Quantifying Variance Components for Repeated Scattering-Parameter Measurements*
Amanda Koepke and Jeffrey A. Jargon
National Institute of Standards and Technology, 325 Broadway, Boulder, CO 80305 USA
Email: amanda.koepke@nist.gov, jeffrey.jargon@nist.gov

Abstract — We quantify random uncertainties for scattering-parameters repeatedly measured with a vector network analyzer, focusing on variations due to multiple calibrations, disconnects, and repeat measurements. We describe a two-stage nested design, which allows us to model the random effects, and present results for a series of coaxial measurements performed by making use of an open-short-load-thru calibration kit with Type-N connectors.

Index Terms — calibration, disconnect, measurement, repeat, scattering-parameters, two-stage nested design, variance.

I. INTRODUCTION

Uncertainties of calibrated scattering parameters (S-parameters) measured with a vector network analyzer (VNA) can be classified as either random or systematic. In previous publications, we have quantified systematic uncertainties resulting from uncertainties in the physical models of our calibration standards [1-2]. In this study, we focus on random uncertainties, specifically on variations due to calibration, disconnect, and repeat measurements.

To quantify these variance components, we set up a two-stage nested, or hierarchical, design [3], depicted in Figure 1. Nesting refers to the imposed structure of the design: within each calibration we have disconnects specific to that calibration, and similarly within each disconnect we have repeat measurements specific to that disconnect. Repeat measurements are taken without breaking any connections, while a disconnect refers to disconnecting and reconnecting the cables before taking a new set of measurements.

In the following sections, we describe the two-stage nested design in detail, and present results for a series of coaxial measurements performed by making use of an open-short-load-thru (OSLT) calibration kit with Type-N connectors.

II. TWO-STAGE NESTED DESIGN

We assume the following random effects model:

\[ Y_{ijk} = \mu + C_i + D_{(ij)} + \varepsilon_{(ijk)} \]  \hspace{1cm} (1)

Here \( Y_{ijk} \) is the multivariate response vector (i.e. calibrated magnitudes and phases of \( S_{21} \) measurements) under calibration \( i (i = 1, ..., I) \), disconnect \( j (j = 1, ..., J) \), and repeat \( k (k = 1, ..., K) \). These vectors have dimension \( F \times 1 \), where \( F \) is the number of measured frequency points. The mean response \( \mu \) is a constant, and is the expected value of \( Y_{ijk} \) over all the calibrations, disconnects, and repeats. We assume the calibration effects \( C_i \) are random variables with mean \( \bar{0} \) and covariance matrix \( \Sigma_C \), the disconnect effects \( D_{(ij)} \) are random variables with mean \( \bar{0} \) and covariance \( \Sigma_D \), and the measurement errors \( \varepsilon_{(ijk)} \) are random variables with mean \( \bar{0} \) and covariance \( \Sigma \). All these effects are assumed to be independent. The notations \( (ij) \) and \( (ijk) \) denote nesting. The diagonal elements of the covariance matrices \( \Sigma_C, \Sigma_D, \) and \( \Sigma \) are the variance component vectors \( \sigma^2_C, \sigma^2_D, \) and \( \sigma^2 \).

We are interested in the variance of the overall mean, \( \bar{Y} \), and the variance components vectors \( \sigma^2_C, \sigma^2_D, \) and \( \sigma^2 \). These can be estimated using a Multivariate Analysis of Variance (MANOVA) method [4]. Using this approach, we partition the sums of squares and cross products (SSCP) into three parts, \( \text{SSCP}_{\text{Total}} = \text{SSCP}_C + \text{SSCP}_D + \text{SSCP}_\varepsilon \) where \( \text{SSCP}_C, \text{SSCP}_D, \) and \( \text{SSCP}_\varepsilon \) are the sums of squares and cross products due to calibration, disconnect, and error. We use these to calculate the estimated mean square errors (MSE) due to the different factors, and from the MSE we estimate the variance of the overall mean and the variance components \( \sigma^2_C, \sigma^2_D, \) and \( \sigma^2 \).

Figure 1. The two-stage nested design consisting of \( I \) calibrations, each including \( J \) disconnects and \( K \) repeated measurements taken at each disconnect. In this figure, \( J=K=5 \) for the \( i^{th} \) calibration.
To calculate the SSCP matrices, we first estimate the group means. For each disconnect \( j = 1, 2, \ldots, J \) and calibration \( i = 1, 2, \ldots, I \), we take the means of all repeat measurements

\[
\bar{Y}_{ij} = \frac{1}{K} \sum_{k=1}^{K} Y_{ijk}.
\]

(2)

Next, for each calibration \( i = 1, 2, \ldots, I \), we calculate the means of all \( J \) disconnects

\[
\bar{Y}_i = \frac{1}{J} \sum_{j=1}^{J} \bar{Y}_{ij}.
\]

(3)

Finally, the overall mean is calculated as

\[
\bar{Y} = \frac{1}{I} \sum_{i=1}^{I} \bar{Y}_i.
\]

(4)

Using these means, we calculate the sums of squares and cross products due to calibration (SSCP\(_c\)), disconnect (SSCP\(_d\)), and error (SSCP\(_e\)).

For the error, we have:

\[
\text{SSCP}_e = \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{K} (Y_{ijk} - \bar{Y}_{ij}) (Y_{ijk} - \bar{Y}_{ij})^T - \bar{Y}_{ij} (\bar{Y}_{ij} - \bar{Y}_{i}) (\bar{Y}_{ij} - \bar{Y}_{i})^T
\]

\[= \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{K} Y_{ijk} Y_{ijk}^T - K \sum_{i=1}^{I} \sum_{j=1}^{J} \bar{Y}_{ij} \bar{Y}_{ij}^T,
\]

(5)

which looks at deviations of individual measurements from their calibration/disconnect level group mean. This matrix has dimension \( F \times F \). Similarly, for disconnect and calibration we have:

\[
\text{SSCP}_d = K \sum_{i=1}^{I} \sum_{j=1}^{J} (\bar{Y}_{ij} - \bar{Y}_i) (\bar{Y}_{ij} - \bar{Y}_i)^T
\]

\[= K \sum_{i=1}^{I} \sum_{j=1}^{J} \bar{Y}_{ij} \bar{Y}_{ij}^T - K \sum_{i=1}^{I} \bar{Y}_i \bar{Y}_i^T,
\]

(7)

and

\[
\text{SSCP}_c = J \sum_{i=1}^{I} (\bar{Y}_i - \bar{Y}_.) (\bar{Y}_i - \bar{Y}_.)^T
\]

\[= J \sum_{i=1}^{I} \bar{Y}_i \bar{Y}_i^T - IJ \bar{Y}_. \bar{Y}_.^T.
\]

(8)

For computational ease and notational clarity, we just consider the \( f^{th} \) diagonal elements of these matrices, as follows:

\[
\text{SSP}_e(f) = \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{K} Y_{ijk}(f) - K \sum_{i=1}^{I} \sum_{j=1}^{J} \bar{Y}_{ij}^2(f),
\]

(11)

\[
\text{SSP}_d(f) = K \sum_{i=1}^{I} \sum_{j=1}^{J} \bar{Y}_{ij}^2(f) - IJ \bar{Y}_i \bar{Y}_i^T(f),
\]

(12)

and

\[
\text{SSP}_c(f) = IJ \bar{Y}_i \bar{Y}_i^T(f) - IJ \bar{Y}_i \bar{Y}_i^T.
\]

(13)

The expected values of these mean squares are given in Table 1. By equating these expected mean squares with their counterparts estimated from the data and solving for the variance components is close to zero. If there is no variability due to error, disconnect, and calibration, and since the variance components are nonnegative, \( E(\text{MS}_c(f)) \geq E(\text{MS}_d(f)) \geq E(\text{MS}_e(f)) \). However, since we are using mean squares estimated from the data to obtain our variance components estimates, there is a chance that \( \text{MS}_c(f) < \text{MS}_d(f) \) or \( \text{MS}_d(f) < \text{MS}_e(f) \), resulting in a negative variance component estimate. When this happens, we set the estimate equal to zero [5].

Notice the expected mean square error for calibration includes variability due to error, disconnect, and calibration, and since the variance components are nonnegative, \( E(\text{MS}_c(f)) \geq E(\text{MS}_d(f)) \geq E(\text{MS}_e(f)) \). However, since we are using mean squares estimated from the data to obtain our variance components estimates, there is a chance that \( \text{MS}_c(f) < \text{MS}_d(f) \) or \( \text{MS}_d(f) < \text{MS}_e(f) \), resulting in a negative variance component estimate. When this happens, we set the estimate equal to zero [5].

We can also use the mean squares to calculate the variance of the overall mean. Typically, the diagonal entries of the covariance matrix for \( \bar{Y}_. \) are calculated as:

\[
\text{Var}(\bar{Y}_.(f)) = \text{MS}_c(f) / IJ.
\]

(17)

However, due to the possibility of obtaining negative estimates for the variance components, this may underestimate the variance. This would occur when one or more of the variance components is close to zero. If there is no variability due to disconnect or calibration, \( E(\text{MS}_c(f)) = E(\text{MS}_d(f)) = E(\text{MS}_e(f)) = \sigma^2 \). In this case, all three estimated mean squares are estimates of the same quantity of interest, and it makes sense to use a weighted average of these terms, with the degrees of freedom as the weights to obtain the variance of the overall average:
### III. MEASUREMENTS

We examined variance components by performing four calibrations ($I=4$) using a single Open-Short-Load-Thru (OSLT) calibration kit with Type-N coaxial connectors. Physical models of the calibration standards were developed and validated with a multiline Thru-Reflect-Line (TRL) calibration within the NIST Microwave Uncertainty Framework [2]. Within each calibration, we connected and disconnected a 20-dB attenuator five times ($J=5$), and made repeated measurements five times ($K=5$) during each connection, as illustrated in Figure 1. Thus, the attenuator was measured 100 (4×5×5) times. All measurements were performed on a frequency grid from 0.2-18 GHz in steps of 0.2 GHz (360 points).

Figures 2 and 3 plot the overall means ($\bar{Y}_-$) of the attenuator’s measured magnitude and detrended phases of the transmission coefficient ($S_{21}$) and Arg{$S_{21}$}. These two figures also display the means of the five disconnects and five repeat measurements for each of the four calibrations ($\bar{Y}_1, \bar{Y}_2, \bar{Y}_3, and \bar{Y}_4$). Variations among the calibrations are clearly visible.

Figures 4 and 5 plot the square root of variance component estimates, which correspond to the variation due to error (σ), the variation due to disconnect (σ_d), and the variation due to calibration (σ_c). For both $|S_{21}|$ and Arg{$S_{21}$}, the variation due to error, $\sigma$, is negligible compared to the variation due to disconnect and calibration, $\sigma_d$ and $\sigma_c$. For $|S_{21}|$, the maximum value of $\sigma_d$ is 0.11 dB and the maximum value of $\sigma_c$ is 0.13 dB. For Arg{$S_{21}$} the maximum value of $\sigma_d$ is 0.48 degrees and the maximum value of $\sigma_c$ is 0.52 degrees. Surprisingly, these estimates do not gradually increase with frequency, but rather rise and fall unexpectedly within the measured frequency range. Furthermore, neither the values of $\sigma_d(f)$ nor $\sigma_c(f)$ consistently dominate the total variability throughout the measured frequency range.

Figures 6 and 7 plot the uncertainties due to random and systematic effects for the attenuator’s $|S_{21}|$ and Arg{$S_{21}$} values. The random portion was calculated as the square root of the maximum of eq. (17) or (18), so $\sqrt{\text{Var}(\bar{Y}_-(f))}$. The systematic portion was determined by the Microwave Uncertainty Framework. The figures illustrate that the magnitude of the uncertainty due to random effects is on the same order as the uncertainty due to systematic effects at frequencies below about 8 GHz, while at higher frequencies the uncertainty due to systematic effects is larger. For $|S_{21}|$, the maximum value of the uncertainty due to systematic effects is 0.21 dB, and for Arg{$S_{21}$}, the maximum value is 1.79 degrees.
Fig. 2. Overall mean of the attenuator’s $|S_{21}|$ values, and means of the $J$ disconnects and $K$ repeats for each $i$th calibration.

Fig. 3. Overall mean of the attenuator’s Arg{$S_{21}$} values, and means of the $J$ disconnects and $K$ repeats for each $i$th calibration.

Fig. 4. Square root of variance component estimates for the attenuator’s $|S_{21}|$ values including error ($\hat{\sigma}$), disconnect ($\hat{\sigma}_D$), and calibration ($\hat{\sigma}_C$).

Fig. 5. Square root of variance component estimates for the attenuator’s Arg{$S_{21}$} values, including error ($\hat{\sigma}$), disconnect ($\hat{\sigma}_D$), and calibration ($\hat{\sigma}_C$).

Fig. 6. Uncertainties due to random and systematic effects for the attenuator’s $|S_{21}|$ values, calculated from repeat measurements (random) and the Microwave Uncertainty Framework (systematic).

Fig. 7. Uncertainties due to random and systematic effects for the attenuator’s Arg{$S_{21}$} values, calculated from repeat measurements (random) and the Microwave Uncertainty Framework (systematic).
Fig. 8. Residuals versus predicted values for the attenuator’s $|S_{21}|$ values. There is no obvious pattern in the residuals and there do not appear to be any outliers.

Fig. 9. Boxplots of residuals for the attenuator’s $|S_{21}|$ values, separated by calibration and colored by disconnect. The spread of the points, summarized by the boxplots, seems comparable among the different disconnects and calibrations.

We should note that it is important to check for any possible violations of the assumption that our data follows the random-effects model described in eq. 1. To do this, we calculate the residuals, $e_{ijk} = Y_{ijk} - \hat{Y}_{ijk}$, where $\hat{Y}_{ijk}$ is the predicted value of $Y_{ijk}$. Formally $\hat{Y}_{ijk} = \mu + \tilde{C}_i + \tilde{D}_{ij}$, where $\mu = \bar{Y}_\ldots$, $\tilde{C}_i = \bar{Y}_i - \bar{Y}_\ldots$, and $\tilde{D}_{ij} = \bar{Y}_{ij} - \bar{Y}_{i\ldots}$. For the two-stage nested design. Thus $\hat{Y}_{ijk} = \bar{Y}_{ij\ldots}$, giving $e_{ijk} = Y_{ijk} - \bar{Y}_{ij\ldots}$. Plotting these residuals versus the predicted values, as shown in Figure 8 for the attenuator’s $|S_{21}|$ measurements, allows us to look for any patterns that might suggest nonlinearity or points that might be outliers. These assumptions do not appear to be violated, as this plot resembles a random cloud of points with no obvious pattern in the residuals. Additionally, for the random effects model in eq. 1, we assume that the within-calibration variability is the same across the four calibrations. This can be checked with a plot of the residuals versus calibration, shown in Figure 9. We use boxplots of the residuals to summarize the many residuals, and separate by disconnect. A boxplot displays the distribution of a set of points. The bottom and top of the box denote the 25th and 75th percentiles of the data, so 50% of the data falls between these lines. The line in the box denotes the median (the 50th percentile). The thin vertical line and the points show the spread of the rest of the data. The points denote data that lies 1.5*IQR away from the ends of the box, where IQR (the interquartile range) is calculated as the distance between the 25th and 75th percentiles. The boxplots in Figure 9 indicate that the spread seems comparable among the different disconnects and calibrations for the attenuator’s $|S_{21}|$ measurements. Residual analysis for the attenuator’s Arg{$S_{21}$} measurements showed similar results.

IV. CONCLUSIONS

We described a two-stage nested design that allows us to quantify random uncertainties for repeated S-parameters measurements, and focused on variations due to multiple calibrations, disconnects, and repeat measurements. Furthermore, we presented results for a series of coaxial measurements performed by making use of OSLT calibrations with Type-N connectors, and discovered our variance estimates did not gradually increase with frequency, but rather rose and fell unexpectedly within the measured frequency range. Neither the estimated calibration or disconnect variances consistently dominated the overall variance throughout the measured frequency range. And in the case presented, the systematic uncertainty values were significantly greater than the random uncertainties values at higher frequencies.

The method we presented may be used for examining a host of other calibration techniques with varying connector sizes and environments, such as waveguide and on-wafer. We conjecture other such combinations could provide dramatically different results.

From our experience with this experiment, because the variation due to error is negligible compared to the variation due to disconnect and calibration, we conclude that random uncertainties can best be captured by making multiple disconnects of any devices-under-test of interest as well as performing multiple calibrations. Simply making repeat measurements on a device without disconnecting and reconnecting it offers little insight regarding the overall random uncertainty.
ACKNOWLEDGEMENTS

*This work was supported by the U.S. government, and is not subject to U.S. copyright.

The authors thank Kevin Coakley, Mike Frey, Paul Hale, Michael Janezic, Aric Sanders, Jolene Splett, Sarah Streett, and Mitch Wallis for their helpful comments.

REFERENCES


