INTERFERENCE DETECTION IN GAUSSIAN NOISE

RAJU BADDI
Raman Research Institute, C.V. Raman Avenue., Bangalore 560 080, India
Received 2010 November 17; accepted 2011 March 27; published 2011 May 9

ABSTRACT

Interference detection in Gaussian noise is proposed. It can be applied for easy detection and editing of interference lines in radio spectral line observations. One does not need to know the position of occurrence or keep track of interference in the band. By using statistical properties of N-channel Gaussian noise it is possible to differentiate interference from normal Gaussian noise. These statistical properties are three quantities which are calculated on the subject spectrum. The first is the expected absolute difference maximum from the mean level across the spectrum. The second is the expected absolute difference maximum (array of adjacent differences of channel values taken across the spectrum), and the third is the expected absolute added difference maximum. For N-channel Gaussian noise, these quantities have a well-defined value. Any deviation across the spectrum that violates an upper limit formed by these expected maxima is attributed to interference. Results obtained on real data by applying this technique have been displayed.

Key words: methods: data analysis – methods: statistical

1. INTRODUCTION

Radio frequency interference (RFI) is a common problem during radio spectral line observations. RFI can be edited manually by inspecting the individual power spectra. RFI lines can appear at various positions in the spectrum depending upon the nature of the interference or the instrument settings. Here a few methods of detection have been proposed which apply to the power spectrum (i.e., after the Fourier transform of the raw data) so that the RFI can be detected and edited using an algorithm. By using such an RFI detection algorithm one can analyze large amounts of data with minimum time and effort. This technique is similar in a way to the already existing RFI mitigation tools (Winkel et al. 2007; Nita & Gary 2010) where a certain threshold on variance/moments is sought to discriminate RFI, but is completely different in approach and method. The method has been described as it applies to radio spectral line observations using dual Dicke switching (Rohlfs & Wilson 2006, Chp. 4).

2. THE RADIO SPECTRAL LINE OBSERVATION

In a typical radio spectral line observation, one first decides upon the frequency at which the line will be observed. The bandwidth required to detect the line(s) and the period for which the observation will be made. This period decides the signal-to-noise ratio in the final spectrum. The line is assumed to be observed with the procedure of dual Dicke switching. In this observational procedure, the band is split into two equal parts. The spectral line is made to appear in these two parts alternately in time by appropriately tuning the telescope. This is usually done by selecting two fixed local oscillator (LO) frequencies, LO1 and LO2, over a period of time so that the gain characteristics of the instrument do not change significantly. Hence, the observation consists of two types of spectra one with the spectral line appearing in the right part and the other with the spectral line in the left part. The first is called the T_on spectrum while the other is called the T_off spectrum. Typical power spectra appear as shown in Figure 1.

To detect the astronomical spectral line, one has to first eliminate the background power (the band profile, Figure 1, which is essentially the background radio power multiplied by the gain of the instrument) and normalize for the gain across the band. This can be achieved by simply subtracting T_off from T_on, since both have the same band profile. In other words, the background radio power in the band is not going to change significantly by a small shift (LO1−LO2) of the center of the band. The gain can be normalized by dividing T_on − T_off by T_off; the noisy features of T_off in the denominator are not going to affect the ratio seriously since the derivative of 1/x is 1/x^2 (here x is T_off). The main portion of the band where the lines are expected to appear has relatively higher magnitudes than the noise features. This can be seen in Figure 1. The resulting output can be written as

\[ PS = \frac{T_{on} - T_{off}}{T_{off}}. \]  

PS consists of two lines (panel (c) of Figure 1), one from T_on and the other from T_off. The spectral line from T_off will be inverted since −ve of T_off was added to T_on. PS is assumed to be Gaussian noise (or approximately Gaussian) with an inherent baseline. Noise characteristics of real data have been displayed in Figure 2. Since there is no difference between addition or subtraction of Gaussian noise due to its symmetry, both lead to Gaussian noise. After this, PS is folded and averaged appropriately so that the inverted line overlaps with the one in the other part. Of course, a negative of the inverted part will be added. This completes the simple data processing. Many such spectra collected over time can be added to improve the signal-to-noise ratio, which eventually results in spectral line detection. It is the property of Gaussian noise that the rms of the averaged spectrum reduces by a factor of 1/√n when n spectra are averaged (Rohlfs & Wilson 2006, Chp. 4).

2.1. Interference

RFI is a common enemy of spectral line observations. It can be produced in a variety of ways, by electrical sparking, computers, electronic gadgets, etc. The magnitude of interference can be both small and large. It is easy to detect the large interference while the small interference is the one which is to be tackled. A typical interference-infected spectrum is as shown in Figure 3.
Figure 1. Typical simulated power spectra with dual Dicke switching at tuning frequencies LO1 and LO2, and their combination $T_{on}/T_{off} − 1$. Abscissa is the channel number and the ordinate is arbitrary power. The spectral lines near channels 50 (for $T_{off}$) and 200 (for $T_{on}$) have been highly exaggerated for the sake of clarity. An interference line in the middle of $T_{on}$ has been introduced for comparison. The bottom two panels, (d) and (e), show the difference array (explained later) for $T_{on}/T_{off} − 1$ and $T_{on}$, respectively. As one can see, the broad spectral lines are suppressed in these arrays. They do not inherit the baseline from their parents, panels (c) and (a), respectively, and are symmetrically distributed about the zero line. Panel (d) or (e) is more suitable for interference detection than panel (c) or (a) due to this elimination of baseline.

(sometimes it is low enough that it is distinguishable only in the ratio $(T_{on} − T_{off})/T_{off}$).

Different sources produce different types of RFIs. For example, the RFI generated by electrical sparking will produce a series of lines. In such cases, the standard deviation of the data itself may increase drastically (Fridmann 2008). By using this information such data can be discarded while some other electronic equipment may produce a single or multiple number of lines at some fixed position(s) in the band. In such cases, one has to devise RFI excision techniques (Fisher 2002).

2.2. RFI Detection

The RFI lines appearing in the power spectrum can be recognized manually by carefully inspecting it, but this would be laborious and time consuming. With the methods described here, interference can be detected within the analysis program to autorecognize them. Using an appropriate algorithm one can also have them edited (this is left to the user). This reduces manual labor and results in speedy data analysis. In the methods proposed here, the following assumptions have been made.

1. The RFI can be observed in a short integration time whereas the detection of the astronomical signal would require a much longer integration time. In other words, the spectrum subjected to interference detection satisfies the criteria that the astronomical line is smaller than or comparable to the standard deviation of the spectra. Also broad features that are characteristic of astronomical lines disappear in the difference array (explained later), which is the one that is actually subjected to interference detection.

2. The RFI is narrowband, i.e., only a small portion or portions of the band are infected with it. It is necessary to use a portion of the band to find the standard deviation of the spectrum. This parameter is necessary in the detection of interference. The standard deviation is an important quantity relevant to noise.

3. It is also not necessary that the RFI be time variable. If the spectrum has interference as per the technique used here,
Figure 2. Noise sampled from a large collection (ORT: 100 sets; WSRT: 23 sets) of $T_{\text{on}}/T_{\text{off}} - 1$ after baseline removal (Figure 1). For ORT $T_{\text{on}}$ and $T_{\text{off}}$ correspond to the 1 s spectrum, whereas for WSRT it is 1 minute.

Figure 3. Typical power spectra with a couple of interference lines appearing in $T_{\text{on}}$ at $\sim 140$ and $\sim 180$. Here the amplitude has been chosen to be small, but in reality it can be either much higher or smaller. The difference array and added difference array for $T_{\text{on}}/T_{\text{off}} - 1$ are shown in Figures 4 and 5.

the associated channels will be flagged or else they will pass unflagged.

The technique used to detect RFI depends on the properties of Gaussian noise. When pure Gaussian noise across $N$ channels is considered, an expected maximum value ($X_{\text{max}}$) exists which can occur across the channels. This expected maximum is the mean of the maximum values from an ensemble of $N$-channel spectra. Similarly, such a maximum for the difference between two adjacent channel values also exists. Lastly, another quantity of interest is the added difference, which is the sum of consecutive differences of the same sign (i.e., the adjacent differences could be either $-ve$ or $+ve$). Suppose these difference signs are $++--++--$. The fifth and sixth differences would be added to form one quantity. These quantities can be calculated analytically and used to inspect the Gaussian noise. Typical interference that is sharp and narrow would produce large channel values or differences between adjacent channel values that deviate considerably from the expected maxima values. When this occurs, the corresponding channel or channels can be flagged for interference. The first quantity, the expected maximum, can be of use when the spectrum has nearly zero mean across the spectrum; in other words, it has zero baseline,
so is of less importance. The other two quantities are more tolerant toward this defect (panels (d) and (e) of Figure 1) since they are differences between adjacent channels which more or less have the same local mean value. The methods using these three quantities have been called direct, difference, and added difference methods. The added difference traps those interference lines that escape the difference method because they are slightly broad. Plots of these quantities signifying their use for detecting interference have been given in the following sections.

2.2.1. Direct Method

The direct method is the simplest method. In this method, the spectrum of \( N_{\text{data}} \) channels is divided into \( n_{\text{blk}} \) number of blocks. Each block has now

\[
N_{\text{blk}} = \frac{N_{\text{data}}}{n_{\text{blk}}} \quad (2)
\]

number of data points. The block with the minimum standard deviation (\( \sigma \)) will be considered to be representative of good data, meaning it is free of any interference. This makes sense because in the regions without any signal one has only noise, so any interference introduced into this data will increase the standard deviation through an increase in the mean value. This can be understood from the formula for standard deviation as follows. Statistically, since Gaussian noise is symmetric about its mean (for argument’s sake, let the mean be 0) one is likely to find a variety of values distributed above and below the mean line with equal probability. Considering one such set of values between \( x \) and \( x + \delta x \) both above and below the mean line, the differences with a mean above (below) the mean line would reduce while with those below (above) the line would increase when the mean deviates due to contamination from interference. In the \( \sigma \) calculation it is the square of this difference that contributes to this change. Since the reduced square and the increased square do not compensate for each other, resulting in a net +ve contribution, this deviation increases the value of \( \sigma \). The minimum standard deviation block will be used as a reference to calculate the required parameters for detecting interference as well as to edit and repair it (one can replace the edited portion appropriately with noise of the same standard deviation as that in the reference block). The absolute maximum value in this block can be used to generate a limit on the maximum absolute value that could be encountered in the other blocks by choosing a suitable multiplicative factor. This can be done since the expected maximum (\( X_{\text{max}} \)) is directly proportional to the standard deviation. Alternatively, it is recommended that one use the formula given below for \( X_{\text{max}} \) along with the standard deviation in the reference block and the total number of channels (\( N_{\text{data}} \)). It should be noted that a constant multiplicative factor (>1, say 1.1–1.3) has to be used with the expected maximum value to allow for realistic departures (rd) from the expected value. If any block has any channel or channels with an absolute value in this block can be used to generate a limit on the maximum absolute value that could be encountered in the other blocks by choosing a suitable multiplicative factor. This can be done since the expected maximum (\( X_{\text{max}} \)) is directly proportional to the standard deviation. Alternatively, it is recommended that one use the formula given below for \( X_{\text{max}} \) along with the standard deviation in the reference block and the total number of channels (\( N_{\text{data}} \)). It should be noted that a constant multiplicative factor (>1, say 1.1–1.3) has to be used with the expected maximum value to allow for realistic departures (rd) from the expected value. If any block has any channel or channels with an absolute value greater than this, then those channels in that block can be flagged for interference. The realistic departure can also be estimated as the Gaussian probability given by \( X_{\text{max}} \) reducing to 1/10th of its value; this is achieved at \( X_{\text{Mr}} \). Hence, \( X_{\text{Mr}} \) would be the sought-after upper limit. The expected maximum absolute value for \( N \)-channel Gaussian noise with zero mean is given by

\[
\langle X_{\text{mx}} \rangle = \sigma \sqrt{2} \int_0^1 \text{erf}^{-1}(s \frac{\hat{x}}{\sigma}) \, ds, \quad (3)
\]

which has a convenient approximation

\[
\langle X_{\text{mx}} \rangle \approx \sigma \sqrt{2} \text{erf}^{-1}(1 - \frac{0.5833}{N}). \quad (4)
\]

Using this, the upper limit (\( X_{\text{Mr}} \)) allowing for rd can be written as

\[
\langle X_{\text{Mr}} \rangle = \sqrt{\langle X_{\text{mx}} \rangle^2 + 4.6\sigma^2}. \quad (5)
\]

The derivation of these equations is discussed in Appendices A and B. This method requires that the spectrum under investigation be symmetrically distributed about the zero line. This can be brought about by calculating the mean of all the channels of the data and subtracting it from each channel value of the spectrum. Also, in some cases, a baseline should be removed from the spectrum before subjecting it to RFI detection. In the case where interference is present, this is difficult since the fitted curve will be biased by the interference line. If one is sure that the spectrum has no baseline then one can use this method or else omit it. However, the following methods are more useful.

2.2.2. Difference Method

The difference method is a simple method of interference detection similar in procedure to the direct method. In this method, adjacent channel differences across a single spectrum are found. For \( N \)-channel Gaussian noise an upper limit on the maximum expected difference exists which will be used as a reference to limit the maximum value that could be encountered across the \( N \) channels. Further, the spectrum is divided into \( n_{\text{blk}} \) blocks. Again the reference block will be the minimum standard deviation block. A typical plot of differences of an interference-infected block with a mean above (below) the mean line would increase the fit to curves will be biased by the interference line. If one is sure that the spectrum has no baseline then one can use this method or else omit it. However, the following methods are more useful.

\[
\langle D_{\text{mx}} \rangle \approx 2\sigma \int_0^1 \text{erf}^{-1}(s \frac{\hat{x}}{\sigma}) \, ds = \sqrt{2}\langle X_{\text{mx}} \rangle. \quad (6)
\]

An upper limit (\( D_{\text{Mr}} \)) taking into account an rd similar to Equation (5) can be written as

\[
\langle D_{\text{Mr}} \rangle = \sqrt{\langle D_{\text{mx}} \rangle^2 + 9.2\sigma^2}, \quad (7)
\]

where \( \sigma \) can be taken from the reference block. The derivation of these equations is discussed in Appendix B. Since for a given observation \( N \) is fixed, one has to calculate the inverse error function only once. This can be easily done in octave invoking the library function \text{erfinv}.

2.2.3. Added Difference Method

The added difference method is similar to the difference method except that consecutive differences with the same sign are added to form a new array of quantities. Suppose the difference signs are + + + + + + + + + + + + + +. In this example, the fifth and sixth differences are added to form one quantity. The plot of such an array for a typical interference-infected spectrum is shown in Figure 5 which is related to Figure 3. The improvement in interference detection criteria is clearly visible in this plot.
The added difference method traps those interference lines that escape the difference method because they are slightly broad. For this method, the constraining limit ($Q_{mx}$) can be obtained by finding the maximum added difference in the reference block or by using the empirical formula below (recommended) with a suitable $rd$ factor ($>1$):

$$Q_{mx} \approx 2 \sigma \text{erf}^{-1} \left( 1 - \frac{0.5833}{2N} \right).$$

Reliable results on Ooty Radio Telescope (ORT) data have been obtained using Equations (6) and (8) with $rd = 1.3$ for both. Using $D_{mx}$ in place of $rd \times D_{mx}$ also produces similar results.

3. OOTY RADIO TELESCOPE

ORT is situated near the town of Ooty in southern India at a longitude of 28°3 and a latitude of 11°23’. ORT (Swarup et al. 1971) is an off-axis parabolic cylinder with a length of 530 m and a width of 30 m. The telescope is located on a hill which...
The interference detection technique described here is quite simple. It makes use of the properties of Gaussian noise to discriminate RFI from background noise. These quantities which can be used to distinguish Gaussian noise from interference are most importantly the expected difference maximum \((D_{mx})\) and the expected added difference maximum \((Q_{mx})\). As discussed in Section 2.2.1 even though the expected maximum can be used as a criterion for tracking interference, it lacks immunity toward inherent baselines. The difference array formed from adjacent differences is more reliable in this respect, since the baseline is almost self-killed. Hence, it is suitable as a quantity to be investigated for interference tracking. The added difference array being formed out of the difference array has this same benefit. The difference maximum is found on the array formed by the differences of adjacent channel values. For an \(N\)-channel spectrum there would be an \(N - 1\) channel difference array. For the \(N\)-channel Gaussian noise spectrum there is always a fixed expected value for the expected difference maximum and likewise an expected added difference maximum. These

The data were analyzed and debugged of interference using the technique described in Sections 2.2.2 and 2.2.3. The results obtained toward a few positions have been displayed in Figure 6.

4. SUMMARY

Figure 6. Results obtained using the difference and added difference methods on spectral line observations at ~328 MHz (\(H271\alpha\)) using the ORT toward four Galactic positions (inset left). WAE: collapse With Automated Editing and baseline removal; AD: collapse of Actual Data. The plots on the right are the auto-edited versions of the plots on the left. Interference detection and editing was performed on 10 s data and again after collapsing 20 of these sets. The collapse on the right consists of a straight line connecting adjacent channels on the two sides of the interference line. It can be noted that the broad astronomical spectral features are intact in WAE except for the absence of sharp interference lines. Here \(rd = 1.3\) (see Section 2.2.1).

has a natural slope of 11.4 equal to the geographical latitude of the place; this gives it the feature of equatorial mount. The operating frequency of the telescope is centered at 326.5 MHz with a maximum bandwidth of 15 MHz at the front-end. The reflecting surface of the cylinder is made of 1100 stainless steel wires running parallel to each other along the entire length of the telescope. An array of 1056 half-wave dipoles in front of a 90° corner reflector forms the primary feed of the telescope. The 1056 dipoles are in groups of 48. The signals received by these groups are added in phase to form 22 group outputs, each known as a module. The telescope is divided into a northern part and a southern part. The northern modules are designated as N1 to N11 and the southern modules are designated as S1 to S11. The beam width due to each module is 2.3 in the east–west direction and 2.2 sec(\(\delta\)) in the north–south direction, where \(\delta\) is the declination. This forms the observing mode and beam for the current project of radio recombination line (RRL) observations.

The RRL observations were made using a new digital backend (Prabu 2010) which provided a bandwidth of 1 MHz for the spectral line mode. This narrow bandwidth restricts the observation of only one recombination line at a time. The transition that was observed was \(H271\alpha\) which has a rest frequency of 328.5958 MHz. The observations were carried out using dual Dicke switching as described earlier in Section 2 with a shift of 300–400 kHz between the two line positions.
The quantities can be directly calculated using the formulae given by Equations (3), (4), and (6). A statistical upper limit allowing for rd for these quantities is given by Equations (5) and (7). Any deviation for these values above the upper limits is used to track interference. It should be noted that the added difference array is not a simple one like the difference array. It requires an auxiliary array to keep a record of the number of channels that have taken part in forming a single added difference channel so that when the added difference maximum upper limit is violated in the added difference array one can trace back to the exact channel numbers in the parent spectrum (T_{on}/T_{off} − 1) to locate interference. A difference array and an added difference array for a typical interference-infected spectrum are shown in Figures 4 and 5, respectively. Notice the reduction in the number of channels for the added difference array. An auxiliary array maintains a record of the number of difference array channels that map to each channel of the added difference array.

Figure 7 shows a comparison between analytical and experimental values of $X_{mx}$ and $D_{mx}$ (using octave’s `randn` function) for different lengths of the Gaussian noise spectrum. Good agreement is seen even at low values of $N$, like 20. As the magnitude of $N$ increases, the agreement between analytical and experimental values becomes stronger. The variation of the magnitude of $X_{mx}$ and $D_{mx}$ with the length of the spectrum in terms of channel number $N$ is shown in Figure 8. The profiles have been overplotted with experimental values to show the performance of Equations (6) and (8).

To support the applicability of the above technique to real data, the nature of noise was investigated for two radio telescopes, the ORT and Westerbork Synthesis Radio Telescope (WSRT). The noise histograms for the two telescopes show believable Gaussian profiles, as shown in Figure 2.

Finally, results obtained on ORT RRL data are displayed in Figure 6 for four positions. It can be noted (right panel) that while the sharp interference lines are edited out, other features elsewhere in the spectrum, including the broad astronomical line, persist and can be compared with sufficient detail with the one that was formed without interference editing (left panel).

It should be mentioned explicitly here that the detection-editing procedure can be applied multiple times (at at least two stages) with regard to cumulation of data. The data subjected to RFI detection–excision–cumulation can be once again subjected to RFI detection–excision. This will eliminate the residual RFI. However, care must be taken that the conditions given in Section 2.2 are still valid at the application stage. This can be easily handled by the knowledge of the amount of data and the expected signal-to-noise ratio in the final spectrum.

The author thanks the referee for comments and suggestions regarding the presentation of this paper.

**APPENDIX A**

**DERIVATION OF THE EXPECTED MAXIMUM VALUE FOR THE GAUSSIAN DISTRIBUTION**

Consider $N$ channels which produce Gaussian noise (Figure 9) of the same standard deviation $\sigma$. The probability that an absolute maximum value of $c$ is produced in these channels is seen even at low values of $N$, like 20. As the magnitude of $N$ increases, the agreement between analytical and experimental values becomes stronger. The variation of the magnitude of $X_{mx}$ and $D_{mx}$ with the length of the spectrum in terms of channel number $N$ is shown in Figure 8. The profiles have been overplotted with experimental values to show the performance of Equations (6) and (8).

$$p(c) = \frac{2^N N e^{-\frac{c^2}{2\sigma^2}}}{(\sigma \sqrt{2\pi})^N} \int_0^c e^{-\frac{x^2}{2\sigma^2}} dx_1 \int_0^c e^{-\frac{x_2^2}{2\sigma^2}} dx_3 \ldots dx_N, \quad (A1)$$

which is nothing but the integral of all possibilities over the remaining $N-1$ channels keeping one of the channels fixed at $c$. The factor $2^N$ is due to the fact that both $+ve$ and $-ve$ values are possible in each channel. The remaining $N-1$ channels are allowed to take all values between $-c$ and $+c$. $N$ appears in the numerator because there are $N$ ways to fix a channel value to $c$. Each of the above integrals evaluates to $\sigma \sqrt{\frac{\pi}{2}} \text{erf}(c/\sigma \sqrt{2})$.
Figure 8. Plots of \( D_{mx} \) and \( Q_{mx} \) as a function of \( N \) given by Equations (14) and (30) for \( \sigma = 1.0 \). The circular dots are experimental values using octave’s \texttt{randn} function.

Figure 9. Calculating the maximum expected value in \( N \)-channel Gaussian noise.

(Abramowitz & Stegun 1972) resulting in

\[
p(c) = \frac{2N e^{-\frac{c^2}{2\sigma^2}}}{\sigma \sqrt{2\pi}} \left[ \text{erf} \left( \frac{c}{\sigma \sqrt{2}} \right) \right]^{N-1}.
\]

Now the maximum expected value can be written as

\[
\langle X_{mx} \rangle = \int_0^\infty c \ p(c) \ dc \tag{A3}
\]

\[
\langle X_{mx} \rangle = \frac{2N}{\sigma \sqrt{2\pi}} \int_0^\infty c \ e^{-\frac{c^2}{2\sigma^2}} \left[ \text{erf} \left( \frac{c}{\sigma \sqrt{2}} \right) \right]^{N-1} dc, \tag{A4}
\]

which can also be written using the derivative of \( S = [\text{erf}(\frac{c}{\sigma \sqrt{2}})]^N \) and expressing \( c \) in terms of \( S \) as

\[
\langle X_{mx} \rangle = \sigma \sqrt{2} \int_0^1 \text{erf}^{-1}(S^{\frac{1}{N}}) dS, \tag{A5}
\]

which has an approximation of

\[
\langle X_{mx} \rangle \approx \sigma \sqrt{2} \text{erf}^{-1} \left( 1 - \frac{0.5833}{N} \right) \tag{A6}
\]

derived by considering the series expansion for \( \text{erf}^{-1}(S) \). To obtain this approximation for \( X_{mx} \), essentially we must find an approximation for the integral in Equation (A5). By substituting the inverse error function’s series expansion we see that

\[
\int_0^1 \text{erf}^{-1}(S^{1/N}) \ dS = \sum_{k=0}^{\infty} \frac{C_k}{2k+1} \left( \frac{\sqrt{\pi}}{2} \right)^{2k+1} \int_0^1 S^{\frac{2k+1}{N}} \ dS \tag{A7}
\]

\[
\int_0^1 \text{erf}^{-1}(S^{1/N}) \ dS = \sum_{k=0}^{\infty} \frac{C_k}{2k+1} \left( \frac{\sqrt{\pi}}{2} \right)^{2k+1} \left[ \frac{S^{\frac{2k+1}{N}}}{N} \right]_0^1
\]

\[
= \sum_{k=0}^{\infty} \frac{C_k}{2k+1} \left( \frac{\sqrt{\pi}}{2} \right)^{2k+1} \frac{N}{2k+N+1}. \tag{A8}
\]

Now comparing the result of Equation (A8) with the general term in series expansion of \( \text{erf}^{-1}(z) \) we see that in place of \( z^{2k+1} \) we have \( N/(2k + N + 1) \). Thus, to seek an approximate value of
Equation (A10) and considering the first five terms. This intro-
duces +ve errors in the lower-order terms but helps to account
for the truncated higher-order terms. In this approximation,
the hand side of Equation (A5) is essentially replaced by erf
and hence the result is Equation (A6).

A better approximation is obtained by binomially expanding
Equation (A10) and considering the first five terms. This intro-
duces +ve errors in the lower-order terms but helps to account
for the truncated higher-order terms. In this approximation,
2k + 1 ∼ N for all the terms. This yields the approximation
\[ z = \left(1 - \frac{0.5833}{N}\right). \]  

Now with this approximation in hand, the integral on the right-
hand side of Equation (A5) is essentially replaced by erf\(^{-1}(z)\)
and hence the result is Equation (A6).

**APPENDIX B**

**DERIVATION OF THE EXPECTED MAXIMUM-DIFFERENCE VALUE FOR THE GAUSSIAN DISTRIBUTION**

To obtain the maximum expected difference between two
adjacent channels in the N-channel Gaussian noise spectrum,
first the probability of a difference of \( d \) occurring in two chan-
nels is calculated (Figure 10). Further, it is assumed that each of
the adjacent channels behaves approximately independently and
the two-channel probability distribution is applied to the N chan-
nels. This approximation becomes better when one considers a large
number of channels, as can be seen by the analysis of three-
channel Gaussian noise:

\[ p_2(d) = \frac{1}{2\pi \sigma^2} \int_{-\infty}^{+\infty} e^{-\frac{x^2 + y^2}{2\sigma^2}} \, dx, \tag{B1} \]

so the probability density for a two-channel difference is

\[ p_2(d) = \frac{1}{2\sigma \sqrt{\pi}} e^{-\frac{d^2}{2\sigma^2}}. \tag{B2} \]

The three-channel probability analysis is as follows.

In Figure 11, \( x_1 \) and \( x_2 \) are fixed such that the difference
between them is \( d \). \( x_3 \) is allowed to take any value such that
its magnitude of difference with \( x_2 \) does not exceed \( d \). Next,
the same is repeated to account for the possibility of \( x_2 \) and \( x_3 \)
being fixed and \( x_1 \) being free to take variable values. The net
probability would be integrating over \( x_2 \) from \(-\infty\) to \(+\infty\). So
we can write this net probability density as

\[ p_{3_2}(d) = \frac{2 \times 2}{(\sigma \sqrt{2\pi})^2} \int_{-\infty}^{+\infty} e^{-\frac{x^2 + y^2}{2\sigma^2}} \int_{-d}^{+d} e^{-\frac{(y-x)^2}{2\sigma^2}} \, dy \, dx \tag{B3} \]

\[ p_{3_3}(d) = \frac{1}{\pi \sigma^2} \int_{-\infty}^{+\infty} e^{-\frac{x^2 + y^2}{2\sigma^2}} \left( \text{erf} \left( \frac{d + x}{\sigma \sqrt{2}} \right) + \text{erf} \left( \frac{d - x}{\sigma \sqrt{2}} \right) \right) \, dx. \tag{B4} \]

The factor of two in the numerator is due to the fact that \( d \)
in the first integrand takes two values +ve and -ve and in both
the cases it integrates to the same amount. Another two is due
to the interchange of roles between \( x_1 \) and \( x_3 \). When a similar
analysis is conducted by considering four channels and taking
the difference in two channels independently, we see that the
probabilities are approximately the same for the occurrence of
a difference \( d \) in the two cases. The probability in the case of
four channels, but pairwise, is

\[ p_{4_2}(d) = \frac{2 \times 2}{(\sigma \sqrt{2\pi})^2} \int_{-\infty}^{+\infty} e^{-\frac{x^2 + y^2}{2\sigma^2}} \int_{-d}^{+d} p_2(y) \, dy \, dx \tag{B5} \]

\[ p_{4_3}(d) = \frac{2}{\pi \sigma^2} \int_{-\infty}^{+\infty} e^{-\frac{x^2 + y^2}{2\sigma^2}} \text{erf} \left( \frac{d}{2\sigma} \right) \, dx \]

\[ = \frac{2}{\sigma \sqrt{\pi}} e^{-\frac{d^2}{2\sigma^2}} \text{erf} \left( \frac{d}{2\sigma} \right). \tag{B6} \]

This calculation can also be performed by applying \( p_2(d) \)
to both pairs. It can be checked (checked numerically here and
the first-order approximation shows that the two agree within a
factor of \( \sqrt{2} \) that for a given value of \( d \),

\[ p_{3_2} \approx p_{4_1}. \tag{B7} \]

Hence, one can assume that the channels behave approxi-
mately independently pairwise in a series of \( N \) channels. By
considering the difference probability distribution for two chan-
nels and treating the \( N-1 \) channel pairs as new single channels,
in which the occurrence value is not “x” but “d,” we can write (using the same derivation as for \(X_{mx}\) for the maximum expected difference value in \(N\)-channel Gaussian noise)

\[
\langle D_{mx} \rangle \approx \frac{N-1}{\sigma \sqrt{\pi}} \int_0^\infty d \ e^{-\frac{d^2}{4\sigma^2}} \left[ \text{erf} \left( \frac{d}{2\sigma} \right) \right]^{N-2} \, dd,
\]

which again using \(S = [\text{erf} \left( \frac{d}{2\sigma} \right)]^{N-1}\) can be written as

\[
\langle D_{mx} \rangle \approx 2\sigma \int_0^1 \text{erf}^{-1} \left( \frac{S}{\pi \sigma} \right) \, dS.
\]

For large \(N\) the \(-1\) in the exponent can be dropped and one can simply use Equation (13):

\[
\langle D_{mx} \rangle \approx \sqrt{2} \langle X_{mx} \rangle.
\]

To account for the rd, one has to use the distribution function \(p_2(d)\), Equation (22). \(D_{mx}\) would give 1/10th of the probability given by \(D_{mx}\) in \(p_2(d)\); this results in Equation (7).

The derivations given here are those of the author. He is unaware of any other such derivations or results.

REFERENCES

Abramowitz, M., & Stegun, I. A. 1972, Handbook of Mathematical Functions (10th ed.; New York: Dover), 297

Fisher, J. R. 2002, in ASP Conf. Ser. 278, Single-Dish Radio Astronomy: Techniques and Applications, ed. S. Stanimirovic et al. (San Francisco, CA: ASP), 433

Fridmann, P. A. 2008, \(AJ, 135, 1810\)

Nita, G. M., & Gary, D. E. 2010, \(MNRAS, 406, L60\)

Prabu, T. 2010, PhD Thesis, Raman Research Inst./Indian Inst. of Science

Robits, K., & Wilson, T. L. 2006, Tools of Radio Astronomy (Astrophys. Space Sci Lib XVI; New York: Springer)

Swarup, G., et al. 1971, \(Nature, Phys. Science (Lond), 230, 185\)

Winkel, B., Kerp, J., & Stanko, S. 2007, \(Astron. Nachr., 328, 68\)