DesignCon® 2022
WHERE THE CHIP MEETS THE BOARD

VNA Calibration Essentials for Practicing Engineers

Jason Sia, Samtec
jason.sia@samtec.com

Travis Ellis, Samtec
travis.ellis@samtec.com

Pete J. Pupalaikis, Nubis Communications
pete.pupalaikis@nubis-communications.com

Gustavo Blando, Samtec
gustavo.blando@samtec.com

Julian Lechner, Samtec
julian.lechner@samtec.com

Istvan Novak, Samtec
istvan.novak@samtec.com

Abstract

Vector Network Analyzers (VNA) have been around for decades. The instrument quickly evolved into a very useful tool for the RF and microwave engineers, but to exploit the full benefits of the instrument, calibrations are required. The practicing engineer doing VNA measurements don't necessarily need to know the inner workings of the calibration process as long as the proper calibration technique is selected for the given measurement task and the pre-defined calibration steps are explicitly followed. Yet, there are a lot of potential misconceptions and misunderstandings about the VNA calibration and what to expect from it, which often could lead to wrong interpretation of the results. Out of the many possible calibration techniques, this paper focuses only on two widely used calibration options, which are commonly called SOLT (Short-Open-Load-Thru) and SOLR (Short-Open-Load-Reflect, also called Unknown thru) calibrations. The paper focuses on one- and two-port VNA calibrations with mechanical standards. This paper is intended to help practicing VNA users with measured and simulated S-parameter data illustrating correct and incorrect expectations and procedures.

Author(s) Biography

Jason Sia is signal integrity practitioner that has been working on product characterization for 3 years with Samtec. He has a bachelor's degree in electrical engineering from Penn State University. In his free time he enjoys tutoring high school students in math and science.

Travis Ellis is a signal integrity practitioner working with customers to successfully deliver their systems to market. He believes signal integrity is critical for success. He has delivered many innovative solutions across multiple industries. He holds a mechanical engineering degree from Portland State University. Travis also enjoys the outdoors and the opportunity to work with many talented peers…

Pete Pupalaikis is a signal integrity engineer with Nubis Communications. Prior to Nubis, he worked for twenty-five years at Teledyne LeCroy designing high speed measurement instrument. He is the author of the book "S-parameters for Signal Integrity" and is an IEEE Fellow.

Gustavo Blando is a Senior Principal SI Architect at Samtec Inc. In addition to his leadership roles, he's charged with the development of new SI/PI methodologies, high speed characterization, tools and modeling. Gustavo has twenty-five years of practical experience in Signal Integrity, high speed circuits design and have participated in numerous conference publications.

Julian Lechner has over two years of experience developing solutions for the function and analysis of high-speed cable tester systems. He works on the design and application of signal processing and machine learning algorithms to tackle signal integrity problems. Julian will receive his MS in Electrical Engineering with a concentration in Signal Processing at Northeastern University in 2022.

Istvan Novak is a Principal Signal and Power Integrity Engineer at Samtec, working on advanced signal and power integrity designs. Prior to 2018 he was a Distinguished Engineer at SUN Microsystems, later Oracle. He served as SUN's representative on the Copper Cable and Connector Workgroup of InfiniBand, and was engaged in the methodologies, designs and characterization of power distribution networks from silicon to DC-DC converters. He is a Life Fellow of the IEEE with twenty-nine patents to his name, author of two books on power integrity, teaches signal and power integrity courses, and maintains a popular SI/PI website. Istvan was named Engineer of the Year at DesignCon 2020.

I. Introduction

Vector Network Analyzers (VNA) have been around for about six decades or more [1]. From the scalar voltmeter, through the vector voltmeter, the instrument quickly evolved into a very useful tool for the RF and microwave engineers. To exploit the full benefits of the instrument, over time, various calibration techniques with various sophistications have been introduced. The practicing engineer doing VNA measurements do not necessarily need to know the inner workings of the calibration process as long as the proper calibration technique is selected for the given measurement task and the pre-defined calibration steps are explicitly followed. Yet, there are a lot of potential misconceptions and misunderstandings about the VNA calibration and what to expect from it, which often could lead to sub-optimum calibration or wrong interpretation of the results. Out of the many possible calibration techniques, this paper focuses only on two widely used calibration options, which are commonly called SOLT (Short-Open-Load-Thru) and SOLR (Short-Open-Load-Reflect, or Unknown thru) calibrations and we further limit the scope to one- and two-port VNA calibrations. Furthermore, though the calibration processes themselves are universally applicable, the measurement illustrations will be limited to mechanical calibration standards.

The SOLT calibration process is well documented, see for instance [2]. Based on the assumption that the thru calibration standard in SOLT is reciprocal, the 'Unknown-thru' or SOLR calibration emerged [3]. A recent overview, explanation and derivation can be found in [4.a], uncertainty analysis can be found for instance in [5] or [10].

After a brief introduction of these calibration techniques, this paper will focus on illustrating some of the simple common misconceptions and misunderstandings and how this may show up in our data and how it can impact the accuracy of the results. We show and illustrate the wrong expectations around the data collected on the calibration standards after calibration. We explain and illustrate with measured and simulated examples of the errors related to the various ways how calibration standards are described for the calibration process, most importantly the polynomial-based and data-based definitions.

For more advanced VNA users who wish to study or implement the calibration flow externally by post-processing raw VNA data, after a brief description of the typical four-receiver VNA topology, we briefly summarize the receiver terms and switch terms. We show and illustrate the differences in the frequency and time domains between two possible ways how raw S parameters can be defined [6]. To keep things simple, we look at only one and two-port data, and we do not include over-constrained calibrations.

This paper is a collection of tidbits, with measured and simulated S-parameter and time-domain data intended to help practicing VNA users so that they can correctly interpret and appreciate the performance of the built-in calibrations offered by the instruments. The practical examples contain illustrative data measured with Keysight E5061B and N5227A network analyzers, post-processed with the Signal Integrity tool [4.b] and additional home-grown scripts.

I.1. About the nomenclature

It should be noted that the nomenclature may vary by instrumentation vendor and authors of publications. Keysight (originally Hewlett-Packard, later Agilent) and Anritsu use the Short-Open-Load (SOL) for reflection calibrations and Short-Open-Load-Thru (SOLT) and Short-Open-Load-Reflect (SOLR) naming convention for full two-port calibrations. Note that Anritsu VNAs call for a ‘Reciprocal’ thru piece in the SOLR calibration process [7], while Keysight uses the ‘Unknown-thru’ nomenclature for that step. Rohde&Schwarz VNAs use the Open-Short-Match (OSM) terminology for reflection calibrations and use Open-Short-Match-Through (TOSM or UOSM) for full two-port calibrations, where Through-Open-Short-Match (TOSM) corresponds to SOLT and UOSM stands for TOSM with Unknown through and it corresponds to SOLR [8]. Beyond the vendor-to-vendor differences, there is further complication in the naming conventions, often due to legacy reasons. As over the decades the network analyzers became more complex, users had access to network analyzers with more ports and have had access to more details to internal data from inside the instruments, some of the commonly used terms have ended up with a multitude of possible ways to refer to them. We need to understand the different options and need to pay attention to the details to correctly interpret measurement parameters and results. As a brief example, let us look at *Figure I.1*. We start with *a* and *b*, which in traditional RF and microwave engineering refer to incident and reflected waves (see *Figure I.1.a*). The complex ratio of the reflected and incident waves at a reference plane is the voltage reflection coefficient, commonly denoted by Γ. *Figure I.1.b* is a reminder that the direction of observation matters: when waves travel through a plane of observation, even if no reflection happens at that location, they switch from incident to reflected wave, and vice versa. And when reflection also occurs at the boundary, the wave magnitudes and phases will change, too. To denote this direction dependency, we could use indices for the waves on the left and right of the boundary, but indices in connection with network analyzers are commonly used to refer to different ports of the instrument, so in *Figure I.1.b* the differentiation is made by colors.

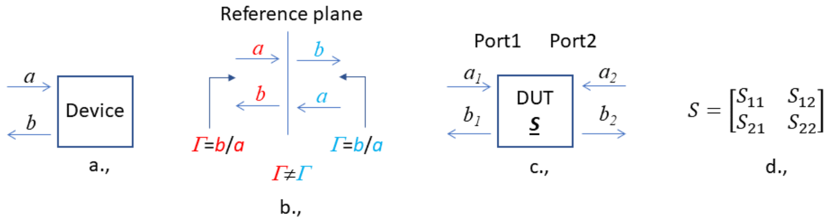
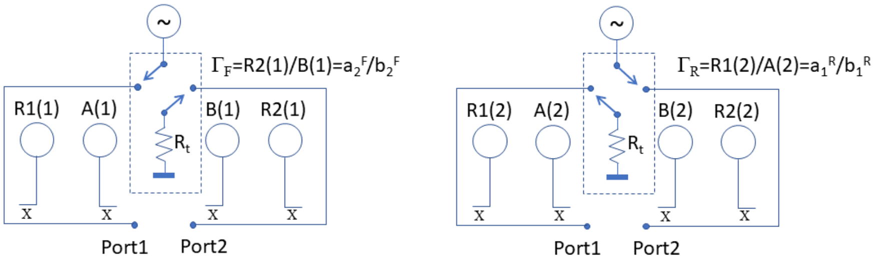


***Figure I.1.**: Waves and scattering matrix definitions.*

Vector-network analyzers measure the scattering (S) parameters of a Device Under Test (DUT). The square S-parameter matrix has a number of rows and columns that equal the number of ports. For a two-port example, it is shown in *Figure I.1.c.* and *I.1.d.* The main diagonal elements (S_{11} and S_{22}) are the reflection coefficients at the two ports and off-diagonal elements (S_{12} and S_{21}) describe wave transfer between ports. Our directional reference is from the DUT’s point of view; *a* waves point into the DUT, *b* waves point outwards. The indices of the waves and the indices of S-matrix elements refer to the ports of the VNA. When we connect a DUT to the VNA ports, we need to define our directional point of view; we need to decide what do we call incident and reflected waves at the VNA port. Probably to avoid possible misinterpretation, when we access receivers in a (two-port Keysight) VNA, the nomenclature changes: receivers measuring the signal coming into the VNA port are called Test receivers and are denoted by A (for port1) and B (for port2).



a., b.,

Figure I.2.: *Receiver and wave naming convention in Keysight VNAs.*

Not to confuse with waves *a* and *b*, receivers are identified by capital letters. Furthermore, receivers measuring the signal going outward at VNA ports are called Reference receivers and are denoted by R1 for Port1 and R2 for Port2. This means up to four ports, the port in test receivers is referenced by the capital letter itself without an index number, whereas in reference receivers the port is referenced by adding its index number. When we describe signals captured by VNA receivers, we also need to include the origin of the signal (which is not included in the generic wave definition of *Figure I.1.a* through *Figure I.1.c*). The port number where the signal is originated is included at the end of the receiver's reference in parenthesis. For instance, R1(2) refers to the reference receiver signal at port1 when the signal is generated by port2. For a two-port VNA with one source and four receivers this is shown when the source is connected to Port1 and Port2 in *Figure I.2.a* and *Figure I.2.b*, respectively. The notation gets further complicated when there are more than two ports on the VNA. Up to and including four ports, the test-port receivers are designated by letters A, B, C and D, but beyond four ports, the notation switches to numbers.

In the PNA family of Keysight VNAs, the user can also access what are called waves. This brings us back to *Figure I.1.c*, except in a VNA we want to identify the source, too, so for instance the wave incident to Port1 (a_1) needs a second index, which specifies the source. The two indices are separated by comma. As such, $a_{1,2}$ refers to the incident wave into Port1 originated at Port2. The port numbering refers to the logical port numbers, whereas the A, B, C, D notation refers to the physical port allocation, though the two can coincide. Note the sequence of indices in the wave denotation: first the destination, second the source, which is just the opposite of the sequence of indices in network matrices, like S-parameters. By comparing the notations and *Figures I.2.a* and *I.2.b* to *Figure 1* and *Figure 2* of [6], it is straightforward to create a translation map: a_1^F=R1(1), a_1^R=R1(2) and so on.

A further convenience but at the same time also complication is that in some VNA models we can reference receiver data in the old Hewlett-Packard '8510-style'. In that notation *a* refers to Reference receivers (measuring waves going out from the port) and *b* refers to Test receivers (measuring waves coming into the port), followed by a single number, which specifies the logical port. Note that here we do not (cannot) specify which port is energized, it has to be ensured implicitly. To sum it up, when for instance we need to read out the signal level exiting port1 while port2 is energized, we can use either R1(2), or $a_{1,2}$ or a_1 interchangeably. And we need to take into account the physical port vs. logical port mapping if that exists.

Section II describes the calibration process, which produces the calibrated (corrected) S parameters from the raw measured data. S parameters, by definition, assume matched terminations on all ports. The raw measurement data from a network analyzer, however, will always have some reflection created by the imperfections of hardware, notably the termination block that is switched between the ports to terminate the undriven port. The R_t termination and the switch are highlighted by the dashed rectangles in *Figure I.2*. These reflection coefficients, called 'switch terms,' Γ_F and Γ_R, are noted in *Figure I.2* first with the Keysight notation, followed by the notation used in [6]. Note that they are in the usual form of reflected-wave/incident-wave, but due to the naming convention applied to the VNA receivers, they show up here in the form of *a/b*, not in the form of *b/a* that one might expect based on *Figure I.1.b*. Because of the non-zero Γ_F and Γ_R parameters, we can define two ratios of waves: a_2/a_1 when port1 is driven and a_1/a_2 when port2 is driven. As shown in Section II, these ratios enable us to establish the connection between the two possible definitions of raw S parameters: raw S parameters which are switch-term corrected and raw S parameters which are not switch-term corrected.

I.2. Instrumentation used

For this project several different VNA models from different manufacturers have been looked at. In this paper we show data and illustrations with two Keysight models: the economy-series E5061 and the PNA series N5227. With the low-frequency extension of the E5061 and the N5292A extender box and N5293AX03 110GHz heads for the N5227 VNA, the two instruments covered the frequency range of 5 Hz to 110 GHz. To keep things simple, data collection with the E5061 VNA was done in the 2 MHz – 3 GHz frequency range with 1500-point linear sweep, using an 85052 mechanical calibration kit. This VNA model has internal SOLT calibration option, but no SOLR. Data collection with the N5227 VNA was in the 10 MHz – 67 GHz frequency range with 6700-point linear sweep and with an 85058B calibration kit, as well as in the 10 MHz – 50 GHz frequency range with an 85056D calibration kit. Many of the measurements shown here were taken multiple times so that we could observe and make sure that the inevitable small variations due to minor cable movements and connecting/disconnecting the components were much smaller than the selected signature we wanted to document. *Figure I.3* shows the typical setup with the instruments, two variants of E5061 VNA on the top and the N5227 VNA with the 85056D calibration kit, 6-inch semirigid DUT and SI test-board DUT on the bottom. Note that with the exception of the 85058B calibration kit, which has four offset short standards and for which only an encrypted data-based calibration kit definition is available, all data shown here was taken with the 85052 and

85056 calibration kits, which have only one short standard and have polynomial definitions. This allowed us to compare side-by-side the results of the internal and external SOLT and SOLR calibration results using the same calibration standard definitions.

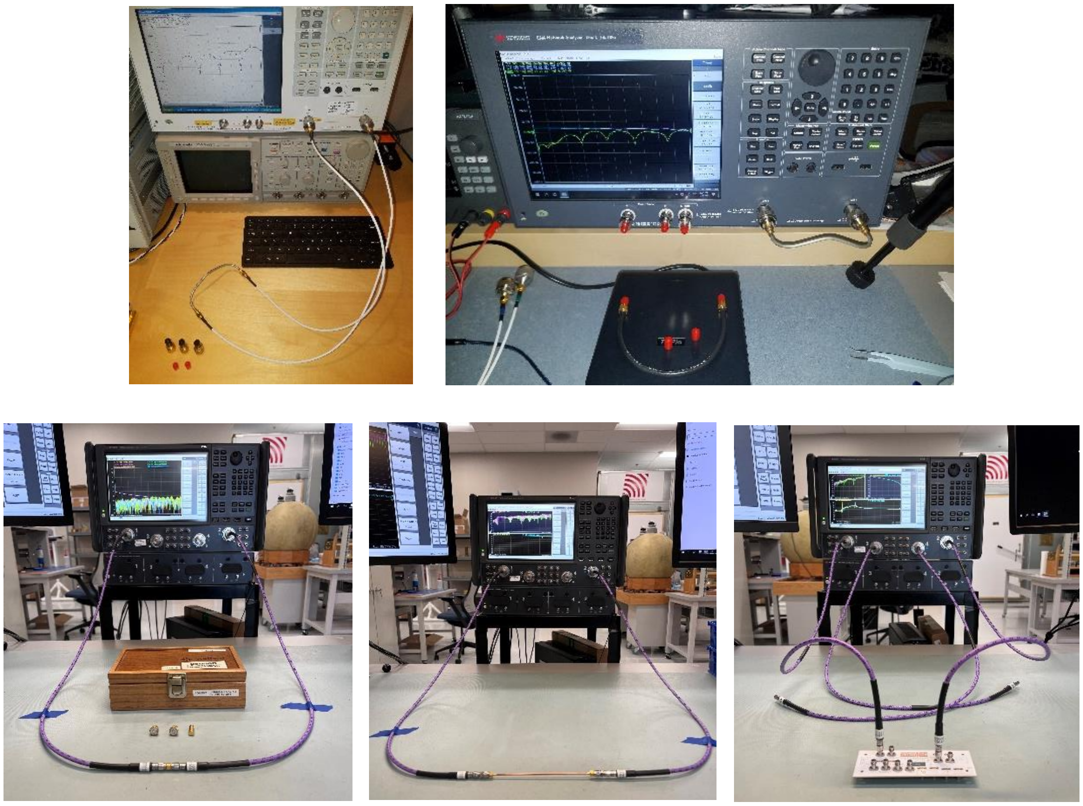


***Figure I.3.**: The typical instrumentation used for collecting data for this paper.*

Part of the learning curve was to figure out what data format we should (and can) read out from the various instruments and how to set the instruments into a state so that they output the data we need. At these details every VNA model was different, sometime also within the same model number, dependent on the configuration and firmware revision. The first question was: how to set the VNA into a state so that we can read out ‘raw’ data. One might think that turning off the calibration correction is the solution, but it turns out that the Keysight VNAs (just as several other vendors’ models) have two levels of corrections. When we turn off the calibration correction, it removes the ‘user’ calibration, but it leaves in place an average correction to the front-panel connectors that is set at the time of manufacturing. To get the raw data, we have to turn off the factory correction as well, which can be accessed from a different menu. And in case of our E5071C VNA model, yet another layer of correction had to be turned off: the ‘Virtual Bridge’ and ‘Automatic Gain Control’. The Keysight PNA series analyzers have different options for calibration correction and a single instruction can be used to bring the instrument into uncorrected state and it can also be double checked on the Status line indicator. The second question was: how to actually read out the raw S parameters. But there was yet another realization: when we turn off all correction in the VNA, it produces the raw wave parameters, S parameters with no switch-term correction, whereas the code we used [4.b] expected switch-term-corrected S parameters. Once this became clear (after [6]), our solution was to read out the raw receiver data (in case of the E5061 VNA) and raw S parameters without switch-term correction (in case of N5227) and follow the calculations outlined in Section II.

Due to HW and FW differences and limitations, we did not find a common process to read out the necessary raw data that would work with both of our chosen VNA models. On the E5061 model we ended up reading out the raw receiver terms. Shown in *Figure I.2*, we have four complex receiver terms in the forward and four in the reverse direction, altogether sixteen scalar numbers at each frequency. As shown on the left of *Figure I.4*, these were collected with four channels and four traces in each channel. Two channels collected magnitudes in the forward and reverse directions, and the other two collected the phases. This ensured phase coherent data for the subsequent processing. For sake of simplicity, we used the same process to read out the raw data for reflection (one-port) and transfer (two-port) S parameters, even though for reflection data, not having connection to the other port, there is no difference between raw S parameters with or without switch-term correction. We made use of this on the N5227 model, where we read out the uncorrected S parameters for reflection, as well as the uncorrected S parameters with the necessary receiver term ratios during transfer measurements, as shown on the upper right screen-capture plot of the VNA's internal calibration during the first pass of thru standard measurements.

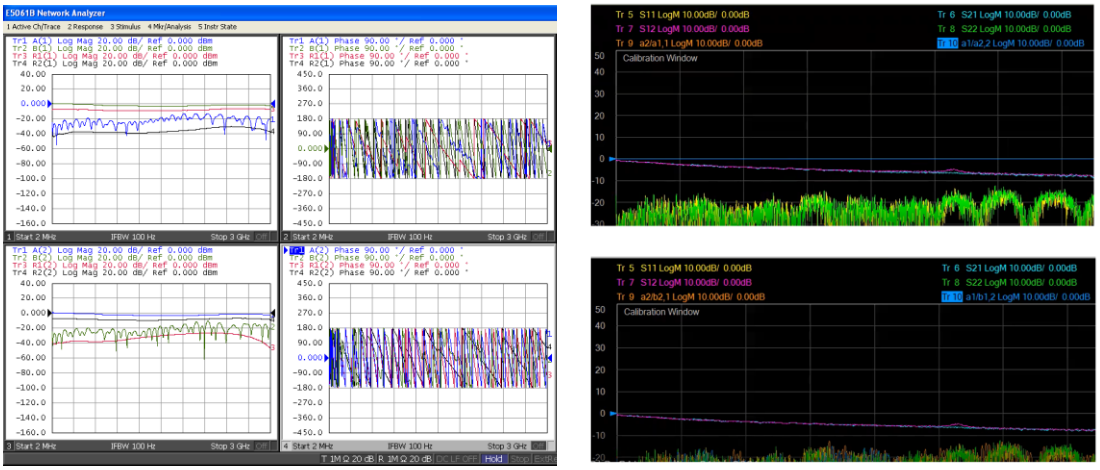


Figure I.4: *Readout of the uncorrected S parameters. E5061 VNA on the left, N5227 VNA on the right.*

II. Summary of SOLT and SOLR calibration processes

S-Parameters are network parameters and are generally expressed, for a P port device, as a $P \times P$ element matrix for each frequency point as

$$\mathbf{S} = \begin{pmatrix} S_{11} & \cdots & S_{1P} \\ \vdots & \ddots & \vdots \\ S_{P1} & \cdots & S_{PP} \end{pmatrix} , \tag{II.1}$$

where the following relationship is implied:

$$\begin{pmatrix} b_1 \\ \vdots \\ b_P \end{pmatrix} = \begin{pmatrix} S_{11} & \cdots & S_{1P} \\ \vdots & \ddots & \vdots \\ S_{P1} & \cdots & S_{PP} \end{pmatrix} \cdot \begin{pmatrix} a_1 \\ \vdots \\ a_P \end{pmatrix} . \tag{II.2}$$

Thus, the s-parameters relate incident waves (labeled as a) to reflected waves (labeled as b). The definition of each element of the s-parameters matrix is given, for a reflected wave at port x due to an incident wave at port y as

$$S_{xy} = \frac{b_x}{a_y} \Big|_{\text{all other } \mathbf{a}=0} \cdot \tag{II.3}$$

In measurement, it is not usually possible to know the values of incident and reflected waves. Part of this problem is solved by the fact that one is always looking for ratios of waves, but the main problem is solved through *calibration* of the network analyzer. This calibration employs a model that looks, mathematically, like an unknown fixture between the device under test (DUT) and the network analyzer [4.a]. Typically, known standards are applied and measured in an effort to arrive at *error terms*. Error terms are incomplete measurements of the fixture device. While incomplete, they are contrived to allow for measurement of the DUT. During the calibration phase, the measurements taken of the known standards are referred to as raw measured s-parameters; they are raw because they are not actually correct measurements but are consistent.

While there are many calibration models, only the commonly used twelve-term model will be discussed here. And, for simplicity, only two-port measurements will be considered.

The twelve-term model consists of two pairs (one per port) of six terms that are summarized in *Table II.1*. Three pairs of error terms are easily obtained through the application and raw measurement of *reflect* standards to each port. The reflect calibration measurements are the short-open-load measurements and obtain the terms E_{D_p}, E_{S_p}, and E_{R_p}. Often, the crosstalk term, $E_{X_{op}}$, is ignored and set to zero. The two remaining terms, $E_{T_{op}}$ and $E_{L_{op}}$, require a *thru* measurement; a two-port measurement of a thru standard. This, and all of the raw two-port measurements, is where the problems begin for VNA users who are performing their own calibrations externally.

(*II.3*) supplies the definition of an s-parameter, which can also be applied to the raw measured s-parameters performed during calibration or applied to the calculation of the DUT. This equation is very misleading, however. It is, in fact, possible to obtain the raw ratios of b_x to a_y from the VNA, but a problem is that all the other values of the incident waves generally cannot be made to be zero.

Table II.1.: *Summary of error terms.*

Term	Name
E_{D_p}	directivity term for port *p*
E_{S_p}	source-match term for port *p*
E_{R_p}	reverse-transmission term for port *p*
$E_{X_{op}}$	crosstalk term for port *o* when port *p* driven
$E_{T_{op}}$	forward-transmission term for port *o* when port *p* driven
$E_{L_{op}}$	load-match term for port *o* when port *p* driven

To exemplify this, imagine that when port 1 is driven, it is actually possible to have $a_2 = 0$, and similarly, when port 2 is driven, it is possible to have $a_1 = 0$. One then has

$$\hat{\mathbf{S}} = \begin{pmatrix} b_{1f} & b_{1r} \\ b_{2f} & b_{2r} \end{pmatrix} \cdot \begin{pmatrix} a_{1f} & 0 \\ 0 & a_{2r} \end{pmatrix}^{-1} = \begin{pmatrix} \frac{b_{1f}}{a_{1f}} & \frac{b_{1r}}{a_{2r}} \\ \frac{b_{2f}}{a_{1f}} & \frac{b_{2r}}{a_{2r}} \end{pmatrix},$$

and the VNA will gladly provide the values of b_{11}/a_1, b_{12}/a_{22}, b_{21}/a_{11}, and b_{22}/a_{22}, but they are not with the values of $a_{12} = a_{21} = 0$. What is really desired is

$$\hat{\mathbf{S}} = \begin{pmatrix} b_{1f} & b_{1r} \\ b_{2f} & b_{2r} \end{pmatrix} \cdot \begin{pmatrix} a_{1f} & a_{2f} \\ a_{1r} & a_{2r} \end{pmatrix}^{-1},$$

which are related through a switch-term correction [6]:

$$\begin{pmatrix} \frac{b_{1f}}{a_{1f}} & \frac{b_{1r}}{a_{2r}} \\ \frac{b_{2f}}{a_{1f}} & \frac{b_{2r}}{a_{2r}} \end{pmatrix} \cdot \begin{pmatrix} 1 & \frac{a_{1r}}{a_{2r}} \\ \frac{a_{2f}}{a_{1f}} & 1 \end{pmatrix}^{-1} = \begin{pmatrix} b_{1f} & b_{1r} \\ b_{2f} & b_{2r} \end{pmatrix} \cdot \begin{pmatrix} a_{1f} & a_{2f} \\ a_{1r} & a_{2r} \end{pmatrix}^{-1}.$$

With a thru measurement $\hat{\mathbf{S}}$ performed of a known thru standard (SOLT), the four error terms, E_{T12}, E_{T21}, E_{L12}, and E_{L21}, are obtained.

Using these error terms, the expressions for two-port-s-parameters is given by

$$\begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} = \begin{pmatrix} \frac{\hat{S}_{11}-E_{D_1}}{E_{R1}} & \frac{\hat{S}_{12}-E_{X_{12}}}{E_{T_{12}}} \\ \frac{\hat{S}_{21}-E_{X_{21}}}{E_{T_{21}}} & \frac{\hat{S}_{22}-E_{D_2}}{E_{R2}} \end{pmatrix} \cdot \begin{pmatrix} 1+E_{S_1}\frac{\hat{S}_{11}-E_{D_1}}{E_{R1}} & E_{L_{12}}\frac{\hat{S}_{12}-E_{X_{12}}}{E_{T_{12}}} \\ E_{L_{21}}\frac{\hat{S}_{21}-E_{X_{21}}}{E_{T_{21}}} & 1+E_{S_2}\frac{\hat{S}_{22}-E_{D_2}}{E_{R2}} \end{pmatrix}^{-1}, \tag{II.4}$$

where $\hat{\mathbf{S}}$ represents the raw measurements of the DUT.

It turns out that, for SOLT calibration, it doesn't matter whether the raw measurements are switch-term corrected or not.[1] For another calibration method employed by the authors during the writing of this paper, the switch-term correction was absolutely required, which caused quite a bit of confusion originally. This other calibration method is the SOLR method, also known as the method of the unknown thru.

Sometimes it is not possible to know the thru standard accurately enough to perform a good thru calibration. In such cases, specifying the thru incorrectly will lead to potentially large errors in the calibration. A remeasurement of the thru with the system will appear to be fine, but measurements of other thru elements are badly affected. SOLR resolves this problem by not requiring perfect knowledge of the thru standard.

When performing the calibration with the thru standard between an initial port p and another port o, one obtains the raw measured s-parameters of the thru as

$$\hat{\mathbf{S}}_{\mathbf{t}} = \begin{pmatrix} \hat{S}_{t_{pp}} & \hat{S}_{t_{po}} \\ \hat{S}_{t_{op}} & \hat{S}_{t_{oo}} \end{pmatrix}.$$

Ferrero and Pisani [3] found that, if the system is assumed reciprocal,

$$\frac{\hat{S}_{t_{po}} - E_{X_{po}}}{\hat{S}_{t_{op}} - E_{X_{op}}} = p$$

for any thru length, loss, or match. Assuming reciprocity,

$$E_{R_p} \cdot E_{R_o} = E_{T_{op}}^2 \cdot p$$

and thus

$$E_{T\,op} = \frac{\sqrt{E_{R_p}} \cdot \sqrt{E_{R_o}}}{p},$$

and

---

[1] The switch-term corrections are not needed to compute a set of error terms and to perform the DUT calculation – the inverse of the switch-term correction gets rolled into the error terms. It does, however, confound any interpretation of the error terms.

$$E_{T_{po}} = \sqrt{E_{R_p}} \cdot \sqrt{E_{R_o}} \cdot p \, ,$$

where, in this calibration, $E_{L_{op}} = E_{S_o}$ and $E_{L_{po}} = E_{S_p}$ are utilized. Only the basic details are provided here. Consult [4.a] for a more in-depth discussion.

III. Typical misconceptions and false expectations about calibrations

If we closely follow the prescribed calibration (and measurement) procedures, we can expect correct measurement results. However, there are still possibilities that we might mis-interpret certain results or may have false expectations towards the (otherwise correct) data. This section looks at some of such possible misinterpretations and false expectations.

As it was noted, a full reflection calibration on a port requires minimum of three standards with sufficiently different impedances across the entire usable frequency range. This gave rise to the commonly-used Short-Open-Load triplet of calibration standards. For a correct reflection calibration, we have to tell the calibration process what those calibrations standards really are. In fact, quoted from [8], the three calibration standards can be Open-Short-Match "…or any other 3 known one-port standards." The traditional way to tell the calibration process the electrical characteristics of the calibration standards is to define simple equivalent circuits for each and feed the parameters of the equivalent circuits into the calibration. *Figure III.1* shows a circuit-equivalent example based on the two Keysight mechanical calibration kits that we use later in this paper.

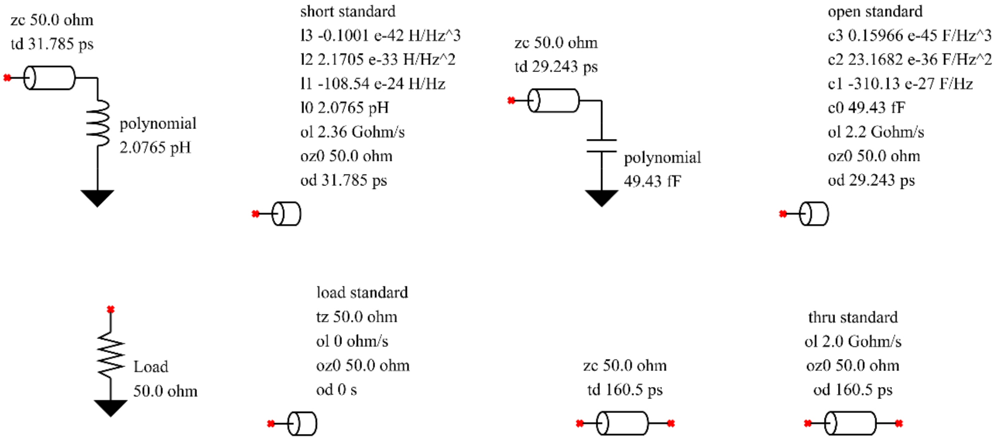


Figure III.1.a: *Drawn in [4.b] of the short (upper left), open (upper right), load (lower left) and thru (lower right) calibration standard of a Keysight 85052C 3.5mm mechanical calibration kit. Equivalent circuit (on the left) and their parameters (on the right).*

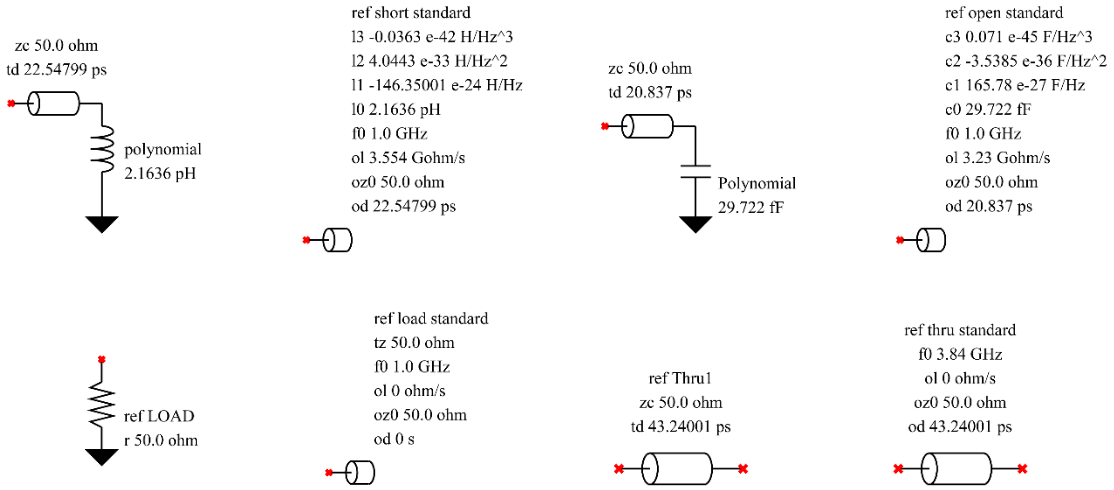


Figure III.1.b: *Drawn in [4.b] of the short (upper left), open (upper right), load (lower left) and thru (lower right) calibration standard of a Keysight 2.4mm 85056D mechanical calibration kit. Equivalent circuit (on the left) and their parameters (on the right).*

As an example of practical values, *Figure III.2* shows the reflection standards of a Keysight 85052 mechanical calibration kit calculated with [4.b] based on their polynomial definitions. These calibration standards are defined by a lossy offset delay, followed by a polynomial capacitance (for open standard) or by a polynomial inductance (for short standard). The calibration kit definition assumes that the load standard is ideal and therefore it is not shown.

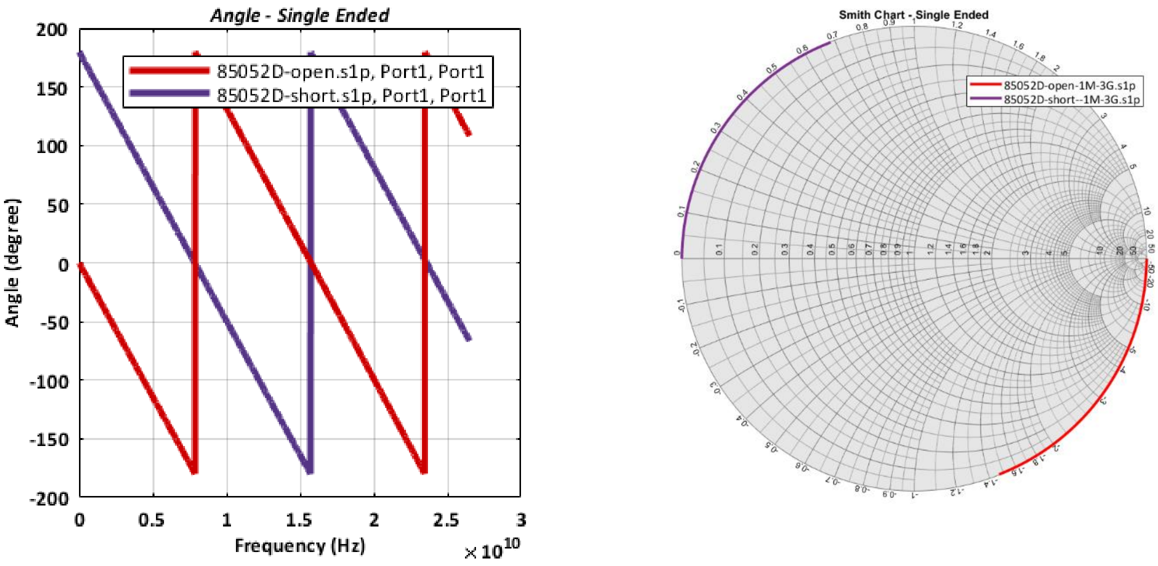


Figure III.2.: *Frequency dependency of reflection of the open standard (blue) and short standard (red) standards of the Keysight 3.5mm 85052 calibration kit. Phase in the full 0 – 26.5 GHz frequency range on the left and 0 – 3 GHz frequency range on Smith plot on the right.*

The traditional equivalent-circuit-based calibration standard definition served the industry well in the early decades of vector network analyzers, but as the frequency range expanded and accuracy expectations were raised, the limitations of the polynomial-based definitions became more obvious. Accurate high-frequency reflection measurements of well-matched DUTs become limited when we assume that our load standard is ideal or described by a simple RLC circuit. Also, the polynomial definition of short standard and open standard assumes zero and infinite impedances at DC, respectively. Assuming infinite resistance for the open standard is not a limiting factor in our every-day SI measurements, but assuming ideal zero DC impedance for the short standard would prevent us from measuring low impedances based on reflection. Note though that the Two-port Shunt-through measurement scheme overcomes this limitation by not relying on the accuracy of reflection measurement. At high frequencies it may be more troublesome that the polynomial-based open standard and short standard definitions tend to be non-causal.

Some of the above limitations can be circumvented by defining the calibration standards through Touchstone files. These data-based or file-based definitions offer the flexibility that their frequency range and accuracy can be improved over time by updating the Touchstone files, without changing the framework of the calibration process. Furthermore, this approach allows us to create our own physical standards and use them in the calibration process. Note, however, that while some network analyzer firmware allows us to use unencrypted Touchstone file definitions for calibration standards (e.g., [9]), the Keysight calibration process allows only encrypted data files. If we know the materials and construction of our reflection standards, we can do simulations to determine their reflection versus frequency.

We can also measure them with a trusted instrumentation, which ideally should have better accuracy than what our target accuracy is. In its simplest form, if we have trusted calibration standards with only encrypted file definitions, we can measure our calibration standards as DUTs after calibration, and the result approximates the encrypted definition within the repeatability limits of our instrument, cables and connections. To illustrate this, *Figure III.3* and *Figure III.4* show the measured reflection standard responses of a Keysight 85058B calibration kit. This calibration kit has four offset shorts to cover the entire DC – 67 GHz frequency range. The calibration standards were measured under different circumstances. Here we show data collected on the front panel with no cable, as well as data with a one-meter Gore phase-stable cable. The left plot of *Figure III.3* compares the phase response of the four short standards and the open standard. This illustrates the reason why we need multiple short standards: the phase difference between the open standard and various short standards varies with frequency to the point that some combinations would not yield proper calibration. The right plot compares the reflection magnitude of the open standard, calibrated and measured with and without a cable. Note that the data with the cable included has slightly more noise.

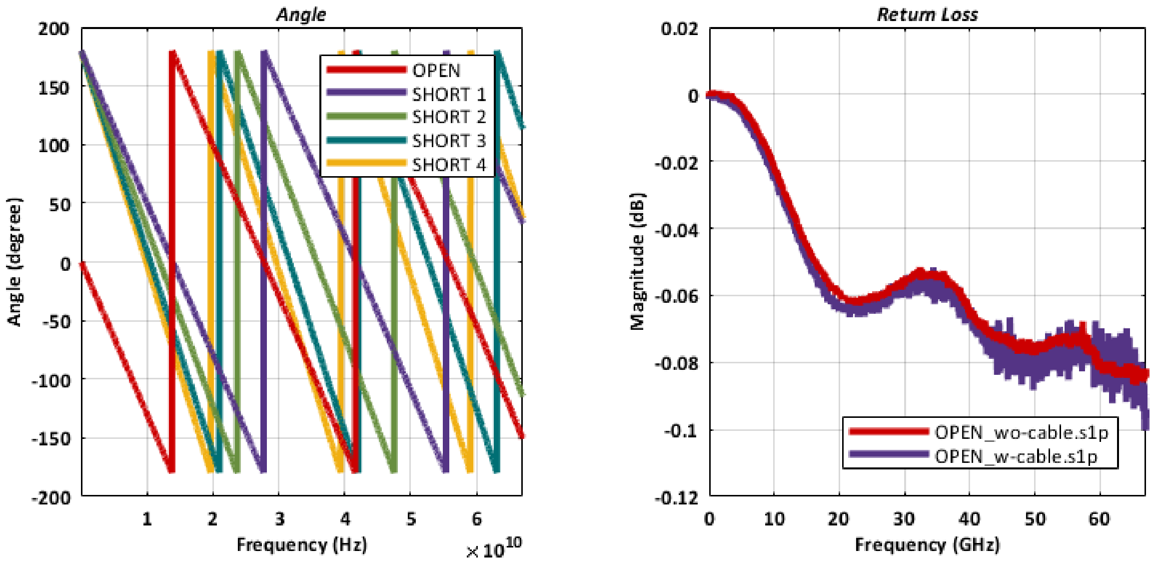


Figure III.3.*: Data measured on the calibration standards of a Keysight 85058B calibration kit open standard and four different* short standard*s.*

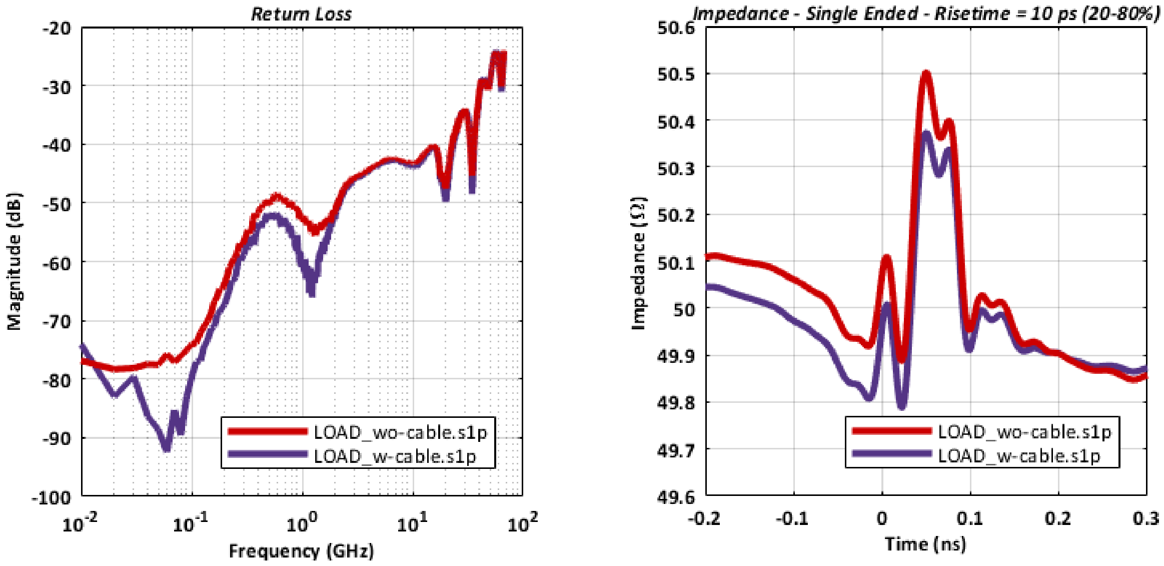


Figure III.4.*: Data measured on the calibration standards of a Keysight 85058B load calibration standard.*

The plots on *Figure III.4* compare the reflection magnitude (on the left) and the TDR responses (on the right) of the load standard with and without cable.

One possible misconception about the reflection standards is assuming that by measuring the open standard and short standards after the calibration we would see the response of ideal infinite or zero impedance. Instead, the correct measured data should match the definitions of the standards. This is true and should be the case even if we know that the definition of the standard has noticeable approximations and simplifications as it is the case with polynomial definitions: if the measuring hardware, cabling and connections had no variability and error from case to case, remeasuring calibration standards after using the same standards for the calibration itself, the measured data would perfectly match the standard definition. In other words, here we merely confirm that the calibration calculations are performed correctly. This means that remeasuring calibration standards does not tell us anything about how effective the calibration is when we measure other DUTs. The same holds true for the load standard as well: if we measure the standard after calibration, we should see what the definition of that standard is: for polynomial-based definition it is usually zero reflection (or minus infinite dB), so any finite value we see is an indication of the repeatability error and drift of the setup, not of the quality of calibration.

But we also have to point out that remeasuring the calibration standards is not totally useless after all: it gives an indication about the stability and repeatability of the setup.

One-port measurements require only reflection calibrations, but two-port (and more than two-port) measurements also need transfer calibrations between the ports. The through-connection calibration step can be done with a known through device (SOLT) or with a reciprocal unknown through device (SOLR). In the second half of the 20th century, before SOLR was invented and became widely used, the SOLT calibration was common. To avoid mis-interpretations and false expectations about the results with SOLT calibrations, we need to pay attention to the definition of the thru device in the calibration kit definition.

Traditionally VNAs were used for RF and microwave measurements, where the typical setup was ready to take insertable devices. Excluding unisex connectors, this assumes a male and a female connector on a two-port DUT and the corresponding female and male connectors at the end of the calibrated VNA setup. This arrangement has two immediate consequences: a) with insertable cabling there is no need for a physical thru piece to make a through connection between the two VNA ports and b) to do the calibration, we need two sets of reflection standards, one with male and another with female connectors. In these days a lot of DUTs are non-insertable, typically having female connectors at both ends. Such non-insertable DUTs require VNA cables that we cannot directly through connect; we need a physical thru device to do it. It is important to remember the above, because legacy calibration kit definitions may default to assuming insertable cabling, so when it comes to the thru calibration step, the internal calibration process may assume zero delay and zero loss for the thru.

Unless we select the correct variant of the calibration kit or manually update the thru standard definition, subsequent measurement may have noticeable error even at relatively low frequencies. All this is illustrated in *Figure III.5*. The data was collected with a Keysight E5061B VNA with a 1500-point linear sweep from 2MHz to 3GHz. The internal SOLT calibration was used with the 85052C calibration kit in two sets of measurements. One calibration and measurement were done with the correct delay and loss values entered for the thru piece and another with the assumption of a zero-length zero-loss thru. After the calibration, the thru device was remeasured to confirm the values together with a few other DUTs. *Figure III.5* shows a relatively clean DUT, a 6-inch long RG405 semirigid cable. On the left the four S parameters are shown with correct thru definition used in the SOLT calibration. On the right you see the results when the thru is assumed to be ideal, zero delay and zero loss. Note that S_{11} and S_{22} already had reflections up to about -30dB, so we don't see much change there. More obvious is the difference in S_{21} magnitude and the group delay. With correct thru definition the S_{21} magnitude has a clean smooth decay as expected and a relatively frequency independent group delay around 730ps. When the thru is assumed to be ideal, both S_{21} magnitude and the group delay have wrong values: we artificially lower losses and delay. In addition, we see a ripple in both responses that we will look at later.

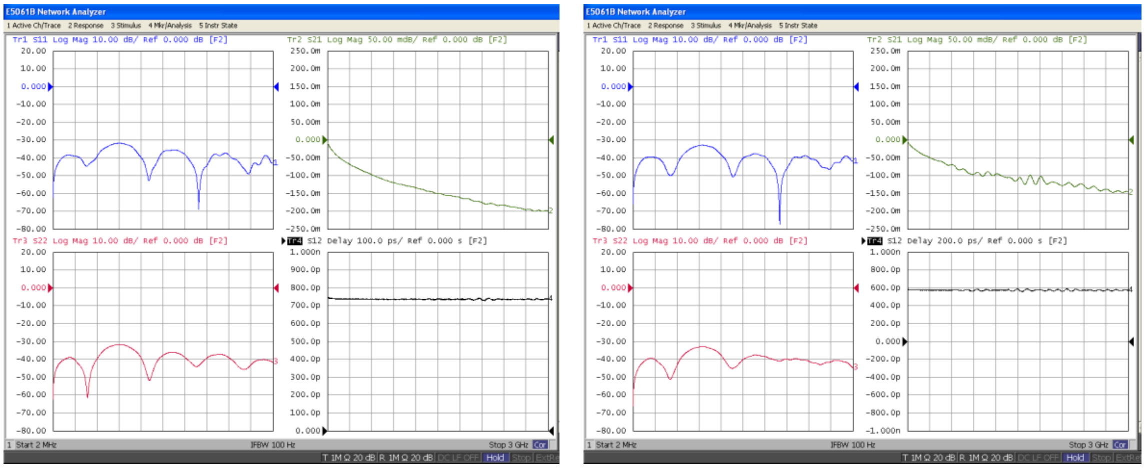


Figure III.5.: *Effect of incorrect definition of* thru *standard in SOLT calibrations on a DUT with low reflections, a 6-inch semirigid coax cable.*

If our DUT has much higher reflection, we will see bigger fluctuations on the incorrect measurement result. This is illustrated with data measured on an 8-inch long 50-ohm surface microstrip trace with a 0.75-inch long 25-ohm and a 0.75-inch long 75-ohm section symmetrically placed along its length. Because of the high reflection to start with, there is no visible difference in S_{11} and S_{22}. S21 magnitude already shows noticeable erroneous ripple and the group delay has very big fluctuations with the wrong thru definition.

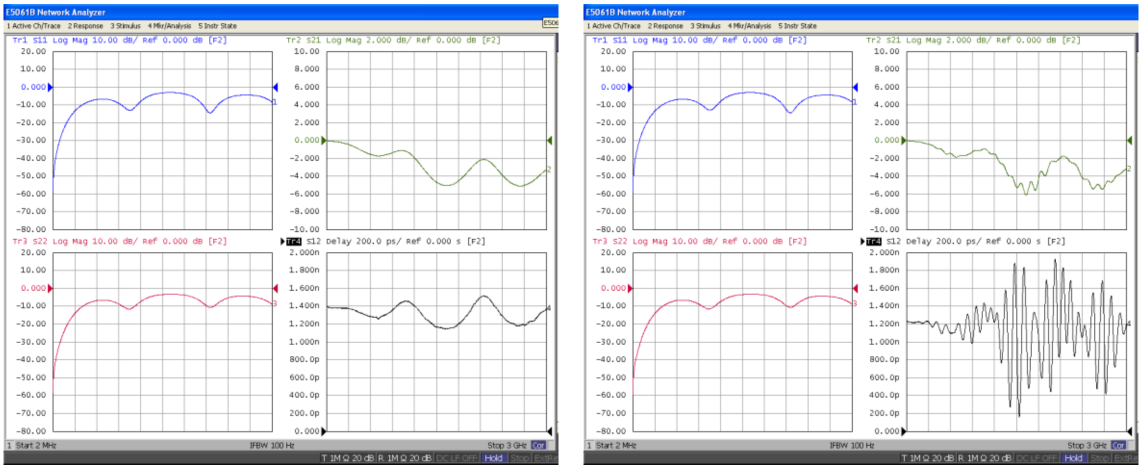


***Figure III.6.:** Effect of incorrect definition of* thru *standard in SOLT calibrations on a DUT with high reflections.*

Ultimately, if we falsely put zero delay and zero loss into the thru definition of the calibration standard but use a physical standard that has 160.5ps delay and 2.3GOhm/s loss (used for *Figures III.5* and *6*), this offset will result in a non-passive and non-causal result for a DUT shorter than the actual thru piece was. This is illustrated in *Figure III.7* where a short SMA barrel was measured with correct and incorrect thru definition during SOLT. Note that the incorrect response (on the right) has $|S_{21}|>1$ and negative delay.

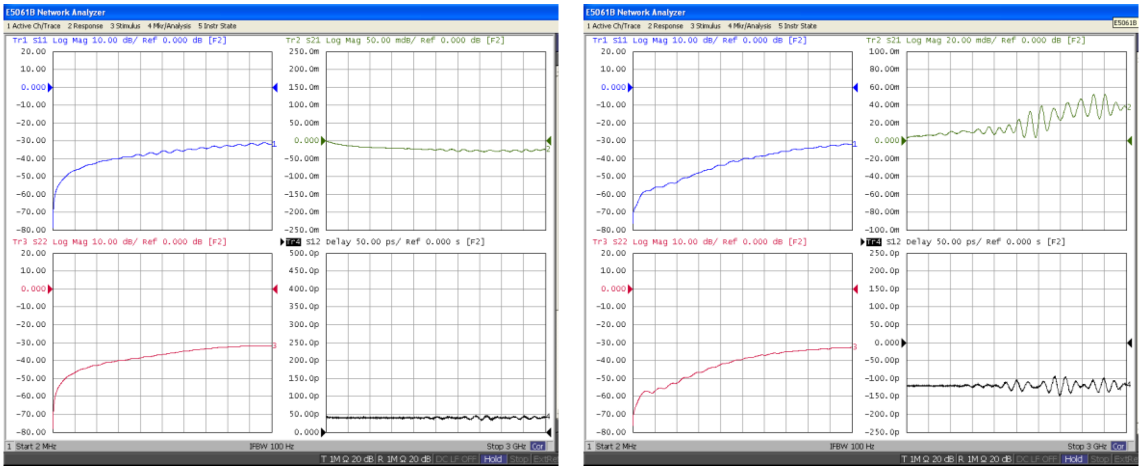


***Figure III.7.**: Erroneous response of a well-matched DUT, a short through barrel.*

To understand the source of the ripple in the frequency-domain S parameters, best is to look at the data in the frequency domain. In *Figure III.8* we show the TDR response of the short barrel from *Figure III-7* (on the left) and the reflective 8-inch microstrip trace from *Figure III-6* (on the right) with correct and incorrect thru definition in SOLT. Note that the short barrel shows the classic non-causal response with the wrong thru definition:

signature at negative times and ‘wrong tilting’ of the response peak. Note also that while in case of the reflective DUT there was very big erroneous ripple in the frequency domain, due to the large impedance swings, we can hardly see any difference in the TDR response.

The thru calibration step and the measured results of the thru standard after a full two-port calibration may be prone to the same misinterpretation and false expectations that we might see in case of reflection calibrations. We may expect zero delay and zero loss when we (re)measure the thru standard. That would be the correct expectation if we had insertable cabling, though in that case there would be no need for a physical thru device. If we have non-insertable cabling, we still may get zero delay and zero loss when we remeasure the physical thru piece, but as it shown and argued above, this is just an echo of the correct or incorrect calibration standard values we entered. And as we said about the reflection calibrations, the result of remeasuring the calibration standard (in this case the thru) is by no means an indication of how well the calibration process will perform under other circumstances; it is merely an echo of the calibration standard definition. When we use incorrect definition of our thru device in SOLT, it will show up as various errors in the different views of the measured data. If we measure a through device other than the thru standard and we happen to know its S parameters, perhaps by having measured the device also on a trusted, more accurate measurement platform, we will see an ‘average’ difference between the true and measured insertion loss and group delay matching the difference between the true loss and delay values of the thru and their entered counterparts in the standard definition.

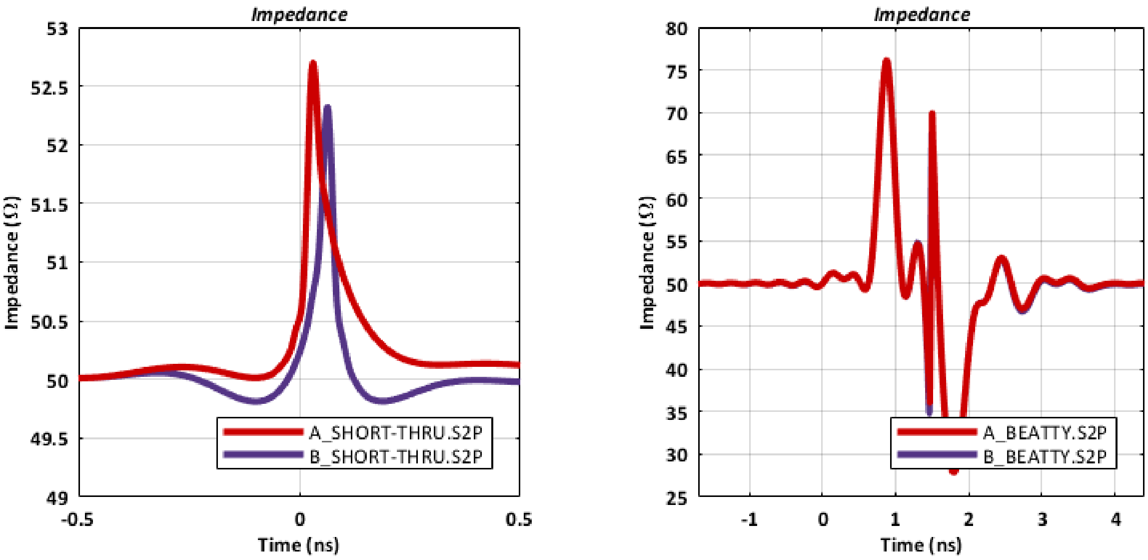


Figure III.8.: *TDR response of the 6-inch semirigid coax with correct and incorrect SOLT thru definition.*

IV. Some calibration errors

To appreciate what the calibration does and what it has to correct, it is instructive to look at some VNA data without correction. As opposed to our naïve expectation, however, when we turn off the calibration correction on most VNAs, it will not remove all corrections. When we turn off the correction of the user calibration, a baseline factory correction that produces an ‘average’ correction to the front-panel connectors, still may apply in some VNA models. To illustrate this, *Figure IV.1* compares the one-port S-parameter views of two devices with proper SOL calibration, calibration correction off and factory correction off, measured with the E5061 VNA and 85052C calibration standards. The plot on the left shows the load standard remeasured, where we need to remember that this measurement after the proper calibration merely echoes the definition of the load standard, which is ideal termination in this case: the reflection with calibration correction turned on is off the chart with the given vertical scale. Looking at the impedance profiles in *Figure IV.2* with factory correction off, we can appreciate the precision and stability we need from the calibration process and from the instrumentation hardware. The plot on the right shows the input impedance of the 8-inch microstrip, terminated, which we used also for *Figure III-6*.

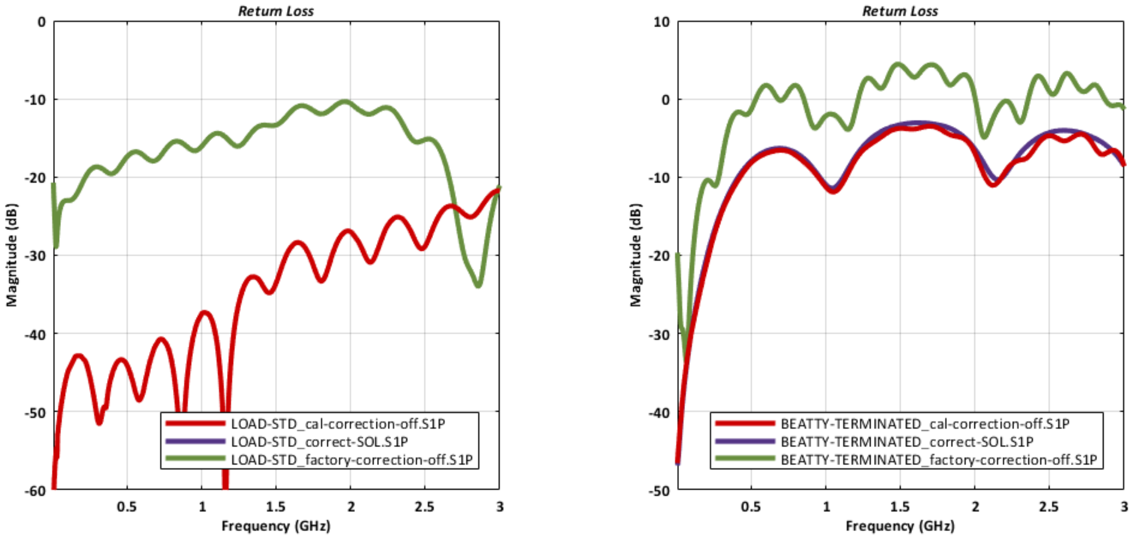


Figure IV.1.: *One-port reflection magnitude of the load standard (on the left) and a reflective transmission line (on the right) with correct SOL calibration, calibration correction off and factory correction off.*

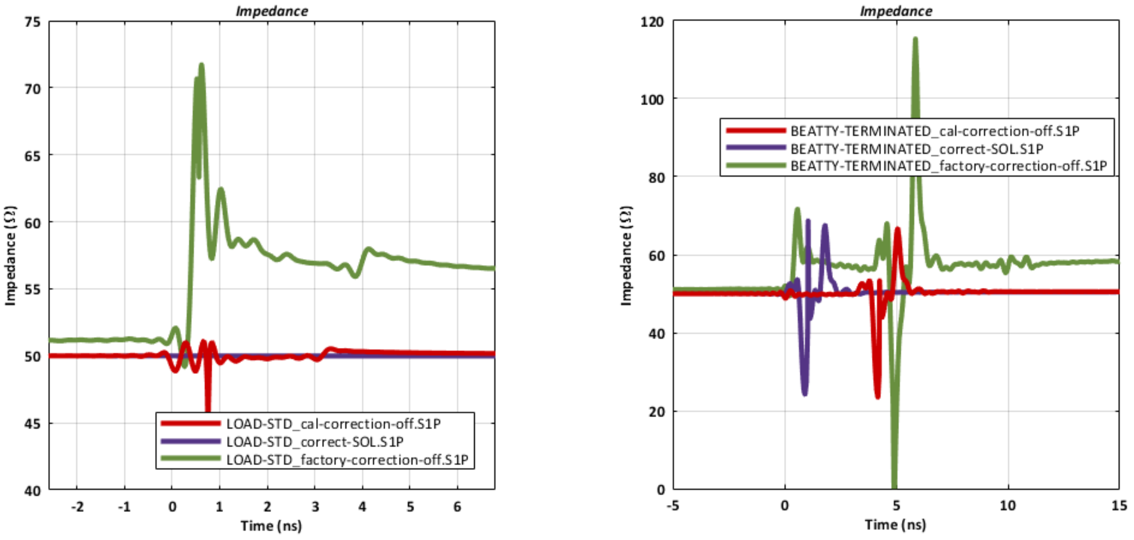


Figure IV.2.: *One-port impedance measurement of the load standard (on the left) and a reflective transmission line (on the right) with correct SOL calibration, calibration correction off and factory correction off.*

Figure IV.3 and Figure IV.4 compare the raw data and SOLT-corrected insertion loss and TDR profile of a thru standard and a 10-inch long semirigid cable.

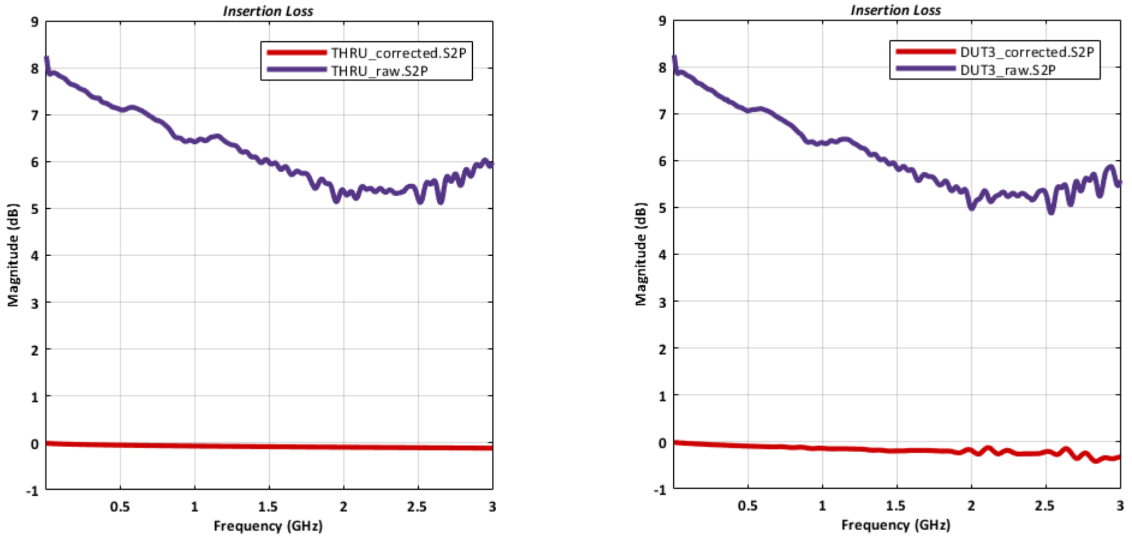


Figure IV.3.: *Two-port insertion-loss magnitude of the thru standard (on the left) and a ten-inch semirigid coax (on the right) with SOLT calibration and factory correction off.*

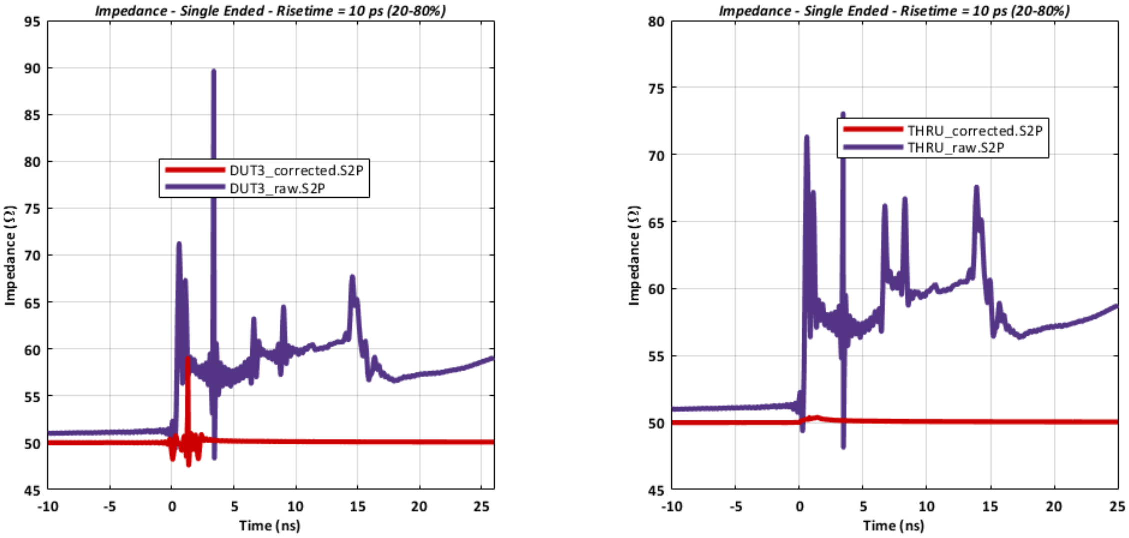


Figure IV.4.: *Impedance profile of the thru standard (on the left) and a ten-inch semirigid coax (on the right) with SOLT calibration and factory correction off.*

V. External calibration

By using the built-in calibration features of network analyzers, we can monitor real-time the quality of our measured data. When we measure a passive interconnect, we have a few simple bounds and characteristics the measured data have to obey, otherwise we know that there is something wrong in our setup, connections or possibly with the measuring instrument. For instance, we know that for a passive and reciprocal and time invariant DUT the transfer elements of the s-parameter matrix should not be bigger than unity at any frequency (though this in itself will not fully guarantee passivity of the constituents), we know that the reflection magnitudes should not be bigger than unity either, and the time-domain response should be causal, no major signature at negative times. Unless we measure a black box with content unknown to us, we may also have some knowledge about the device we measure, which further bounds the expected s-parameter response. Once we calibrate the setup, we can look at the corrected measured data as it is being collected. This is a huge convenience and benefit, but by doing so we lose the possibility to experiment with different calibration options on the same exact raw data set: if we want to remeasure the DUT with a different calibration (for instance with a different calibration standard, or different standard definition), we have to retake the data. Also, in automated bulk measurements, unless artificial intelligence is used to monitor the data real time, the immediate feedback about the data quality is not that important. In such cases we can do raw data collection with the VNA and do corrections externally, when we have the option to apply different calibration corrections to the same raw data set. If we have a known thru standard, we can do SOLT, SOLR and in SOLR we can use difference data pieces to recover the error terms.

Following the process outlined in Section II and doing SOLT calibration with the same standards and same standard definitions, with all VNA models we tried, the internal and external calibration processes produced the same result and therefore not shown here. It was also found that for SOLT calibration with correctly specified thru standard we got the same results with raw s-parameters regardless whether those were switch-term corrected or not. The external SOLR calibration is different. The process described in Section II requires switch-term corrected raw S parameters; we showed it in Section I how we did it on the two specific VNA models. *Figure V.1* shows the thru standard from the 85052C calibration kit recovered externally with SOLT and SOLR calibration with the E5061 VNA (a reminder again that the SOLT result just echoes the standard's definition).

We can notice that the recovered thru standard approximates the standard definition fairly closely, but in the frequency domain there is a noticeable (though not big) unexplained fluctuation. The reason becomes clear when we look at the TDR plot: there is a small spike around 7.5 ns (and the ripple in the frequency domain corresponds to this period), which cannot be part of the measured DUT response. The DUT is a well-matched transmission line with 160.5 ps delay.

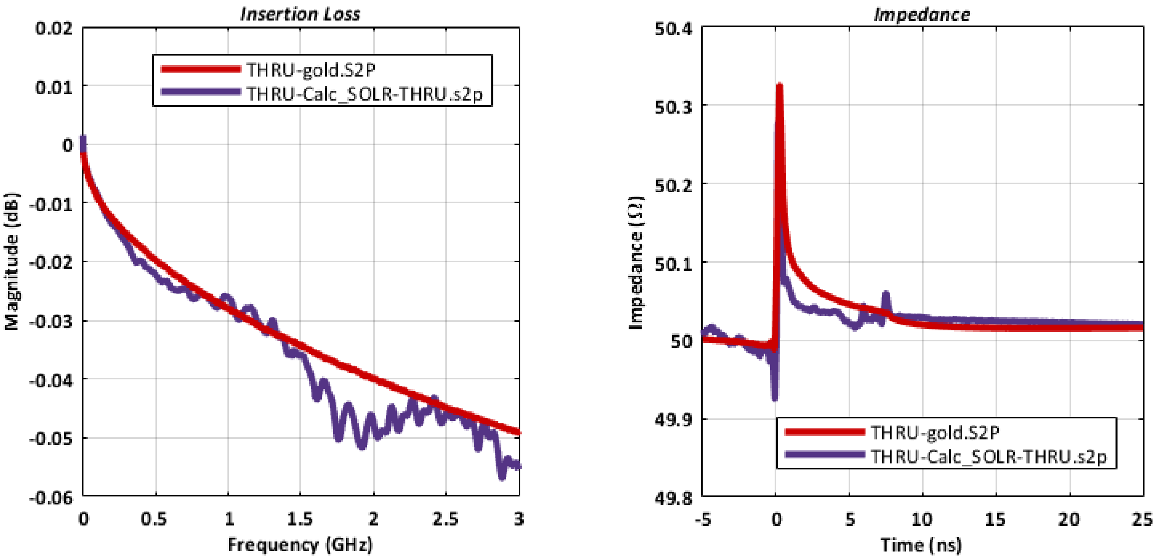


Figure V.1: *Thru calibration standard measured with external SOLT and SOLR calibration. S_{21} magnitude shown on the left and TDR profile is shown on the right.*

What we see is a residual error after the external SOLR calibration. Since this VNA model does not have internal SOLR calibration, we cannot compare this result to what the VNA would do with the same kind of calibration. We do that comparison with the N5227 VNA, shown in *Figure V.2*, where red trace shows the internal and blue the external SOLR result.

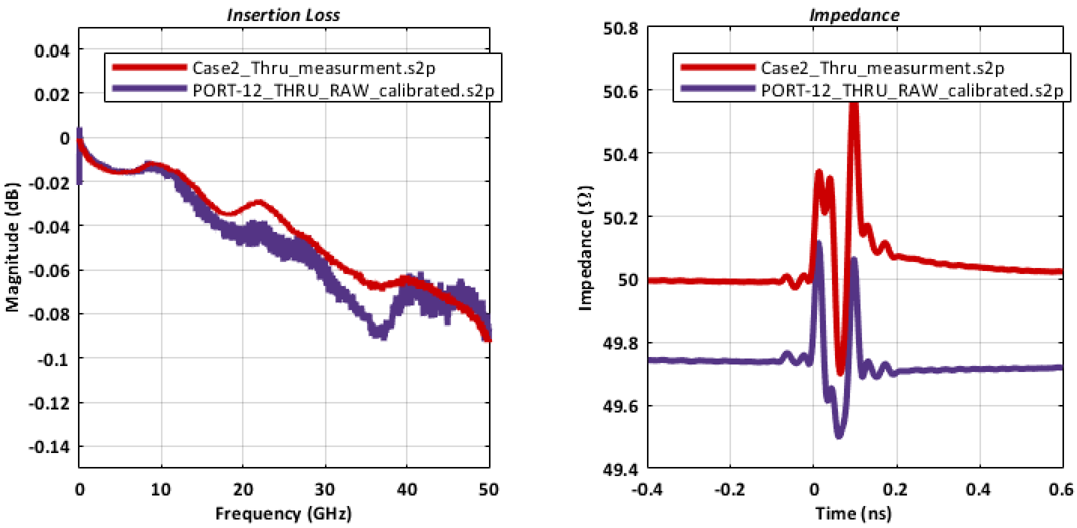


Figure V.2: *Thru calibration standard of the 85056D calibration kit measured with internal and external SOLR calibration. S_{21} magnitude shown on the left and TDR profile is shown on the right.*

Note that in this case both the internal and external SOLR results show a small ripple in the recovered thru response, which corresponds to the time difference between the minor discontinuity peaks in the TDR profile, but otherwise the TDR responses are free of erroneous residual reflections even if we open up the horizontal time scale. On *Figure V.3* we illustrate the impact of minor cable movements during calibration and measurement. Using the 85058B calibration kit with its file-based standard definition, SOLR calibration was done and subsequently the thru standard was measured. Two cases are shown: one with absolute minimum cable movement and another with a few inches more cable movements. Note that the difference, though not very big, well noticeable.

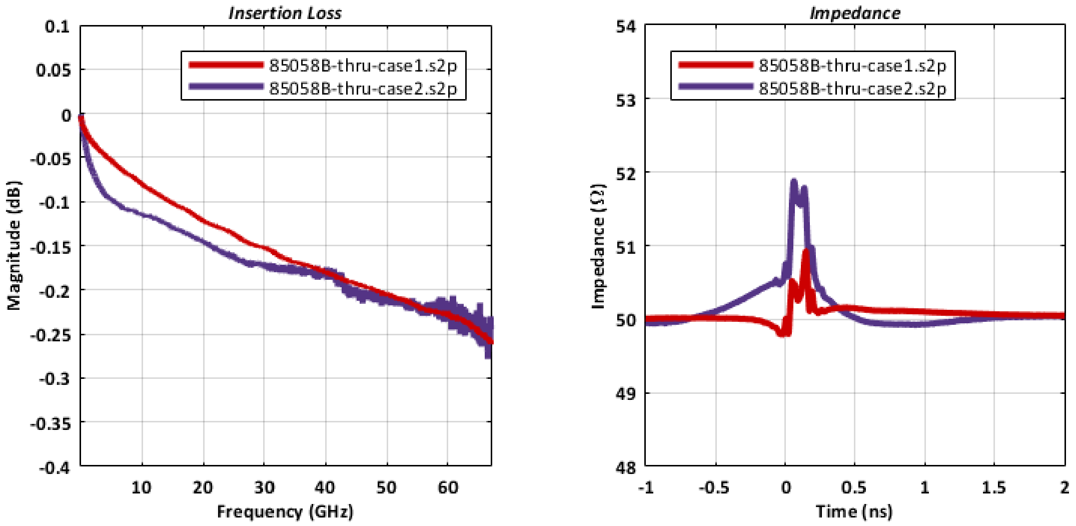


Figure V.3: *Thru calibration standard of the 85058B calibration kit measured with internal SOLR calibration with different degrees of small cable movements. S_{21} magnitude shown on the left and TDR profile is shown on the right.*

Finally, in *Figure V.4* we illustrate the possibility that with SOLR calibration we can use any reasonable two-port device to calculate the error term and correct other DUT results. Note that even using a relatively good semirigid cable as a standard, because of the higher reflections around its connectors, the unknown-thru calibration accuracy already shows more errors. And clearly the highly reflective microstrip trace is not usable as a thru calibration standard. A 6-inch semirigid cable is measured with three unknown thru standards: the thru standard of the 85052C calibration kit, the 6-inch semirigid itself and the 8-inch reflective microstrip shown in *Figure IV.2*. Measured with the E5061 VNA and external SOLR calibration. Full vertical scale on the left, zoomed vertical scale on the right.

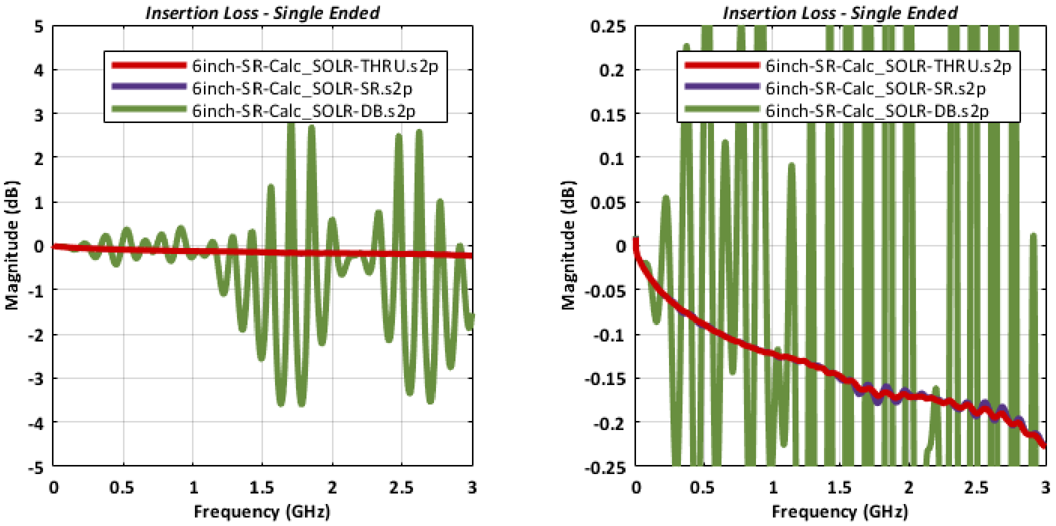


Figure V.4: *Insertion loss of a 6-inch semirigid cable with three different SOLR calibrations.*

Summary and conclusions

In this paper we summarized a few key practical points about SOLT and SOLR calibrations of two-port VNAs. Possible misconceptions were looked at starting with the simple and trivial to the more sophisticated that advanced users may also find useful. The key points and tips we showed are

- remeasuring calibration standards after calibration echoes the calibration standard definitions and therefore it is an indication of the stability and repeatability of the measurement setup, not an indication of the quality of calibration
- to check the quality of calibration we need to measure a known golden standard that was not part of the calibration
- wrong thru standard definition in SOLT calibrations results in non-causal and potentially non-passive results. It tends to manifest itself as random fluctuation in the frequency-domain response

- professional network analyzers tend to have two layers of corrections: a factory correction and a user calibration. To get raw data from the network analyzer, we need to turn off all corrections
- there are two possible definitions of raw S parameters: one as the actual network acquires the data, with no switch-term correction, and one based on the S-parameter definitions, which assumes and requires the correction for the switch terms. Switch terms do not matter in one-port measurements
- SOLT and SOLR calibrations can be done externally on the appropriate raw data read out from the instrument
- external SOLR calibration following published processes will remove the majority of errors. More complex and/or second-tier calibration may be needed to remove the residual error
- SOLR calibrations can be done with a variety of unknown-thru devices, but as their return loss (and insertion loss) get worse, the accuracy of the calibration suffers.

Acknowledgments

The authors wish to thank the following individuals for their help, support and advice: Joe Frank, Stefaan Sercu, Adam Gregory, Zak Speraw, Kyle Cronin of Samtec, Al Neves of Wild River Technologies, Yasuhiro Mori, OJ Danzy, Bob Schaefer, Mark Robinson, Mike Resso, Rusty Myers of Keysight, Brian Walker and John Hollowell of Copper Mountain Technologies, Jon Martens of Anritsu.

References

[1] Doug Rytting, "ARFTG 50 Year Network Analyzer History," 71st ARFTG Microwave Measurement Conference, Atlanta, GA, June 2008
[2] Doug Rytting, "An Analysis of Vector Measurement Accuracy Enhancement Techniques," Hewlett-Packard, RF & MW Symposium, 1980
[3] Ferrero, "Two-Port Network Analyzer Calibration Using an Unknown THRU," IEEE Microwave and Guided Wave Letters, Vol.2, No. 12, December 1992
[4.a] Pupalaikis, S-Parameters for Signal Integrity, Cambridge University Press, 2020
[4.b] SignalIntegrity: Signal and Power Integrity Tools. https://pypi.org/project/SignalIntegrity/
[5] Andrej Rumiantsev, "On Wafer S-Parameters & Uncertainties"
[6] Jargon, Williams, "The Relationship Between Switch-Term-Corrected Scattering-Parameters and Wave-Parameters Measured With a Two-Port Vector Network Analyzer," IEEE MICROWAVE AND WIRELESS COMPONENTS LETTERS, VOL. 28, NO. 10, OCTOBER 2018
[7] VectorStar™ MS464xB Series Microwave Vector Network Analyzer Calibration and Measurement Guide, Anritsu 10410-00318
[8] R&S ZNL/ZNLE Vector Network Analyzers User Manual. Rohde Schwarz User Manual 1178.5966.02 ─ 10
[9] Copper Mountain Technologies S2 VNA software, https://coppermountaintech.com/s2-vna-soft/
[10] METAS VNA Tools. https://www.metas.ch/metas/en/home/fabe/hochfrequenz/vna-tools.html