

4 Evaluating Uncertainty in Vector Network Analyser Measurements

4.1 Introduction

The dramatic growth of radio-based devices and applications over the last 50 years has led to the VNA becoming a critical instrument in most laboratories. Many of these applications required both accurate and reliable measurements from these VNAs. This is particularly so in areas such as manufacturing, calibration and testing. This is often driven by requirements given in international Quality Management documents such as the ISO 9000 series of standards [1] (for manufacturing and process control) and the ISO 17025 standard [2] (for calibration and testing).

The requirements given in these international standards are for measurements that can be demonstrated as fit-for-purpose (in terms of the achievable level of accuracy, etc) and made traceable to the international system of units [3], [4]. These requirements are not trivial for a VNA due to the complicated nature of the VNAs operating principles, for example the calibration mathematics. Combined with the available computing power and cost at that time, a full or rigorous evaluation of uncertainty for VNA measurements was difficult and time-consuming. This led to much work by experts in this field to develop easier methods that addressed these needs in ways that were suitable for use by end-users in the manufacturing, calibration and testing communities. Much of this work was undertaken by the ANAMET Technology Group (www.npl.co.uk/anamet) during the 1990s. This resulted in a series of reports [5]–[7] describing the development of a guidance document that gave a procedure for assessing the performance of calibrated VNAs. The resulting guidance document [8] was published by the European co-operation for Accreditation (EA, www.european-accreditation.org) so that laboratories operating to the ISO 17025 standard and/or ISO 9000 series of standards could implement the method

for their own purposes. Ownership of this EA document was later transferred to the European Association of National Metrology Institutes (EURAMET) and re-published [9] as part of their Calibration Guides series of documents. This document, along with the recent updated version, is available as a free download from the EURAMET web-site: www.euramet.org.

In addition to the EURAMET guide, VNA manufacturers have also produced their own advice for users to estimate the combined standard uncertainty in their measurements [10] and provided software tools in some cases [11], [12]. Often this advice is based on the same methods presented in the EURAMET guide, which work by evaluating the residual error in VNA calibrations.

4.2 Evaluating Residual Error in VNA Calibrations

4.2.1 Method

The EURAMET Guide [9] presents a process for evaluating the uncertainty of measurements performed on a calibrated VNA, allowing users to verify that values measured using the instrument are of acceptable accuracy. This process involves measuring a selection of dominant contributions to measurement uncertainty and combining them appropriately. Contributions include both systematic errors, which remain constant over the period of measurements, and random errors, which do not. The error model for voltage reflection coefficient (VRC) measurements performed with a VNA is represented in [9] by the following equations for one-port (1) and two-port (2) measurements:

$$U_{\text{VRC}} = D + T\Gamma + M\Gamma^2 + R_{\text{VRC}} \quad (4.1)$$

$$U_{\text{VRC}} = D + T\Gamma + M\Gamma^2 + R_{\text{VRC}} + S_{21}^2\Gamma_L \quad (4.2)$$

where U_{VRC} is the combined uncertainty in the measurement, D is the residual directivity, T represents the effect of tracking and nonlinearity, M is the residual test port match (TPM), Γ is the measured voltage reflection coefficient, R_{VRC} is the sum of all the random contributions, S_{21} is the nominal attenuation of the device-under-test (DUT), and Γ_L is the residual load match. The most significant systematic error contributors to the measurement uncertainty are, in most cases, the directivity and TPM.

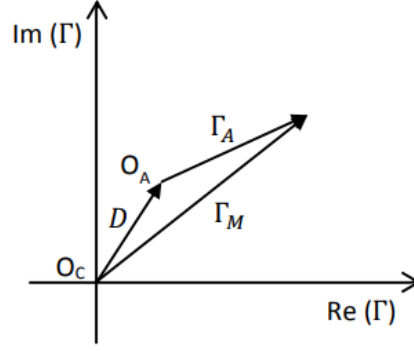
To perform a VRC measurement, a VNA must be able to separate reflected and incident voltage waves by their direction of travel and then sample using complex receivers. However,

various components in the signal path may cause a portion of the incident wave to leak into the reflected wave receiver without having reached the DUT. This directivity error should ideally be removed by applying correction terms extracted during the VNA calibration. However, as no calibration will be perfect, some residual directivity error will remain (referred to as effective directivity in [9]). To measure the residual directivity, a matched load can be connected to the test port being assessed. This should theoretically reflect no amount of the incident wave and therefore the only voltage present at the reflected wave receiver should be due to the residual directivity. In practice, the matched load will not be perfectly matched (for the same reason the calibration will never be perfect) so it is likely the residual directivity will typically be either over- or underestimated. An improved method, used in [9] and widely accepted for use with coaxial measurements, is the ripple extraction technique. This uses a similar principle to measure the residual directivity, but significantly increases the accuracy of the residual error evaluation. An illustration of its method is provided in Figure 4.1.

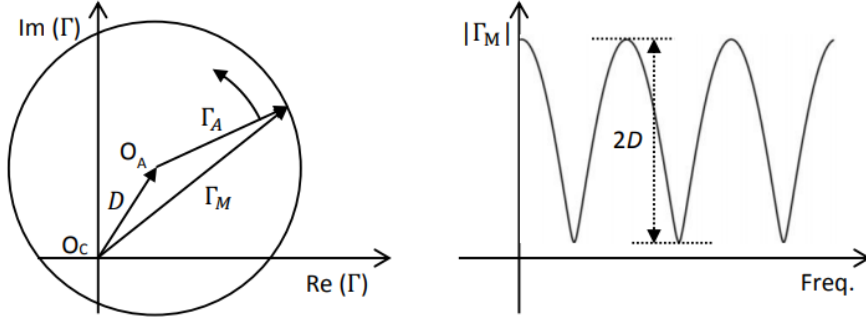
To perform the evaluation, a short length of line is connected to the test port, to the end of which is added a matched load covering the frequency range under test. The line section should be traceable to national standards and have a characteristic impedance identical to that of the VNA setup. For these reasons a beadless airline is suggested in [9]. The load can be either the same as that used for calibration or another with $0.1 \geq \Gamma \geq 0.2$ to ensure that $\Gamma \geq D$. If $\Gamma < D$, then the measured residual directivity will be underestimated as explained by Figure 4.1. If the calibration matched load is used for the measurement, the small reflection from a second connection and any loss in the airline will cause the VRC to be greater than the residual directivity from the original measurement of the load. Alternatively, because $\Gamma < 0.1$ for the matched load used for calibration, using another load with a known higher Γ ensures that there is no underestimate. Once the instrument has been configured, Γ is measured and the magnitude plotted against frequency using a linear scale. A ripple should be clearly visible on the trace, from which the residual directivity can be calculated from:

$$D = \frac{\text{Maximum Ripple Amplitude}}{2} \quad (4.3)$$

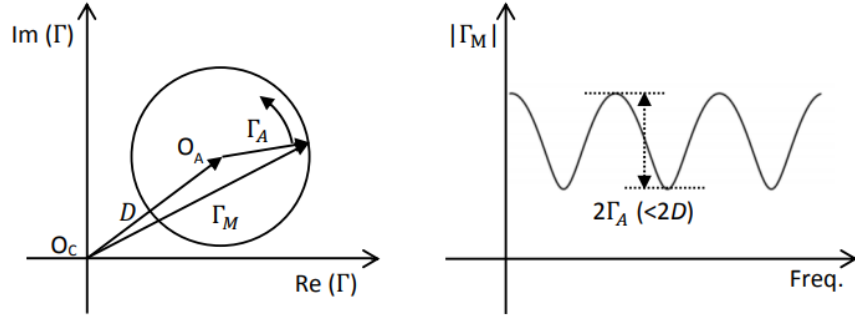
For coaxial measurements as specified in [9], there is a high probability that the condition required to avoid underestimation is met. However, in order to assess the suitability of the technique in waveguide a method of assessing this condition has been used. By examining either a complex plot (polar or Smith chart) or a phase plot, the geometric symptom shown in Figure 4.1 can be identified. When using a complex plot, the origin should lie within the circumference of



(a) When measured on the calibrated VNA, the VRC of a perfect matched load would reveal the true origin (O_T) of the plot as offset from the calibrated origin (O_C) by the residual directivity (D). If a realistic matched load offset by a line section is instead measured, the VRC as measured by the VNA (Γ_M) will be the sum of the residual directivity and the true VRC (Γ_A).



(b) As Γ_M is measured across a swept frequency range, the electrical length of the line increases causing the phase to sweep also. This rotates Γ_T , resulting in ripples in the plot of $|\Gamma_M|$ against frequency. The magnitude of the ripples is equal to $2D$.



(c) However, if $\Gamma_A < D$, then the ripple magnitude is now $2\Gamma_A$ instead of $2\Gamma_D$ and the residual directivity as evaluated using the ripple extraction technique would be underestimated.

Figure 4.1: An illustration of the ripple technique measuring residual directivity.

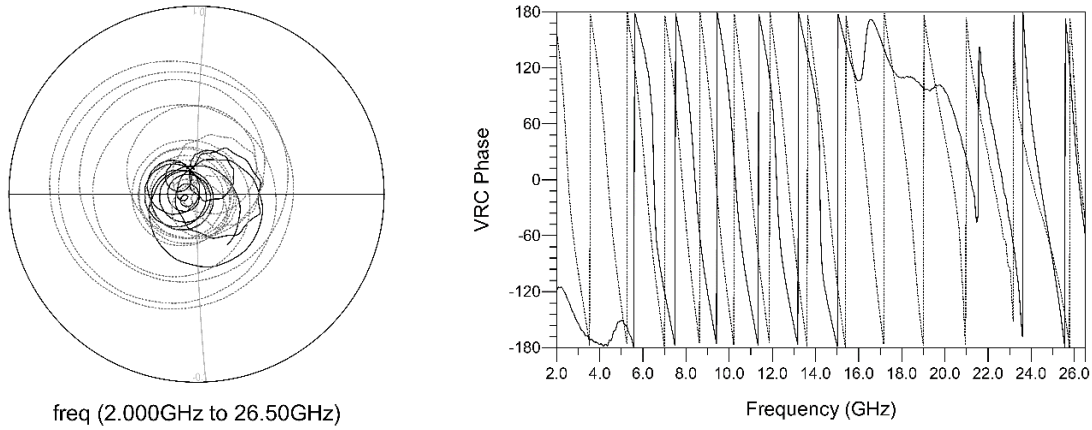


Figure 4.2: Smith chart and phase plot of a residual directivity measurement performed in coaxial transmission line using the calibration load (solid line) and a load from a different calibration kit (dotted line). The intersection of the grid lines on the Smith chart represents the origin of the plot ($\Gamma = 0$). The measurement which used the different load indicate an accurate residual directivity estimate across the entire measured spectrum. The measurement which used the calibration load suggests an underestimate of the residual directivity is likely between 16–22 GHz.

the reflection coefficient trace for the residual error to be accurately measured. For any frequency range where it does not, the ripple technique provides an underestimation of the residual error. When using a phase plot, there will be regular wrapping of the reflection coefficient phase for frequency ranges where the residual error is correctly measured, whereas when underestimation occurs the phase will vary by less than 180 degrees per period. An example of both plots are shown in Figure 4.2. Either of these methods can be used to identify when a calibration and ripple measurement needs to be repeated. If the repeat measurements still fail the test, then the choice of loads may need to be altered.

TPM is caused by imperfections in the impedance match between components in the VNA setup. This causes delayed reflections that interfere with the DUT measurement and can provide false values. Calibration also corrects for TPM, but as with directivity some residual error will remain. To measure residual TPM, a short circuit is connected to the test port being assessed. This should reflect the entire incident signal and therefore maximise any reflections in the VNA setup. If residual TPM error is present then the measured VRC will be less than 1. However, the short circuit may not provide a perfect reflection and so the ripple extraction technique is

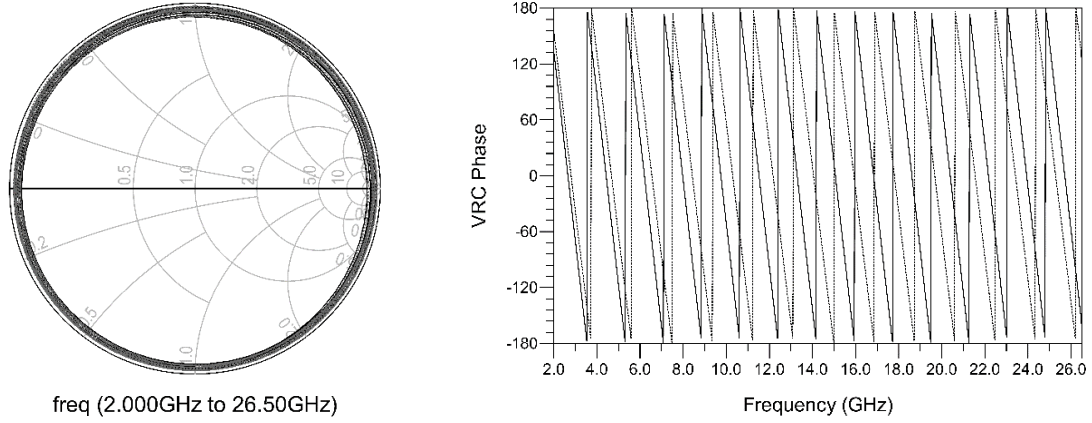


Figure 4.3: Smith chart and phase plot of a residual TPM measurement performed in coaxial transmission line using the calibration short circuit (solid line) and a short circuit from a different calibration kit (dotted line). Both measurements indicate an accurate residual TPM estimate across the entire measured spectrum. Residual TPM measurements avoid the risk of underestimation via the mechanism described for residual directivity measurements due to the typical VRC being close to one (far from the origin of the Smith chart).

favoured for this measurement also.

To measure residual TPM using the ripple extraction technique, the same procedure as for residual directivity is followed but the matched load at the end of the line is replaced by a short circuit. A similar plot is acquired and the residual TPM is given by:

$$M = \frac{\text{Maximum Ripple Amplitude}}{2} \quad (4.4)$$

Because the reflection coefficient for this measurement should be close to 1, there is no risk that this value will be greater than the true residual TPM and cause an underestimate as can be the case for residual directivity. This is shown in Figure 4.3, where the origin of the Smith chart is clearly inside the circular trace and the phase consistently wraps across the entire measured spectrum.

To perform the technique for both described residual error sources requires just three components: A short circuit, matched load, and a short section of line. These components are realizable in both coaxial and rectangular waveguide, so the technique should be physically possible to perform in waveguide setups. The technique can be applied to any type of calibration for example

three-known-loads and thru-reflect-line (TRL). In coaxial, the short-open-load-thru (SOLT) variant of the former is used. However, in waveguide an open circuit is not straightforward to realise or widely adopted, so a common variant of SOLT calibration which uses an offset short (SOSLT) can be used instead.

4.2.2 Investigation of ripple technique effectiveness with waveguide VNA calibrations

Although both the EURAMET guide and VNAs using rectangular metal waveguide test ports have both existed for many years, there was no published evidence that the methods employed in the guide (the ripple technique) could be successfully applied to those measurements. The author undertook an investigation into this during the degree where they compared residual error measurements in coaxial line to those in waveguides at frequencies ranging from 8.2 GHz to 750 GHz (submillimetre-wave). This section presents the results of the subsequent papers [13], [14].

4.2.2.1 Coaxial Transmission Line and Microwave Waveguide

The ripple extraction technique was first performed on coaxial line in accordance with the EURAMET Guide [9] instruction as described in the previous section. The guide itself provides a range of typical values for both residual directivity and TPM ripple measurements with which our results can be compared to ensure that the process was followed correctly. All measurements presented in this investigation were acquired using a Keysight 5247A PNA-X instrument fitted with 1.85 mm ports attached to flexible port extender cables with rugged connectors. The coaxial measurement setup used a 7.5 cm beadless airline and the calibration kit matched load and short circuit. Figure 4.4 shows the ripple trace obtained by plotting VRC magnitude against frequency for a residual directivity measurement on both ports using a SOLT calibration. Figure 4.5 shows a similar ripple measured with a short circuit instead of a matched load, which is caused by the residual TPM. The results of the measurements, along with the expected ranges provided in [9], are presented in Table 4.1.

It can be seen that the measured values for the coaxial line setup fall well within the typical ranges. The same method was applied to two types of centimetre band waveguide, WG16 (WR90) and WG20 (WR42). These waveguides have usable frequency ranges of 8.212.4 GHz and 18.026.5 GHz respectively. In order to avoid the effects of non-propagating (evanescent) modes created by the waveguide test adapter, a section of standard line with length equal to several times the

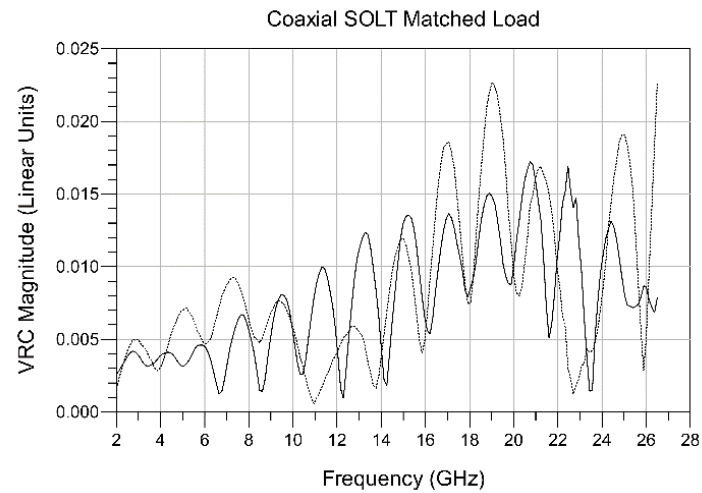


Figure 4.4: Magnitude of the reflection measurement of a matched load offset by a short line length in coaxial media. The VNA was calibrated using the SOLT method, and measurements were performed on both port 1 (solid line) and port 2 (dotted line). Variations apart from the dominant ripple are due to the imperfect response of the matched load and the beadless line.

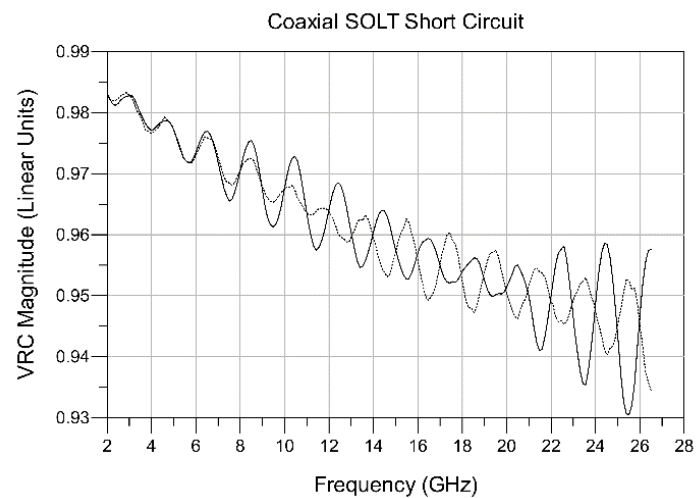


Figure 4.5: Magnitude of the reflection measurement of a short circuit offset by a short line length in coaxial transmission line. The VNA was calibrated using the SOLT method, and measurements were performed on both port 1 (solid line) and port 2 (dotted line). The increasing attenuation with frequency is characteristic of the beadless airline.

	Port 1	Port 2	Port 1	Port 2
	Residual Directivity	Residual Directivity	Residual TPM	Residual TPM
SOLT	0.0075	0.0085	0.0135	0.0100
TRL	0.0075	0.0070	0.0015	0.0015
cg-12 Reference	0.0020.02	0.0020.02	0.0050.02	0.0050.02

Table 4.1: Residual directivity and TPM values of 3.5mm coaxial line VNA calibrations as measured by the ripple extraction method. Both SOLT and TRL calibration techniques were measured.

	Port 1	Port 2	Port 1	Port 2
	Residual Directivity	Residual Directivity	Residual TPM	Residual TPM
WG16 SOSLT	0.0045	0.0040	0.0070	0.0060
WG16 TRL	0.0030	0.0060	0.0020	0.0015
WG20 SOSLT	0.0030	0.0040	0.0095	0.0050
WG20 TRL	0.0020	0.0020	0.0015	0.0010

Table 4.2: Residual directivity and TPM values of WG16 and WG20 waveguide VNA calibrations as measured by the ripple extraction method. Both SOSLT and TRL calibration techniques were measured.

aperture width was attached to each adapter where possible and the measurement planes defined at the end of the lines. The results of the ripple extraction are shown in Table 4.2:

4.2.2.2 Millimetre-wave Waveguides

To perform measurements above the 67 GHz upper limit of the VNA used for the investigation, a range of VDI frequency extender heads were attached to the instrument. These instruments include on their test port a line section of suitable length to allow the dissipation of evanescent modes. To study the performance of the ripple extraction technique at millimetre wavelengths, WG25 (WR15) and WG30 (WR05) waveguides were chosen. These waveguides have operating frequency ranges of 5075 GHz and 140220 GHz respectively. The results of the ripple extraction are shown in Table 4.3:

	Port 1	Port 2	Port 1	Port 2
	Residual Directivity	Residual Directivity	Residual TPM	Residual TPM
WG25 SOSLT	0.0015	0.0015	0.0175	0.0165
WG25 TRL	0.0015	0.0020	0.0090	0.0070
WG30 SOSLT	0.0075	0.0090	0.0185	0.0235
WG30 TRL	0.0070	0.0080	0.0210	0.0150

Table 4.3: Residual directivity and TPM values of WG25 and WG30 waveguide VNA calibrations as measured by the ripple extraction method. Both SOSLT and TRL calibration techniques were measured.

4.2.2.3 Submillimetre-wave Waveguides

The final part of the investigation studied the ripple extraction technique when applied to sub-millimetre wavelength VNA setups. The waveguide chosen for these measurements was in the 500750 GHz band (WR1.5) for which only one frequency extender head was available. Therefore, the TRL calibration was omitted and only port one was measured. The line section used was part of the VDI WR1.5 calibration and verification kit and was 2.54 mm (1 inch) in length. Figure 4.6 shows the setup used during the measurement of residual directivity, with a line and match connected to the frequency extender head. The results from the ripple measurements in this waveguide are presented in Table 4.4. The large difference in residual TPM error between calibrations suggests that there is poor connection repeatability between the waveguide components, an issue which is explored further in the following section. This problem also created difficulty in accurately measuring the residual directivity error, as studying the VRC phase revealed that there was no wrapping present and on average only 90 degree total phase deviation. Figure 4.7 shows the result of this measurement, which suggests that the residual directivity will be underestimated using the ripple extraction technique with this data. The calibration kit included two sets of standards, so the load was swapped but did not improve the problem. Rotation of the load about 180 degrees, which is possible with the UG-387 flange design, also did not increase the phase deviation. The same issue was found when the measurements were performed using the second calibration, which used the entire second set of calibration standards in the kit. This behaviour suggests that the connection repeatability issue is also likely to blame for the residual directivity underestimation. If the matched load was connected with a flange misalignment that caused more reflection than when the line was subsequently added, then $\Gamma < D$ and the ripple

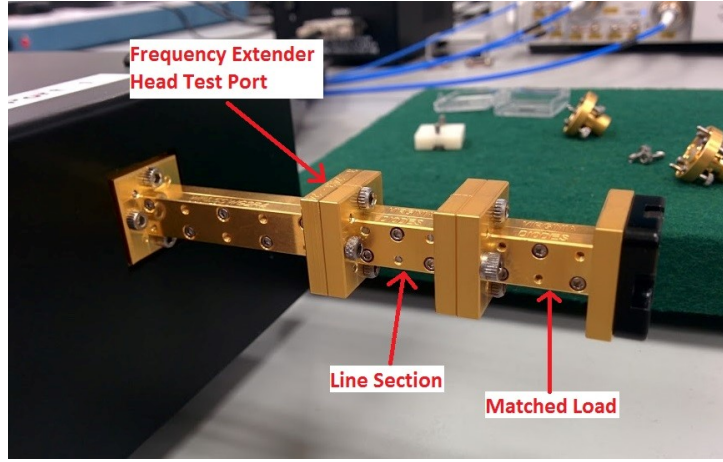


Figure 4.6: The setup used for residual directivity measurements in WR1.5 waveguide.

	Port 1	Port 1
	Residual Directivity	Residual TPM
WR1.5 SOSLT Cal. 1	0.0205	0.1420
WR1.5 SOSLT Cal. 2	0.0250	0.0650

Table 4.4: Residual directivity and TPM values of two WR1.5 waveguide VNA calibrations as measured by the ripple extraction method. Two similar calibrations were performed using different standards from the same kit.

technique will fail.

4.2.2.4 Waveguide Discontinuities

When a discontinuity is present between two sections of rectangular waveguide, an increased reflection will be seen at the location of the join. There are three types of discontinuity possible in rectangular waveguide: E-plane and H-plane lateral displacements, angular displacement, and corner rounding. A report produced by Bannister [15] presented the effects of these discontinuities at centimetre and millimetre wavelengths with measured data, and subsequent work by Kerr extended this using simulations [16], [17].

The error in VRC contributed to by this effect is proportional to wavelength and therefore the aperture size of the waveguide. At submillimetre wavelengths, the error has been shown to be considerable [18], [19]. Recently, efforts have been made to improve the connection repeatability

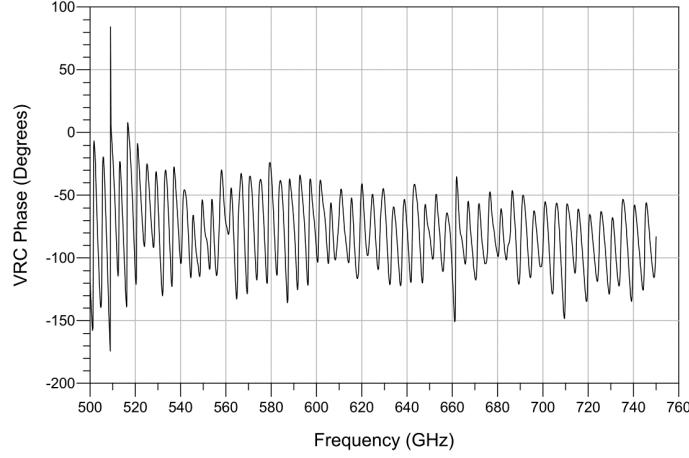


Figure 4.7: Phase plot for a residual directivity measurement performed in WR1.5 waveguide using a SOLT calibration. The lack of phase wrapping over the entire operating bandwidth indicates that the ripple extraction technique will underestimate the residual error.

for waveguide at these wavelengths, and a new IEEE draft standard [20] presents three new connector types which significantly improves the alignment. Another potential improvement attempts to reduce misalignment errors during calibration by replacing the waveguide offset short with a radiating open, and new calibration algorithms have been developed to accompany this [21]. The open standard must be very well characterised, and the technique has not yet seen widespread uptake.

4.2.2.5 Discussion

In centimetre wavelength band waveguides (WG16, WG20) the ripple extraction technique provided similar measurements of residual directivity and TPM than those performed in coaxial line. This is not surprising as providing a good quality match was used for calibration and the connections were well made and aligned, both errors should not have significant contributions outside of the VNA itself. When extended to millimetre wavelength waveguide, the residual errors were larger but still within the recommended values for coaxial lines. At sub-millimetre wavelengths the two calibrations resulted in significant differences in the residual error. A likely cause is the effect of discontinuities in the waveguide components during calibration, especially the matched load and offset short. Additionally, by studying the phase of the measured VRC, it

was shown that the residual directivity measurements were underestimating the true value and therefore the ripple technique was failing. This may also be due to the poor repeatability of the waveguide connection causing the VRC of the matched load used for the ripple extraction to be lower than that used for the calibration. Steps were taken to attempt to resolve this issue, but the extent of the misalignment errors prevented an accurate measurement of residual directivity from being performed. New improvements to sub-millimetre wavelength waveguide flanges should significantly reduce this problem and allow the ripple technique to work more consistently with these waveguides.

A useful assessment method for the ripple technique was presented earlier in this section and should be performed whenever the technique is used. When viewing the measured VRC on a polar or Smith chart, if the origin of the chart lies within the circumference of the trace, the technique is accurately measuring the residual error. However, if the trace is circling outside of the origin of the chart, the ripple half-height only represents the magnitude of the reflection measurement used for the technique. If the technique has been deemed failed by this assessment, the VNA should be recalibrated and the test attempted again. If it fails after this, then the test artefacts should be connected in a way that exaggerates the mismatch, for example by inverting a connection, and the test performed again. Only when the origin of the Smith chart lies within the reflection trace are the results from the technique accurate.

4.2.2.6 Conclusion

The investigation studied the effectiveness of the ripple extraction technique when applied to rectangular waveguide measurements up to 750 GHz. Typical values of residual directivity and test port match were provided for users to compare against. The effect of discontinuities at waveguide interconnections were explored and some recent efforts to mitigate these problems were mentioned. Finally, the investigation showed that the ripple extraction technique may not currently be a reliable way of measuring residual error in sub-millimetre wavelength systems.

4.3 Sources of VNA Measurement Uncertainty

4.3.1 Calibration Standards

4.3.2 Sources of Random Error

4.3.3 Additional Sources

4.4 Software Frameworks for VNA Uncertainty Evaluation

4.4.1 Standalone Vendor Tools

4.4.2 Keysight PNA-X Dynamic S-Parameter Uncertainty Option

4.4.3 METAS VNA Tools II

4.4.4 NIST Microwave Uncertainty Framework

4.5 Sources of MHVNA Measurement Uncertainty

4.5.1 Power Calibration

4.5.2 Phase References

4.6 Evaluation of Uncertainty in MHVNA Measurements

4.6.1 Measurement Models

4.6.2 Input Quantities

4.6.3 Processing Structure

4.6.4 Results

4.7 Conclusions

4.8 Introduction

The dramatic growth of radio-based devices and applications over the last 50 years has led to the VNA becoming a critical instrument in most laboratories. Many of these applications required

both accurate and reliable measurements from these VNAs. This is particularly so in areas such as manufacturing, calibration and testing. This is often driven by requirements given in international Quality Management documents such as the ISO 9000 series of standards [1] (for manufacturing and process control) and the ISO 17025 standard [2] (for calibration and testing).

The requirements given in these international standards are for measurements that can be demonstrated as fit-for-purpose (in terms of the achievable level of accuracy, etc) and made traceable to the international system of units [3], [4]. These requirements were not trivial for a VNA due to the complicated nature of the VNAs operating principles, for example the calibration mathematics. Combined with the available computing power and cost at that time, a full or rigorous evaluation of uncertainty for VNA measurements, per for example the ISO GUM document, was difficult and time-consuming. This led to much work by experts in this field to develop easier methods that addressed these needs in ways that were suitable for use by end-users in the manufacturing, calibration and testing communities. Much of this work was undertaken by the ANAMET Technology Group (www.npl.co.uk/anamet) during the 1990s. This resulted in a series of reports [5]–[7] describing the development of a guidance document that gave a procedure for assessing the performance of calibrated VNAs. The resulting guidance document [8] was published by the European co-operation for Accreditation (EA, www.european-accreditation.org) so that laboratories operating to the ISO 17025 standard and/or ISO 9000 series of standards could implement the method for their own purposes. Ownership of this EA document was later transferred to the European Association of National Metrology Institutes (EURAMET) and re-published [9] as part of their Calibration Guides series of documents. This document, along with the recent updated version, is available as a free download from the EURAMET web-site: www.euramet.org.

In addition to the EURAMET guide, VNA manufacturers have also produced their own advice for users to estimate the combined standard uncertainty in their measurements [10] and provided software tools in some cases [11], [12]. Often this advice is based on the same methods presented in the EURAMET guide.

In more recent years rigorous evaluations of VNA uncertainty have become possible, through both the efforts of NMIs and access to greater computing resources. The difference between the previous approaches (which we will call “residual error” evaluations) and rigorous evaluations concerns the way in which uncertainty contributed by the VNA calibration is estimated and included in the measurement model. The two methods will be discussed later in this Chapter. Figure 4.8 illustrates the general structure of VNA uncertainty sources and evaluation, from

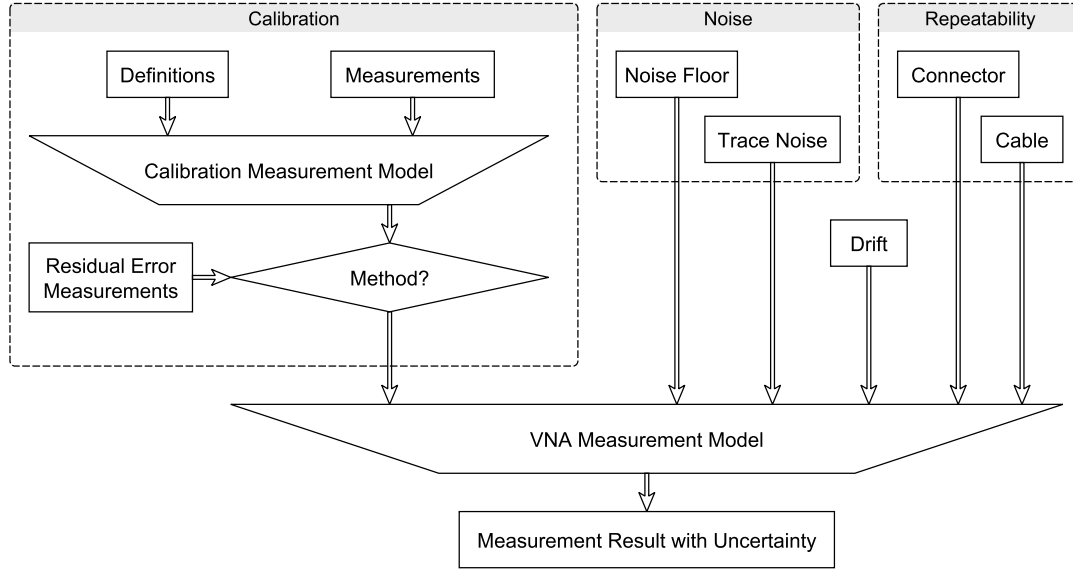


Figure 4.8: Structure of VNA uncertainty evaluation. Input quantities are shown at the top of the diagram, grouped where applicable. Uncertainties from the calibration are either evaluated directly (rigorous/full evaluation) or are estimated after the calibration has been performed (“residual error” evaluation). Typically the calibration measurement model is processed as part of the VNA measurement model, but it is shown separately here to distinguish the difference between VNA uncertainty evaluation methods.

which each component will now be explained.

4.9 VNA Measurement Model Input Quantities

4.9.1 Calibration Standards

In order to perform the calibration, or error correction, of a VNA as described in Chapter 2, we must compare measurements of impedance standards to definitions of their true values in order to obtain error coefficients. It is interesting that although sources of error in the calibration cover both systematic and random types, when the calibration is performed and these quantities are measured, any random errors are captured in the evaluated combined uncertainty, which itself is purely systematic.

4.9.1.1 Definitions

Because the definitions of impedance standards are based on prior measurements, and all measurements include uncertainty, they are also included as a source of uncertainty in the VNA measurement model. We will now look at how uncertainty can be included in the three types of standard definition commonly used in VNA calibration routines.

Databased Definitions The simplest definition of an impedance standard is the databased definition. The standard is characterised by measurement on a VNA which has been calibrated to a suitable accuracy, i.e. using standards which are at an appropriate position in the traceability chain (closer to the primary standard). The combined standard uncertainty of the characterisation is then provided with the estimates as the definition of the standard.

It is important to characterise the standard across a suitable frequency range, usually the entire range for which it can be applied. In addition to the combined standard uncertainty, the covariances between the real and imaginary components of the measured s-parameters can also be included in the standard definition for more accurate uncertainty evaluations in measurements using the standard [RidlerSalterEtc]. If the standard will be used to calibrate VNAs producing results to be used in time-domain studies or multi-harmonic studies (i.e. nonlinear work), then ideally covariances between s-parameters for each frequency should be included. However, this is something which would consume a lot of data storage and the saving of this information is not supported by VNAs at this time.

Polynomial Model Definitions In order to reduce the information required to accompany the standard, one-port models are often defined as coefficients of a polynomial fit between reflection coefficient and frequency. This characterisation is performed by the manufacturer and the coefficients are available in the user manual or specification sheet for the impedance standards. Testing is performed by the manufacturer to ensure that the supplied standards do not deviate from this characterisation. Similar to the databased definition, care must be taken to ensure the characterised polynomial fit is valid across a suitable frequency range.

An example of a short-circuit polynomial fit, as defined in [22], is

$$L_S = L_0 + L_1 f + L_2 f^2 + L_3 f^3, \quad (4.5)$$

$$Z_S = j2\pi f L_S, \quad (4.6)$$

$$\Gamma_S = \frac{Z_S - Z_r}{Z_S + Z_r}, \quad (4.7)$$

where L_S and Z_S are the respective inductance and impedance of the short-circuit, L_n are coefficients provided by the definition of the standard, Z_r is the reference impedance and Γ_S is the reflection coefficient of the short-circuit at frequency f . A similar model is used for open-circuits, with the inductive component replaced with a capacitive one.

Although no manufacturers are yet providing uncertainties with the polynomial coefficients supplied in their standard definitions, some have included the ability for future incorporation in the data file structure. In addition to uncertainty originating from the measurement of the standard, there will also be uncertainties relating to the error of the polynomial fit. Because of this, the polynomial definition is typically the least accurate type of standard definition, however it is popular due to the portability of the small number of coefficients which can be defined for the entire set of manufactured models.

Physical Definitions Typically understood as the most accurate, physically defined standards use robust geometric models to calculate their impedances from dimensional measurements. Models are available in coaxial transmission line for short-circuits, open-circuits and arbitrary line lengths. Good examples of these models and their derivations can be found in [22], [23, Appendix C]. Matched loads are more difficult to accurately model and so they are typically included as a databased standard, even if the other standards are physically defined. In order to avoid this issue, the TRL calibration can be used, which requires standards that can all be physically defined. This combination of calibration model and standard definition is commonly used by NMIs when performing traceable characterisations of other standards (i.e. databased definitions).

To obtain uncertainties in the reflection coefficients of physically defined standards, which are the measurands which we would then like to use as input quantities to the VNA uncertainty evaluation, an additional uncertainty evaluation must be performed. This takes the estimates and uncertainties of the dimensional measurements of the standards and propagates them through the measurement model (the geometric model relating dimension and frequency to reflection coefficient) to obtain the reflection coefficients with uncertainties required.

Dimensional measurements of the standards are usually supplied by the manufacturer for TRL calibration kits, but like the other definition types can be supplanted by recent measurements performed by NMIs or calibration providers. Another benefit of physically defined impedance standards is that their definition is valid over their entire valid frequency range, removing the risk associated with databased and polynomial definitions.

Covariance information is not required to be supplied by the definition because the dimensional measurements should be uncorrelated (they are separately manufactured parts). During the uncertainty evaluation to obtain the reflection coefficients of the standards, covariances between both the complex impedance components and those at different frequencies are calculated. Therefore physical definitions are well suited to both portability, accuracy and use with time-domain or nonlinear measurements.

4.9.1.2 Measurements

During the calibration procedure the impedance standards are measured. Although the VNA is not yet calibrated, it is still performing a measurement (with reference to an arbitrary impedance) and so all of the sources of uncertainty additional to those from the calibration will be included. Sources such as noise are therefore counted several times during a VNA measurement due to their inclusion in the prior calibration measurements. Because of this, it is important to include correlations where possible for these sources as they will have a greater impact on the combined standard uncertainty of the result.

4.9.2 Noise

Noise plays an important role in RF and microwave engineering. It is a random effect which occurs at a fundamental level and cannot be corrected for, both in communication systems and instrumentation. Wireless transmission systems lose a large amount of signal through propagation and receivers must be designed to handle signals with a considerable amount of noise. Likewise, VNA transmission measurements of devices with high isolation or reflection measurements of devices with low insertion loss also exhibit a low signal-to-noise ratio. To obtain accurate results, it is therefore important for VNA manufacturers to provide instruments with low noise sources and receivers. To ensure that measurements which are more susceptible to noise are accurate, it is also important to quantify the noise associated with the measurements so that it can be included as a contribution to their combined standard uncertainties.

The amount of electrical noise in VNA measurements is a sum of thermal noise and contributions from the VNA components. Thermal noise, caused by the random motion of free electrons in a conducting material by heat, is specified as 4×10^{-21} W/Hz (-174 dBm) [10]. The intermediate frequency bandwidth (IFBW) setting on the VNA controls the frequency range of signals being measured at each discrete point in the sweep, and can typically be set from 1 Hz to more than 10 kHz. It is this figure which multiplies the thermal noise to provide a theoretical minimum noise floor if we assumed the VNA did not contribute any noise itself. The disadvantage of a very low IFBW is that the VNA takes longer to perform the measurement, so there is a compromise between speed and accuracy. This provides a good example of where uncertainty evaluation can have a direct perceivable benefit to measurement accuracy - if the IFBW is reduced the uncertainty due to noise should also reduce. By quantifying the uncertainty the engineer can make an informed decision as to what value to set the IFBW, which could have expensive time implications in a test environment.

In addition to the thermal noise, the VNA contributes to the noise level from the source, receiver and other test set components (i.e. local oscillator phase noise). Setting attenuators to lower values will improve noise as resistors add additional thermal noise to the measurement. The noise figure (NF) of a VNA is typically quoted in the specifications, and can be used with the thermal noise to calculate the theoretical noise level of the VNA measurement [10]:

$$L_N = -174\text{dBm} + NF + 10 \log(S_F)\text{dB} + 10 \log\left(\frac{B_{IF}}{\text{Hz}}\right) \text{dB} \quad (4.8)$$

For the purposes of uncertainty evaluation, VNA noise level is measured on the instrument itself and is typically included in the measurement model as two input quantities, noise floor and trace noise. It is important that the VNA settings (IFBW, test port power, averaging) are the same as will be used for the DUT measurements, and that no calibration is applied.

4.9.2.1 Noise Floor

The noise floor is a term used to describe the noise present in the measurement with no external signal present. It can be measured with a matched load connected directly to the test port and should not include any noise contribution from the VNA source. Alternatively it can be measured from the transmission measurements of a 2-port VNA while short-circuits or open-circuits are connected for the trace noise measurement described below. Many measurements are made ([9] suggests a few hundred) and the standard deviation is calculated for each frequency point.

4.9.2.2 Trace Noise

Trace noise includes noise contributions from the VNA source and is a function of the measured power at the receiver. It is measured with a short-circuit or open-circuit connected directly to the test port.

4.9.3 Repeatability

Both the DUT connections and movement of the VNA port extension cables contribute uncertainty to the repeatability of VNA measurements. These errors vary when a different device is connected to the test setup, and are included for both the DUT measurement and the measurements of the impedance standards during calibration.

4.9.3.1 Connections

Connector repeatability is a fundamental quality of precision RF and microwave connectors. It can have a significant impact on DUT measurements if the response of the connector varies between multiple connections, and this perturbs the reference impedance of the calibrated VNA away from 50 Ohms if this occurs between connecting calibration standards.

In order to measure connector repeatability uncertainty, a short-circuit can be connected to a calibrated VNA and the reflection coefficient measured. Measurements are then repeated, reconnecting the short at a different azimuthal rotation each time. This should be done at least three times with rotations of 120 deg, although some guidance recommends up to 16 times[9].

4.9.3.2 Cable Stability

When test port cables are flexed, their physical dimensions are perturbed, which in turn affects their s-parameter response. Specialist VNA test port extension cables are provided to mitigate the effects of cable flexure, although they do not remove them. It is therefore advisable to restrict the movement of these cables using clamps and supports, and also to wait a suitable time (typically 30 seconds or more) for stresses in the internal dielectric to settle.

Cable stability can be characterised in different ways depending on the VNA measurement model used. The general method is to perform repeated measurements while moving the cable between a range of positions which cover the maximum extent the cable will move during measurements of the calibration standards and DUTs. Specifically, some methods connect a

short-circuit to each test port cable, and others connect the cables together to perform the characterisation (if used for a two-port measurement).

4.9.4 Drift

Drift does not strictly fit into the random or systematic error category, because it can vary over the course of measurements but is not random in nature. The dominant cause of drift is environmental conditions, especially temperature variations. Temperature affects the physical dimensions and s-parameters of RF components in the test setup and also the behaviour of the electronic devices inside the source and receiver - both of which perturb the result of the measurement.

Drift is an important source of uncertainty to include for measurements occurring over a long period of time after calibration, especially those in a test environment or in standards labs while performing a large number of repeat measurements. In order to minimise the effects of drift, the room which the VNA is located in should have good temperature control and all the equipment should have been powered on for at least 24 hours for their internal temperatures to settle.

In order to measure uncertainty due to drift, many repeat measurements are performed over a long period (24 hours or more). This procedure can be automated so that an operator is not required for the entire period. An alternative method of including uncertainty due to drift can be achieved by performing calibration measurements of standards both before and after the DUT measurements. The calibrations can be averaged and a crude standard deviation obtained.

4.9.4.1 VNA Linearity

One source of VNA measurement uncertainty which is still included in some evaluations is that from the nonlinearity of the receivers. However, in modern VNAs automatic level control corrections are sufficiently accurate to allow this error to be neglected [24], [25].

4.10 Residual Calibration Error VNA Uncertainty Evaluation

4.10.1 Method

4.10.2 Application to Waveguide VNAs

4.11 Rigorous VNA Uncertainty Evaluation

4.11.1 Method

4.11.2 Software Frameworks

4.11.2.1 Keysight PNA-X Dynamic S-Parameter Uncertainty Option

4.11.2.2 METAS VNA Tools II

4.11.2.3 NIST Microwave Uncertainty Framework

4.12 NVNA Uncertainty Evaluation

4.12.1 Phase Reference

4.12.2 Power Meter

4.12.3 Uncertainty Evaluation using MUF

4.13 Conclusion

References

- [1] International Organization for Standardization, *Iso 9000: International standards for quality management*, Genève, Switzerland: International Organization for Standardization, 2015. [Online]. Available: <https://www.iso.org>.
- [2] —, *Iso 17025: General requirements for the competence of testing and calibration laboratories*, Genève, Switzerland: International Organization for Standardization, 2005. [Online]. Available: <https://www.iso.org>.
- [3] International Bureau of Weights and Measures (BIPM), *The International System of Units (SI)*, 8th ed. 2019. [Online]. Available: <https://www.bipm.org>.
- [4] —, *A concise summary of the International System of Units, the SI*, 8th ed. 2019. [Online]. Available: <https://www.bipm.org>.
- [5] R. N. Clarke, J. P. Ide, G. R. Orford, and N. M. Ridler, “Draft eal procedure for the assessment of vector network analysers (vna), anamet report 002,” National Physical Laboratory, Tech. Rep. 002, Sep. 1996, Available from the National Physical Laboratory.
- [6] J. P. M. de Vreede, “Draft procedure for the assessment of vector network analysers (vna), anamet report 019,” National Physical Laboratory, Tech. Rep. 019, Oct. 1998, Available from the National Physical Laboratory.
- [7] —, “Draft ea guidance document: Assessment of calibrated vector network analysers (vna), anamet report 026,” National Physical Laboratory, Tech. Rep. 026, Aug. 1999, Available from the National Physical Laboratory.
- [8] European co-operation for Accreditation, “Ea guidelines on the evaluation of vector network analysers (vna),” Tech. Rep. EA-10/12, May 2000.
- [9] European Association of National Metrology Institutes (EURAMET), “Guidelines on the evaluation of vector network analysers (vna), version 2,” Tech. Rep. cg-12, Mar. 2011.

- [10] M. Hiebel, *Fundamentals of Vector Network Analysis*. Rohde & Schwarz, 2008, ISBN: 9783939837060.
- [11] Keysight Technologies. Downloadable vector network analyzer uncertainty calculator, [Online]. Available: <https://www.keysight.com/main/software.jsp?ckey=1000000418:epsg:sud&cc=GB> (visited on 07/16/2019).
- [12] Anritsu. Exact uncertainty calculator, [Online]. Available: <https://www.anritsu.com/en-GB/test-measurement/support/downloads/software/dwl19572> (visited on 07/16/2019).
- [13] L. Stant, P. Aaen, and N. Ridler, “Evaluating residual errors in waveguide network analyzers from microwave to submillimetre-wave frequencies,” in *IET Colloquium on Millimetre-Wave and Terahertz Engineering & Technology 2016*, Institution of Engineering and Technology (IET), 2016. DOI: 10.1049/ic.2016.0016.
- [14] —, “Evaluating residual errors in waveguide VNAs from microwave to submillimetre-wave frequencies,” *IET Microwaves Antennas Propag.*, vol. 11, no. 3, pp. 324–329, Feb. 2017. DOI: 10.1049/iet-map.2016.0455.
- [15] D. Bannister, E. Griffin, and T. Hodgetts, *On the dimensional tolerances of rectangular waveguide for reflectometry at millimetric wavelengths*. National Physical Laboratory, 1989.
- [16] A. Kerr, E. Wollack, and N. Horner, “Waveguide flanges for alma instrumentation,” *ALMA Memo*, vol. 278, 1999.
- [17] A. Kerr, “Mismatch caused by waveguide tolerances, corner radii, and flange misalignment,” *Electronics Div. Tech. Note*, vol. 215, 2010.
- [18] D. F. Williams, “500 ghz–750 ghz rectangular-waveguide vector-network-analyzer calibrations,” *IEEE Transactions on Terahertz Science and Technology*, vol. 1, no. 2, pp. 364–377, 2011.
- [19] H. Li, A. R. Kerr, J. L. Hesler, and R. M. Weikle, “Repeatability of waveguide flanges with worst-case tolerances in the 500–750 ghz band,” in *79th ARFTG Microwave Measurement Conference*, IEEE, 2012, pp. 1–8.
- [20] “Ieee standard for rectangular metallic waveguides and their interfaces for frequencies of 110 ghz and above—part 2: Waveguide interfaces,” *IEEE Std 1785.2-2016*, pp. 1–22, Sep. 2016. DOI: 10.1109/IEEESTD.2016.7564020.

- [21] A. Arsenovic, R. M. Weikle, and J. Hesler, “Two port calibration insensitive to flange misalignment,” in *84th ARFTG Microwave Measurement Conference*, Dec. 2014, pp. 1–7. DOI: 10.1109/ARFTG.2014.7013416.
- [22] Keysight Technologies, “Specifying calibration standards and kits for keysight vector network analyzers,” Tech. Rep. 5989-4840EN, Jun. 2016.
- [23] A. Lewandowski, “Multi-frequency approach to vector-network-analyzer scattering-parameter measurements,” PhD thesis, Warsaw University of Technology, 2010.
- [24] D. K. Rytting, “Network analyzer accuracy overview,” in *58th ARFTG Conference Digest*, vol. 40, Nov. 2001, pp. 1–13. DOI: 10.1109/ARFTG.2001.327486.
- [25] J. Martens, “On quantifying the effects of receiver linearity on vna calibrations,” in *2007 70th ARFTG Microwave Measurement Conference (ARFTG)*, Nov. 2007, pp. 1–6. DOI: 10.1109/ARFTG.2007.8376223.