



Consiglio Nazionale delle Ricerche

Free-space measurement
of low-reflectivity
targets

A.Simonetto, E.Sassolini

FP 16/04

November 2016

ISTITUTO DI FISICA DEL PLASMA
"Piero Caldirola"

Via R. Cozzi 53 - 20125 Milano (Italy)

Introduction

The measurement of free-space reflectivity of well matched and relatively small targets is a challenging task. The technique described in [1] and [2] was implemented in our laboratory using a FORTRAN program for data processing. The program is available in the same location as this report. This document serves as documentation to the software, and describes data processing and validation tests. Both measurements with good data and with several problems are shown, allowing assessment of the robustness of the technique.

Theory

When a load ρ_L is measured with a VNA through a two-port error matrix \underline{S} , i.e. a virtual scattering matrix including all the non-ideal instrument behaviour, the measured quantity ρ_M is:

$$\rho_M = S_{11} + \frac{S_{12}S_{21}}{1 - S_{22}\rho_L} \rho_L = \frac{(S_{12}S_{21} - S_{11}S_{22})\rho_L + S_{11}}{-S_{22}\rho_L + 1} \equiv \frac{a\rho_L + b}{c\rho_L + 1} \quad (1)$$

where S_{nm} are the elements of the matrix and the load is connected to port 2.

Assuming that the phase of the load is varied at constant amplitude, the measured value describes a circle in complex space, with center X and radius R given by (see e.g.[3]):

$$X = \frac{b - ac^*|\rho_L|^2}{1 - |c|^2|\rho_L|^2} = \frac{S_{12}S_{21}S_{22}^*|\rho_L|^2}{1 - |S_{22}|^2|\rho_L|^2} + S_{11} \quad (2)$$

$$R = \frac{|a - bc||\rho_L|}{1 - |c|^2|\rho_L|^2} = \frac{|S_{12}S_{21}||\rho_L|}{1 - |S_{22}|^2|\rho_L|^2} \quad (3)$$

If the load is a variable short, $|\rho_L|=1$, then

$$X_1 = \frac{S_{12}S_{21}S_{22}^*}{1 - |S_{22}|^2} + S_{11} \quad (4)$$

$$R_1 = \frac{|S_{12}S_{21}|}{1 - |S_{22}|^2} \quad (5)$$

If the load is instead a variable match, i.e. $|\rho_L| \ll \min(1, 1/|S_{22}|)$ (the last term is not superfluous, since the error matrix is not physical and in principle does not satisfy the usual conservation laws)

$$X_0 = \frac{S_{12}S_{21}S_{22}^*|\rho_L|^2}{1 - |S_{22}|^2|\rho_L|^2} + S_{11} \xrightarrow{|\rho_L| \rightarrow 0} S_{11} \quad (6)$$

$$R_0 = \frac{|S_{12}S_{21}||\rho_L|}{1 - |S_{22}|^2|\rho_L|^2} \xrightarrow{|\rho_L| \rightarrow 0} |S_{12}S_{21}||\rho_L| \quad (7)$$

The quantity of interest, $|\rho_L|$, can then be obtained to first order from

$$|\rho_L| \equiv \frac{R_0}{R_1} \times \frac{1}{1 - \left| \frac{X_1 - X_0}{R_1} \right|^2} \quad (8)$$

One should note that the second term in (8) is a small correction accounting for the finite value of S_{22} . In all the measurements performed it was negligibly different from 1, so it was ignored.

Error bar

Standard error propagation from the quantities extracted from data shows

$$\begin{aligned} \sigma_{|\rho_L|}^2 = & (\partial_{R_1} |\rho|)^2 \sigma_{R_1}^2 + (\partial_{R_0} |\rho|)^2 \sigma_{R_0}^2 + (\partial_{x_1} |\rho|)^2 \sigma_{x_1}^2 + \\ & + (\partial_{x_0} |\rho|)^2 \sigma_{x_0}^2 + (\partial_{y_1} |\rho|)^2 \sigma_{y_1}^2 + (\partial_{y_0} |\rho|)^2 \sigma_{y_0}^2 \end{aligned} \quad (9)$$

where x_1, x_0, y_1, y_0 are real (x) and imaginary (y) parts of X_1 and X_0 .

Computing all derivatives and discarding terms of order $(|X_1 - X_0|/R_1)^2$, one obtains

$$\sigma_{|\rho_L|}^2 \equiv |\rho_L|^2 \left\{ \frac{\sigma_{R_0}^2}{R_0^2} + \frac{\sigma_{R_1}^2}{R_1^2} \right\} \quad (10)$$

Software

A small FORTRAN program was written to process the data. The program reads two files, `short.txt` and `load.txt`, simply consisting of lists of VNA data files for a mirror and for a good load (the device under test) at different positions along the incident beam. The positions are not restricted, but should cover as uniformly as possible a half-wavelength at the minimum frequency of interest. The range of spatial scan should not be excessive, to avoid modifications to the geometry.

Complex numbers are mapped on pairs of real numbers

$$X = [\text{Re}(X), \text{Im}(X)] \equiv [x, y] \quad (11)$$

and circles in this space are fitted using a non-iterative algorithm [4]: a standard least-squares fit is applied to

$$f(x_0, y_0, R_{sq}) = \sum_{i=1}^N \left[(x_i - x_0)^2 + (y_i - y_0)^2 - R_{sq} \right]^2 \quad (12)$$

where $R_{qs} = R^2$.

As a result, the radius R of the best fitting circle is

$$R = \sqrt{\frac{1}{N} \sum (x_i - x_0)^2 + (y_i - y_0)^2} \quad (13)$$

with (x_0, y_0) minimizing the sum of deviations. The center (x_0, y_0) is the solution of equation

$$\underline{G} \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} = \underline{C} \quad (14)$$

where the matrix G and the vector C are given by

$$G_{11} = 8 \sum_{i=1}^N x_i^2 - \frac{8}{N} \left(\sum_{i=1}^N x_i \right)^2 \quad (15)$$

$$G_{12} = G_{21} = 8 \sum_{i=1}^N x_i y_i - \frac{8}{N} \sum_{i=1}^N x_i \sum_{i=1}^N y_i \quad (16)$$

$$G_{22} = 8 \sum_{i=1}^N y_i^2 - \frac{8}{N} \left(\sum_{i=1}^N y_i \right)^2 \quad (17)$$

$$C_1 = -\frac{4}{N} \sum_{i=1}^N x_i^2 \sum_{j=1}^N x_j - \frac{4}{N} \sum_{j=1}^N y_j^2 \sum_{k=1}^N x_k + 4 \sum_{j=1}^N y_j^2 x_i + 4 \sum_{i=1}^N x_i^3 \quad (18)$$

$$C_2 = -\frac{4}{N} \sum_{i=1}^N y_i^2 \sum_{j=1}^N y_j - \frac{4}{N} \sum_{j=1}^N x_j^2 \sum_{k=1}^N y_k + 4 \sum_{j=1}^N x_j^2 y_j + 4 \sum_{i=1}^N y_i^3 \quad (19)$$

This algorithm is very efficient, but only working for data that approximately lie on a circle, giving unexpected results otherwise. For example, if all data lie on a circle except a very far one, the best fitting circle passes exactly through the external datum and approximately through the center of the circle formed by the other data, which is indeed the best-fitting circle, but not the intended one.

Such data can be present in frequency sweeps when the VNA fails to lock (more details below). To prevent garbage data to affect the output, the fit is made in a loop, discarding at each step all the data whose distance from the center exceeds a threshold. The threshold is set at present to 1.8 times the radius. To prevent the algorithm to run away, the loop is limited to three iterations.

To help spotting the occurrence of garbage data, a different algorithm, based on standard statistical quantities, was also implemented.

If data are uniformly distributed on a circle, its center is the barycenter of the data

$$\begin{bmatrix} x_0 \\ y_0 \end{bmatrix} = \frac{1}{N} \sum_{i=1}^N \begin{bmatrix} x_i \\ y_i \end{bmatrix} \quad (20)$$

and its radius is the standard deviation of the data as in (13).

If the distribution of data on the circle is non-uniform, (20) is only approximately true, but is robust to the presence of very noisy data.

The program writes frequency-dependent results with both approaches in two tab-separated files (`fitresult.txt`, `statresult.txt`).

Each file contains a header and records for all frequencies. Each line contains

<i>frequency</i> [GHz]	
<i> ρ </i> [dB]	from (8) without correction term
<i> ρ (1+σ/ρ) </i> [dB]	upper 67% error bar, computed using (10)
<i> ρ (1-σ/ρ) </i> [dB]	lower 67% error bar, computed using (10)
<i>Max</i> [dB]	upper peak-to-peak value (i.e. Max(R _{0i})/min(R _{1j}))
<i>min</i> [dB]	lower peak-to-peak value (i.e. min(R _{0i})/Max(R _{1j}))
<i> ρ </i> [dB]	from (8) including the correction term
<i>correction</i>	correction term in (8) (normally very close to 1)
<i>Re(X₁)</i>	center of circle through short data, real part
<i>Im(X₁)</i>	center of circle through short data, imaginary part
<i>R₁</i>	radius of circle through short data
<i>Re(X₀)</i>	center of circle through load data, real part
<i>Im(X₀)</i>	center of circle through load data, imaginary part
<i>R₀</i>	radius of circle through load data
<i>ε(R₁)/R₁</i>	fractional error on radius of circle through short data, square root of of last term in (10)
<i>ε(R₀)/R₀</i>	fractional error on radius of circle through load data, square root of of last-but-one term in (10)

Experimental results show that the circle-fitting algorithm always works very well for mirror data, so that another file (`result.txt`) is written, holding results obtained with the fit for short data and average, standard deviation for the load.

This file contains a header and records for all frequencies. Each line contains

frequency[GHz]

$|\rho| [dB]$ from (8) without correction term
 $|\rho (1+\sigma/\rho)| [dB]$ upper 67% error bar, computed using (10)
 $|\rho (1-\sigma/\rho)| [dB]$ lower 67% error bar, computed using (10)
 $Max [dB]$ upper peak-to-peak value (i.e. $\text{Max}(R_{0i})/\text{min}(R_{1j})$)
 $min [dB]$ lower peak-to-peak value (i.e. $\text{min}(R_{0i})/\text{Max}(R_{1j})$)

All results are mutually consistent for good data.

Further output files (`s###.#.txt`, `l###.#.txt`) are written for short and load data at 10 equidistant frequencies (for the short) and at each frequency (for the load), representing the circles in the complex plane.

Each file contains a header line and then 360 (or more) lines containing

i	line number	
$R_{0,i}\cos(i \text{ deg})+\text{Re}(X_{0,i})$		Best-fitting circle from fit
$R_{0,i}\sin(i \text{ deg})+\text{Im}(X_{0,i})$		
$R_{0,i}(1+\varepsilon/R)\cos(i \text{ deg})+\text{Re}(X_{0,i})$		Upper error bar of best-fitting circle from fit
$R_{0,i}(1+\varepsilon/R)\sin(i \text{ deg})+\text{Im}(X_{0,i})$		
$R_{0,i}(1-\varepsilon/R)\cos(i \text{ deg})+\text{Re}(X_{0,i})$		Lower error bar of best-fitting circle from fit
$R_{0,i}(1-\varepsilon/R)\sin(i \text{ deg})+\text{Im}(X_{0,i})$		
$R_{0,i}(1+Max/R)\cos(i \text{ deg})+\text{Re}(X_{0,i})$		Upper peak-to-peak error bar from fit
$R_{0,i}(1+Max/R)\sin(i \text{ deg})+\text{Im}(X_{0,i})$		
$R_{0,i}(1+min/R)\cos(i \text{ deg})+\text{Re}(X_{0,i})$		Lower peak-to-peak error bar from fit
$R_{0,i}(1+min/R)\sin(i \text{ deg})+\text{Im}(X_{0,i})$		
<i><same quantities computed with average, standard deviation></i>		
x_i		i-th measurement
y_i		

where $0,1$ is 0 for `l###.#.txt` files and 1 for `s###.#.txt` ones.

A `gnuplot` script is generated to plot all the circles from these files.

Finally, another text file, `NLoadsUsed.txt`, is written to document the actual amount of data used in fits. It contains a header including the number of positions available and then one line per frequency point, containing the frequency and the number of positions used. For position sweeps at fixed frequency the number is always the same as the number of positions available, for frequency sweeps it is normally different, because of the out-of-range data ignored in fits.

The table below summarizes inputs and outputs.

INPUT	DESCRIPTION
<code>short.txt</code>	list of files for <i>short</i> measurements
<code><files></code>	text files (in VNA format) named in <code>short.txt</code>
<code>load.txt</code>	list of files for <i>load</i> measurements
<code><files></code>	text files (in VNA format) named in <code>load.txt</code>
OUTPUT	DESCRIPTION
<code>fitresult.txt</code>	result obtained with fit (13-19)
<code>statresult.txt</code>	result obtained with average, standard deviation (13, 20)
<code>result.txt</code>	result obtained using fit for short and average, standard deviation for load
<code>s###.#.txt</code>	at frequency ###.# (10 across the band): measured short data, fitted circles in complex plane
<code>l###.#.txt</code>	same as above, at all frequencies, for load data
<code>gnuplotcmd</code>	gnuplot script to plot <code>s###.#.txt</code> and <code>l###.#.txt</code> files into <code>.png</code> files
<code>NLoadsUsed.txt</code>	number of load datasets used at each frequency

Measurements

A mirror with the same size as the samples to be measured was mounted on a linear translation stage directly controlled by the VNA software. Two-dimension sweeps in space/frequency or frequency/space were made with the setup shown in Figure 1.

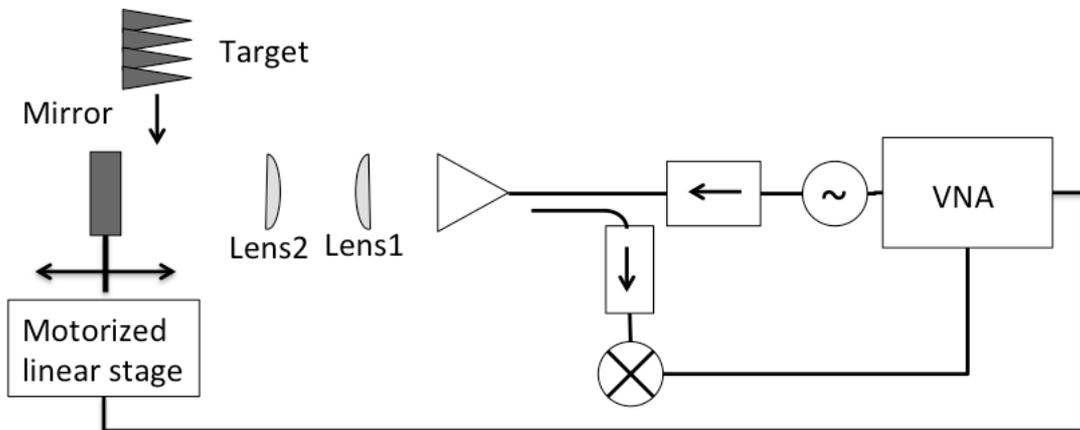


Figure 1: measurement scheme

The frequency step cannot be made too small, to keep measurement times within practicable limits. But the length of the free-space path, while partially compensated with proper cable lengths, still demands high frequency stability. As a consequence, due to the finite response time of the VNA, the frequency sweep must be either slowed down to impracticable limits, or made by phase-locking the VNA to an external reference at each step. Still, the acquisition time should be kept at a minimum, and occasionally the phase-lock may fail, providing data with a stray phase, that must be accounted for in processing.

Because of this limitation, spatial scans at fixed frequency, i.e. frequency/space 2D sweeps were found more efficient. This may vary depending on the frequency range.

Results

The figures below show the output of the *gnuplot* script for a case of good data at 32.5 GHz in a 5mm scan.

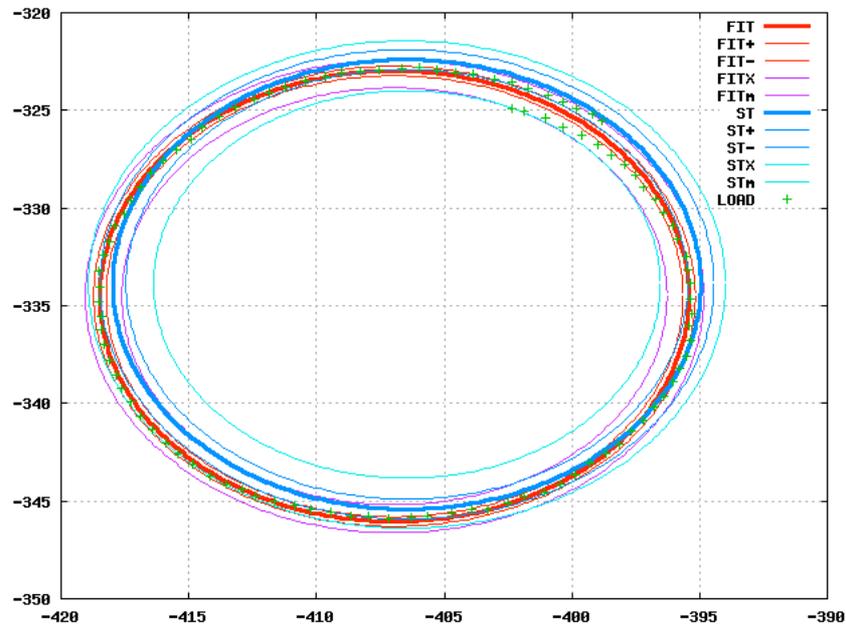


Figure 2: low reflectivity target. Measured data (crosses), best-fit circle (red), circle from statistical quantities (blue)

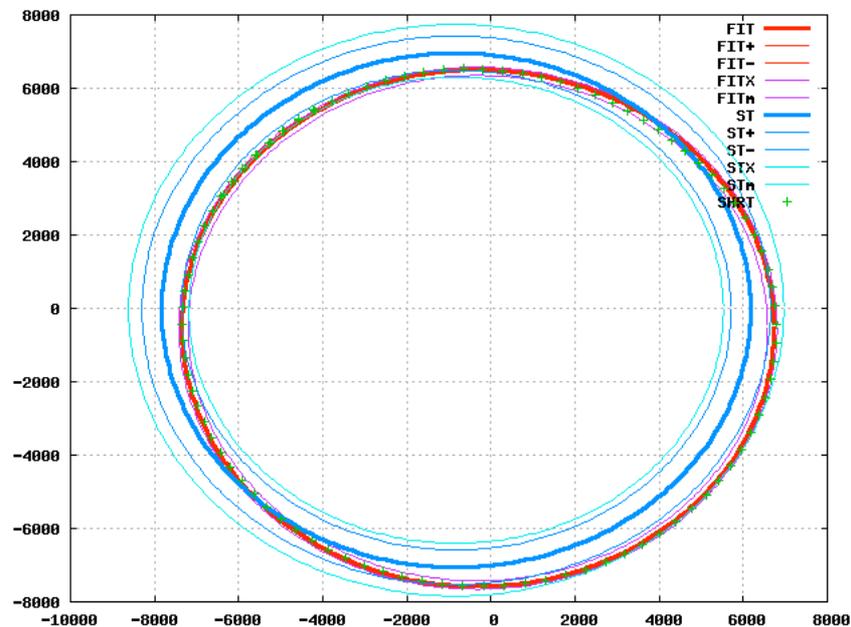


Figure 3: reference mirror. Measured data (crosses), best-fit circle (red), circle from statistical quantities (blue)

The crosses are the experimental data, thick red and blue circles represent the best fit and the estimate from statistical parameters, the light red and blue represent the outer, inner and peak to peak error bars.

One can see that the estimate of center from the fit procedure (red) is more accurate in this case, as it should. But the estimated radius is very similar, so that the result is similar in both

cases, as shown in Figure 4, where the values in all three output files is plotted as a function of frequency.

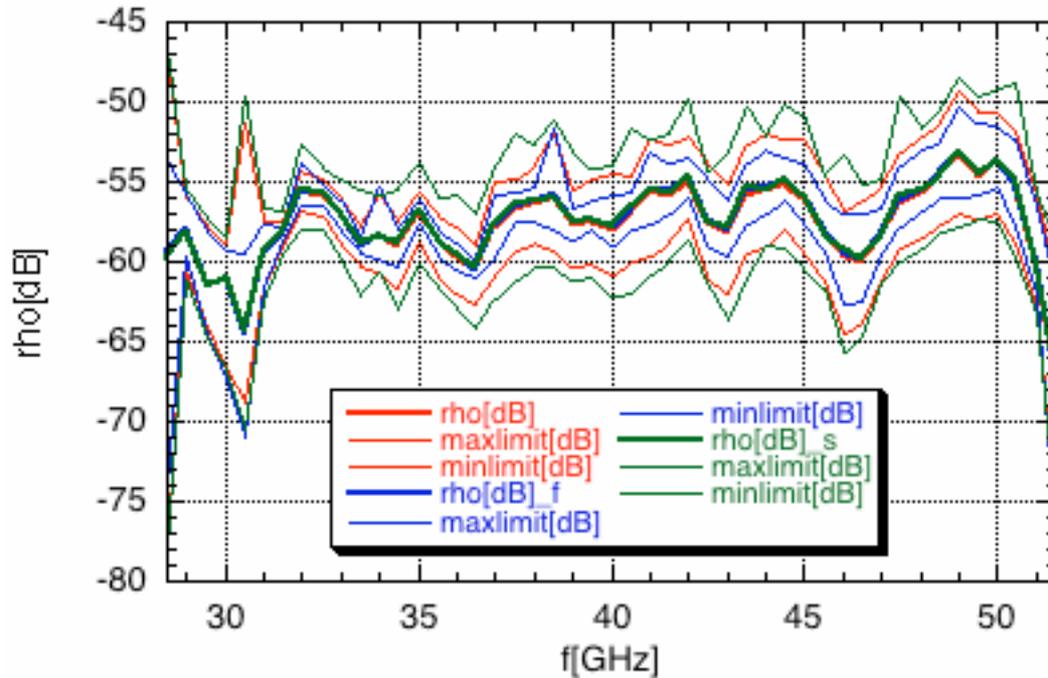


Figure 4: estimated target reflectivity. Red: from file result.txt, blue: from file fitresult.txt, green: from file statresult.txt. Peak-to-peak error bars (thin lines) are shown in the same colour as the reflectivity (thick lines)

As one can see, the three traces overlap. Peak-to-peak error bars are shown because the 1- σ ones are too small to see on this scale. Estimating the center of circles from the average is robust but inaccurate, and thus the (peak-to-peak) error bars are largest in this case. At the very edges of the band (Q-band in this case, 33-50GHz nominal range) the beam parameters are quite different, which might explain the unusual behaviour. Figure 5 below shows the same data as Figure 2, but at 28.5 GHz, i.e. the point at extreme left in Figure 4. One can see that systematic errors affect the measurement, so that the circle is a sort of spiral.

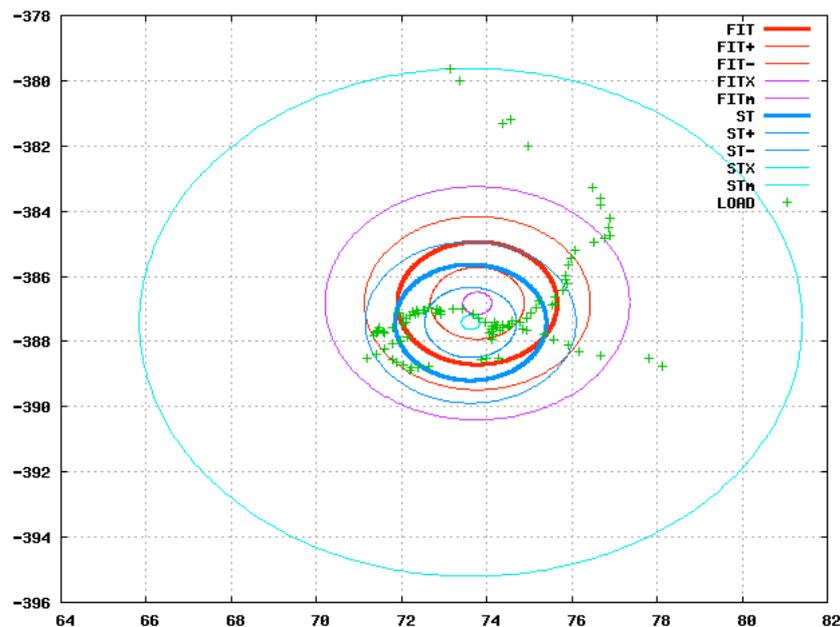


Figure 5: same as Figure 2, at 28.5 GHz

Presumably, the beam is so distorted at this frequency that the effectiveness of anechoic screening is insufficient. These systematic errors explain the large (peak-to-peak) error bars, but the estimated radius has the proper order of magnitude, so that the result is in line with the rest of the data.

Figure 6 below shows a case of frequency sweep affected by locking problems.

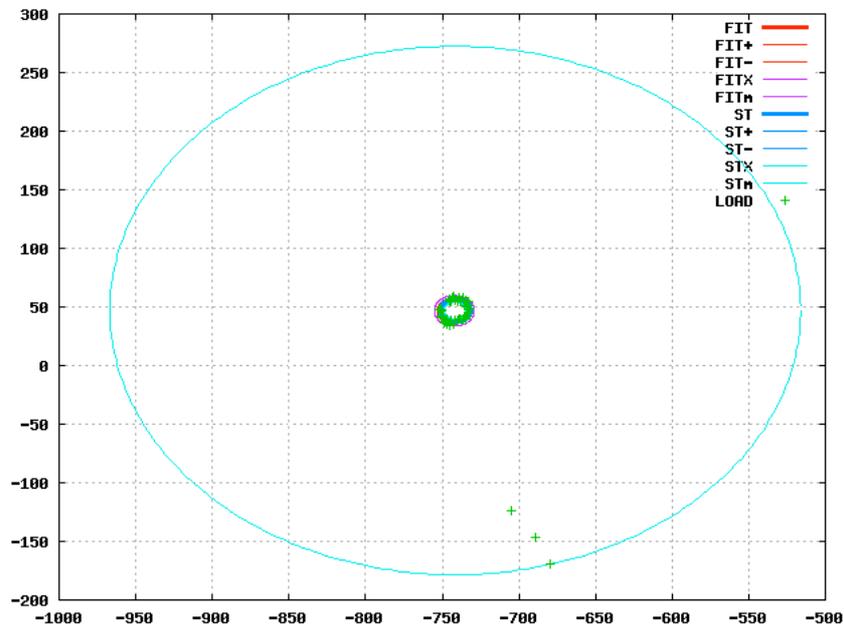


Figure 6: same as Figure 2, at 45 GHz in a V-band frequency sweep

As one can see, the three data points at the bottom of the figure are ignored in the analysis, so that the results are similar to those measured in Q-band, but the peak-to-peak error bars become huge because of the garbage data.

Figure 7 and Figure 8 show the regions of overlap between Ka-Q and Q-V bands on the same target.

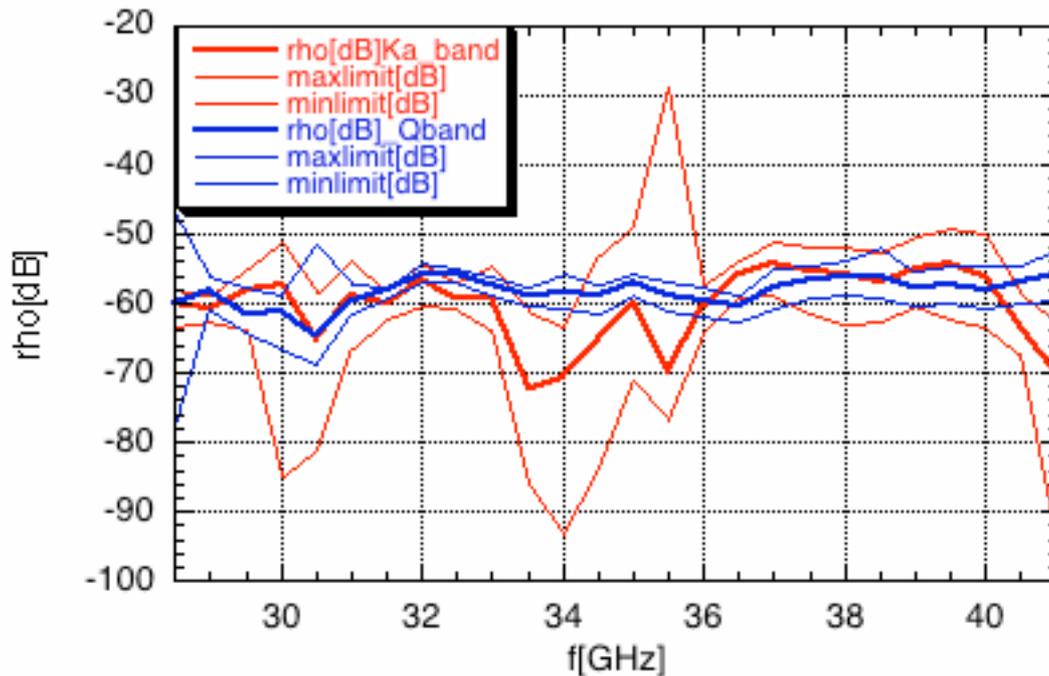


Figure 7: overlap between Ka and Q bands, on the same target.

As one can see in Figure 7, the Ka-band data are rather poor. Close examination of the plots in complex plane shows a high level of systematic errors. The antenna was later found misaligned, which may indeed be the cause of the problem. The peak-to-peak error bars are extremely sensitive to this. Nonetheless, when data are reasonably good, the measurements in two bands (with different antennas and different beam parameters) give similar results.

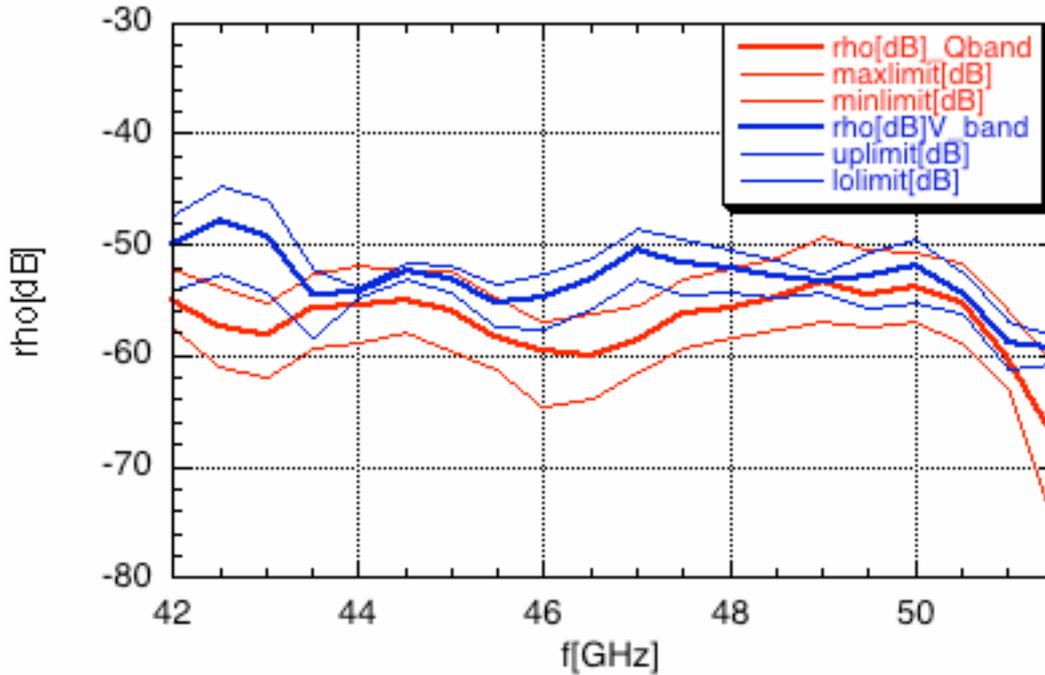


Figure 8: overlap between Q and V bands, on the same target

Figure 8 shows the region of overlap between Q and V bands, for the same target. V-band measurements were taken as frequency sweeps, and are thus affected by garbage data due to missed phase-lock. This is why the peak-to-peak error bars are replaced with the $1\text{-}\sigma$ error bars for the blue traces in the plot. Again, one can see that where error bars are lower, the estimates are in good agreement, whereas close examination of curves in the complex plane allows to pin-point problematic measurements.

Discussion

This measurement technique is efficient, mostly automated, and potentially faster than the one we normally use, i.e. virtual time gating on calibrated data. It also requires no calibration tools. But measurement times can only be kept short by using relatively large frequency steps, which prevents the use of time gating. As a result, there is no way to remove multipath effects from the data, and extreme care should be taken in the anechoic screening of metal structures around the targets, unless they are sufficiently large to prevent spillover losses. The achievable level of screening sets a limit to precision. Moreover, when the targets are spatially not uniform -as the one shown in Figure 1- unavoidable multipath effects can be present when the probing beam is not much larger than the size of target sub-structures. A combination of virtual time gating and this technique (on a *very* small set of spatial positions, otherwise measurement times become excessive) may provide the best precision.

References

- [1] A. Murk, A. Duric and F. Patt *Characterization of ALMA Calibration Targets*, 19th International Symposium on Space Terahertz Technology, Groningen, 28-30 April 2008. [Online] <http://www.nrao.edu/meetings/issst/papers/2008/2008530533.pdf>
- [2] A. Murk, A. Duric *ALMA Calibration Device Prototype Calibration Load Test Report* FEND-40.06.04.00-005-A-REP [Online] <https://safe.nrao.edu/wiki/pub/ALMA/CalAmp/FEND-40.06.04.00-005-A-REP.pdf>
- [3] G. F. Engen *Microwave circuit theory an foundations of microwave metrology* Peter Peregrinus Ltd on behalf of IEE, ISBN 0 86341 287 4, Chapter 9
- [4] L. Moura, R. Kitney, *A direct method for least-squares circle fitting*, Computer Physics Communications 64 (1991) 57–63