

MEASUREMENT UNCERTAINTY IN NON-LINEAR
BEHAVIOURAL MODELS OF MICROWAVE AND
MILLIMETRE-WAVE AMPLIFIERS

Laurence Stant



A thesis submitted in partial fulfillment for the degree of Doctor of
Philosophy

in the
Advanced Technology Institute and Department of Electronic
Engineering
Faculty of Engineering and Physical Sciences
University of Surrey

September 2017

Declaration of Authorship

I confirm that the submitted work is my own work and that I have clearly identified and fully acknowledged all material that is entitled to be attributed to others (whether published or unpublished) using the referencing system set out in the programme handbook. I agree that the University may submit my work to means of checking this, such as the plagiarism detection service Turnitin UK. I confirm that I understand that assessed work that has been shown to have been plagiarised will be penalised.

The authors confirm that data underlying the findings are available without restriction. Details of the data and how to request access are available from the University of Surrey publications repository at <http://epubs.surrey.ac.uk/blahTBD>

Laurence Stant (Author)

Date

“What error drives our eyes and ears amiss? Until I know this sure uncertainty I’ll entertain the offered fallacy.”

William Shakespeare, The Comedy of Errors

“That’s right!” shouted Vroomfondel, “we demand rigidly defined areas of doubt and uncertainty!”

Douglas Adams, The Hitchhikers Guide to the Galaxy

Abstract

Abstract goes here

Research Outcomes

Publications

- [1] H. Votsi, L. Stant, N. Ridler, and P. Aaen, “Microwave impedance characterization platform for extreme impedance devices,” *IET Microwaves Antennas Propag.*, 2018.
- [2] L. Stant, D. Root, N. Ridler, and P. Aaen, “Uncertainty evaluation of predicted optimum load match using x-parameters of microwave amplifiers,” *IEEE Microw. Wireless Compon. Lett.*, 2018.
- [3] M. Salter, L. Stant, K. Buisman, and T. Nielsen, “An inter-laboratory comparison of NVNA measurements,” in *2011 Workshop on Integrated Nonlinear Microwave and Millimetre-Wave Circuits*, Brive, Jul. 2018.
- [4] L. Stant, M. Salter, N. Ridler, D. Williams, and P. Aaen, “Propagating measurement uncertainty to microwave amplifier nonlinear behavioural models,” *IEEE Trans. Microw. Theory Techn.*, 2018.
- [5] L. Stant, P. Aaen, and N. Ridler, “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.
- [6] —, “Comparing methods for evaluating measurement uncertainty given in the JCGM ‘evaluation of measurement data’ documents,” *Measurement*, vol. 94, pp. 847–851, Dec. 2016. DOI: 10.1016/j.measurement.2016.08.015.
- [7] —, “Evaluating residual errors in waveguide network analysers 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.

Acknowledgements

I want to thank...

Contents

	Page
List of Figures	ix
List of Tables	xi
1 Introduction	3
1.1 Wireless Communications	3
1.2 Amplifier Measurement and Modelling	3
1.3 The Role of Uncertainty in Measurement	3
1.4 Thesis Structure	3
Bibliography	3
2 Radio Frequency and Microwave Measurements	5
2.1 Introduction	5
2.2 Electromagnetic Wave Parameters	6
2.2.1 Wave Definitions	6
2.2.2 Derived Metrics and Figures of Merit	7
2.3 Measurements of Nonlinear Devices	10
2.4 Vector Network Analysers	13
2.4.1 Architecture	13
2.4.2 Error Models	16
2.4.3 Calibration	17
2.5 Large Signal Network Analysers	17
2.5.1 Absolute 8-Term Error Model	17
2.5.2 Power Meter Calibration	17
2.5.3 Phase References	17
2.6 Conclusions	17
Bibliography	17
3 Measurement Uncertainty	19
3.1 Introduction	19
3.2 The Measurement Process	24
3.2.1 Modelling the Measurement	24
3.2.2 Evaluating Standard Uncertainty of Input Quantities	25
3.2.3 Evaluating Combined Standard Uncertainty	34
3.2.4 Expanded Uncertainty and Coverage Intervals	38
3.3 Sensitivity Analysis	38

3.4	Conclusions	40
	Bibliography	40
4	Evaluating Uncertainty in Vector Network Analyser Measurements	42
4.1	Introduction	43
4.2	Evaluating Residual Error in VNA Calibrations	43
4.2.1	Imperfect Calibration Standards	43
4.2.2	The Ripple Technique	43
4.2.3	Application of The Ripple Technique at Sub-Millimetre Wavelength in Rectangular Metallic Waveguide	43
4.3	Sources of VNA Measurement Uncertainty	43
4.3.1	Calibration Standards	43
4.3.2	Sources of Random Error	43
4.3.3	Additional Sources	43
4.4	Software Frameworks for VNA Uncertainty Evaluation	43
4.4.1	Standalone Vendor Tools	43
4.4.2	Keysight PNA-X Dynamic S-Parameter Uncertainty Option	43
4.4.3	METAS VNA Tools II	43
4.4.4	NIST Microwave Uncertainty Framework	43
4.5	Sources of MHVNA Measurement Uncertainty	43
4.5.1	Power Calibration	43
4.5.2	Phase References	43
4.6	Evaluation of Uncertainty in MHVNA Measurements	43
4.6.1	Measurement Models	43
4.6.2	Input Quantities	43
4.6.3	Processing Structure	43
4.6.4	Results	43
4.7	Conclusions	43
	Bibliography	43
5	Propagating Measurement Uncertainty into Nonlinear Behavioural Models	45
5.1	Introduction	45
5.2	The X-Parameter Model	45
5.2.1	Model Definition	45
5.2.2	Extraction Procedure	45
5.2.3	Applications of X-Parameters	45
5.3	Design and Simulation using Nonlinear Behavioural Models Incorporating Measurement Uncertainty	45
5.4	Conclusions	45
	Bibliography	45
6	Applications of Nonlinear Behavioural Models Incorporating Measurement Uncertainty	47
6.1	Introduction	47
6.2	Prediction of Optimum Load Match and Delivered Power using X-Parameters Incorporating Measurement Uncertainty	49
6.3	Conclusions	49
	Bibliography	49

7	Conclusions	51
7.1	Further Work	51
	Bibliography	51

List of Figures

2.1	The frequency dependence of the magnitude (red dotted trace) and phase (blue solid trace) of S_{21} for a bandstop filter.	8
2.2	An short example Touchstone file showing a two-port measurement at two frequencies. The rows continue to the right of the figure.	9
2.3	Gain compression occurs when an amplifier is driven into a nonlinear operating regime.	11
2.4	Intermodulation products from two tones within the cellular channel bandwidth f_1 and f_2 . The second order products, and upper third order products, can be easily filtered out. However, the lower third order products $2f_1 - f_2$ and $2f_2 - f_1$ are located within the channel bandwidth and interact with the useful data, increasing EVM and BER Hall2013	12
2.5	A one-port simple reflectometer. a_1 is the incident wave generated by the source, which is admitted to the DUT while also being sampled by the directional coupler and sent to the reference receiver via a_{1REF} . The reflected wave, b_1 , is also sampled by another directional coupler and sent to the test receiver as b_{1REF} , with the remaining power dissipated at the matched source.	13
2.6	A modern two-source mixer-based VNA, which employs heterodyning to allow measurements at microwave frequencies. Two directional couplers are located between each source and the DUT and are connected back to back. These sample waves travelling in both directions and are connected to mixers which downconvert the microwave frequencies (R) into intermediate frequencies (I) which can be sampled by the complex receivers. The shared local oscillator (LO) feeding the mixers preserves phase coherence between the receivers. This configuration is known as a two-port double-reflectometer VNA. Figure adapted by author from Root2013	15
2.7	The 8-term error model for a two-port measurement. E_D , E_S , and E_T are the same as for the one-port model, except there are now sets of each for both ports. These extra terms account for different error values when the incident signal is sourced from each port. Additionally for the two-port case, a transmission term E_{TN} has been added for each direction.	17
3.1	Egyptian royal cubit rod of Maya (treasurer of King Tutankhamun) 1336–1327 BC. The cubit is thought to be the earliest attested standard measure of length, first used in the 3rd or 4th Millennium BC.	20

3.2	The traceability chain, where the pyramid shows the number of instances of standards in each tier. Secondary standards are held at NMIs and used to periodically calibrate working standards, which are sent by manufacturers and laboratories. User measurements are made using instruments calibrated with these working standards, so they number the greatest and are at the bottom of the traceability chain.	21
3.3	Scaling factor to convert from a GUM standard uncertainty to a GUM Supplement.	28
3.4	The minimum number of observations, n , required to calculate the standard uncertainty of a full set of S-parameters for a microwave device with m ports using the GUM-S2 approach. The number of input quantities, N , for each device is also shown.	32
3.5	An illustration of the propagation of distributions from three input quantities g_{X1} , g_{X2} , g_{X3} , through the measurement model, Y , to the measurand, g_Y GUM'S1	35
3.6	An example of results from a sensitivity analysis which reveal the origins of the complicated behaviour of the combined standard uncertainty with respect to a variable (in this case frequency).	39

List of Tables

3.1	The indication values from the example “Simultaneous Resistance and Reactance Measurement” and their statistical properties as evaluated by the approaches given in GUM’2008 and GUM’S2	30
3.2	A comparison of the results obtained for the example “Simultaneous Resistance and Reactance Measurement” using the approaches given in GUM’2008 and GUM’S2	31
3.3	The difference in standard uncertainties obtained using the GUM ($u(\bar{x})_{GUM}$) and the GUM-S1/S2 ($u(\bar{x})_{SUPP}$) approaches to measure a full set of scattering parameters for microwave devices with various numbers of ports, m . Each device has $2m^2$ input quantities, N , and requires a minimum of $N + 3$ repeat observations, n , in order for $u(\bar{x})_{SUPP}$ to be defined.	31

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetur adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

Fusce mauris. Vestibulum luctus nibh at lectus. Sed bibendum, nulla a faucibus semper, leo velit ultricies tellus, ac venenatis arcu wisi vel nisl. Vestibulum diam. Aliquam pellentesque, augue quis sagittis posuere, turpis lacus congue quam, in hendrerit risus eros eget felis. Maecenas

eget erat in sapien mattis porttitor. Vestibulum porttitor. Nulla facilisi. Sed a turpis eu lacus commodo facilisis. Morbi fringilla, wisi in dignissim interdum, justo lectus sagittis dui, et vehicula libero dui cursus dui. Mauris tempor ligula sed lacus. Duis cursus enim ut augue. Cras ac magna. Cras nulla. Nulla egestas. Curabitur a leo. Quisque egestas wisi eget nunc. Nam feugiat lacus vel est. Curabitur consectetur.

Suspendisse vel felis. Ut lorem lorem, interdum eu, tincidunt sit amet, laoreet vitae, arcu. Aenean faucibus pede eu ante. Praesent enim elit, rutrum at, molestie non, nonummy vel, nisl. Ut lectus eros, malesuada sit amet, fermentum eu, sodales cursus, magna. Donec eu purus. Quisque vehicula, urna sed ultricies auctor, pede lorem egestas dui, et convallis elit erat sed nulla. Donec luctus. Curabitur et nunc. Aliquam dolor odio, commodo pretium, ultricies non, pharetra in, velit. Integer arcu est, nonummy in, fermentum faucibus, egestas vel, odio.

Sed commodo posuere pede. Mauris ut est. Ut quis purus. Sed ac odio. Sed vehicula hendrerit sem. Duis non odio. Morbi ut dui. Sed accumsan risus eget odio. In hac habitasse platea dictumst. Pellentesque non elit. Fusce sed justo eu urna porta tincidunt. Mauris felis odio, sollicitudin sed, volutpat a, ornare ac, erat. Morbi quis dolor. Donec pellentesque, erat ac sagittis semper, nunc dui lobortis purus, quis congue purus metus ultricies tellus. Proin et quam. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos hymenaeos. Praesent sapien turpis, fermentum vel, eleifend faucibus, vehicula eu, lacus.

1 Introduction

1.1 Wireless Communications

1.2 Amplifier Measurement and Modelling

1.3 The Role of Uncertainty in Measurement

1.4 Thesis Structure

Testing, testing[1], [2].

References

- [1] L. Stant, P. Aaen, and N. Ridler, “Evaluating residual errors in waveguide network analysers 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.
- [2] —, “Comparing methods for evaluating measurement uncertainty given in the JCGM ‘evaluation of measurement data’ documents,” *Measurement*, vol. 94, pp. 847–851, Dec. 2016. DOI: 10.1016/j.measurement.2016.08.015.

2 Radio Frequency and Microwave Measurements

2.1 Introduction

To characterise nonlinear behavioural models, the radio frequency (RF) response of a device to electromagnetic wave stimuli must be measured. When compared with DC (and low frequency) measurements, RF and microwave measurements present significant additional challenges. For DC systems, it is desirable to propagate voltages through a circuit with minimal loss in amplitude. To achieve this effectively, components are typically designed with high input impedance and low output impedance. With RF systems, circuit components and interconnects can be of the order of a quarter-wavelength in length, and therefore signals must be treated as electromagnetic waves to account for different behaviour at these frequencies. When a travelling wave encounters a discontinuity in impedance, such as a cable connector or on-wafer structure, some of the power in the wave is reflected. The amount of reflected power is proportional to the size of the impedance mismatch between each side of the discontinuity. Hence, for RF systems, the transmission of power is the focus of the circuit designer. The measurement of power flowing through a transmission line is complicated by three key factors. Firstly, because the waves are travelling, the instantaneous voltage at any point on the transmission line will vary between the peak-to-peak values of the wave. Secondly, there are waves travelling in both directions along the transmission line which must be measured separately. Finally, the power of the wave is a complex quantity which consists of both magnitude and phase. To perform these measurements, a specialist instrument called a vector network analyser (VNA) can be used. In this chapter, the concepts and measurements associated with this instrument are introduced, which will be used later in the thesis to understand the uncertainty contributions from measurements to nonlinear

behavioural models.

2.2 Electromagnetic Wave Parameters

2.2.1 Wave Definitions

To describe the power of electromagnetic waves propagated through a transmission line, several definitions are in use in industry and academia for either accuracy or convenience. To avoid confusion in this document, these will now be defined. Information presented in this section has been obtained from [1-5].

2.2.1.1 Travelling Waves

Travelling waves represent a solution to Maxwell's equations along a transmission line. They are physical and measurable via slotted line experiments or thru-reflect-line calibrations [6] (see Chapter X). Travelling waves are defined by the total transverse electric and magnetic fields \mathbf{E}_t and \mathbf{H}_t of a single propagating mode at each frequency:

$$\mathbf{E}_t = c^+ e^{-\gamma z} \mathbf{e}_t + c^- e^{+\gamma z} \mathbf{e}_t, \quad \mathbf{H}_t = c^+ e^{-\gamma z} \mathbf{h}_t - c^- e^{+\gamma z} \mathbf{h}_t \quad (2.1)$$

where, following the notation of [5], \mathbf{e}_t and \mathbf{h}_t are the un-normalized electric and magnetic fields of the modal solution of Maxwell's equations in transmission line, $\gamma = a + ib$ is the complex propagation constant of the mode, z is the direction of propagation, and c^+ and c^- are complex quantities representing the un-normalized forward and backward amplitude of the mode, respectively.

2.2.1.2 Equivalent-Circuit Voltage and Current

To represent travelling waves as equivalent low frequency circuit parameters such as voltage and current, a normalisation is chosen to derive a characteristic impedance for the transmission line. This normalisation takes the form

$$\mathbf{E}_t(z) = \frac{v(z)}{v_0} \mathbf{e}_t, \quad \mathbf{H}_t(z) = \frac{i(z)}{i_0} \mathbf{h}_t, \quad (2.2)$$

where v_0 and i_0 are normalisation constants that allow v and i to take units of root-mean-square voltage and current, respectively [5].

2.2.1.3 Pseudowaves

Equivalent voltages and currents cannot be used in lossy transmission lines where the electric and magnetic fields are out of phase. To account for this and provide a solution which can be used with conventional circuit design methodologies (e.g. Smith chart techniques [7]) and simulators, pseudowaves can be used. This representation is defined with a reference impedance, Z_{ref} , which can be chosen by the user, but is typically 50- Ω in conventional measurements. The forward and backward pseudowaves a and b can be written as:

$$a(Z_{\text{ref}}) = \left[\frac{|v_0|}{v_0} \frac{\sqrt{\Re(Z_{\text{ref}})}}{2|Z_{\text{ref}}|} \right] (v + iZ_{\text{ref}}), \quad b(Z_{\text{ref}}) = \left[\frac{|v_0|}{v_0} \frac{\sqrt{\Re(Z_{\text{ref}})}}{2|Z_{\text{ref}}|} \right] (v - iZ_{\text{ref}}) \quad (2.3)$$

2.2.1.4 Power Waves

Finally, power waves are defined so that the relationship $P = |a|^2 - |b|^2$ is true for any reference impedance, where P is the power transmitted through the transmission line and a and b are the forward and backward power waves, respectively. They are defined as:

$$a(Z_{\text{ref}}) = \frac{v + iZ}{2\sqrt{\Re(Z_{\text{ref}})}}, \quad b(Z_{\text{ref}}) = \frac{v - iZ}{2\sqrt{\Re(Z_{\text{ref}})}}. \quad (2.4)$$

Data taken from Keysight instruments used later in this work is presented in power wave format, with units of square-root Watts. To convert these values into decibels referenced to 1 milliwatt, the following formula is used:

$$P(\text{dBm}) = 10 \log_{10}(P(\sqrt{W})^2) + 30 \quad (2.5)$$

2.2.2 Derived Metrics and Figures of Merit

The behaviour of a linear microwave device can be completely defined by the complex ratio of electromagnetic waves which are scattered at each port to those which are incident at each port. The combination of these ratios constitutes the scattering parameters (s-parameters) of a microwave device and are used extensively in the design and measurement of microwave systems[ref]. The formal definitions of the s-parameters for a two-port device are

$$S_{11} = \left. \frac{b_1}{a_1} \right|_{a_2=0}, \quad S_{12} = \left. \frac{b_1}{a_2} \right|_{a_1=0}, \quad S_{21} = \left. \frac{b_2}{a_1} \right|_{a_2=0}, \quad S_{22} = \left. \frac{b_2}{a_2} \right|_{a_1=0}, \quad (2.6)$$

where both a and b can be expressed in either pseudowave or power wave representation. The term “scattered” can be interchanged with “transmitted” and “reflected” depending on if the scattered wave is output on a different port, or the same port, to the incident wave, respectively.

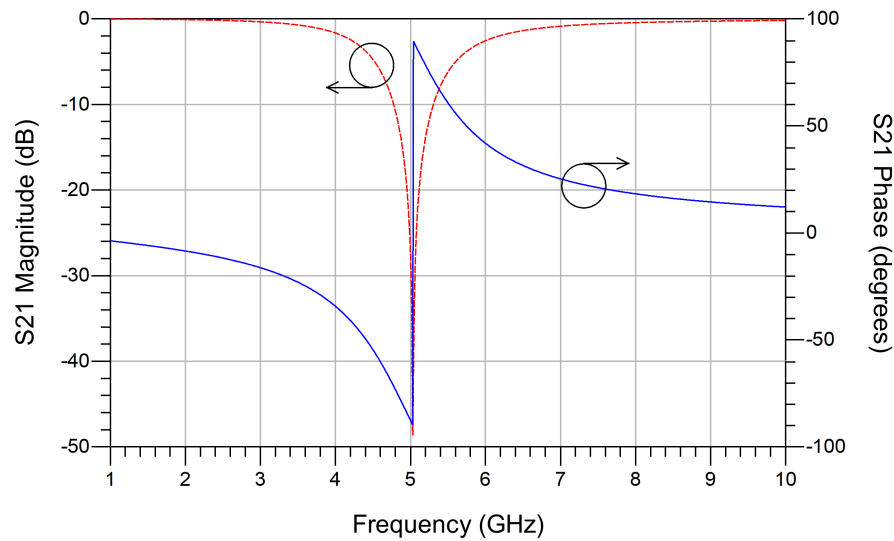


Figure 2.1: The frequency dependence of the magnitude (red dotted trace) and phase (blue solid trace) of S_{21} for a bandstop filter.

A signal flow diagram is provided in Fig. 1 showing the relationship between equivalent-circuit voltage and current, pseudowaves/power waves and s-parameters for a two-port device. The s-parameters of all microwave devices will exhibit some degree of frequency dependence. This effect originates from physical processes occurring in the device and can either be a benefit or hinderance to a design. Most passive components (including cables) will have a usable bandwidth which is an unwanted limitation, whereas microwave filters are a ubiquitous component where the same fixed bandwidth is the main purpose of the device. To capture this frequency dependence, s-parameters are measured across a frequency range and stored in a table, usually in Touchstone format (see Fig.2). An example of the frequency dependence of a filter is shown in Fig. 3. For a device operating in the linear regime, if multiple stimuli at different frequencies are incident on the device, they will not interact with each other. The scattered waves will have the same frequency components as if the stimulus at each frequency was applied separately. This is called the frequency superposition principle and does not apply to nonlinear operating regimes, which will be discussed later in this chapter.

Scattering parameters are often expressed in matrix form, where the column index is the scattered port, and the row index is the incident port. For a two-port device, the s-parameter

```
!Hewlett Packard, E8753E, U538432892
!Date: Saturday, February 03, 2018 19:29:08
!Correction: S11(Off) S21(Off) S12(Off) S22(Off)
!S2P File: Measurements: S11, S21, S12, S22:
# Hz S RI R 50
3.00000000000000E+4 9.34997558593750E-1 5.69458007812500E-2 -1.73002481460571E-5 -1.6689300537109
6.00000000000000E+9 1.53465270996094E-2 -6.19812011718750E-2 1.51917338371277E-5 5.52944839000702
```

Figure 2.2: An short example Touchstone file showing a two-port measurement at two frequencies. The rows continue to the right of the figure.

matrix would be

$$S = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \quad (2.7)$$

The most interesting characteristic of a two-port microwave device is often the effect which it has on a transmitted wave in the forward direction (S_{21}). If the device increases the magnitude of the incident signal this metric is called gain, otherwise it is called insertion loss. Typically gain is associated with active devices (those which are powered from an external source separate to the incident microwave signals) such as amplifiers, and insertion loss is associated with passive devices (those with no external power source) such as attenuators, splitters and mixers. The power gain (operating gain) and insertion loss relating to S_{21} can be calculated using

$$\text{Power Gain} = 10 \log_{10} |S_{21}|^2 \text{ dB}, \quad (2.8)$$

and

$$\text{Insertion Loss} = -10 \log_{10} |S_{21}|^2 \text{ dB}, \quad (2.9)$$

respectively.

Optimal transmission in microwave systems requires impedance matching between components, and it is inevitable that this matching will not be perfect and so some power will be reflected in a two-port device. Therefore, the match of a device is another important measurement, which is dependent on the voltage reflection coefficient (Γ) of the device and can be related to the impedance of a source and load by

$$\Gamma_{xx} = S_{xx} = \frac{Z_L - Z_S}{Z_L + Z_S}, \quad (2.10)$$

where x is a port index. A more thorough definition of voltage reflection coefficient for a two-port device includes any effect from the impedance seen at the other port, and for the case of input match is calculated as

$$\Gamma_{11} = S_{11} + \frac{S_{12}S_{21}\Gamma_L}{1 - S_{22}\Gamma_L}, \quad (2.11)$$

where Γ_L is the voltage reflection coefficient of the load connected to the device. For amplifiers, the amount of isolation (reduction of S_{12}) is an important characteristic of the device, whereby a fully isolated amplifier ($S_{12} = 0$) is said to be unilateral and equations 2.10 and 2.11 are equivalent.

For active devices, such as amplifiers, it can also be useful to consider the power reflected at the input when calculating the power gain of the device. The transducer gain of a device accounts for this potential loss of power at the input and provides a more portable metric which is not dependent on the impedance of the measurement setup. It is defined as

$$G_T = \frac{1 - |\Gamma_S|^2}{1 - |\Gamma_{in}\Gamma_S|^2} |S_{21}|^2 \frac{1 - |\Gamma_L|^2}{1 - |S_{22}\Gamma_L|^2}, \quad (2.12)$$

where Γ_{in} is the input match of the device.

For all devices operating in the linear regime, any reflected or transmitted wave will have a frequency equivalent to that incident to the device. In addition, the stimulus power that was used to measure the s-parameters is not important as the ratio of scattered to incident wave magnitude is not dependent on this quantity. However, when microwave devices operate in the nonlinear regime, these conditions no longer apply, and s-parameters cannot be used to capture the full behaviour of the device.

2.3 Measurements of Nonlinear Devices

Microwave devices operating in the nonlinear regime exhibit three differences from their linear counterparts which are significant to the designer:

1. The amplitude of electromagnetic waves scattered from the device are not linearly dependent on the amplitude of waves incident. This is the cause of features such as gain compression and gain expansion in amplifiers. Some of these effects are solely due to the nonlinear sources inside the device, while others are a symptom of the combined response of the nonlinearity and the power supply. A typical gain compression curve is shown in Fig. 4.
2. The frequency superposition principle does not apply, and instead the frequency spectrum of scattered waves contains components at frequencies other than those incident upon it. Rather than the incident signals purely summing inside the device, they are also multiplied

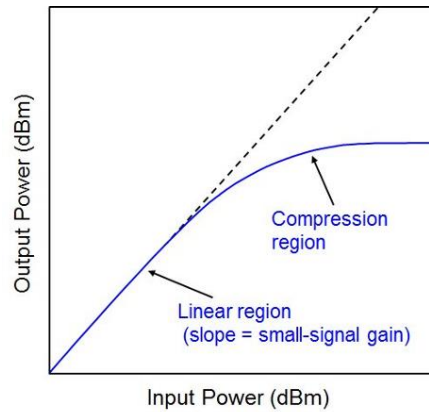


Figure 2.3: Gain compression occurs when an amplifier is driven into a nonlinear operating regime.

with each other (frequency mixing), as shown by

$$b = c_0 + c_1 a + c_2 a^2 + c_3 a^3 + \dots, \quad (2.13)$$

$$\alpha = \beta = 2\pi\omega t, \quad (2.14)$$

$$a(t) = A \cos(\alpha), \quad (2.15)$$

$$\cos(\alpha) \cos(\beta) = \frac{1}{2} (\cos(\alpha + \beta) + \cos(\alpha - \beta)), \quad (2.16)$$

$$a^2(t) = \frac{1}{2} A^2 [\cos(2\pi(2\omega)t) + 1], \quad (2.17)$$

$$a^3(t) = \frac{1}{4} A^3 [\cos(2\pi(3\omega)t) + 3 \cos(2\pi\omega t)]. \quad (2.18)$$

Here, a and b are the incident and scattered waves for the device, c_i are coefficients of the device's nonlinear transfer function, and $a(t)$ is a wave in the time domain with amplitude A and frequency ω . For stimuli with a single frequency ($\alpha=\beta$, as above), integer multiples of that frequency will be scattered from the device (harmonics). For stimuli with multiple tones ($\alpha \neq \beta$), additional products from combinations of the incident tone frequencies will be scattered (intermodulation). If the nonlinear device is incident with a fixed bandwidth of frequencies, such as the case for communications signals, then sidebands will be produced around the harmonics of the oscillator frequencies. This effect can be troublesome in practical designs where the unwanted sidebands overlap with the useful microwave bandwidth, distorting the signal. For this reason, it is important for designers to be able to accurately measure and characterise this nonlinear effect. Fig. 5 shows example spectra of these effects.

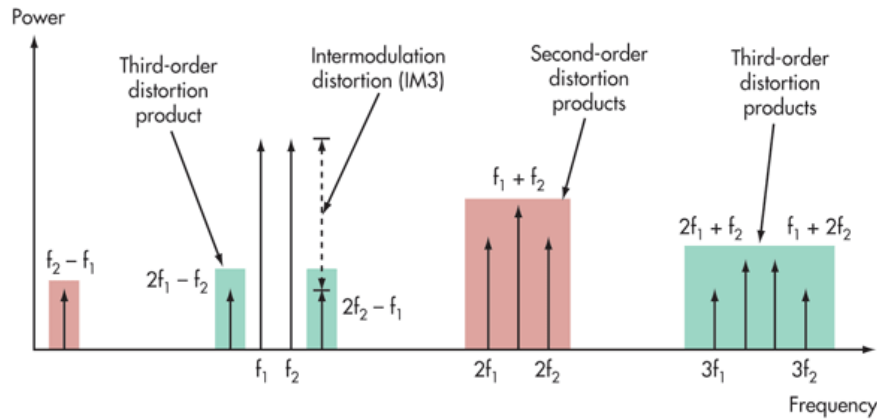


Figure 2.4: Intermodulation products from two tones within the cellular channel bandwidth f_1 and f_2 . The second order products, and upper third order products, can be easily filtered out. However, the lower third order products $2f_1 - f_2$ and $2f_2 - f_1$ are located within the channel bandwidth and interact with the useful data, increasing EVM and BER **Hall2013**.

3. The amplitude of scattered waves with multiple incident waves is dependent (nonlinearly) on the phase of the incident waves. In the linear regime the superposition principle prevents this, but now there is a nonlinear dependence which can have significant effects on the amplitude of scattered waves. Designers must consider this when building efficient nonlinear amplifiers, which leads to the practice of accurately terminating scattered harmonic frequencies at an optimum phase. This will be covered in more detail in chapter X when we discuss nonlinear device models.

The result of these differences is that the measurement requirements for nonlinear devices are considerably larger than for linear devices. The nonlinear dependencies on stimulus power and phase means that ratioed measurements no longer fully capture the device response, and absolute measurements of the magnitude and phase of both the incident and scattered waves is required. The production of scattered waves at frequencies different to those in the stimulus demands an additional dimension of measurements. In contrast to These complications must be met with changes to both the measurement system and the method of storing the results.

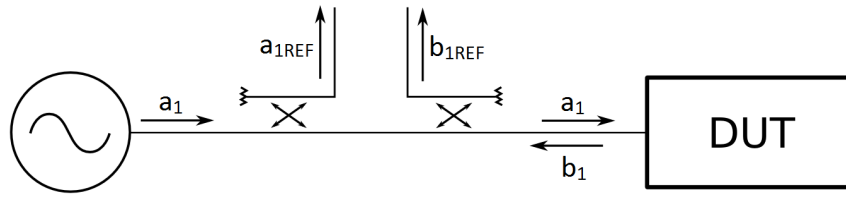


Figure 2.5: A one-port simple reflectometer. a_1 is the incident wave generated by the source, which is admitted to the DUT while also being sampled by the directional coupler and sent to the reference receiver via a_{1REF} . The reflected wave, b_1 , is also sampled by another directional coupler and sent to the test receiver as b_{1REF} , with the remaining power dissipated at the matched source.

2.4 Vector Network Analysers

To measure the incident and scattered waves for a DUT and calculate the s-parameters as in (5), a vector network analyser (VNA) is typically used. The VNA is a quintessential piece of RF and microwave instrumentation and is found in most if not all such laboratories. Due to the challenging nature of measurements at these frequencies, it is a complicated instrument with many internal parts. This section explains how the VNA functions and the procedures behind its calibration. For a good history of VNA architecture and product development please refer to [Teppati Camb, Dunsmore Wiley].

2.4.1 Architecture

The origin of the VNA lies in an early instrument called a reflectometer. Designed in 1947 by Parzen and Yalow [x], it became an invaluable tool for characterising transmission lines used in telecommunication systems. Shown in Figure x, the incident signal is generated by a swept signal source and passes through the directional coupler before arriving at the DUT. The voltage reflection coefficient of the DUT will cause an amount of incident power to be reflected, which passes back through the coupler before being absorbed by the source (which has very low reflection). The directional couplers allow the waves travelling between the source and the DUT to be sampled by complex receivers, filtering the two waves by their direction of travel thus allowing the incident and scattered waves to be separated for measurement.

The limitation of a single reflectometer is that it can only measure waves at one port of

a DUT, therefore preventing transmission measurements. By adding a second reflectometer and synchronising the stimuli and measurements, it is possible to measure all s-parameters of a two-port device. This is the fundamental structure of a VNA, and most variations consist of changing the number of sources or receivers to optimise the instrument for cost or performance. Many older designs use an economical single source which is switched between both ports, whereas now the price of sources has fallen, there are instruments available with two independent sources, which allows two-tone and some types of nonlinear measurements. These more versatile units often also expose more connections between internal components (e.g. the couplers and receivers) to allow the user to perform non-standard measurements or to add attenuation or preamplification for extreme stimulus powers. Modern VNAs also offer the option of measuring more than two ports, which are referred to as “multi-port” measurements. Several manufacturers offer four-port instruments which include four reflectometers (with usually two sources), although with external switching networks it is possible to expand this up to 48 ports [<http://www.microwavejournal.com/articles/21785-vector-network-analysis-with-up-to-48-ports>]. The basic block diagram of a modern two-port double-reflectometer VNA is shown in Fig. 9. To measure both stimulus conditions for the two-port S-parameter equations in (5), the sources alternate between delivering power and acting as a load for each measurement. As the source is swept the a and b waves for all ports are measured against frequency, from which the VNA software calculates the S-parameters. The receivers sampling the incident waves are known as the reference receivers and those sampling the scattered waves are called measurement or test receivers.

To perform S-parameter measurements using a VNA, the user must set both the frequency span and number of frequency points. They may also change settings of intermediate frequency bandwidth (IFBW) and numerical averaging, both of which reduce measurement noise by applying digital filtering but can consequently increase acquisition time. The user will then perform a calibration, which corrects for any response present in the measurement setup that is not caused by the DUT. When the system is calibrated physical ‘measurement planes’ are defined, where only effects of the signal path on the DUT side of the planes are incorporated in the measurement results. This is illustrated in Fig. 10. Once this step is complete, the VNA is ready for use. However, it is good practice to first check that calibration was successful by measuring some known devices (verification), or to use techniques such as ripple extraction (discussed in Chapter 4) to measure the residual uncertainty. This process characterises remaining error which the calibration failed to correct.

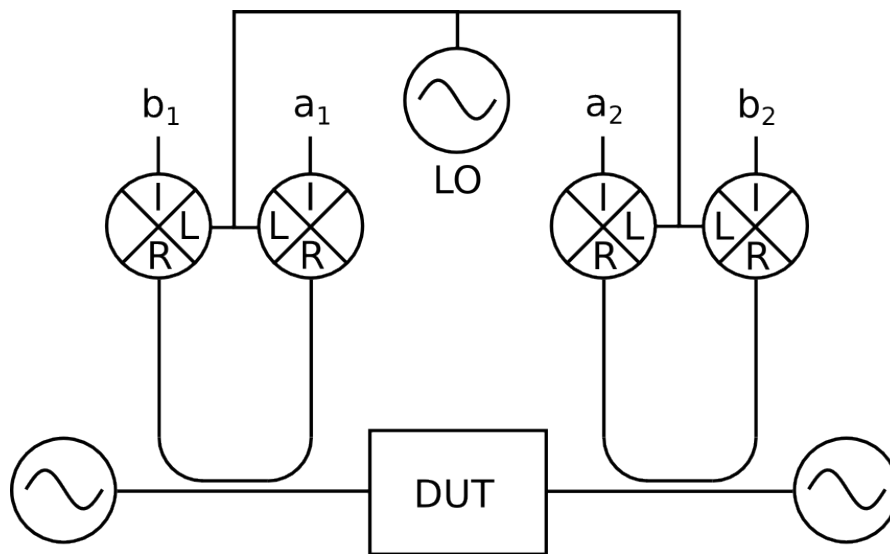


Figure 2.6: A modern two-source mixer-based VNA, which employs heterodyning to allow measurements at microwave frequencies. Two directional couplers are located between each source and the DUT and are connected back to back. These sample waves travelling in both directions and are connected to mixers which downconvert the microwave frequencies (R) into intermediate frequencies (I) which can be sampled by the complex receivers. The shared local oscillator (LO) feeding the mixers preserves phase coherence between the receivers. This configuration is known as a two-port double-reflectometer VNA. Figure adapted by author from **Root2013**.

2.4.2 Error Models

To remove the effect of the measurement setup from the measured device response, an error model is formed to capture the response of the measurement setup during calibration. These error models are stored in the memory of the VNA and are typically de-embedded from the measured device response before the results are presented to the user (although the raw measurements can still be obtained for separate post-processing). Because the measurement setup response is frequency dependent, the error model coefficients are characterised across the measurement bandwidth and are either applied at each measurement frequency or linearly interpolated.

2.4.2.1 One-Port Model

The classic one-port error model can be obtained through analysis of the signal flow diagram of a one-port VNA shown in Fig. 6. One can write the relationship between the measured (Γ_M) and absolute (Γ_A) reflection coefficients as

$$\Gamma_M = D + \frac{T\Gamma_A}{1 - M\Gamma_A}, \quad (2.19)$$

where D , M and T are error coefficients which capture the unwanted response of the measurement setup. For this model, the three coefficients each have a physical meaning as they are caused by separate physical effects (illustrated in Fig. 7):

- **Directivity** (D) is caused by the nonideal operation of the directional couplers used to separate the incident and reflected waves inside the VNA. In practice, some amount of incident wave will travel into the test receiver port (vica versa??), reducing the measured gain of the device under test.
- **Test port match** (M) results from the impedance of the VNA test port (either the original test port or the extended measurement plane including any cables or other components in the setup) being different from the characteristic impedance of the measurement, which is typically 50- Ω . This effect will cause some of the incident wave to be reflected at the test port which is not due to the device response.
- **Reflection tracking** (T) characterises the insertion loss of the couplers and other measurement components between the reference receiver and the test receiver.

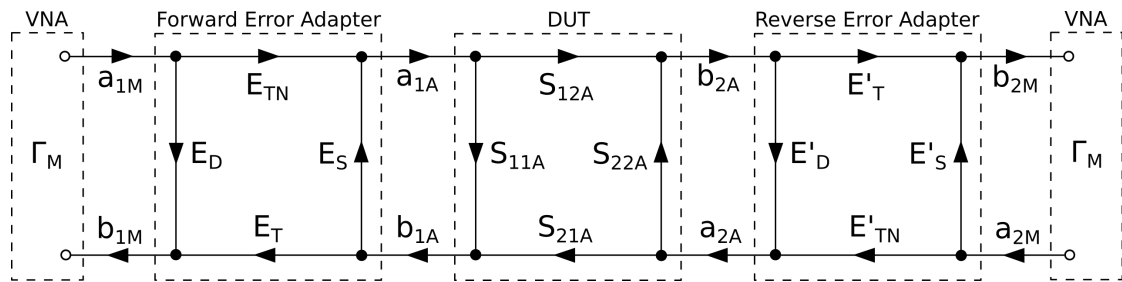


Figure 2.7: The 8-term error model for a two-port measurement. E_D , E_S , and E_T are the same as for the one-port model, except there are now sets of each for both ports. These extra terms account for different error values when the incident signal is sourced from each port. Additionally for the two-port case, a transmission term E_{TN} has been added for each direction.

2.4.2.2 8-Term Model

Devices with two or more ports require transmission measurements in addition to the reflection measurements which can be corrected using the one-port model. A popular two-port error model, the 8-term model, adds two transmission terms to the method used for the one-port model. This is shown in Fig. blah.

2.4.3 Calibration

2.5 Large Signal Network Analysers

2.5.1 Absolute 8-Term Error Model

2.5.2 Power Meter Calibration

2.5.3 Phase References

2.6 Conclusions

Testing, testing[1], [2].

References

- [1] L. Stant, P. Aaen, and N. Ridler, “Evaluating residual errors in waveguide network analysers 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.
- [2] —, “Comparing methods for evaluating measurement uncertainty given in the JCGM ‘evaluation of measurement data’ documents,” *Measurement*, vol. 94, pp. 847–851, Dec. 2016. DOI: 10.1016/j.measurement.2016.08.015.

3 Measurement Uncertainty

3.1 Introduction

A measurement is an observation of a physical effect or quantity which provides useful information. This information, through the ages, has been used to facilitate advancement of both scientific knowledge and industrial development - from the production of standardised stone blocks to build the pyramids of ancient Egypt, to the production of standardised car parts to build Henry Ford's Model T. In the scientific realm, advanced measurement techniques at laboratories such as CERN are used to convince the world that new subatomic particles exist.

To communicate information about a measurement, the recipient needs to be able to either make or imagine a similar observation to that of the original measurer (or metrologist). The simplest way of doing this is to provide the recipient with the same physical effect or quantity for which to make their own observation (if you require a new nut for a bolt from a hardware shop, you might intuitively take the bolt with you), however, this can be inconvenient or impractical with larger objects, or if the recipient is located far away. Instead, you might substitute a more portable representation. For example, if you were to measure the size of a doorway to see if a new piece of furniture may fit through it, you might cut a piece of string to the same length and use this as the representation of the width of the item. However, this approach is very wasteful and also impractical for many physical effects (temperature, flow, pressure).

A solution widely thought to have been first established in the 3rd or 4th Millennium BC (see Figure 3.1), is a system of units. In such a system, a discretised value of a quantity is standardised and knowledge of its value is disseminated to all people who wish to use it. Typically, a range of discrete values are chosen, such that the system of units can be conveniently used to represent all measurements. Knowledge of the discretised values is obtained from a primary standard which becomes the definition of the unit and is used to create copies of the standard which can be



Figure 3.1: Egyptian royal cubit rod of Maya (treasurer of King Tutankhamun) 1336–1327 BC. The cubit is thought to be the earliest attested standard measure of length, first used in the 3rd or 4th Millennium BC.

given to users of the unit system to perform measurements with. The most common method of performing measurements with a unit system is to use a standard to calibrate a measuring instrument, which can then be used to measure an arbitrary value of a quantity in the units defined by the standard.

The introduction of a regulated system of units enables commerce, as traded goods can be reliably valued between merchants across cities. This application is encountered by all citizens, and so there is a high demand for standards to be produced from the primary standard and physically distributed. It becomes impractical to create all standards by copying the primary standard directly (in some cases because the value of the primary standard is perturbed each time it is measured), and so a tiered organisational structure of standards is used. In this structure, there is a tier consisting of a small number of standards which are created directly from measurements of the primary standard, followed by subsequent tiers of larger numbers of standards which are derived from measurements of those in the previous tier. For any standard produced, it should be possible to trace the lineage back to a measurement of the primary standard. This is referred to as a traceability chain (see Figure 3.2) and it is a fundamental tenet of metrology. Measurements with a shorter traceability chain are considered more traceable than those with longer chains.

Today, the primary standards are maintained in most countries by a National Measurement Institute (NMI) and co-ordinated by the Bureau of International Weights and Measures (BIPM). To accommodate international trade and compatibility, a routine process of inter-comparisons is undertaken to ensure that the values of the primary standards between countries are in agreement.

Secondary standards are also kept by the NMIs and are used to reduce excessive wear to

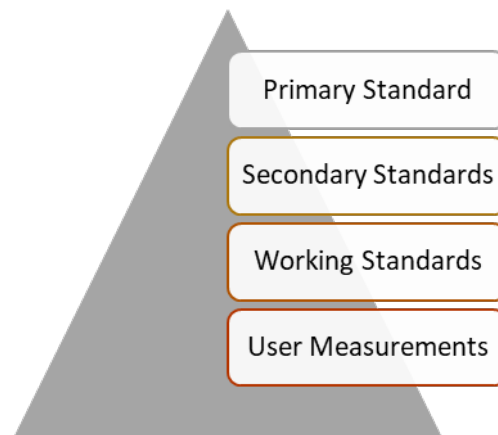


Figure 3.2: The traceability chain, where the pyramid shows the number of instances of standards in each tier. Secondary standards are held at NMIs and used to periodically calibrate working standards, which are sent by manufacturers and laboratories. User measurements are made using instruments calibrated with these working standards, so they number the greatest and are at the bottom of the traceability chain.

the primary standard caused by frequent measurements (and also to reduce bottlenecks caused by having a single standard). They are calibrated against the primary standard as infrequently as possible, again to reduce wear. Secondary standards are used by the NMI to characterise working standards which are sent to them by manufacturers and research institutes. Another important task of each NMI is to perform investigations to discover new and improved methods of measurement, which make use of secondary standards to better compare the accuracy of different methods.

Working standards are used, for example, by instrumentation manufacturers who may use them to calibrate their products before shipping to the customer, and more generally the standards can be used to calibrate test equipment to identify faulty products. Larger research institutes typically use working standards to recalibrate instrumentation prior to performing very sensitive measurements. To ensure that product specifications and scientific measurements are traceable and of high quality, accreditation services such as the United Kingdom Accreditation Service (UKAS) exist to certify manufacturers and laboratories that demonstrate good measurement practice and use traceable measurements [1].

The selection of quantities for which primary standards are kept is only a subset of those for which recognised units exist. This is because many units are derived quantities, where

their value can be obtained by calculation using definitions of other units. For example, the definition of the unit of resistance (R , ohms) can be derived from that of voltage (V , volts) and current (I , amperes), because $R = V/I$. The eight fundamental “base” units which make up the International System of Units (SI), are the metre, kilogram, second, ampere, kelvin, candela and mole. From these unit definitions, it is possible to define any other derived unit in use. NMIs will usually keep secondary standards of most derived quantities that users may wish to calibrate against, which are traceable to one or more primary standards of the base units. Although traditionally all primary standards were defined by physical artefacts (e.g. metallic weights, burning candles), these are being gradually replaced by definitions involving physical constants (e.g. Planck, Boltzmann), which do not degrade over time or use. The “Ninth SI Units” [2], a proposition recently accepted by the BIPM, covers the redefinition of four of the SI units (the ampere, the kilogram, the kelvin and the mole) which is scheduled for May 2019.

The crucial effect of traceability on measurements is the confidence in their results. Measurements with poor traceability (longer chains) will produce results which are likely to be less accurate than those with better traceability (shorter chains). The reason for this is measurement uncertainty, which will now be explained.

It is impossible to know the true value of a quantity being measured as many undesirable physical effects typically occur during the measurement process. These effects contribute error (an unwanted perturbation) to the measured value, causing a reduction in accuracy (the deviation of the measured value from the true value). Typical sources of error in measurement include thermal noise, imperfect calibration and drift of environmental conditions from those at which a measuring instrument was calibrated. In some cases, it is possible to quantify and correct for these errors, but there are often many sources (some of which contribute very small errors) which cannot be corrected for. This is because either the error cannot be quantified or the value of the error will change over the duration of the measurement process (random errors). Any source of error which cannot be removed from a measurement becomes a source of uncertainty, because the deviation of the measured value from the true value due to this source of error is uncertain. If it is possible to quantify the amount of uncertainty in a measurement, then a degree of confidence can be formed about its value. If every measurement has an associated uncertainty in its value, then any measurement involving the results of previous measurements will include uncertainty contributions from both measurements. Measurements with good traceability involve fewer sources of uncertainty than those with poor traceability, leading to a higher degree of confidence in the former. It is because of this fact that NMIs strive to reduce the uncertainties

in their primary standard definitions, which in turn reduces the uncertainty in all traceable measurements.

Because it is impossible to know the amount of error in a source of uncertainty, probability and statistical theories are used to instead describe the amount of uncertainty associated with it. By the nature of these theories there are often several methods which can be used to obtain a result, which sometimes provide different values. To ensure consistency and portability of uncertainty definitions, measurement guides were created in each industry and area of science, which specialised in processing the results of typical measurements. In addition, different guides were produced depending on the level of accuracy required - as more accurate measurements often require more effort to complete. Although this practice allowed suitable measurement comparisons within each field (e.g. chemistry, mechanical engineering), ambiguities still existed in uncertainty definitions between fields. To address this, a landmark document was published in 1993 by the International Organisation for Standardisation (ISO), the Guide to the Expression of Uncertainty in Measurement (GUM) **GUM'1993**. This document was the work of representatives from seven international organisations: the BIPM, the International Organisation of Legal Metrology (OIML), the International Electrotechnical Commission (IEC), the ISO, the International Federation of Clinical Chemistry and Laboratory Medicine (IFCC), the International Union of Pure and Applied Chemistry (IUPAC), and the International Union of Pure and Applied Physics (IUPAP). The GUM, updated in 2008 **GUM'2008**, is still used today as a reference for the evaluation of measurement uncertainty in many laboratories and industries across the world. The seven original organisations which wrote the GUM, together with the International Laboratory Accreditation Cooperation (ILAC, of which UKAS is a member), form the Joint Committee for Guides in Metrology (JCGM), who maintain the GUM and subsequent additional documents. These additional documents consist of the International Vocabulary of Metrology (VIM) [5] and two supplements to the GUM **GUM'S1**, **GUM'S2**: Supplement 1 covers the use of a Monte Carlo method **Metropolis'1949** in uncertainty evaluation; Supplement 2 is used where more than one quantity is measured at the same time (multivariate).

Throughout this dissertation, the methodologies presented in the GUM will be used. The international authority of the guide, developed by seven international organisations (including the two global standardisation bodies IEC and ISO), gives strong motivation to use it as a basis for a framework to evaluate uncertainty in measurement.

This Chapter describes the evaluation of uncertainty prescribed in the GUM and highlights an inconsistency in the current version of the GUM and associated documents (which can have

a profound effect on electromagnetic measurements).

3.2 The Measurement Process

In contrast to basic evaluations of uncertainty, where only repeat measurements of the quantity of interest are analysed, the GUM prescribes a more rigorous approach, which defines a mathematical model of the measurement process (measurement model) and propagates uncertainty through that model to the result (measurands). This allows any uncertainties from previous measurements, including those involving standards in the traceability chain, to be included in the result. The measurement model can be simple, such as measuring resistance using input quantities of voltage and current, or complicated and multivariate, requiring many input quantities and producing many output quantities. In some cases, the measurement model may not be known and can be defined as a black box, but this has certain limitations discussed later with Monte Carlo methods.

The GUM defines a process that is to be followed when evaluating uncertainty in measurement. It consists of the following steps:

1. Modelling the measurement.
2. Evaluating standard uncertainty of input quantities.
3. Determining combined standard uncertainty of the measurands.
4. Determining expanded uncertainty of the measurands.

where standard uncertainty is an uncertainty expressed as a standard deviation and expanded uncertainty is used to define a coverage interval encompassing a large fraction of the distribution of values that could reasonably be attributed to the measurand.

3.2.1 Modelling the Measurement

The VIM document [5] defines a measurement model¹ as “a mathematical relation among all quantities known to be involved in a measurement”. In many cases, where an explicit relation can be written, it is possible to further define a measurement function. We can represent this

¹The definition of model used in “measurement model” is different to that used when describing models of amplifiers seen elsewhere in this dissertation.

generally as a set of measurands Y having a functional relationship, $f(\cdot)$, depending on N input quantities X_1, X_2, \dots, X_N :

$$Y = f(X_1, X_2, \dots, X_N) \quad (3.1)$$

The estimates of the measurands \bar{Y} can be found by evaluating the measurement model using the estimates of each input quantity $\bar{x}_1, \bar{x}_2, \dots, \bar{x}_N$:

$$\bar{Y} = f(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_N) \quad (3.2)$$

Each input quantity could either be observed during the present measurement, a result from a previous measurement, or another source of information such as a datasheet or specification. An example of a measurement model could be for a temperature measurement, where the input quantities would include the value observed from the meter, the previously measured values of two calibration temperatures, and the assumed values of those calibration temperatures. Using this method, uncertainty from the calibration can be included in the evaluation. This is especially true for uncertainties caused by systematic errors, which do not vary during the measurement process and cannot be evaluated purely by performing repeat measurements.

3.2.2 Evaluating Standard Uncertainty of Input Quantities

Sources, or components, of uncertainty in measurement can be divided into two categories: Category A uncertainty components are those that are evaluated using statistical analysis of a series of observations (i.e. repeats); Category B components are those that are evaluated using other means, for example using information from datasheets.

The GUM presents methods that include the use of both Bayesian and classical probabilistic methods to evaluate the uncertainty in the input quantities for a measurement model. In particular, classical methods **Neyman'1937** are used for the treatment of Category A uncertainty components and Bayesian methods **Gelman'2013** are used for the treatment of Category B uncertainty components. An informative discussion on these types of method can be found in **White'2016**. Since the publication of the GUM, some authors have stated (for example, in **Kacker'2006**, **Kacker'2005**, **Kacker'2003**, **Bich'2014**) that this combination of different probabilistic methods (i.e., Bayesian and classical) represents an inconsistency in the GUM methodology for evaluating measurement uncertainty. The author has published a paper considering the effects of this inconsistency on electromagnetic wave measurements at radio frequencies

Stant'2016, which forms the basis for this section of the chapter.

The supplements to the GUM **GUM'S1**, **GUM'S2** resolve the above-mentioned inconsistency by introducing a method for treating the Category A uncertainties that follows a Bayesian approach **Elster'2007**. Therefore, the two supplements no longer contain the inconsistency found in the original GUM document. However, as a consequence of this change, there is now inconsistency between the method used to evaluate uncertainty described in the GUM and that described in the two supplements. In many situations, these different methods do not have a significant impact on the overall uncertainty that is evaluated. For situations where a considerable number of input quantities are observed simultaneously, the two different approaches can produce significantly different values of uncertainty. Such situations often occur in the area of high-frequency electromagnetic metrology, which is the topic of this dissertation.

3.2.2.1 Category A Evaluation

GUM Method The classical statistical technique **Neyman'1937** applied to Category A uncertainties in the current GUM is based on a series of observations of a randomly varying input quantity. After n observations x_1, x_2, \dots, x_n , the best available estimate (arithmetic mean of measured values), \bar{x} , and standard deviation, s , of a randomly varying input quantity, X , is written as

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, \quad (3.3)$$

$$s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2} \quad (3.4)$$

respectively, where x_i is the result of the i th observation. Importantly, a minimum of two observations must be made ($n = 2$) in order for \bar{x} and s to be defined. The standard uncertainty of the best estimate of X , $u(\bar{x})_{\text{GUM}}$ can be found by dividing s by the square root of the number of observations:

$$u(\bar{x})_{\text{GUM}} = \frac{s}{\sqrt{n}} \quad (3.5)$$

If there are correlated (mutually dependent) input quantities present in the measurement model, the covariances of each pair of input quantities must also be calculated before the propagation of uncertainty. Both the standard uncertainties and the covariances for N input quantities

can be represented in a symmetric $(N \times N)$ matrix containing the variance of each quantity (s^2) along the diagonal and the covariance between x_i and x_j in the i, j th element. This is called the “uncertainty matrix” in the GUM and the “measurement covariance matrix” in the GUM Supplement 2. An example given in the GUM and described later in this Chapter, demonstrates this scenario using the example of a simultaneous measurement of resistance and reactance with voltage, current and phase as correlated input quantities **GUM’2008**. Once the uncertainties of the input quantities have been evaluated, they are propagated through the measurement model. This requires the sensitivities of the measurand to each input quantity to be calculated to at least a first order approximation (i.e. the partial derivatives of the measurement model known). The estimates of the input quantities are used in the measurement model to obtain the estimate of the measurand. The variances and covariances of the input quantities are combined with the sensitivity coefficients in order to obtain the variance of the measurand. The combined standard uncertainty of the measurand is equal to the positive square root of this value. The result of the measurement process is then presented as the measurand estimate and combined standard uncertainty. Alternatively, the combined standard uncertainty can be multiplied by a positive factor to form an expanded measurement uncertainty. From this value, a coverage interval can be derived which states a particular probability that the true value of the measurand is within that range.

GUM Supplement Method Both GUM supplements (GUM-S1/S2) **GUM’S1**, **GUM’S2** use a Bayesian approach **Klaunberg’2012** to assign a probability density function (PDF) to all input quantities. This approach results in the choice of a t -distribution to characterize Category A input quantities, in contrast to the Gaussian distribution used in the GUM **GUM’S2**. Of particular relevance is the inclusion of the degrees-of-freedom parameter, ν , in the definition of the standard uncertainty and covariances of a t -distribution. Whereas for the Gaussian distribution ν is used as a measure of reliability of the standard uncertainty, it is explicitly required when using the t -distribution in order to obtain the standard uncertainty, $u(\bar{x})_{\text{SUPP}}$:

$$u(\bar{x})_{\text{SUPP}} = \frac{s}{\sqrt{n}} \times \sqrt{\frac{\nu}{\nu - 2}}, \quad (3.6)$$

where $\nu = n - N$, with n being the number of observations and N being the number of input quantities. In the GUM-S1 only a univariate t -distribution is offered, which represents $N = 1$ input quantities. For this case (3.6) can be rewritten as:

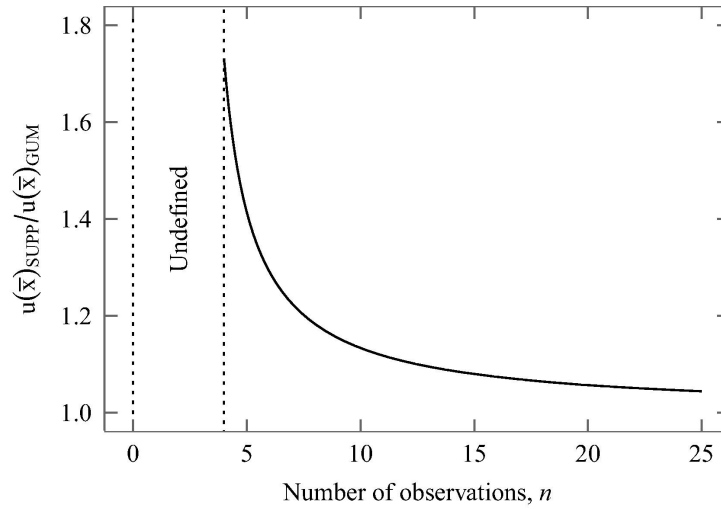


Figure 3.3: Scaling factor to convert from a GUM standard uncertainty to a GUM Supplement.

$$u(\bar{x})_{\text{SUPP}} = \frac{s}{\sqrt{n}} \times \sqrt{\frac{n-1}{n-3}}. \quad (3.7)$$

Equation 3.7 is undefined if n is less than four, in which case the standard uncertainty cannot be calculated for a single input quantity according to the guidance given in the GUM-S1 (and the GUM-S2). Figure 3.3 illustrates the ratio between the standard uncertainty values calculated for different numbers of observations of a single Category A input quantity using the GUM and the GUM-S1/S2 approaches. It can be seen that when $n = 4$, $u(\bar{x})_{\text{SUPP}} = \sqrt{3} \times u(\bar{x})_{\text{GUM}}$, and as the number of observations increases the results from both approaches converge: If n tends to infinity, the t -distribution tends towards a Gaussian distribution. However, most commercial laboratories would avoid making large numbers of measurements as this is often time-consuming and therefore expensive.

For measurements involving multiple input quantities, such as the measurement of a vector quantity, a multivariate/joint distribution should be used as suggested in the GUM-S2. The variances and covariances between all pairs of input quantities are obtained using a matrix form of (3.6) (**GUM'S2**):

$$\mathbf{V}(\mathbf{X}) = \frac{\nu}{(\nu - 2)} \frac{\mathbf{S}(\mathbf{X})}{n} = \frac{1}{n(n - N - 2)} \sum_{i=1}^n (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^\top \quad (3.8)$$

$$\mathbf{S}(\mathbf{X}) = \frac{1}{\nu} \sum_{i=1}^n (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^\top \quad (3.9)$$

$$\mathbf{V}(\mathbf{X}) = \begin{bmatrix} u(\mathbf{x}_1)^2 & u(\mathbf{x}_1, \mathbf{x}_2) & \dots & u(\mathbf{x}_1, \mathbf{x}_n) \\ u(\mathbf{x}_2, \mathbf{x}_1) & u(\mathbf{x}_2)^2 & \dots & u(\mathbf{x}_2, \mathbf{x}_n) \\ \vdots & \vdots & \ddots & \vdots \\ u(\mathbf{x}_n, \mathbf{x}_1) & u(\mathbf{x}_n, \mathbf{x}_2) & \dots & u(\mathbf{x}_n)^2 \end{bmatrix} \quad (3.10)$$

where $\mathbf{U}_\mathbf{X}$ is the uncertainty matrix, \mathbf{x}_i is a sample from the array of vectors containing input quantity indications and $\bar{\mathbf{x}}$ is the arithmetic mean of that array. For this multivariate case, the minimum value of n will increase linearly with N , such that the standard uncertainty is undefined unless $n > N + 2$.

Comparison of GUM and GUM Supplements approach using example H.2/9.4 Both the GUM and the GUM-S2 provide an identical example which can be used to compare the different standard uncertainties. The example is a simultaneous measurement of resistance and reactance, which uses a measurement model with multiple input quantities and multiple output quantities (measurands). The input quantities are voltage V , current, I , and phase, ϕ , and the measurands are resistance R , reactance, X , and impedance, Z . The measurement model is defined as:

$$R = \frac{V}{I} \cos \theta, \quad X = \frac{V}{I} \sin \theta, \quad Z = \frac{V}{I} \quad (3.11)$$

Six sets of observations [5] ($n = 6$) of V ; I ; ϕ are obtained independently by measurement. The version of this example given in the GUM uses only $n = 5$ sets, but one additional set of values of V ; I ; ϕ has been added for the GUM-S2 example to allow (3.6) to be defined for $N = 3$ input quantities, a condition which was explained at the end of the previous section. These values, together with their arithmetic means and standard uncertainties as calculated from the two approaches using (3.5) and the matrix form of (3.6) (which is applicable to measurements involving multiple input quantities), are presented in Table 3.1. The ratios of the standard uncertainties from each approach is also included in the table, which are identical for all these input quantities due to their dependence only on n and N , which are also equal for all these input quantities (e.g. when $n = 6$ and $N = 3$, $\sqrt{(\nu/(\nu - 2))} = \sqrt{((n - N)/(n - N - 2))} = \sqrt{3}$. This

Value	V/V	I/A	ϕ/rad
x_1	5.007	19.663	1.0456
x_2	4.994	19.639	1.0438
x_3	5.005	19.640	1.0468
x_4	4.990	19.685	1.0428
x_5	4.999	19.678	1.0433
x_6	4.999	19.661	1.0445
\bar{x}	4.9990	19.6610	1.04446
$u(\bar{x})_{\text{GUM}}$	0.0026	0.0077	0.00061
$u(\bar{x})_{\text{SUPP}}$	0.0045	0.0134	0.0011
$\frac{u(\bar{x})_{\text{GUM}}}{u(\bar{x})_{\text{SUPP}}}$	1.732	1.732	1.732

Table 3.1: The indication values from the example “Simultaneous Resistance and Reactance Measurement” and their statistical properties as evaluated by the approaches given in **GUM’2008** and **GUM’S2**.

explains why standard uncertainties evaluated with Category A methods using the minimum number of observations following the GUM-S1/S2 approach are always 1.732 times larger than the standard uncertainties calculated following the GUM approach.

This difference in the input quantity uncertainties calculated from the two approaches propagates through the measurement model and therefore significantly affects the combined standard uncertainties of the measurands. Table 3.2 presents the combined standard uncertainties of the measurands for the described example as evaluated by both approaches, together with a ratio of the uncertainty values. For all three measurands the combined standard uncertainty calculated using the GUM-S1/S2 method is more than double the equivalent values calculated using the GUM method. For other measurement models with higher sensitivities to the input quantities, this difference could be even greater.

Comparison of GUM and GUM Supplements approach using microwave scattering parameters example High-frequency electromagnetic metrology often involves using multiple complex-valued quantities. Common input quantities for this type of measurement, measured using instruments such as vector network analysers (VNA), are scattering parameters (S-parameters), as described in Chapter 2. Because each S-parameter is a complex-valued quan-

Method	$u(R)/\Omega$	$u(X)/\Omega$	$u(Z)/\Omega$
GUM	0.058	0.241	0.193
GUM-S2	0.130	0.540	0.431
$\frac{\text{GUM-S2}}{\text{GUM}}$	2.241	2.241	2.233

Table 3.2: A comparison of the results obtained for the example “Simultaneous Resistance and Reactance Measurement” using the approaches given in **GUM’2008** and **GUM’S2**.

Ports, m	Input quantities, N	Required minimum number of repeat observations, n , for $u(\bar{x})_{SUPP}$ to be defined	$\frac{u(\bar{x})_{GUM}}{u(\bar{x})_{SUPP}}$
1	2	5	1.732
2	8	11	1.732
3	18	21	1.732
4	32	35	1.732
\vdots	\vdots	\vdots	1.732
8	128	131	1.732

Table 3.3: The difference in standard uncertainties obtained using the GUM ($u(\bar{x})_{GUM}$) and the GUM-S1/S2 ($u(\bar{x})_{SUPP}$) approaches to measure a full set of scattering parameters for microwave devices with various numbers of ports, m . Each device has $2m^2$ input quantities, N , and requires a minimum of $N + 3$ repeat observations, n , in order for $u(\bar{x})_{SUPP}$ to be defined.

tity ($S = (S_{\text{Re}}, S_{\text{Im}})$), there are $2m^2$ input quantities required in a measurement model for the complete response of an m port device. All these quantities are correlated, so a multivariate distribution should be used to represent them. It has been shown previously that for a Category A evaluation of uncertainty, both the number of repeat observations and the number of input quantities have a significant effect on the difference in uncertainty as calculated from the two approaches presented in the GUM and the GUM-S1/S2. Table 3.3 shows the ratio of uncertainties calculated from both approaches when applied to a measurement using scattering parameters obtained from the minimum number of repeat observations, n , for devices with m ports.

It can be seen that for devices with multiple ports, n can become large in order for 3.6 to be defined and calculate the standard uncertainty. It is often the case that the user will not always

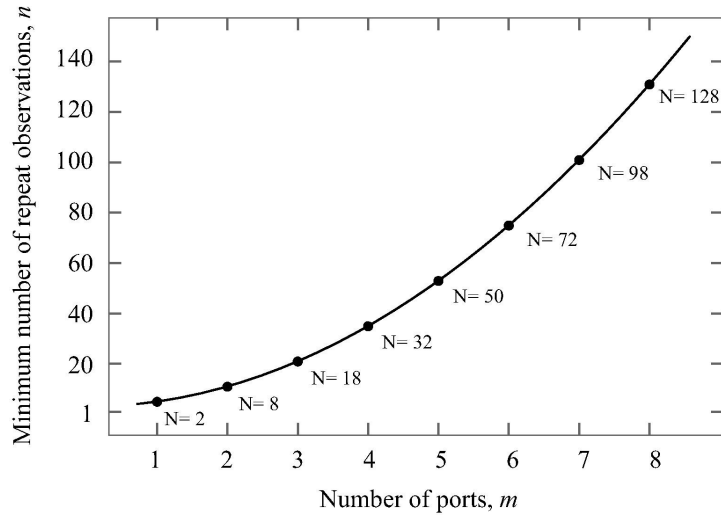


Figure 3.4: The minimum number of observations, n , required to calculate the standard uncertainty of a full set of S-parameters for a microwave device with m ports using the GUM-S2 approach. The number of input quantities, N , for each device is also shown.

have the time or resources available to perform such a quantity of measurements. In microwave measurement environments, connections are typically made by hand using coaxial connectors. A typical measurement may include a Category A evaluation of uncertainty due to connection repeatability. Considering the specific example of a 4-port device, this requirement would result in the need for a minimum of $35 \times 4 = 140$ repeat coaxial connections to be made in order to perform a Category A evaluation of the standard uncertainty using the GUM-S1/S2 approach. By contrast, the classical approach used in the GUM is defined with just 2 repeat observations, which would require only $2 \times 4 = 8$ repeat coaxial connections to be made. Figure 3.4 shows the minimum number of repeat observations required when using the GUM-S1/S2 approach, n , in order to be able to calculate a Category A evaluation of the standard uncertainty of a full set of S-parameters for a microwave device with m ports. In all cases, the standard uncertainty obtained using the GUM-S1/S2 approach is approximately 1.7 times larger than that obtained using the GUM approach.

Discussion The inconsistency of the approaches used in the GUM and its supplements to calculate the standard uncertainty of Category A input quantities of a measurement has two noticeable consequences:

1. There can be a large difference in the standard uncertainties reported by each approach. It is not straightforward to decide which is the correct approach to use, however the GUM approach is likely to be more attractive to commercial laboratories and test engineers since this leads to achieving smaller uncertainties in their results.
2. For situations involving multiple Category A input quantities, the Bayesian approach introduced in the GUM-S1/S2 can require a large number of observations before the standard uncertainty can be defined. Although the standard uncertainty calculated using the GUM approach will become less reliable with fewer observations, it is still possible to obtain a result with only two observations of any number of input quantities. In a commercial laboratory the additional measurements required by the GUM-S1/S2 approach can be impractical, with many laboratories typically using only two or three measurements per device following the GUM approach. For a single input quantity this would require a potential doubling of the number of observations and therefore the test duration, which would either slow throughput or require more test stations to be added. If implemented, the additional time or financial investment would then produce uncertainties that are significantly larger than those obtained using the GUM approach.

This inconsistency is yet to be resolved, and the draft of an updated GUM which replaced much of the remaining classical approach with Bayesian techniques received many poor reviews when circulated for discussion. Work is now being carried out to find solutions to the issues raised by converting to a fully Bayesian GUM. Specific to the example presented in this Section, an article was recently published which offers a way to use Bayesian statistics to evaluate uncertainty in Category A input quantities with $n \geq 2$ repeat observations, which is the same number required by the classical approach **Cox'2017**.

For the work in this thesis, which is based on multivariate electromagnetic measurement problems, the GUM approach (instead of the Supplement 2 approach) is used. In addition, an existing software framework, introduced later, which is included as part of the complete framework presented in this work, also uses the GUM approach for processing Category A uncertainty components.

3.2.2.2 Category B Evaluation

Category B uncertainty components are those which have not been obtained by repeated measurements. Possible sources include previous measurement data, experience or knowledge of

relevant materials and instruments, manufacturer's specifications, data provided by calibration and other certificates and reference data from handbooks.

Values obtained from these sources will typically be an estimate accompanied by either a standard uncertainty or an expanded uncertainty. The latter can be converted to a standard uncertainty, the process of which is described in Section 3.6. Category B uncertainty components are not restricted to Gaussian or t-distributions, and could for example be normal (rectangular), beta, or Cauchy distributions. Unless the combined standard uncertainty is determined via a Monte Carlo method, as explained in the following section, the standard uncertainty must be known for the value to be used as an input quantity.

3.2.3 Evaluating Combined Standard Uncertainty

In order to determine the standard uncertainty of the measurand (the combined standard uncertainty) the uncertainties of the input quantities must be propagated through the measurement model. The GUM offers several methods to achieve this, which will be described in this section.

3.2.3.1 Monte Carlo Methods

Supplement 1 of the GUM **GUM'S1** covers the use of a Monte Carlo technique to determine combined standard uncertainty in the measurand. The Monte Carlo technique has three important benefits for the propagation of uncertainty:

1. The measurement model does not need to be known explicitly. In some cases, the algorithm used to obtain a measurement result is proprietary and cannot be made available to the metrologist. Alternatively, the measurement model may be very complicated or involve numerical solutions which cannot be differentiated as required by other propagation methods.
2. Full knowledge of the probability distributions of the input quantities are used and preserved through the uncertainty propagation. Because the input quantity distributions are sampled directly, the complete probability distribution of the measurand can be obtained (see Figure 3.5). This can be very useful when more exotic distributions such as u-shaped distributions are used for input quantities, or if the measurement model is strongly nonlinear, when one cannot make assumptions about the probability distribution of the measurand.

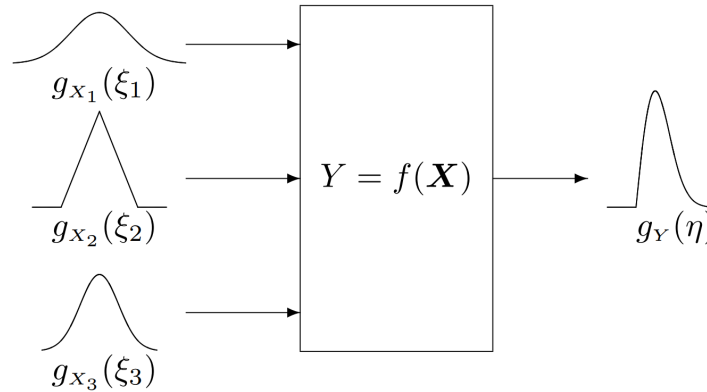


Figure 3.5: An illustration of the propagation of distributions from three input quantities g_{x_1} , g_{x_2} , g_{x_3} , through the measurement model, Y , to the measurand, g_Y **GUM'S1**.

3. The uncertainty propagation preserves nonlinearities in the measurement model. Alternative propagation methods presented in the GUM cause the measurement model to be linearised around the estimate. In most cases where a nonlinear measurement model is used, however, the uncertainty values are sufficiently small that a linear approximation is valid **GUM'2008**. Often, an initial Monte Carlo propagation is used to validate this assumption.
4. All correlations between input quantities are preserved. For many measurements involving multiple input quantities (especially in electromagnetic measurements), the uncertainties of one or more input quantities may be correlated. This means that when the value of one quantity changes, it affects the values of others. This can both increase or decrease the combined standard uncertainty in the measurand significantly. Chapter 4 will discuss the impact of correlations on VNA measurements.

The primary disadvantage of Monte Carlo methods is the time required to process them. For an accurate evaluation of uncertainty, the number of samples must be sufficiently large. Generally, the GUM recommends 106 samples for a 95% coverage interval accurate to one or two significant digits **GUM'S1**. The number of samples increases with the size of the desired coverage interval of similar accuracy. For many measurements today, the processing power of modern computers is sufficient for the duration of uncertainty propagations using Monte Carlo methods to be acceptable. However, in situations where the measurement model is very time-consuming to process, or where the uncertainty evaluation must be very fast, linear propagation

techniques may be preferred. A detailed explanation of the steps involved in performing a Monte Carlo propagation can be found in Section 7 of **GUM'S1**.

3.2.3.2 Law of Propagation of Uncertainty

The primary propagation method presented in the GUM is the Law of Propagation of Uncertainty (LPU). This method uses first-order derivatives of the measurement model, together with the variances (and co-variances) of the input quantities, to determine a value for the combined standard uncertainty. The use of first-order derivatives means that the measurement model is linearised, which in many applications can be a valid assumption. The LPU provides different equations for combining independent (uncorrelated) and correlated input quantities. For independent input quantities,

$$u_c^2(\bar{y}) = \sum_{i=1}^N \left(\frac{\partial f}{\partial x_i} \right)^2 u^2(\bar{x}_i), \quad (3.12)$$

where $u_c^2(\bar{y})$, the combined variance, is the square of the combined standard uncertainty, f is the function describing the measurement model, and x_i is an input quantity with variance $u^2(\bar{x}_i)$. For correlated input quantities (whose covariances always form a symmetric and positive semi-definite matrix),

$$u_c^2(\bar{y}) = \sum_{i=1}^N \sum_{j=1}^N \frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_j} u(\bar{x}_i, \bar{x}_j) = \sum_{i=1}^N \left(\frac{\partial f}{\partial x_i} \right)^2 u^2(\bar{x}_i) + 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^N \frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_j} u(\bar{x}_i, \bar{x}_j). \quad (3.13)$$

Supplement 2 to the GUM offers a matrix formulation of (3.13) **GUM'S2**, which handles the multivariate case where multiple measurands are encountered. If the covariance matrix of dimension $N \times N$ associated with $\bar{\mathbf{x}}$ is

$$\mathbf{U}(\bar{\mathbf{x}}) = \begin{bmatrix} u(\bar{\mathbf{x}}_1, \bar{\mathbf{x}}_1) & \dots & u(\bar{\mathbf{x}}_1, \bar{\mathbf{x}}_n) \\ \vdots & \ddots & \vdots \\ u(\bar{\mathbf{x}}_n, \bar{\mathbf{x}}_1) & \dots & u(\bar{\mathbf{x}}_n, \bar{\mathbf{x}}_n) \end{bmatrix} \quad (3.14)$$

the covariance matrix of dimension $m \times m$ associated with $\bar{\mathbf{y}}$ is

$$\mathbf{U}(\bar{\mathbf{y}}) = \begin{bmatrix} u(\bar{\mathbf{y}}_1, \bar{\mathbf{y}}_1) & \dots & u(\bar{\mathbf{y}}_1, \bar{\mathbf{y}}_n) \\ \vdots & \ddots & \vdots \\ u(\bar{\mathbf{y}}_n, \bar{\mathbf{y}}_1) & \dots & u(\bar{\mathbf{y}}_n, \bar{\mathbf{y}}_n) \end{bmatrix} \quad (3.15)$$

and the sensitivity matrix $\mathbf{C}_{\bar{\mathbf{x}}}$ of dimension $m \times N$ containing the first-order partial derivatives of the measurement model to each input quantity (the Jacobian of the measurement model) is given by evaluating

$$\mathbf{C}_{\bar{\mathbf{x}}} = \begin{bmatrix} \frac{\partial f_1}{\partial X_1} & \cdots & \frac{\partial f_1}{\partial X_N} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial X_1} & \cdots & \frac{\partial f_m}{\partial X_N} \end{bmatrix} \quad (3.16)$$

at $\mathbf{X} = \bar{\mathbf{x}}$, then $\mathbf{U}_{\bar{\mathbf{y}}}$ is given by

$$\mathbf{U}_{\bar{\mathbf{y}}} = \mathbf{C}_{\bar{\mathbf{x}}} \mathbf{U}_{\bar{\mathbf{x}}} \mathbf{C}_{\bar{\mathbf{x}}}^\top \quad (3.17)$$

The LPU does not provide any information about the shape of the probability distribution of $\mathbf{U}_{\bar{\mathbf{y}}}$ or its components. The results of the measurement are obtained from the estimates and the combined standard uncertainties – the positive square roots of the diagonal terms of the covariance matrix.

3.2.3.3 Finite Difference Methods

Included in the definition of the LPU is an alternative method of determining the sensitivity coefficients in (3.16), without the need to know the measurement model, f , explicitly. This technique can be described as a finite difference method and involves measuring the change in Y while varying a particular X_i and holding all other input quantities constant. This is often used when there may not be a model available for a particular process but a rudimentary uncertainty analysis is required, or if the computational power is available to recompute the finite differences when the estimates change. Typically, the sum of the estimate and the standard uncertainty of each input quantity $\bar{x}_i + u(\bar{x}_i)$ is used, although a more rigorous version also includes the standard uncertainty subtracted from the estimate $\bar{x}_i - u(\bar{x}_i)$ to check for asymmetry. Because only two points are used to solve for each sensitivity coefficient (the estimate and the estimate plus standard uncertainty), this uncertainty propagation also linearises the measurement model.

If all of the input quantities are considered independent and the standard uncertainty was chosen as the value with which to perturb the input quantities, then by subtracting the estimate of the measurand from each sample and adding the results in quadrature, the combined standard uncertainty in the measurand can be obtained:

$$u_c(\bar{y}) = \sqrt{\sum_{i=1}^N ((\bar{x}_i + u(\bar{x}_i)) - \bar{x}_i)^2} \quad (3.18)$$

3.2.4 Expanded Uncertainty and Coverage Intervals

Although it is recommended to express a result with combined standard uncertainty $u_c(\bar{y})$, it is often required, especially in safety critical applications, for the uncertainty to encompass a larger fraction of the distribution of values that could reasonably be attributed to the measurand. An expanded uncertainty, U , is instead used and is related to the combined uncertainty by $U = k u_c(\bar{y})$ **GUM'2008**. The multiplying factor, k , is termed the coverage factor and is typically in the range 2 to 3, often either of those two integer values and is defined by specifications or standards relating to the application. Using expanded uncertainties, the result can be expressed as $\bar{y} \pm U$, which is a popular format for datasheets and specifications.

To obtain a coverage factor that states a probability (e.g. 95%) that the true value of a measurand is within the associated interval is not straightforward, and depends on the probability distribution of the measurand. If all input quantities are Category A uncertainty components and the measurement model is linear, then the measurand distribution can be assumed to be Gaussian. In this case, the coverage interval is known as a confidence interval and can be given as a percentage by $\text{erf}(z/\sqrt{2}) \times 100$, where $\text{erf}(x)$ is the Gauss error function of x .

In situations where the above conditions cannot be met, a level of confidence can be obtained by calculating the effective degrees of freedom ν_{eff} of the distribution of the measurand. This process is explained in Annex G of [5]. For Monte Carlo propagations with sufficient samples, the confidence interval can be found by analysing the distribution of the measurand and obtaining the deviation from the estimated value which encompasses the desired percentage of samples (e.g. 95%).

3.3 Sensitivity Analysis

A benefit of propagating uncertainties through the measurement model is that an analysis of the sensitivity of the measurands to each input quantity can be performed. The sensitivity coefficients obtained from the measurement model can either be compared directly or multiplied by the standard uncertainty of the respective input quantity, in order to obtain an uncertainty figure for the measurand which can be compared with those calculated for other input quantities. This

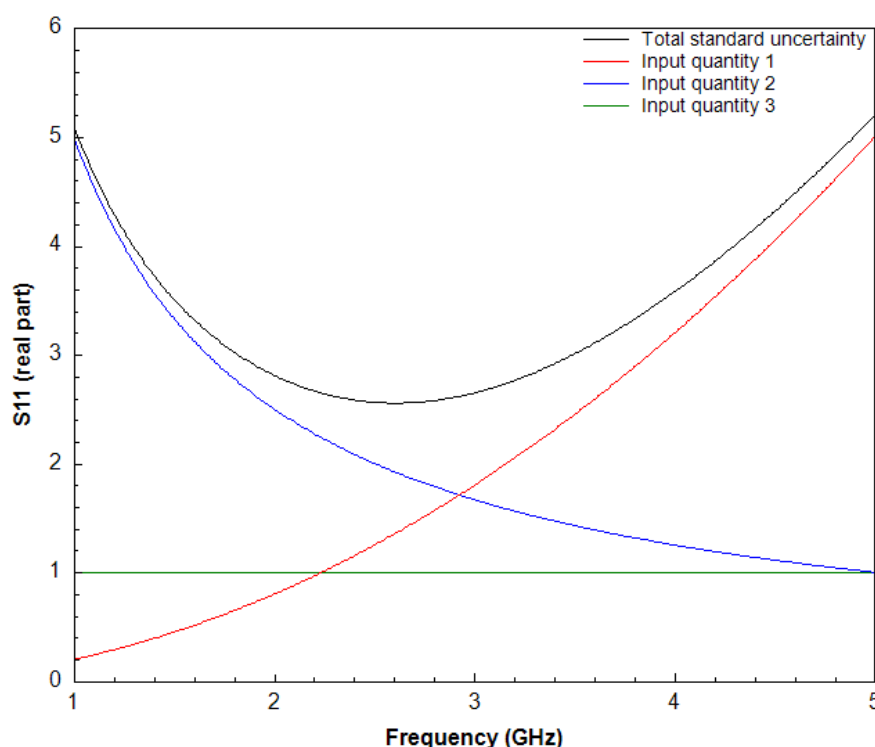


Figure 3.6: An example of results from a sensitivity analysis which reveal the origins of the complicated behaviour of the combined standard uncertainty with respect to a variable (in this case frequency).

method is similar to the finite difference propagation technique described in 3.2.3.3, which can also be used to perform a sensitivity analysis. Because the input quantities are perturbed from their estimate sequentially (while all others are held at their estimate), this form of sensitivity analysis is termed “sequential perturbation”.

The results of the sensitivity analysis can be very useful to the metrologist. Not only can the relative impact of different input quantity uncertainties be reviewed, but also complicated behaviour in the combined standard uncertainty may be better understood; Figure 3.6 shows an example. Sensitivity analyses are also an efficient approach to reducing combined standard uncertainty. Once input quantities with dominant contributions have been identified they can be targeted for improvement – or in some cases an alternative measurement model can be used which avoids them.

3.4 Conclusions

This chapter presented how measurements underpin modern life, supporting trade and commerce and facilitating new discoveries in science and engineering. Through traceability and the unit system, the evaluation and management of uncertainty in measurements generates confidence and trust. In an attempt to standardise the definition and representation of measurement uncertainties, an internationally-used guidance document, the GUM, offers rigorous methods to evaluate these uncertainties. However, the GUM continues to be developed, and recently an inconsistency was created in the evaluation of Category A uncertainty components. This chapter has reviewed the inconsistency from the objective of electromagnetic measurements, an area of metrology where the effects have been shown to be potentially significant.

Three methods for propagating uncertainty through a measurement model to determine the combined standard uncertainty of the measurands were described. Although the Monte Carlo method preserves the most information about both the uncertainties of the input quantities and the measurement model, the higher computational effort can be prohibitive in some cases. Instead, the LPU provides two linear alternatives, which are often much more efficient but require validation to ensure that the measurement model can be treated as linear.

The idea of expanded uncertainty and coverage intervals was introduced, these being met frequently in Category B uncertainty components defined from datasheets and specifications. A confidence interval is straightforward to calculate if the measurement involves only Category A uncertainty components and has a linear measurement model, or if a Monte Carlo propagation is used and the probability distribution of the measurand can be attributed to a standard type. In other cases, a coverage interval can be calculated using knowledge of the input quantities and further guidance from the GUM.

Finally, this chapter described sensitivity analysis, which can be carried out using results from the LPU procedure. The framework presented in this thesis utilises a sensitivity analysis to allow the user to examine and attempt to minimise significant sources of uncertainty, which is especially important in sensitive electromagnetic measurements such as those made on-wafer.

References

- [1] UKAS. Our role, [Online]. Available: <http://www.ukas.com/about/our-role/> (visited on 01/07/2019).
- [2] BIPM. (Nov. 10, 2016). Draft of the ninth SI brochure, [Online]. Available: <http://www.bipm.org/utis/common/pdf/si-brochure-draft-2016b.pdf> (visited on 01/07/2019).
- [3] International Organization for Standardization, *Guide to the expression of uncertainty in measurement*, 1st ed., Geneva, 1993.
- [4] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML, *Guide to the Expression of Uncertainty in Measurement JCGM 100:2008 (GUM 1995 with minor corrections)*, 2008. [Online]. Available: <http://www.bipm.org/utis/common/documents/jcgm/JCGM%20100%202008%20E.pdf>.
- [5] —, *International Vocabulary of Metrology Basic and General Concepts and Associated Terms (VIM) JCGM 200:2012*, 2012. [Online]. Available: <http://www.bipm.org/utis/common/documents/jcgm/JCGM%20200%202012.pdf>.
- [6] —, *Supplement 1 to the “Guide to the Expression of Uncertainty in Measurement” – Propagation of distributions using a Monte Carlo method JCGM 101:2008*, 2008. [Online]. Available: <http://www.bipm.org/utis/common/documents/jcgm/JCGM%20101%202008%20E.pdf>.
- [7] —, *Supplement 2 to the “Guide to the Expression of Uncertainty in Measurement” – Extension to any number of output quantities JCGM 102:2011*, 2011. [Online]. Available: <http://www.bipm.org/utis/common/documents/jcgm/JCGM%20102%202011%20E.pdf>.

4 Evaluating Uncertainty in Vector Network Analyser Measurements

4.1 Introduction

4.2 Evaluating Residual Error in VNA Calibrations

4.2.1 Imperfect Calibration Standards

4.2.2 The Ripple Technique

4.2.3 Application of The Ripple Technique at Sub-Millimetre Wavelength in Rectangular Metallic Waveguide

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

References

- [1] L. Stant, P. Aaen, and N. Ridler, “Evaluating residual errors in waveguide network analysers 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.
- [2] —, “Comparing methods for evaluating measurement uncertainty given in the JCGM ‘evaluation of measurement data’ documents,” *Measurement*, vol. 94, pp. 847–851, Dec. 2016. DOI: 10.1016/j.measurement.2016.08.015.

5 Propagating Measurement Uncertainty into Nonlinear Behavioural Models

5.1 Introduction

5.2 The X-Parameter Model

5.2.1 Model Definition

5.2.2 Extraction Procedure

5.2.3 Applications of X-Parameters

5.3 Design and Simulation using Nonlinear Behavioural Models Incorporating Measurement Uncertainty

5.4 Conclusions

Testing, testing[1], [2].

References

- [1] L. Stant, P. Aaen, and N. Ridler, “Evaluating residual errors in waveguide network analysers 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.
- [2] —, “Comparing methods for evaluating measurement uncertainty given in the JCGM ‘evaluation of measurement data’ documents,” *Measurement*, vol. 94, pp. 847–851, Dec. 2016. DOI: 10.1016/j.measurement.2016.08.015.

6 Applications of Nonlinear Behavioural Models Incorporating Measurement Uncertainty

6.1 Introduction

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus

adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetur adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

Fusce mauris. Vestibulum luctus nibh at lectus. Sed bibendum, nulla a faucibus semper, leo velit ultricies tellus, ac venenatis arcu wisi vel nisl. Vestibulum diam. Aliquam pellentesque, augue quis sagittis posuere, turpis lacus congue quam, in hendrerit risus eros eget felis. Maecenas eget erat in sapien mattis porttitor. Vestibulum porttitor. Nulla facilisi. Sed a turpis eu lacus commodo facilisis. Morbi fringilla, wisi in dignissim interdum, justo lectus sagittis dui, et vehicula libero dui cursus dui. Mauris tempor ligula sed lacus. Duis cursus enim ut augue. Cras ac magna. Cras nulla. Nulla egestas. Curabitur a leo. Quisque egestas wisi eget nunc. Nam feugiat lacus vel est. Curabitur consectetur.

Suspendisse vel felis. Ut lorem lorem, interdum eu, tincidunt sit amet, laoreet vitae, arcu. Aenean faucibus pede eu ante. Praesent enim elit, rutrum at, molestie non, nonummy vel, nisl. Ut lectus eros, malesuada sit amet, fermentum eu, sodales cursus, magna. Donec eu purus. Quisque vehicula, urna sed ultricies auctor, pede lorem egestas dui, et convallis elit erat sed nulla. Donec luctus. Curabitur et nunc. Aliquam dolor odio, commodo pretium, ultricies non, pharetra in, velit. Integer arcu est, nonummy in, fermentum faucibus, egestas vel, odio.

Sed commodo posuere pede. Mauris ut est. Ut quis purus. Sed ac odio. Sed vehicula hendrerit sem. Duis non odio. Morbi ut dui. Sed accumsan risus eget odio. In hac habitasse platea dictumst. Pellentesque non elit. Fusce sed justo eu urna porta tincidunt. Mauris felis odio, sollicitudin sed, volutpat a, ornare ac, erat. Morbi quis dolor. Donec pellentesque, erat ac sagittis semper, nunc dui lobortis purus, quis congue purus metus ultricies tellus. Proin et quam. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos hymenaeos. Praesent sapien turpis, fermentum vel, eleifend faucibus, vehicula eu, lacus.

6.2 Prediction of Optimum Load Match and Delivered Power using X-Parameters Incorporating Measurement Uncertainty

6.3 Conclusions

Testing, testing[1], [2].

References

- [1] L. Stant, P. Aaen, and N. Ridler, “Evaluating residual errors in waveguide network analysers 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.
- [2] —, “Comparing methods for evaluating measurement uncertainty given in the JCGM ‘evaluation of measurement data’ documents,” *Measurement*, vol. 94, pp. 847–851, Dec. 2016. DOI: 10.1016/j.measurement.2016.08.015.

7 Conclusions

7.1 Further Work

Testing, testing[1], [2].

References

- [1] L. Stant, P. Aaen, and N. Ridler, “Evaluating residual errors in waveguide network analysers 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.
- [2] —, “Comparing methods for evaluating measurement uncertainty given in the JCGM ‘evaluation of measurement data’ documents,” *Measurement*, vol. 94, pp. 847–851, Dec. 2016. DOI: 10.1016/j.measurement.2016.08.015.