CARDIFF UNIVERSITY
PRIFYSGOL CAERDYDD

ORCA – Online Research @ Cardiff

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository:https://orca.cardiff.ac.uk/id/eprint/173881/

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Baddeley, Alexander, Quaglia, Roberto and Tasker, Paul J. 2024. Calibration error reduction in millimeter-wave load-pull systems measuring highly-reflective loads. IEEE Transactions on Microwave Theory and Techniques

Publishers page:

Please note:
Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies. See http://orca.cf.ac.uk/policies.html for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.

Calibration Error Reduction in Millimeter-Wave Load-Pull Systems Measuring Highly-Reflective Loads

Alexander Baddeley, *Graduate Student Member, IEEE,* Roberto Quaglia, *Member, IEEE,* Paul J. Tasker, *Fellow, IEEE,*

***Abstract*—This paper investigates sources of calibration error inherent in the load-pull measurement of millimeter-wave transistors when high reflection coefficient loads are applied. In these conditions, the calibration error on relative metrics such as power gain and power added efficiency can be significant and strongly dependent on the reflection coefficient. Using both simulations and experiments, the paper tries to discriminate the source of uncertainty between noise and "operator" errors. While the former can be minimized by maximizing the dynamic range and improving the filtering on the receiver, the latter, linked to the small mechanical perturbations that occur when changing the system configuration from calibration to measurement, is unavoidable in most load pull systems.**

To address this issue, the paper proposes a method, based on the TRL calibration, that consists of load-pulling the thru and line calibration standards using the measurement system in its measurement configuration and calculating their scattering parameters by means of least squares on the measured dataset. This enables a recalculation of the calibration coefficients previously obtained with the measurement system in calibration configuration and allows for a significant reduction in calibration error. This method has been tested and applied to measurements of mm-wave devices in both active and passive load-pull systems, demonstrating a significant impact on the performance metrics being measured. Calibration error, at 82.5 GHz, is shown to have resulted in a power gain inaccuracy as high as 0.7 dB at reflection coefficient of 0.7, which led to an underestimation of power added efficiency of 8.9 % for a gallium arsenide pHEMT.

***Index Terms*—Calibration, High-electron mobility transistors, Galium nitride, Galium arsenide, Load-pull, Microwave measurements, Millimeter wave, Power amplifier (PA).**

I. Introduction

Load-pull (LP) systems play a pivotal role in the large-signal characterization of transistors in assessing device output power, efficiency, or linearity when they are loaded with terminations across the Smith Chart. This experimental data, which underpins the design and extraction/validation of large-signal models [1], [2], becomes crucial at mm-wave frequencies, where the probability of achieving successful design of integrated circuits in a single tape-out cycle is heavily dependent on the accuracy of the transistor models. Therefore, improving the accuracy of mm-wave transistor models could profoundly impact the speed and cost-effectiveness of product development for new mm-wave devices.

The authors are with the Centre for High Frequency Engineering, Cardiff University, Cardiff CF24 3AA, U.K. (e-mail: baddeleya@cardiff.ac.uk; quagliar@cardiff.ac.uk; tasker@cardiff.ac.uk).

Driven by rapid advancements in GaN processes, the use of GaN transistors in power amplifiers operating at Ka-band is becoming commonplace due to its power density and ample gain [3]. Conversely, at E-band, GaAs-based technologies remain the preferred technology due to their comparatively higher f_{T} and $f_{\max}$ and the lower power requirement for point-to-point radio systems $< 30\,\mathrm{dBm}$ [4]. However, there is a growing emergence of GaN devices capable of power amplification at E-band [5], [6].

Notwithstanding the technology of choice, mm-wave devices commonly require a load impedance for an optimal power match that corresponds to a $|\Gamma_{\mathrm{L}}| > 0.7$. This has an impact on the accuracy of the measurements, as it is well reported in literature that there is increased uncertainty in high frequency measurements as the load impedance moves towards the edge of the Smith Chart [7], [8], [9], [10], [11]. In [11] the authors use the NIST calibration uncertainty tool [12] to propagate uncertainties on measured waves from numerous sources of systematic error through to LP measurements. The uncertainty is quantified as a function of load impedance termination. Addressing this load-dependent uncertainty was the primary focus of [9] which expanded upon the method introduced in [8]. It presents a weighted minimization algorithm that adjusts the small-signal calibration error coefficients to reduce the calibration uncertainty across the Smith Chart, using the power gain from a LP measurement of a zero-length thru as the reference for calibration error.

This work aims to identify and address primary sources of calibration error in practical mm-wave LP measurements and provide a correction methodologically based on the thru-reflect-line (TRL) or multi-line TRL (mTRL) calibration technique. Section II details the measurement system and methodology employed, section III introduces the definition and previous work that focuses on uncertainty characterization and correction, along with establishing a baseline uncertainty level in two practical load-pull systems. Section IV offers theoretical and experimental characterization of errors present in the load-pull systems analyzed in the previous section. Section V presents a method to minimize error. Finally, section VI demonstrates the effectiveness of the calibration error reduction technique on the measurement of GaN and GaAs devices at 30 GHz and 82.5 GHz respectively.

II. System Description

A generalized real-time vector network analyzer based load-pull system is shown in Fig. 1.

A pair of reflectometers are used to sample power waves that are measured at the VNA receivers as a_0, b_0, a_3 and b_3. Through large signal calibration, these measured waves are translated to calibrated waves at the device under test (DUT) measurement plane denoted a_1, b_1, a_2 and b_2. By actively or passively controlling the load impedance, the system enables the measurement of a DUT large-signal parameters such as output power, gain, efficiency and Γ_{in} as a function of the measured Γ_{L}.

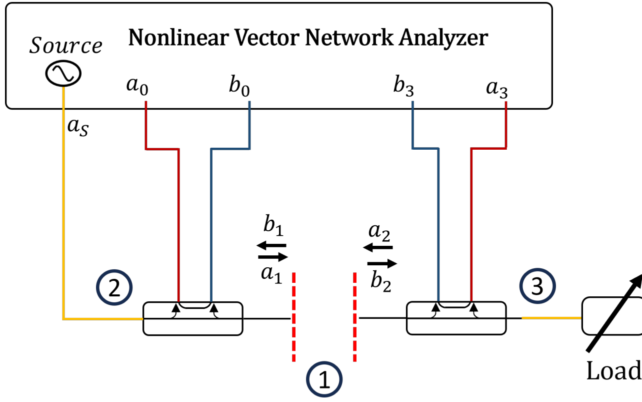

Fig. 1: Generalized real-time load-pull system.

This work uses two different mm-wave LP measurement systems to illustrate the most common sources of error that arise in the large-signal calibration and measurement of mm-wave devices spanning from Ka-band to E-band.

The first system, identified henceforth as the *Active LP system*, is a real-time vector-receiver active LP system, capable of measurements up to 67 GHz, shown in Fig. 2. It consists of a 67 GHz Rohde & Schwarz (ZVA 67) VNA, using an external test set based on Marki 2-67 GHz (C-0265) directional couplers. The VNA is connected to the measurement set-up using 1.5 m semi-rigid 1.85 mm coaxial cables. For this work, on-wafer measurements were conducted using the 100 μm 67 GHz MPI ground-signal-ground (GSG) probes.

In this paper, a maximum $|\Gamma_{\text{L}}| = 0.95$ was imposed to limit the reflection coefficient to a realistic measurement conditions.

The second system, identified henceforth as the *Passive LP system*, is a passive on-wafer LP system, capable of measurements up to 110 GHz, and is shown in Fig. 4 and detailed in [13]. This works is tailored for fundamental frequency measurements therefore the system is not set-up to use a harmonic phase reference (HPR). The system uses a 67 GHz Rhode & Schwarz VNA (ZVA 67) , Focus Microwave 110 GHz Delta tuners, R&S ZD-110 source, and a R&S mm-wave diplexer for the receiver chain, all integrated onto the MPI TS2000 probe station. The on-wafer probes used for this work were the 100 μm 110 GHz MPI GSG probes. The maximum reflection coefficient is limited by the tuner loss, with the maximum $|\Gamma_{\text{L}}| = 0.7$ at 82.5 GHz.

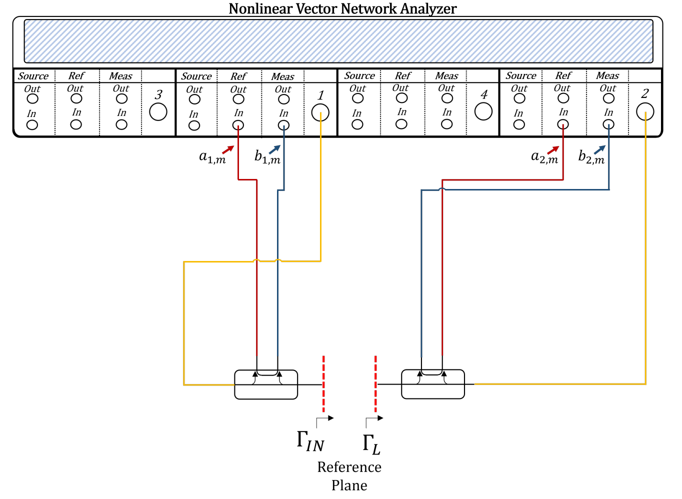

Fig. 2: Active LP measurement system diagram

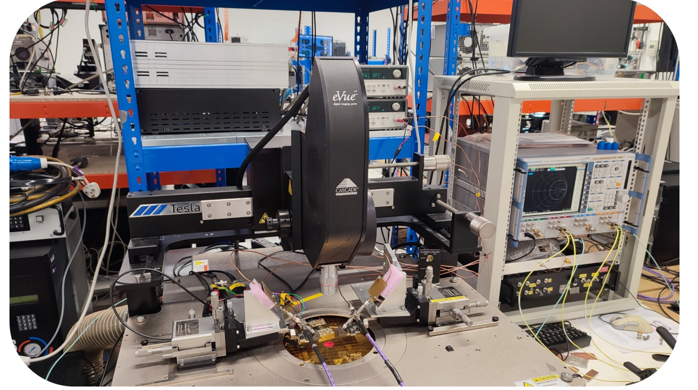

Fig. 3: Active LP measurement system

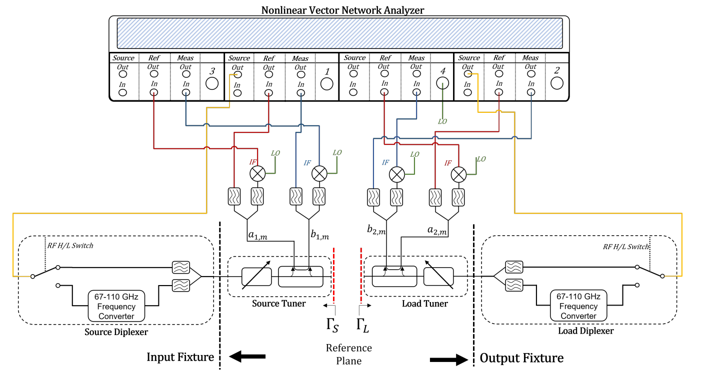

Fig. 4: Passive LP measurement system diagram

A. Calibration and Measurement procedure

Large-signal calibration follows the same procedure, independent of measurement topology. It consists of a two-port vector calibration to move the measurement reference plane from the VNA receivers to the DUT plane, as well as an absolute calibration that scales the error terms to allow for the measurement of absolute quantities. For this work, the TRL calibration algorithm [14] was used for vector calibration. The choice was motivated by requiring an accurate calibration

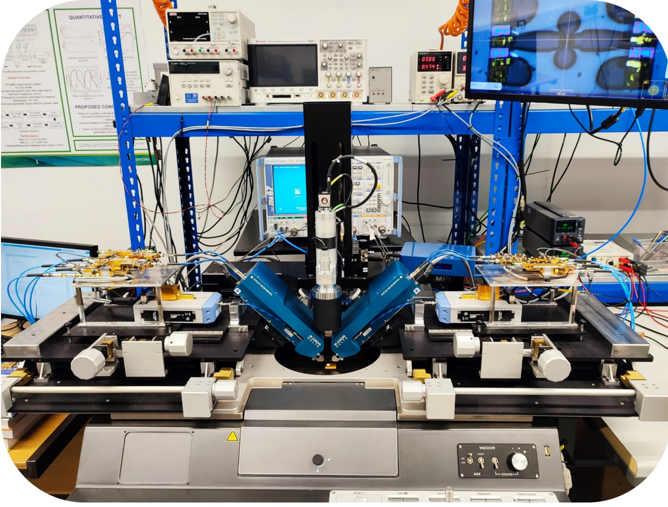

Fig. 5: Passive LP measurement system

technique for mm-wave on-wafer measurements [15], but not requiring the wide bandwidth calibration offered by mTRL [16], as this work focuses on fundamental frequency LP measurements.

Using the generalized real-time load-pull system diagram in Fig. 1 as reference, the following steps outline the fundamental frequency calibration and measurement procedure:

1) Configure the measurement system.
2) Perform on-wafer vector calibration at the DUT reference planes (position 1 in Fig. 1). Using a TRL calibration kit.
3) Verification of vector calibration.
4) Perform absolute calibration at the extended output port (position 3 in Fig. 1) Using a SOL coaxial calibration and power meter.
5) Add bias-tee, circulator and external amplifiers outside the reflectometers (position 2 and 3 in Fig. 1)
6) Verification of large-signal calibration.
7) Perform LP measurement.

III. Uncertainty Evaluation

A. Background

The contributions of uncertainty in load-pull (LP) measurements have been extensively documented, with a comprehensive investigation into sensitivity and propagation to measured waves detailed in [11].

The uncertainty from absolute calibration and its impact on absolute measured quantities has been reported in [7].

This work focuses on the uncertainty associated with the vector calibration coefficients and therefore its effect on the measured ratio quantities.

In [8], [9] and [17] correction methods have been developed to reduce residual calibration uncertainty, related to the vector calibration, by optimizing or updating the set of calibration coefficients based on a LP measurement of a known standard.

The method for evaluating calibration residual uncertainty [7], [8], [9], [10] involves the measurement of power gain from the LP measurement of a zero-length thru. The power gain is given by:

$$G_{\mathrm{P}} = \frac{P_{\mathrm{out}}}{P_{\mathrm{in}}} = \frac{|b_2|^2 - |a_2|^2}{|a_1|^2 - |b_1|^2} = \frac{|b_2|^2(1-|\Gamma_{\mathrm{L}}|^2)}{|a_1|^2(1-|\Gamma_{\mathrm{in}}|^2)} \tag{1}$$

where P_{in}, P_{out} are the delivered power to the input and output, respectively.

If the calibration is perfect, no errors, the power gain should equal one (0 dB) and be insensitive to variations in the reflection coefficient. However, calibration-induced measurement errors will result in a residual uncertainty that tends to rapidly increase moving towards the edge of the Smith Chart. In research, residual uncertainty as high as 1.5 dB at $|\Gamma_{\mathrm{L}}|$ = 0.95 is shown [9], [10].

This paper does not directly calculate uncertainty metrics using [12] or [11], as these have been covered extensively. Instead, it employs a combination of theoretical and experimental methods to identify and attribute sources of error in practical mm-wave LP systems. Errors identified solely from experimental measurements may be system specific, yet an attempt is made to provide explanations for similar error profiles reported in other papers that remain unexplained [8], [9], [10]. Consequently, the authors argue that a significant proportion of residual calibration uncertainty observed in mm-wave LP measurements stems from unavoidable operator error. The calibration and measurement process described in the previous section cannot be treated as discrete sections if high-accuracy measurements are to be achieved. The following sections outline a practical approach to identify sources of unavoidable calibration error and propose a method to minimize this error by updating the calibration coefficients whilst the measurement system is configured in its final state.

B. Reference uncertainty

To demonstrate this idea, a baseline uncertainty was established for both measurement systems in Fig. 2 and 4. This provides insight into the best-case calibration uncertainty achievable with these measurement systems before introducing error sources, such as random error and mechanical alteration, that arise when using commercial LP benches.

Random error associated with the systems dynamic range were mitigated by minimizing the level of attenuation between the DUT and measurement receivers and conducting measurements at a low IF bandwidth (1 Hz).

To minimize mechanical movement, absolute calibration was performed before vector calibration. Although this deviates from standard procedure, the accuracy of this step is not of concern as only ratio quantities are evaluated. The thru calibration standard was probed last for TRL calibration, with the probes remaining landed for the subsequent LP measurement to determine residual uncertainty.

Large-signal calibration was conducted ten times on both systems to establish a typical level of uncertainty under ideal conditions using the previously mentioned steps.

Fig. 6 and 7 illustrate the best-case residual uncertainties for these measurement systems. At 30 GHz, the active LP system exhibits a maximum uncertainty of $\pm$ 0.1 dB, while the passive

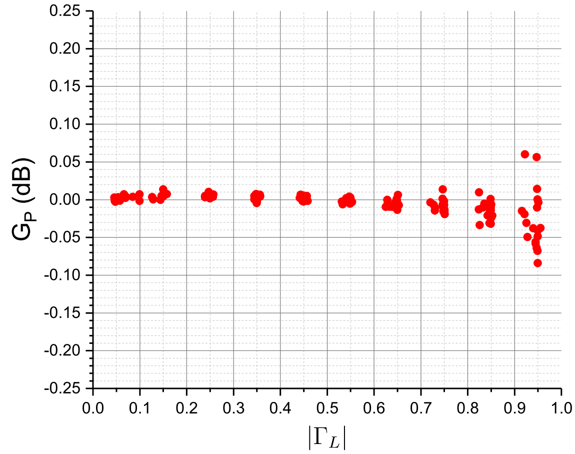

Fig. 6: LP measurement showing typical best-case residual power gain uncertainty as measured on a zero-length thru on the active LP system at 30 GHz over an equally distributed set of load impedances across the Smith Chart up to $|\Gamma_\mathrm{L}|$ = 0.95.

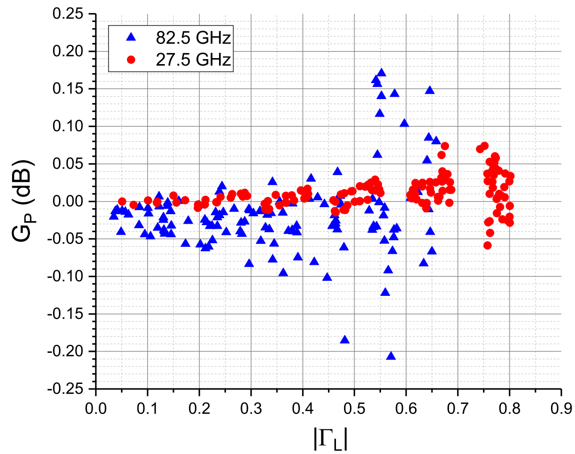

Fig. 7: LP measurement showing typical best-case residual power gain uncertainty as measured on a zero-length thru on the passive LP system at 82.5 GHz over an equally distributed set of load impedances across the Smith Chart up to $|\Gamma_\mathrm{L}|$ = 0.65.

LP system shows uncertainties of ± 0.1 dB at 27.5 GHz and ± 0.2 dB at 82.5 GHz, respectively. In both measurement systems a similar levels of uncertainty are observed at Ka-band. However, an increased level of uncertainty is shown at E-band. The uncertainty shown in these figures fall within the category of measurement repeatability uncertainty, which has been well reported in [18] and [19], therefore, further analysis is not provided. These uncertainties are used as a baseline reference for the calibration correction method described in section V.

IV. Calibration Error

This section will investigate the primary factors contributing to errors introduced by a realistic calibration procedure. First, it will provide a theoretical and practical demonstration of the impact of calibration dynamic range on LP measurement errors. Second, it will examine the contribution to error from mechanical movement after vector calibration.

To emulate a real-world measurement scenario, the systems will be appropriately configured : include high levels of attenuation between the DUT and receiver ports (Section IV A) along with mechanical alteration (Section IV B).

The application space for the active LP system is the measurement of a GaN device at Ka-band [20], while the passive LP system is used for the measurement of a GaAs device at E-band [21].

A. Random Error

The ability to measure a DUT with greater output power comes with a trade-off with measurement accuracy, owing to the finite dynamic range of the VNA and the associated measurement hardware [22]. In modern VNAs, the maximum output power available at the source port and dynamic range are frequency-dependent, but are typically around 10 dBm and 120 dB, respectively [23].

The challenge in such measurements stems from the degradation of an instrument's dynamic range due to added attenuation at the measurement receivers, aimed at preventing compression (the maximum linear power is -5 dBm for the VNA receivers), and set-up related loss. This results in a reduced Signal-to-Noise (SNR) Ratio at the measurement receivers. Therefore, this section aims to quantify this trade-off through ideal simulations and practical measurements. This will demonstrate the critical selection of the calibration IF bandwidth to minimize calibration error for LP measurements.

To quantify the trade-off between device output power and measurement accuracy, an investigation into the effect of dynamic range during vector calibration has on LP measurement accuracy was performed in script. The methodology involved applying a noise source to simulated waves of the TRL standards, performing calibration and propagating errors through to an ideal LP simulation. The steps are as follows:

1) Define the s-parameter matrix for TRL standards.
2) Generate ideal waves for each TRL standard (including switch correction).
3) Apply Gaussian white noise to each set of waves to replicate system dynamic range.
4) Perform calibration on the simulated waves.
5) Simulate LP using each calibration set (Ideal LP with infinite dynamic range).
6) Calculate power gain.
7) Repeat steps 3) - 6) N times using uncorrelated noise realizations.

In a real measurement system, the impact of the IF bandwidth, attenuation and set-up related loss lead to degradation of the system's dynamic range. In simulation, the dynamic range was varied from 90 dB to 50 dB, a reasonable practical range for high power LP measurements.

With N representing the number of noise iterations and μ the mean of the power gain distribution with respect to the noise iterations, the simulation was conducted with N = 10,000 and $\mu = 0$. The standard deviation, σ, of the power gain from LP due to a given dynamic range during calibration

is shown for each magnitude of reflection coefficient. This approach demonstrates the sensitivity of vector calibration and its impact on LP measurement uncertainty.

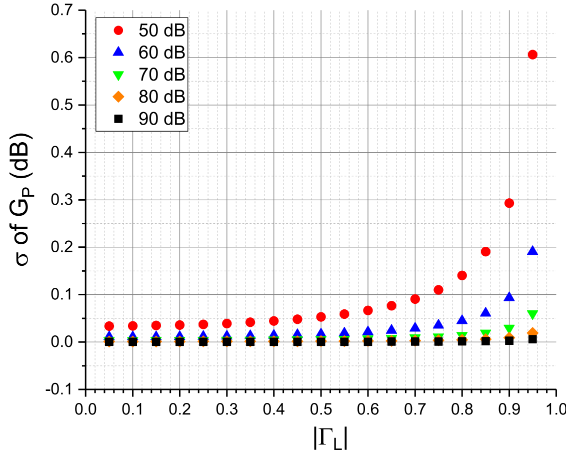

Fig. 8: Simulated standard deviation of power gain for a LP measurement with 10,000 noise iterations, N for each reflection coefficient and dynamic range.

The effect of dynamic range on LP measurements is shown in Fig. 8. It is shown that the standard deviation of calibration error remains marginal, at less than 0.05 dB, for all displayed dynamic range conditions up to $|\Gamma_L| = 0.4$. However, this error becomes significant when measuring highly reflective loads.

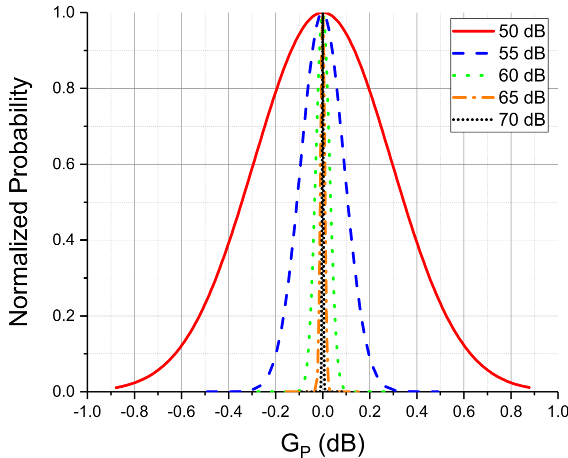

Fig. 9: Simulated normalized probability distribution function of power gain for a LP measurement where the number, N, of calibrations per dynamic range condition was 10,000.

Examining the probability distribution for the case when $|\Gamma_L| = 0.9$, in Fig. 9. Calibration error at $4\,\sigma$ reaches 0.1 dB when dynamic range is degraded to 65 dB.

A practical implementation of these simulations was applied to the active LP measurement system at 30 GHz, as this option offered the greatest tuning range.

The measurement system was configured to use both the measurement and reference receivers, as per Fig. 2. The disadvantage of this configuration is the reduced dynamic range performance of the reference receiver. For the VNA used in this work, the measurement and reference receiver have a dynamic range of 125 and 98 dB, respectively. The VNA was configured to use the reference receivers for incident signal, a, and the measurement receiver for the response signal, b.

To emulate a real-world measurement of a Watt-level device, attenuators were added to achieve 35 dB of attenuation between the DUT and the receivers. Using only the VNA sources for input power, the system's dynamic range for an IF bandwidth of 1 Hz at 30 GHz is approximately 90 dB for the measurement of b_0 and b_3 and 63 dB for a_0 and a_3.

With $N = 100$, a set of calibrations at different IF bandwidths ranging from 1 Hz to 100 Hz was conducted. Following the calibration, a LP measurement was performed on a zero-length thru with an IF bandwidth of 1 Hz with a maximum error in repeatability at $|\Gamma_L| = 0.9$ of less than 0.02 dB. In post-processing, the calibration was removed from the corrected waves and all 100 calibrations were applied to the raw measured data. The standard deviation and probability distribution function was calculated for G_P for each Γ_L condition and IF bandwidth.

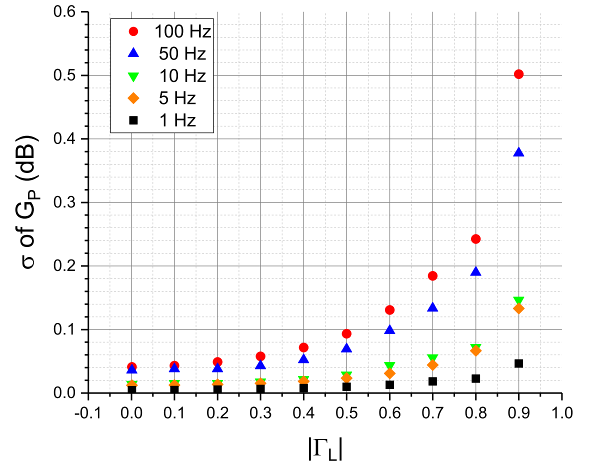

Fig. 10: Measured standard deviation of power gain from a LP measurement, where the number of separate calibrations at each IF bandwidth, $N = 100$. Measured on the active LP system at 30 GHz.

From Fig. 10, the sensitivity of calibration error to system dynamic range is directly reflected, as shown in simulations, with the error rapidly increasing in highly reflective load impedance terminations. For an IF bandwidth of 1 Hz, the standard deviation of error is below 0.05 dB for all Γ_L up to 0.9. As IF bandwidth is increased beyond 1 Hz, the error becomes significant. Considering a single impedance condition where $|\Gamma_L| = 0.9$, in Fig. 11. The power gain error has been normalized to a $\mu = 0$, to isolate the effect on error caused by the dynamic range in calibration and avoid combining the error with the baseline uncertainty in Fig. 6. The maximum error, shown in the probability distribution function, for an IF bandwidth of 10 Hz is 0.5 dB at $3\,\sigma$. Both figures practically demonstrate the findings from the simulations and highlight practical limitations on the measurement accuracy of high power devices.

B. Systematic Error

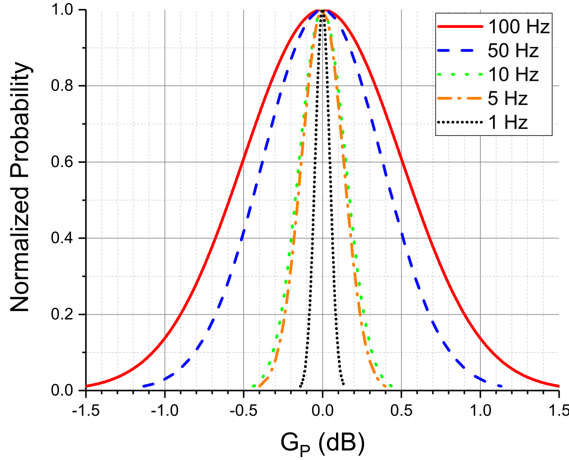

Fig. 11: Measured normalized probability distribution function of power gain, where the number of separate calibrations at each IF bandwidth, N = 100, Measured on the active LP system at 30 GHz with $|\Gamma_{\mathrm{L}}|$ = 0.9.

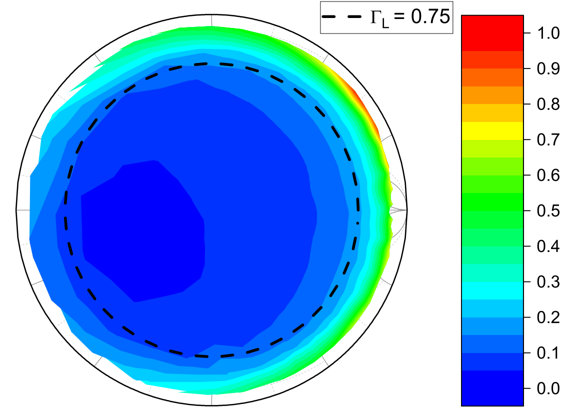

Fig. 12: LP measurement of power gain in decibels for a zero-length thru on the active LP system at 30 GHz. A continuous Γ_{L} = 0.75 circle (dashed line) is marked for measurement system comparison.

This subsection demonstrates the impact that mechanical reconfiguration (if required) of the system after vector calibration has on overall calibration error, with reference to the process flow described in Section II. Mechanical reconfiguration is required in both measurement systems highlighted in this paper. Specifically, the power calibration (step 4) involves an extended port coaxial calibration and connection of power meter to position 3 in Fig. 1. Additionally, these systems require the addition of measurement fixtures (step 5) such as circulators, bias tee and external amplifiers for the measurement condition.

These effects are demonstrated experimentally by isolating the effect caused by mechanical alteration. The random error was limited by removing attenuators between the measurement reference plane and VNA receivers. A 1 Hz IF bandwidth was used, thus optimizing accuracy over measurement speed. Without mechanical alterations, a LP measurement verified that the baseline calibration uncertainty in Fig. 6 and 7 was attained. Subsequently, an absolute calibration was performed and the measurement fixtures connected into the system. Following this, a load-pull was performed to determine any change in calibration error. This was performed across a total of ten calibrations to determine a typical error for both the passive and active measurement systems.

Fig. 12, 13 and 14 illustrate the measurements conducted on both the active and passive measurement systems of a zero-length thru after the mechanical alterations associated with the absolute calibration and the addition of the bias tie, circulator, and external amplifiers. These plots show the standard margin of error resulting from these alterations in both systems.

In each figure, the error profile shows a rapid increase as $|\Gamma_{\mathrm{L}}|$ moves towards the edge of the Smith Chart.

Specifically: Using Fig. 12 and 13, the observed calibration errors across both active and passive measurement systems remain consistent for ka-band, with errors typically around $\pm$0.3 dB at $|\Gamma_{\mathrm{L}}|$ = 0.75. In the active system, the error

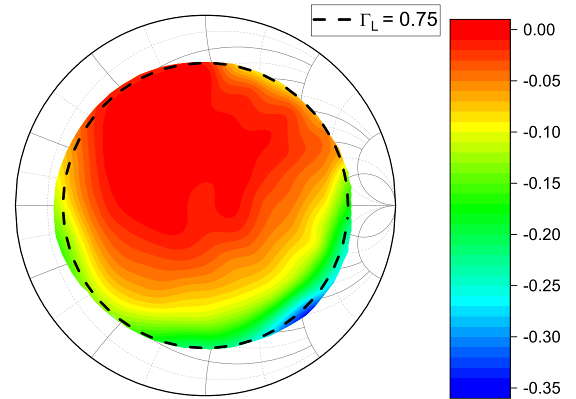

Fig. 13: LP measurement of power gain in decibels for a zero-length thru on the passive LP system at 27.5 GHz. A continuous Γ_{L} = 0.75 circle (dashed line) is marked for measurement system comparison.

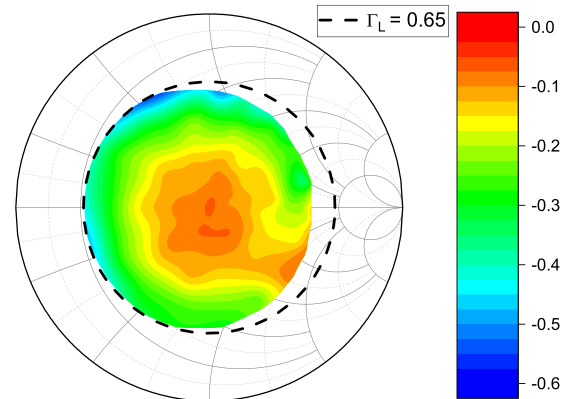

Fig. 14: LP measurement of power gain in decibels for a zero-length thru on the passive LP system at 82.5 GHz. A continuous Γ_{L} = 0.65 circle (dashed line) is marked for measurement system comparison.

significantly increases beyond $|\Gamma_L|$ = 0.8, to a typical maximum error of ± 1 dB at $|\Gamma_L|$ = 0.95, though a direct comparison with passive tuners cannot be provided due to the limited tuning range.

Using Fig. 13 and 14 to compare the error on the passive system changing only frequency, there is an indication of increased error from mechanical alterations as frequency increases. At $|\Gamma_L|$ = 0.65, the typical maximum errors were ± 0.2 dB and ± 0.6 dB for 27.5 GHz and 82.5 GHz, respectively.

These errors stem from mechanical alterations post-vector calibration, from processes such as absolute calibration and the addition of the bias-tee, circulator, and amplifier. In theory, a vector calibration is valid for any change to source/load fixtures. However, despite appearing stable visually, these adjustments introduce minor mechanical perturbations which result in a small change of the network parameters of the path from DUT to VNA receivers, hence leading to an error in the calibration coefficients. Consequently, the small errors on the calibration coefficients are magnified when measuring at a large reflection coefficient.

V. Calibration Technique

The errors discussed in this paper can be avoided if the system can be calibrated in its final measurement state. However, in commercial LP benches, this requirement cannot always be met.

The use of driver amplifiers, circulators and bias tees during the large signal calibration procedure is not common practice due to the limited calibration frequency range (harmonic measurements) and the need for precise control of power at the extended port plane to prevent damage to the power meter during power calibration.

Unless absolute calibration can be performed prior to vector calibration, the system must be reconfigured to perform the extended coaxial calibration and power measurement.

In such cases, this section offers a method to use when the systems exhibits a calibration error during verification: load-pull of a zero-length thru standard. The method updates the vector calibration coefficients of a TRL/mTRL calibration with the system in its final measurement state, mitigating issues of random error and mechanical error.

The methodology is detailed below.

A. Procedure

The procedure follows these steps:

1) Perform large-signal calibration.
2) Configure system into its final measurement state.
3) LP measurement on the thru standard.
4) LP measurement on the line/s standard.
5) De-embed calibration from the LP measurements to get uncorrected measured waves.
6) Perform least squares operation on the uncorrected measured waves to calculate two new "equivalent" s-parameter matricies for the thru and line standards.
7) Re-compute the TRL/ mTRL calibration coefficients using the new thru and line s-parameter matrices and the original reflect s-parameter matrix.

The collection of a large dataset across Γ_L values for the computation of "equivalent" thru and line s-parameter matrices was proposed in [17]. In this approach, the load-pull on the thru and line standards was performed *before* setting up the system in its final configuration, so that the first large-signal calibration already contained this information. This meant that it was applied to a mechanically unperturbed system, hence residual error was already very small, less than 0.06 dB. However, there was no demonstration of the impact of the correction on real measurements, nor an attempt to identify the sources of the residual error.

The method differs significantly in its approach compared to the other correction methods from [8], [9], [10]. These methods involved directly using the measurement of the zero-length thru to optimize the vector calibration coefficients using a non-linear algorithm to minimize the power gain across the Smith Chart. Although effective, there are no constraints on the coefficients and the method optimizes for only a single parameter, the power gain. This can lead to a divergence from the theory of the original calibration algorithm and potentially from physical reality.

In this paper, a two-step calibration solution is proposed. First, the standard calibration 1) followed by a second calibration (3-7) that incorporates the method proposed in [17] when the system is mechanically stable. It is important to note that step 1) must be performed for at least two reasons. Firstly, for on-wafer systems, the absolute power calibration inevitably requires mechanical changes to the system, so it would not make sense to do it after. Secondly, the vector calibration provides an estimate of the corrected Γ_L at the DUT, allowing the LP in steps 3-4) to be performed with a reasonable Smith Chart coverage, limited only in passive LP systems by the maximum achievable Γ_L. Having knowledge of the Smith Chart coverage is beneficial for maximising the accuracy of the technique. In [17], the accuracy improves as the radius of reflection coefficient increases.

B. Method

The method is based on the underpinning condition that the *physical* network parameters of the thru and line standards are independent of Γ_L , while the *measured* network parameters can depend on Γ_L [17]. The method uses a least squares algorithm to generate "equivalent" measured s-parameters for the thru and the line standards, which will have the minimum deviation from the set of Γ_L conditions. In detail, if a number n of different Γ_L conditions are measured, we can re-arrange the corresponding uncorrected waves as:

$$\begin{bmatrix} b_{0,1} & \dots & b_{0,n} \\ b_{3,1} & \dots & b_{3,n} \end{bmatrix} = \begin{bmatrix} S_{Raw} \end{bmatrix} \begin{bmatrix} a_{0,1} & \dots & a_{0,n} \\ a_{3,1} & \dots & a_{3,n} \end{bmatrix} \tag{2}$$

where S_{Raw} is the s-parameter matrix of either the thru or line.

S_{Raw} can be derived by means of least squares method. In particular, the Moore-Penrose pseudo-inverse on A allows for a least squares operation to generate the s-parameter matrix:

$$S_{\mathrm{Raw}} = A^{+}B \tag{3}$$

where

$$A^{+} = (A^{T}A)^{-1}A^{T} \tag{4}$$

Since the system is already calibrated, the uncorrected waves must be generated from the measured waves and the original calibration coefficients. Although this might seem a complication, it is actually a practical approach as most LP systems will provide the calibrated waves and the error box coefficients directly, while accessing the raw waves would require the user to interact with the VNA receivers directly, with a clear difference from model to model.

The large-signal calibration equation that converts uncorrected measured waves at the VNA receivers to corrected waves at the DUT can be written as:

$$\begin{bmatrix} b_1 \\ a_1 \\ a_2 \\ b_2 \end{bmatrix} = e_{10} \begin{bmatrix} [R_{\mathrm{A}}]^{-1} & 0 \\ 0 & [R_{\mathrm{B}}] \end{bmatrix} \begin{bmatrix} b_0 \\ a_0 \\ a_3 \\ b_3 \end{bmatrix}. \tag{5}$$

This contains the absolute calibration scaling term (e_{10}) and the 2x2 wave cascading matrices, R_{A} and R_{B}:

$$R_{\mathrm{A}}^{-1} = \begin{bmatrix} \dfrac{1}{e_{01}e_{10}} & -\dfrac{e_{00}}{e_{01}e_{10}} \\ \dfrac{e_{11}}{e_{01}e_{10}} & -\dfrac{\Delta_{E_{\mathrm{A}}}}{e_{01}e_{10}} \end{bmatrix} \tag{6}$$

$$R_{\mathrm{B}} = \begin{bmatrix} -\dfrac{\Delta_{E_{\mathrm{B}}}}{e_{32}e_{10}} & \dfrac{e_{22}}{e_{32}e_{10}} \\ -\dfrac{e_{33}}{e_{32}e_{10}} & \dfrac{1}{e_{32}e_{10}} \end{bmatrix} \tag{7}$$

which are composed of the s-parameter error box terms e_{xy}, and the determinants ΔE_{A} and ΔE_{B} of the error boxes. In this method, only these terms are modified, while the absolute scaling term, e_{10}, remains the same. Measured waves can be obtained from the known calibrated waves and original error coefficients by inverting the system of (5). The measured waves for the load-pulled thru and line can be used to calculate their respective S_{Raw} using (3). New R_{A}^{-1} and R_{B} matrices can be calculated with the TRL algorithm, and (5) can be applied to any new measurement using the original scaling factor e_{10}.

For the methodology to be self-consistent with the TRL algorithm, it can be proved that LP measurements of both the line and thru standards are required. To verify this, the quality factor for the measured standards [17] will be calculated:

$$Q = det\left(\left[R_{\mathrm{d}}\right]\left[R_{\mathrm{t}}\right]^{-1}\right), \tag{8}$$

and must be ideally unitary. According to the notation in [14], $[R_{\mathrm{d}}]$ is the wave cascading matrix for the raw uncorrected line:

$$\left[R_{\mathrm{d}}\right] = \left[R_{\mathrm{A}}\right]\left[R_{\mathrm{l}}\right]\left[R_{\mathrm{B}}\right] \tag{9}$$

with $[R_{\mathrm{l}}]$ being the cascade matrix of an ideal line, while $[R_{\mathrm{t}}]$ is the cascade matrix of the raw uncorrected thru:

$$\left[R_{\mathrm{t}}\right] = \left[R_{\mathrm{A}}\right]\left[R_{\mathrm{B}}\right]. \tag{10}$$

$[R_{\mathrm{A}}]$ and $[R_{\mathrm{B}}]$ are the error boxes.

C. Application of method

To demonstrate the impact of the proposed correction, the methodology follows as per Section V. A). The active and passive LP systems were configured for 30 GHz and 82.5 GHz measurements, respectively, with 35 dB of attenuation between the DUT and receivers and a calibration IF bandwidth of 1 Hz.

The graphs presented show a comparison of uncertainty metrics between the LP measurement of a zero-length thru using the *original* and *second step* calibration coefficients.

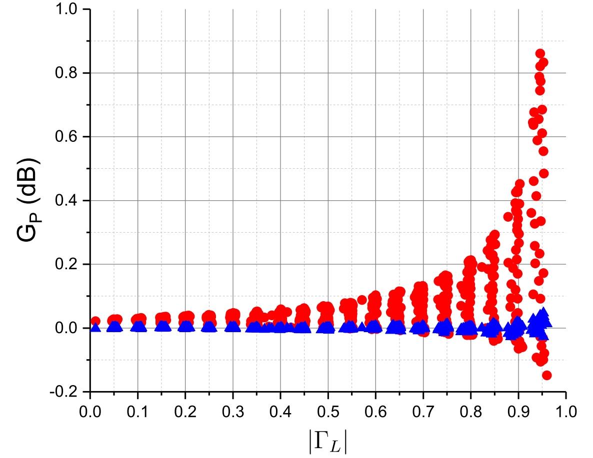

Fig. 15: LP measurement showing calibration error on the power gain of a zero-length thru on the active LP system at 30 GHz using the original (Circle) and second step (Triangle) calibration coefficients.

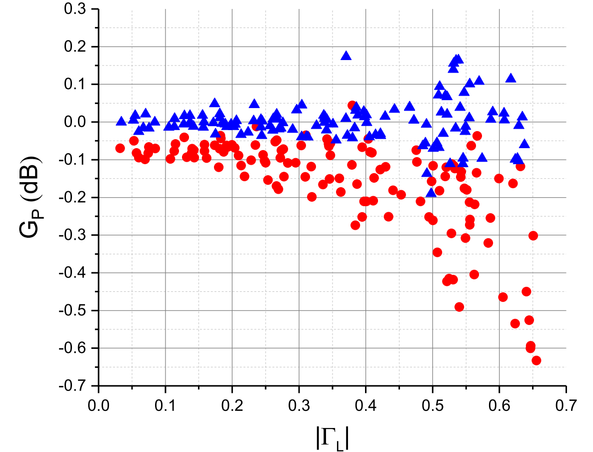

Fig. 16: LP measurement showing calibration error on the power gain of a zero-length thru on the passive LP system at 82.5 GHz using the original (Circle) and second step (Triangle) calibration coefficients.

In Fig. 15 and 16, the comparison of the measured power gain of a zero-length thru with the original and second step calibration coefficients for both measurement systems. It can be seen that the calibration error is reduced to the reference baseline uncertainty of ± 0.1 dB at 30 GHz for the active LP system and ± 0.2 dB at 82.5 GHz for the passive LP system.

The correction methodology is shown to be self-consistent with other known performance metrics, the gain ratio b_2/a_1 and the magnitude and phase ratios of the reflection coefficients (Γ_{L} and Γ_{in}) are shown in Fig. 17, 18 and 19.

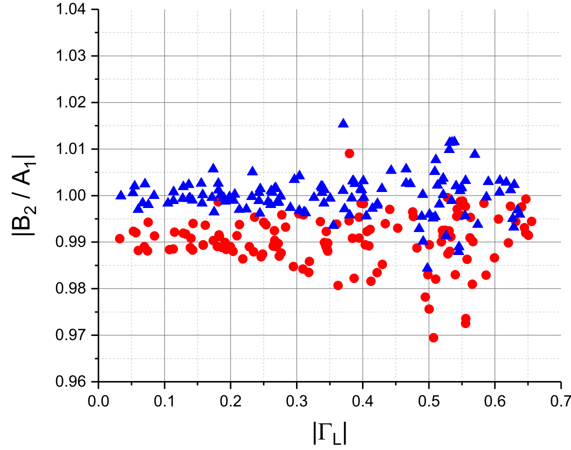

Fig. 17: LP measurement showing calibration error on the gain of a zero-length thru at 82.5 GHz using the original (Circle) and Second step (Triangle) calibration coefficients.

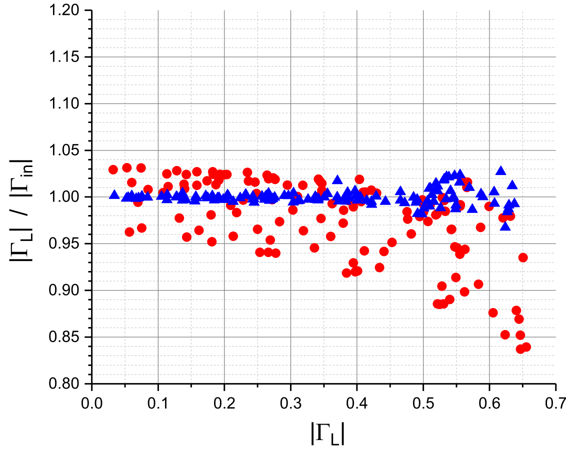

Fig. 18: LP measurement showing calibration error on the magnitude ratio of Γ_L and Γ_in of a zero-length thru at 82.5 GHz using the original (Circle) and second step (Triangle) calibration coefficients.

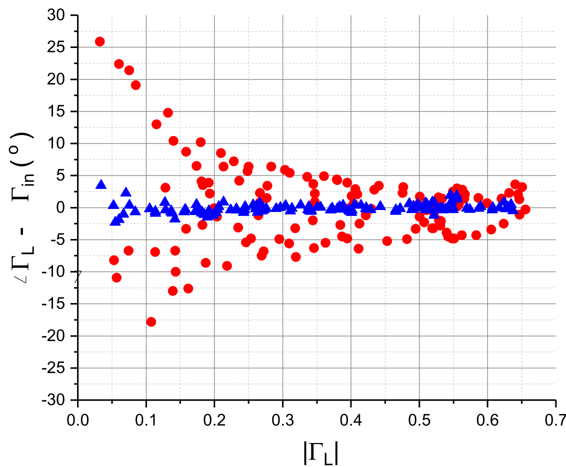

Fig. 19: LP measurement showing calibration error on the phase of Γ_L and Γ_in of a zero-length thru at 82.5 GHz using the original (Circle) and second step (Triangle) calibration coefficients.

There is a reduction in calibration error from 3 % to 2 % in gain, 17 % to 4 % in reflection coefficient magnitude ratio and 26 ° to 5 ° in the phase difference between reflection coefficients.

The quality factor from Eq. 8 is applied to this example in Table I which shows the necessity to perform LP measurements on both thru and line standards in this method. The exclusion of the line standard does not affect the relative calibration error in G_P presented in this section; its effect would manifest as an error in the reference impedance of the calibration.

TABLE I: Calibration Quality Factor

Standards	Quality Factor
Ideal	1 + j0
Thru	1.042 - j0.0004
Thru + Line	0.998 + j0.008

VI. Device Measurement

This section applies the correction methodology to the measurement of mm-wave devices for both systems and illustrates the effect on relative metrics (PAE and G_P) used to quantify a device's large-signal performance. The calibration was performed on a TRL calibration kit on the same wafer as the DUT, with the reference plane at the centre of the thru standard, to avoid de-embedding of feed lines when evaluating device performance at the gate / drain manifold [24], [25]..

A. Active LP System - Ka-band GaN HEMT

For the active LP system, the LP characterization of a Qorvo 0.15 µm GaN HEMT performed at 30 GHz. The calibration error, as shown by a measurement of a zero-length thru and the subsequent correction is shown in Fig. 15. The measurement of the active device was performed with the original calibration coefficients and the second step calibration coefficients were applied in post-process.

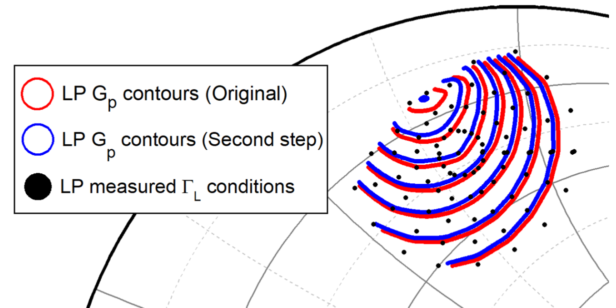

Fig. 20: LP measurement of the Qorvo 0.15 µm GaN HEMT at 30 GHz with V_D = 3.5 V, I_DQ = 50 mA/mm and P_Av = 24.3 dBm using the active LP system. Showing a comparison of G_P using the original (red) and second step (blue) calibration coefficients. Maximum gain contour is 11 dB with 0.5 dB contours.

In Fig. 20, the G_P contour show a negligible change in the position of the contours. The power sweep, in Fig. 21, at the optimum for a power match show the calibration error resulted in an overestimation of the power gain and PAE by up to 0.12 dB and 1.5 % respectively.

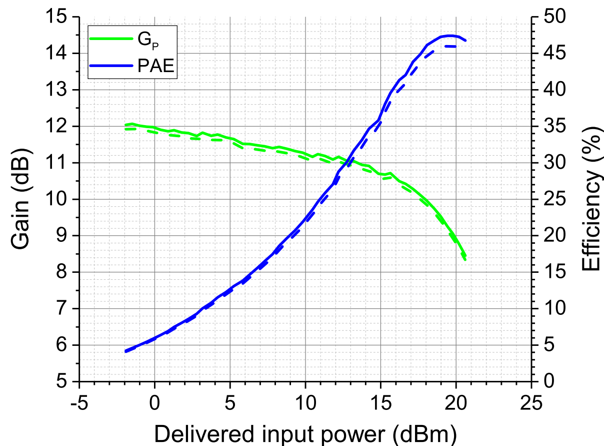

Fig. 21: Power sweep measurement of G_P and PAE using the original (Solid Line) and second step (Dotted Line) calibration coefficients for a GaN HEMT at 30 GHz with V_D = 24 V, I_DQ = 50 mA/mm at the optimum load impedance for a power match of $\Gamma_\mathrm{L} = 0.8\angle 93\,°$.

B. Passive LP system - E-band GaAs pHEMT

Measurement of Qorvo 90 nm GaAs pHEMT at 82.5 GHz on the passive LP system. The calibration error as shown by a measurement of a zero-length thru and the subsequent correction is shown in Fig. 16. The measurement of the active device was performed with the original calibration coefficients and the second step calibration coefficients were applied in post-process. The LP measurement of the GaAs device in a backed-off state, with the device operating in the linear region, can be compared to the gain circles derived from a small signal measurements of the device in the same operating condition, using the original and second step calibration coefficients to evaluate the validity of the correction method.

The gain circles are derived from an s-parameter measurement of the same device using a mTRL calibration [16] with the device biased in the same operating condition as the LP measurement.

In an ideal scenario, the LP gain contours should overlap perfectly with the gain circles.

Fig. 22 shows the translation of s-parameter derived gain circles and the measured gain contours from LP in a backed-off condition (available input power of -2.5 dBm, compression is at approximately 20 dBm). It shows a good match between the measurement using the second step calibration coefficients. The circularity of the outer contours is effected by interpolation between measured LP conditions. Using the original calibration coefficients, the contours have an approximate error of 0.25 dB.

In Fig. 23, the power gain contours from a LP measurement of the GaAs device. Closed contours were not possible due to the maximum $|\Gamma_\mathrm{L}| = 0.7$ of the passive tuners at this frequency. The comparison shows a small change in the contour shape and an approximate offset in contours by 0.25 dB.

From the power sweep measurement in Fig.24, the load impedance was tuned to the optimal condition for a power match realizable with passive tuners. It can be seen that the

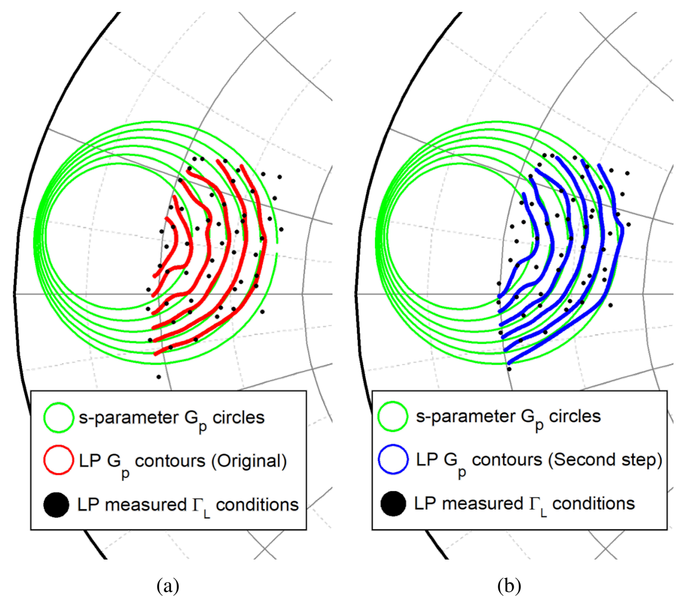

Fig. 22: LP measurement in a backed-off state showing a comparison of G_P and small signal power gain circles (green) for a.) original and b.) second step calibration coefficients. The device was the Qorvo 90 nm GaAs pHEMT measured at 82.5 GHz with V_D = 3.5 V, I_DQ = 50 mA/mm and P_Av = -2.5 dBm using the passive LP system. Maximum gain contour is 1.75 dB with 0.25 dB contours.

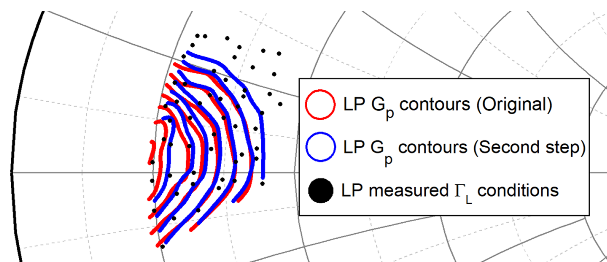

Fig. 23: LP measurement of the Qorvo 90 nm GaAs pHEMT at 82.5 GHz with V_D = 3.5 V, I_DQ = 50 mA/mm and P_Av = 20.2 dBm using the passive LP system. Showing a comparison of G_P between the original (red) and second step (blue) calibration coefficients. Maximum gain contour is 1.5 dB with 0.25 dB contours.

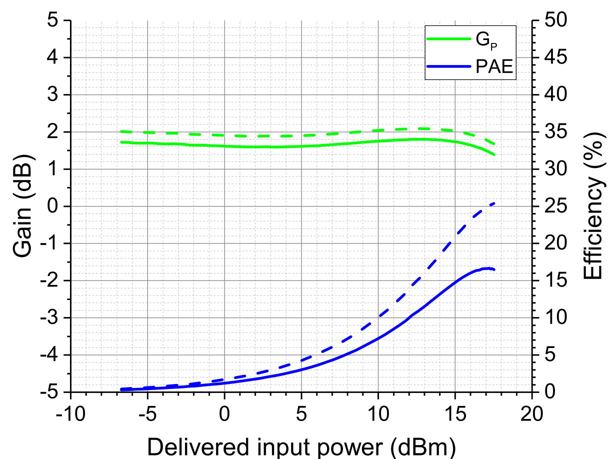

Fig. 24: Power sweep measurement of G_P and PAE using the original (Solid Line) and second step (Dotted Line) calibration coefficients for a GaAs pHEMT at 82.5 GHz with V_D = 3.5 V, I_DQ = 50 mA/mm at $\Gamma_\mathrm{L} = 0.7\angle 179\,°$

calibration error resulted in an underestimation of the power gain and PAE by up to 0.3 dB and 8.9 % respectively.

C. Error correction analysis

The calibration error observed in the measurement of a zero-length thru does not directly translate to the measurement of an active device. To provide a complete picture, the extent of correction due to calibration error is compared in both cases.

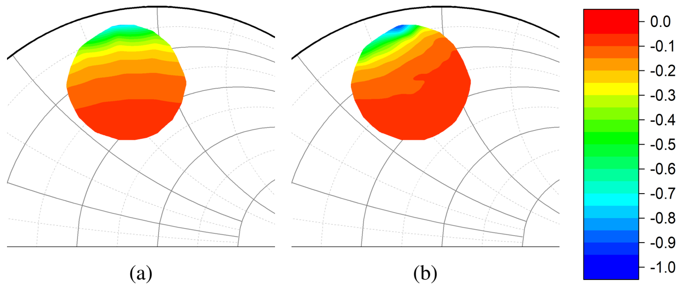

Fig. 25: LP measurement showing the power gain using the original and second step calibration coefficients for a measurement of the a.) thru standard and b.) GaN device at 30 GHz. 0.1 dB contours.

Fig. 25 illustrates the difference between the power gain measured on a DUT with the original and second step calibration coefficients. It can be seen that the contours differ in size and shape. The explanation for this load dependence of correction can be reasonably concluded using Eq. 1 as reference by considering three measurement cases.

- Zero loss DUT: $\Gamma_{\mathrm{in}} = \Gamma_{\mathrm{L}}$, the error associated with both parameters becomes increasingly sensitive as the load reflection coefficient moves towards the edge of the Smith Chart. This maximizes the worst-case error.
- Lossy DUT: $\Gamma_{\mathrm{in}} < \Gamma_{\mathrm{L}}$, the effective impedance coverage on the input side of the device is reduced, hence the maximum error is also reduced. In the case where Γ_{L} approaches the edge of the Smith Chart, the error is predominately influenced by Γ_{L}.
- Active DUT: The feedback is normally low. This results in a weak correlation between Γ_{in} and Γ_{L}. Consequenctly, the error depends on the position of the DUT's Γ_{in}.

These cases help illustrate how the nature of the DUT affects the amount of correction on measured DUT quantities.

VII. Conclusion

In this paper, we have presented a practical approach to evaluating the primary sources of error in the calibration and LP measurement of mm-wave devices at high reflection coefficients. The most significant sources of error identified include random error in vector calibration related to the system dynamic range and systematic errors arising from mechanical alterations to the measurement set-up post-vector calibration. It has been demonstrated that the systematic errors are particularly challenging to avoid in systems that are unable to be calibrated in their final measurement configuration.

We have proposed a two step calibration procedure that consists of a least squares correction methodology consistent with the TRL algorithm based on the LP measurement of thru and line calibration standards when the system is configured in its final measurement state. The method was applied to the measurement of GaAs and GaN devices on both active and passive load-pull systems in Ka and E-band showing an error correction of up to 0.3 dB in power gain and 8.9 % in power added efficiency.

References

[1] I. Angelov, M. Thorsell, D. Kuylenstierna, G. Avolio, D. Schreurs, A. Raffo, and G. Vannini, "Hybrid Measurement-based Extraction of Consistent Large-signal Models for Microwave FETs," in *2013 European Microwave Conference*, 2013, pp. 267–270.

[2] H. Qi, J. Benedikt, and P. J. Tasker, "Nonlinear Data Utilization: From Direct Data Lookup to Behavioral Modeling," *IEEE Transactions on Microwave Theory and Techniques*, vol. 57, no. 6, pp. 1425–1432, 2009.

[3] H. Wang *et al.* "Power Amplifiers Performance Study 2000 - Present". [Online]. Available: https://ideas.ethz.ch/research/surveys/pa-survey.html

[4] *Fixed Radio Systems; Characteristics and Requirements for Point-to-Point Equipment and Antennas*, Std. ETSI EN 302 217-2 V3.3.0 , 2021.

[5] B. Cimbili, C. Friesicke, F. Van Raay, S. Wagner, M. Bao, and R. Quay, "High-Efficiency Watt-Level E-band GaN Power Amplifier with a Compact Low-loss Combiner," in *2023 IEEE Topical Conference on RF/Microwave Power Amplifiers for Radio and Wireless Applications*, 2023, pp. 42–44.

[6] E. Ture, S. Leone, P. Brückner, R. Quay, and O. Ambacher, "High-Power (>2 W) E-Band PA MMIC Based on High Efficiency GaN-HEMTs with Optimized Buffer," in *2019 IEEE MTT-S International Microwave Symposium (IMS)*, 2019, pp. 1407–1410.

[7] A. Ferrero, V. Teppati, and A. Carullo, "Accuracy Evaluation of On-Wafer Load-Pull Measurements," *IEEE Transactions on Microwave Theory and Techniques*, vol. 49, no. 1, pp. 39–43, 2001.

[8] V. Teppati, A. Ferrero, D. Parena, and U. Pisani, "Accuracy Improvement of Real-Time Load-Pull Measurements," *IEEE Transactions on Instrumentation and Measurement*, vol. 56, no. 2, pp. 610–613, 2007.

[9] S. Bonino, V. Teppati, and A. Ferrero, "Further Improvements in Real-Time Load-Pull Measurement Accuracy," *IEEE Microwave and Wireless Components Letters*, vol. 20, no. 2, pp. 121–123, 2010.

[10] V. Teppati and C. R. Bolognesi, "Evaluation and Reduction of Calibration Residual Uncertainty in Load-Pull Measurements at Millimeter-Wave Frequencies," *IEEE Transactions on Instrumentation and Measurement*, vol. 61, no. 3, pp. 817–822, 2012.

[11] K. Łukasik, J. Cheron, G. Avolio, A. Lewandowski, D. F. Williams, W. Wiatr, and D. M. M.-P. Schreurs, "Uncertainty in Large-Signal Measurements Under Variable Load Conditions," *IEEE Transactions on Microwave Theory and Techniques*, vol. 68, no. 8, pp. 3532–3546, 2020.

[12] D. Williams. NIST Microwave Uncertainty Framework. [Online]. Available: https://www.nist.gov/servicesresources/software/wafer-calibration-software

[13] A. Baddeley, S. Woodington, D. Gecan, A. Sheikh, J. Lunn, P. Tasker, and R. Quaglia, "Millimeter-Wave On-Wafer Large Signal Characterization System for Harmonic Source/Load Pull and Waveform Measurements," in *2023 IEEE/MTT-S International Microwave Symposium - IMS 2023*, 2023, pp. 1101–1104.

[14] G. Engen and C. Hoer, "Thru-Reflect-Line: An Improved Technique for Calibrating the Dual Six-Port Automatic Network Analyzer," *IEEE Transactions on Microwave Theory and Techniques*, vol. 27, no. 12, pp. 987–993, 1979.

[15] S.-H. Shin, N. Ridler, A. Tucker, T. Williams, and X. Shang, "Comparison of On-Wafer Calibrations for Measurements of Active and Passive Devices at Millimeter-Wave Frequencies," 11 2023.

[16] R. Marks, "A Multiline Method of Network Analyzer Calibration," *IEEE Transactions on Microwave Theory and Techniques*, vol. 39, no. 7, pp. 1205–1215, 1991.

[17] A. Aldoumani, T. Williams, J. Lees, and P. J. Tasker, "Enhanced Vector Calibration of Load-Pull Measurement Systems," in *83rd ARFTG Microwave Measurement Conference*, 2014, pp. 1–4.

[18] A. C. Stelson, A. M. Hagerstrom, J. A. Jargon, and C. J. Long, "Quantifying Receiver Nonlinearities in VNA Measurements for the WR-15 Waveguide Band," *IEEE Transactions on Microwave Theory and Techniques*, vol. 70, no. 5, pp. 2743–2749, 2022.

[19] A. Lewandowski and D. Williams, "Characterization and Modeling of Random Vector Network Analyzer Measurement Errors," in *MIKON 2008 - 17th International Conference on Microwaves, Radar and Wireless Communications*, 2008, pp. 1–4.

[20] N. Estella, E. Camargo, J. Schellenberg, and L. Bui, "High-Efficiency, Ka-band GaN Power Amplifiers," in *2019 IEEE MTT-S International Microwave Symposium (IMS)*, 2019, pp. 568–571.

[21] A. Bessemoulin, J. Tarazi, M. Rodriguez, M. G. McCulloch, A. E. Parker, and S. J. Mahon, "Reduced-size E-band GaAs power amplifier MMIC," in *2015 10th European Microwave Integrated Circuits Conference (EuMIC)*, 2015, pp. 25–28.

[22] A. Ferrero, "High power Load Pull Measurements: Today Technologies and Tomorrow Challenges," in *78th ARFTG Microwave Measurement Conference*, 2011, pp. 1–3.

[23] Rohde and Schwarz, *ZVA Vector Network Analyzer Specifications*, October 2020. [Online]. Available: https://www.rohde-schwarz.com/brochure-datasheet/zva

[24] A. Rumiantsev, P. Sakalas, F. Pourchon, P. Chevalier, N. Derrier, and M. Schroter, "Application of On-Wafer Calibration Techniques for Advanced High-Speed BiCMOS Technology," in *2010 IEEE Bipolar/BiCMOS Circuits and Technology Meeting (BCTM)*, 2010, pp. 98–101.

[25] K. Yau, E. Dacquay, I. Sarkas, and S. P. Voinigescu, "Device and IC Characterization Above 100 GHz," *IEEE Microwave Magazine*, vol. 13, no. 1, pp. 30–54, 2012.