Noise Figure Measurement – A Reality Check

David Stockton GM4ZNX and Ian White GM3SEK

Many LNAs now seen in EME stations and tested at conferences have a Noise Figure (NF) that is significantly lower than the measurement uncertainty of the test equipment. This raises some very serious questions, not only about the measurements but also about claimed specifications, system performance calculations and bragging rights!

This paper will identify some of the major uncertainties in NF measurements, including some little-known problems in the older generation of test equipment (HP8970/HP346). We will also see that even modern Noise Figure Analysers have not made a major impact on measurement uncertainties, because the problems are inherent in the measurement technique.

Every measurement technique has its own characteristic problems. The problems with conventional NF measurements are discussed at length in Agilent AN57-2*, Noise Figure Measurement Accuracy – The Y-Factor Method,* and in the other references provided.[1] AN57-2 also warns against many types of "avoidable operator errors"; but this paper concentrates on the measurement uncertainties that are inherent in the technique itself.

Another problem about NF measurement is that amateur radio has a very poor collective memory. Even though most of the information in this paper has been known for 20 years or more, relatively few amateurs seem to be actively aware of it. This paper will provide a much-needed 'memory refresh' about the reality of NF measurements.

The Y-Factor Measurement Method

Almost all noise figure measurements use the 'Y-Factor method', so first we need to explain how this method *should* operate. Then we can move on to the problems.

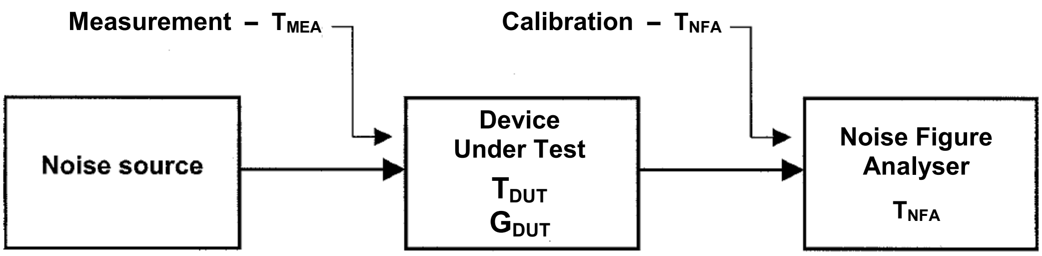

Figure 1: Test setup for Y-factor noise measurements

The Y-factor technique uses a **Noise Source** and a **Noise Figure Analyser (NFA)**. Initially the two items are connected directly together to calibrate the system. Then the Device Under Test (**DUT**) is inserted between them to measure its NF and gain.

The noise source can be switched either ON or OFF. The ON state generates noise at a calibrated power level; and even when switched OFF, the noise source still generates a

1. Most of the key references are provided as PDF files on the Conference DVD.

smaller amount of thermal noise power. The equivalent thermal noise temperatures are usually written as T_{ON} and T_{OFF}, or alternatively T_{HOT} and T_{COLD}.

The Noise Figure Analyser is an instrument for measuring *relative* noise power levels. The NFA does not need to be calibrated in terms of absolute power levels (that calibration resides in the noise source) but the NFA must be able to measure noise power *ratios* with very high accuracy and over a wide dynamic range.

The **Y factor** is simply the ratio of two noise power levels, one measured with the noise source ON and the other with the noise source OFF:

$$Y = N_{ON}/N_{OFF} \tag{1a}$$

Because noise power is proportional to noise temperature, we can also say:

$$Y = T_{ON}/T_{OFF} = T_{HOT}/T_{COLD} \tag{1b}$$

Other variants of the Y-factor method involve switching between two different noise sources, one 'hot' and the other 'cold', or even changing the physical temperature of the source. We discuss some of these options in the longer version of this paper on the Conference DVD. For now, though, we will concentrate on the most common kind of Y-factor measurements using a single switchable noise source and the companion NFA (either commercial or home made).

The advantages of this technique for a general-purpose test instrument are that the noise source can provide two different, calibrated power levels covering the entire frequency range. The use of broadband noise also makes it unnecessary to measure the noise bandwidth of either the NFA or the DUT (in most cases; but see later).

The penalty of using noise as a test signal is that noise is a statistical phenomenon, so this introduces large and unavoidable uncertainties. Accurate measurements of noise power take time. To compute an averaged value of the Y factor with a reasonably low statistical uncertainty will typically require several hundred sample measurements of N_{ON} and N_{OFF}. A modern NFA will control the repeated on/off switching of the noise source, and then do all the necessary computations to determine the NF and gain of the DUT.

NF definitions

Noise factor F is a dimensionless ratio:

$$F = (S/N \text{ at input port}) / (S/N \text{ at output port}) \tag{2a}$$

Noise Figure NF is measured in dB:

$$NF = 10\log_{10} \{(S/N \text{ at input port}) / (S/N \text{ at output port})\} \tag{2b}$$

where the S/N quantities are linear ratios between signal power and noise power.

In other words, **NF is defined as a ratio between two Y-factors**.

The engineering definitions of F and NF, derived from the basic definitions above, are:

$$F = 1 + (T_E/T_0) \tag{3a}$$

$$NF = 10 \log_{10}\{1 + (T_E/T_0)\} \quad \text{dB} \tag{3b}$$

where T_E is the effective (or equivalent) input noise temperature of the device.

Equation 3 also introduces the important reference temperature T_0 which is defined as 290 K (16.8°C, 62.2°F).

Calculations in Y-factor measurements

NF is defined at the input of the DUT but the noise power is being measured at the output, so any measurement of NF must include an accurate measurement of the device gain G_{DUT}. We can see this by developing some equations for the two-step measurement method shown in Figure 1:

Calibration (NFA only): $Y_{CAL} = N_{NFA\ ON} / N_{NFA\ OFF}$ (4)

Measurement with DUT: $Y_{MEA} = N_{MEA\ ON} / N_{MEA\ OFF}$ (5)

To find the true gain and noise temperature of the DUT, we now need to remove the effects of the NFA. Full details are given in AN57-2.

The noise source parameters T_{HOT} and T_{COLD} are available from the ENR (Excess Noise Ratio) calibration of the noise source, plus a measurement of the physical temperature T_{COLD}. ENR is formally defined as:

$$ENR = (T_{HOT} - T_{COLD}) / T_0 \quad (6a)$$

$$ENR = 10 \log_{10} \{(T_{HOT} - T_{COLD}) / T_0\} \quad dB \quad (6b)$$

where T_0 is the standard reference temperature of 290 K. ENR calibration data are normalized to T_{COLD} = 290 K, so T_{HOT} must always be corrected to account for the actual physical temperature T_{COLD} which can only be measured inside the noise source itself.

$$T_{HOT} = (ENR \times T_0) + T_{COLD} \quad (7)$$

From the calibration step which measured Y_{CAL} the NFA can calculate:

$$T_{NFA} = (T_{HOT} - Y_{CAL} T_{COLD}) / (Y_{CAL} - 1) \quad (8)$$

And the same again for the measurement with the DUT in place:

$$T_{MEA} = (T_{HOT} - Y_{MEA} T_{COLD}) / (Y_{MEA} - 1) \quad (9)$$

From the four different noise power measurements the NFA can calculate the gain of the DUT:

$$G_{DUT} = (N_{MEA\ ON} - N_{MEA\ OFF}) / (N_{NFA\ ON} - N_{NFA\ OFF}) \quad (10)$$

And finally:

$$T_{DUT} = T_{MEA} - (T_{NFA} / G_{DUT}) \quad (11)$$

The NFA will normally convert T_{DUT} into NF, and display NF_{DUT} and G_{DUT} in the normal engineering units of dB.

That is how it's all done, through quite a long chain of calculations. So where do the problems arise?

Generic Problems with Y-factor Measurements

The problems with Y-factor measurements are largely hidden in the algebra, so let's try a different view, and look at this problem through some simple graphics. Then we can go back to those equations with a better understanding of what is really happening.

The simplified case

To make this even simpler, let us temporarily assume a noise-free NFA. This removes the need for a calibration step, so we can obtain both T_{DUT} and G_{DUT} from just a single Y-factor measurement. Figure 2 illustrates this simplified case. The horizontal axis is the 'input power' to the DUT using a scale of noise temperature. The vertical axis is the 'output power' measured within the NFA, on a linear scale.

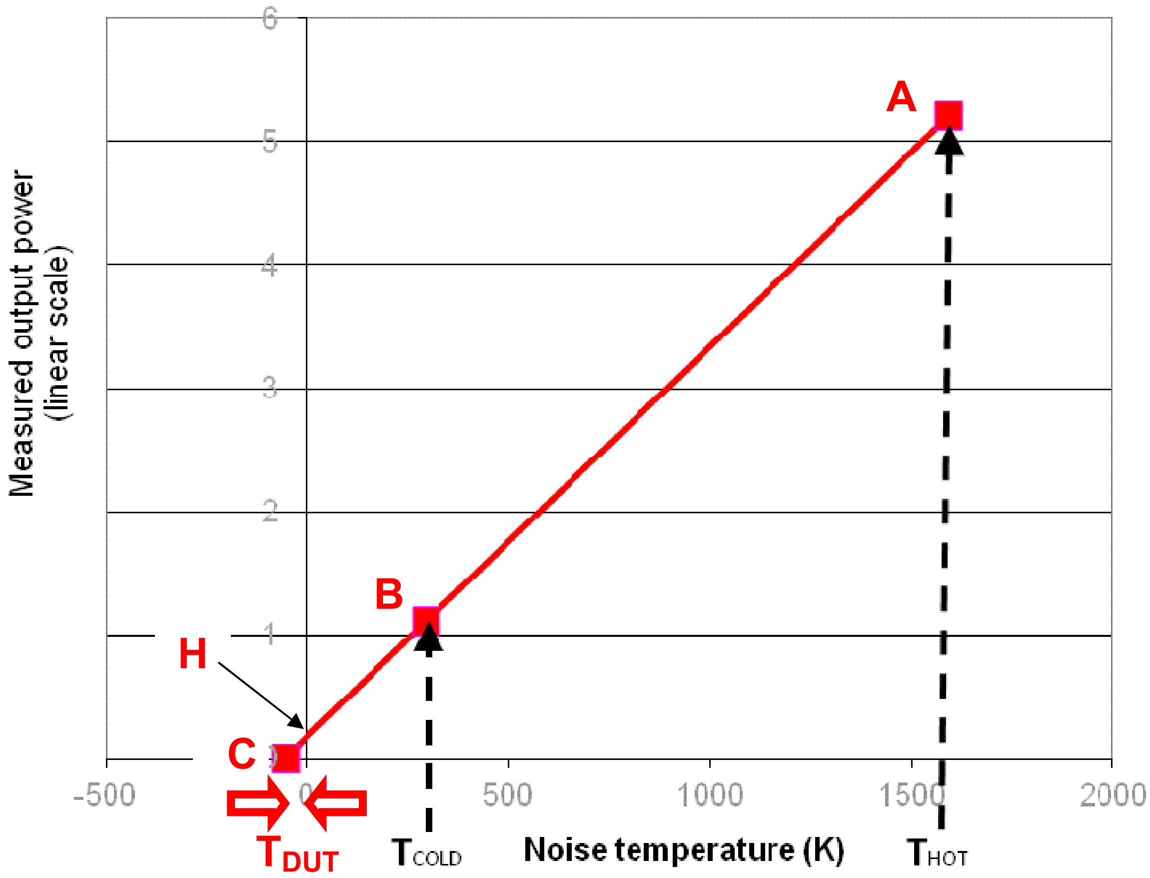

*Figure 2: Y-factor measurement requires a long extrapolation to find **T_{DUT}** at point **C**. (Simplified case with a noiseless NFA.)*

The Y-factor method always requires *pairs* of measurements. With the noise source switched on, the input power into the DUT corresponds to T_{HOT} and the measured output gives us point **A** in Figure 2. With the noise source switched off, the noise input power falls to T_{COLD} and the output power falls to give us point **B**.

Now extrapolate line **A**–**B** backward towards the origin at (0, 0). Because the DUT is generating internal noise, there would be some output power even when the input noise temperature was zero (this is point **H**). Extrapolating even further backward, the line eventually reaches zero output power at point **C**. From the geometrical construction, the intercept point **C** corresponds to $-(T_{DUT})$. Notice the minus-sign; T_{DUT} itself is positive, of course.

The slope of the line **A**–**B** gives us the gain of the DUT as shown by equation 10. We will need that value of G_{DUT} later in the calculations.

Figure 2 is drawn to scale using realistic values, so now we can see the fundamental problems with the accuracy of this method.

We are trying to measure T_{DUT} as a very small distance along the noise temperature axis, by making a long extrapolation from two measurement points **A** and **B** that are a very long way away. Any small errors in the co-ordinates of points **A** and **B** will be greatly magnified at point **C**, causing very significant errors in the final measured value of T_{DUT}.

Here is a very brief summary of the potential errors in drawing that line, **A**–**B**–**C**. Many more details are given in the longer version of this paper on the Conference DVD.

Error in noise source ENR

The **Excess Noise Ratio** (ENR) of the noise source is defined as:

$$ENR = (T_{HOT} - T_{COLD}) / T_0 \qquad (6a)$$

This value defines the horizontal distance between points **A** and **B** in Figure 2. ENR data are obtained from a factory calibration, and have a typical uncertainty of **± 0.2 dB or worse** (see later). Any error in the ENR calibration will affect the slope of line **A**–**B** and hence the intercept at point **C**.

Error in T_{COLD}

T_{COLD} is the physical temperature of the noise source. Through the definition of ENR, any error in T_{COLD} will affect T_{HOT} as well, so the whole line **A**–**B**–**C** moves sideways. Every 1°C error in T_{COLD} will cause an approximately equal error in the measured value of T_{DUT}.

When attempting to measure noise figures on the order of 0.2 dB, which is a noise temperature of only 13 K, we simply cannot afford to guess at T_{COLD}. We need to measure it accurately, **inside** the noise source if at all possible.

Gain errors in the DUT

Non-linearity in the DUT is a very serious error because it would mean that **A**–**B** is not a straight line but a curve. We also assume perfect linearity when we customarily refer all of the physical noise sources within the DUT to a single *equivalent* noise source at the input, so that the correct amount of magnified input noise will appear at the output. Curvature in the DUT transfer characteristic would mean that the linear extrapolation is no longer valid… but we don't know by how much.

Linearity is a reasonable assumption for a single-stage DUT handling very small signals; but it begins to fall apart with multi-stage LNAs that have a very high gain, and is certainly suspect when measuring complete receivers. When the overall gain is very high, some gain compression could easily occur in later stages of the DUT. We always tend to underestimate the range of peak values that are present in noise signals, and underestimate the amount of 'headroom' required to avoid peak compression that will then alter the mean value.[2]

Gain change in the DUT can be a very serious error when measuring certain types of LNA. The gain of the DUT is calculated from the measured slope of line **A**–**B**; but this depends on the hidden assumption that G_{DUT} does not change when the noise source is switched on and off. This point requires more detailed discussion later.

NFA and calibration errors

Even the simplified case of Figure 2 has revealed several serious sources of error or uncertainty. When we add in a real-life NFA, we find even more.

As shown in Figure 1, before making a real-life measurement we have to do a calibration: a hot/cold Y-factor measurement with the noise source connected directly to the NFA. Figure 3 (next page) shows this in graphical form – and once again, everything is drawn to true scale.

2. If any AGC system is active in the DUT, the measurement will be ruined – so that is another operator error to be avoided.

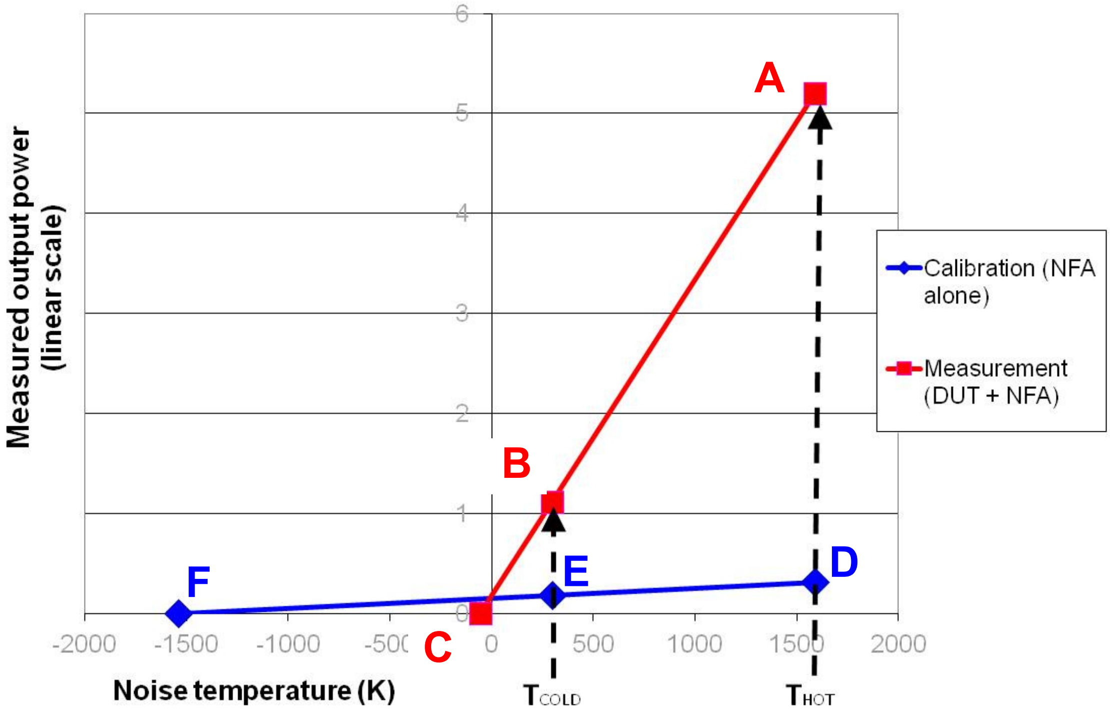

*Figure 3: A real-life measurement also includes the calibration **D–E–F**.*

The calibration step gives us a new line **D–E–F**, where point **F** represents the noise temperature of the NFA itself.[3] Because the NFA is a general-purpose broadband instrument, its NF is quite high, typically about 8 dB or 1500 K. This measurement of T_{NFA} is subject to quite large errors because of the very shallow slope and the long extrapolation to point **F**, but in most cases it plays a relatively small part in the final calculation of T_{DUT}.[4]

We are now ready to look at the errors inside a real-life NFA. Many of these errors apply particularly to the older-generation NFAs like the HP8970 series, as detailed in the longer version of this paper on the Conference DVD.

Non-linearity errors in the NFA

We can see from Figure 3 that measurements **A** and **B** take place at two quite different output levels, while **D** and **E** are at a lower level still. Non-linearity errors within the NFA can therefore be quite important, especially with a high-gain DUT which pushes point **A** even higher. Professional NFAs are designed to extreme standards of linearity. Typically they divide their IF amplification into blocks, with attenuators switched in and out automatically to extend the dynamic range. If the user makes reasonable efforts to avoid unnecessary demands on the dynamic range, the linearity error in a professional NFA should be less than ± 0.05 dB across the range required. Amateur NFAs are often much worse that that.

Noise power measurement errors

NF measurements are based on the accurate measurement of noise power. The older HP8970 series of analysers used diode detectors which had some small errors in

---

3. Point **F** represents $-(T_{NFA})$. Once again, notice the minus-sign!
4. In the final calculation of T_{DUT} using equation 11, T_{NFA} (and any associated error) will be divided by G_{DUT} which is usually a large number.

measuring noise signals with a high crest factor. The more modern Agilent NFA series use DSP which is much more accurate. An ADC takes samples of the noise voltage, which are then accurately converted to noise power density using a root-mean-square algorithm.

Frequency errors

The HP8970 and similar NFAs of that generation are liable to considerable frequency errors and drift, which can be important if the DUT is a narrowband device.

Bandwidth

The HP8970, the Maury/Ailtech equivalents and the (obsolete) Agilent base model N8972 only make measurements in a fixed 4 MHz bandwidth. This is fine for measuring wideband DUTs because the wide IF bandwidth gives rapid averaging of the noise power levels and little jitter in the results.

But amateur-band preamps and receivers may be much narrower than 4 MHz, and the problem comes when you calibrate the NFA at its own bandwidth of 4 MHz and then *change* the measurement bandwidth when you insert the narrowband DUT. This produces a measurement error because it defeats the assumptions of the NFA calculations.[5]

Summary so far

We have looked at the Y-factor measurement method from two different viewpoints, equations and graphics. The equations are definitive, but in some ways the graphics are more revealing about the elusive nature of T_{DUT} – hiding down in a corner, close to the origin (0,0) but far away from the points **A**, **B**, **D** and **E** that we actually measured.

Specific Problems with Some DUTs

All the problems identified above are generic, and in principle they can affect measurements on all different types of DUT. But the practical question is: ***How big are these errors for this LNA that I'm measuring?***

In this section we will identify specific problems that affect some DUTs much more than others.

Noise and gain contours

Noise figure, noise factor, and noise temperature are all based on a simplified model of what is really happening inside the device. This model ignores all the different physical mechanisms that create noise within the device, and replaces the whole lot by an **equivalent** amount of noise that would be created by a single noise source at the input port. For your continuing sanity, it is important to remember that this input noise source is wholly fictitious!

The main thing that we lose in referring noise parameters to the input port is that the device NF is affected by the source and load impedances. We 'patch' this problem by making the fictitious input noise source dependent on the source impedance. Device manufacturers usually construct a contour map on a Smith chart showing lines of constant NF depending on the input impedance presented to the device, and yet again we simplify reality by assuming that the contours are circular (Figure 4). On the map of complex impedance, the NF contours look like a shallow circular depression. This also

5. This error is fully discussed in AN57-2. For the purposes of this paper we regard it as an avoidable 'operator error'.

neglects the lesser effect of the load impedance on NF, which is a reasonable approximation *except* when working at frequencies where available gains are low and S_{12} is large.

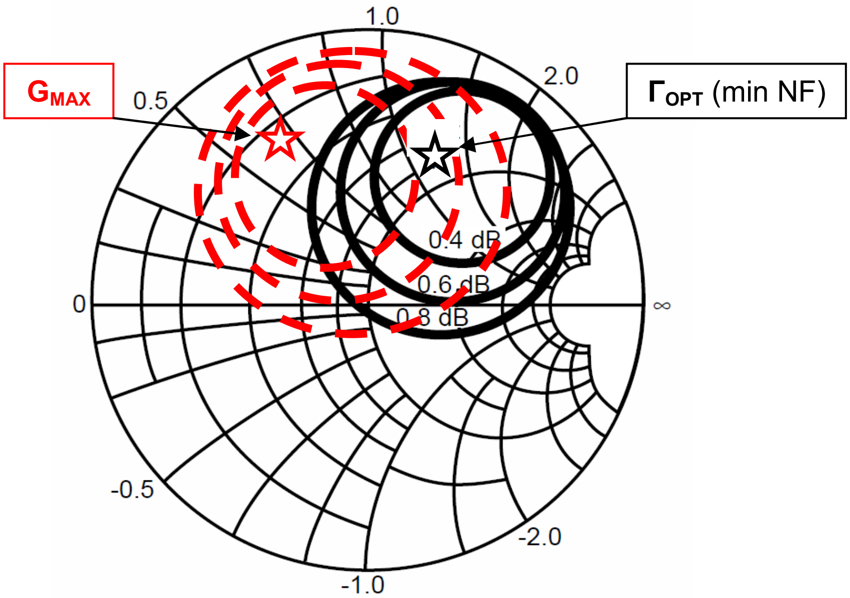

Figure 4: Typical contours of NF (assumed to be circular) on the Smith chart.

But on that same map of input impedance we also need to notice the contours of gain. Where NF contours look like a shallow depression, gain contours often look like quite a steep hill. Maximum gain does not occur at the same input impedance as minimum NF, so whenever we are adjusting the input circuit to find Γ_{OPT} at the bottom of the NF depression, we are also moving across the *slopes* of the gain peak which are much steeper. Small changes in the input circuit can often lead to quite large changes in gain; and this makes the NF measurement quite vulnerable to gain-related errors.

DUT-related gain errors

In *DUBUS* issues 4/1988 and 4/1990, Rainer Bertelsmeier DJ9BV [6] explained very clearly how NF measurements can be affected by the difference in the output impedance of a noise source between its ON and OFF states. The output impedance of the noise source becomes the input source impedance for the DUT, so any changes will affect both the internal noise and the gain of the DUT. This in turn will falsify the computation of NF. Many more details are given in the longer version of this paper on the Conference DVD.

Input tuning can cause the *indicated* NF to vary both above and below the true value. It is vital to understand all the consequences of this:

1. **The minimum indicated NF is not the true minimum.**

   The NFA is no longer helping you to find the optimum noise performance – most likely, it is leading you **away** from the true minimum.

2. **Professional noise figure analysers can deliberately tell lies**.

   Although negative values of NF are physically impossible, in some cases a negative error should genuinely drive the *displayed* value of NF below 0 dB. It is very unhelpful that some NFAs quietly suppress any negative values in the

6. SK 2008. Rainer's presence at these EME Conferences is sadly missed.

NF readout! The user needs to see those impossible results, as a warning that the measurement has gone seriously wrong.

Most people now understand that accurate measurements on LNAs require a noise source like the HP346A which has extra built-in attenuation to minimize the change in impedance between the ON and OFF states. Less well understood is that this isn't a complete cure: the errors are reduced but they didn't go away completely. Barely understood at all is the role of the DUT design in all this.

Device NFs in the 1980s were about 0.3–0.4 dB, so changing to a more suitable noise source made this source of error small enough to be ignored. But not forever: when modern ultra-low-noise devices are used in the same 1980s-style 'retro' designs, even those smaller errors can become important again. By tuning the preamp input, the displayed value of NF can easily be 'sweetened' by 0.02 dB, while at the same time making the true NF *worse*. This potentially fatal error been known by EMEers since the 1980s, or even before… but 30 years later it seems to have been forgotten.

When the same ultra-low-noise device is used in a modern LNA design with reasonably good input matching, the phase angle effect once again becomes small and well under control. The NF measurement is still subject to all the *other* errors and uncertainties described in the rest of this paper, but at least the indicated minimum value will be very close to the true minimum.

ENR Calibration

ENR calibration is probably the largest unavoidable source of uncertainty in NF measurements because it isn't something you can do for yourself. Even large corporations struggle to deliver the accuracy we need today. There are many more details about ENR calibrations in the longer version of this paper on the Conference DVD.

We are all aware that modern LNAs should be measured using a noise source with an ENR of about 5 dB. Figures 2 and 3 are calculated using this value, which is about the right magnitude to minimize calculational uncertainties[7] with most low-noise DUTs.

Unfortunately the trend towards lower-noise devices means that the majority of the commercial noise sources coming onto the surplus market are 15 dB ENR types like the HP346B. It is possible to make a 5 dB source by attaching a high quality 10 dB attenuator to the output of an existing 15 dB source – but you cannot casually subtract 10 dB from the published ENR and expect accurate results! There are many corrections to be applied, requiring detailed measurements with a Vector Network Analyser. As well as a very accurate measurement of the attenuation value in a 50 Ω reference environment, you also need the vector reflection coefficients Γ_{ON} and Γ_{OFF} for the noise source, and Γ_{IN} and Γ_{OUT} for the attenuator. Then you must grind through the calculations… and at the end of it all, the uncertainty in ENR will be worse than that of the bare noise source.

A purpose-made 5 dB ENR noise source like the HP346A will have a very similar attenuator built in – but crucially, the ENR calibration of the HP346A will already *include* that attenuator. The specified uncertainty is still ± 0.2 dB, the same as for the HP346B.

And one more question: even if you do own a factory calibrated noise source, how many years ago was that calibration done?

---

7. The measured numbers in an NF calculation are already 'fuzzed' by the statistical properties of noise, so it is very important to avoid any further uncertainties that arise when subtracting two numbers that are almost equal, or when calculating the ratio between two extremely different numbers. The best situation is usually when all the numbers involved are similar in magnitude, but still distinctly different from each other.

Total Uncertainty Budget

Fire-up the uncertainty calculator on the Agilent website. We recommend the online Java version[8] because it gives useful guidance about input values that are instrument-specific. Now plug in the numbers for your system… and find out how little you can **really** trust your results.

Here are two examples:

1. **DUT NF 0.3 dB, DUT gain 30 dB**

   A modern LNA design with input return loss **10** dB, output return loss 10 dB

   Measurement uncertainty:

   ± 0.29 dB (HP346A, 8970B)

   ± 0.23 dB (N4000A, N8973A)

2. **DUT NF 0.3 dB, DUT gain 30 dB** – same NF and gain as case 1, but…

   A 1980s-style LNA design with input return loss **1** dB, output return loss 10 dB

   Measurement uncertainty:

   ± 0.60 dB (HP864A, 8970B)

   ± 0.51 dB (N4000A, N8973A)

We can see three major conclusions:

1. The modern-style LNA has measurement uncertainties comparable with the true NF, and most of this (about 0.2 dB) comes from the ENR calibration. With care, other errors amount to about ± 0.1 dB.
2. The old-style tuneable LNA has 2x the measurement uncertainties, and most of that comes from the highly mismatched input circuit. The uncertainty amounts to 2x the true NF, and it isn't even possible to find the optimum input settings.[9]
3. The modern N4000A/N8973A lineup offers a little lower uncertainty than the older HP346A/8970 but it does not cure any of the problems that we have been discussing.

Conclusions

With modern LNAs using devices that have extremely low noise figures, conventional NF measurements are coming close to the end of the road. There are too many sources of uncertainty that simply cannot be avoided.

However, **that should not prevent us from minimizing the <u>avoidable</u> uncertainties**, not only by careful measurement technique but also by avoiding LNA designs that increase those measurement uncertainties even more.

References

The longer version of this paper on the Conference DVD contains more information and discussion, and most of the key references are provided as PDF files.

---

8. http://sa.tm.agilent.com/noisefigure/NFUcalc.html
9. Do not misquote us on this! We did **NOT** say or wish to imply that 1980s-style LNAs "don't work". But we **do** say that measurements on those LNAs using a conventional noise source and NFA cannot be *meaningful*.