by Paul Kiciak, N2PK, paul@n2pk.com

A Comparison of VNA Transmission Calibration Methods

This article describes an experimental comparison of VNA transmission calibration methods using the N2PK homebrew vector network analyzer (VNA) [1].

OSL calibration, commonly in use in the N2PK VNA test programs, is the 'gold standard' for reflection measurements. Two calibration methods are described here that improve accuracy for transmission measurements over the commonly used response calibration. However, a dual detector VNA is required for these two new calibration methods.

The intent here is to aid N2PK VNA users in determining which calibration method best suits their accuracy needs and also to encourage other N2PK VNA software developers to add these two transmission calibration methods to their software.

Transmission Calibration Methods

Three VNA calibration methods are used for this comparison:

- Response Calibration (RC)
- Enhanced Response Calibration (ERC)
- 10-term Calibration (10-term)

All of these calibrations are considered in the context of the recent discussion on the N2PK-VNA Yahoo group regarding Correlated Double Sampling (CDS). CDS offers the option of eliminating the Isolation calibration that has been required previously to account for the detector DC offset. CDS effectively eliminates detector offset. What remains is actual coupling (or lack of isolation) between the DDS sources and the detector(s). For the testing here, stray coupling is not an accuracy issue, so CDS calibration without isolation will be used in all cases.

Response Calibration

Response Calibration [2] is the simplest. Response Calibration takes the complex ratio of the Device Under Test (DUT) transmission gain with respect to that of a Through reference. It assumes that the source and load impedances are some nominal impedance like 50-ohms. The measured S21 errors due to source & load mismatch with response calibration can be assessed using the equation for Gm in [3], if the actual DUT S-parameters and the VNA source and load matches are known:

$$G_m := \frac{S_{21}\cdot(1-\Gamma_L\cdot\Gamma_S)}{(1-S_{22}\cdot\Gamma_L)\cdot\left(1-S_{11}\cdot\Gamma_S-\dfrac{S_{21}\cdot S_{12}\cdot\Gamma_L\cdot\Gamma_S}{1-S_{22}\cdot\Gamma_L}\right)} \quad (1)$$

or equivalently

$$G_m := \frac{S_{21}\cdot(1-\Gamma_L\cdot\Gamma_S)}{(1-S_{22}\cdot\Gamma_L)\cdot(1-S_{IN}\cdot\Gamma_S)} \quad (2)$$

where

G_m is the measured complex voltage insertion gain,
S_{11} – S_{22} are the DUT S-parameters,
Γ_S and Γ_L are respectively the Source and Load Match of the VNA test ports at the DUT terminals.
S_{IN} is the reflection coefficient at port 1 with Γ_L terminating port 2.

Note that the terms in the above equations are vectors, so G_m is a vector as well, unlike the worst case scalar approximations generally made in commercial VNA documents for error assessments.

The following plots shows the RF DDS return loss and the Detector return loss respectively vs. frequency:

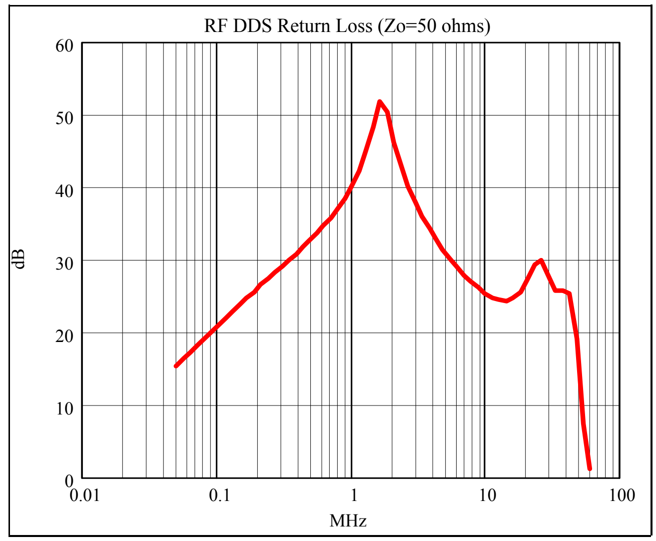

Figure 1. Typical RF DDS Return Loss

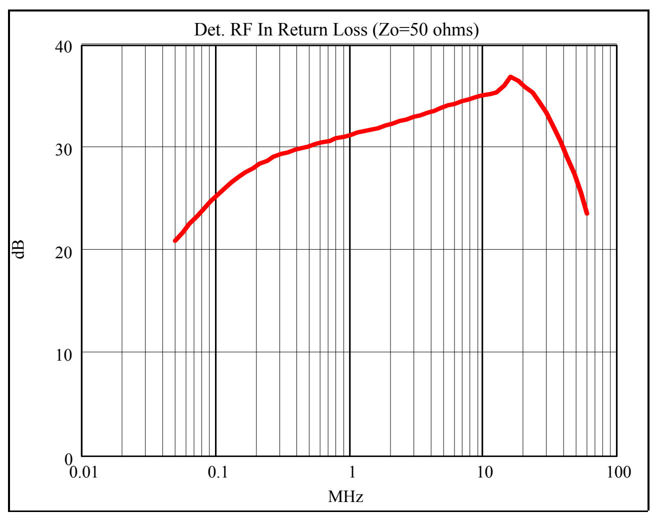

Figure 2. Typical Detector RF In Return Loss

While the detector return loss is greater than 20 dB over the VNA's 0.05-60 MHz range, the DDS return loss is significantly degraded above 50 MHz.

Even at 50 MHz, the error can be significant. As a sample calculation using the G_m equation at 50 MHz, assume:

Γ_S=0.1,
Γ_L=0.05,
S_{11}=S_{22} = -0.3,
S_{21}=S_{12}=0.6,

then G_m=0.953*0.6.

Normalizing G_m to S_{21} and converting to dB means the error in S_{21} is –0.4 dB. Considering that S_{21} is –4.4 dB, -0.4 dB is a fairly large error. And the error gets even larger at 60 MHz where $|\Gamma_S|$ is about 0.8.

One-port OSL reflection calibration is used with response calibration to collect the S_{11} and S_{22} data, where needed.

MathCAD worksheets that automate the G_m calculation for simple and Monte-Carlo evaluations can be provided to the interested reader by e-mail request.

Enhanced Response Calibration

Enhanced Response Calibration [4] takes source and load impedances into account but in a limited way. Both source and load impedances are accounted for during the through calibration. But only the source impedance is accounted for during the DUT measurement because the effect of the load impedance cannot be accounted for since the DUT S-parameters are measured only in the 'forward' direction and the effect of the load impedance depends on the 'reverse' direction parameters.

According to [5], the accuracy of enhanced response calibration is improved by the uncertainty in the source match after calibration. Ideally and as an approximation, that would mean that Γ_S=0 reducing the G_m equation to:

$$G_m = S_{21}/(1-S_{22}\Gamma_L) \quad (3)$$

This would reduce the G_m error in the earlier 50 MHz example data from 0.4 dB with response calibration to about 0.13 dB with enhanced response calibration.

10 and 12-term Calibration

12-term calibration [6] fully accounts for source and load impedances and also for coupling (Isolation) that is independent of port impedances. However, DUT direction reversal is required to make calculations of ANY DUT S-parameter. For example:

$$S_{21} = \frac{b_2}{a_1} = \frac{\left(\frac{S_{21M} - E_{XF}}{E_{TF}}\right)\left[1 + \left(\frac{S_{22M} - E_{DR}}{E_{RR}}\right)(E_{SR} - E_{LF})\right]}{D}$$

where

$$D = \left[1 + \left(\frac{S_{11M} - E_{DF}}{E_{RF}}\right)E_{SF}\right]\left[1 + \left(\frac{S_{22M} - E_{DR}}{E_{RR}}\right)E_{SR}\right] - \left(\frac{S_{21M} - E_{XF}}{E_{TF}}\right)\left(\frac{S_{12M} - E_{XR}}{E_{TR}}\right)E_{LF}E_{LR}$$

and the "E's" are the 12-terms established by calibration:

E_{DF}	Directivity	E_{DR}	Directivity
E_{SF}	Port-1 Match	E_{SR}	Port-2 Match
E_{RF}	Reflection Tracking	E_{RR}	Reflection Tracking
E_{XF}	Leakage	E_{XR}	Leakage
E_{LF}	Port-2 Match	E_{LR}	Port-1 Match
E_{TF}	Transmission Tracking	E_{TR}	Transmission Tracking

DUT reversal can happen in one of two ways – manual or relay/solid-state switch controlled via an S-parameter Test Set. Manual DUT reversal is used here.

With manual DUT reversal, the six forward error terms are re-used as the six reverse error terms. Test Set controlled DUT reversal means that the forward and reverse error terms are generally different and must be separately measured.

While constant coupling is a good approximation for some test set-ups (including the one used here), it is not always a good approximation. More elaborate coupling models, such as 16-term correction [7], have been developed to handle these cases.

Assuming that the constant coupling model is valid, 12-term accuracy is primarily limited by the accuracy of the OSLT calibration standards, linearity, connector and test set repeatability, drift, and noise.

Figure 14 on pg. 12 of [5] shows an error estimate for 12-term calibration of the Agilent 8753D VNA. Figure 14 also has an equation that can be used to determine 12-term accuracy in the general case if the right data is available. Simply plugging applicable data from Figure 14 into the G_m equation here yields a worst-case 0.03 dB transmission error for a 1 dB insertion loss & 16 dB return loss DUT vs. 0.05 dB calculated using the Agilent equation. But it is not surprising that the G_m equation here is not an accurate representation of the more complex 12-term calibration. In any case, it is clear based on these calculations that 12-term calibration generally offers significant accuracy improvement over both response and enhanced response calibration.

10-term calibration is a subset of 12-term calibration in that it does not use the Isolation calibration steps or the E_{XF} and E_{XR} error terms. The stray coupling will cause less than 0.3 dB error as long as it is at least 30 dB or more below the lowest DUT level. For the tests here, the stray coupling is at least 50 dB below the DUT signal levels.

Calibration Summary

As I noted on pg 11 of [3]:

"Response calibration is not as accurate as full 12 error term two-port correction mainly because it does not account for source and load match in the VNA. But it is much simpler and quicker to perform. Its accuracy is acceptably good for most purposes since the source and load return loss in this VNA is better than 25 dB over the HF range (1.8-30 MHz). As noted earlier, external pads can be used where needed to improve transmission measurement accuracy. Judicious use of external amplifiers, on one or both sides of the DUT, can be used to offset the loss of dynamic range normally associated with pads. Selective use of external amplifiers with known gains, with or without pads, can even be used to augment dynamic range"

The object of the testing here is to attempt to quantify and demonstrate the accuracy improvements that are possible with the slower and more complex enhanced response and 10/12-term calibrations without adding external buffers or amplifiers, that may drift or distort, and pads that result in dynamic range loss.

Devices Tested

The DUTs tested here are:

- 1.7 dB SMA attenuator (pad)
- T-Check DUT
- 20 dB SMA attenuator
- Hi-Z tap

The SMA pads are off-the-shelf with a male SMA on one end and female SMA at the other end.

The T-Check DUT here is a three-way female SMA tee with a high quality male 50-ohm termination connected to the middle leg. The T-Check DUT has about 9.5 dB return loss at both ports and about 3.5 dB insertion loss. The T-Check methodology and the Ct VNA 'figure-of-merit' were developed by Rohde & Schwarz [8].

The Hi-Z tap is a 10 kohm, 1/8 w resistor in series between ports 1 and 2. There is also a shunt 47 ohm, 1/8 w resistor from port 2 to ground. Both DUT ports have a female SMA connector. The Hi-Z tap has about 46 dB insertion loss, 0.09 dB return loss at port 1 and 30 dB return loss at port 2. There will be a decrease in insertion loss as 60 MHz is approached due to capacitance shunting the series 10 kohm resistor.

The SMA pads are examples of well-matched DUTs with low and medium insertion losses respectively.

The T-Check DUT and the Hi-Z tap are examples of DUTs that are mis-matched on at least one-port and again low and medium insertion loss respectively. The Hi-Z tap is mis-matched at port 1 and fairly well matched at port 2.

All of these DUTs are wideband and their insertion losses are low enough that DDS spurs and harmonics do not significantly affect accuracy.

All DUTs are easily reproduced so that any interested reader can make comparable measurements.

Test Configurations

After some preliminary tests, the test configuration is the same for all three calibration methods so that the accuracy improvements gained by the more complex calibration methods can be clearly seen.

The preliminary tests have the DUTs connected directly between the RF DDS and the Detector RF Input to highlight the effects of source mismatch on S_{21}. A VNA with better return loss above 50 MHz is also used. And a 10 dB pad is added between the RF DDS and the DUT on the standard VNA to show the accuracy improvement due to better source match.

The rest of the tests all have the standard T1-6T bridge between the RF DDS and Detector #1 RF In. The DUT is connected between the bridge DUT port and the Detector #2 RF In. The bridge generally improves the source match at the expense of 6 dB loss in dynamic range.

All DUTs have female SMA connectors. Homebrew female SMA OSL calibration standards, as described in [9], are used for all reflection measurements.

A commercial female-female SMA adapter is used as the Through calibration standard for all transmission measurements.

The VNA.CFG file for the test programs contains the stray values for these OSL standards. Currently, there are no stray values assigned to the Through standard. These stray values establish the reference plane at the back of the DUT SMA connectors.

The VNA is equipped with male SMA connectors on both the bridge DUT port and also on semi-rigid coaxes that connect to the RF DDS Out and Detector RF In as needed.

Other external hardware, such as the source matched RF DDS buffer or the load matched Detector Pre-amp and the use of pads as described in [9], was not evaluated here. While these hardware additions to the VNA significantly improve VNA source and load match over the full 0.05-60 MHz range, the focus here is on software not hardware corrections.

Software

The potential for significant improvements to N2PK VNA software with Correlated Double Sampling was quickly recognized after the author first saw a reference to it on a German QRP forum [10], where Ralf, DL4MW, wrote:

"In the amateur radio was once a stand-alone VNA presented with graphic display, which offset the problem by two other measurements with 180 and 270 degree phase situation has been resolved. That is yes in principle with the N2PK-VNA also, and would only have to rein in the software."

Ralf didn't call it Correlated Double Sampling - that name was adopted after Claudio, IN3OTD, did an N2PK-VNA Yahoo group post that referenced a Linear Technology application note that used that terminology [11].

CDS has been introduced into two of my DOS transmission programs. One is a beta version of the TRANS program and called TRRG5G. The other is also a beta version of my simultaneous reflection and transmission program called RETR3c. Both programs can optionally include the Isolation calibration with CDS, although Isolation was not used here as noted above.

The correction to any measured transmission value for Isolation takes the same form with or without CDS. An Isolation vector value is measured without a Through or DUT in place and that vector quantity is subtracted from all measured DUT transmission values and also from the Through value for response calibration. For enhanced response and 12-term calibrations, the Isolation vector values define the E_{XF} and E_{XR} error terms.

What is different about the Isolation values with CDS is their magnitudes. Without CDS, the Isolation values are typically on the order of one or two millivolts as they represent the detector DC offset. With CDS, the Isolation values are typically one microvolt or less – some 60 dB lower! Since the Through calibration and DUT transmission magnitudes are at least 5,000 microvolts, ignoring a one microvolt isolation offset with CDS results in little error, saves time, and results in calibrations that have better long-term stability.

TRRG5G uses response calibration and RETR3C provides either enhanced response calibration without DUT reversal or 10/12-term calibration with DUT reversal.

The reflection measurements needed in some cases to provide S_{11} and S_{22} data to supplement response calibrated S_{21} and S_{12} measurements were collected using a beta version of the REFL program that supports both (fast) detectors. This program is called REF12. REF12 does not have CDS capability but that is generally less critical for reflection measurements.

Once these programs are released, their names will be shortened to TRANS, RETR, and REFL respectively.

For all tests, the detectors were operated at their slowest ADC conversion rate to maximize dynamic range and accuracy. As noted, CDS calibrations without Isolation were used for all transmission measurements.

Test Methodology

The test methodology is straightforward. For the standard configuration tests, the OSL standards are connected to the T1-6T bridge DUT port. The Through and DUT are connected between the T1-6T bridge DUT port and the Detector #2 RF Input for all transmission and all enhanced response and 10-term calibrated reflection measurements.

Where reflection data is needed in combination with response calibrated transmission measurements, the DUT port #2 is connected either to the Detector #2 RF In to correspond to ERC and 10-term or to a precision 50-ohm load.

Since the DUTs all have female SMA connectors on both ports, DUT reversal is simplified and no errors are introduced as happens with 'insertable' DUTs that have a male connector on one port and a female connector on the other port. The N2PK budget does not allow for the purchase of hermaphroditic connectors such as the 7 mm or 14 mm connector systems [12], which are both insertable and reversible without errors!

Measured test data are stored in files and brought into Lotus 123 for plotting.

Test Results

Some preliminary test results are first considered using non-standard test configurations and shown on Figures 3-5.

Figure 3 shows the DUT transmissions with the DUT connected between the RF DDS and the Detector RF In using my standard dual detector VNA clocked at 148 MHz. This plot shows deviations from expected above 50 MHz.

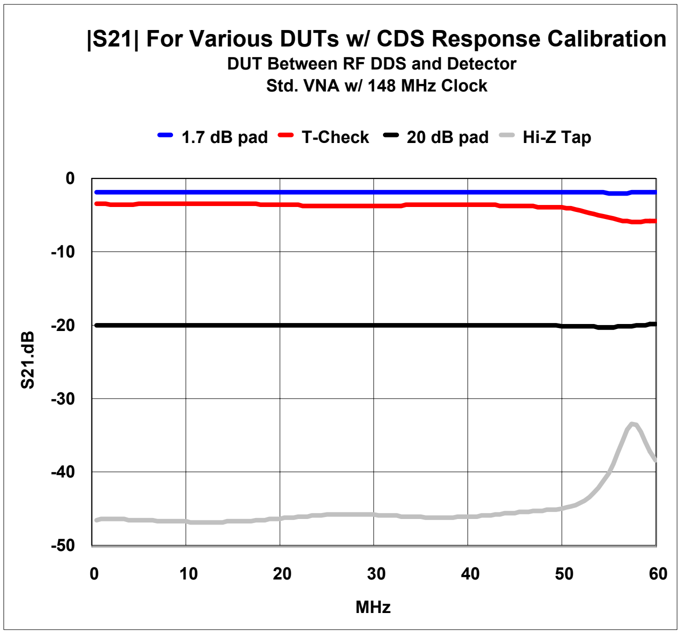

Figure 3. DUT transmissions without a bridge

Figure 4 shows the DUT transmissions for my standalone VNA PCB that uses the Connor Winfield 156.25 MHz master oscillator and re-designed anti-alias filters that have better return loss from 50 to 60 MHz. This plot shows far smaller deviations from expected over the 50-60 MHz range.

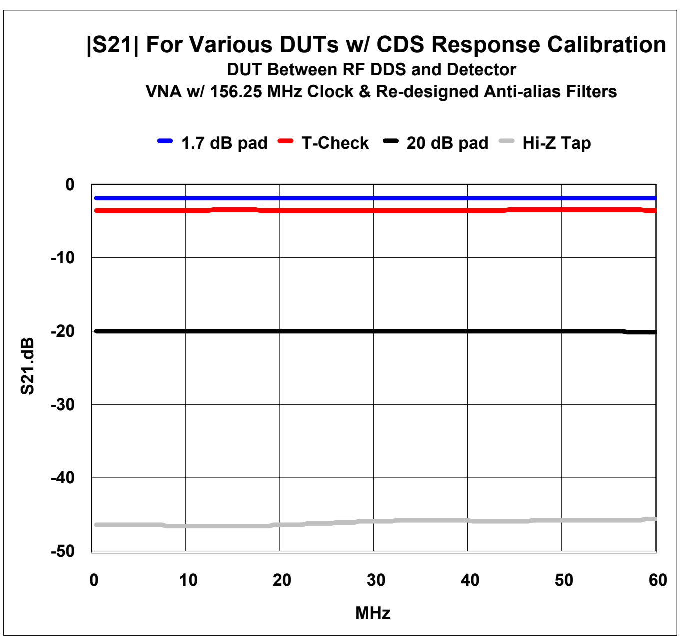

Figure 4. DUT transmissions without a bridge and higher cutoff DDS filters

Figure 5 shows the DUT transmissions again with my standard dual detector VNA but now with a 10 dB pad between the RF DDS and the DUT. This plot also shows far smaller deviations from expected over the 50-60 MHz range.

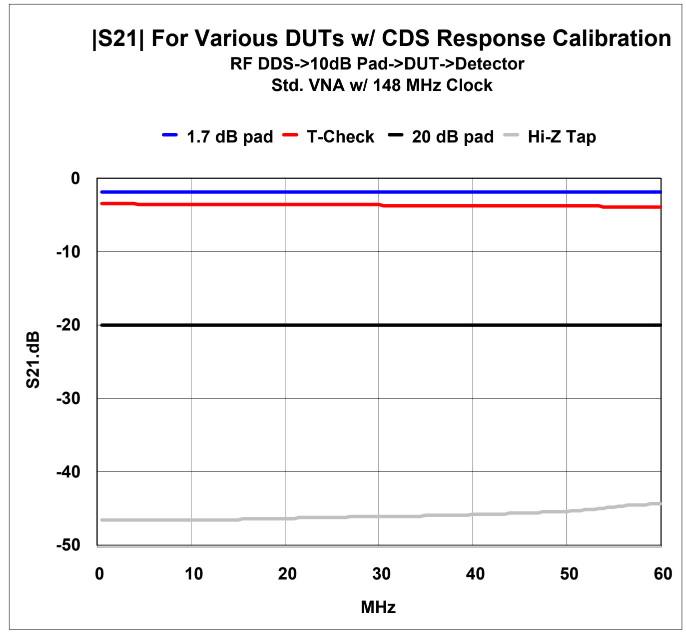

Figure 5. DUT transmissions without a bridge and a pad between the RF DDS and the DUT

Considered together, Figures 3 to 5 demonstrate the effect of degraded source match on transmission.

The remaining test results were performed with the standard test configuration with the T1-6T bridge present between the RF DDS and the Detector # 1 RF Input. The DUT is connected between the bridge DUT port and Detector #2 RF In. A dual detector VNA is required for this test configuration.

Figures 6 through 9 show the forward transmission (S_{21}) in dB for the 1.7 dB pad, the T-Check DUT, the 20 dB pad, and the Hi-Z tap respectively. Each plot contains test results for one DUT using response calibration (RC), enhanced response calibration (ERC), and 10-term calibration (10-term). In addition, Figure 10 shows the reverse transmission (S_{12}) in dB for the Hi-Z tap. Unlike the prior plots, each of these plots has an expanded vertical scale to highlight the differences in test data for the three calibration methods.

Figure 6 shows the calibration comparison for the 1.7 dB pad. Response calibration deviates up to about 0.1 dB at 60 MHz from the enhanced response and the 10-term calibration, which are virtually identical. The correlation between ERC and 10-term is expected based on equation 3 since this DUT is well matched at both ports which means that $|S_{22}\Gamma_L|<<1$ and $G_m \sim= S_{21}$. The deviations for RC from 10-term are largely due to VNA source mismatch.

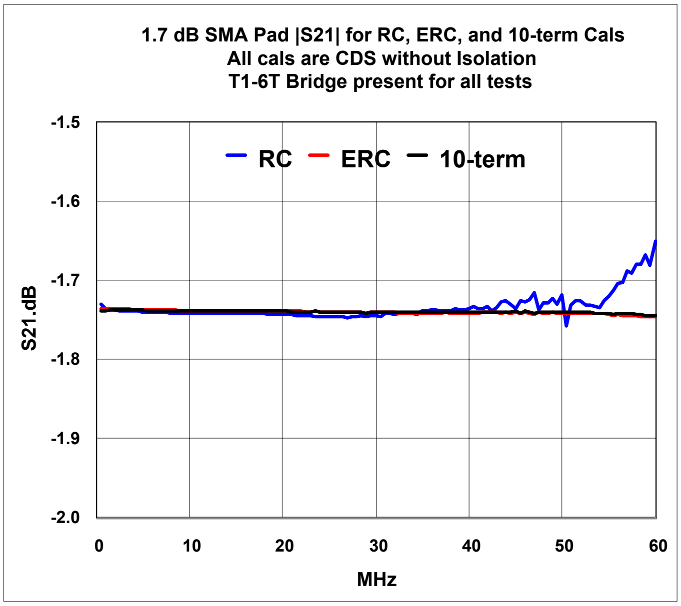

Figure 6. 1.7 dB pad transmission with RC, ERC, and 10-term calibrations

Figure 7 shows the calibration comparison for the T-Check DUT. RC again shows the largest deviations from 10-term but now ERC is somewhat different than 10-term. Again though, equation (3) shows why there are differences between ERC and 10-term. The T-Check DUT has mediocre return loss at both ports, so that $|S_{11}|$~=$|S_{22}|$~=0.33. So, there is a stronger dependency on the VNA load match, $|\Gamma_L|$, which is on the order of 0.1. The expected deviation then between ERC and 10-term is on the order of 0.26 dB vs. 0.16 dB measured. The 0.26 dB estimate from equation (3) assumes worst case vector alignment which isn't always met in practice. Apparently, the deviations for RC from 10-term are due to a combination of VNA source and load mismatch and DUT port mismatches.

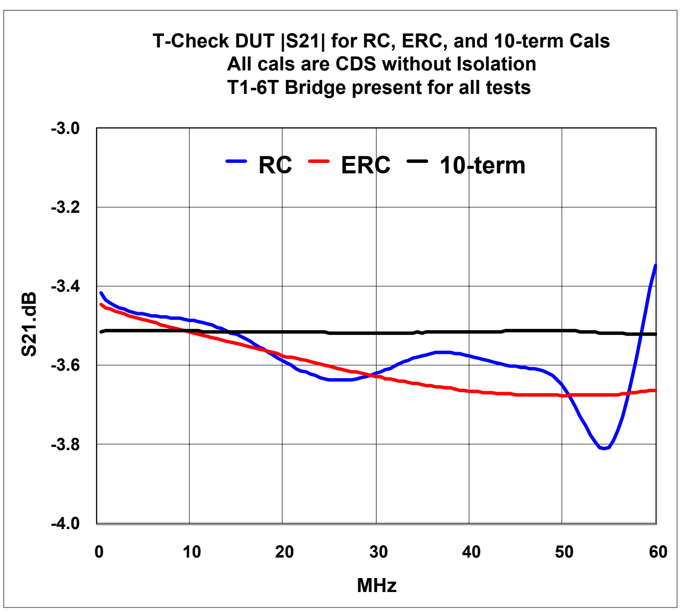

Figure 7. T-Check DUT transmission with RC, ERC, and 10-term calibrations

Figure 8 shows the calibration comparison for the 20 dB pad. RC deviates up to about 0.26 dB at 60 MHz from 10-term. ERC differs from 10-term by about 0.02 dB maximum. As with the 1.7 dB pad, the correlation between ERC and 10-term is expected based on equation 3 since this DUT is well matched at both ports which means that $|S_{22}\Gamma_L|$<<1 and G_m~=S_{21}. The deviations for RC from 10-term are largely due to VNA source mismatch.

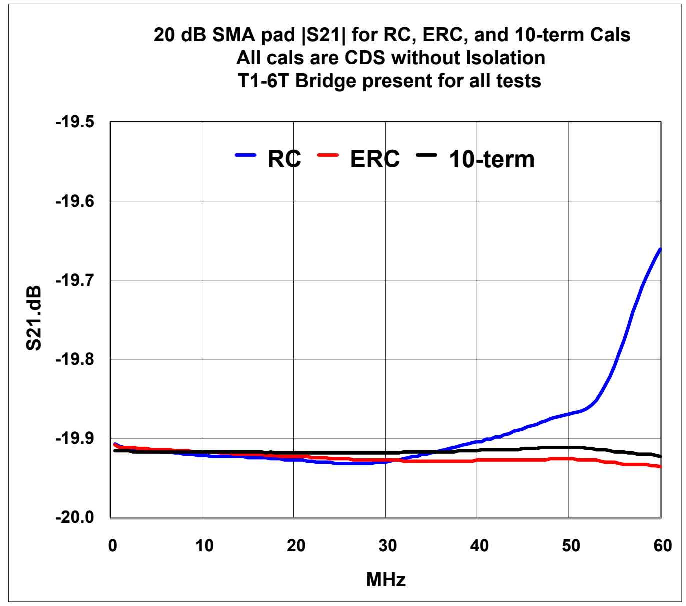

Figure 8. 20 dB pad transmission with RC, ERC, and 10-term calibrations

Figure 9 shows the calibration comparison for the Hi-Z tap forward transmission (S_{21}). RC deviates up to about 0.94 dB at 60 MHz from 10-term. ERC differs from 10-term by 0.01 dB maximum. The correlation between ERC and 10-term is expected based on equation 3 since this DUT is well matched at port 2 which means that $|S_{22}\Gamma_L|$<<1 and G_m~=S_{21}. Apparently, the deviations for RC from 10-term are due to a combination of VNA source mismatch and DUT port 1 mismatch.

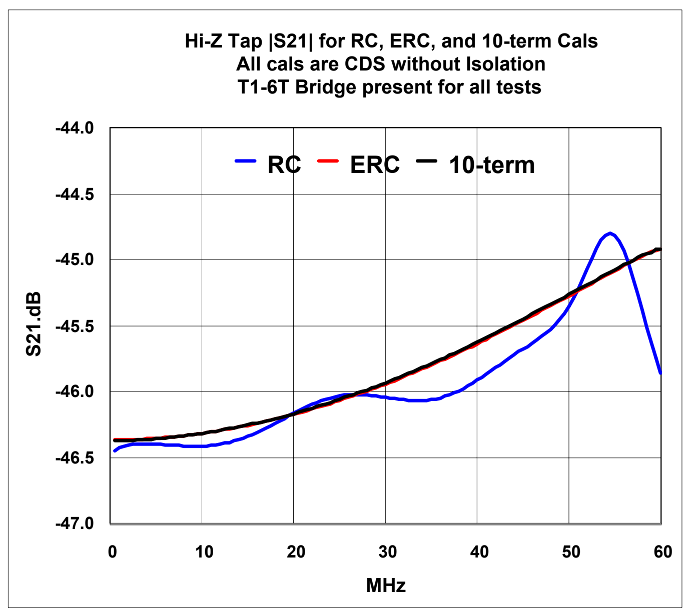

Figure 9. Hi-Z Tap forward transmission with RC, ERC, and 10-term calibrations

Figure 10 shows the calibration comparison for the Hi-Z tap reverse transmission (S_{12}). While the measured reverse transmission for the first three DUTs is about the same as forward, that is not true for the Hi-Z tap. There are differences for RC and ERC vs. 10-term since port 1 is mismatched while port 2 is well matched. As a result, the transmission errors are different, as can be easily seen from equation (3) for ERC.

Compare Figure 9 and 10 to see these differences. While Figure 9 show ERC and 10-term to be almost identical for forward transmission, Figure 10 shows that RC and ERC are very close for reverse transmission. Note also that $|S_{21}|$ and $|S_{12}|$ for 10-term are virtually identical.

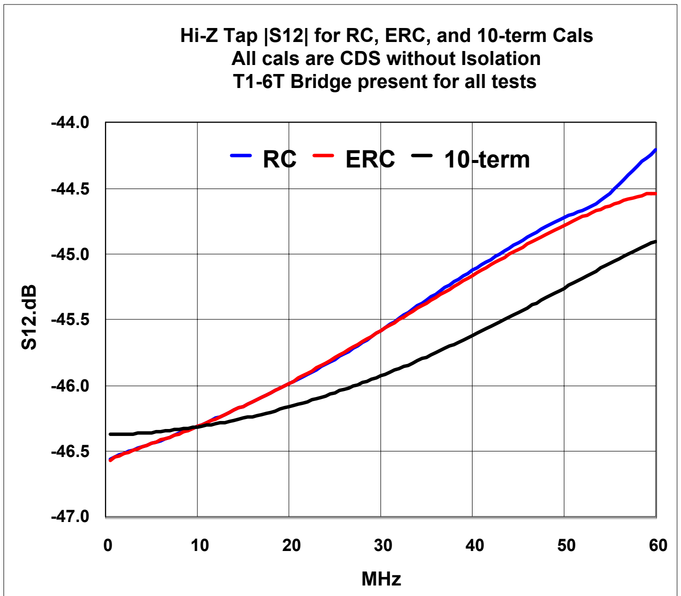

Figure 10. Hi-Z Tap reverse transmission with RC, ERC, and 10-term calibrations

Figure 11 shows the forward reflection (S_{11}) for the T-Check DUT. The Rohde & Schwarz Ct figure of merit requires all four S-parameters for its calculation. Due to symmetry, the reverse reflection (S_{22}) is virtually identical to the forward. The same is true for S_{21} and S_{12}. While data for all four parameters were collected, they are not shown here.

Since RC does not require a reflection measurement, a separate one-port OSL calibration is used to collect reflection data to be used with RC transmission data for the Ct calculation. But there are two ways to collect this reflection data. One way is to terminate the non-driven DUT port by the Detector RF In, as is done with ERC and 10-term. The other way is to terminate the non-driven DUT port by a high-quality 50-ohm termination. Both are shown in Figure 11.

Figure 11 shows that one-port and ERC reflections are virtually identical as long as the DUT non-driven port is terminated the Detector RF In. This is not surprising since both basically measure S_{IN} in equation (2). Figure 11 also shows that one-port and 10-term reflections are virtually identical as long as the DUT non-driven port is terminated by a high-quality 50-ohm termination. This too is not surprising since 10-term basically measures S_{11} and S_{IN} is also essentially S_{11} since Γ_L is nearly zero.

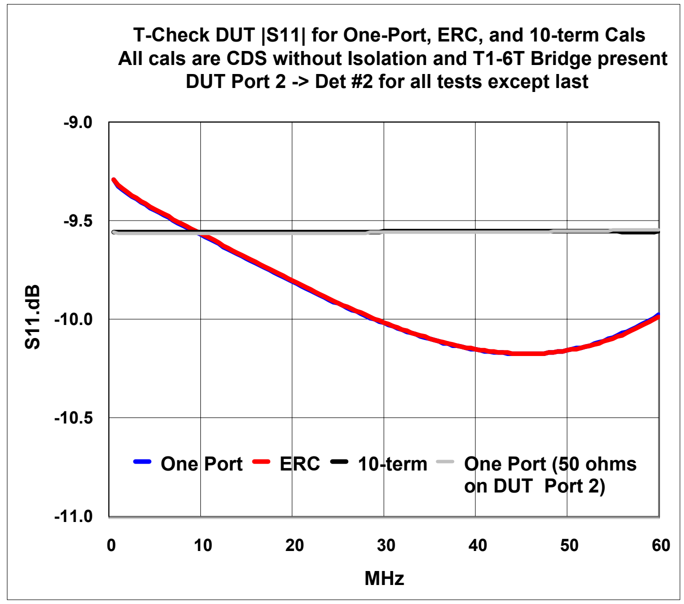

Figure 11. T-Check DUT forward reflection with One-port, ERC, 10-term, and One-port w/ 50 ohms on port 2

The C_T figure of merit using the T-Check DUT is shown in [8] as:

$$\frac{|S_{11} S^*_{21} + S_{12} S^*_{22}|}{\sqrt{(1 - |S_{11}|^2 - |S_{12}|^2)(1 - |S_{21}|^2 - |S_{22}|^2)}}$$

C_T is ideally 1 or 100% and is intended to give a figure of merit or performance measure for the VNA hardware and, of course, software being tested. Rohde & Schwarz in [8] also has:

"Deviations of up to +/-10% are considered as minor (green range). Deviations between 10% and 15% are still acceptable (yellow range) and those more than 15% should not occur in a good vector network analyzer after careful system error calibration (red alert)."

Figure 12 shows C_T vs. frequency for the four different calibrations tested here. The maximum deviation from 100% for 10-term calibration is about 0.3%, so that meets Rohde & Schwarz's 10% criterion by a wide margin. RC with a high quality 50-ohm termination on the non-driven DUT for reflection measurements just meets the 10% criterion. ERC just meets the 15% criterion. And RC with Detector #2 terminated reflection data fails the 15% criterion at 54 MHz by 2.3%. Note the similarity in shape for the two RC curves. Since the 50-ohm reflection terminated RC has very good reflection coefficients, what's left there is largely due to transmission errors. That means then that the degradation to the detector terminated RC is largely due to the poorer reflection coefficients.

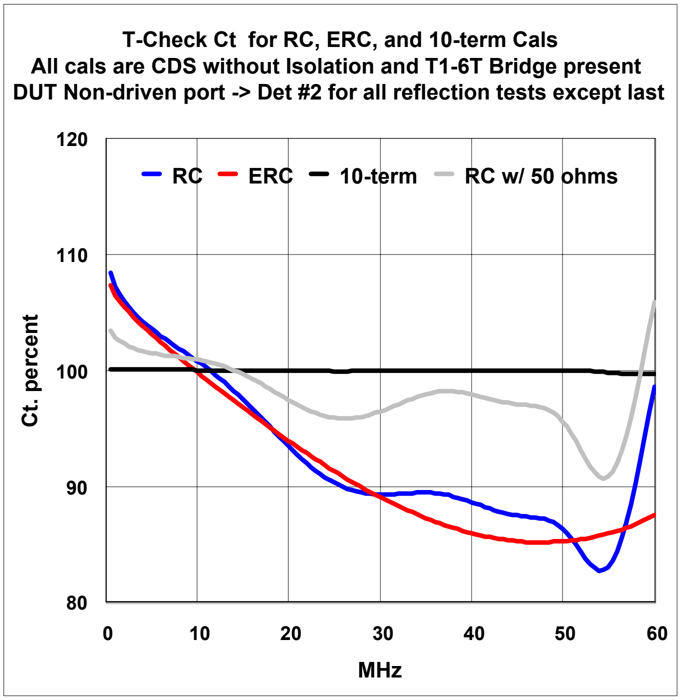

Figure 12. T-Check DUT Ct figure of merit with RC, ERC, 10-term, and RC w/ 50 ohms on port 2

Conclusions

Response calibration is reasonably accurate as long as the VNA source and load matches are good.

Enhanced response calibration has significantly better accuracy than response calibration as long as the DUT output match is good.

Enhanced response calibration accuracy suffers if the DUT output match is not good but the accuracy is no worse than with response calibration.

10-term (or 12-term) calibration has the best accuracy.

Analytic expressions are available for each calibration method to assess accuracy for specific test cases.

Notes

Clicking on a URL below with an active Internet connection, in most cases, will bring it up in your web browser if you have the Adobe plug-in. If any documents cannot be found (the web is a dynamic place), please contact me by e-mail.

1. The N2PK VNA is described at:
   http://n2pk.com
   and:
   http://n2pk.com/VNA/VNAarch.html
2. For response calibration, see page 11 here:
   http://cp.literature.agilent.com/litweb/pdf/5968-5329E.pdf
   which has:
   "The simplest form of direct measurement is a response calibration, which is a form of normalization. A reference trace is placed in memory and subsequent traces are displayed as data divided by memory. A response calibration only requires one standard each for transmission (a thru) and reflection (a short or open). However, response calibration has a serious inherent weakness due to the lack of correction for source and load mismatch and coupler/bridge directivity. Mismatch is especially troublesome for low-loss transmission measurements (such as measuring a filter passband or a cable), and for reflection measurements. Using response calibration for transmission measurements on low-loss devices can result in considerable measurement uncertainty in the form of ripple. Measurement accuracy will depend on the relative mismatch of the test fixture in the network analyzer compared to the DUT."
3. The measured S21 error for response calibration can be determined using the equation for Gm on pg 15 here:
   http://n2pk.com/VNA/n2pk_vna_pt_1_ver_c.pdf
4. For enhanced response calibration, see Joel Dunsmore's July 13, 2007 post here:
   http://forums.tm.agilent.com/phpBB2/viewtopic.php?p=3980
   where he writes:
   "Now, enhanced response cal does a 1 port cal on port one, then during the through, it also measures the load match, and thus it can correct the transmission tracking term (which is used to correct S21) for the mismatch between the source match and the load match. But, during the measurement, the load match information is not used. However, the S11 measurement is used to correct the S21 trace for the effect of mismatch between the source of the VNA and the input match of the DUT. So, in that way it is about 2/3 as good as a full 2 port cal. Why 2/3? Well, during the calibration there is an error that occurs between the source and the load match during the thru measurement. During the DUT measurement, there is an error due to the input match, and another due to the output match. For enhance response cal, all but the output match is corrected." (Typos corrected here)
5. An accuracy estimate of enhanced response calibration can be found in Agilent AN 1287-3 on pg. 11 at:
   http://cp.literature.agilent.com/litweb/pdf/5965-7709E.pdf
6. A description of 12-term calibration can be found in the Appendix at:
   http://www.boulder.nist.gov/div818/81801/NonlinearDevCharPublicationsOnly/MetrologyPublications/ARFTG99F_OSLT.pdf
   There is also some interesting information about calibration standards in the body of the above document.
7. 16-term calibration (or error correction) can be found here::
   http://cpd.ogi.edu/IEEE-MTT-ED/Network%20Analyzer%20Error%20Models%20and%20Calibration%20Methods.pdf
8. The T-Check accuracy test can be found at
   http://www.rohde-schwarz.dk/www/downcent.nsf/ANFileByANNoForInternet/6F6136DE64DA92DEC1256B4A0044F093/$file/1ez43_0e.pdf
9. An RF DDS buffer and a Detector pre-amp, which significantly improve VNA source and load match respectively over the full 0.05-60 MHz range, are described on pgs 26-28 here:
   http://n2pk.com/VNA/n2pk_vna_pt_2_ver_b2.pdf
10. Nov. 28, 2007 German QRP forum post where Ralf, DL4MW, describes the essence of Correlated Double Sampling:
    http://tinyurl.com/4bu6sh
    "In the amateur radio was once a stand-alone VNA presented with graphic display, which offset the problem by two other measurements with 180 and 270 degree phase situation has been resolved. That is yes in principle with the N2PK-VNA also, and would only have to rein in the software."
11. Linear Technology application note describing Correlated Double Sampling:
    http://www.linear.com/pc/downloadDocument.do?navId=H0,C1,C1154,C1009,C1099,P1560,D4960
12. For examples of hermaphroditic connectors such as the APC7 (7mm) or the GR900 (14 mm) connectors, see:
    http://ece-www.colorado.edu/~kuester/Coax/connchart.htm