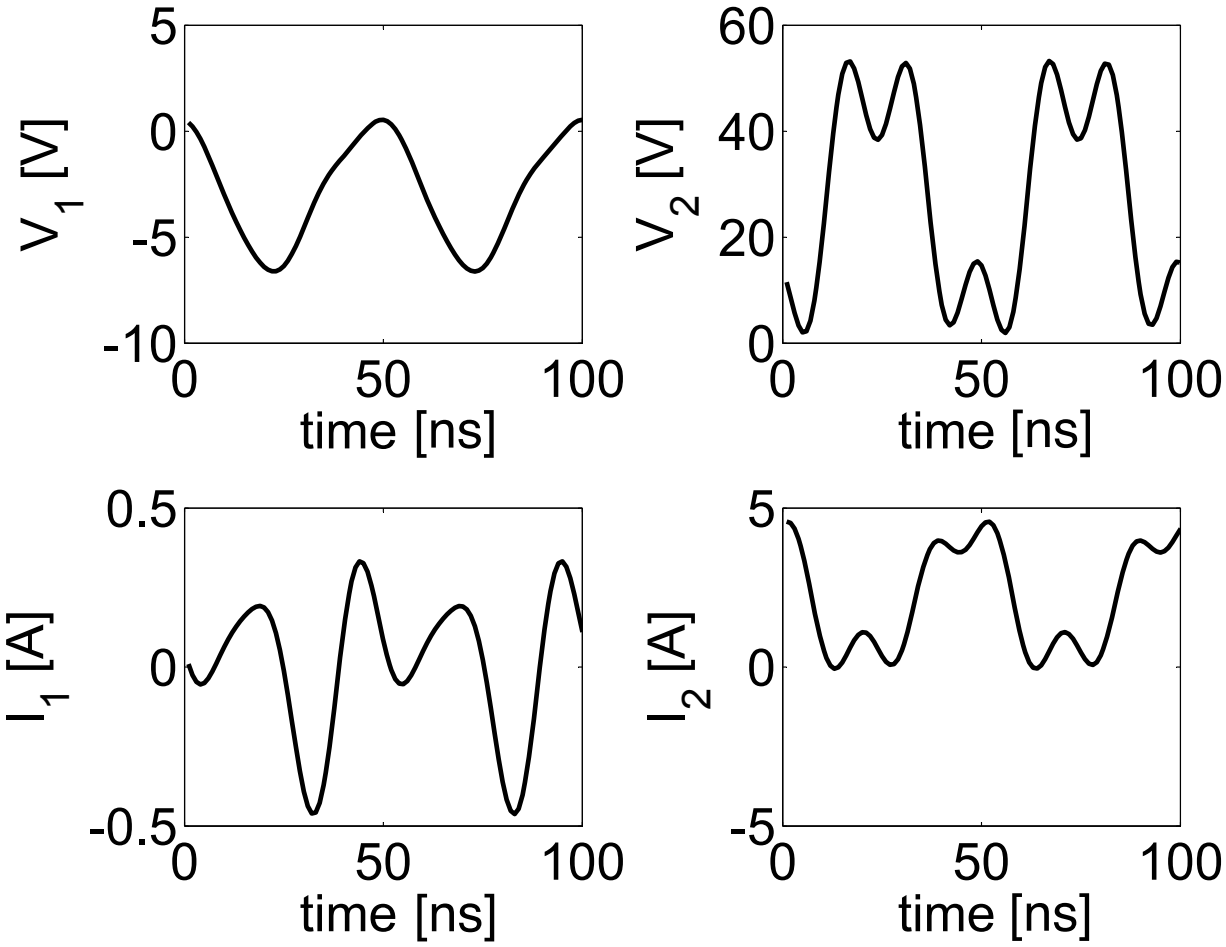


Figure 3.23: Single Measurement Time-Domain Voltage and Current Response

3.3.4 Power-Sweep Nonlinear Distortion Measurements

The power-sweep configuration, in Fig. 3.24, provides useful information about the AM-AM and AM-PM compression (shown in Fig. 3.25), assuming constant bias and matching conditions. Although this measurement can be performed under CW and pulse-modulated signals, it is traditionally preferable to present a digitally modulated signal to better emulate the final performance of the finished PA.

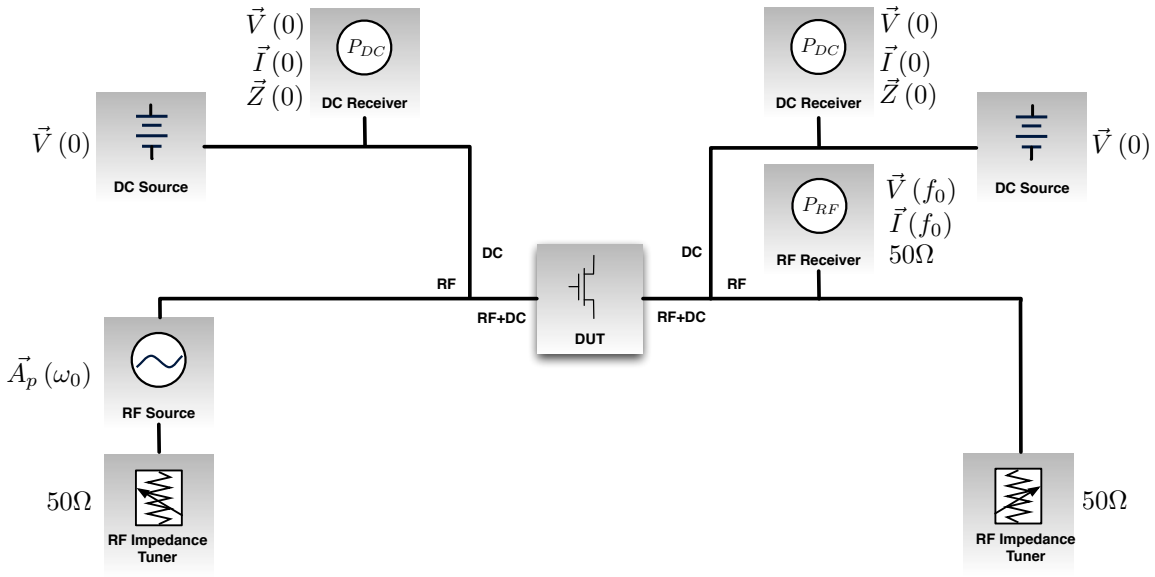


Figure 3.24: Power Sweep Test Configuration

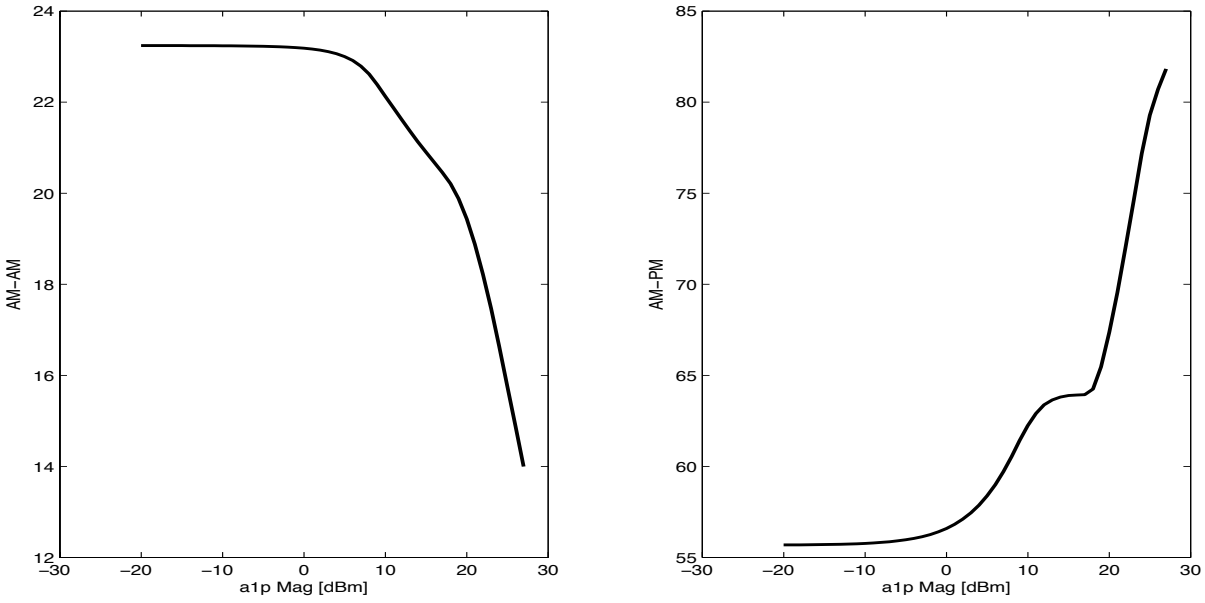


Figure 3.25: AM-AM and AM-PM Power-Sweep Response

3.4 Towards Automated Nonlinear Modeling and Design

The NCS represents a modular framework that provides seamless integration between measurement and simulation environments by operating in three different modes:

- Measurement Mode.
- Simulation Mode.
- Model Mode.

Measurement Mode connects to instruments on a measurement test-bench based on the hardware instruments available. This modular approach means that the NCS can continue to perform measurements even when some of the instruments have not been connected. For example, if an RF impedance tuner is not specified, the NCS will assume an *ideal* 50Ω impedance termination (or the actual constant port impedance when port mismatch correction is applied).

The Simulation Mode allows the functionality of the Measurement Mode to be *emulated* using automated simulations from any Harmonic Balance simulator. This can be used to test software processes without performing time-consuming calibration techniques or scheduling measurement system down-time. The Simulation Mode uses the exact same interface and algorithms as the Measurement mode, enabling the design process flow in Fig. 3.26 by *seamlessly* integrating the measurement and simulation environments.

Finally, the Model Mode calculates the output response by applying the specified inputs to a mathematical behavioural model. This feature ensures that behavioural models can be evaluated *directly* without using the Harmonic Balance nonlinear (iterative) circuit-based solver that may result in convergence errors.

This system forms the back-bone for developing CAD-based, automated design techniques, using the process-flow shown in Fig. 3.26. Hence, it allows application specific

characterization routines, like load-pull, to be synthesized on top of a common measurement framework. Fig. 3.5 described the block diagram of the proposed NCS, which emphasizes the utilization of:

- Multi-harmonic passive impedance tuners in order to improve the matching, the bi-directional power-transfer between the measurement system and unmatched DUTs.
- High-Power multi-harmonic active signal sources to synthesize the large-signal operating condition quickly, and without significantly degrading power-transfer within the measurement system.

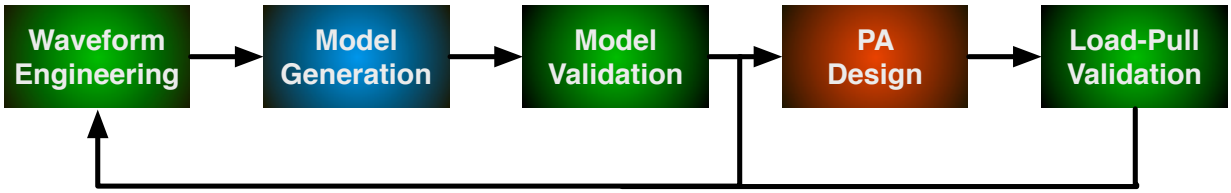


Figure 3.26: Feedback Design Process

While the proposed NCS can be customized to emulate specific design conditions, perhaps by repositioning the impedance tuners or by integrating a design-specific bias network, *it is generally best to implement a characterization system that is conjugately matched over as many frequencies as possible. This type of system will maximize stability, minimize the power requirements of active signal injection, and will simplify the behavioural model extraction by ensuring that all of the incident power is delivered to the DUT, as further described in Chapter 4.*

Chapter 4

Multi-Harmonic Behavioural Modeling

Chapter 3 provided the characterization backbone for studying multi-harmonic behavioural model formulation and extraction procedures in a controlled, repeatable environment. Although characterization systems have be used to facilitate traditional PA design procedures "on the bench", as the PA design requirements increase and circuit topologies advance, CAD tool based design has become crucial to maximizing PA performance. This shift in design methodology requires a nonlinear behavioural model to replicate the DUT response inside the CAD environment, based on measurement data collected from the NCS. However, it is critical to develop behavioural models that balance a trade-off between mathematical formulation complexity and model prediction robustness.

Chapter 3 created a flexible measurement environment that was optimized for model extraction. While characterization applications are typically focused on debugging a problem by exactly replicating a nonlinear operating condition on the test-bench, modelling is more concerned with extracting data quickly. For example, Passive Load-Pull measurements illustrate how the DUT will behave under different *load impedances*, as shown in Fig 3.19. As this information can be directly applied to the design of a matching network, it is the optimal choice for test-bench characterization. Unfortunately, passive impedance tuners are slow mechanical circuits that may induce oscillation without careful reconfigura-

tion. Alternatively, Active Load-Pull measurements describe how the DUT behaves under different *incident power-waves* (at both ports), as shown in Fig 3.21. As PA designers only inject power at the fundamental frequency on the input port, this configuration is over-complicated and unintuitive for the needs of a PA designer. However, the open-loop active load-pull architecture of the NCS facilitates rapid measurement extraction using power sweeps, rather than impedance sweeps, so that the model extraction procedure closely represents the model formulation.

This chapter starts with the model requirements imposed by a high-power, unmatched, broadband DUT, and uses preliminary characterization analysis to establish the parameters of a controlled behavioural model comparison. A high-level comparison of several behavioural formulations investigates how each solution scales with respect to nonlinear order, and the number of system inputs. It also compares the implication each model formulation places on the extraction procedure and the implied complexity of measurement test-bench. An further in-depth comparison proposes two complimentary solutions: i) a first-order expansion PHD model for weakly nonlinear systems with simple hardware requirements, and ii) a higher-order MHV-based formulation with more sophisticated hardware requirements. The first solution maximizes the accuracy of the existing PHD model formulation by focussing instead on optimizing the model extraction procedure for unmatched, broadband devices. The second solution provides a higher-order model formulation for applications where the poly-harmonic superposition assumption is no longer valid. Both models are extracted under the same LSOP, using simultaneous multi-harmonic active load-pull measurements, which are essential for testing the assumption of harmonic source superposition. The modular nature of the NCS makes it easy to integrated hardware resources based on the RF bandwidth and port impedances of the DUT, it provides programmable extensions to extract the PHD or MHV models based on the nonlinear order of the DUT, and it provides an automated measurement environment for model validation.

4.1 Preliminary Characterization Modeling Space Definition

4.1.1 Definition of the Model Variables Determined by the Device Under Test

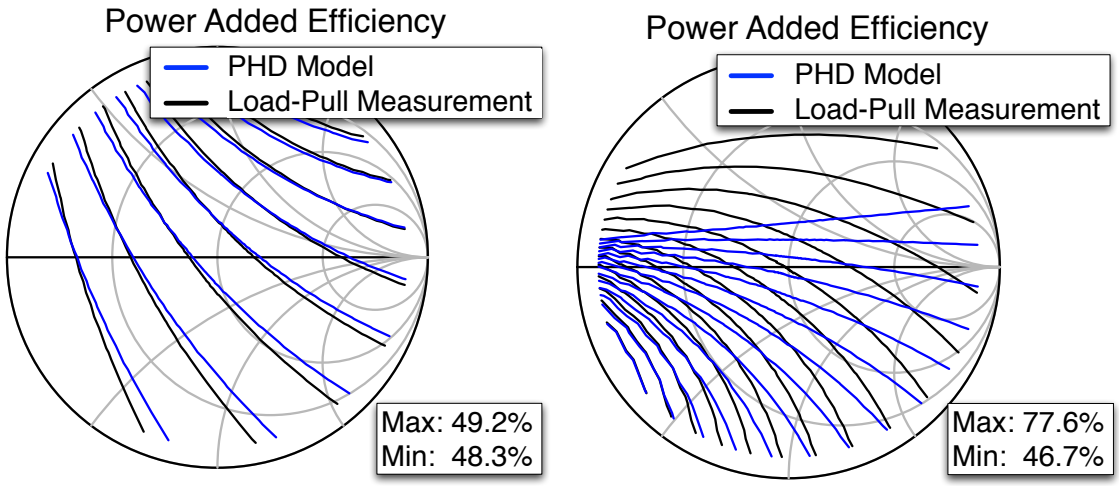
It's difficult to find a model topology that provides an optimal solution for all DUTs, therefore the following characteristics of the DUT will impact the selected model:

- Number of harmonics
- Matched or unmatched (number of stimulus ports)
- Transistor conduction angle (transistor mode of operation)

The design frequency and the available bandwidth of the DUT will determine how many harmonics are below the threshold frequency (f_t) and therefore impact DUT performance. Harmonics above the f_t are irrelevant because they have a lower output power, hence they are insensitive to model extraction stimuli. The study in Fig 4.1 demonstrates the relative sensitivity to the second harmonic source impedance by looking at the $X^S_{21,12}$ PHD model coefficient, when modelling a band-limited pre-matched transistor (Fig. 4.1a) and a broadband general purpose transistor (Fig. 4.1b). This analysis shows that the PAE of the general purpose transistor is highly sensitive to the second harmonic source impedance. As a result, the PHD model predicts the weak nonlinearity of the pre-matched transistor in Fig. 4.1a, but it does not have the ability to predict the higher-order nonlinear mixing products in Fig. 4.1b[61].

While compact circuit models divide-and-conquer a model using equivalent circuit components, the accuracy of a behavioural model is inherently limited by the number system of inputs and their linear-independence. As nonlinear systems operating under large-signal conditions produces cross-harmonic mixing-products, multi-harmonic system inputs inevitably become linearly dependent as nonlinearity increases. In device-level models this problem is further exacerbated by the number DUT ports. Whereas system-level models

operate under the assumption of a fixed load impedance termination (50Ω), device-level models have unknown impedance terminations, hence matched PAs represent a SISO system (shown in Fig. 2.10a) and unmatched transistors represent a MIMO (shown in Fig. 2.10b). Although the more generalized MIMO model increases the likeliness of a global maximum solution, it is unlikely that this type of model could produce the same accuracy of a more constrained SISO behavioural model. Never-the-less approximate limits on harmonic power-sweeps can be enacted when the impedance matching network is to be implemented using a passive circuit.



(a) Pre-Matched Transistor

(b) General-Purpose Transistor

Figure 4.1: Accuracy of PHD model when predicting 2nd Harmonic Source-Pull of a Pre-Matched Transistor and Broadband General-Purpose Transistor

The conduction angle of the transistor represents the fraction of the carrier period spent in cut-off and is determined by the DC class of operation. Under large-signal operation, another fraction of the carrier period is spent operating in triode mode. Although precise knowledge of how often the transistor operates in cut-off or triode mode would require detailed information about device packaging and parasitic capacitances (invaliding the black-box model assumption), significant information can be inferred by performing preliminary characterization before extracting a model. For example, the DC I/V char-

acteristic plot, and the quiescent DC operating point largely determine the time spent in cut-off, whereas triode-operation can be slightly more difficult to predict.

To illustrate the "sources of nonlinearity" inside a nonlinear system, a Class-AB, two-harmonic, two-port, conjugately matched transistor is characterized at the intrinsic device plane. Fig. 4.2 demonstrates the onset of cut-off nonlinearity (Fig. 4.2b), and triode nonlinearity (Fig. 4.2c) while sweeping A_{12} and $|A_{11}|$. The device is more prone to operating in cut-off than triode because is is biased between Class-A and Class-B. The load-line, shown in Fig. 4.2, represents the cumulative time-domain response of the individual "harmonic load-lines" at f_0 and $2f_0$, whose slope is proportional to the frequency-dependent load conductance (G_{21} and G_{22}) shown in Fig. 4.4a. As these load-lines have similar slopes in a conjugate matched network, the relative phase $\phi(A_{12})$ mostly determines the length of the cumulative load-line. For example, the nonlinear B_{21} response in Fig. 4.2a and Fig. 4.2b are different, because the shape of the first nonlinearity is attributed to A_{12}, while the second response attributed to both A_{11} and A_{12}. Since the nonlinear response in Fig. 4.2b and Fig. 4.2c are primarily attributed to A_{11} and A_{12}, we can clearly distinguish the conduction angle spent in cut-off from the conduction angle spent in triode. We can conclude that when the device begins to operate in cut-off mode, it becomes a simultaneous nonlinear function of A_{11} and A_{12}.

Applying A_{21} (with $A_{11} = 0$) creates a large variation in V_{DS}, but small changes in V_{GS} and I_{DS} which are negligible when the transistor satisfies the small-signal (or small-mismatch) unilateral transistor condition[62]. This implies that injecting a signal at any harmonic on the drain produces a horizontal load-line, as shown in Fig. 4.4b, that increases the likelihood of triode operation. Unfortunately, we will see this statement is only partially correct. Fig. 4.3 demonstrates the nonlinear response (B_{21}) to signals simultaneously applied at A_{11} and A_{21}, with varying $\phi(A_{21}) - \phi(A_{11})$ phase shift. By varying the relative phase $\phi(A_{21}) - \phi(A_{11})$, a wide distribution of load-lines can be synthesized. Although the load-line in Fig. 4.3a encompasses many shapes, it does not enter cut-off, hence the response to signals injected on the drain is linear until the device enters triode operation (as shown in Fig. 4.3b).

Preliminary characterization applied to an unknown black-box nonlinear system determines how many dedicated RF signal sources must be provided based on the following

inferences:

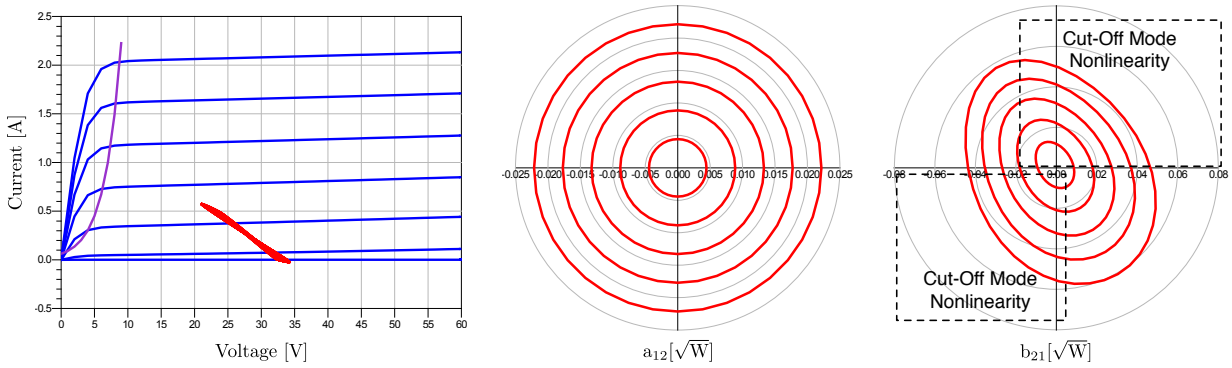
- Non-linearity due to cut-off will increase the nonlinear order with respect to the signals applied to the gate.
- A conjugately matched DUT is unlikely to operate in triode.
- If load-mismatch must be generated to synthesize an LSOP (which is often is the case for high-power PAs), the nonlinear order will increase with respect to all signals applied to the system.
- Nonlinearity is maximized at a specific $\phi(A_{21}) - \phi(A_{11})$ phase relationship, where the transistor operates in both cut-off and triode mode.

Although conjugate impedance terminations provide the best measurement conditions for model extraction, for high-power PAs (when I_{DSmax} rating prevents conjugate matching), a constant A_{21} signal must be injected at a specific phase to maximize both the current *and voltage* swing. Under the maximum output power operating condition, the transistor will operate in both cut-off and triode mode, increasing the nonlinear order with respect to all inputs into the system.

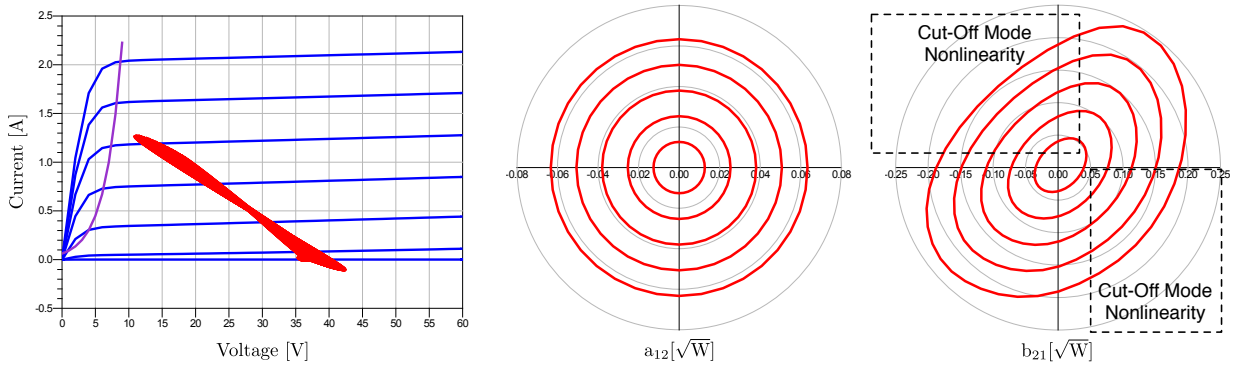
4.1.2 Large-Signal Operating Condition Synthesis

Using the information from the previous section, a LSOP will be constructed to compare multi-harmonic behavioural models under the strongest possible nonlinear conditions, where output power is maximized. A Cree CGH60060D 60W GaN is biased in Class-AB to approximate most design applications. The fundamental frequency ($2.06GHz$) is chosen, such that 3-harmonics are below the threshold frequency, $f_t = 6GHz$. The rest of the constraints are determined by using some preliminary characterizations.

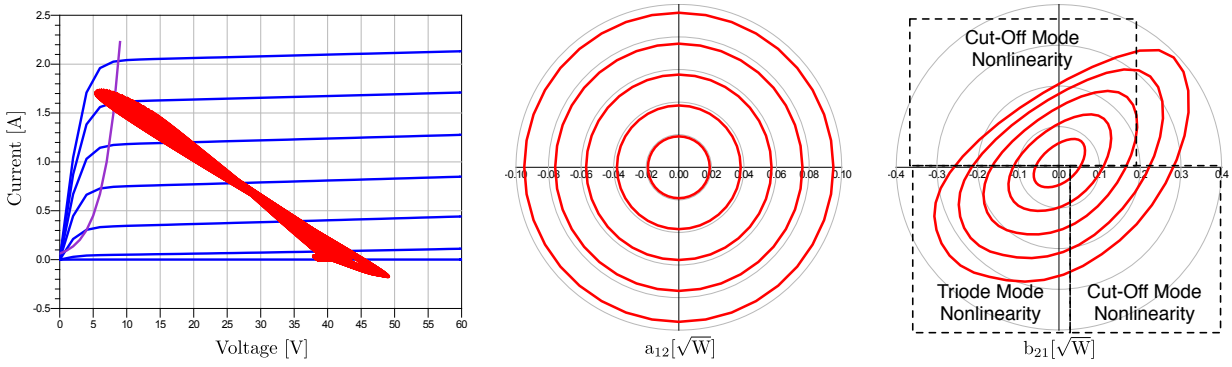
Matching the DUT with a multi-harmonic passive impedance tuner maximizes the power delivered to the transistor, producing a better representation of the model inputs, and reducing the order of the model by eliminating system feedback. Preliminary S-Parameter measurements at harmonic frequencies provide a useful approximation of the



(a) $|A_{11}| = 4\ dBm$

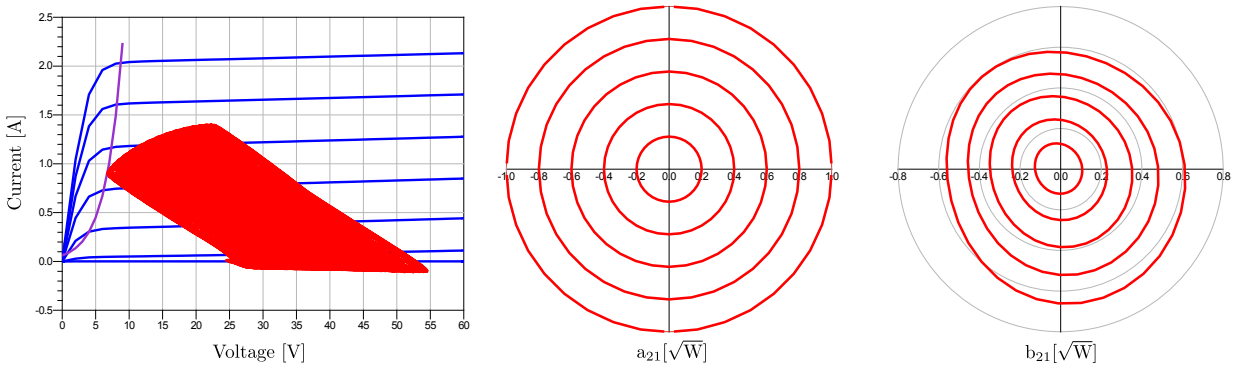


(b) $|A_{11}| = 16\ dBm$

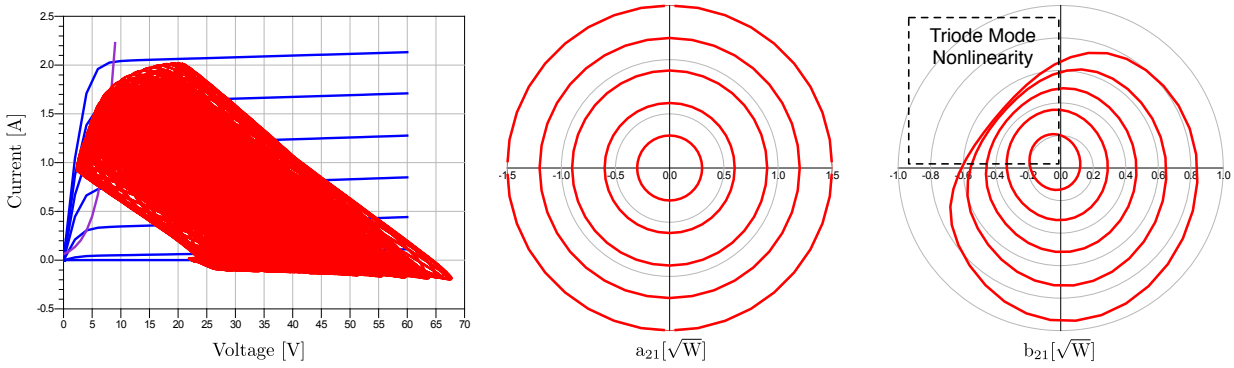


(c) $|A_{11}| = 19\ dBm$

Figure 4.2: Nonlinear Sensitivity of the DUT to $|A_{12}| < 10\ dBm$ and $0 < \phi(A_{12}) < 2\pi$

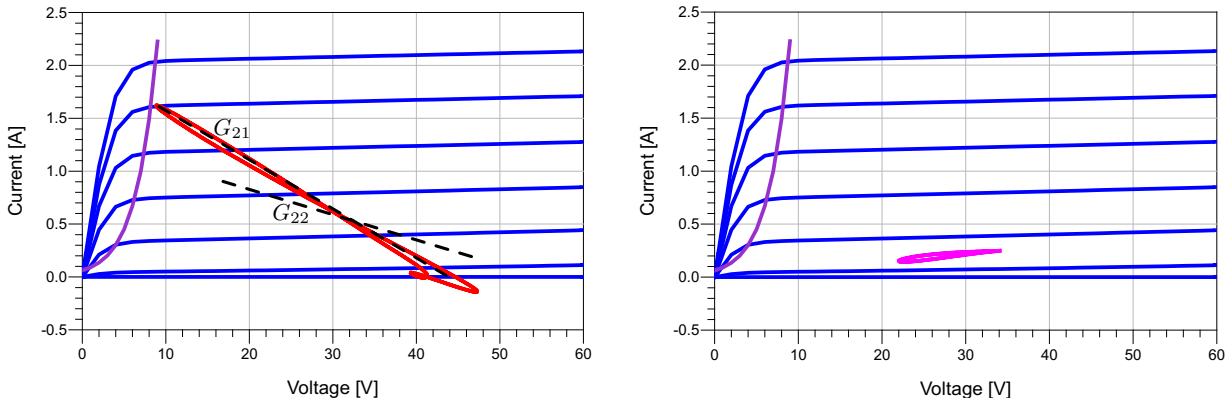


(a) $|A_{11}| = 16\ dBm$



(b) $|A_{11}| = 19\ dBm$

Figure 4.3: Nonlinear Sensitivity of the DUT to $|A_{21}| < 33.5\ dBm$ and $0 < \phi(A_{21}) < 2\pi$



(a) Forward Injected Signal Response

(b) Reverse Injected Signal Response

Figure 4.4: Load-Line Response to Forward and Reverse Multi-Harmonic Injected Signal

conjugate match (Z_{RF}) by using (4.1-4.2). Under nonlinear operation, further optimization of the match is achieved by maximizing b_{pik}, $\forall ik \neq 11$. For devices that do not draw current under DC bias conditions (Class B - Class C), multi-harmonic load-pull search algorithms must be used to maximize b_{pik}, $\forall ik \neq 11$.

$$\Gamma_S = \Gamma^*_{IN} = \left[S_{11} + \frac{S_{12}\Gamma_L S_{21}}{1 - S_{22}\Gamma_L} \right]^* \tag{4.1}$$

$$\Gamma_L = \Gamma^*_{OUT} = \left[S_{22} + \frac{S_{12}\Gamma_S S_{21}}{1 - S_{11}\Gamma_S} \right]^* \tag{4.2}$$

Since conjugate impedance terminations do not maximize output power in a high-power PA, a $A_{21} \neq 0$ must be injected to maximize the drain current/voltage swing. Based on the analysis in Fig. 4.3b, the strongest nonlinear operating condition can be empirically found by sweeping $|A_{p11}|$, $|A_{p21}|$, and $\phi(A_{p21})$, without knowledge of the transistor package and extrinsic device model. The harmonic incident power-wave conditions should not be included n the LSOP. Assuming the future matching networks are designed using passive components, the multi-harmonic model extraction range (aRFStep) should be em-

pirically approximated as the amplitude of the maximum harmonic reflected power-wave. Since the harmonic impedances are conjugately matched, sweeping the harmonic incident power-waves from $0 - aRFStep$ will ensure nearly uniform coverage of the harmonic smith charts. The final time-domain voltage/current waveforms of the LSOP used for future model comparison is shown in Fig. 4.5, while the load-line is shown in Fig. 4.6. The resulting LSOP settings are summarized in Table 4.1.

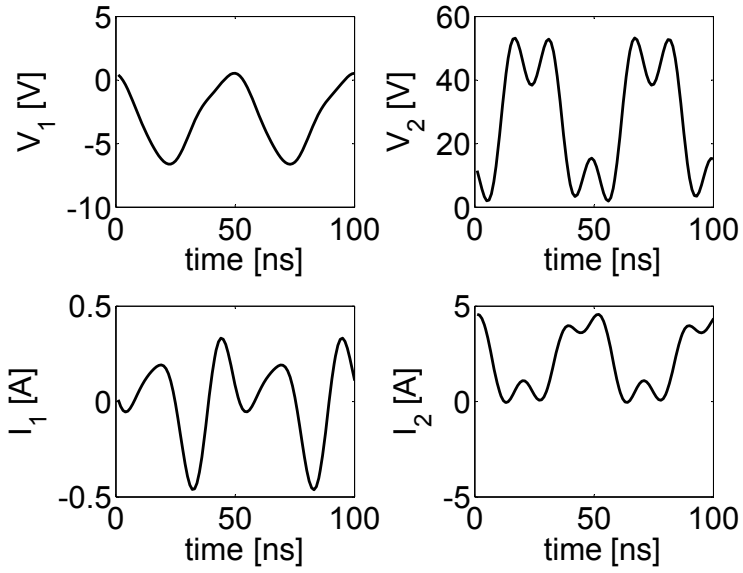


Figure 4.5: Intrinsic Transistor Current/Voltage Waveforms

4.2 Defining the Model Formulation Based on the Nonlinear Order of the LSOP

The model formulation must be chosen based on the nonlineaity of the LSOP defined in section 4.1, such that the desired model accuracy is achieved using the minimum model complexity. This section discusses the limitations of three behavioural models in dealing with unmatched nonlinear transistors, namely:

- The Multi-Harmonic Volterra (MHV) Series Model.

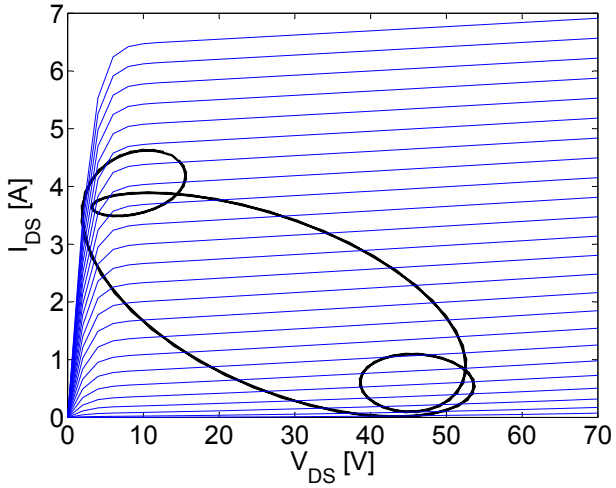


Figure 4.6: Dynamic RF Load-Line Operating in Triode and Cut-off

Table 4.1: Initial Multi-Harmonic Model Settings

Setting	Description	CGH60060F 60W	Unit
Freq	Fundamental frequency	2.06	GHz
NumPorts	Number of device ports	2	
aHarmonics	Number incident harmonic power-waves	3	
bHarmonics	Number reflected harmonic power-waves ($\geq aHarmonics$)	3	
V_{DC}	V_{GS} and V_{DS} DC LSOP	$[-2.92, 28]$	V
Z_{RF}	Multi-harmonic RF port impedance (approximate conjugate match)	$[16.7 - 0.725i, 5.01 + 0i;$ $1.28 + 0i, 2.64 - 2.18i;$ $2.63 + 0i, 2.82 - 13.4i]$	Ω
A_{RF}	A_{11} and A_{21} RF LSOP	$[0.708, 3.16\angle 180]$	$\sqrt{W}$
aRFStep	Maximum RF multi-harmonic extraction power range	$[0.2, 2.5]$	$\sqrt{W}$

- The Poly-Harmonic Distortion (PHD) Model
- The Cardiff Model

4.2.1 Limitations of the Multi-Harmonic Volterra Model

Limitation of the Multi-Harmonic Volterra Model Formulation

The MHV model in (2.10-2.12) offers a complete description of the static nonlinear system. Unfortunately, describing the response as a summation of all possible mixing products in the frequency domain requires a large number of coefficients, and the size of the model is dictated by number of harmonics and number of ports in the system. Alternatively, we could formulate the Volterra Series in the time-domain, however the time-step (sampling rate) would also be determined by the number of harmonics. Additionally, the time-domain measurement of a black-box system is not a relevant depiction of the true time-domain signal at the intrinsic drain of the transistor.

By definition, the Volterra series defines the response of weakly nonlinear systems with respect to non-commensurate input signals at a constant DC bias. The MHV model circumvents this limitation by defining each RF kernel ($V_{ik,m}$) as a nested function of several DC mixing products ($\vec{x}_n(0)$), thus multiplying the number of kernels by the number of DC mixing products. The resulting formulation only requires one nested layer of functions and is similar to (4.3).

$$y = \sum f_i\left(x_i, g_j\left(x_{DC}\right)\right) \tag{4.3}$$

Limitation of the Multi-Harmonic Volterra Model Extraction

While the MHV formulation provides infinite accuracy in theory, it is complex in terms of the number of coefficients and the resulting number of measurements. There are two types of extraction methods:

1. Direct Synthesis

2. Expansion

The direct extraction method uses linear regression techniques to extract the Volterra Series under the assumption that the kernels are linearly independent. Since the kernels are not orthogonal, they must all be extracted simultaneously, thus problems with a large number of coefficients will become poorly conditioned. Alternatively, approximating the Volterra series with an orthogonal expansion (as is done by the PHD Model), will result in a better conditioned problem where model coefficients can be extracted with greater certainty.

Another practical limitation of the direct synthesis is that coefficients are dependent on the available source power (A_{jl}), whereas the nonlinearity of the system is dependent on the delivered input power ($A_{jl} - B_{jl}$) (a feedback equation). Unmatched systems may result in higher-order mixing products and poor accuracy. The nonlinear and bilateral properties of the DUT imply the input and output impedance is dependent on the stimulus condition, thus it is impossible to ensure that the DUT will remain conjugately matched during the entire model extraction. Alternatively, an expansion model formulation based on differential changes in stimulus, such as the PHD model, would have increased immunity to port matching because *changes* in available source power would be proportional to the *change* in delivered input power.

4.2.2 Limitation of the PHD Model

Limitation of the PHD Model Formulation

The PHD model formulation relies on the superposition of low-order expansion of mixing products to approximate the final behaviour of the system, as shown in (4.4).

$$y = \sum f_i(x_i) \tag{4.4}$$

The PHD model isolates the contributions of different incident power-waves to the reflected power-waves. The inherent fault in this formulation is that the principle of source-superposition does not apply to nonlinear systems. Therefore, the PHD model should describe the system as function of its inputs, as well as all possible input combinations.

The PHD model in (2.23) represents a low-order expansion of the MHV model as described in [43]. To summarize the results of the derivation, the PHD model is a first-order polynomial expansion of the MHV formulation around a LSOP. While the DUT is biased at the LSOP, non-LSOP variables are modelled by stimulating the DUT with a sinusoidal input, simultaneously injecting a linear signal at $+l\omega_0$ (a_{jl}) and $-l\omega_0$ (a^*_{jl}) on port j. By sweeping the phase of the sinusoidal stimulus, the individual inputs will be swept in a circular pattern, shown in Fig. 4.2a, however a_{jl} and a^*_{jl} rotate in separate directions. The combined response of both inputs is assumed to form a ellipse, represented by $b_{21} = X^S_{21,jl}a_{jl} + X^T_{21,jl}a^*_{jl}$, as shown in 4.2a. Unfortunately, as the system becomes more nonlinear, as shown in Fig. 4.2c and Fig 4.3b, the response takes on a distorted shape that must be described by higher-order functions. Therefore PHD models are intrinsically limited to weakly nonlinear systems.

In an attempt to use the PHD to model higher-order nonlinearities, $|A_{21}|$, $\phi(A_{21})$, or other variables that induce highly nonlinear behaviour can be added to the list of LSOP variables as done in the Keysight Load-Dependent X-Parameter model (2.24) and (2.25). Designating more variables as the LSOP constrains the modelling region to a smaller portion of the measurement space where a first-order expansion is still valid. To continue offering device prediction under different LSOP values, we need to create a piecewise model, a multi-dimensional look-up table of PHD models, that are defined over a range of each LSOP variable. Some questions that come to mind are:

- Which variables should be considered LSOP terms?
- What is the minimum LUT density?
- How do these choices impact the number of measurements, and the time needed extract the model?

Based on these questions, it is difficult to quantitatively derive how the Load-Dependent X-Parameter model can be modified to predict higher nonlinearities because it is always theoretically possible to decompose the measurement space into a LUT of progressively smaller PHD models[61]. The underlying assumption behind the Load-Dependent X-Parameter model is that changes to LSOP variables result in a strongly nonlinear change in

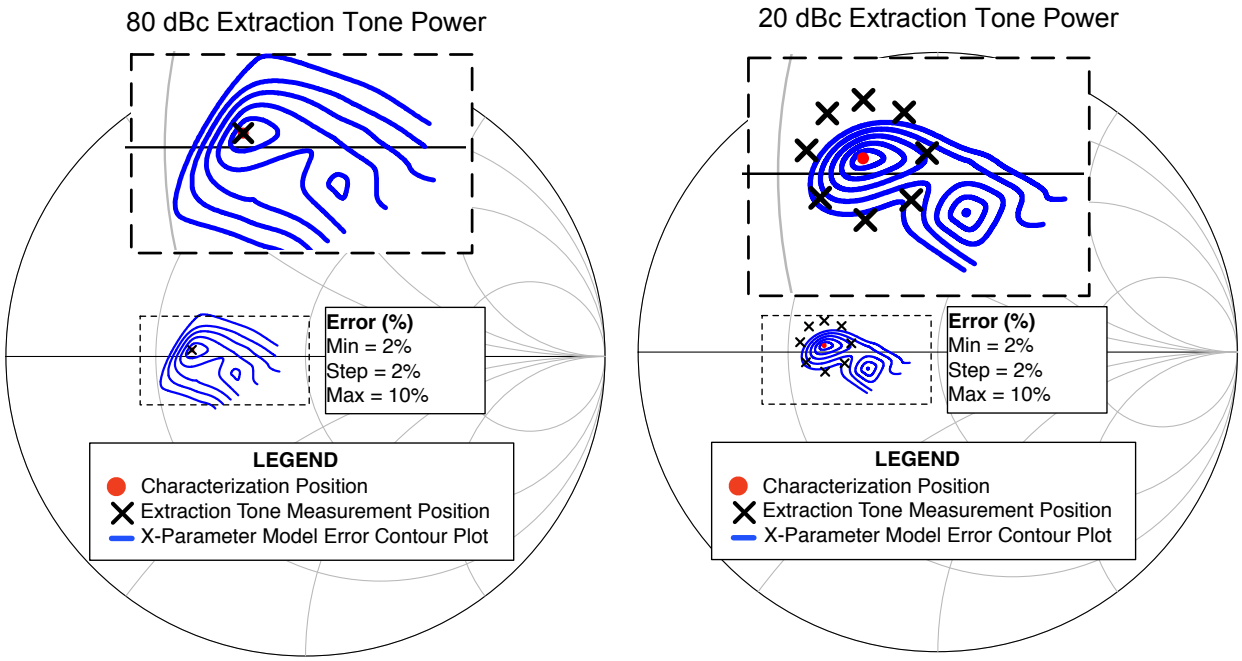
performance, while all other variables produce a weakly nonlinear response. However, Fig. 4.4 hypothesizes that all RF inputs contribute to the cumulative load-line behaviour, and Fig. 4.2 and Fig. 4.3 clearly show that $|A_{11}|$, A_{12} and A_{21} all contribute to the nonlinear behaviour when the DUT operates in cut-off or triode mode. This suggests that if only one input variable generates a strongly nonlinear response, *all* other RF inputs will also become strongly nonlinear, and all variables become LSOPs immediately. If all variables become LSOPs, the modelling prediction becomes extremely localized, and the Load-Dependent X-Parameter model will degenerate to a multi-dimensional source/load-pull LUT. Furthermore, increasing the number of dimensions in the LUT (LSOPs), will demand an increased density in the LUT to compensate for limitations of inter-model interpolation techniques.

Obviously, the accuracy of the PHD approximation depends purely on the intended application as demonstrated in Fig. 4.1. In this example, the low-pass pre-matching networks in Fig. 4.1a reduce the impact of harmonic tuning, thus decreasing the nonlinear response to all harmonic stimuli. Unfortunately, the general purpose transistor in Fig 4.1b suggests a_{12} (and perhaps other harmonic inputs) should be added to the list of LSOPs and this device may be difficult to predict using the PHD model.

Limitation of PHD Model Extraction

The practical extraction of the PHD model, implemented by the X-Parameter model, uses a multi-harmonic injection signal called the Extraction Tone (ET) to resolve the model coefficients. Based on the formulation in (2.23), a ET must be applied sequentially at all harmonics and must swept over multiple values so that two unknowns can be solved by applying a single ET. In theory, the ET signal is an infinitesimal signal that is applied to approximate a first-order numerical derivative. In reality, the ET signal and its response must be large enough to be generated by the signal-source and detectable by the receiver. On the other hand, the ET signal should be small enough so that is does not produce a nonlinear response and so that extraction measurements remain within the valid proximity of the LSOP. The impact of the ET magnitude on model accuracy was simulated in Fig. 4.7. First, an infinitesimal ($-80dBc$) ET was used to extract the $X^S_{21,21}$ and $X^T_{21,21}$ model coefficients by accurately approximating a numerical first-order derivative. Although the

resulting output power error in Fig. 4.7a is minimized around the LSOP the error increases rapidly when the load-impedance diverges from the LSOP. Alternatively, when the ET is $-20dBc$ below the output signal, the average error is reduced because the model extraction distribution is more evenly distributed over the model measurement space. When using a low-degree function to approximate a high-order nonlinearity, average error can be improved by extracting the secant of a curve, rather than the tangent, however the secant model coefficients do not reflect the first-order expansion formulation of the PHD model. This approach is dangerous because it reduces the average error by manipulating the error distribution, as shown in Fig. 4.8, so that it no longer reflects the normal error distribution of the intended application.



(a) 80dBc Extraction Tone

(b) 20dBc Extraction Tone

Figure 4.7: X-Parameter Model Output Power Error vs Γ_{21}

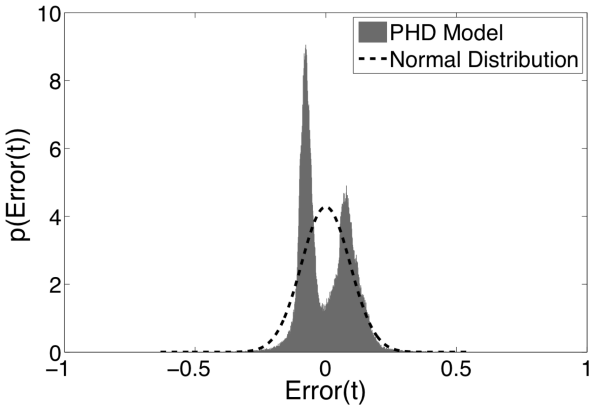


Figure 4.8: PHD Model Output Power Error Distribution using a $-20dBc$ Extraction Tone

4.2.3 Limitations of the Cardiff Model

Limitation of the Cardiff Model Formulation

Unlike the superposition formulation of the PHD model, which is conceptually difficult to apply to stronger nonlinear systems, the Cardiff model describes the system as a nested set of functions, as shown in (4.5).

$$y = f_1\left(x_1, f_2\left(x_2, ...\right)\right) \tag{4.5}$$

This formulation is more suitable for describing high-order non-linearities, however each outer function is faced with the increasingly daunting task of conditioning the parameters of its nested functions. Hence, this formulation is best suited to problems that do not have many inputs, and in a multi-harmonic frequency-domain problem, the number of inputs will be dictated by the available bandwidth of the DUT and the number of ports.

Delving deeper into the variants of the Cardiff model, it is apparent that nested functions are split into two categories: i) Fourier Series functions that describe the response to phase variations of incident-waves ($\phi(A_{jl})$), and ii) polynomial functions that describe the response to magnitude variations of incident-waves ($|A_{jl}|$). The phase model function in (2.27), describes the active load-pull reflected-wave response in Fig. 4.2 and Fig. 4.3 using a n^{th} order Fourier Series for each magnitude of ($|A_{jl}|$). The number of coefficients (n) in

the phase model is determined by the nonlinearity at the highest magnitude, however some of the coefficients will degenerate to zero at lower magnitudes of A_{jl}. A subsequent magnitude model, describes the variation of inter-phase model coefficients using the polynomial function described in (2.28). As each phase-model relies on the same number of coefficients (n), over describing the phase model at the highest magnitude can possibly result in poorly conditioned models at low magnitudes. Thus accurately modeling a high-order nonlinearity would result in poor extrapolation at lower powers and vice-versa.

Extending the Cardiff model to multi-harmonic and multi-port problems is accomplished by increasing the number of nested functions in (4.5). While this is theoretically possible, the number of model coefficients would increase dramatically with the number of inputs unless some of these nested functions were assumed to be separable. The nested functions could only be separable if the response of some inputs was assumed to be linear, however Fig. 4.2 and Fig. 4.3 demonstrates that the load-line response to each input is inter-related once the transistor enters cut-off or triode operation. As the system becomes more nonlinear, the functions would no longer be separable and the number of model coefficients would increase dramatically. Therefore, although the Cardiff model formulation provides superior higher-order modeling, it is difficult to scale the model to a larger number of harmonics and ports.

Limitation of Cardiff Model Extraction

Extraction of this model is complicated due to the calibrated multi-harmonic active source/load-pull system that is needed to sweep the incident-waves in a polar sweep distribution. The order of the nonlinear system will determine the shape of the active load-pull contours, thus determining the density of points that are needed during the model extraction. Above all, the greatest limitation of the Cardiff model is the dramatic increase in measurements when modeling a higher number of harmonics or ports. A more thorough comparison of the PHD model and Cardiff Model was presented by my colleague Amir Amini and can be found in [63].

4.2.4 Model Limitations Comparison

Each of the model formulations described in this section has different limitations, therefore a summary of the trade-offs between each model formulation is provided in Table 4.2.

Table 4.2: Model Formulation Comparison

Functionality	PHD Model	Load-Dependent X-Parameter Model	Multi-Harmonic Volterra Model	Cardiff Model
# of Coefficients	Low	High	High	Highest
# of Swept Measurements	Low	High	High	Highest
Complexity Increase with # of inputs (N)	(N)	$(N)\#\ of\ Loads$	???	$(N)^{\#ofInputs}$
Formulation Type	$\sum f_i(x_i)$	$LUT\left[\sum f_i(x_i)\right]$	$\sum f_i(x_i, g_j(x_{DC}))$	$f_1(x_1, f_2(x_2, ...))$
Order of Nonlinearity	Lowest	High	Highest	Highest
Condition Number (κ)	Lowest	Lowest	Highest	Moderate
Extrapolation	Good	Good	Moderate	Poor
Active Load-Pull Complexity	Lowest	Lowest	High	Highest

The PHD model and the Cardiff model represent two extremes: the PHD model provides mild nonlinear prediction using the least number of measurements, while the Cardiff model provides strong nonlinear prediction using the largest number of measurements. The

PHD model is not optimal because it doesn't predict higher-order nonlinearities, while the Cardiff Model is not optimal because it would require too many measurements. The Load-Dependent X-Parameter model may seem like an easy compromise, however Fig. 4.7 demonstrates that it cannot be practically implemented for higher-order nonlinear systems. While the MHV model theoretically offers the same accuracy as the Cardiff Model, it's formulation is primarily based on the superposition of mixing products. Due to its *variable-order* nature, it is theorized that the MHV model will require less kernels (and extraction measurements) than the Cardiff model. Therefore, the PHD model and the MHV model have been selected as complimentary solutions, providing weak and strong nonlinear prediction, using simple and complex test-bench architectures, as illustrated in Fig. 4.9

It is possible that a realistic approximation of the MHV model could be extracted over a localized measurement space if we could find a practical algorithm for removing insensitive mixing terms from the formulation. To synthesize a successful model, we must limit the number of linearly independent model kernels and the number of extraction measurements for a given number of system inputs. The number of system inputs is the product of the number of ports of the DUT, and the number of harmonics below f_t, as shown in Fig. 2.10b. Including a model kernel with many inputs will produce a poorly conditioned problem matrix and will exponentially increase the number of measurements that are required to extract that coefficient. Including a model kernel with a high-order input exponent will also increase the number of measurements that are needed to extract a stable numerical coefficient. The equation in (4.6) is an example kernel that consists of of three kernel inputs with an order of five.

$$\text{Mixing Term} \quad = \quad a_{11}a_{12}^{*2}a_{21}^{2} \tag{4.6}$$

Fig. 4.10 plots the number of model kernels versus input harmonics for the PHD model and the MHV model. Fig. 4.10b prunes the kernels based on the maximum order and the maximum number of inputs. A database of these kernels has been generated using search algorithms so that pre-determined models can be loaded during runtime based on the choice of $NumPorts$, $aHarmonics$, $bHarmonics$, $MaxOrder$, and $MaxInputs$. By comparing the number of kernels in each model, it is apparent that exploring higher-order

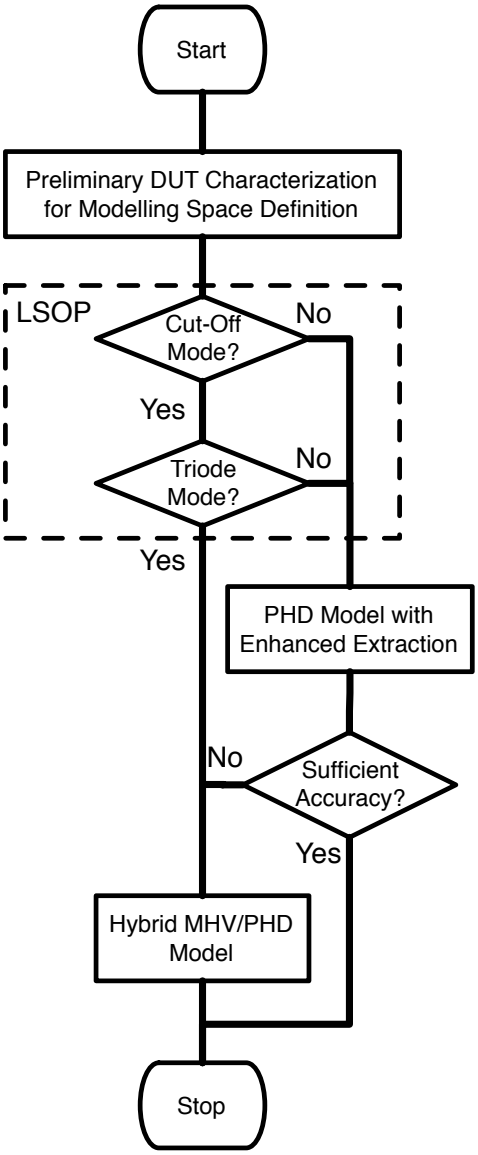


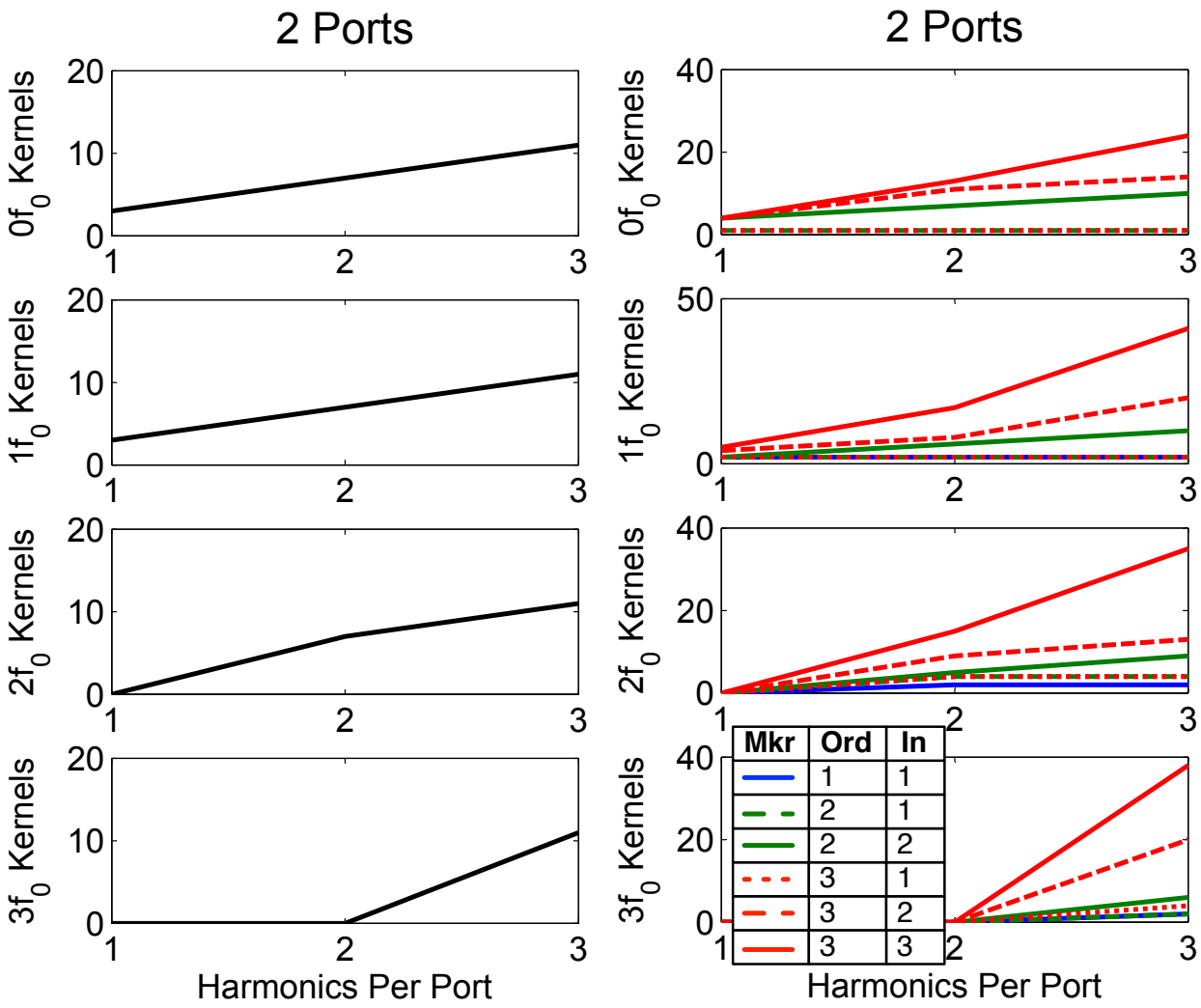
Figure 4.9: Proposed Model Formulation Selection Process Flow Diagram

mixing behaviour of the DUT will require many more kernels. As the number of inputs is pre-determined by the f_t of the DUT, we need to constrain the number of kernels for larger numbers of inputs by using Parameter Subset Selection to prune kernels with a large order or many input dependencies. This practical compromise is to develop a localized model that approximates the nearby behaviour using a lower number of kernels.

Each MHV model consists of multiple sub-models for each output harmonic, on each port ($\#SubModels = bHarmonics \times NumPorts$), therefore it would be inefficient to generate separate measurements for each sub-model. As the number of input harmonics and ports increases, it is also impractical to generate a single multi-dimensional sweep of the input parameters, because this would require $(MaxOrder + 1)^{2 \times aHarmonics \times NumPorts - 1}$ measurements. Therefore, a search algorithm is used to translate the kernels into a series of sweep plans, so that the number of measurements is minimized, resulting in the number of measurements shown in Fig. 4.11b. As before, the number of measurements increases dramatically with the number of inputs, however this can be mitigated by filtering the kernels based on the order and the number of kernel inputs. For the sake of comparison, the PHD model only consists of kernels with one input, hence the number of measurements will increase linearly with the number of system inputs, as shown in Fig. 4.11a.

4.3 PHD Model Accuracy Enhancement by Improving Harmonic Response Superposition During Extraction

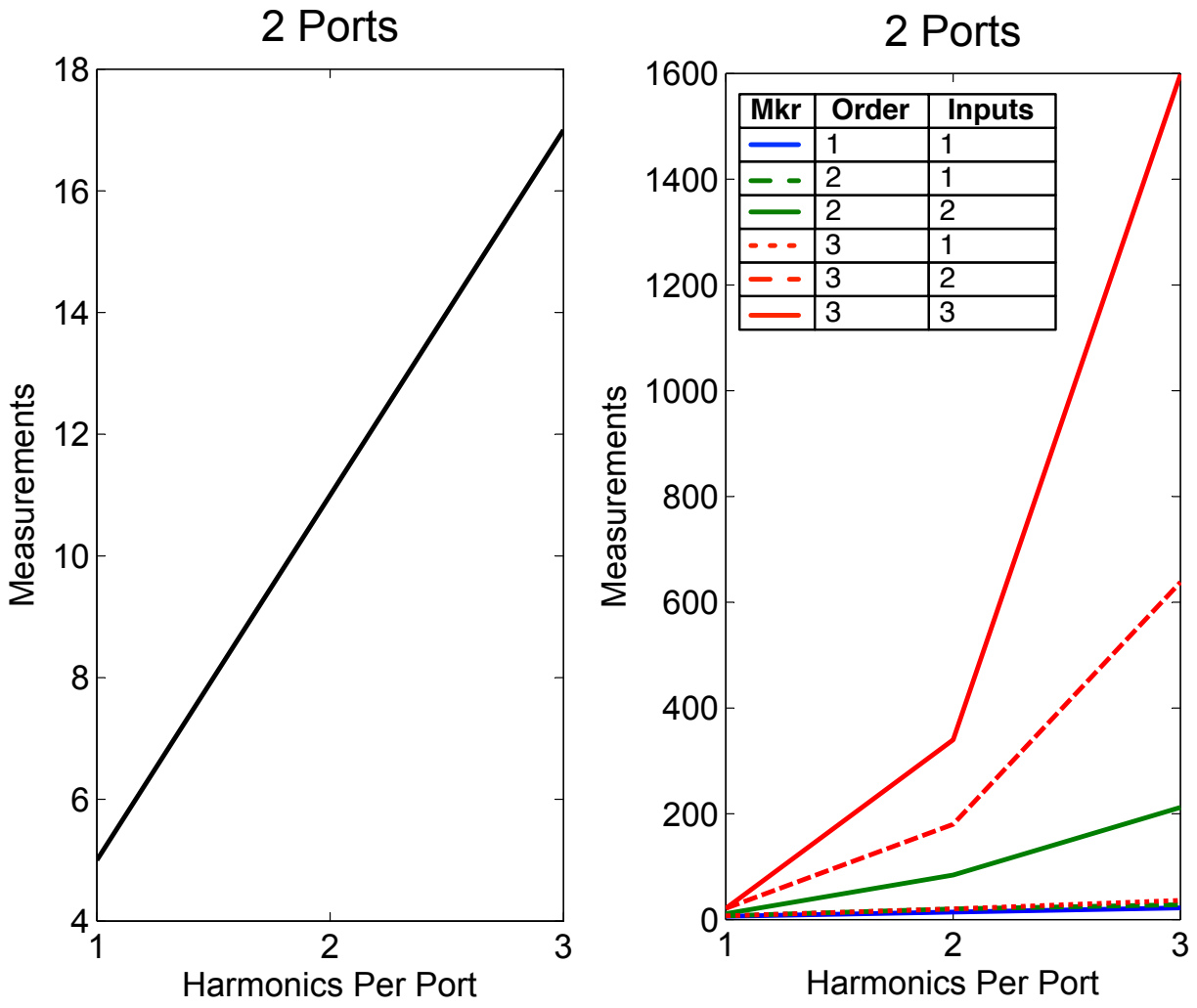
The harmonic superposition assumption greatly decreases the number of measurements by assuming that each harmonic matching condition is separable, however poor matching during model extraction can result in output power-waves reflecting back into the system as hidden variables. To demonstrate this problem, the LSOP in section 4.1.2 will be used to compare a matched and unmatched PHD model extraction. The impedances synthesized during extraction of a 2-port, 3-harmonic PHD model are tracked and the coverage of Γ_{12} (normalized to the passive port impedance Z_{12}) is plotted in Fig. 4.12. Under conjugate



(a) PHD Kernels

(b) Multi-Harmonic Volterra Kernels

Figure 4.10: Number of Kernels per Harmonic Input



(a) PHD Measurements

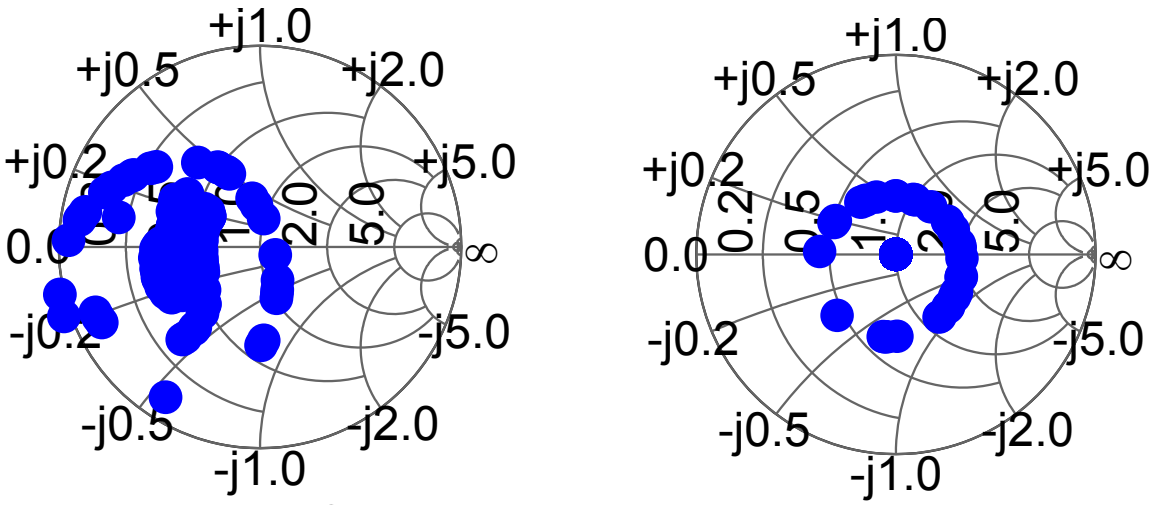
(b) Multi-Harmonic Volterra Measurements

Figure 4.11: Number of Measurements per Harmonic Input

matched conditions in Fig. 4.12b, $\Gamma_{1l} = 0$ when all other harmonic/port extraction signals are applied, resulting in a concentration of points in the center of the Smith Chart. When a_{12} is applied at a constant amplitude, variable phase shift, the b_{12} response is weakly non-linear, resulting in a circular-like impedance modulation. Alternatively, when the source impedance is mismatched ($\Gamma_{1l} = 0.5$), in Fig. 4.12a, the load-modulation occurs even when the extraction signal is applied at other harmonics/ports. Also, when a_{12} is applied, the impedance modulation appears to be more distorted than in Fig. 4.12b. Therefore, this study suggests that the PHD model formulation is no longer satisfied when any harmonic impedance is mismatched, because $a_{12} \neq 0$ is injected into the system as a hidden variable when other input signals are applied. To quantitatively prove this theory, the PHD model was extracted using the same LSOP synthesized with: i) $\Gamma_{1l} = 0$ and ii) $\Gamma_{1l} = 0.5$, over 3-harmonics by adjusting A_{1l} to correct for the mismatch. The results shown in Table 4.3 demonstrate the accuracy of the PHD model can be improved by $\sim 5dB$ when extracted under a multi-harmonic conjugate matched source impedance. Although it is more difficult to synthesize the same LSOP under multi-harmonic load impedance mismatch, it is obvious that further degradation of model accuracy would occur.

Table 4.3: Model Extraction Validation

	Accuracy [dB]	**Existing PHD Model Extraction**	**Proposed PHD Model Extraction**
I	$0f_0$	-20.3	**-25.5**
P_{OUT}	$1f_0$	-23.4	**-28.3**
	$2f_0$	-7.38	**-10.0**
	$3f_0$	-18.0	**-23.0**



$Z_0 = 50\Omega$

(a) Unmatched ($\Gamma_{1l} = 0.5$)

$Z_0 = 16.7\Omega$

(b) Matched

Figure 4.12: PHD Model Extraction Measurement Coverage of Γ_{12}

4.4 High-Order Model Looking Beyond the First-Order Harmonic Superposition Assumption

The MHV model can predict higher-order nonlinear systems where the PHD model is no longer valid. The change in model formulation comes at the cost of added hardware complexity, namely dedicated RF sources are required at each harmonic input on all ports. The additional coefficients of the MHV model are less linearly independent that the well-conditioned PHD model, therefore care should be taken to fit the nonlinear system with the lowest-order formulation possible. The LSOP described in section 4.1.2 was used to compare the accuracy of the PHD model in section 4.3 with the MHV model. The model extraction settings in Table 4.4, demonstrate that the MHV model requires almost $10X$ more measurements to extract.

Table 4.4: Model Comparison Extraction Settings

Settings	PHD Model	MHV Model	Unit
Freq	2.06	2.06	GHz
NumPorts	2	2	
aHarmonics	3	3	
bHarmonics	3	3	
VDC	$[-2.92, 28]$	$[-2.92, 28]$	V
ZRF	$[16.7 - 0.725i, 5.01 + 0i;$ $1.28 + 0i, 2.64 - 2.18i;$ $2.63 + 0i, 2.82 - 13.4i]$	$[16.7 - 0.725i, 5.01 + 0i;$ $1.28 + 0i, 2.64 - 2.18i;$ $2.63 + 0i, 2.82 - 13.4i]$	Ω
ARF	$[0.708, 3.16\angle 180]$	$[0.708, 3.16\angle 180]$	$\sqrt{W}$
aRFStep	$[0.2, 2.5]$	$[0.2, 2.5]$	$\sqrt{W}$
Oversampling	10	1	
NumMeasurements	170	1598	

Once a multi-harmonic behavioural model has been extracted, validating the model is not a straightforward process. Although the model can be validated using power sweeps or load-pull contours, each of these validations exercise a fraction of the model coefficients.

Furthermore, the harmonic superposition assumption of the PHD model is not tested if validations only vary one input variable at a time. To compare the accuracy of the MHV model with the PHD model, we need to generate a validation procedure that is independent of both model extractions[64]. To approximate a practical design application the verification must must include the following characteristics:

- All input conditions are equally likely due to the randomness of the unspecified matching conditions.
- It is not sufficient to validate against each input separately.
- To avoid extrapolated results, the model comparison must remain within the Time-Domain and Frequency domain measurement space of where the model was extracted.

The last condition is difficult to achieve because the time-domain and frequency-domain measurement space is primarily determined by the number of simultaneous system inputs that are stimulated during model extraction. As the number of simultaneous inputs increases, the amount of variation in the time-domain reflected power-waves increases, however the frequency-domain Smith Chart coverage remains the same.

To compare the model accuracy, each model has been extracted around the same LSOP using the settings in Table 4.4. While it is traditionally customary to validate a model using an input power-sweep or load-pull measurement, each of these measurements would only exercise a fraction of the model coefficients. Hence, an independent characterization is performed to compare the model prediction to the actual measurement using a multi-dimensional input power sweep outlined in Table 4.5. This validation procedure exercises all model coefficients simultaneously, validating the assumptions of harmonic signal superposition, hence it more accurately approximates the final PA design application.

The NMSE of each sub-harmonic model is listed in Table 4.6, and the condition number, representing the instability of the model coefficients, is presented in Table 4.7. The results suggest that the MHV model can achieve up to $4dB$ improvement in fundamental frequency output power prediction and up to $5dB$ improvement in DC power prediction. Unfortunately, the accuracy of the harmonic models degrades rapidly because important

Table 4.5: Model Comparison Sweep Plans

Sweep Name	Sweep Plan
harmonic	TYPE: linear, START: 0, STOP: 3, STEP: 1, POINTS: 4
port	TYPE: linear, START: 1, STOP: 2, STEP: 1, POINTS: 2
$\Re(a_{11})$	TYPE: linear, START: -0.05, STOP: 0.05, STEP: 0.10, POINTS: 2
$\Re(a_{12})$	TYPE: linear, START: -0.035, STOP: 0.035, STEP: 0.07, POINTS: 2
$\Im(a_{12})$	TYPE: linear, START: -0.035, STOP: 0.035, STEP: 0.07, POINTS: 2
$\Re(a_{13})$	TYPE: linear, START: -0.035, STOP: 0.035, STEP: 0.07, POINTS: 2
$\Im(a_{13})$	TYPE: linear, START: -0.035, STOP: 0.035, STEP: 0.07, POINTS: 2
$\Re(a_{21})$	TYPE: linear, START: -0.35, STOP: 0.35, STEP: 0.7, POINTS: 2
$\Im(a_{21})$	TYPE: linear, START: -0.35, STOP: 0.35, STEP: 0.7, POINTS: 2
$\Re(a_{22})$	TYPE: linear, START: -0.35, STOP: 0.35, STEP: 0.7, POINTS: 2
$\Im(a_{22})$	TYPE: linear, START: -0.35, STOP: 0.35, STEP: 0.7, POINTS: 2
$\Re(a_{23})$	TYPE: linear, START: -0.35, STOP: 0.35, STEP: 0.7, POINTS: 2
$\Im(a_{23})$	TYPE: linear, START: -0.35, STOP: 0.35, STEP: 0.7, POINTS: 2

high-order, multi-input kernels have been pruned using Parameter Subset Selection. Therefore, a Hybrid model has been synthesized using the DC/fundamental MHV sub-models and the PHD harmonic sub-models.

Table 4.6: Normalized Mean-Squared Error Model Comparison

		$NMSE$ $[dB]$	**Proposed PHD Model**	**MHV Model**	**Hybrid Model**	**Hybrid* Model**
Frequency	I	$0f_0$	-24.5	**-29.7**	**-29.7**	-29.4
	P_{OUT}	$1f_0$	-21.7	**-23.8**	**-23.8**	-22.9
		$2f_0$	**-2.25**	2.26	**-2.25**	**-2.25**
		$3f_0$	**-18.7**	-10.7	**-18.7**	**-18.7**
Time	P_{OUT}		-20.6	-19.2	**-21.4**	-20.8

Table 4.7: Condition Factor (κ) Model Comparison

	κ	**Proposed PHD Model**	**MHV Model**	**Hybrid Model**	**Hybrid* Model**
P_{OUT}	$0f_0$	**75.9**	379	379	150
	$1f_0$	**75.9**	480	480	99
	$2f_0$	**75.9**	289	**75.9**	**75.9**
	$3f_0$	**75.9**	2120	**75.9**	**75.9**

It is difficult to determine how the accuracy of each frequency-domain model will impact the overall accuracy of the time-domain waveforms, hence an Inverse Discrete Fourier Transform ($DFT^{-1}(b_{pik})$) is used to calculate the RF time-domain NMSE (4.7 - 4.10), shown in Table 4.6. Due to the higher relative amplitude of the fundamental frequency signal, the $NMSE_i(t)$ is weighted strongly by the accuracy of the fundamental frequency

models. The relative time-domain error is plotted in Fig. 4.14 for all models, and the results suggest the Hybrid model provides superior prediction in most cases due to the MHV fundamental frequency model. The results in Table 4.6, suggest the RF time-domain output signal NMSE is improved by $\sim 2.1dB$ based on 95% of the measurement data, based on the error distribution presented in Fig. 4.13. Unlike the PHD model error distribution in Fig. 4.8, the Hybrid model contains a normal error distribution more accustom to random utilization inside a realistic design application. All of theses models are valid inside the frequency-domain coverage region described in Fig. 4.15. Whereas most harmonic Smith Charts contain full passive impedance coverage, the input 2^{nd} harmonic (Z_{12}) is an exception because the reflected power-wave B_{12} is proportional to the fundamental input power (A_{11}). This happens because the equivalent circuit C_{gs} is strongly nonlinear. Hence, applying the proposed model to predict the global measurement space would require a LUT in terms of A_{11}, A_{21}, and A_{12}. As this model was extracted under the strongest nonlinear LSOP, it is suggested that other LUT models could either encompass a larger measurement space coverage (adaptive LUT distribution), or use an adaptive order model (adaptive LUT model formulation). Either one of these LUT methods could be used to capture the entire measurement space. The number of measurements used to extract each model is shown in Table 4.8.

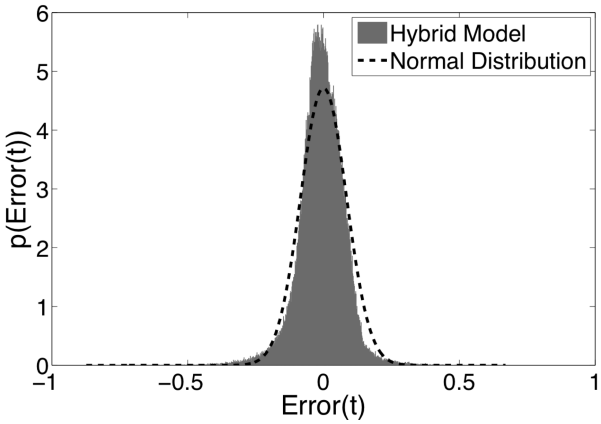


Figure 4.13: Hybrid Model Output Power Error Distribution

$$error_i(t) = \frac{\left|\left|\hat{b}_i(t) - b_i(t)\right|\right|_2^2}{||b_i(t)||_2^2} \tag{4.7}$$

$$error_i(t)\,[dB] = 10 * \log(error_i(t)) \tag{4.8}$$

$$NMSE_i = mean(error_i(t), t) \tag{4.9}$$

$$NMSE_i\,[dB] = 10 * \log(NMSE_i) \tag{4.10}$$

Where,

i is the port index.

$\hat{b}_i(t)$ is the time-domain reflected power-wave predicted by the model.

$b_i(t)$ is the time-domain reflected power-wave measured during validation.

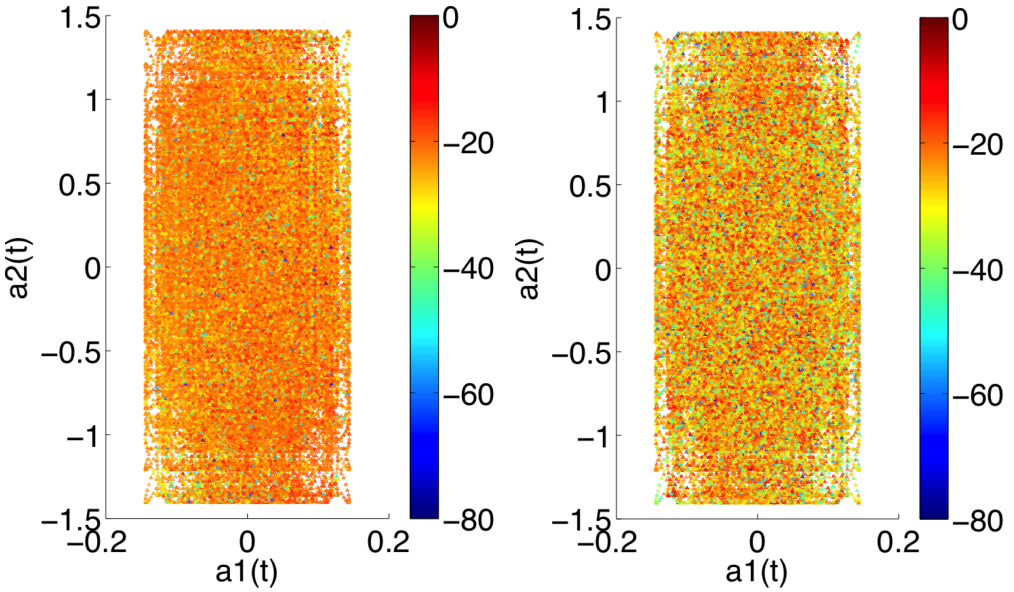
$error_i(t)$ is the time-domain relative error in the output power.

$NMSE_i$ is the Normalized Mean Squared Error of the time-domain reflected power-wave.

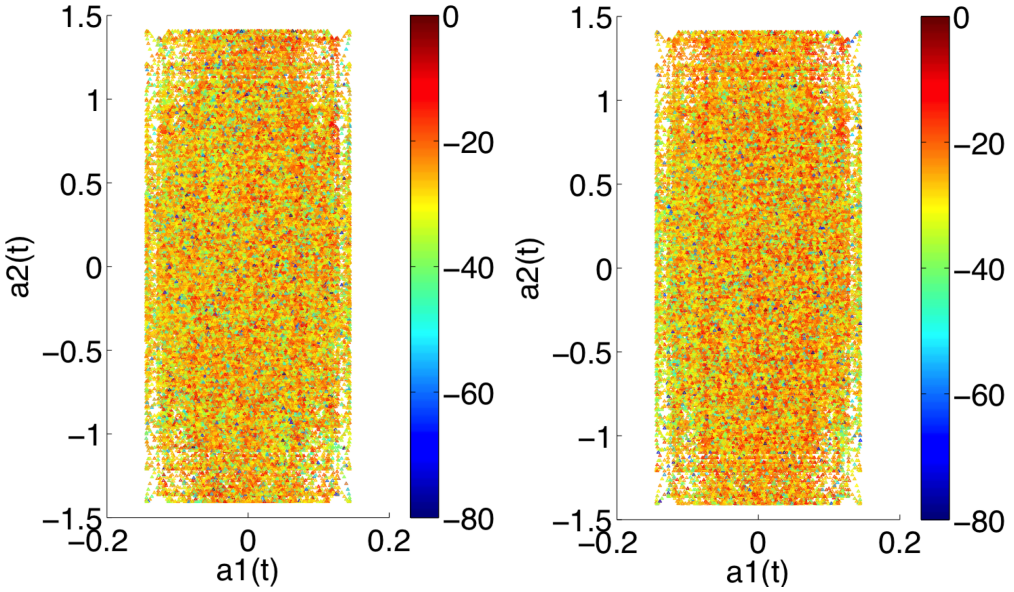
Table 4.8: Model Extraction Measurements

	PHD	**MHV**	**Hybrid**	**Hybrid***
Measurements	17	1598	1192	1192

Since the kernels are not orthogonal, adding additional mixing terms will result in a degradation of the condition number of the problem matrix. Even if the additional mixing terms provide better curve-fitting of the output data, a very large condition number implies a wider variation (larger error) in the model coefficients that predict the output. Fig. 4.16 illustrates that error in the output, such as measurement noise, produces forwards error (4.11), alternatively the same error in the output can be attributed to backwards error (4.12), the error in model coefficients. A well-conditioned model offers stable prediction of



(a) PHD (b) MHV



(c) Hybrid (d) Hybrid*

Figure 4.14: Time-Domain Normalized Reflected Power-Wave Error in [dB]

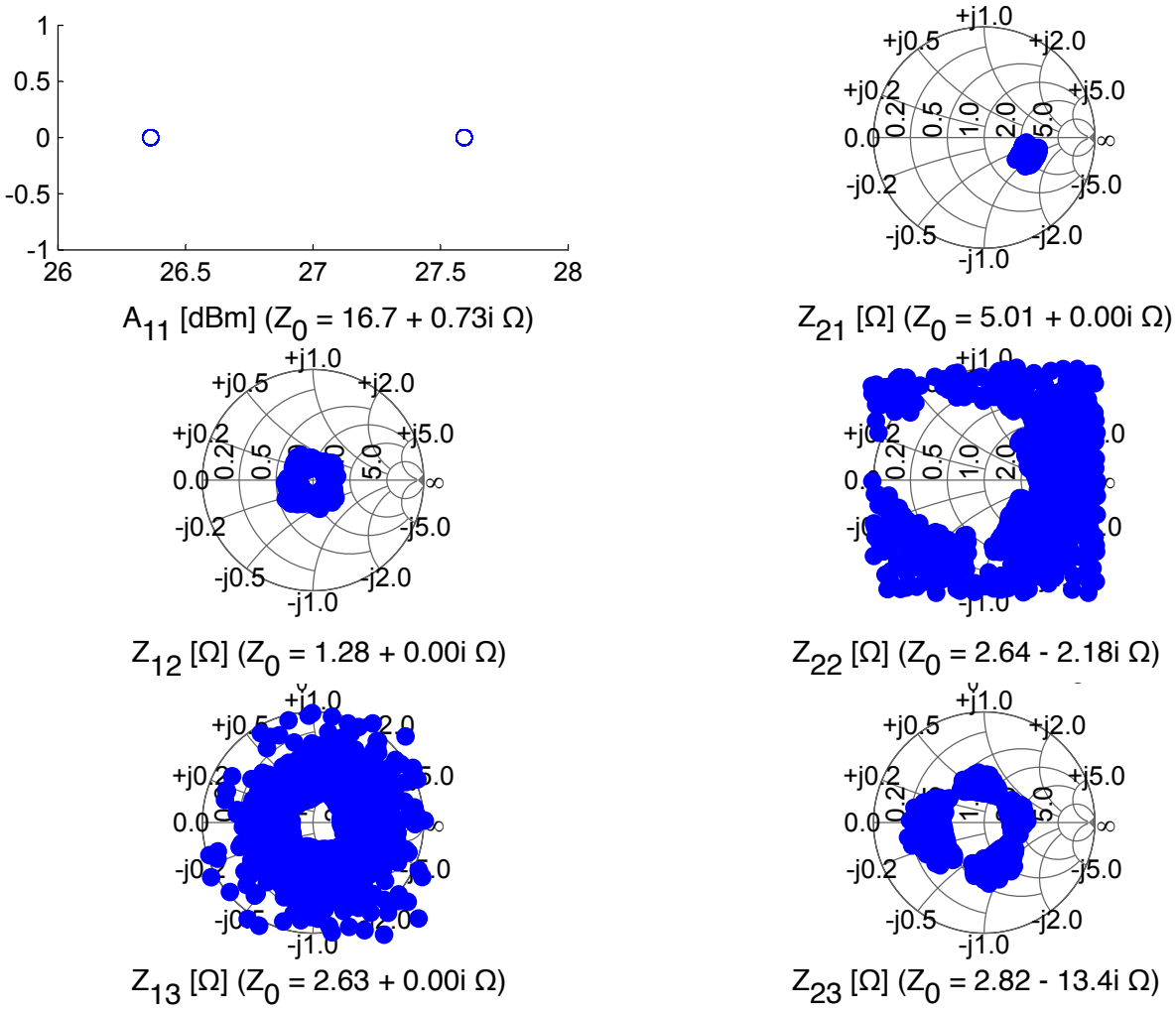


Figure 4.15: Frequency-Domain Coverage Model Comparison

.

signals that were not used to train the model and it provides better extrapolative qualities that are essential for this type of localized model. Hence, the backwards error was minimized by reducing the product of the output residual error multiplied by the condition number (4.14). Minimizing the inequality in (4.14) does not necessarily minimize (4.11), hence some output error must be sacrificed to train better model coefficients as illustrated in Fig. 4.17. The kernels in Fig. 4.17 are sorted from most relevant to least relevant, hence the last two kernels have little impact on the accuracy of the model. The moderate depredation of accuracy is illustrated under the Hybrid* column of Fig. 4.6 and the improvement in condition number is shown in Fig. 4.7. The proposed backwards minimization technique reduces over-fitting of higher-order models, providing the model with better ability to extrapolate.

$$\Delta y = \hat{y} - y \tag{4.11}$$

$$\Delta x = \hat{x} - x, \text{ where } f(\hat{x}) = \hat{y} \tag{4.12}$$

$$\kappa(A) <= \frac{||\Delta x||}{||x||}\frac{||y||}{||\Delta y||} \tag{4.13}$$

$$\frac{||\Delta x||}{||x||} <= \kappa(A)\frac{||\Delta y||}{||y||} \tag{4.14}$$

Where,

x is the actual value of the input.

$\hat{x}$ is the predicted value of the input.

y is the actual value of the output.

$\hat{y}$ is the predicted value of the output.

$\kappa(A)$ is the condition number of the LSE problem matrix.

This chapter described a MIMO behavioural modelling solution that is built on top of the characterization system described in Chapter 3. It proposed a methodology for

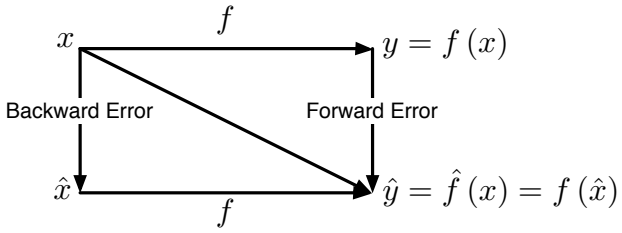


Figure 4.16: Forward and Backward Error in a Least Squares Solution

Backwards Error Minimization	**Disabled**	**Enabled**
MHV Kernels	Error [%]	Error [%]
A_{21}	3.9	2.02
A_{11}	80.7	2.05
$A_{11}^2 A_{22}^*$	33.9	10.55
$A_{21}^* A_{21}^2$	79.7	10.70
$A_{21}^* A_{22}$	130.6	24.37
$A_{11}^* A_{22}$	104.0	24.53
$A_{11}^* A_{12}$	91.6	89.70
$A_{12} A_{21}^*$	104.2	106.39
Average	**78.6**	**33.8**

Figure 4.17: Backwards Error Minimization of $B_{2,1}$ Model Coefficients of 2-Harmonic Hybrid Model

applying the PHD model formulation to predict unmatched nonlinear devices by using multi-harmonic impedance tuners to minimize mismatch within the system. The model extraction procedure outlined in section (4.1.2) and section (4.3).

Accurate transistor models, based on realistic large-signal measurements, are vital for designing modern PAs that are implemented using complex design architectures. Chapter 5 provides an example of a concurrent dual-band PA, designed using filter theory, whose design can benefit extensively from the nonlinear characterization and modelling methods that have been developed thus far.

4.5 Towards Nonlinear Power Amplifier Design

The complimentary PHD model (section 4.3) and Hybrid MHV model (section 4.4) can be used in a PA design application. These models can be imported into a circuit simulation CAD tool and the expected variation in design performance is outlined in Table 4.9, based on the equations derived in (4.15-4.24).

$$\frac{\delta Y_1/\delta Y_2}{Y_1/Y_2} = \frac{\delta Y_1}{Y_1} + \frac{\delta Y_2}{Y_2} \tag{4.15}$$

$$\frac{\delta P_L}{P_L} = 10 * \log\left(\sqrt{NMSE(B_{21})^2 + NMSE(A_{21})^2}\right) \tag{4.16}$$

$$= 10 * \log\left(NMSE(B_{21})\right) \tag{4.17}$$

$$\delta P_L = 10 * \log\left(mean(P_L)\right) \pm 10 * \log\left(1 + \frac{\delta P_L}{P_L}\right) + 30 \tag{4.18}$$

$$\frac{\delta G_P}{G_P} = 10 * \log\left(\sqrt{NMSE(B_{21})^2 + NMSE(A_{11} - B_{11})^2}\right) \tag{4.19}$$

$$= 10 * \log\left(\sqrt{NMSE(B_{21})^2 + NMSE(B_{11})^2}\right) \tag{4.20}$$

$$\delta G_P = 10 * \log\left(mean(G_P)\right) \pm 10 * \log\left(1 + \frac{\delta G_P}{G_P}\right) \tag{4.21}$$

$$\frac{\delta \eta}{\eta} = \sqrt{NMSE(B_{21})^2 + NMSE(I_{20}V_{20})^2} \tag{4.22}$$

$$= \sqrt{NMSE(B_{21})^2 + NMSE(I_{20})^2} \tag{4.23}$$

$$\delta \eta = mean(\eta) \pm mean(\eta)\frac{\delta \eta}{\eta} \tag{4.24}$$

Where,

Y is any frequency-domain output predicted by the model.

$NMSE(Y_{ik})$ is the normalized mean squared error (the standard deviation in W) of a voltage/current/power (in $V/I/\sqrt{W}$) at port i, harmonic k.

$\frac{\delta P_L}{P_L}$ is the relative error in the power delivered to the load in dB.

δP_L is the absolute error in the power delivered to the load in dBm.

$\frac{\delta G_P}{G_P}$ is the relative error in the power gain in dB.

δG_P is the absolute error in the power gain in dB.

$\frac{\delta \eta}{\eta}$ is the relative error in the drain efficiency in %.

$\delta \eta$ is the absolute error in the drain efficiency in %.

Table 4.9: Design Performance Model Comparison

Performance Metric	**Proposed PHD Model**	**MHV Model**	**Hybrid Model**	**Hybrid* Model**
$\frac{\delta P_L}{P_L}$ $[dB]$	-21.7	**-23.8**	**-23.8**	-22.9
$\delta P_L[dBm]$	46.7 ± 0.0295	**46.7 ± 0.0181**	**46.7 ± 0.0181**	46.7 ± 0.0222
$\frac{\delta G_P}{G_P}$ $[dB]$	-18.5	**-21.9**	**-21.9**	-20.6
$\delta G_P[dB]$	21.3 ± 0.0610	**21.3 ± 0.0280**	**21.3 ± 0.0280**	21.3 ± 0.0378
$\frac{\delta \eta}{\eta}$ [%]	0.770	**0.430**	**0.430**	0.520
$\delta \eta$[%]	76.0 ± 0.580	**76.0 ± 0.330**	**76.0 ± 0.330**	76.0 ± 0.400

The relative error in Table 4.9 demonstrates that the MHV/PHD Hybrid model provides roughly a $2X$ improvement in power delivered to the load, power gain, and drain efficiency prediction. The absolute error in Table 4.9 provides specific error tolerances for a $60W$ PA with a gain of $20dB$, however absolute error tolerances do not transfer to other design applications. Therefore, it is more scientifically correct compare behaviour models using relative tolerances. While the MHV Hybrid model provides drain efficiency prediction below 0.5%, the greatest benefit is the $3.4dB$ improvement in vector power-gain prediction

that results in better pre-fabrication prediction of PA linearization techniques. Continued improvements in the MHV model extraction, utilizing higher-order approximations, will enable complete end-to-end design emulation inside a CAD simulation tool by using the complimentary modelling solution outlined in Fig. 2.9 to ensure "first-pass" nonlinear circuit designs.

Chapter 5

Power Amplifer Design Example: Concurrent Dual-Band Power Amplifier

5.1 Multi-Standard Power Amplifier Design Techniques

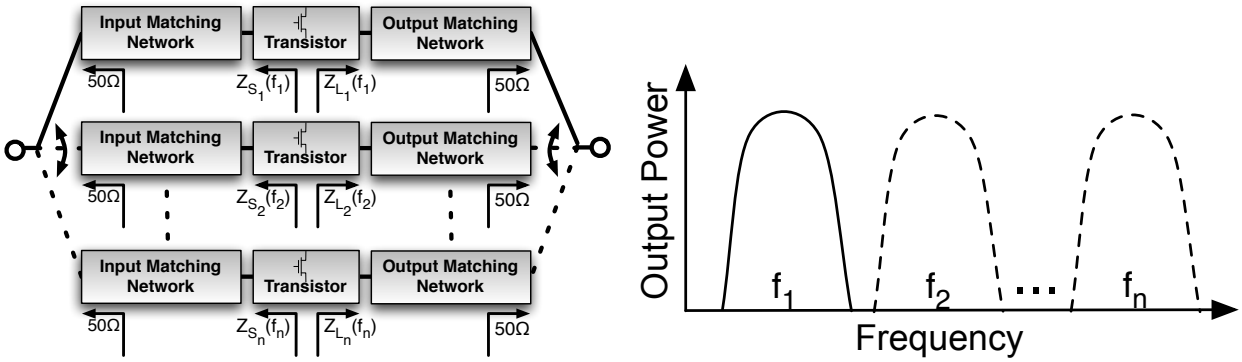
5.1.1 Multi-Band Power Amplifier Design Techniques

The proliferation of wireless standards has led network providers to simultaneously broadcast signals over multiple radio bands. The mandate to support both old and new standards has made it expensive for network providers to upgrade base-stations to support the latest technology. There has been a lot of work directed towards building multi-band components for RF front-ends. Most of this work has focused on developing multi-band passive components such as filter theory and power dividers, but little work has been done on multi-band RF power amplifiers. Some the approaches towards designing frequency-agile RF front-ends are described in the following section [65].

Switched Single-Band Power Amplifiers

In this design, a PA is designed for each frequency band and a switch is used to change between them. This design has the following problems:

- Non-concurrent frequency operation
- Large power consumption from multiple DC-biased PAs
- Large size
- Efficiency degrading losses in the output MN switch



(a) Circuit Diagram

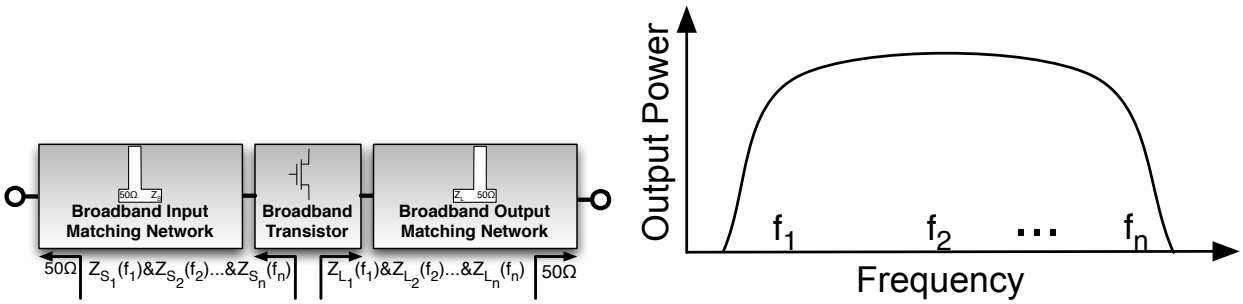
(b) Spectrum Performance

Figure 5.1: Multi-band Switched Power Amplifier Design

Broadband Power Amplifiers

This design consists of a wide-band PA operating over the range of two frequency bands [66]. While this design allows for concurrent operation and uses only a single PA, it also creates the following setbacks:

- Out-of-Band power amplification and spurious emissions
- Poor performance away from the center frequency



(a) Circuit Diagram (b) Spectrum Performance

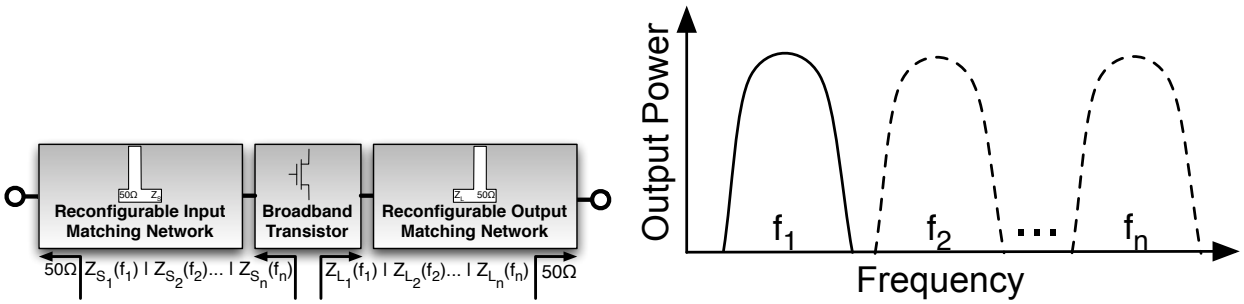
Figure 5.2: Broadband Power Amplifier Design

Reconfigurable Power Amplifiers

In this design, a single PA is used and the input and output MNs are designed so that they can be dynamically tuned to provide matching at a different frequency band [67][68][69]. This design also has several drawbacks:

- Non-concurrent frequency operation
- Complex design of Microelectromechanical systems (MEMS) devices
- Power limitation of MEMS devices
- Repeatability of MEMS devices

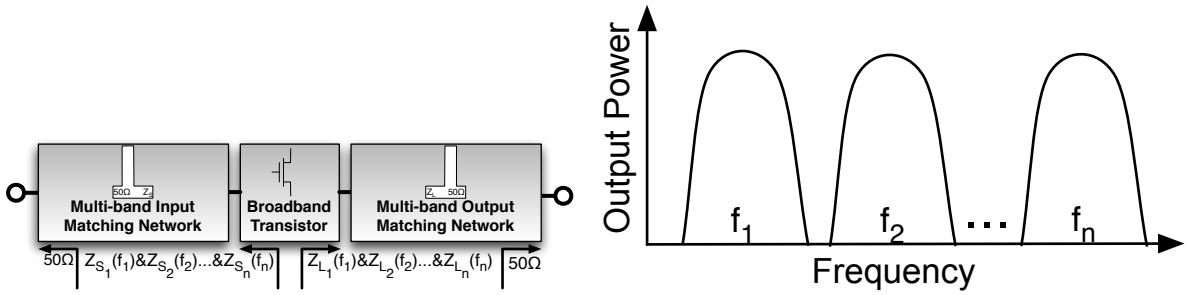
While multi-band PAs based on multiple PAs or reconfigurable MNs do not provide simultaneous multi-band coverage, the broadband solution is more likely to suffer matching degradation at the outer edges of the useful bandwidth. As a result, there is a significant demand for a single power amplifier that will operate with optimal matching at multiple frequencies concurrently as shown in Fig. 5.4. This design uses a single static passive matching network to simultaneously provide the optimal matching impedance transformation at two unrelated frequency bands. The moderate complexity of this design is justified because it requires a small bill of materials and it dramatically reduces the cost by using



(a) Circuit Diagram (b) Spectrum Performance

Figure 5.3: Reconfigurable Power Amplifier Design

only one transistor. The goal of this research was to develop an automated procedure to design a concurrent dual-band PA.



(a) Circuit Diagram (b) Spectrum Performance

Figure 5.4: Concurrent Multiband Power Amplifier Design

5.1.2 Multi-Band Matching Network Design Techniques

In the literature, there are few examples of concurrent multi-band amplifiers. The author in [70] created a dual-band Class-F PA design at f_1 and f_2 by augmenting a Class-F PA at f_1 with additional stubs that are connected by using narrowband LC resonators at f_2, as shown in Fig. 5.5. Since a Class-F PA provides second and third harmonic tuning, the (self-resonant frequency) SRF of the lumped resonators must be at least 6.5 GHz,

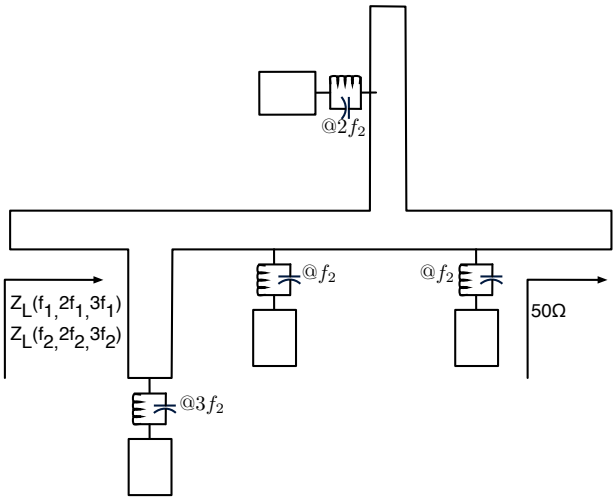


Figure 5.5: Dual-Band Power Amplifier with Resonator Stub Extensions

which is difficult to achieve. Due to the SRF of the LC resonator, this circuit provides significantly worse performance at the lower-band than at the higher-band and achieves a drain efficiency of 50% at 2.14GHz and 30% at 1.7GHz. Clearly this solution does not achieve the same performance as an equivalent single-band PA at each frequency.

Another publication describes an "impedance buffer" solution that can separate the MN into sequential blocks where each block determines the matching impedance of a single frequency and is isolated from the MN of the other frequencies with an impedance buffer as shown in Fig. 5.6 [71]. In this sense, a dual-band solution could be easily extended to a multi-band solution by simply cascading additional matching blocks. An impedance buffer consists of either a short or open-circuit, and this implies that each matching block must provide a transformation from a short/open-circuit to the proper matching impedance using only reactive components. Hence, this solution is only relevant when synthesizing a reactive full-reflection matching impedance at harmonic frequencies. Unfortunately, this solution is not applicable to the design of multi-band MNs at the fundamental frequency where a 50Ω impedance must be transformed to a complex impedance at the transistor port.

Another paper utilizes metamaterials to design a dual-band Class-E PA[72]. While it is

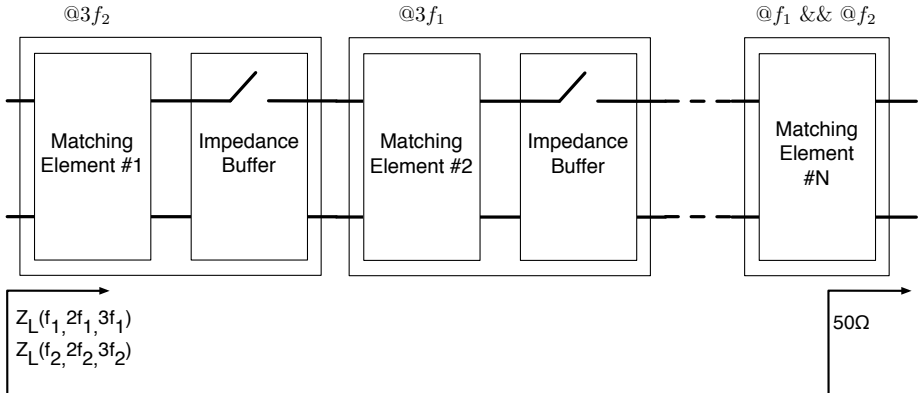


Figure 5.6: Dual-Band Power Amplifier with Impedance Buffer Technique

reasonable that a repetitive structure could provide an equivalent dual-band behaviour, it is unclear whether metamaterials are a mature technology that can be implemented inside high-power MN and further development in this field is required. For example, the paper presented in [72] achieves an efficiency of approximately 50% in a low-power transistor.

While there has been limited work on dual-band amplifiers, there has been extensive work on dual-band filters, power dividers and antennas. Several publications have provided different approaches for implementing a multi-band $\lambda/4$ transmission line[73][74][75]. Examples of dual-band transmission-lines have been implemented in dual-band power dividers and dual-band filters[76][77][78][75]. Since a PA MN is a trans-impedance filter, several publications on multi-band filters and other passive components are of considerable interest. The work presented in [79] demonstrates how a trans-impedance filter can be realized using a homogeneous multi-section $\lambda/4$ transmission line as shown in Fig. 5.7. In this paper, the Chebychev function ($\sec(\theta_m)\cos(\theta)$) was replaced with a function ($acos(\theta)^2 + b$, where a and b are constants) that inherently exhibits dual-band behaviour. Compared to a bandpass chebychev filter, the dual-band chebychev filter splits the pole locations between the two radio bands as demonstrated in Fig. 5.7. This filter provides an impedance transformation from R_L to 50Ω and since the impedance stepping is homogeneous, it guarantees that the variation of the characteristic impedance of each section will be contained to $R_L < Z_{0i} < 50\Omega$. Calculating the characteristic impedance of each line is a complicated process and the values can be computed approximately using the theory of small-signal

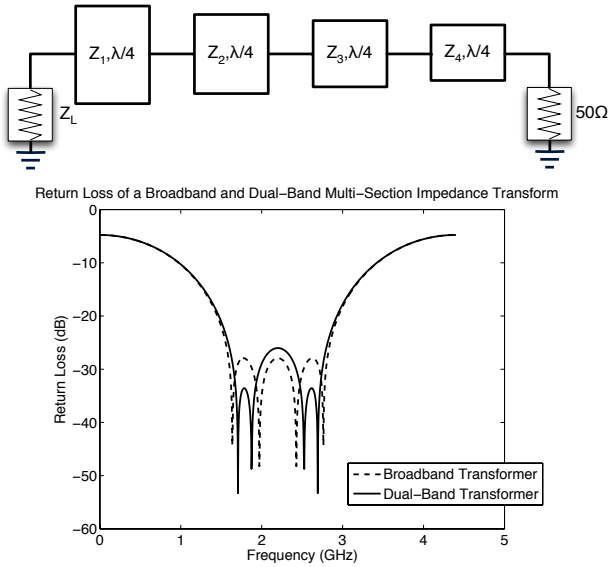


Figure 5.7: Multi-Section Impedance Transformer

reflections[24]. Overall, this procedure provides a good solution for impedance transformations from purely real impedances to 50Ω, however a second-order (4-pole) impedance transformer requires four $\lambda/4$ sections and may result in a very large matching circuit. In contrast the dual-band filter described in [80][75] uses $\lambda/4$ open and short-circuit resonator stubs. Since the dual-band stubs resonate at both frequencies, an n^{th} order filter will achieve n-poles in each radio-band, indicating that this circuit will be half of the size of the multi-section $\lambda/4$ transformer.

Most multi-band circuits depend on two fundamental building blocks:

- Multi-band Impedance/Admittance Inverters (K/J-Inverters)
- Multi-band Impedance Resonators

The common solutions found in the literature are discussed in the following sections.

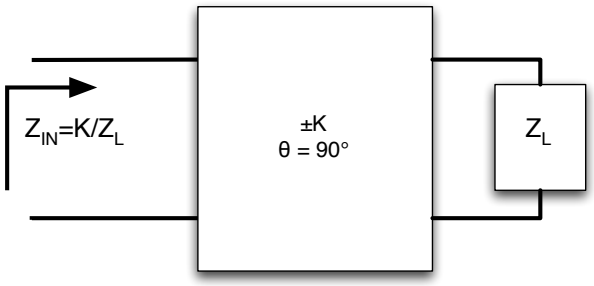


Figure 5.8: Function of an Impedance Inverter

5.1.3 Multi-Band Impedance Inverter Design Techniques

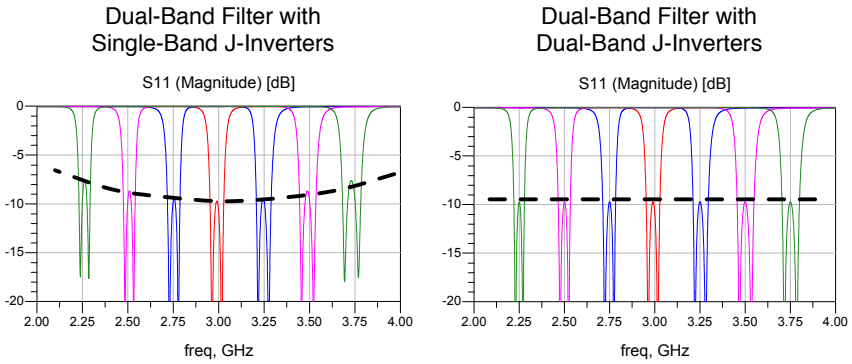
The function of the impedance inverter is demonstrated in Fig. 5.8. The impedance looking into the transformer is proportional to the inverse of the load impedance that is seen at the other end of the transformer.

The two-port network impedance inverter can be characterized by using ABCD parameters as shown in (5.1).

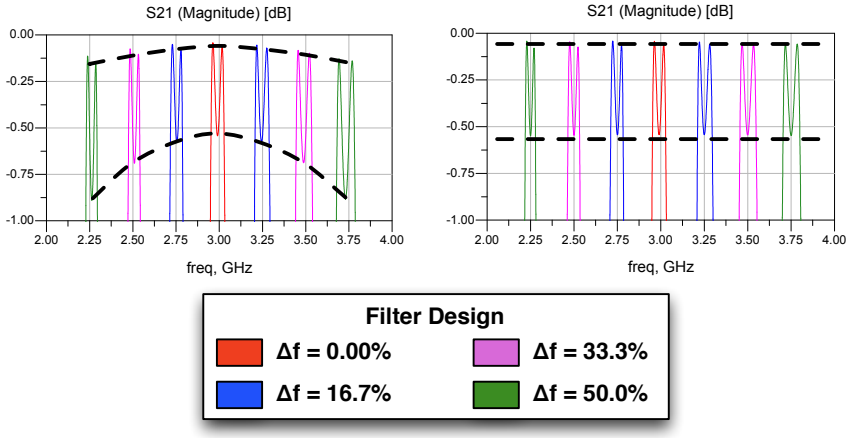
$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} 0 & \pm jK \\ \pm\frac{j}{K} & 0 \end{bmatrix} \tag{5.1}$$

According to the theory presented above, an impedance inverter appears to be frequency independent, however any physical realization of an impedance inverter relies on reactive components (lumped or distributive) that are frequency dependant. Furthermore, most distributive implementations of an impedance inverter are largely based around a $\lambda/4$ transmission line, and have a narrowband frequency range where the theoretical impedance inverter transformation is valid. The importance of realizing this transformation at two frequencies is demonstrated in Fig. 5.9 where the performance of several filters were constructed using either single-band or dual-band impedance inverters. As the frequency separation of the dual-bands is increased, the performance of the filter with the single-band J-inverters diminishes, however the dual-band J-inverter allows the optimal filter performance to be achieved regardless of the frequency separation.

Further investigation has resulted in three existing solutions that have been used in



(a) Magnitude of S_{11} of Different Dual-band Filters



(b) Magnitude of S_{21} of Different Dual-band Filters

Figure 5.9: Different Dual-Band Filters with Increasing Frequency Separation That Demonstrate the Usefulness of Dual-band Impedance Inverters

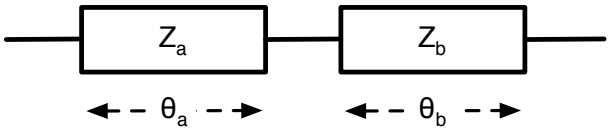


Figure 5.10: Step-model Impedance Inverter

dual-band applications. These circuits are designed to simultaneously satisfy the constraints in (5.1) at two arbitrary frequencies, denoted f_1 and f_2. To examine the design constraints related to the separation of f_1 and f_2, the fractional frequency separation of these frequencies (Δ_f) is defined in (5.3).

$$\Delta_f = \frac{f_2 - f_1}{f_0} \tag{5.2}$$

where,

$$f_0 = \frac{f_2 + f_1}{2} \tag{5.3}$$

Step-Model The technique discussed in [81] demonstrates how two sequential $\lambda/4$ transmission lines with different characteristic impedances can be combined to form a dual-band impedance inverter. The topology of the circuit is presented in Fig. 5.10, where Z_a and Z_b represent the characteristic impedance of each line, while θ_a and θ_b represent the electrical length at the center frequency, f_0. The ABCD parameters are shown in the following equations.

$$A = cos\,(\theta_a)\,cos\,(\theta_b) - \frac{Z_a}{Z_b} sin\,(\theta_a)\,sin\,(\theta_b) - \frac{Z_a}{Z_b} cos\,(\theta_a)\,sin\,(\theta_a)\,tan\,(\theta_b) \tag{5.4}$$

$$B = jZ_a sin\,(\theta_a)\,cos\,(\theta_b) + jZ_b cos\,(\theta_a)\,sin\,(\theta_b) \tag{5.5}$$

$$C = j\frac{1}{Z_a} sin\,(\theta_a)\,cos\,(\theta_b) + j\frac{1}{Zb} cos\,(\theta_a)\,sin\,(\theta_b) \tag{5.6}$$

$$D = cos\,(\theta_a)\,cos\,(\theta_b) - \frac{Z_b}{Z_a} sin\,(\theta_a)\,sin\,(\theta_b) - \frac{Z_a}{Z_b} cos\,(\theta_a)\,sin\,(\theta_a)\,tan\,(\theta_b) \tag{5.7}$$

By equating (5.5-5.7) to (5.1), we can solve for K, Z_a and Z_b as follows.

$$\pm K = Z_a \frac{sin\,(\theta_a)}{cos\,(\theta_b)} \tag{5.8}$$

$$Z_a = \pm K \frac{cos\,(\theta_b)}{sin\,(\theta_a)} \tag{5.9}$$

$$\tag{5.10}$$

$$Z_b = \pm K \frac{sin\,(\theta_b)}{cos\,(\theta_a)} \tag{5.11}$$

For this relationship to hold at both frequencies, $\theta_a = \theta_b = n\frac{\pi}{2}$ at f_0. The final equations for Z_a and Z_b are given in terms of Δ_f where $\theta_a\,(f_1) = \theta_b\,(f_1) = n\frac{\pi}{2}\left(1 - \frac{\Delta_f}{2}\right)$ and $\theta_a\,(f_2) = \theta_b\,(f_2) = n\frac{\pi}{2}\left(1 + \frac{\Delta_f}{2}\right)$ in equations (5.10) and (5.11).

$$Z_a = \frac{\pm K}{tan\left(n\frac{\pi}{2}\left(1 - \frac{\Delta_f}{2}\right)\right)} \tag{5.12}$$

$$\tag{5.13}$$

$$Z_b = \pm K tan\left(n\frac{\pi}{2}\left(1 - \frac{\Delta_f}{2}\right)\right) \tag{5.14}$$

T-Model The technique discussed in [73] demonstrates how two $\lambda/4$ transmission lines that are separated by a single open circuit stub will encompass the performance of an impedance inverter at two separate frequencies. The topology of this circuit is presented in Fig. 5.11, where the characteristic impedances, Z_a and Z_b, denote the characteristic impedance of the transmission lines and stub respectively. The electrical lengths, θ_a and θ_b are given for the center frequency (f_0) and the ABCD parameters of this circuit are given in the following equations.

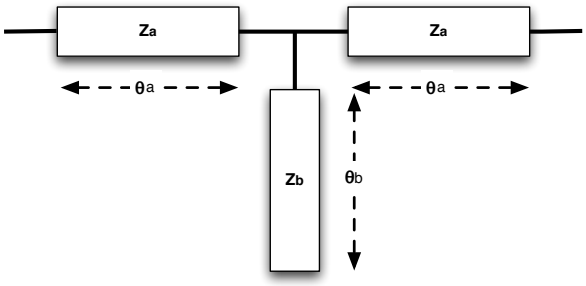


Figure 5.11: T-model Impedance Inverter

$$A = cos^2(\theta_a) - sin^2(\theta_a) - \frac{Z_a}{Z_b} cos(\theta_a) sin(\theta_a) tan(\theta_b) \tag{5.15}$$

$$B = j2Z_a cos(\theta_a) sin(\theta_a) - j\frac{Za^2}{Zb} sin^2(\theta_a) tan(\theta_b) \tag{5.16}$$

$$C = j\frac{2}{Z_a} cos(\theta_a) sin(\theta_a) - j\frac{1}{Zb} cos^2(\theta_a) tan(\theta_b) \tag{5.17}$$

$$D = cos^2(\theta_a) - sin^2(\theta_a) - \frac{Z_a}{Z_b} cos(\theta_a) sin(\theta_a) tan(\theta_b) \tag{5.18}$$

By equating (5.16-5.18) to (5.1) we can solve for K, Z_a and Z_b as follows.

$$\pm K = Z_a tan(\theta_a) \tag{5.19}$$

$$Z_a = \frac{\pm K}{tan(\theta_a)} \tag{5.20}$$

$$Z_b = \frac{Z_a cos(\theta_a) sin(\theta_a) tan(\theta_b)}{cos^2(\theta_a) - sin^2(\theta_a)} \tag{5.21}$$

For this relationship to hold at both frequencies, $\theta_a = n\frac{\pi}{2}$ at f_0 so that $tan(\theta_a(f_1)) = \pm tan(\theta_a(f_2))$. Given this condition, $\theta_b = m\frac{\pi}{2}$ at f_0 so that $A = D = 0$ at both frequencies. The final equations for Z_a and Z_b are given in terms of Δ_f where $\theta_a(f_1) = n\frac{\pi}{2}\left(1 - \frac{\Delta_f}{2}\right)$, $\theta_a(f_2) = n\frac{\pi}{2}\left(1 + \frac{\Delta_f}{2}\right)$, $\theta_b(f_1) = m\frac{\pi}{2}\left(1 - \frac{\Delta_f}{2}\right)$ and $\theta_b(f_2) = m\frac{\pi}{2}\left(1 + \frac{\Delta_f}{2}\right)$ in equations (5.20) and (5.21).

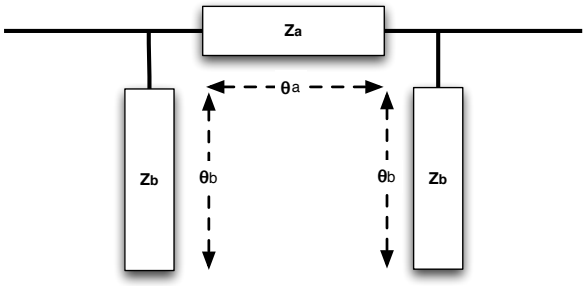


Figure 5.12: Pi-model Impedance Inverter

$$Z_a = \frac{\pm K}{tan\left(n\frac{\pi}{2}\left(1-\frac{\Delta_f}{2}\right)\right)} \tag{5.22}$$

(5.23)

$$Z_b = \frac{Z_a cos\left(n\frac{\pi}{2}\left(1-\frac{\Delta_f}{2}\right)\right) sin\left(n\frac{\pi}{2}\left(1-\frac{\Delta_f}{2}\right)\right) tan\left(m\frac{\pi}{2}\left(1-\frac{\Delta_f}{2}\right)\right)}{cos^2\left(n\frac{\pi}{2}\left(1-\frac{\Delta_f}{2}\right)\right) - sin\left(n\frac{\pi}{2}\left(1-\frac{\Delta_f}{2}\right)\right)} \tag{5.24}$$

Pi-Model The Pi-model is derived in the same way as the T-model, except it is constructed from a $\lambda/4$ transmission line at the center frequency and can be connected with a parallel open circuit stub at each end[75]. The ABCD parameters are given in (5.25)-(5.28).

$$A = cos\left(\theta_a\right) - \frac{Z_a}{Z_b} sin\left(\theta_a\right) cot\left(\theta_b\right) \tag{5.25}$$

$$B = jZ_a sin\left(\theta_a\right) \tag{5.26}$$

$$\begin{aligned} C &= j\frac{1}{Zb} cos\left(\theta_a\right) cot\left(\theta_b\right) + j\frac{1}{Z_a} sin\left(\theta_a\right) + j\frac{Z_a}{Z_b^2} sin\left(\theta_a\right) cot^2\left(\theta_b\right) \\ &+ j\frac{1}{Z_b} cos\left(\theta_a\right) cot\left(\theta_b\right) \end{aligned} \tag{5.27}$$

$$D = cos\left(\theta_a\right) - \frac{Z_a}{Z_b} sin\left(\theta_a\right) cot\left(\theta_b\right) \tag{5.28}$$

By equating (5.25-5.28) to (5.1) we can solve for K, Z_a and Z_b as follows.

$$\pm K = Z_a sin(\theta_a) \tag{5.29}$$

$$Z_a = \frac{\pm K}{sin(\theta_a)} \tag{5.30}$$

$$Z_b = Z_a tan(\theta_a) tan(\theta_b) \tag{5.31}$$

Again, $\theta_a = n\frac{\pi}{2}$ and $\theta_b = m\frac{\pi}{2}$ at f_0 to satisfy the equations for Z_a and Z_b at both frequencies. The final design equations for the Pi-model are shown in (5.30) and (5.31).

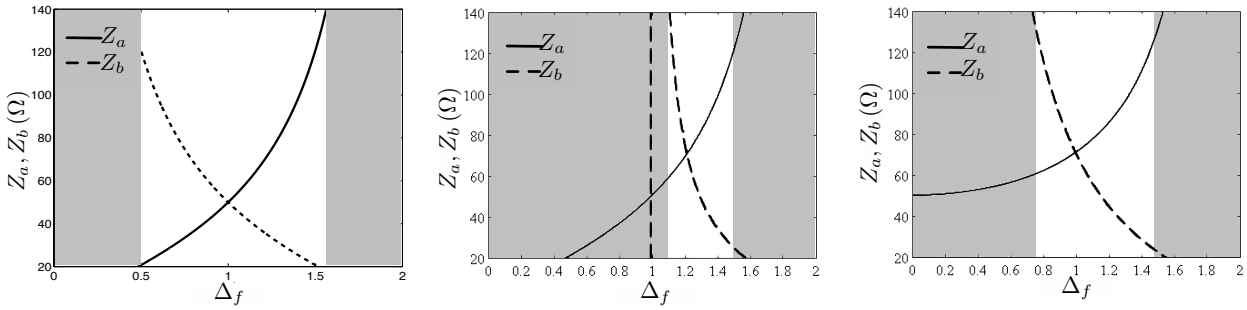
$$Z_a = \frac{\pm K}{sin\left(n\frac{\pi}{2}\left(1-\frac{\Delta_f}{2}\right)\right)} \tag{5.32}$$

$$\tag{5.33}$$

$$Z_b = Z_a tan\left(n\frac{\pi}{2}\left(1-\frac{\Delta_f}{2}\right)\right) tan\left(m\frac{\pi}{2}\left(1-\frac{\Delta_f}{2}\right)\right) \tag{5.34}$$

A comparison of the Step-Model, T-Model, and Pi-Model versus the dual-band frequency separation (Δ_f) is presented in Fig. 5.13. In this comparison, $K = 50$ and the characteristic impedance of Z_a and Z_b are plotted with respect to frequency separation, Δ_f. The results shown in Fig. 5.13 demonstrate that the values of Z_a and Z_b will deviate in a nonlinear fashion with respect to Δ_f. To illustrate the potential limitations associated with each J-inverter, each model has been limited to situations where $25\Omega < Z_0 < 135\Omega$, and the permissiable dual-band frequency separations (Δ_f) are shown in Fig. 5.13 and . Table 5.1. Since the derivation of Z_a and Z_b in all impedance inverters is linearly dependent on K, the analysis could be linearly scaled to any other value of K. For example, a trans-impedance filter ($50\Omega \leftarrow 5\Omega$) can be designed using decreasing values of K between subsequent resonators.

The Step-Model has the largest frequency band separation, while the T-Model has the smallest. This suggests that the Step-Model is the optimal solution, however the Pi-model has the shortest electrical length (least insertion loss). While the Step-Model is the only circuit that can be extended from dual-band to multi-band applications, the Pi-Model is best suited for PA applications because it is important to minimize insertion loss. In



(a) Step-Model (b) T-Model (c) Pi-Model

Figure 5.13: Frequency Band Separation Comparison of Impedance Inverter Circuit Topologies

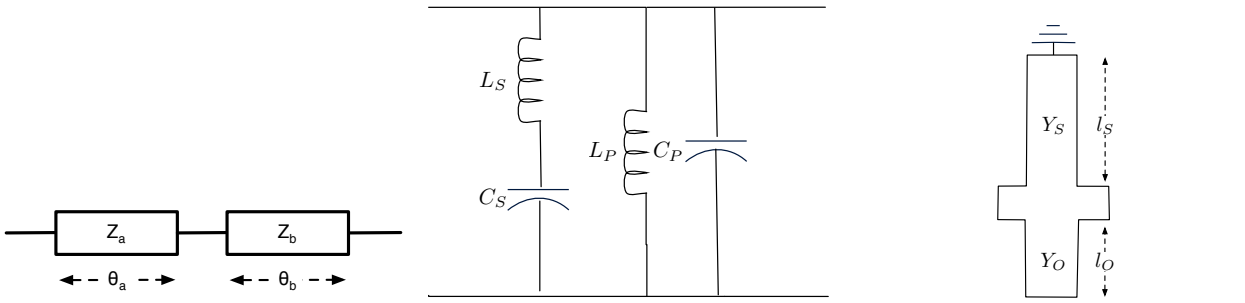
Table 5.1: Comparison of Impedance Inverter Circuit Topologies

Impedance Inverter Type	Frequency Band Separation Range
Step-Model	$0.5 < \Delta_f < 1.55$
T-Model	$1.12 < \Delta_f < 1.51$
Pi-Model	$0.75 < \Delta_f < 1.46$

situations where the Pi-Model can not be realized, modifications to the peripheral shunt connected components can be achieved to extend the frequency range of the Pi-Model

5.1.4 Multi-Band Impedance Resonator Design Techniques

A dual-band resonator can consist of series-connected $\lambda/4$ transformers, such as those used in [79] or it may also be constructed using shunt-connected lumped-circuit or distributive-stub resonators as demonstrated in [73][75], and shown in Fig 5.14. While the $\lambda/4$ series resonators can be extended to multi-band applications, the shunt-connected resonators can be used in a circuit that is half the size. This section will focus on the implementation of the shunt-connected resonators.



(a) Series $\lambda/4$ Resonator (b) Shunt Lumped Resonator (c) Shunt Distributive Resonator

Figure 5.14: Dual-Band Resonator Lumped-to-Distributive Equivalent Circuit Transformation

A plot of the admittance of the dual-band resonator as a function of frequency is presented in Fig. 5.15 and both the lumped-element and an equivalent distributive-element implementation is compared. Both the microstrip and lumped-element circuits resonate at the same frequencies separated by asymptotes, however the slope of the x-intercept is not the same. Theoretically, there are enough degrees of freedom in these circuits to control both the resonant frequency (x-axis intercept) as well as the fractional bandwidth (the slope at the x-axis intercept). Since the filter theory has been derived using lumped components, an equivalent circuit transformation must be made between these circuits

such that the resulting circuit is equivalent at the design frequencies and over the required bandwidth. This implies that the function and first-order derivative of the microstrip circuit must be equivalent to the lumped-element circuit at both frequencies. When using an open circuit and short circuit stub, the circuit equivalence can be represented with the following equations.

$$\omega_1 C_P - \frac{1}{\omega_1 L_P} - \frac{1}{\omega_1 L_S - \frac{1}{\omega_1 C_S}} = Y_O tan\left(\frac{\omega_1 l_O}{v}\right) + Y_S cot\left(\frac{\omega_1 l_S}{v}\right) \quad (5.35)$$

$$\omega_2 C_P - \frac{1}{\omega_2 L_P} - \frac{1}{\omega_2 L_S - \frac{1}{\omega_2 C_S}} = Y_O tan\left(\frac{\omega_2 l_O}{v}\right) + Y_S cot\left(\frac{\omega_2 l_S}{v}\right) \quad (5.36)$$

$$\frac{\partial}{\partial\omega}\left[\omega_1 C_P - \frac{1}{\omega_1 L_P} - \frac{1}{\omega_1 L_S - \frac{1}{\omega_1 C_S}}\right] = \frac{\partial}{\partial\omega}\left[Y_O tan\left(\frac{\omega_1 l_O}{v}\right) + Y_S cot\left(\frac{\omega_1 l_S}{v}\right)\right] \quad (5.37)$$

$$\frac{\partial}{\partial\omega}\left[\frac{\partial}{\partial\omega}\omega_2 C_P - \frac{1}{\omega_2 L_P} - \frac{1}{\omega_2 L_S - \frac{1}{\omega_2 C_S}}\right] = \frac{\partial}{\partial\omega}\left[Y_O tan\left(\frac{\omega_2 l_O}{v}\right) + Y_S cot\left(\frac{\omega_2 l_S}{v}\right)\right] \quad (5.38)$$

In these equations, L_S and C_S represent the series resonator lumped element values, L_P and C_P represent the parallel resonator lumped element values, as shown in Fig. 5.14b. Y_O and Y_S represent the characteristic admittance of the open and short-circuit microstrip stubs, and l_O and l_S represent the length of the open and short circuit stubs, as shown in Fig. 5.14c. It is impossible to analytically solve these equations, therefore a numerical solver is used to calculate the distributive-element parameters.

Alternatively we could derive the entire MN theory directly in terms of the final distributive resonator formulation. This solution involves deriving the microstrip stub dimensions directly from the low-pass filter prototype. Compared to the other solutions, this approach will optimize both the resonator and the adjacent impedance inverters to create the desired admittance slope around both of the design frequencies One of solution results in a $\lambda/4$ open and short-circuit stub[75], while the other manipulates both the width and the length of the stubs [73] to achieve the desired admittance profile.

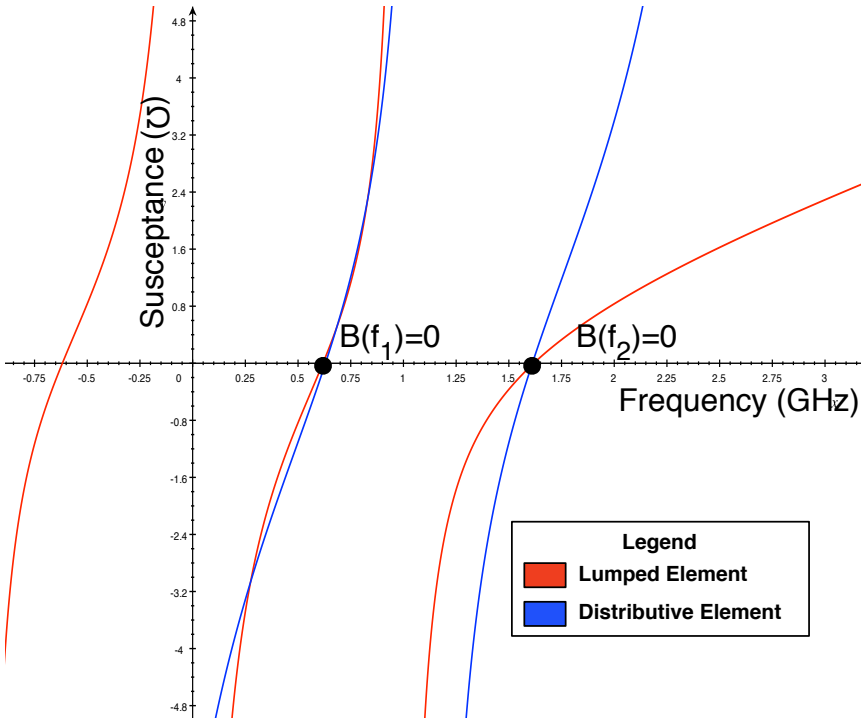


Figure 5.15: Lumped Element and Distributive Element Dual-Band Admittance Resonator

5.1.5 Proposed Concurrent Multi-band Power Amplifier Techniques

Two parallel MN solutions have been investigated for the purpose of designing a concurrent dual-band PA.

1. A concurrent dual-band PA using simple matching network theory
2. A concurrent dual-band PA using filter-based matching network theory

The first idea is much simpler than the second solution. This circuit was designed as an exercise to determine the major issues surrounding multi-band PA design. The filter-based MNs provide analytical control over the fractional bandwidth at each radio band and represent a much more theoretically robust solution to the problem. Each of these solutions will be described in the following sections.

The goal of this project was to develop a sequential procedure, based on basic matching theory, that could be used to design a dual-band MN. As distributive MNs are often difficult to intuitively describe, the foundation of this theory is based on lumped-element MN design. In order to determine the final circuit design, the following equivalent circuit transformations are needed:

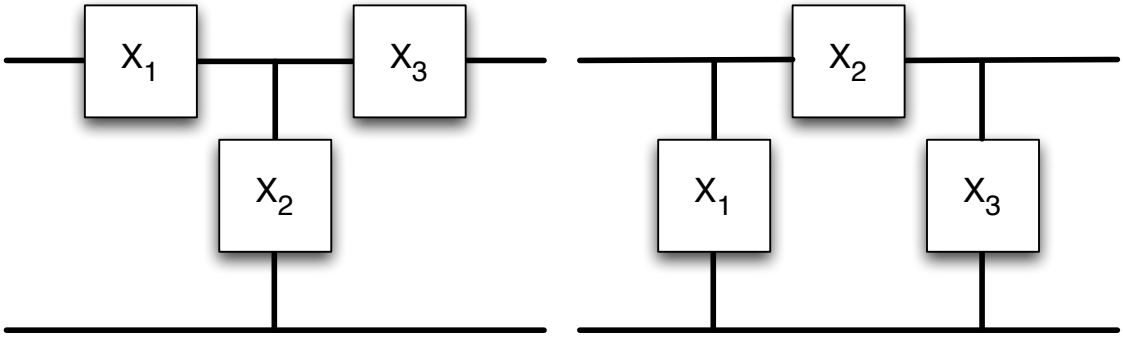
1. Lumped-Element Single-Band MN $\rightleftharpoons$ Lumped-Element Dual-Band MN
2. Lumped-Element Dual-Band MN $\rightleftharpoons$ Distributive-Element Dual-Band MN

5.2 Concurrent Dual-band Power Amplifier Based on Matching Network Theory

Dual-Band Lumped-Element Transformation

As a starting point, a single-band MN is designed at each frequency. While a two component MN satisfies the minimum degrees of freedom, three components are needed to maintain empirical control over the matching bandwidth (the Q-factor), while providing additional flexibility. A Pi-model or T-model MN, as shown in Fig 5.16, can be chosen to provide the required impedance transformation. In general, the Pi-model provides a more suitable impedance transformation from high impedances, while the T-model provides an optimal solution for lower impedances. The Pi-model or T-model MNs designed at each frequency can consist of any possible combination of inductors or capacitors, however both models must be either a Pi-model or a T-model for the single-band to dual-band circuit transformation to work. Therefore, either of the circuits shown in Fig. 5.16 can be used as a starting point for the dual-band MN design.

Up to this point the MNs at frequency, f_1 and f_2, are designed independently to present the load impedance,Z_{L1} or Z_{L2}, to the port of the transistor. The dual-band circuit transformation is completed by replacing each component with an LC resonator that has the equivalent impedance at the design frequencies. There are two types of resonator transformations that can take place:



(a) T-Model (b) Pi-Model

Figure 5.16: Single-Band Matching Network Circuit Diagram

- If $X_n(f_2) > X_n(f_1) \Longrightarrow$ A series resonator must be used
- If $B_n(f_2) > B_n(f_1)(X_n(f_1) > X_n(f_2)) \Longrightarrow$ A parallel resonator must be used

The conditions described above can be explained by analyzing the behaviour of a series and parallel resonator as shown in Fig. 5.17. Assuming that $f_2 > f_1$ and that $f_1, f_2 > 0$ we can see that the reactance (X) of a series resonator is a monotonically increasing function, while the susceptance (B) of a parallel resonator is also a monotonically increasing function above $0Hz$. Therefore, the function of the resonator must be defined such that it intersects the impedance used in the single-band MN at f_1 and f_2 as described in (5.39-5.42). It is important to note that this version of the dual-band resonator differs from the dual-band resonator described in Section 5.1.4 because it was not derived from filter theory and it does not have sufficient degrees of freedom to define the fractional bandwidth.

Single-band to Dual-band Series Resonator Transformation:

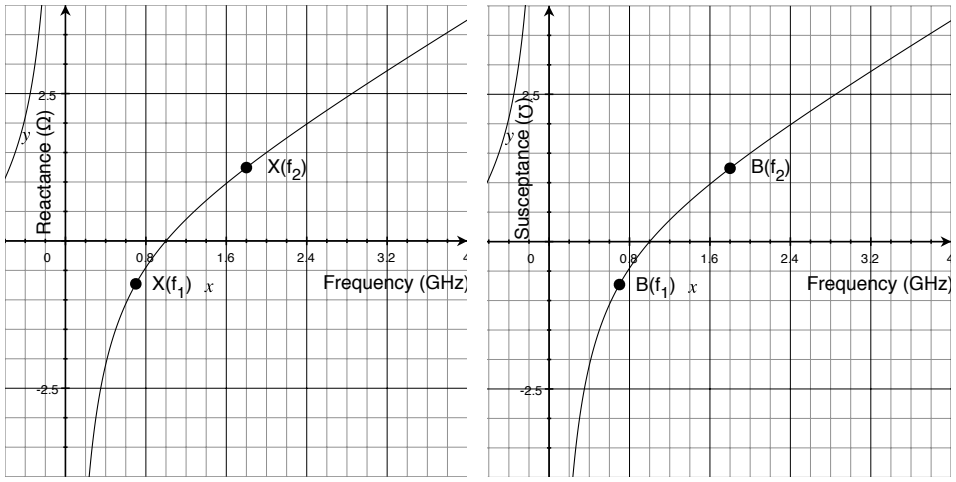
$$X_n(\omega_1) = \omega_1 L_n - \frac{1}{\omega_1 C_n} \tag{5.39}$$

$$X_n(\omega_2) = \omega_2 L_n - \frac{1}{\omega_2 C_n} \tag{5.40}$$

Single-band to Dual-band Parallel Resonator Transformation:

$$B_n(\omega_1) = \omega_1 C_n - \frac{1}{\omega_1 L_n} \tag{5.41}$$

$$B_n(\omega_2) = \omega_2 C_n - \frac{1}{\omega_2 L_n} \tag{5.42}$$



(a) Series Resonator (b) Parallel Resonator

Figure 5.17: Dual-Band Resonators Behavior vs. Frequency

An example of a dual-band matching circuit based on a three-element T-model is shown in Fig. 5.18. In this figure, the first and third resonators are in series, while the second resonator is based on a parallel transformation.

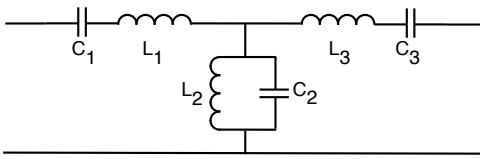


Figure 5.18: Lumped Element Dual-Band Matching Network

In microstrip MN theory, it is difficult to realize series connected stubs. As a result, the first and third resonators in Fig. 5.18 cannot be directly transformed into an equivalent distributive microstrip circuit. Instead, they must be converted into equivalent shunt resonators through the use of impedance/admittance inverters discussed in Section 5.1.3, as shown in Fig. 5.19.

A lumped-element equivalent of the dual-band J-inverter based on the Pi-model was implemented to get the exact circuit behavior described in (5.1) at frequencies, f_1 and f_2,

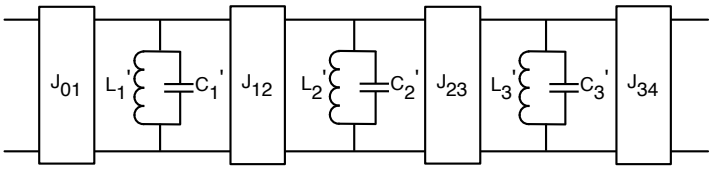


Figure 5.19: Lumped Element Dual-Band Matching Network with Impedance Inverters

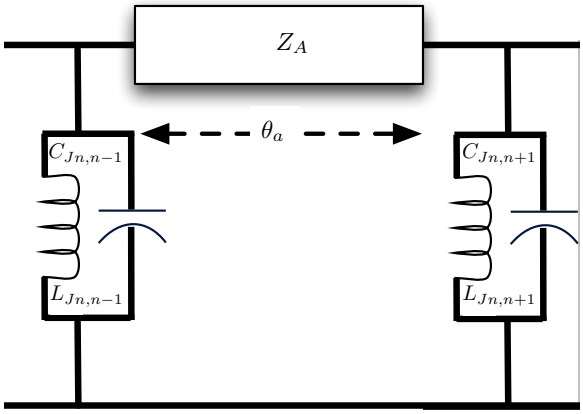


Figure 5.20: Lumped Element Pi-Model Impedance Inverter

and is shown in Fig. 5.20. The resulting matching network is represented in Fig. 5.21 and the component values are calculated as follows:

$$L''_{n,n+1} = \frac{1}{\frac{1}{L'_n} + \frac{1}{L_{Jn,n-1}} + \frac{1}{L_{Jn,n-1}}} \tag{5.43}$$

$$C''_{n,n+1} = C'_n + C_{Jn,n-1} + C_{Jn,n+1} \tag{5.44}$$

Given that the resulting circuit contains four $\lambda/4$ transmission lines, the total length of the MN is λ, which is approximately 2-4 times larger than traditional single-band MNs. This comparatively large size of the dual-band MN, results in larger conduction loss due to the tangential loss ratio ($tan\delta$) of the microstip substrate and this is detrimental to the PAE of the PA. The technique proposed in [82] was used to replace the transmission line of impedance Z_J, with three equivalent transmission lines of characteristic impedance Zm,

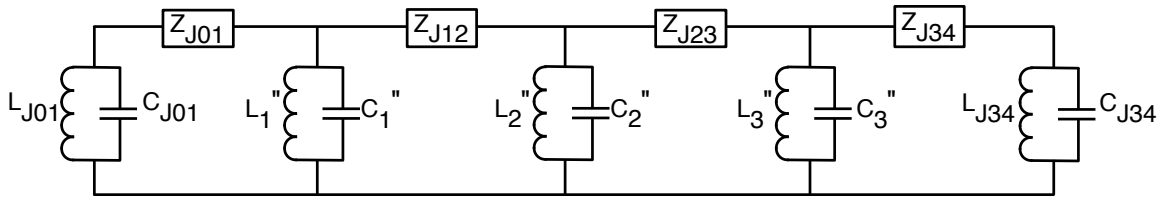


Figure 5.21: Lumped Element Dual-Band Matching Network with Dual-band Mircostrip Impedance Inverters

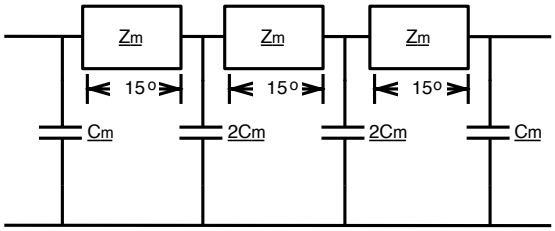


Figure 5.22: Compact Impedance Inverter

separated by shunt capacitance's Cm, as shown in Fig 5.22. The choice of three sections was made to maximize compactness while still meeting the frequency requirements of the design. Compacting the J-inverters yielded significant reduction of the insertion loss and the size of the dual-band MNs was reduced by more than 50%. The resulting circuit will also have modified resonators that are calculated using the following equations:

$$L_n''' = L_n'' \tag{5.45}$$

$$C_n''' = C_n'' + C_{M_{n-1,n}} + C_{M_{n,n+1}} \tag{5.46}$$

Dual-Band Distributive-Element Transformation

While the distributive portion of the dual-band impedance inverters have been implemented using microstrip transmission line, the other components have been combined into the adjacent resonators. In the final equivalent circuit transformation, these parallel lumped-element resonators must be converted into distributive-element resonators.

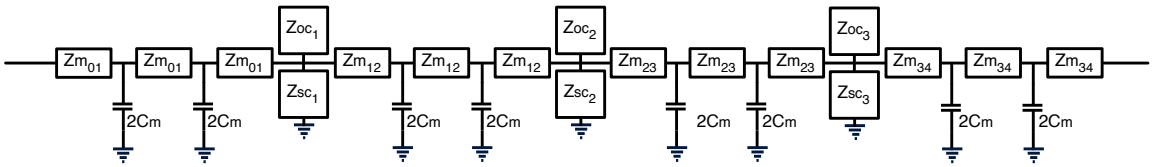


Figure 5.23: Dual-Band Distributive Matching Network with Compact Impedance Inverters

This transformation is completed using the following equations and the resulting MN is described in Fig. 5.23.

$$\frac{tan\,(\theta_{OC})}{Z_{OC}} - \frac{cot\,(\theta_{SC})}{Z_{SC}} = \omega_1 C''' - \frac{1}{\omega_1 L'''} \tag{5.47}$$

$$\frac{tan\,(\theta_{OC})}{Z_{OC}} - \frac{cot\,(\theta_{SC})}{Z_{SC}} = \omega_2 C''' - \frac{1}{\omega_2 L'''} \tag{5.48}$$

Results

The methodology defined above was used to design a dual-band 10W PA using the Cree CGH40010F transistor at 2.5GHz and 3.5GHz [83]. The input and output MN were designed to provide an impedance transformation from $Z_S(f_1), Z_S(f_2) \rightarrow 50\Omega$ and $Z_L(f_1)$, $Z_L(f_2) \rightarrow 50\Omega$ respectively, where the optimal impedances were determined using a simulated source/load-pull measurement. The resulting circuit was simulated using the Harmonic Balance Simulator in ADS and was compared with equivalent single-band PA designs at 2.5GHz and 3.5GHz as shown in Table 5.2. While there is a small trade-off in the PAE (around 5%) and output power (about 0.4 dB), this can be attributed to the fact that the dual-band MNs are electrically longer than those utilized in the single-band applications. Therefore, these small drops in performance are acceptable when considering that the PA works at both operating frequencies concurrently.

Bandwidth Problem Although the simulation results at 2.5 GHz and 3.5 GHz were satisfactory, it was difficult to fabricate and measure this circuit because it was inher-

Table 5.2: Simulated Performance Comparison Between the Dual-Band PA and equivalent Single-Band PAs

	Single-Band		*Dual-Band*	
	2.5 GHz	3.5 GHz	2.5 GHz	3.5 GHz
P.A.E (%)	68.7	55.8	64.0	51.7
Pout (dBm)	41.15	41.03	40.73	40.64

ently narrowband. Since the procedure was only defined at the operating frequencies, the fractional bandwidths around f_1 and f_2 were completely uncontrolled.

Negative Impedance Problem The dual-band impedance inverters are constructed using inductors and capacitors that have unrealizable negative values. This produces a problem when the adjacent resonator cannot absorb these values or when the dual-band impedance inverter is placed on the outer edges of the impedance MN. To eliminate this problem, several solutions are presented:

- Using an even order T-model MN or an odd order Pi-model will ensure that there is no impedance inverter next to the 50Ω port.
- The negative component values could also be combined with the optimal matching impedance (Z_L) to form a new optimal impedance (Z'_L) as the new design goal. This would require an iterative calculation procedure that would be implemented in programmable manor.

Impractical Impedance Values Problem Since there are several equivalent circuit transformations that are taking place, it is very difficult to make good decisions at an early stage that will ensure that practical microstrip transmission line and stub dimensions will be calculated at the end of the process. The characteristic impedance values of the microstrip elements plays a big role on the choice of substrate and a wide variation in the characteristic impedance cannot be realized using a single substrate.

The general consensus from this analysis is that the behaviour of the circuit is significantly dictated by design decisions that are made throughout the entire design process. While a sequential solution produces a result that is guaranteed to provide the correct matching at the design frequencies, this solution is not unique and there may be a more optimal solutions for maximizing bandwidth and minimizing the insertion loss. To improve on this design, additional design constraints must be included in the definition of the problem to ensure that the final design meets specific bandwidth requirements.

5.3 Concurrent Dual-band Power Amplifier Based on Filter Theory

Based on the experiences of the previous section, a dual-band MN and PA based on dual-band trans-impedance filter theory, was developed by my colleague, Xin Fu. The *bandwidth problem* was circumvented by using filter theory to define the FBW as a design constraint and the *negative impedance problem* was eliminated by using an odd-order Pi-model for the dual-band J-Inverters The *impractical impedance problem* was eliminated harmonically separating the frequency bands ($f_1 = 800MHz$ and $f_2 = 1900MHz$), however exact harmonic separation was avoided so that harmonic terminations of both bands could be optimized for Class J operation. Overall, all problems were mitigated by modifying the dual-band filter technique in [75] to provide a direct synthesis from the Low-Pass Prototype filter to the dual-band distributive element circuit. The resulting dual-band MN schematic is shown in Fig. 5.24, and the synthesis equations are summarized in (5.49-5.52).

$$Z_A = \frac{1}{4f_0C_n} sec^2\left(\frac{pi}{2}\Delta_f\right) \tag{5.49}$$

$$Z_B = \frac{1}{4f_0C_n} csc^2\left(\frac{pi}{2}\Delta_f\right) \tag{5.50}$$

$$Z_C = \frac{1}{J_{n-1,n} cos\left(\frac{pi}{2}\Delta_f\right)} \tag{5.51}$$

$$Z_D = \frac{1}{J_{n-1,n} sin\left(\frac{pi}{2}\Delta_f\right) tan\left(\frac{pi}{2}\Delta_f\right)} \tag{5.52}$$

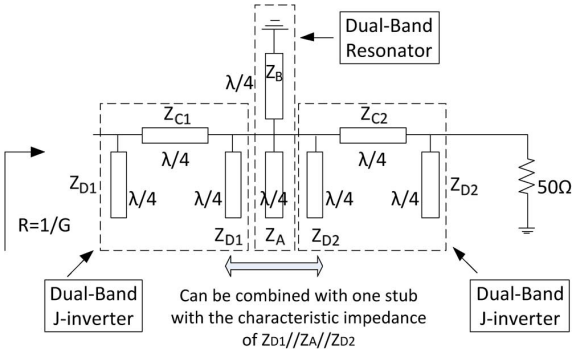


Figure 5.24: First Order Dual-Band Trans-Impedance Filter[3]

Dual-band Low-Impdendance Matching Network

The input/output impedance of FET transistor can be approximated as shunt RC circuit, however a packaged transistor also contains bond wires, transistor leads, and parasitic elements that distort this ideal equivalent circuit. Therefore, in widely separated dual-band applications the ideal matching impedance will sometimes be inductive and will sometimes be capacitive. In order to utilize the trans-impedance filter, we devise a low-impedance matching network that translates Z_{L1} and Z_{L2} to a constant conductance, G. To account for all permutations of inductive and capacitive loads, both a shunt open-circuit stub and shunt short-circuit stub of variable widths and lengths are used to transform the complex loads to a constant resistance circle (below the real-axis), such that a transmission line will translate the impedance to G at both frequencies. As this circuit is difficult to define theoretically, the dimensions of these three elements are solved numerically using an optimization goal that minimizes the Q-factor of the MN.

Dual-band Harmonic Matching Considerations

Another important of consideration of a dual-band PA is the impedance termination of all harmonics of f_1 and f_2. Since explicitly tuning the harmonic impedances for each band would result in a large MN consisting of many stubs, the work of Xin Fu utilized the a Class-J design to diminish the sensitivity to the harmonic matching impedance. This

created a condition where the second harmonic impedance could be sufficiently tuned by controlling the electrical length of the bias network. This is done by adding a transmission line ($Y_0 = G$) between low-impedance MN and the dual-band trans-impedance filter. The final MN design in Fig. 5.25 shows how the dual-band low impedance MN, the bias network (with harmonic offset transmission line), and the dual-band trans-impedance filter can be designed as three separate modules and integrated into a combined MN.

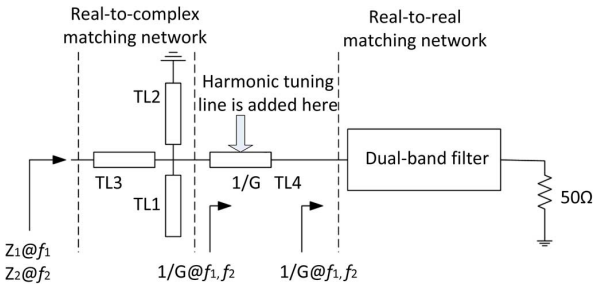


Figure 5.25: Dual-Band Filter-Based Matching Network[3]

A concurrent dual-band filter-based Class-J 45W PA was designed at $0.8GHz$ and $1.9GHz$ and is shown in Fig. 5.26. The maximum performance in the lower-band and upper-band achieved a drain efficiency and output power of at-least 68% and $45dBm$ respectively, as shown in Fig. 5.27.

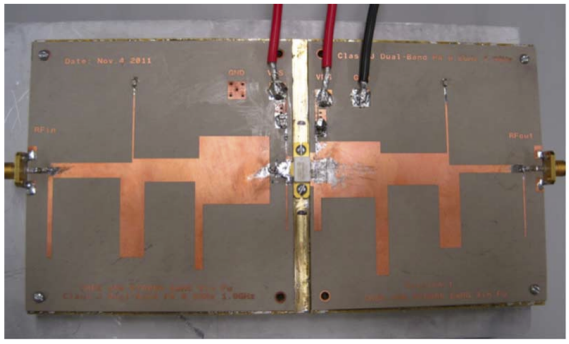
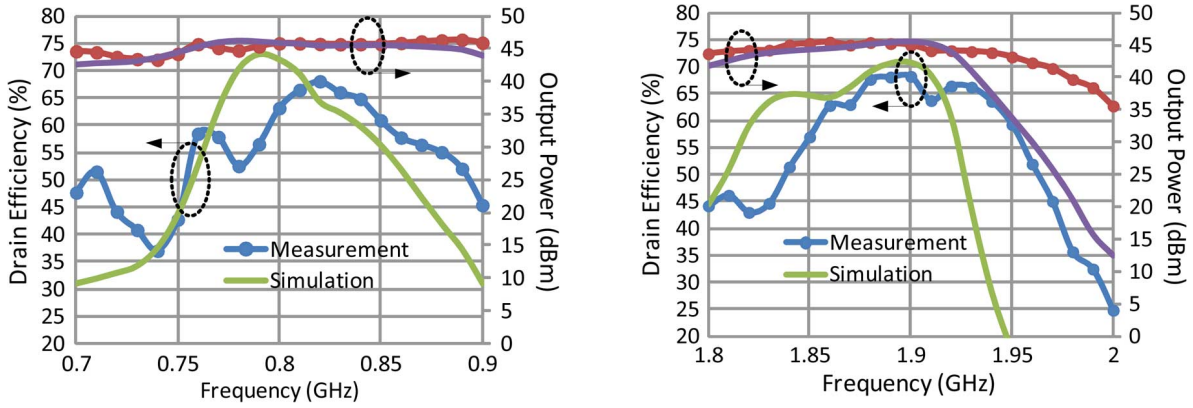


Figure 5.26: Dual-Band Filter-Based Power Amplifier[3]



(a) Lower-Band Performance

(b) Higher-Band Performance

Figure 5.27: Measurement and Simulation Results of a Concurrent Dual-Band Filter-Based Class-J 45W Power Amplifier Performance[3]

5.4 Optimizing Concurrent Operation using Nonlinear Characterization and Modeling Techniques

Chapter 4 demonstrates that small changes in the phase relationship between MIMO signals in a nonlinear system can alter the time-domain load-line operation and cause the transistor to enter cut-off or triode operation, as shown in Fig. 4.4. Although the PAs designed Section 5.2 and Section 2.1.1 are capable of operating simultaneously at two-frequencies, they are not optimized for concurrent operation, therefore controlling the relative phase relationship between the signal at f_1 and f_2 is crucial to synthesizing the most optimal time-domain waveform behaviour. Furthermore, the output of a dual-band PA produces rich frequency spectrum that is not considered in PA design theory. Most specifically, the efficiency enhancement techniques, predominant in modern PA designs, are derived from load-line theory or Waveform Engineering that assumes single-band operation. The study of concurrent dual-band operation demands Waveform Engineering design techniques that can be implemented using the Characterization system in Chapter 3. Additionally, compact-circuit models, typically optimized for single-band operation, can

be supplemented with the behavioural modelling algorithm proposed in Chapter 4, that will model DUT under concurrent dual-band operation.

The Characterization system in Chapter 3 can synthesize the phase coherent MIMO signals needed to measure and validate a concurrent dual-band PA, however the system only supports harmonicly related signals. Hence, further development of this system is needed because high-efficiency dual-band PAs cannot be designed at harmonic frequency band separations. The behavioural modelling algorithm in Chapter 4 can also be extended to support non-commensurate dual-band PA applications.

Hopefully, this design example has illustrated the complexity behind modern nonlinear circuit design. The quest to design high-efficiency, high-PAPR, frequency-agile PAs produces complex analog circuit designs that are nearly impossible to optimize after fabrication. Hence, modern PA deign techniques are only as good as the nonlinear characterization and modelling techniques that enable electronic design automation. Although this Chapter studied dual-band PA design, it is a single example of complex analog circuits that are driving the need for improvements to nonlinear characterization and modeling techniques.

Chapter 6

Conclusions

6.1 Progress

Increasing PA linearity and power-efficiency can only be achieved by operating transistors under stronger nonlinear conditions. Although linearity and power-efficiency traditionally represent opposing design requirements, advancements in PA design theory provide solutions that achieve both. However, designing PAs to attain an optimum trade-off between efficiency and linearity over extended RF bandwidth can no longer be achieved empirically, hence CAD tools are crucial to the successful deployment of modern PAs. To successfully emulate the PA electro-thermal behaviour, CAD tools utilize transistor models whose accuracy is limited by three factors:

1. The formulation of the model.
2. The model extraction procedure.
3. The accuracy of the measurement data.

This thesis improved the accuracy of measurement data by constructing a modular, sequentially calibrated Nonlinear Characterization System (NCS) that replicates the large-signal operating conditions of the DUT under CW or pulsed stimulus. The NCS is designed

to measure high-power, unmatched, broadband nonlinear transistors and is designed to seamlessly integrate measurement and simulation data. As such, advanced characterization or modelling algorithms can be equivalently synthesized in an ideal Harmonic Balance simulation without performing calibration or incurring measurement system downtime. Finally, the NCS provides post-measurement analysis integrated into the MATLAB environment that can perform traditional measurement techniques, such as DC I/V Characterization plots or Load-Pull contours. It can also calculate traditional figures-of-merit, such as P_L, or PAE, using post-measurement equations.

The NCS was used to study how the accuracy of the existing PHD model could be improved by optimizing the characterization system for behavioural model extraction. The resulting model extraction improved the NMSE output power prediction of the PHD model by $5dB$. By studying the formulation of various behavioural models, the MHV model was identified as a suitable candidate for strongly nonlinear systems which also could be scaled to approximate systems with many inputs. Consequently, a practical extraction of the MHV model was created to mimic the behaviour of strongly nonlinear unmatched transistors, which produced an *additional* $5dB$ and $2dB$ NMSE improvement in DC drain current and fundamental frequency output power. A hybrid solution combined the DC and fundamental accuracy of the MHV model with the harmonic PHD models, to improve accuracy where it counts, while reducing the number of measurements needed to extract the model. The MHV Hybrid model improved the prediction of vector power-gain (G_P) by $3.4dB$, thereby significantly improving the emulation of linearization techniques within a CAD simulation.

A design example of a concurrent dual-band PA was used to prove the merits of the NCS. This studied the design of a concurrent dual-band PA using matching network theory, and more sophisticated filter theory. Although a dual-band PA can be synthesized using matching network theory, it is very difficult to design for a specific RF bandwidth. Alternatively, the dual-band PA synthesized using filter-theory maintains the desired performance over a sequential design process that uses equivalent circuit transformations. The resulting filter-based dual-band PA is designed at $0.8GHz$ and $1.9GHz$ achieves a maximum performance in each band of 68% drain efficiency and $45dBm$ output power in both radio bands during non-concurrent single-band operation. As the matching networks are

designed to operate simultaneously at both radio bands, post-fabrication optimization is extremely difficult, hence CAD driven first-pass simulation accuracy is a requirement for this design application. The characterization system in Chapter 3 was developed to study multi-band PA devices, however it is currently limited to commensurate frequencies, hence further development of this platform is needed to study the concurrent operation of the dual-band PA. Nonetheless, this design example exemplifies the complexity associated with PA designs that target multiple objectives to maximize power-efficiency and linearity over a specific RF bandwidth. It demonstrates the interoperability between nonlinear characterization, nonlinear modelling, and nonlinear circuit design that is needed to meet the performance requirements of next generation wireless technology.

6.2 Publications and Invited Presentations

Refereed Journal Publications

1. **D. T. Bespalko**, and S. Boumaiza, "Towards Higher-Order Multi-Harmonic Behavioural Models that Enable High-Efficiency Power Amplifier Design," *IEEE Transactions on Microwave Theory and Techniques*, [Submitted Aug. 2015].

2. X. Fu, **D. T. Bespalko**, and S. Boumaiza, "Novel Dual-Band Matching Network for Effective Design of Concurrent Dual-Band Power Amplifiers, *IEEE Transactions on Circuits and Systems I*, vol. 61, no. 1, pp. 293-301, Jan. 2014.

3. L. Betts, **D. T. Bespalko** and S. Boumaiza, "Application of Agilent's PNA-X NVNA and X-Parameters in Power Amplifier Design," *Microwave Journal*, pp.1–16, May 2011.

4. **D. T. Bespalko** and S. Boumaiza, "Concurrent Dual-Band GaN Power Amplifier with Compact Microstrip Matching Network," *Microwave and Optical Technology Letters*, vol. 51, no. 6, pp.1604–1607, Jun. 2009.

Refereed Conference Publications

1. **D. T. Bespalko**, A. Amini, and S. Boumaiza, "A High-Order Model Looking Beyond the First-Order Harmonic Superposition Assumption," *IEEE Radio Wireless Week*, Jan. 2016.

2. **D. T. Bespalko**, and S. Boumaiza, "Enhanced PHD Model Extraction by Improving Harmonic Response Superposition During Extraction," *86th ARFTG Conference*, Atlanta, GA, USA, Dec. 2015.

3. A. Amini, **D. Bespalko**, and S. Boumaiza,"On the Poly-Harmonic Distortion Modeling Framework and its Practical Realization," *IEEE Topical Symposium on Power Amplifiers for Wireless Communications*, Sep. 2012.

4. X. Fu, **D. T. Bespalko**, and S. Boumaiza, "Novel dual-band matching network topology and its application for the design of dual-band Class J power amplifiers," *IEEE MTT-S International*, Jun. 2012.

5. **D. Bespalko**, and S. Boumaiza, "Nonlinear Device Characterization and Modeling for High Power Amplifier Design," *12th annual IEEE Wireless and Microwave Technology (WAMI) Conference*, Apr. 2011.

6. **D. T. Bespalko** and S. Boumaiza, "X-parameter Measurement Challenges for Un-Matched Device Characterization," *Microwave Measurements Conference (ARFTG)*, 2010 75th ARFTG. IEEE, Jun. 2010.

Invited Presentations

1. **D. T. Bespalko**, S. Boumaiza, "Systematic Improvement of PHD Model Prediction by Improving Harmonic Source Superposition During Model Extraction," *ARFTG NVNA Open Forum*, Nov. 2015.

2. **D. T. Bespalko**, S. Boumaiza, "Evolution from Nonlinear Device Characterization Test-bench to Behavioural Model Extraction of High Power Transistors," *ARFTG NVNA Open Forum*, Nov. 2011.

3. L. Betts, **D. T. Bespalko**, S. Boumaiza, "Power Amplifier Design Utilizing the NVNA and X-Parameters," Micro-Application, *International Microwave Symposium*, Jun. 2011.

4. **D. T. Bespalko**, D. Wu, and S. Boumaiza, "Load-Pull based X-Parameter for Waveform Engineered High Efficiency Power Amplifiers," *Focus Microwave Annual User Seminar Nonlinear and Noise Device Characterization*, Jun. 2009.

6.3 Future Research Projects

Investigation and optimization of nonlinear behaviour is built around characterization tools that enhance our understanding of high-frequency circuits. This understanding is built on layers of integrated hardware and software that performs automated parametric sweeps, optimization, and modeling. Additional calibration and modeling routines can be implemented to enhance the NCS by performing sophisticated automated measurement routines. Further hardware and software development is necessary to support broadband modulated signal waveforms that more accurately represent the final PA design, permitting the study of memory effects and real-time load-pull[84][30].

The success of a multi-harmonic behavioural model is dictated by improving the speed of characterization and by improving the conditioning of the LSE problem matrix. Future support of real-time load-pull systems can dramatically increase the model extraction speed, allowing for more sophisticated models to be generated based on larger measurement datasets. By adding support for envelope load-pull measurements, models can be designed to capture electrical and thermal memory effects.

Increasing PAPR and efficiency requirements dictate that future PA designs will need to consider and exploit the nonlinear operation of RF transistors. For dual-band PA performance to be simultaneously maximized at both frequencies, much of the existing high-efficiency PA design techniques must be re-evaluated, and future design must be based solely around time-domain voltage and current stochastic Waveform Engineering.

APPENDICES

Appendix A

Characterization Software

The NCS is a control software utility that was primarily developed in MATLAB/Java to control the calibration, characterization and modeling of nonlinear devices[85].

A.1 Instrument Manager

User-defined instrument drivers can be displayed in a Measurement Dashboard, as shown in Fig A.1. Instrument drivers that are unspecified (do not exist in the test-bench) exhibit a default measurement behaviour.

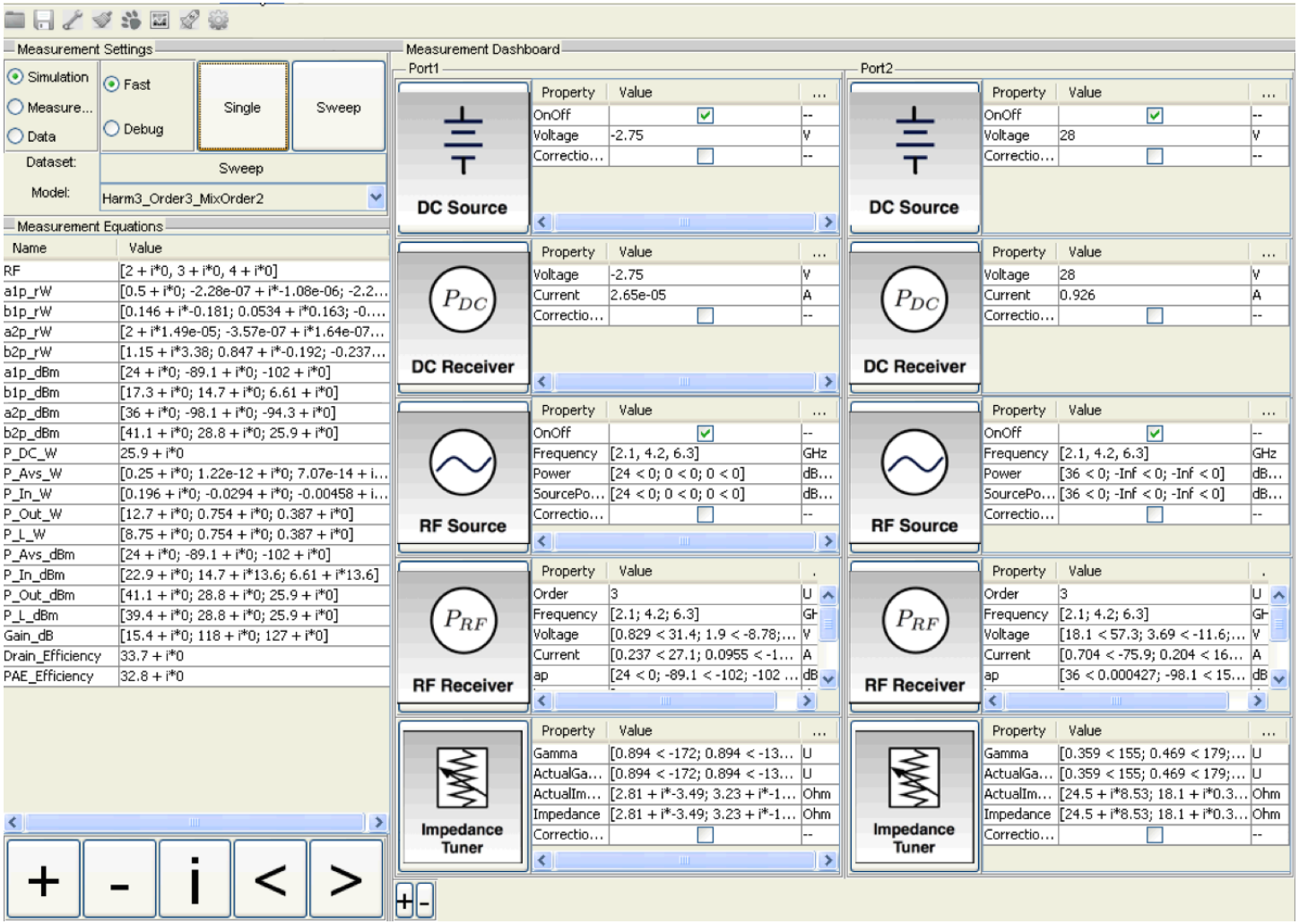


Figure A.1: Measurement Dashboard

A.2 Variable Sweep Generator

A sequence of linear, logarithmic, random, and relative sweep plans are used to perform an automated measurement. The supported sweep variables are shown in Table A.1.

A.3 Measurement Viewer

This is a flexible multi-tab window consisting of tiled plotting axes shown in A.2. The *Settings* button allows the user to edit plots using the built-in MATLAB plotting tools as shown in Fig. A.2. Custom plot functions can be added by modifying the *Plot Categories* and *Plot Signatures* databases.

Table A.1: Supported Variable Sweep Plans

	Sweep	Sweep Description
	$harmonic$	Measurement harmonic index [0...# of harmonics]
	$port$	Measurement port index [1...# of ports]
	$ssphase$	Small-signal extraction tone phase index [1...# of phase points]
	$ssfreq$	Small-signal extraction tone harmonic index [1...# of harmonics]
	$ssport$	Small-signal extraction tone port index [1...# of ports]
	$\|a_{jl}\|$	Amplitude of small-signal RF power on port j at harmonic l
	$\angle A_{jl}$	Phase of small-signal RF power on port j at harmonic l
	$Re\,(a_{jl})$	Real component of small-signal RF power on port j at harmonic l
	$Im\,(a_{jl})$	Imaginary component of small-signal RF power on port j at harmonic l
	$\|A_{jl}\|$	Amplitude of large-signal RF power on port j at harmonic l
	$\angle A_{jl}$	Phase of large-signal RF power on port j at harmonic l
	$Re\,(A_{jl})$	Real component of large-signal RF power on port j at harmonic l
	$Im\,(A_{jl})$	Imaginary component of large-signal RF power on port j at harmonic l
	V_{j0}	DC voltage on port j
	$\|\Gamma_{jl}\|$	Amplitude of reflection coefficient on port j at harmonic l
	$\angle \Gamma_{jl}$	Phase of reflection coefficient on port j at harmonic l
	$Re\,(\Gamma_{jl})$	Real component of reflection coefficient on port j at harmonic l
	$Im\,(\Gamma_{jl})$	Imaginary component reflection coefficient on port j at harmonic l
	$\|Z_{jl}\|$	Amplitude of impedance on port j at harmonic l
	$\angle Z_{jl}$	Phase of impedance on port j at harmonic l
	$Re\,(Z_{jl})$	Real component of impedance on port j at harmonic l
	$Im\,(Z_{jl})$	Imaginary component impedance on port j at harmonic l
	$fund_j$	Fundamental frequency on port j

	Required		Model Extraction		Large-Signal Operating Point

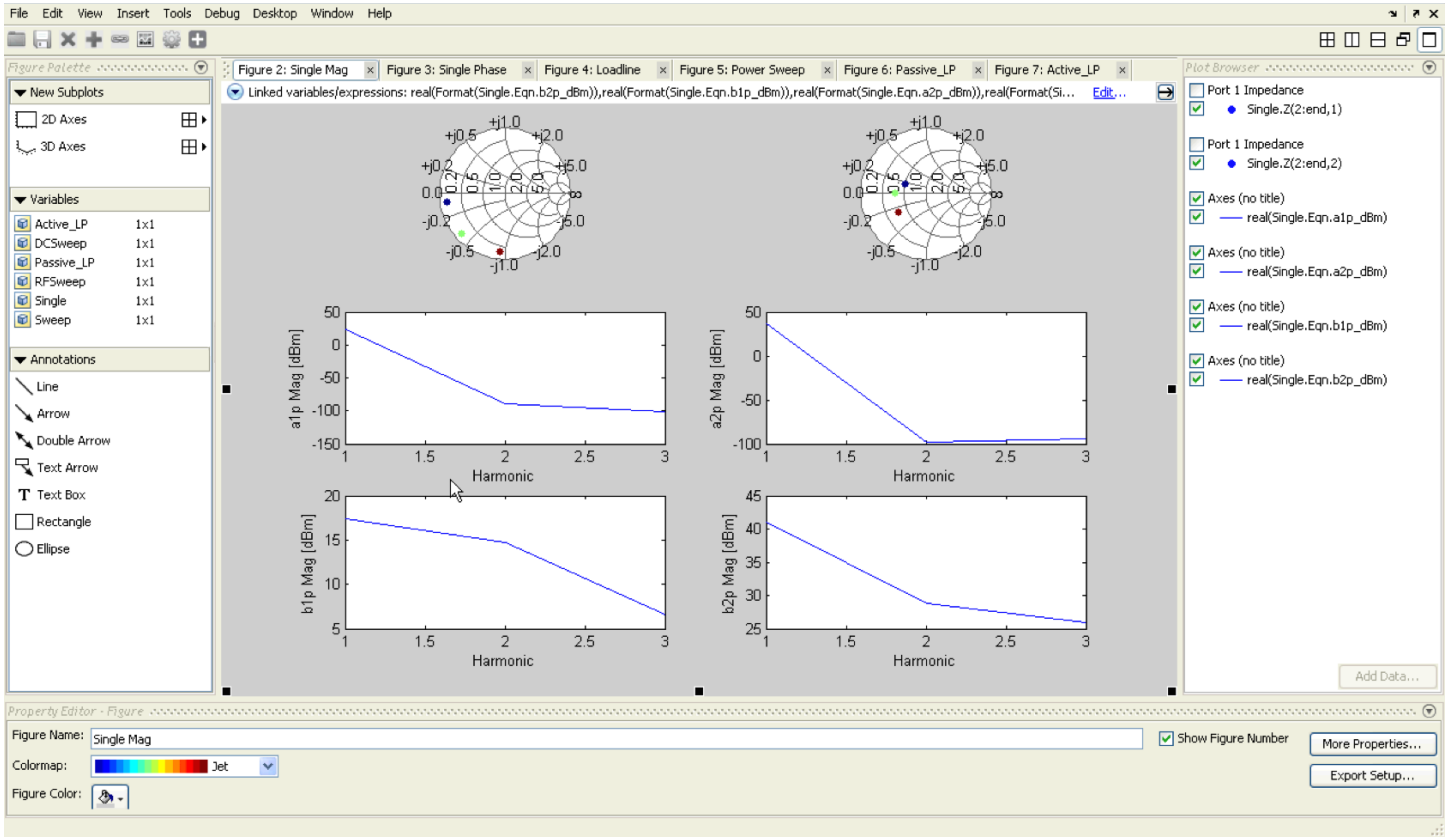


Figure A.2: Measurement Viewer with MATLAB plotting tools

Appendix B

Multi-Harmonic Volterra Model Results

Several Multi-Harmonic Volterra Models were extracted and compared to independent measurements that used a similar random distribution. While the validation of the model is not as robust as the comparison shown in Section 4.4, it gives valuable insight into the model input parameter trade-offs that influence the model accuracy. Table B.1 - B.3 demonstrates the predictive capability of the model using different input parameters and evaluates the model extraction in both simulation and measurement mode. The simulations are extracted from the compact circuit model of the Cree CGH40010F and CGH40045F 10W and 45W GaN transistors.

Table B.1: Multi-Harmonic Volterra Model Notes

ID #	Part	Ports	Num Harm.	Max Order	Max Inputs	A_{11} $[\sqrt{W}]$	A_{21} $[\sqrt{W}]$	Notes
1	10W	1	2	3	2	0.5		Cannot predict DC gate current of compact circuit model
2	10W	2	2	3	2	0.5	0	2-Port model has similar accuracy to 1-port model, suggesting there is good isolation between the ports.
3	10W	2	2	3	2	0.5	0	2X extrapolation during validation increases the error by 10X indicating poor extrapolation.
4	10W	2	2	3	2	0.5	0	2X oversampling provides little improvement to the model validation.
5	10W	2	2	3	3	0.5	0	A 2X increase in model coefficients, degrades accuracy of all harmonics, especially those with a large increase in coefficients.
6	10W	2	2	3	2	0.5	2	$1f_0$ load mismatch results in all port/harmonic mismatch, thus reducing coverage. It is impossible to know which $\angle A_{21}$ results in strongest nonlinearity without de-embedding to the intrinsic device.
7	10W	2	2	3	2	0.5	$2\angle 90$	
8	10W	2	2	3	2	0.5	$2\angle 180$	
9	10W	2	2	3	2	0.5	$2\angle$-90	
10	10W	2	3	3	2	0.5	0	Added coefficients reduce the accuracy of lower-harmonics. $3rd$ harmonic is modelled poorly due to decreased dynamic range.
11	10W	2	3	3	3	0.5	0	Increase in coefficients deteriorates accuracy.
12	45W	2	2	3	2	1	0	45W device model is more accurate because it is only sensitive to two harmonics.
13	45W	2	2	3	3	1	0	Increase in coefficients deteriorates accuracy.
14	45W	2	3	3	3	1	0	See Test #10 Notes.

Table B.2: Multi-Harmonic Volterra Model Extraction

	Simulation								Measurement							
	$0f_0$		$1f_0$		$2f_0$		$3f_0$		$0f_0$		$1f_0$		$2f_0$		$3f_0$	
ID #	Port1 $[dB]$	Port2 $[dB]$	Port1 $[dB]$	Port2 $[dB]$	Port1 $[dB]$	Port2 $[dB]$	Port1 $[dB]$	Port2 $[dB]$	Port1 $[dB]$	Port2 $[dB]$	Port1 $[dB]$	Port2 $[dB]$	Port1 $[dB]$	Port2 $[dB]$	Port1 $[dB]$	Port2 $[dB]$
1	-3.7		-24.3		-20.6				-54.0		-41.1		-25.4			
2	-9.1	-31.7	-31.1	-28.9	-29.3	-22.3			-28.5	-37.1	-37.1	-31.0	-26.5	-22.3		
3	-9.1	-31.7	-31.1	-28.9	-29.3	-22.3			-28.5	-37.1	-37.1	-31.0	-26.5	-22.3		
4	-0.5	-30.4	-27.5	-27.1	-26.6	-20.0										
5	2.2	-31.9	-29.5	-27.6	-24.0	-18.0			-28.2	-35.2	-37.3	-29.6	-25.3	-18.7		
6	-6.3	-33.6	-38.8	-33.2	-26.3	-24.9			-29.3	-39.5	-40.9	-31.9	-26.8	-26.4		
7	-15.2	-29.9	-35.2	-36.6	-31.6	-33.2			-27.7	-29.7	-28.2	-32.1	-25.1	-26.3		
8	-10.5	-30.5	-35.8	-30.9	-24.1	-26.1			-29.2	-30.3	-27.3	-29.9	-25.5	-26.9		
9	-4.5	-37.0	-35.3	-31.1	-24.6	-27.5			-27.8	-37.7	-39.1	-34.6	-26.7	-27.5		
10	1.0	-33.7	-30.2	-29.0	-24.6	-18.4	-14.3	-19.2								
11	-1.2	-32.1	-27.2	-26.3	-23.8	-18.1	-11.5	-18.2								
12	-14.6	-26.3	-48.3	-35.6	-31.6	-22.1			-29.9	-32.5	-45.1	-34.5	-28.1	-19.5		
13	4.1	-25.2	-44.3	-32.2	-26.6	-14.0			-31.2	-28.4	-42.7	-31.4	-27.3	-10.6		
14	-14.0	-27.7	-47.4	-36.9	-29.9	-13.0	-17.9	-22.8								

Table B.3: Multi-Harmonic Volterra Model Verification

	Simulation								Measurement							
	$0f_0$		$1f_0$		$2f_0$		$3f_0$		$0f_0$		$1f_0$		$2f_0$		$3f_0$	
ID #	Port1 [dB]	Port2 [dB]	Port1 [dB]	Port2 [dB]	Port1 [dB]	Port2 [dB]	Port1 [dB]	Port2 [dB]	Port1 [dB]	Port2 [dB]	Port1 [dB]	Port2 [dB]	Port1 [dB]	Port2 [dB]	Port1 [dB]	Port2 [dB]
1	-6.6		-26.2		-21.6				-49.8		-35.0		-24.0			
2	-2.3	-26.6	-29.2	-28.3	-24.2	-16.7			-25.9	-27.6	-33.5	-28.2	-25.5	-20.6		
3	5.1	-14.7	-17.7	-17.1	-14.0	-4.6			-15.5	-14.7	-24.4	-19.2	-10.4	-6.6		
4	-3.0	-27.6	-28.2	-28.0	-25.2	-16.8										
5	7.9	-24.0	-22.8	-21.3	-19.4	-13.8			-26.0	-29.9	-32.9	-26.3	-23.2	-17.1		
6	-6.3	-33.6	-38.8	-33.2	-26.3	-24.9			-21.9	-37.7	-37.2	-29.4	-25.2	-23.7		
7	-7.7	-26.1	-26.1	-28.1	-24.8	-27.6			-23.5	-28.8	-26.9	-30.3	-23.6	-21.8		
8	-0.6	-18.8	-29.4	-26.2	-21.7	-21.3			-27.1	-28.5	-24.9	-28.7	-23.8	-25.2		
9	-3.8	-30.7	-29.1	-26.9	-22.6	-24.4			-23.1	-36.3	-36.4	-32.6	-26.8	-25.9		
10	6.4	-25.7	-27.3	-26.9	-23.2	-16.3	-11.7	-19.3								
11	2.8	-20.1	-25.0	-24.5	-22.1	-15.2	-8.4	-13.7								
12	9.2	-23.8	-47.7	-37.6	-25.3	-17.2			-22.9	-28.1	-37.8	-31.4	-26.8	-17.6		
13	-8.9	-21.9	-38.9	-29.3	-20.2	-13.7			-20.9	-27.4	-38.9	-30.0	-26.8	-16.0		
14	2.9	-21.5	-44.8	-34.9	-26.3	-14.8	-14.9	-13.7								

References

[1] S. Woodington, T. Williams, H. Qi, D. Williams, L. Pattison, A. Patterson, J. Lees, J. Benedikt, and P. J. Tasker, "A novel measurement based method enabling rapid extraction of a RF Waveform Look-Up table based behavioral model," *Microwave Symposium Digest, 2008 IEEE MTT-S International*, pp. 1453–1456, 2008.

[2] S. Woodington, R. Saini, D. Williams, J. Lees, J. Benedikt, and P. J. Tasker, "Behavioral model analysis of active harmonic load-pull measurements," *Microwave Symposium Digest (MTT), 2010 IEEE MTT-S International*, pp. 1688–1691, 2010.

[3] X. Fu, D. T. Bespalko, and S. Boumaiza, "Novel Dual-Band Matching Network for Effective Design of Concurrent Dual-Band Power Amplifiers," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 61, no. 1, pp. 293–301, Jan. 2014.

[4] S. C. Cripps, "Efficiency Enhancement Techniques," in *RF Power Amplifiers For Wireless Communications*. Artech House, 2006, pp. 285–334.

[5] B. Kim, J. Kim, I. Kim, and J. Cha, "The Doherty Amplifier," *IEEE Microwave Magazine*, pp. 1–9, Oct. 2006.

[6] J. Moon, J. Kim, I. Kim, J. Kim, and B. Kim, "A Wideband Envelope Tracking Doherty Amplifier for WiMAX Systems," *IEEE Microwave and Wireless Components Letters*, vol. 18, no. 1, pp. 49–51.

[7] V. Teppati, A. Ferrero, and M. Sayed, "Modern RF and Microwave Measurement Techniques." Cambridge: Cambridge University Press, 2009, pp. 130–157.

[8] F. Verbeyst and E. Vandamme, "Large-Signal Network Analysis. Overview of the measurement capabilities of a Large-Signal Network Analyzer," in *ARFTG Conference Digest-Fall, 58th.* IEEE, 2001, pp. 1–14.

[9] D. Vye, "Fundamentally changing nonlinear microwave design," *Microwave Journal*, 2010.

[10] D. E. Root, "Measurement-Based Mathematical Active Device Modeling for High Frequency Circuit Simulation," *IEICE Transactions on Electronics*, vol. E82-C, no. 6, pp. 924–936, Jun. 1999.

[11] P. Aaen, J. A. Plá, and J. Wood, "Modeling the Active Transistor," in *Modeling and Characterization of RF and Microwave Power FETs.* Cambridge: Cambridge University Press, 2009, pp. 183–257.

[12] J. P. Dunsmore, "Handbook of Microwave Component Measurements: with Advanced VNA Techniques," Oct. 2012, pp. 66–123.

[13] Agilent Technologies, "Agilent Vector Signal Analysis Basics," Agilent Technologies, Tech. Rep. Application Note 150-15, Jul. 2004.

[14] S. A. Maas, "Harmonic-Balance Analysis and Related Methods," in *Nonlinear Microwave and RF Circuits.* Artech House, Jan. 2003, pp. 119–212.

[15] J. Arrillaga and N. Watson, *Power Systems Electromagnetic Transients Simulation.* The Institution of Engineering and Technology, Michael Faraday House, Six Hills Way, Stevenage SG1 2AY, UK: IET, 2003.

[16] D. Williams, P. Hale, and K. A. Remley, "The Sampling Oscilloscope as a Microwave Instrument," *Microwave Magazine, IEEE*, vol. 8, no. 4, pp. 59–68, 2007.

[17] Z. Aboush, C. Jones, G. Knight, A. Sheikh, H. Lee, J. Lees, J. Benedikt, and P. J. Tasker, "High Power Active Harmonic Load-Pull System for Characterization of High Power 100Watt Transistors," pp. 1–4, Aug. 2000.

[18] J. Verspecht, "The return of the sampling frequency convertor," in *62nd ARFTG Microwave Measurements Conference*. IEEE, 2003, pp. 155–164.

[19] Hewlet Packard, "Microwave Transition Analyzer," pp. 1–6, 1991.

[20] NMDG Engineering and Maury Microwave Corporation, "A Network Analyzer for Complete Active Component Characterization and Real-Time Harmonic Load Pull," *Microwave Journal*, pp. 1–3, Nov. 2004.

[21] J. Verspecht, J.-P. Teyssier, and D. F. De Groote, "SWAP-X402," pp. 1–4, Dec. 2009.

[22] Agilent Technologies, "Agilent Nonlinear Vector Network Analyzer (NVNA)," Tech. Rep., 2008.

[23] NMDG Engineering, "Product Note : NM320," pp. 1–8, Jan. 2013.

[24] D. M. Pozar, "Microwave Engineering 3rd Edition." John Wiley & Sons, 2005, pp. 258–261.

[25] C. Tsironis, R. Meierer, B. Hosein, T. Beauchamp, and R. Jallad, "MPT, Universal Multi-Purpose Tuner," pp. 1–5, Nov. 2005.

[26] Y. Takayama, "A New Load-Pull Characterization Method for Microwave Power Transistors," in *MTT-S International Microwave Symposium Digest*. MTT005, 1976, pp. 218–220.

[27] P. J. Tasker, "RF IV Waveform Measurement and Engineering Systems," *armms.org*, 2010.

[28] V. Teppati, A. Ferrero, and U. Pisani, "Recent Advances in Real-Time Load-Pull Systems," *IEEE Transactions on Instrumentation and Measurement*, vol. 57, no. 11, pp. 2640–2646, Nov. 2008.

[29] A. Alghanim, J. Lees, T. Williams, J. Benedikt, and P. J. Tasker, "Reduction of electrical baseband memory effect in high-power LDMOS devices using optimum termination for IMD3 and IMD5 using active load-pull," *Microwave Symposium Digest, 2008 IEEE MTT-S International*, pp. 415–418, 2008.

[30] M. Marchetti, M. J. Pelk, K. Buisman, W. Neo, M. Spirito, and L. C. N. De Vreede, "Active Harmonic Load–Pull With Realistic Wideband Communications Signals," *IEEE Transactions on Microwave Theory and Techniques*, vol. 56, no. 12, pp. 2979–2988, 2008.

[31] Y. Tsividis and C. McAndrew, *Operation and Modeling of the MOS Transistor*, Oct. 2010.

[32] B. D. Huebschman, P. B. Shah, and R. Del Rosario, "Theory and Operation of Cold Field-effect Transistor (FET) External Parasitic Parameter Extraction," Tech. Rep., May 2009.

[33] J. Wood and D. E. Root, "Bias-dependent linear scalable millimeter-wave FET model," *IEEE Transactions on Microwave Theory and Techniques*, vol. 48, no. 12, pp. 2352–2360, Dec. 2000.

[34] D. E. Root, "Future Device Modeling Trends," *IEEE Microwave Magazine*, vol. 13, no. 7, pp. 45–59, 2012.

[35] J. Sirois, S. Boumaiza, M. Helaoui, G. Brassard, and F. M. Ghannouchi, "A robust modeling and design approach for dynamically loaded and digitally linearized Doherty amplifiers," *IEEE Transactions on Microwave Theory and Techniques*, vol. 53, no. 9, pp. 2875–2883, 2005.

[36] B. Fehri and S. Boumaiza, "Systematic Estimation of Memory Effects parameters in PowerAmplifiers' Behavioral Models," pp. 1–4, Mar. 2011.

[37] J. C. Pedro and N. B. Carvalho, "Nonlinear Analysis Techniques for Distortion Prediction," in *Intermodulation Distortion in Microwave and Wireless Circuits*. Artech House, Dec. 2002, pp. 73–194.

[38] J. Hu, "Radio Frequency Front-End Measurement and Modeling," Ph.D. dissertation, North Carolina State University, 2011.

[39] M. Schetzen, *The Volterra and Wiener Theories of Nonlinear Systems*. Krieger Publishing Company, Jan. 2006.

[40] C. Fernandes, J. Mota, and G. Favier, "MIMO Volterra modeling for nonlinear communication channels," *... and Nonlinear Models*, 2010.

[41] S. A. Maas, "Volterra-Series and Power-Series Analysis," in *Nonlinear Microwave and RF Circuits*. Artech House, Jan. 2003, pp. 215–276.

[42] M. Steer, P. J. Khan, and R. Tucker, "Relationship between Volterra series and generalized power series," in *Proceedings of the IEEE*, 1983, pp. 1453–1454.

[43] P. Roblin, "Behavioral Modeling," in *Nonlinear RF Circuits and Nonlinear Vector Network Analyzers*. Cambridge University Press, Jun. 2011, pp. 160–200.

[44] F. Verbeyst and M. V. Bossche, "The Volterra input-output map of a high frequency amplifier as a practical alternative to load-pull measurements," in *Instrumentation and Measurement Technology Conference, 1994. IMTC/94. Conference Proceedings. 10th Anniversary. Advanced Technologies in I & M., 1994 IEEE*. IEEE, 1994, pp. 283–286.

[45] J. Verspecht, "Everything you've always wanted to know about Hot-S22 (but we're afraid to ask)," Tech. Rep., 2002.

[46] J. Verspecht and D. E. Root, "Polyharmonic distortion modeling," *Microwave Magazine, IEEE*, vol. 7, no. 3, pp. 44–57, 2006.

[47] D. E. Root, J. Verspecht, J. Horn, and M. Marcu, "X-Parameters - Fundamental Concepts," in *X-Parameters*. Cambridge University Press, Sep. 2013.

[48] J. Horn, D. E. Root, and G. Simpson, "GaN Device Modeling with X-Parameters," *Compound Semiconductor Integrated Circuit Symposium (CSICS), 2010 IEEE*, pp. 1–4, 2010.

[49] P. J. Tasker, "Practical waveform engineering," *IEEE Microwave Magazine*, vol. 10, no. 7, pp. 65–76.

[50] Agilent Technologies, "High Power Amplifier Measurements Using Agilent's Nonlinear Vector Network Analyzer," pp. 1–27, Feb. 2010.

[51] G. Gonzalez, "Representations of Two-Port Networks," in *Microwave Transistor Amplifiers: Analysis and Design*. Prentice Hall, 1996, pp. 1–86.

[52] P. Roblin, "Large-Signal Vector Measurement Techniques with NVNAs," in *Nonlinear RF Circuits and Nonlinear Vector Network Analyzers*. Cambridge University Press, Jun. 2011, pp. 17–63.

[53] D. T. Bespalko, D. Wu, and S. Boumaiza, "Load-Pull based X-Parameter for Waveform Engineered High Efficiency Power Amplifiers," in *Focus Microwave Annual User Seminar Nonlinear and Noise Device Characterization*, Boston, MA, Jun. 2009.

[54] M. L. Edwards and J. H. Sinsky, "A new criterion for linear 2-port stability using a single geometrically derived parameter," *IEEE Transactions on Microwave Theory and Techniques*, vol. 40, no. 12, pp. 2303–2311, Dec. 1992.

[55] J. Verspecht, "Calibration of a measurement system for high frequency nonlinear devices," *Brussels*, 1995.

[56] J. P. Dunsmore, *Handbook of Microwave Component Measurements*, ser. with Advanced VNA Techniques. Chichester, UK: John Wiley & Sons, Aug. 2012.

[57] M. Wollensack, J. Hoffmann, J. Ruefenacht, and M. Zeier, "VNA Tools II: S-parameter uncertainty calculation," in *Microwave Measurement Conference (ARFTG), 2012 79th ARFTG*. IEEE, 2012, pp. 1–5.

[58] D. F. Williams, J. C. M. Wang, and U. Arz, "An optimal vector-network-analyzer calibration algorithm," *IEEE Transactions on Microwave Theory and Techniques*, vol. 51, no. 12, pp. 2391–2401, Dec. 2003.

[59] Focus Microwaves Inc., "MPIV command language V8," pp. 1–4, Mar. 2014.

[60] Agilent Technologies, "Agilent PNA-X Series Microwave Network Analyzers," pp. 1–42, Mar. 2014.

[61] D. T. Bespalko and S. Boumaiza, "X-parameter measurement challenges for unmatched device characterization," in *Microwave Measurements Conference (ARFTG), 2010 75th ARFTG*. IEEE, 2010, pp. 1–4.

[62] G. Gonzalez, *Microwave Transistor Amplifiers: Analysis and Design.* Prentice Hall, 1996.

[63] A. Amini, D. Bespalko, and S. Boumaiza, "On the Poly-Harmonic Distortion Modeling Framework and its Practical Realization," in *IEEE Topical Symposium on Power Amplifiers for Wireless Communications*, San Diego, CA, Sep. 2012.

[64] D. T. Bespalko, A. Amini, and S. Boumaiza, "A Higher-Order Multi-Harmonic Behavioural Model for High-Efficiency Power Amplifier Design," *IEEE Transactions on Microwave Theory and Techniques.*

[65] Y. Takayama, K. Uchida, T. Fujita, and K. Maenaka, "Microwave dual-band power amplifiers using two-frequency matching," *Electronics and Communications in Japan (Part II: Electronics)*, vol. 89, no. 5, pp. 17–24, 2006.

[66] K. Krishnamurthy, "Ultra-broadband, Efficient, Microwave Power Amplifiers in Gallium Nitride HEMT Technology," Ph.D. dissertation, Bell & Howell Information and Learning Company, Santa Barbara, Jun. 2000.

[67] A. Fukuda, T. Furuta, H. Okazaki, and S. Narahashi, "A 0.9-5- GHz Wide-Range 1W-Class Reconfigurable Power Amplifier Employing RF-MEMS Switches," *Microwave Symposium Digest, 2006. IEEE MTT-S International*, pp. 1859–1862, 2006.

[68] W. C. E. Neo, Y. Lin, X. D. Liu, L. C. N. De Vreede, L. E. Larson, M. Spirito, M. J. Pelk, K. Buisman, A. Akhnoukh, A. De Graauw, and L. K. Nanver, "Adaptive Multi-Band Multi-Mode Power Amplifier Using Integrated Varactor-Based Tunable Matching Networks," *IEEE Journal of Solid-State Circuits*, vol. 41, no. 9, pp. 2166–2176, Sep. 2006.

[69] G. Fischer, W. Eckl, and G. Kaminski, "RF-MEMS and SiC/GaN as enabling technologies for a reconfigurable multi-band/multi-standard radio," *Bell Labs Technical Journal*, vol. 7, no. 3, pp. 169–189, Mar. 2003.

[70] R. Negra, A. Sadeve, S. Bensmida, and F. M. Ghannouchi, "Concurrent Dual-Band Class-F Load Coupling Network for Applications at 1.7 and 2.14 GHz," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 55, no. 3, pp. 259–263.

[71] P. Colantonio, F. Giannini, R. Giofre, and L. Piazzon, "A Design Technique for Concurrent Dual-Band Harmonic Tuned Power Amplifier," *IEEE Transactions on Microwave Theory and Techniques*, vol. 56, no. 11, pp. 2545–2555.

[72] S. H. Ji, G. S. Hwang, C. S. Cho, J. W. Lee, and J. Kim, *836 MHz/1.95GHz Dual-Band Class-E Power Amplifier Using Composite Right/Left-Handed Transmission Lines*. IEEE, Dec. 2006.

[73] H. Zhang, "Compact, Reconfigurable and Dual-Band Microwave Circuits," Ph.D. dissertation, Hong Kong, Jan. 2007.

[74] C. Monzon, "Analytical derivation of a two-section impedance transformer for a frequency and its first harmonic," *Microwave and Wireless Components Letters, IEEE*, vol. 12, no. 10, pp. 381–382, 2002.

[75] H. Y. A. Yim and K. K. M. Cheng, "Novel dual-band planar resonator and admittance inverter for filter design and applications," in *IEEE MTT-S International Microwave Symposium Digest, 2005.* IEEE, pp. 2187–2190.

[76] G. Siso, J. Bonache, and F. Martin, "Dual-band Y-junction power dividers implemented through artificial lines based on complementary resonators," *Microwave Symposium Digest, 2008 IEEE MTT-S International*, pp. 663–666, 2008.

[77] A. S. Mohra and M. A. Alkanhal, "Dual Band Wilkinson Power Dividers Using T-Sections," *Journal of Microwaves, Optoelectronics and Electromagnetic Applications*, pp. 1–8, Sep. 2008.

[78] W.-P. Liu and Y.-W. Tung, "Third-order dual-band filter with dual-band admittance inverter," *Microwave and Optical Technology Letters*, vol. 53, no. 1, pp. 140–143, Nov. 2010.

[79] G. Castaldi, V. Fiumara, and I. M. Pinto, "A Dual-Band Chebyshev Impedance Transformer," *Wiley Periodicals, Inc. Microwave and Optical Technology Letters*, vol. 39, no. 2, pp. 141–145, 2003.

[80] X. Guan, Z. Ma, P. Cai, Y. Kobayashi, T. Anada, and G. Hagiwara, "Synthesis of dual-band bandpass filters using successive frequency transformations and circuit conversions," *IEEE Microwave and Wireless Components Letters*, vol. 16, no. 3, pp. 110–112.

[81] C. Monzon, "Analytical derivation of a two-section impedance transformer for a frequency and its first harmonic," vol. 12, no. 10, pp. 381–382, 2002.

[82] Y.-H. Chun and J.-S. Hong, "Compact wide-band branch-line hybrids," *IEEE Transactions on Microwave Theory and Techniques*, vol. 54, no. 2, pp. 704–709, 2006.

[83] D. T. Bespalko and S. Boumaiza, "Concurrent dual-band GaN power amplifier with compact microstrip matching network," *Microwave and Optical Technology Letters*, vol. 51, no. 6, pp. 1604–1607, Jun. 2009.

[84] M. Akmal, "An Enhanced Modulated Waveform Measurement System," Ph.D. dissertation, Jan. 2012.

[85] "Undocumented Secrets of MATLAB-Java Programming," 2011.