Theoretical Background on Noise

Pablo Collado Soto

Observatorio Astronómico de Yebes

1 What's noise?

We can generally state that noise in the realm of electronics refers to *unwanted* electric signals that hinder a system's operation. The key word here is *unwanted*: in an ideal setting we would experience no noise whatsoever.

The sources of this noise are both artificial and natural. By artificial we mean that any equipment we use will introduce some kind of noise just by acting on the signal. This is something we cannot avoid but that we can indeed try to minimize. The object of our work is estimating what's the noise introduced by the amplifiers built at the *Yebes Astronomical Laboratory*. Whilst this process is fairly straightforward for industry level equipment, this is not the case for amplifiers used in the field of radio astronomy and we'll need to look into how we can indeed measure what we need.

When we say noise is also naturally generated we mean that it's a consequence of the quantized nature of both matter and light. A clear example of this noise is the *Brownian Motion* that's modeled as a *Wiener Process*, a specific *stochastic process* which is nothing more than a function assigning each realization of a random experiment an entire function of time. Another example is the so called *black body radiation* which is related with electro-magnetic radiation such as light. In any case, we can see how these two examples are related with **temperature**. The hotter a fluid is, the larger the amplitude of the movement of the particles within it and, the larger the temperature, the more radiation that's produced.

As this natural noise sources are the consequence of a lot of interactions we can assume thanks to the central limit theorem that their effects will follow a normal or gaussian distribution. What's more, if we model our noise as an *Additive White Gaussian Noise* we can indeed notice that our noise distribution, Z, is given by:

$$Z \sim N(0, \sigma)$$

The above applies to each realization of the noise follows a normal distribution with zero mean and σ variance. Each measurement will then produce an output at a given time $Y_i = X_i + Z_i$ where X_i is the input to the system we are studying at that same instant and Z_i is a realization of the noise. What's more, if we take a look at the definition of the expectation of a random variable we can see that it's defined as:

$$E[X] = \int_S x f(x) dx \to E[X+Y] = E[X] + E[Y]$$

In other words, the output's mean is the superposition of the input's mean and the noise's mean. What's more, if we try to take the average of the noise we'll observe the following:

$$\sum_{i=1}^{N} Z_i \sim N(N\mu, \sqrt{N\sigma^2}) \rightarrow \frac{1}{N}\sum_{i=1}^{N} Z_i \sim N(\mu, \frac{\sigma}{\sqrt{N}})$$

The above suggests that the variance of the average of the noise tends to zero as N increases. This in turn implies that if we take more and more averages we can limit the effect of the noise as the width of the associated normal curve will continue to shrink as N increases and, as the average of the noise is zero, its contribution to the $Y_i = X_i + Z_i$ expression diminishes. In the limit where $N \rightarrow \text{inf}$, the variance would become 0 and we would end up at a distribution resembling the *Dirac delta* function $\delta(x)$. The "bad" part is that in order to increase the number of noise realizations we take into account we need to increase the integration time. We can also increase the number of data points we take into account by increasing the bandwidth we take into account. The more frequencies we add up, the more noise realizations we are taking into account and the lower the average will become. Then, both integration time and bandwidth work towards the same goal. If we increase them we take more values into account and we are ultimately decreasing the noise's variance, this reducing it's effect on our signal. At the end of the document we'll also find out that this variance is related with the *noise temperature* in such a way that the larger the *noise temperature* is, the larger the variance. Thus, we'll want to minimize that noise temperature at all costs.

This can all be summarized by stating that any signal interfering in the transmission between an emitter and a receptor is noise. This noise will in turn determine the sensitivity of our system, that is, the minimum input signal that we can actually appreciate. If we think about "literal" noise we should picture us in a room full of people talking. If the noise around us is too loud we won't be able to whisper to someone across the room but we can try to shout instead and that's likely to be heard at the other end. In this scenario, the existing noise level lowers the sensitivity of the system and so the input signals we must present to it need to be larger in amplitude or they'll go unheard... Given the nature of this noise we can "fight back" by taking the average of what we receive to minimize its impact and get a clearer look of whatever static value we are trying to observe. The catch is that the more averaging we do, the longer the measurements take and so we could suffer changes in the system's condition that could compromise the accuracy of our measures. In the end, noise is the "bad guy" and we need to try and prevent it by whatever methods we can devise.

2 Characterizing the noise

We have looked into what noise is but, how can we express it's effect quantitatively? We will begin by exploring the noise factor F, but we first need to get some definitions out of the way:

2.1 The signal to noise ration (S/N)

As its name implies, the signal to noise ratio is defined as the quotient between the signal's power and the noise's power at a given point in the system. This quantity being a ratio it's dimensionless and often express in decibels. That is:

$$S/N_{db} = 10 \cdot log(S/N_{lin}),\ where\ S/N_{lin} = \frac{P_{signal}}{P_{noise}}$$

From the very definition of this factor we can conclude that:

1. $S/N_{lin} \to$ inf: The signal power is very large when compared to the noise power or we have no noise at all. This is a best case scenario we'll seldom experience.

2. $S/N_{lin} \to 0$: The noise power is very large when compared to the signal power. The noise in this scenario is "deafening", we can hardly detect the input signal. This is the worst case scenario.

If there is something we need to take away is that the larger the S/N ratio is, the better the system's operation.

Now, knowing what the signal to noise ration is we can define the noise factor of a system or device as the quotient of the signal to noise ratios at the input and output, respectively:

$$F = \frac{S/N_i}{S/N_o},\ where\ T_s = 290\ K$$

We define the noise factor for a source temperature of $290\ K \approx 16,8°$. This implies that this parameter is useful for characterizing systems at room temperature, which is true for most systems. In any case, the noise factor is telling us what the effect of a system is on the signal to noise ration. That is, if $F = 1$ then the system won't be adding any noise (an ideal scenario) and if $F \to$ inf then the system is decreasing the signal to noise ratio by a great deal: it's adding a lot of noise. The best noise factor we can aim for is indeed 1. Then, the lower the noise factor of a system, the better that system behaves in terms of noise.

Note that, just like with the signal to noise ratio, the noise factor is dimensionless as it's a ratio between dimensionless quantities. What's more, we often refer to the noise figure, which is nothing more than the noise factor expressed in logarithmic units:

$$NF\ [db] = 10 \cdot log(F\ [lin])$$

We can play around a little with the expression of the noise factor to take the system's gain into account as well. As the gain (G) acts both on the input signal and the input noise we can clearly see that:

$$F = \frac{\frac{P_{S_i}}{P_{N_i}}}{\frac{P_{S_o}}{PN_o}} = \frac{\frac{P_{S_i}}{P_{N_i}}}{\frac{G \cdot P_{S_i}}{P_{N_a} + G \cdot P_{N_i}}} = \frac{P_{N_a} + GP_{N_i}}{GP_{N_i}}$$

The key factor here is the noise power added by the system or device we are trying to characterize, P_{N_a}. We would also like to mention that the device we're analyzing will often be referred to as the *Device Under Test* (DUT).

If we consider the input noise to the system to be *thermal* or *Johnson* noise (the one present in every resistor no matter its bias point) we can see that it's power is given by:

$$P_{N_i} = k \cdot T \cdot B$$

Where $k = 1,38 \cdot 10^{-23}\ \frac{J}{K}$ is *Boltzmann's constant*, T is the resistor's temperature in $Kelvin$ and B is the noise bandwidth we are considering in $Hertz$. If we carry out the dimensional equation we can indeed see how:

$$[\frac{J}{K} \cdot K \cdot Hz] = [J \cdot Hz] = [J \cdot \frac{1}{s}] = [\frac{J}{s}] = [W]$$

Now, inputting noise to the system means we'll drive its input with a resistor. This subtlety can seem a little awkward at first but it makes complete sense: all in all, our noise source is the resistor itself!

When working with the noise factor we showed how the output power noise is given by $P_{N_o} = P_{N_a} + GP_{N_i}$. If we now assume the input power is indeed *thermal noise* we conclude that:

$$P_{N_o} = P_{N_a} + GkBT_s \rightarrow P_{N_o}(T_s) = P_{N_a} + GkBT_s$$

That is, we can regard this output power as a function of the source's temperature T_s. Note that for $T_s = T_{real}$ we get the power of the output noise for the real temperature the noise source is at. Anyway, considering the output noise power as a function of the source temperature we find out how $P_{N_o}(T_s)$ describes a straight line with slope $m = GkB$ and with a value of $b = P_{N_a}$ at the origin. When trying to characterize a system we are trying to find the noise it's adding, that is, P_{N_a}. Now, knowing the shape of the output noise power is a straight line we can easily compute it with just knowing two points within it. That is, if we take two measurements we can find the slope $m = \frac{\Delta P_{N_o}(T_s)}{\Delta T_s}$ and we could then solve for a point to find out the noise power added by the DUT with $P_{N_a} = b = P_{N_o}(T_s = \alpha) - kGB\alpha$, where α is one of the values of the source temperature we used to derive the slope.

All in all, knowing the shape of the output power noise we can easily extract the noise added by the system by taking a couple of points from the graph and solving for it.

2.2 Enter the noise temperature (T_e)

The noise factor and noise figure parameters are suitable for characterizing the noise added by a system in industrial grade systems where the noise levels are fairly high. We should nonetheless remember the lab's amplifiers are to be used for radioastronomy, a field where the signals we are to detect are incredibly soft. This means that the amplifiers we are to analyze introduce extremely little amounts of noise and thus the noise factor becomes an unsuitable unit. Whilst we could technically use this unit as the definition is totally applicable to any system, we'll often find ourselves working with factors that are above one by very small quantities. This becomes cumbersome rather fast.

In the same way decibels came into the picture to facilitate engineer's life and make handling units easier, we can regard the noise introduced by a system in terms of what we call the *noise temperature*, T_e. This quantity is defined as the amount by which we should increment the temperature of the noise source (a.k.a input load) so that the observed output power of an ideal system that adds no noise (that is, $P_{N_a} = 0$) would be the same as the one observed on a the non-ideal version of the system where $P_{N_a} \neq 0$ with the input load at the original temperature.

The above implies we'll now regard noise in terms of temperatures instead of powers. We shan't forget that the thermal noise equation ($P_N = kBT$) gives us a direct correspondence between temperature and noise power. We have deliberately omitted the discussion on the particulars of amplifiers for radioastronomy up until now so as to keep the discussion as general as possible. Just as a teaser, we can state that this way of regarding power is way more suitable for systems with very low added power noises as well as those whose input loads are not at room temperature ($T_s \not\approx T_o = 290\ K$). What's more, temperature is a way more "physical" unit than the noise factor, so it helps us get a more intuitive feel of the system.

The above can seem pretty intimidating, but if we look at the math behind this all we'll see that if we are regarding noise temperatures, our output noise power becomes:

$$P_{N_o}(T_s,\ T_e) = G \cdot kB(T_s + T_e) = G \cdot kBT_s + G \cdot kBT_e$$

Where T_s is the original temperature of the source and T_e is the noise temperature. This noise power output has to be the same as the one we saw before, by definition. That is, $P_{N_o}(T_s) = P_{N_o}(T_s,\ T_e)$. This implies:

$$P_{N_o}(T_s) = P_{N_o}(T_s,\ T_e) \rightarrow P_{N_a} = G \cdot kBT_e$$

If we take that into account and plug it into the expression of the noise factor we find:

$$F = \frac{P_{N_a} + G \cdot kBT_s}{G \cdot kBT_s} = \frac{G \cdot KBT_e + G \cdot kBT_s}{G \cdot kBT_s} = \frac{T_e}{T_s} + 1 \rightarrow T_e = (F-1)T_s$$

We can clearly see how the noise temperature and the noise factor are related. If we recall the noise factor is defined for source temperatures $T_s = T_o = 290\ K$ we can particularize the above to show:

$$F = \frac{T_e}{T_o} + 1 \rightarrow T_e = (F-1)T_o$$

We would also like to point out how $P_{N_o}(T_s,\ T_e)$ is now a line that crossed the cartesian plane in the origin. What we need to find however is that value of T_e, which is the one characterizing our system.

2.3 Cascading systems

When defining the noise factor we stated that it could be used to characterize single devices or entire systems. If we do the latter, we need to take into account the gain of both systems and how it acts on the different noise powers. We'll begin, as always, by defining the input noise power $P_{N_i} = kBT_s$, where T_s is the temperature of the input load. Now, after traversing the first amplifier we'll observe that intermediate output power noise is given by $P_{N_{o_1}} = P_{N_{a_1}} + G_1 \cdot kBT_s$. Now, all this noise will be amplified by the second stage's gain and this second stage will also contribute with it's own noise so that the output noise power is (note we can add them up because the noise powers are uncorrelated):

$$P_{N_o} = G_2 \cdot (P_{N_{a_1}} + G_1 \cdot kBT_s) + P_{N_{a_2}}$$

If we go back to the equations presented in the previous subsection we'll find a convenient relation between the added noise power of a system and it's noise factor. Taking into account that $P_{N_a} = G \cdot kBT_e \rightarrow T_e = \frac{P_{N_a}}{G \cdot kB}$ we find that:

$$T_e = (F-1)T_s \rightarrow \frac{P_{N_a}}{G \cdot kB} = (F-1)T_s \rightarrow P_{N_a} = (F-1)T_sGkB$$

If we apply that to the expression for the whole output noise power we reach:

$$P_{N_o} = G_2 \cdot ((F_1 - 1)G_1 kBT_s + G_1 kBT_s) + (F_2 - 1)G_2 kBT_s =$$
$$= F_1 G_1 G_2 kBT_s - G_1 G_2 kBT_s + G_1 G_2 kBT_s + (F_2 - 1)G_2 kBT_s =$$
$$= kT_s BG_2 G_1 \cdot (F_1 + \frac{F_2 - 1}{G_1})$$

Now, if we apply the definition of noise factor to the above we'll get the noise factor for the whole system:

$$F_{sys} = \frac{P_{N_o}}{G_{sys} \cdot P_{N_i}} = \frac{kT_s BG_2 G_1 \cdot (F_1 + \frac{F_2 - 1}{G_1})}{G_1 G_2 kBT_s} = F_1 + \frac{F_2 - 1}{G_1}$$

The key takeaway from the above is that the larger the gain of the first stage (G_1), the smaller the contribution of the second stage to the noise factor ($\frac{F_2-1}{G_1}$) will become. In other words, we are interested in having a large gain in the first stage.

Given we'll characterize single amplifiers one might think this discussion on cascading systems is not interesting. The things is we **will** consider the measuring equipment (i.e. the noise figure analyzer) to be a second stage as it'll univocally add some noise to the measurements. This means that in order to minimize it's impact we need to provide a large gain in the DUT.

We would like to end this section by noting that the noise factor and noise temperature of a system vary with frequency. This implies that in order to fully characterize the system we'll have to sweep a range of frequencies within the amplifier's working bandwidth.

3 Measuring procedure: the Y factor

If we recall the previous expressions we'll quickly notice that the function describing the output noise power is a straight line. This means that by taking two measurements we can interpolate the line's slope and find out any point within it so that we can indeed find out the noise parameters of our system. If we go back to the system we employed when defining the *noise temperature* we'll recall that the output noise power was given by: $P_{N_o} = kGB(T_e + T_s)$, where T_s was the input load's temperature and T_e was the noise temperature itself. Now, if we want to find out the output noise temperature we need only normalize the previous expression by kB so that we end up at $T_{N_o} = G \cdot (T_e + T_s)$. Now, to find out the value of T_e we could find a point within the line described by the previous equation and solve for T_e:

$$T_e = \frac{T_{N_o}(T_s = \alpha) - G \cdot \alpha}{G}$$

Now, if we take two measurements we can indeed find out the value of that gain G:
$G = \frac{T_{N_o}(T_s=T_h) - T_{N_o}(T_s=T_c)}{T_h - T_c} \rightarrow \frac{G(T_e+T_h) - G(T_e+T_c)}{T_h - T_c} = G$

We have shown we can indeed get the slope by just carrying out the computation shown in the first quotient. If we plug the second equation into the first and we take the function's value for $T_s = T_h$ we find that:

$$
\begin{aligned}
T_e &= \frac{T_{N_o}(T_h) - \frac{T_{N_o}(T_h) - T_{N_o}(T_c)}{T_h - T_c} \cdot T_h}{\frac{T_{N_o}(T_h) - T_{N_o}(T_c)}{T_h - T_c}} \\
&= \frac{T_{N_o}(T_h)T_h - T_{N_o}(T_h)T_c - T_{N_o}(T_h)T_h + T_{N_o}(T_c)T_h}{T_{N_o}(T_h) - T_{N_o}(T_c)} \\
&= \frac{T_{N_o}(T_c)T_h - T_{N_o}(T_h)T_c}{T_{N_o}(T_h) - T_{N_o}(T_c)}
\end{aligned}
$$

Now, we can express the above in terms of the so called Y factor. This factor is just the ratio between the output noise power measured with the noise source at temperatures T_h and T_c, that is, a "hot" and a "cold" temperature. Then, applying that $Y = \frac{T_{N_o}(T_h)}{T_{N_o}(T_c)}$ and dividing both numerator and denominator by $T_{N_o}(T_c)$ we find that:

$$
T_e = \frac{T_{N_o}(T_c)T_h - T_{N_o}(T_h)T_c}{T_{N_o}(T_h) - T_{N_o}(T_c)} = \frac{T_h - YT_c}{Y - 1}
$$

As always, we can extract the noise factor from the noise temperature employing the relation stating that:

$$
F = \frac{T_e}{T_o} + 1, \ where \ T_o = 290 \ K
$$

Then, the conclusion we reach is that as long as we have two measurements we can work the numbers to find both the gain and the noise figure of an amplifier. Getting this data requires us to change the noise of the input load. That can be accomplished in several ways, depending on the technology we employ:

1. **Avalanche diode**: We can reverse-bias diodes into their avalanche region with steady currents to generate noise. We'll observe that when the diode is not biased, the noise diminishes and so we find the two data points we need. Working with diodes has the advantage of being able to almost instantaneously switch the "load temperature", that is, we can change the input noise with a high frequency. That is why, when using these noise sources, we'll usually sweep the entire frequency range we are interested in once and we'll switch the input noise for each frequency. That is, we'll fix a frequency ω_1 and take the measurements with $T_s = T_c$ and $T_s = T_h$ to then move to a new frequency ω_2. This ensures that the system's conditions remain constant in the instants we measure both input temperatures. Note this type of noise source usually comes with an attenuator at the output so that the impact of the reflections caused by the constant impedance changes is minimized. This attenuator can be "taken out" of the noise source itself and cooled down for cryogenic applications. When this is indeed done, we are trying to get the source temperatures we use as a reference (that is, T_h and T_c) close to the ones we expect to find in our amplifier. By getting them close to T_e we are minimizing the measurement's error. This can be shown with some tricky math that we don't know how to handle, but the intuitive idea is that if we have a reference close to the measured value we're minimizing measuring inaccuracies. This follows from how we are taking points in the output noise power that are closer to the Y axis. Then, the line we build from them will be closer to the "real thing". One can also think about a multimeter used for measuring voltages. If we calibrate it for large voltages when we intend to measure small ones we can be sure our results won't be as accurate as they could be...

2. **Hot & cold load**: As the name implies, we'll use a heating circuit to physically change the temperature of an input load. Owing to the thermal noise equation we can be sure these changes in temperature correspond to different noise powers, which is what we need at the input of the DUT. The problem is that temperature can't vary as fast as we could switch the diode before: we need a non-negligible amount of time to get the input to the desired temperature. This makes us take another approach at measuring: we'll get the load to a given temperature (i.e. $T_s = T_c$) and then sweep all the frequencies we want to take into account for the measurement. We'll then vary the load's temperature ($T_s = T_h$) and sweep the frequency range once more. Even though one might think this procedure is equivalent to the one described before, it's not. Given the non-negligible time it takes us to carry out the measurements we can't guarantee the state of the system is the same for both measurements with $T_s = T_c$ and $T_s = T_h$ for a given frequency ω...

3.1 The need for calibration

Before making any measurements, we need to have a baseline against which to compare the results. This is what we call the calibration step, in which we plug the noise source directly to the noise figure analyzer. Whilst external conditions do vary with time, it's enough to calibrate the equipment once or twice per day.

After the calibration step, we'll plug the DUT in between the noise source and the noise figure analyzer to carry out the desired tests.

3.2 The effects of the analyzer

As we said before, the measuring equipment introduces some noise which we want to avoid at all costs. The way of doing so is increasing the gain of the DUT as explained when analyzing the cascade noise factor before in this document. As the gain of the DUT is fixed by design, we can employ pre-amplifiers between the DUT and the measuring equipment so that the noise figure analyzer's effect is less noticeable. This pre-amplifier is designed so that it has a very low noise temperature and, as it's known, we can correct its contribution when working with the gathered results. We need to be careful however, as we risk saturating the measuring equipment if the combined gain of the DUT and the pre-amplifier (they're multiplied in linear units and added in logarithmic ones) is too high. This saturation would destroy any meaningful data, so we must avoid it at all costs.

When using these pre-amplifiers we can regard them as a part of the receiver as long as we calibrate the receiver with the pre-amplifier attached. That way we will still obtain accurate measurements without increasing the data treatment overhead.

4 The particulars of radioastronomy

With equipment such as noise figure analyzers already existing, one might think why we would need to write custom software to carry out that very procedure. Whilst existing machines are suitable for industrial applications, radioastronomy deals with extremely weak signals where the precision needed when characterizing equipment must be very accurate. At the end of the day we are studying a noise source (space) whose noise temperature we need to characterize extremely well. The more noise that our equipment introduces, the worst...

What's more, in the laboratory we employ auxiliary equipment to make more precise measurements such as *cold attenuators* which aren't taken into account by the existing noise figure analyzers. This means we need to manually correct results taken it into account.

Throughout the document we have seen how there's a close link between noise power and temperature. The larger the temperature, the larger the noise levels we need to deal with. That's why these amplifiers are *cryogenic*, that is, they operate at temperatures close to 0 K. This lowers the noise level of the amplifier, which is our goal after all.

In the end, we are trying to increase the system's sensibility so that we can study the received power of a band as well as its spectral distribution with the utmost precision. This sensibility is related to the variance of the noise we observe at the receiver and that variance is related to the system's noise temperature, the bandwidth we are studying and the integration time on the data:

$$\Delta T = \frac{T_{sys}}{\sqrt{B\tau}}\ Note:\ Var = \frac{\sigma}{\sqrt{N}};\ \sigma\ \alpha\ T_{sys};\ N\ \alpha\ B\tau\ (\alpha \equiv Directly\ proportional)$$

Then, the lower the system's noise temperature, the better our sensibility will be. The same is true if we increase either the integration time (τ) or the bandwidth (B) or both. Note that $[\tau] = s \rightarrow [\tau \cdot B] = [s \cdot Hz] = [s \cdot \frac{1}{s}] = \emptyset$, that is, this is a dimensionless quantity.

All in all, we'll try to decrease the system's noise temperature T_{sys} so that we enhance the overall sensitivity. This will let us study ever weaker signals. As amplifiers add up to this noise temperature, lowering it will in turn lower T_{sys}.

References

[1] J.D. Gallego; R. Amils, I. López C. Diez González, and I. Malo. Using keysight pna-x and nfa noise receivers for noise temperature measurements with cryogenic variable temperature loads.

[2] M. Carmen Diez. Concepto de ruido en señales eléctricas y su medida. Curso de Formación IGN.

[3] Juan Daniel Gallego Puyol. Radioastronomía y tecnología.

[4] Agilent Technologies. Fundamentals of rf and microwave noise figure measurements. Application Note 57-1.