(Draft) ANITA technical note: Proper signal-to-noise estimates for ANITA receivers

Peter Gorham
UH Manoa

December 16, 2005

1 Introduction

Understanding ANITA's trigger sensitivity is crucially dependent on accurate models for the thermal noise that dominates the background. Equally important is a model for the statistics of the received amplitude fluctuations of a signal in the presence of the thermal noise background.

Currently, the UH Monte Carlo is using an oversimplified and thus inaccurate approach to modeling the latter case. Once the signal level (voltage received from ambient signal peak field strength) at the antenna terminals is determined, a value for the instantaneous thermal Gaussian noise is separately drawn from a Gaussian random number generator and added to the signal value to give the resultant amplitude. This procedure is accurate only in the limit of large signals, and is particularly inaccurate in the application for ANITA, where it is used to determine whether a signal, at the level of only a few times the thermal noise standard deviation, will trigger. I suspect that all of the other ANITA Monte Carlos are doing something similar.

The correct approach derives from first recognizing that the statistics of the amplitude of signal in the presence of noise are properly described by Rician, rather than Gaussian, distributions. In the absence of a signal, the electric field due to polarized[1] thermal noise can be represented as a sum of random phasors

[1] in our case by the antenna.

(eg., complex numbers) whose projections on the real and imaginary axes are independent Gaussian random variables, thus giving circular Gaussian statistics. Summing these phasors and taking the real part to get a realized voltage results in a Gaussian two-sided voltage distribution, or equivalently a Rayleigh distribution of amplitudes, and a uniformly random phase distribution.

Once a signal is present, the phasor sum involves now a sum of both random phasors and a constant phasor representing the signal (ignoring intrinsic fluctuations in the signal). This leads to a different statistical distribution of amplitudes, the Rician distribution, which contains the Rayleigh distribution as a special case. The purpose of this note is to describe the correct statistical properties of a signal in the presence of thermal noise for ANITA applications. The discourse will largely follow several sections in J. Goodman's *Statistical Optics* [1], which I find to be a particularly clear presentation of this, despite its wider optics (rather than radio) focus.

2 Noise amplitude distribution

For ideal polarized[2] thermal noise treated as a random phasor sum, circular Gaussian statistics apply to the resultant phasor (amplitude and phase) that is realized at any given moment at the receiving terminals of an antenna. Since the most general sort of amplitude trigger depends on how the received voltage exceeds some voltage threshold in absolute value, the normalized probability distribution relevant to noise amplitude A is the Rayleigh distribution:

$$p_A = \frac{a}{\sigma^2}\, e^{-a^2/(2\sigma^2)} \tag{1}$$

where $a = \sqrt{r^2 + i^2}$ is the measured amplitude at some instant, depending on the real and imaginary components r, i of the underlying field phasor, and σ^2 is the variance of either r or i, thus $\sigma^2 = \bar{r^2} = \bar{i^2}$. We can also conclude that $\sigma = V_{rms}$, the root-mean-square voltage at the antenna terminals. Figure 1 plots the Rayleigh distribution as a function of σ.

While the two-sided Gaussian voltage distribution is zero-mean, the Rayleigh amplitude distribution has zero probability at zero amplitude. The mean of the Rayleigh distribution is

$$\bar{a} = \sqrt{\pi/2}\,\sigma$$

[2]It is polarized by the antenna acting as a polarizer in our case

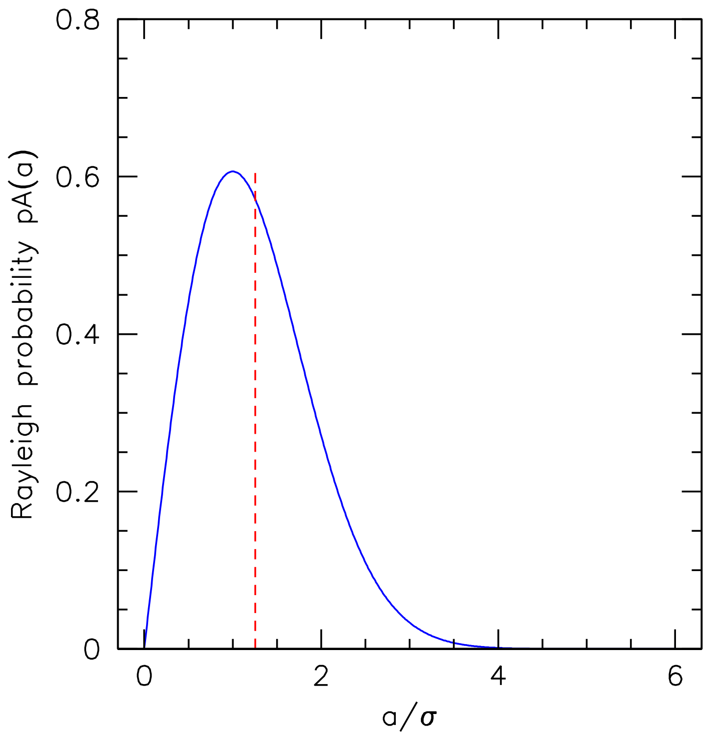

Figure 1: Rayleigh distribution as a function of the standard deviation of the components of the random phasor, which is equivalent to the σ of a realized voltage at an antenna in a thermal noise bath. The red-dashed line indicates the mean of the distribution.

plotted as a red dashed line in the figure, and the variance of the distribution is given by

$$\sigma_a^2 = \left[2 - \frac{\pi}{2}\right] \sigma^2 \ .$$

2.1 Cumulative Rayleigh distribution.

Since what we often care about in assessing thermal noise is the rate of accidentals, where the thermal noise exceeds a given amplitude in the absence of any signal, the cumulative distribution function (CDF) of the Rayleigh distribution is the relevant quantity, measuring the cumulative probability that the noise will exceed some amplitude threshold k:

$$P_A(a > k) = \int_\infty^k \frac{a}{\sigma^2} e^{-a^2/(2\sigma^2)} \, da \tag{2}$$
$$= e^{-k^2/(2\sigma^2)} .$$

This distribution is pure Gaussian. It differs from the CDF of a Gaussian since it involves the distribtion of the absolute value of the amplitude. (Some people refer to this form of the CDF as the "complementary CDF", since it involves the integral from infinity down to the threshold, rather than the integral below the threshold. But the two are always related by CDF = 1-(complementary CDF) so it does not matter much.

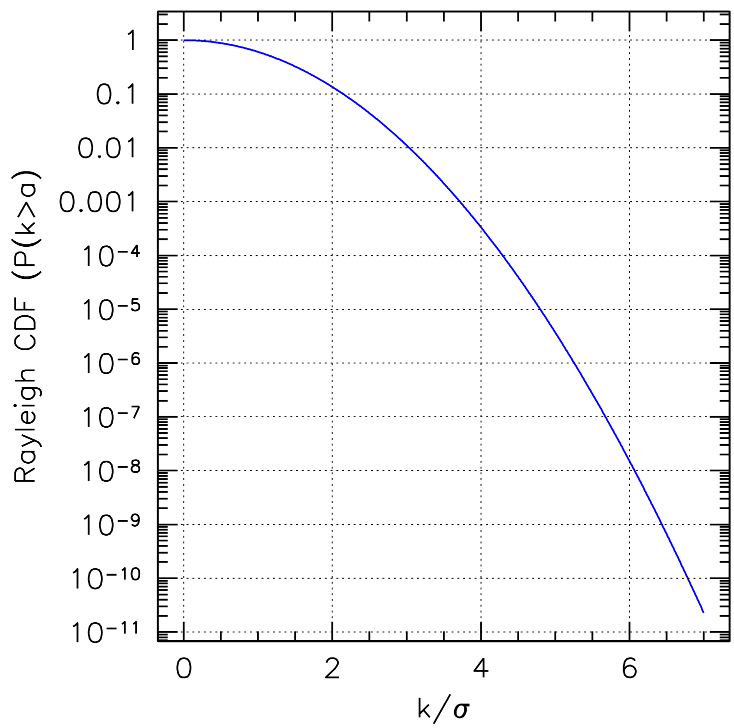

Figure 2: CDF of Rayleigh distribution.

2.2 The Intensity

Since the instantaneous intensity of the noise or a signal is defined as the squared amplitude of the field present, eg. $I = |A|^2$, we can immediately identify the statistics of the intensity from a transformation of variables in the Rayleigh distribution:

$$p_I = p_A(A = \sqrt{I}) \times \left|\frac{dA}{dI}\right| \tag{3}$$

which gives after some manipulation

$$p_I = \frac{1}{2\sigma^2}\, e^{-I/(2\sigma^2)} \tag{4}$$

and, recognizing that the mean intensity $\bar{I} = 2\sigma^2$,

$$p_I = \frac{1}{\bar{I}}\, e^{-I/\bar{I}}\ , \tag{5}$$

a negative exponential distribution as expected.

The CDF of the intensity as a function of some threshold intensity $\kappa_I = I_{thresh}/\bar{I}$ also follows immediately from the CDF for the Rayleigh distribution above

$$P(I > \kappa_I) = e^{-\kappa_I} \tag{6}$$

which gives the pure negative exponential distribution that is well known for radio power.

3 Signal+Noise amplitude distribution

The presence of a continuous signal of constant amplitude, or a pulse with a given peak amplitude causes a shift in the centroid of the sum of phasors to a location in the complex plane that is close to the peak of the fixed phasor that represents the imposed signal. The probability distribution of this resultant phasor, the sum of the random noise phasors and the signal phasor, in the Rice distribution, also called Rician (and sometimes Ricean). The joint amplitude and phase distribution is given as:

$$p_{A\Phi}(a,\phi) = \frac{a}{2\pi\sigma^2}\ \exp\left(-\frac{(a\cos\theta - s)^2 + (a\sin\theta)^2}{2\sigma^2}\right) \tag{7}$$

where s is the magnitude of the signal strength, and we require $a > 0$ and $-\pi < \phi \leq \pi$. The separate amplitude and phase distributions are then found by intregrating over the phase and amplitude separately, and are given as:

$$p_A(a) = \frac{a}{\sigma^2}\ \exp\left(-\frac{a^2+s^2}{2\sigma^2}\right) I_0\left(\frac{as}{\sigma^2}\right) \tag{8}$$

where I_0 is a modified Bessel function of the first kind, zero order.

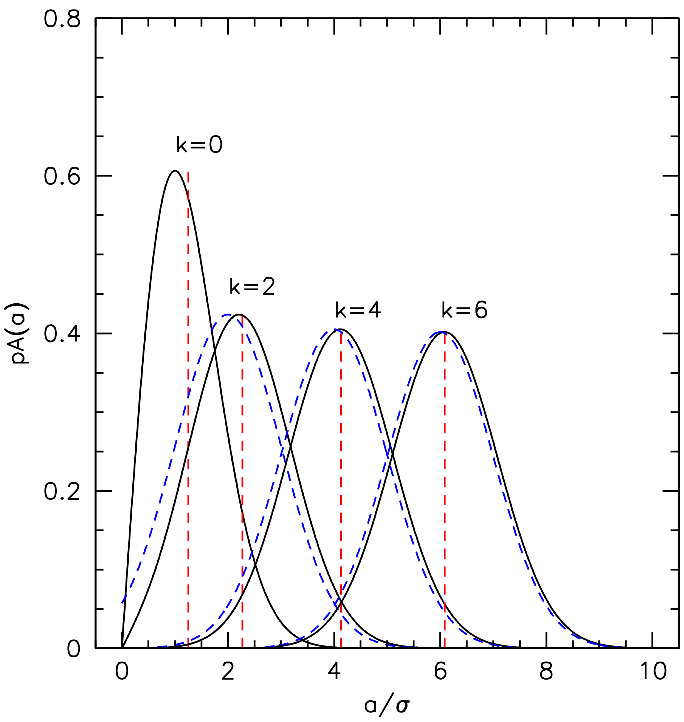

Figure 3: Rician distributions for various values of $k = s/\sigma$. The red dashed linen indicate the mean of each distribution, and the blue dotted lines are Gaussian distributions centered on each k with variance σ^2.

The corresponding mean and second moment of the amplitude distribution are

$$\bar{a} = \sqrt{\pi/2}\sigma\ e^{-k^2/4}[\ (1+k^2/2)I_0(k^2/4) + k^2/2\ I_1(k^2/4)\ ] \tag{9}$$

$$\bar{a^2} = \sigma^2[2+k^2]\ . \tag{10}$$

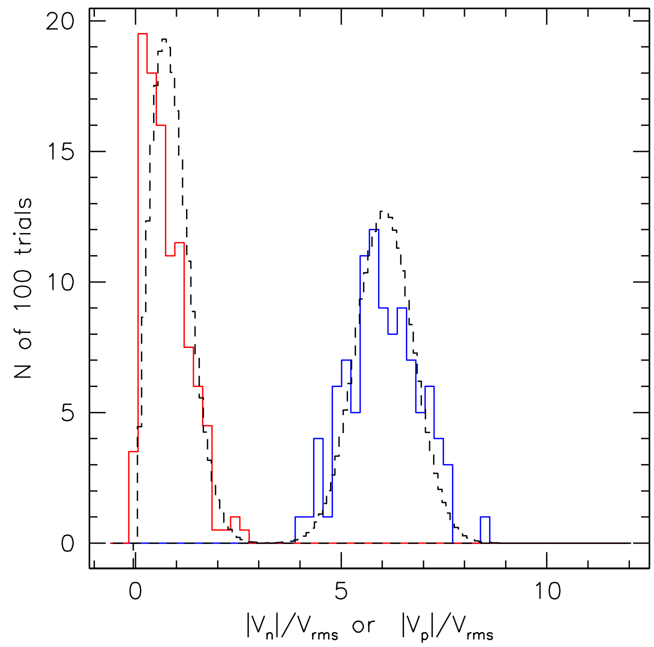

Figure 4: Data measured from 100 events using ANITA front-end receivers with a $\sim 6\sigma$ signal injected onto a background of ~ 450K thermal noise. The expected theoretical Rician distributions for the both the noise ($k = 0$ and signal $k = 6.3$ are shown.

where I_1 is a modified Bessel function of the first kind, order one.

Figure 3 plots Rician distributions vs. normalized amplitude a/σ for various values of the imposed signal's signal-to-noise ratio (SNR) $k = s/\sigma$. The red-dashed lines indicate the mean for each value of k. Also plotted are Gaussian distributions that would be obtained by using the naive approximation currently employed in the ANITA UH Monte Carlo, where the random Gaussian amplitudes are added to the fixed signal amplitude to simulate the superposition of the two.

It is evident that, although the Gaussian-coadd method is not a bad approximation at large signal levels, it is a poor approximation at the lower levels considered here, in several ways. First, it overestimates the small amplitudes, giving a near-flat distribution at zero amplitude, in strong disagreement with the true distribution which goes smoothly to zero at zero amplitude. Second, it underestimates the large fluctuations, mainly by virtue of understimating the mean value of the resultant distribution.

Figure 4 shows the result of a laboratory experiment with ANITA receivers and an impulse generator, bandpassed to look like an ANITA signal-of-interest, with a net amplitude of just over 6σ relative to the rms thermal noise voltages. We have sampled the time series of $N = 100$ events in two locations: at the peak of the signal (blue), and about 25 ns prior to it at a location which indicates pure thermal noise (red). In each the Rician theoretical curve $N\Delta a p_A(a)$ is also plotted. The agreement is good except for the lowest values of the noise measurements, where the scope digital resolution caused an excess of values equal to zero.

3.1 ANITA threshold

Since ANITA simulations use a threshold of typically 2.3σ in amplitude for detection within the sub-bands, the Rician distribution is very important to implement for signals that are close to this level.

To estimate the error made by using the Gaussian-coadd method compared to the proper Rician CDF for estimating the threshold-crossing probability for a given signal level with respect to noise, we use a quasi-CDF for the Gaussian coadd (computed by folding back the absolute value of the negative amplitudes for low signal levels in the Gaussian coadd case), and compare it to the Rician CDF for various signal levels in Figure 5. Table 1 gives the estimated probabilities for exceeding 2.3σ in several comparable cases.

Table 1: *Probabilities to exceed a* 2.3σ *threshold in the two cases of a Rician CDF or Gaussian coadd quasi-CDF.*

$k = s/\sigma$	Gaussian-coadd CDF	Rician CDF
0.5	0.6%	8.8%
1.0	4%	16.5%
1.5	15.6%	29%
2.0	34%	46%
2.5	56%	65%

It is evident that the Gaussian coadd method greatly understimates the threshold-crossing probability for a signal, compared to the Rician distribution, at low but

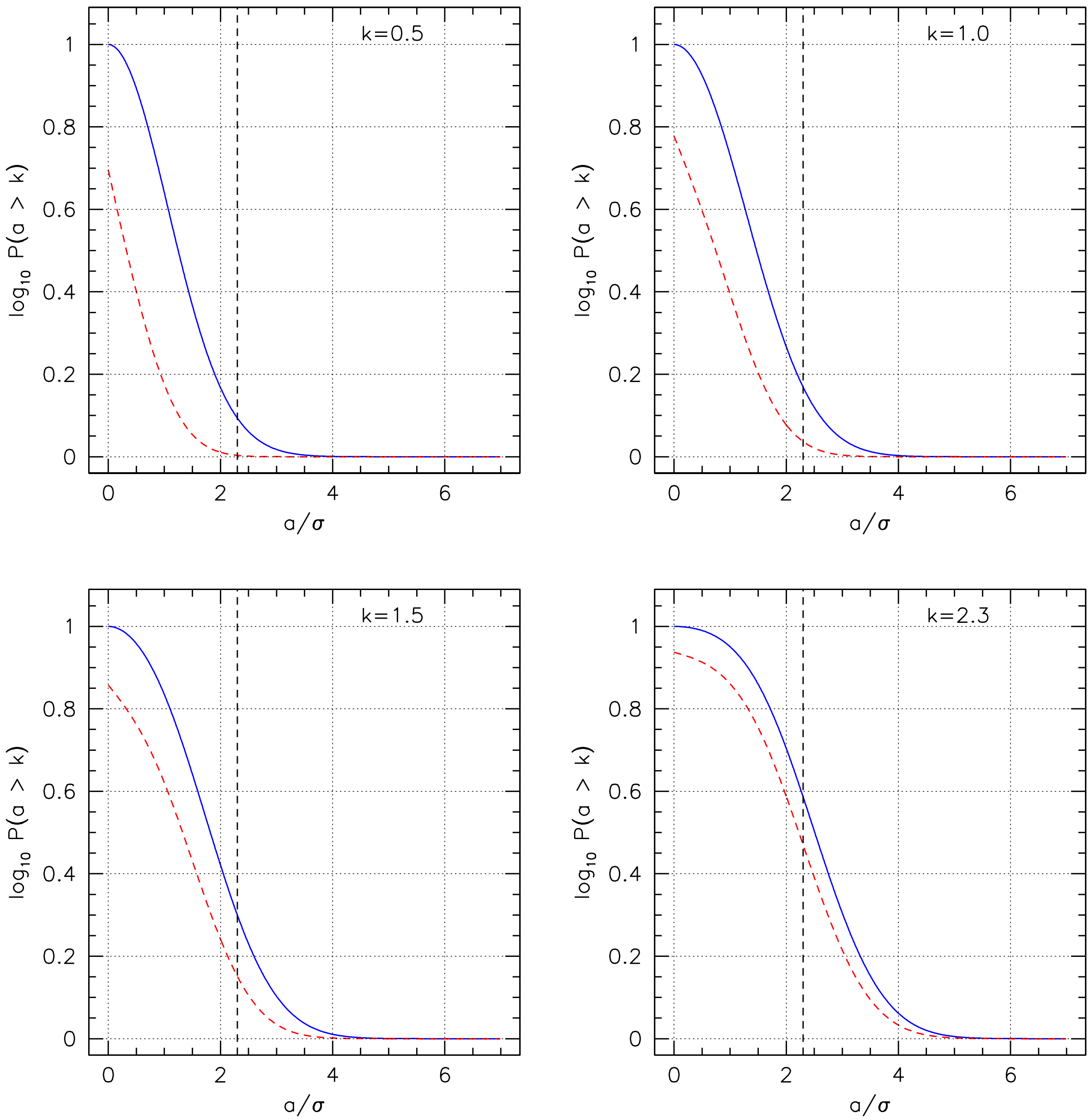

Figure 5: Rician vs. Gaussian-coadd CDFs for signals at the level indicated in the figure.

still interesting signal levels. For example, a signal with a roughly 6σ amplitude in one polarization produces levels of order 2σ per sub-band in an ANITA antenna, and one would expect us to trigger on such a signal with reasonable efficiency. For a 2σ signal, the Gaussian coadd method underestimates the trigger probability by 35%. Given that the subsequent combinatorics of the trigger require high powers of these probabilities, the Gaussian coadd method yields poor simulations of the actual trigger efficiency at signal levels close to our expected threshold.

References

[1] J. Goodman, *Statistical Optics*, 1985, (Wiley & Sons: New York), first edition.