

# Performance Study of a Cross-Frequency Detection Algorithm for Pulsed Sinusoidal RFI in Microwave Radiometry

B. Güner, N. Niamsuwan, and J. T. Johnson

**Abstract**—An analysis of the performance of a cross-frequency detection algorithm for pulsed sinusoidal radio frequency interference (RFI) is performed. The performance obtained is compared to that of pulse and kurtosis detection methods that have been previously reported. The results of the study show that the cross-frequency algorithm provides good performance in detecting pulsed sinusoidal RFI at high duty cycles, including the case of continuous sinusoidal interference. Use of the cross-frequency algorithm requires choice of a detection threshold that in practice can be estimated using measured data. The effect of this threshold estimation on the detector performance is also examined.

## I. INTRODUCTION

SEVERAL recent works have documented the detrimental effects of radio frequency interference (RFI) on Earth observing passive microwave radiometer systems [1]-[5]. Methods for the detection and mitigation of RFI [6]-[15] have also received attention, and several radiometric systems capable of RFI detection (and mitigation) with varying performance against different RFI types have been implemented. This paper is focused on the detection of RFI in measured radiometer data; the impact of mitigating detected RFI is not considered.

Most of the systems implemented for RFI detection to date have focused on finding signals concentrated in either time or frequency. Such signals are generally caused by anthropogenic emissions rather than the broadband thermal noise that radiometers are intended to measure. Detectors that search for energy outliers in the time-domain are called “pulse” detectors [9]-[11], while those that search in the frequency domain here are called “cross-frequency” detectors [6],[11]. Such methods are effective against RFI sources that are narrow in time or frequency, respectively. RFI detection can also be performed by testing the distribution of the fields received by the radiometer, which should be Gaussian for thermal noise; a deviation from the Gaussian distribution indicates the presence of RFI. Although numerous methods exist to test Gaussianity (e.g. [16]-[17]), to date only the kurtosis test of Gaussianity has been used in practical implementations [12]-[14]. The kurtosis test has been shown to be an effective tool against a wide variety of RFI types, but a blind spot in detecting pulsed sinusoidal interference has also been reported [12]. Later studies have shown that it is possible to remove this blind spot and to improve detection performance

by sub-dividing the integration period into multiple shorter time intervals (i.e. subsampling in time) and by sub-dividing the instrument bandwidth into multiple frequency bands (i.e. subsampling in frequency) and applying the kurtosis test to each time/frequency sub-sample [18],[19].

A central question for future radiometric systems is the selection of a detection method (or methods) that results in the best performance for a specific application. Such a selection of course depends on the RFI environment that is to be observed, but present knowledge of the RFI environment remains limited. One RFI type that has been considered in previous studies is pulsed sinusoidal RFI, due to the ability of this type to describe radar-like emissions (low duty cycles) as well as continuous (i.e. very narrowband or “CW”) sinusoidal interference. Comparative studies of several algorithms for pulsed sinusoidal RFI have been performed [19]-[21]. However, the performance of cross-frequency detection algorithms was not analyzed in the past. Because cross-frequency detection only uses brightness temperatures recorded in multiple frequency channels, it can be performed for radiometers having multiple frequency channels without additional requirements on hardware or data rate, and also can be used together with other methods such as kurtosis algorithms. Therefore, it is of interest for future systems to determine the detection performance of the cross-frequency approach, particularly in cases where other detection methods perform poorly.

This paper reports a theoretical performance model for cross-frequency detection of pulsed sinusoidal RFI with the motives outlined above. Results for pulse and time/frequency subsampled kurtosis methods are also provided for comparison purposes. In a practical implementation of the cross-frequency detector, the detection threshold should vary with respect to the brightness temperature of the observed scene in order to maintain approximately constant false alarm and detection rates. A method for estimating this threshold and the impact of this estimation on detection performance is also studied. Notations and the signal model used in the paper are introduced in the next Section, as well as the performance models used for each of the three detectors considered. Results comparing the performance of these detectors for different duty cycles and RFI strengths are then presented in Section 3, along with an experiment conducted to confirm the cross-frequency detector model developed. Special attention is given to the CW case where the pulse and subsampled kurtosis algorithms perform poorly. Effects of the system temperature estimation on the cross-frequency detector are studied in Section 4, and final

conclusions are provided in Section 5.

## II. FORMULATIONS

### A. Signal model and notations

The signal model employed in [20] (adapted from [22]) is used in most cases in this study when analyzing the cross-frequency approach. This model is based on a discretized representation of the fields observed by the radiometer, assuming sampling at the Nyquist rate of the observed bandwidth. However, a simplified signal model from [19] is used when comparing with the pulse and kurtosis methods because modeling kurtosis algorithm performance is more difficult under the signal model of [20]. The signal model of [20] is reviewed below; the simplifications employed in [19] are explained when the subsampled kurtosis method is discussed in Section II-D.

In the model of [20], a radiometer integration period of  $Q$  samples is divided into a set of  $I$  “frames” of length  $N$  samples each such that  $Q = IN$ . For the cross-frequency detector,  $N$  corresponds to the length of an FFT operation (in samples) used to produce the multiple frequency channels examined by the detector. Because the number of Nyquist samples  $Q$  in an integration period is typically a very large number for radiometer systems, it is impractical to perform an FFT of size  $Q$ , so that a smaller FFT size  $N$  is necessary, resulting in  $I$  FFT outputs per integration period. In addition, it is expected that most radiometer systems having multiple frequency channels (whether obtained through an FFT operation as considered here or through analog means) will record brightnesses in all the channels, causing an increase in the system data rate as  $N$  is increased. Values of  $N$  up to 1024 have been used in some experiments [11], but here smaller values of  $N$  are utilized that are more realistic for potential future satellite-based systems.

For illustration,  $Q$  is chosen as 768000 samples in most of the results considered along with  $N$  values of 8, 16 and 32. For a radiometer sampling, for example, at 20 nsec (i.e. the Nyquist rate of a 50 MHz bandwidth), the value of  $Q$  used results in a 15.36 msec radiometer integration period.

When pulsed sinusoidal RFI is present, sampled received fields can be written as [20]:

$$x_i[n] = A \cos(2\pi f_0 [(i - i_0)N + n] + \phi) \mathcal{I}(n, i) + w_i[n] \\ n = 0, 1, \dots, N - 1, \quad i = 0, 1, \dots, I - 1 \quad (1)$$

where  $w_i[n]$  refers to independent identically distributed (i.i.d.) Gaussian samples (i.e. thermal noise) with zero mean and standard deviation  $\sigma$ . Pulsed sinusoidal interference is added to this thermal noise, with  $i_0$  representing the pulse arrival “frame”,  $A$  the pulse amplitude,  $f_0$  the pulse frequency, and  $\phi$  the initial phase. The function

$$\mathcal{I}(n, i) = \begin{cases} 1 & i_0 N + N_s \leq iN + n < i_0 N + N_s + N_p \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

describes the location (including the arrival sample  $N_s$  within frame  $i_0$ ) and duration ( $N_p$  samples) of the RFI pulse. In the signal model of [20], it is assumed that frequency  $f_0$ , phase  $\phi$ , and arrival sample  $N_s$  of successive pulses are uniformly distributed over the values that they may take (i.e. between 0

to  $1/2$  for  $f_0$ , 0 to  $2\pi$  for  $\phi$ , and 0 to  $N - 1$  for  $N_s$ ). The resulting RFI duty cycle,  $d$ , is equal to

$$d = \frac{N_{pulse} N_p}{IN} \quad (3)$$

where  $N_{pulse}$  is the number of pulses in the radiometer integration period (assumed to be an integer.)

It is of interest to represent the amplitude of the RFI sine waves in terms of the uncertainty in the radiometer power estimate because RFI that is comparable to the radiometer uncertainty is difficult to detect for traditional systems. For a radiometer measuring the noise power in bandwidth  $B$ , it is well known that the uncertainty in the radiometer power estimate is  $T_{sys}/\sqrt{BT}$ , where  $T$  is the duration of the radiometer integration period and  $T_{sys}$  is the radiometer system temperature. For Nyquist sampling of bandwidth  $B$ , a sample is recorded every  $1/(2B)$  seconds, and the number of samples recorded in integration period  $T$  is  $Q = 2BT$ . The radiometer uncertainty is thus  $T_{sys}\sqrt{2/Q}$  in terms of  $Q$ . RFI amplitudes are described in what follows in terms of the ratio  $R$  of the average RFI power contribution ( $dA^2/2$ ) to the uncertainty in the radiometer power estimate:

$$R = \frac{dA^2}{2T_{sys}} \sqrt{\frac{Q}{2}} \quad (4)$$

$$A = \sqrt{\frac{2RT_{sys}}{d}} \sqrt{\frac{2}{Q}} \quad (5)$$

Note the variance  $\sigma^2$  of the Gaussian random variables is identical to  $T_{sys}$ , so that

$$\frac{A}{\sigma} = \sqrt{\frac{2R}{d}} \sqrt{\frac{2}{Q}} \quad (6)$$

For a given RFI power ratio  $R$ , duty cycle  $d$ , and integration period  $Q$ , this equation determines the associated  $A/\sigma$  ratio, which for higher duty cycles and longer integration times can become very small. For example, for  $d = 1$  (continuous interference) and  $Q = 768000$ ,  $A/\sigma$  is approximately  $1/17.6$  when  $R = 1$ , implying a low signal-to-noise ratio interference. For digital radiometers employing analog-to-digital (A/D) converters, interference having small  $A/\sigma$  values may fall below the smallest power measurable by the A/D converter (i.e. fail to toggle any A/D converter bits), and therefore be unimportant. Equation (6) can be used to assess these effects if the placement of the noise power level ( $\sigma$ ) within an A/D converter’s dynamic range is specified. In what follows, it is assumed that sufficient dynamic range exists to observe both thermal noise and interference regardless of the associated  $A/\sigma$  value.

### B. Cross-frequency detection model

A schematic of the cross-frequency detection algorithm considered here is shown in Figure 1. Incoming time domain field samples  $x_i[n]$  are passed through a non-overlapping  $N$  point FFT operation to obtain FFT outputs  $X_i[k]$ . The power in these FFT outputs ( $|X_i[k]|^2$ ) is then computed; this quantity is called a power spectrum. A total of  $I$  power spectra is obtained

in the radiometer integration period of  $Q$  samples; these  $I$  spectra are averaged to obtain the average power spectrum  $Y[k]$ . The cross frequency detector chooses the maximum over  $k$  (as described below) of  $Y[k]$  and declares detection if the maximum exceeds a threshold value  $T_{\text{cross}}$ . This threshold is related to the expected scene brightness temperature if it is assumed that the detector is applied after calibration of the radiometer measurements. If the scene temperature is not known a-priori (as is the case in radiometry), methods for estimating the threshold can be applied as described in Section 4.

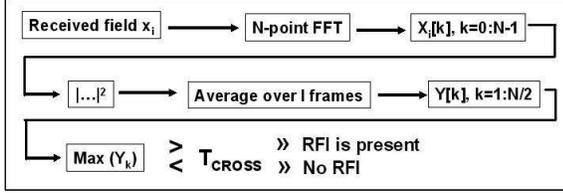


Fig. 1. Schematic of the cross-frequency detection algorithm

In the absence of RFI the FFT operation produces  $N$  complex Gaussian random variables for  $X_i[k]$ ; see Appendix one for details.  $N/2 - 1$  of these  $N$  values are positive frequencies that have corresponding conjugate negative frequency points; these negative frequencies are discarded. The square of the amplitude of each of these  $N/2 - 1$  positive frequency components is a chi-squared random variable with two degrees of freedom. The FFT operation also produces DC ( $k = 0$ ) and Nyquist ( $k = N/2$ ) frequency outputs that are purely real. For consistency with the positive frequency points, the power in the DC and Nyquist frequency outputs is averaged to obtain an additional chi-squared random variable with two degrees of freedom (additional details in Appendix 1). As a result,  $N/2$  such random variables are obtained for each frame; the algorithm is described as using  $N/2$  “channels” henceforth. The averaging over  $I$  spectra then produces  $N/2$  chi-squared random variables with  $2I$  degrees of freedom ( $Y[k]$ ). If these  $N/2$  random variables are independent, which is the case in the absence of RFI when a rectangular FFT “window” is used, the probability that the maximum of the  $Y[k]$ ’s is less than  $T_{\text{cross}}$  can be calculated by multiplying the probabilities that each of the  $N/2$  random variables is less than  $T_{\text{cross}}$ . The resulting quantity then allows computation of the detector false alarm rate as a function of  $T_{\text{cross}}$ ; see Appendix 2 for details.

When RFI is present, the detector involves the maximum of  $N/2$  non-central chi-squared random variables with  $2I$  degrees of freedom, as described in detail in Appendix 3. The non-centrality parameters of these random variables can be obtained by computing the Fourier transform of a specified pulsed source. These random variables remain independent so that the maximum operation can be computed as in the false-alarm case. An average probability of detection for a specified threshold can then be obtained by numerically averaging the probability of detection over the frequency, phase, arrival sample, and other properties of the RFI signal. This process includes the influence of “scalloping” loss in the FFT (i.e. the RFI pulse frequency is not aligned with an FFT bin) as well

as partial filling of FFT computations by an RFI pulse.

Since the cross frequency algorithm uses only frequency resolution and no time resolution in its detector, its performance should be best for RFI that is localized in frequency. Such cases concentrate RFI power into one of the frequency bins of the algorithm and increase the chance of detection.

### C. Pulse detection model

A schematic of the pulse detection model considered in this paper is shown in Figure 2. In this model, it is assumed that the square of time domain measured fields  $x_i[n]$  are summed over  $N$  samples. Here for convenience the number of time domain samples integrated is chosen to be  $N$  so that the total radiometer integration period  $Q$  is divided in a manner similar to the cross-frequency method. The resulting sum, denoted as  $PD_i$  in the schematic, is a chi-squared random variable with  $N$  degrees of freedom when there is no RFI. When RFI is present, a non-central chi-squared random variable with  $N$  degrees of freedom is obtained instead. The non-centrality parameter for this case can again be computed for specified RFI pulse properties. During a radiometer integration period of  $Q = IN$  samples,  $I$  such random variables are attained, and the detection statistic is the maximum of these  $N$  random variables. The probabilities of false alarm and detection for a specified threshold are calculated as the probability that the test statistic is greater than the threshold ( $T_{\text{PULSE}}$ ) in the absence and presence of RFI, respectively. See [20] for more detailed information on this performance model.

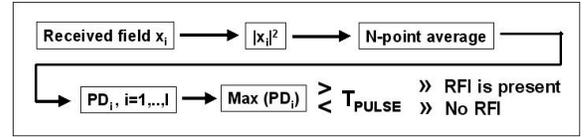


Fig. 2. Schematic of the pulse detection model

Typically the pulse detection algorithm performs best for pulse durations in samples that are comparable to  $N$ . Choosing  $N$  to be identical for the pulse and cross-frequency algorithms is motivated by the fact that pulse detection algorithms are often designed to detect interference that is much shorter than the radiometer integration period (i.e. small  $N$  values.) This is because high amplitude but very low duty cycle pulses that are easily detected by the pulse algorithm when  $N$  is small can produce low-level interference that is difficult to detect in brightnesses averaged over the entire radiometer integration period.

### D. Kurtosis detection model

A schematic of the kurtosis detector is shown in Figure 3. In this case,  $N/2$  frequency sub-channels are assumed to be obtained by filtering in hardware rather than through an FFT operation. In the absence of RFI, received fields in each channel are independent Gaussian random variables having a variance that is reduced from that of the total channel ( $\sigma^2$ ) by a factor of  $N/2$ . Additionally, each of the kurtosis sub-channels originally has  $2Q/N$  time samples to maintain the

Nyquist rate. This set of samples can further be split into a set of time sub-samples, shown as  $t_{ss}$  in the schematic, if desired. The kurtosis estimator  $K_n$  is computed for each time and frequency sub-sample by dividing the fourth central moment of the received field ( $m_4$ ) by the field's second central moment ( $m_2$ ) squared. For a sufficiently large number of field samples in the computation,  $K_n$  can be modeled as a Gaussian random variable with known mean and variance values as given in [14]. In the presence of RFI, the mean kurtosis can either become larger or smaller than its mean value in the absence of RFI (which is three), hence a two-sided test is used as shown in the schematic. The overall probabilities of detection and false alarm are then computed as the probability that the kurtosis values of at least one of the time/frequency subsamples exceeds the threshold of the two-sided test.

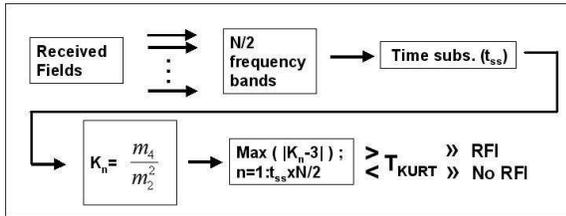


Fig. 3. Schematic of the kurtosis detection model

The process used here is based on that described in [19], in which a simplified version of the pulsed sinusoidal RFI signal model is used. In this model, pulsed RFI is assumed to occur in only one of the frequency sub-samples (i.e. no scalloping effects.) It is also assumed that RFI pulses, when present, begin at the beginning of the observations, and only one RFI pulse is present in each radiometer integration period (i.e.  $N_{pulse}$  is 1). To provide a fair comparison among the detectors, this simplified signal model is used in the results shown whenever the three detection methods are compared. See [19] for additional information on the kurtosis detector performance model.

The results to be shown use four time sub-samples for the kurtosis algorithm throughout; this implies that the  $2Q/N$  time samples in each radiometer channel are further subdivided into four groups of  $Q/(2N)$  samples, and the kurtosis computed for each sub-interval and frequency channel. Other numbers of time sub-samples were also examined (including no time sub-sampling) but only modest changes were observed in the results for the higher duty cycle interference of primary interest here.

### III. DETECTION PERFORMANCE RESULTS

#### A. Experimental validation of cross-frequency performance model

An experimental confirmation of the cross-frequency detector performance model was conducted using a digital radiometer sampling at 200 MSPS; raw A/D converter samples were recorded in the experiment in order to allow maximum flexibility in analyzing the results. A white noise input was produced using a cascade of amplifiers with the input terminated, and a sinusoidal source of controllable amplitude was combined

with the noise input. When included, the sinusoidal source was operated continuously (i.e. 100% duty cycle interference), and was tuned at a rate of 1.11 kHz/msec through the 100 MHz bandwidth sampled by the radiometer. This configuration reproduces the fact that the performance model averages the probability of detection over uniformly distributed RFI frequencies. An integration period  $Q$  of 524288 samples (2.62 msec) was used in the experiment because the radiometer had a 512K FIFO system for storing recorded data. The change in frequency of the sinusoidal source within this time interval (approximately 2.91 kHz) is negligible compared to the 12.5 MHz bandwidth of the FFT channels obtained with the  $N = 16$  value used. Three RFI source power levels (no RFI,  $R = 0.77$ , and  $R = 2.88$ ) were selected, with approximately 4800 integration periods recorded in each case. Each experiment required approximately twelve minutes of real time due to time delays associated with transferring and recording A/D converter samples. The experiment utilized a 10-bit A/D converter (samples range from -511 to 512), and the noise standard deviation was set to approximately 100 in order to allow very low amplitude RFI to be resolved. A small number of noise samples (average of one per integration period) were found to saturate the A/D converter, but the results will show that this small number of samples apparently had a negligible impact. Variations in the white noise power level with time were also found to be insignificant for the time scales of interest, so no calibration operations were required.

The cross frequency algorithm was implemented on the dataset as described in Section II-B, using  $N = 16$ . This implies that 32,768 16 point FFT operations are performed for each 512K measurement, and the resulting powers averaged over the 32K FFT outputs. DC and Nyquist frequency bins were combined as discussed previously to produce 8 power outputs for each integration period. The maximum of these eight channels was selected; both the statistics of this quantity (the detection statistic) and the detector output as a function of the threshold were computed.

Figure 4 compares histograms of the measured cross-frequency detection statistic (symbols) with those predicted by the performance model, for the RFI free,  $R = 0.77$ ,  $R = 2.88$  cases. The horizontal axis is expressed in terms of power level standard deviations from the mean, with both the standard deviation and mean value equal to those of the RFI free case. The vertical axis represents the probability density function for the specified horizontal axis. Agreement between theory and measurement is good, indicating the accuracy of the performance model. The change in the detection statistic as the RFI power increases is apparent, along with the fact that significant changes occur even for the  $R = 0.77$  case.

Detector performance is often presented using receiver operating characteristic (ROC) curves. A ROC curve is a plot of the probability of detection ( $P_d$ ) vs. the probability of false alarm ( $P_{fa}$ ) as the threshold is changed. A good detector achieves a high probability of detection with a low false alarm rate, while a poor detector has  $P_{fa} \approx P_d$  so that the ROC curve is a 45 degree line. Figure 5 compares measured and predicted ROC curves for the experiment, using a linear scale in the left plot and a logarithmic scale in the

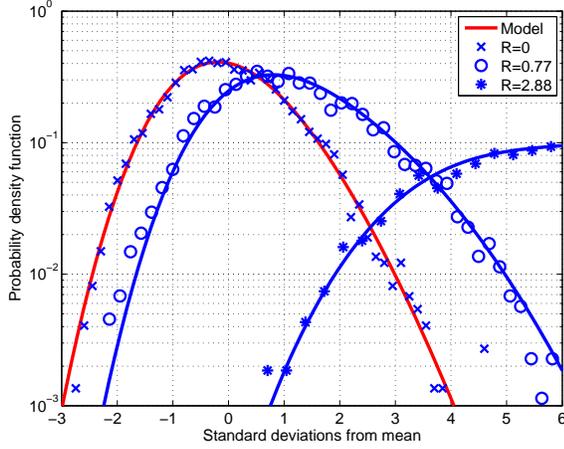


Fig. 4. Measured (symbols) and modeled probability density functions of the detector statistic for the RFI free,  $R = 0.77$ , and  $R = 2.88$  cases with  $Q = 524288$ ,  $N = 16$  (8 channels), and  $d = 1$ . Horizontal axis is expressed in terms of deviations of the statistic using the RFI free mean and standard deviation for all cases.

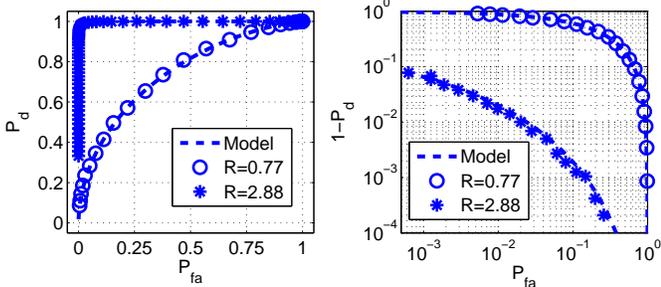


Fig. 5. Measured (symbols) and modeled ROC curves corresponding to Figure 4. Left uses linear axes while right uses logarithmic scales with one minus the probability of detection on the vertical axis.

right (one minus the probability of detection on the vertical axis). The comparison again shows a good match between model and measurement. Small differences that are present are believed to be due primarily to slight variations in the sinusoidal source amplitude as it tuned; the values of  $R$  used are averages over the recording period. The fact that the cross-frequency algorithm can provide good detection performance at low RFI power levels is apparent in these results.

### B. Comparison with other RFI detection algorithms

For  $Q = 768000$ , ROC curves of the cross-frequency, pulse, and kurtosis detectors are plotted for the case of CW interference (i.e. 100 percent duty cycle) in Figure 6 with an 8 channel radiometer (or  $N = 16$  for the pulse detector). ROC curves corresponding to  $R = 1$ ,  $R = 1.5$  and  $R = 2$  interference power levels are shown for the cross-frequency detector, and for  $R = 2$  only for the pulse and kurtosis detectors. ROC curves for the pulse and kurtosis detectors are almost a straight line indicating that these detectors are insensitive to CW RFI at this power level. In contrast, the cross-frequency detector obtains excellent performance at  $R = 2$  for CW RFI.

Following [19], a single “area under the curve” (AUC)

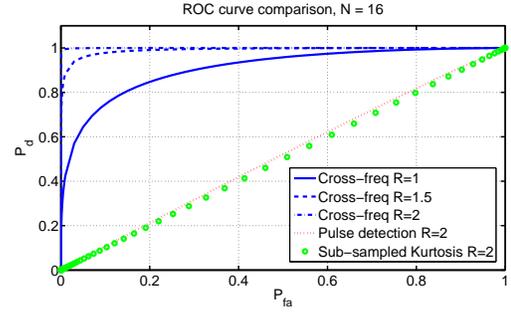


Fig. 6. ROC curves for the cross-frequency, pulse, and subsampled kurtosis algorithms: CW RFI, 8 channels ( $N = 16$  for the pulse detector),  $Q = 768000$ .  $R = 1$ ,  $R = 1.5$  and  $R = 2$  cases shown for the cross-frequency case,  $R = 2$  only for the pulse and kurtosis cases.

metric can be used to represent overall detection performance; this quantity corresponds to the area under the ROC curve. As seen in Figure 6, in the worst case the probability of detection is equal to the probability of false alarm, yielding an area under the ROC curve of 0.5. The best case produces perfect detection for any false alarm rate, yielding an area under the ROC curve of 1. The AUC is defined such that the AUC in the worst case is equal to 0 and equal to unity in the best case; this definition requires subtracting 0.5 from the true area under the ROC curve and multiplying the result by 2.

Figure 7 presents an image of AUC values for the cross-frequency detector using four frequency channels (i.e.  $N = 8$ ). The horizontal axis of the image is the ratio  $R$  that describes the RFI strength, and the vertical axis is the duty cycle of pulsed sinusoidal interference on a logarithmic scale. The logarithmic color scale is defined so that “blue” areas indicate AUC values near unity while “red” areas correspond to lower AUC values where the detector is less effective. The image makes clear that the performance of the cross-frequency algorithm depends only weakly on the RFI duty cycle. This is due to the fact that the non-centrality parameters in the cross-frequency algorithm depend primarily on the total RFI power contained within a radiometer measurement, regardless of the duty cycle. The results show that it is possible to obtain good detection performance ( $AUC \approx 0.95$  or more) for  $R \geq 2$  at all duty cycle values.

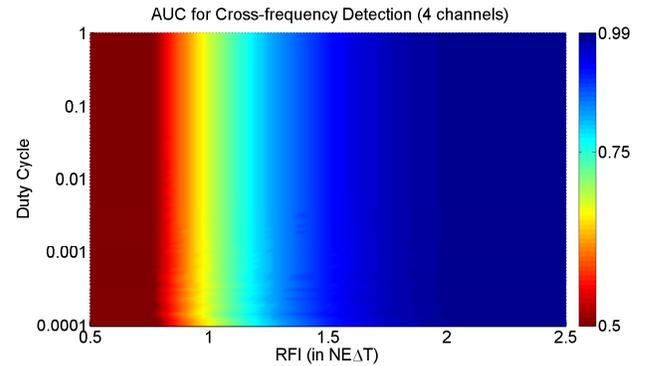


Fig. 7. AUC image for the cross-frequency detector vs. RFI strength and duty cycle, four channels

Results from the cross-frequency detector (upper) are com-

pared with those of the subsampled kurtosis (middle) and pulse (lower) detectors in Figure 8 (four frequency subchannels and four time subsamples for the kurtosis detector and an  $N = 8$  sample integration for the pulse detector.) Although the pulse and kurtosis methods provide a performance improvement for RFI with low duty cycles compared to the cross-frequency detector, they become insensitive to RFI at higher duty cycles.

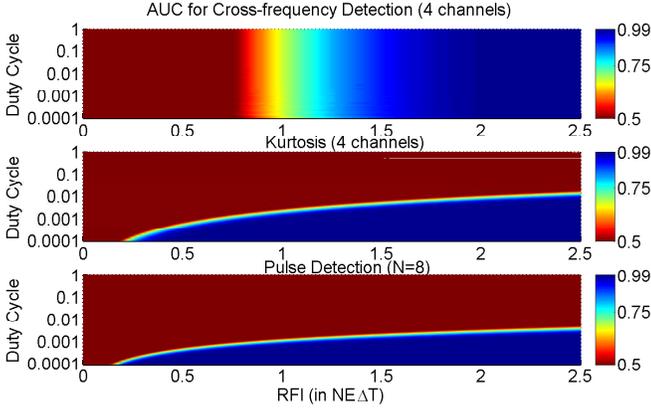


Fig. 8. AUC image comparison for the cross-frequency (upper), subsampled kurtosis (middle) and pulse (lower) algorithms vs. RFI strength and duty cycle, four channels ( $N = 8$ )

Figure 9 provides a similar AUC image comparison for a 16 channel radiometer ( $N = 32$ ). The results show improved detection performance for the cross-frequency algorithm compared to  $N = 8$ , such that good performance is achieved at  $R \geq 1$ , while the performance of the sub-sampled kurtosis method is also improved compared to  $N = 8$  but remains insensitive to higher duty cycle RFI. Similar conclusions are obtained as the number of time sub-samples in the kurtosis algorithm is varied from one to four to eight; none of these cases provide an AUC greater than 0.38 at the highest RFI power level of the image once the RFI duty cycle exceeds 10%, and none have an AUC greater than 0.005 at duty cycle 100%.

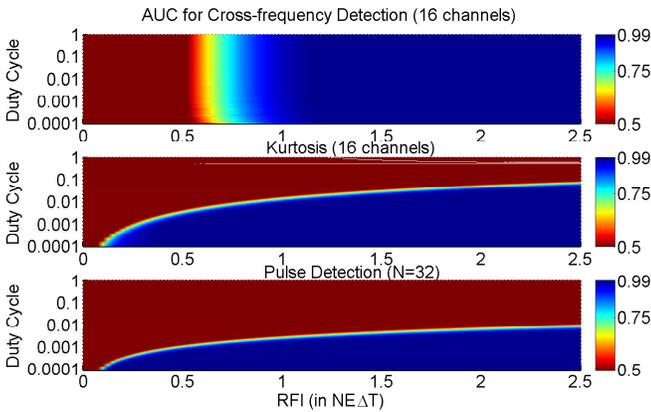


Fig. 9. AUC image comparison for the cross-frequency (upper), subsampled kurtosis (middle) and pulse (lower) algorithms vs. RFI strength and duty cycle, sixteen channels ( $N = 32$ )

The results presented to this point have used the signal model of [19], which assumes that the RFI is centered in

a single frequency channel. To quantify the impact of non-centered RFI frequencies (i.e. scalloping loss), plots of the probability of detection versus the RFI strength are provided in Figure 10 for a constant  $P_{fa}$  of 1%; note the inverted and logarithmic nature of the vertical axis in the Figure. Curves using the signal model of [19] and of Section II-A are compared for 8 and 16 frequency channels in the case of continuous sinusoidal interference. The results show that detector performance degrades when scalloping loss is included. For example, a 99 percent probability of detection is not achieved until  $R \approx 2.3$  with 16 channels when scalloping loss is considered, versus  $R \approx 1.4$  neglecting scalloping loss. These variations are significant, but the subsampled kurtosis algorithm would also be affected in a similar manner when non-centered frequencies are considered, so that the relationship between the algorithms is not expected to be significantly impacted.

For a further inspection of scalloping loss effects, an AUC image of cross-frequency algorithm results with scalloping loss is provided in Figure 11 (upper) for sixteen channels. The lower plot is a reproduction of cross-frequency algorithm results from Figure 9, which neglect scalloping loss, included here for comparison purposes. It can again be seen that scalloping loss causes a moderate degradation in performance. The RFI strength needed for good detection performance ( $AUC \approx 0.95$ ) increases from  $R \approx 1$  to  $R \approx 1.5$ . However, the behavior of the cross-frequency detector's performance versus changing duty cycle and RFI strength is very similar in both cases. In general, the cross frequency algorithm appears to be very useful for detecting pulsed sinusoidal RFI with high duty cycles, and use of such an algorithm (in addition to other approaches) in any radiometer system having multiple frequency channels is recommended.

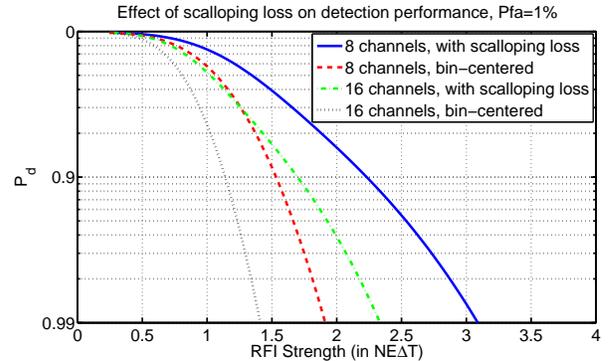


Fig. 10. Probability of detection vs. RFI strength for a constant  $P_{fa}$  of 1%, CW RFI. Results when RFI is assumed to be centered in a channel are compared with the general case of random frequency RFI for 8 and 16 frequency channels.

#### IV. SYSTEM TEMPERATURE ESTIMATION ISSUES

The previous results have assumed that a fixed threshold  $T_{cross}$  is used in the detection algorithm, corresponding to a fixed variance of the radiometer thermal noise. In reality, the variance of the radiometer observed thermal noise can vary significantly with the geophysical scene observed, so that some

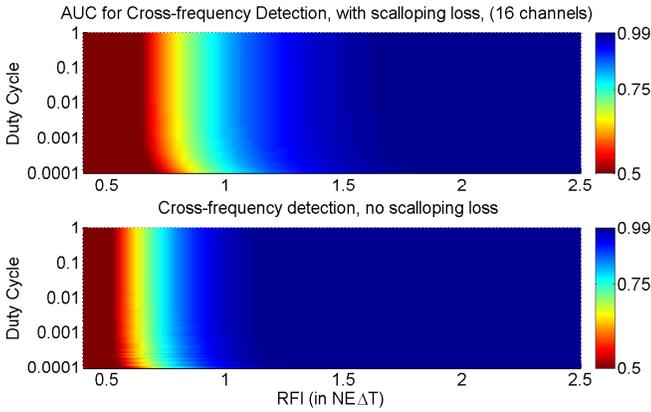


Fig. 11. AUC image comparison for the cross-frequency algorithm with scalloping loss (upper) and no scalloping loss (lower) vs. RFI strength and duty cycle, sixteen channels ( $N = 32$ )

estimate of the current thermal noise variance is required in order to maintain a known relationship between the threshold value and the detector probability of false alarm. One method for estimating this variance involves taking the mean of the observed brightnesses versus frequency, excluding a specified number of the largest brightnesses in the computation of the mean. Detection can then be declared if any of the frequency channels exceed some function of this system temperature estimate by a specified threshold.

To model the estimation process, consider a radiometer having  $M$  channels ( $M = N/2$  in the notation used) so that the cross frequency blanker considers  $M$  independent chi-squared random variables with  $2I$  degrees of freedom in the RFI free case. The estimation procedure requires sorting these measurements, so that  $M$  sorted measurements are obtained, and the largest  $M_{drop}$  values are excluded in the computation of the system temperature estimate.

One quantity of interest is the probability density function of the system temperature estimate, which is the mean of the smallest  $M - M_{drop}$  of the chi-squared random variables. This pdf in theory could be obtained using properties of order statistics, but in practice is very difficult to evaluate since the computation involves an  $M - M_{drop}$  dimensional space. To simplify the computation, it is assumed that the pdf of the system temperature estimate can be modeled as a Gaussian random variable; this is likely to be reasonable as long as  $2I$  is large. Once this approximation is applied, knowledge is required only of the mean and variance of the system temperature estimate. This mean and variance can be obtained by combining means and covariances of the lower  $M - M_{drop}$  order statistics; the required means and covariances of these order statistics can be obtained using the one and two point order statistic pdfs in [23]. It is also assumed that the pdf of the system temperature estimate so obtained is applicable both in the presence and absence of RFI; this is equivalent to assuming that any RFI lies entirely within the discarded  $M_{drop}$  largest channels when computing the system temperature estimate.

As an illustration of this procedure, Figure 12 plots the change in mean and the standard deviation of the system temperature estimate with respect to the number of channels

discarded for the 4, 8, and 16 channel cases. Here it is assumed that the mean value of  $T_{sys}$  is 590 K and its standard deviation is 0.95 K (corresponding to  $Q = 768000$  from Section II-A) so that the results shown are in units of Kelvin. The negative bias in the mean system temperature estimate is less than 1.5 K if 2 channels are discarded even for a four channel system. The dependence of the standard deviation of the system temperature estimate on the number of channels discarded is also modest, confirming that the effect of discarding channels in the system temperature estimation should be small, especially as the number of channels becomes larger.

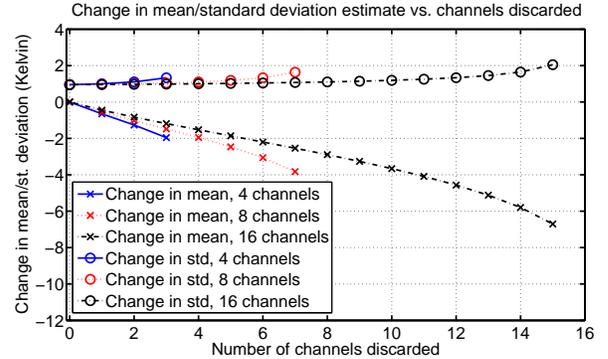


Fig. 12. Change in mean and the standard deviation of the system temperature estimate as channels are discarded in the threshold estimation procedure of the cross-frequency algorithm, for 4, 8, and 16 frequency channels.

To incorporate the system temperature estimation process into the detector performance model, the difference between the maximum brightness temperature ( $T_{MAX}$ ) and the mean of the smaller  $M - M_{drop}$  frequency channels ( $T_{MEAN}$ ) can be taken to be a single random variable and compared with a fixed threshold. However, to model this detector analytically, the joint pdf of  $T_{MAX}$  and  $T_{MEAN}$  must be known, but this quantity is difficult to compute. One possible approach involves approximating  $T_{MAX}$  and  $T_{MEAN}$  as independent, but since  $T_{MAX}$  is always greater than  $T_{MEAN}$ , this assumption is problematic.

An alternate approach creates a new random variable

$$T_N = cT_{MAX} + T_{MEAN}. \quad (7)$$

and defines  $c$  so that the correlation between  $T_N$  and  $T_{MAX}$  is zero in the RFI free case:

$$c = \frac{-cov(T_{MAX}, T_{MEAN})}{var(T_{MAX})} \quad (8)$$

Here,  $cov$  and  $var$  represent the covariance and variance operations respectively. The new detection rule is to compare the difference between  $T_{MAX}$  and  $T_N$  to a threshold, so that the probability of detection is

$$P(T_{MAX} - T_N > \Delta T_{CROSS}) \quad (9)$$

where  $P$  denotes the probability and  $\Delta T_{CROSS}$  is a fixed threshold.  $T_{MAX}$  and  $T_N$  are now assumed to be independent, so that the pdf of  $T_{MAX} - T_N$  can be obtained analytically. Note that while  $T_N$  and  $T_{MAX}$  are uncorrelated, they are

clearly not independent, however the assumption that  $T_N$  and  $T_{MAX}$  are independent appears to be less restrictive than assuming that  $T_{MEAN}$  and  $T_{MAX}$  are independent.

Using the assumption of independence, the pdf of the difference random variable can be evaluated as a convolution between the original pdf of the maximum brightness temperature and a scaled version of the Gaussian pdf of the system temperature estimate. The resulting random variable is then compared with a scaled version of the original threshold.

To assess these approximations, Figure 13 compares ROC curves obtained with this method to Monte Carlo simulations that include system temperature estimation effects without approximation. CW RFI having  $R = 2$  is considered, and the estimation procedure discards the 2 channels with the highest brightnesses. Comparisons are shown for 4, 8, and 16 channels, and 16384 realizations were used in the Monte Carlo simulation. The comparisons show that the accuracy of the approximation improves as the number of channels increases, and that the approximation is reasonable for 8 channels and very good for 16 channels. Other simulations also confirmed that results from this procedure can be used without a significant loss in accuracy if the number of frequency channels is greater than or equal to 8.

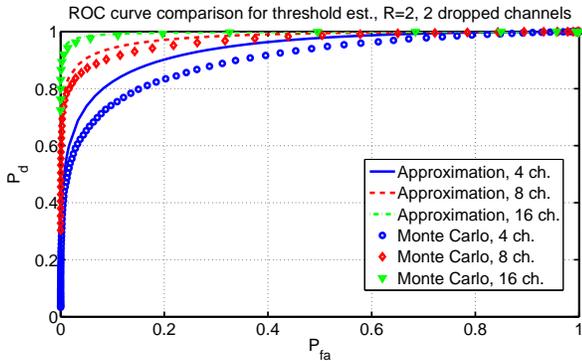


Fig. 13. ROC curves obtained using the analytical approximation to threshold estimation effects compared with results from Monte Carlo simulations. CW RFI having  $R = 2$ ,  $M_{drop} = 2$ , and for 4, 8 and 16 channels.

To show the effect of the system temperature estimation on detection performance, the probability of detection versus CW RFI strength is plotted in Figure 14 for  $P_{fa} = 1\%$  when the system temperature is estimated with 2 channels discarded in the 8 and 16 channel cases using the approximate model. For comparison, results from Figure 10 when the system temperature is exactly known are also included. These results confirm that estimation of the system temperature causes only a modest degradation in detection performance for high duty cycle sinusoidal interference.

## V. CONCLUSIONS

The performance of a cross-frequency detection algorithm was analyzed for pulsed sinusoidal RFI and compared with that achieved by the pulse and sub-sampled kurtosis algorithms. It was shown that the performance of the cross-frequency detector improves as the number of frequency channels increases, and that its performance is only weakly

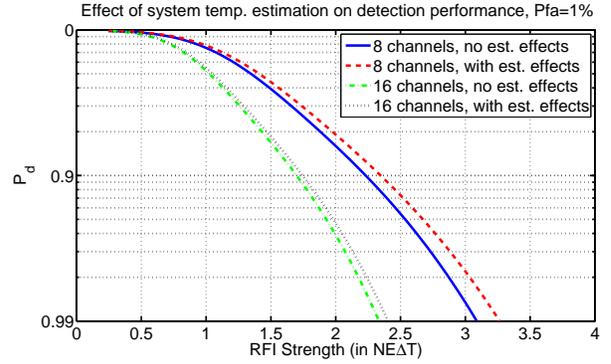


Fig. 14. Results of Figure 10 are compared with curves including system temperature estimation effects. CW RFI

sensitive to the RFI duty cycle at a fixed  $R$  RFI power level. A reasonable sensitivity to pulsed sinusoidal RFI is achieved even with only four radiometer channels. The performance of the pulse and sub-sampled kurtosis methods was found to exceed that of the cross-frequency algorithm for low duty cycle pulsed RFI, but the cross-frequency method offers significant performance improvements over these methods for higher duty cycle pulses including CW interference. A method for modeling the impact of system temperature estimation effects on detector performance was also provided, and it was shown that system temperature estimation is expected to result in only a modest degradation for pulsed sinusoidal RFI sources.

It should be noted that, since pulsed sinusoidal RFI sources are localized in frequency by definition, it is not surprising that the cross-frequency detector performs very well against this class of RFI. Because this class of RFI, especially the CW case which represents any very narrowband emission, is expected to be encountered in Earth observations, use of cross-frequency detection algorithms appears to be warranted in future Earth observing radiometer systems. Such use is compatible with any other algorithms such as the sub-sampled kurtosis or pulse methods that can provide enhanced performance for low duty cycle pulses and/or other RFI source types.

The RFI detection performance that is ultimately achieved will be a function of the threshold utilized (which is a tradeoff between the probability of false alarm and the probability of detection), the nature of the RFI encountered, and radiometer system parameters such as the number of frequency channels and the integration period. The examples shown were intended to be representative of potential future satellite system configurations, and indicate that a cross-frequency approach can provide reasonable detection of continuous sinusoidal interference at power levels of a few times the radiometer uncertainty. The implications of such performance for science measurements will depend on the geophysical quantity of interest. For high sensitivity measurements such as soil moisture at L-band, this performance is likely acceptable, while for low sensitivity applications such as ocean-salinity at L-band, additional performance gains (which could be obtained, for example, by using a larger number of frequency channels or by accepting a larger number of false alarms) may be necessary.

## APPENDIX I

## PROPERTIES OF FFT OUTPUTS IN THE ABSENCE OF RFI

Assuming no RFI is present, equation (1) becomes

$$x_i[n] = w_i[n] \quad (10)$$

where  $w_i[n]$  is a Gaussian random variable with zero mean and variance of  $\sigma^2$ . The FFT of  $x_i[n]$  is

$$\begin{aligned} X_i[k] &= \sum_{n=1}^N x_i[n] e^{-j2\pi nk/N} \\ &= \sum_{n=1}^N x_i[n] \cos(2\pi nk/N) - j \sum_{n=1}^N x_i[n] \sin(2\pi nk/N) \end{aligned} \quad (11)$$

$$= A_i[k] - jB_i[k] \quad (12)$$

The real and imaginary parts of  $X_i[k]$  ( $A_i[k]$  and  $B_i[k]$ ) are sums of zero mean Gaussian random variables, and are therefore also zero-mean Gaussians with variance:

$$\begin{aligned} \text{var}(A_i[k]) &= \sum_{n=1}^N \cos^2(2\pi nk/N) \sigma^2 \\ &= \begin{cases} N\sigma^2 & \text{if } k = 0 \text{ or } N/2 \\ N\sigma^2/2 & \text{otherwise} \end{cases} \\ \text{var}(B_i[k]) &= \sum_{n=1}^N \sin^2(2\pi nk/N) \sigma^2 \\ &= N\sigma^2/2 \quad k \neq 0, N/2 \end{aligned}$$

Note  $B_i[k]$  vanishes for  $k = 0$  or  $N/2$ . Since  $A_i[k]$  and  $B_i[k]$  are both Gaussian, they can be shown to be independent because their covariance vanishes.

Computing the power in bin  $k$  gives

$$|X_i[k]|^2 = (A_i[k])^2 + (B_i[k])^2 \quad (13)$$

which, for  $k$  not 0 or  $N/2$ , is a sum of the squares of two zero-mean Gaussian random variables each having variance  $N\sigma^2/2$ . This implies that

$$Z_i[k] = \frac{2}{N\sigma^2} |X_i[k]|^2 \quad (14)$$

for  $k$  not 0 or  $N/2$  is a standard chi-squared random variable with two degrees of freedom. When  $k = 0$  or  $N/2$ , we have

$$|X_i[k]|^2 = (A_i[k])^2 \quad (15)$$

where  $A_i[k]$  has variance  $N\sigma^2$ ; this is a scaled chi-squared random variable with one degree of freedom.

While some references [22] neglect  $k = 0$  and  $N/2$  since they are different statistically from the other  $k$  values, it is desirable to include information from these bins in order to retain sensitivity to RFI across the entire radiometer bandwidth. Properties of the FFT show that these bins in fact contain information from a bandwidth half as large as those contained in  $k = 1$  to  $N/2 - 1$ . For example, the  $k = 0$  bin in some sense includes power at frequencies slightly above and slightly below zero frequency, but information below zero frequency is simply the conjugate of that above zero frequency. In addition, it can be shown that  $X_i[0]$  and  $X_i[N/2]$  are independent.

Combining these ‘‘half channels’’ to create a ‘‘full channel’’ that is a chi-squared random variable with two degrees of freedom analogous those at other  $k$  values is therefore possible. It is necessary to average the two in order to account for the fact that the variance of  $A_i[k]$  is  $N\sigma^2$  when  $k = 0$  or  $N/2$  instead of  $N\sigma^2/2$  for other  $k$  values. Defining

$$Z_i[N/2] = \frac{2}{N\sigma^2} \left[ \frac{1}{2} (|X_i[0]|^2 + |X_i[N/2]|^2) \right] \quad (16)$$

results in  $Z_i[k]$  for  $k = 1$  to  $N/2$  representing the  $N/2$  frequency channels, with each a standard chi-squared random variable with two degrees of freedom.

## APPENDIX II

## DERIVATION OF THE PROBABILITY OF FALSE ALARM

In Appendix I, it was shown that after FFT operation and the averaging of the  $k = 0$  and  $N/2$  bins,  $N/2$  chi-squared random variables with 2 degrees of freedom are obtained scaled by  $N\sigma^2/2$ . When these variables are averaged over  $I$  frames, each is a chi-squared random variable with  $2I$  degrees of freedom scaled by  $N\sigma^2/(2I)$ . If the cumulative distribution function of a chi-squared variable with  $2I$  degrees of freedom is denoted by  $CDF_{2I}$ , the probability of false alarm for the cross-frequency case at a threshold value of  $T_{\text{cross}}$  can be written as

$$P_{fa} = 1 - \left[ CDF_{2I} \left( \frac{T_{\text{cross}} 2I}{NT_{sys}} \right) \right]^{N/2} \quad (17)$$

where the fact that the powers in each bin  $k$  are independent allows the CDF's for each channel to be multiplied when determining the CDF of the maximum.

## APPENDIX III

## CALCULATION OF THE NON-CENTRALITY PARAMETER AND THE PROBABILITY OF DETECTION

When pulsed sinusoidal RFI is present, the samples  $x_i[n]$  are given by equation (1). For this type of RFI, the real and imaginary parts of the FFT operation given in equation (11) are still normally distributed, but have non zero means  $\mu$  given by

$$\begin{aligned} \mu_{A,i}[k] &= \sum_{n=1}^N \{ A \cos(2\pi f_0 [(i - i_0)N + n] + \phi) \\ &\quad \times \mathcal{I}(n, i) \cos(2\pi nk/N) \} \\ \mu_{B,i}[k] &= \sum_{n=1}^N \{ A \cos(2\pi f_0 [(i - i_0)N + n] + \phi) \\ &\quad \times \mathcal{I}(n, i) \sin(2\pi nk/N) \} \end{aligned}$$

Note that mean values can be obtained in all bins  $k$  even for a sinusoidal source due to scalloping, partial frame filling, or pulsed source effects.

$A_i[k]$  and  $B_i[k]$  retain the same variances as in the RFI free case, and remain independent as well. The quantities  $Z_i[k]$  are now non-central chi-squared random variables with two degrees of freedom, with the non-centrality parameter given by  $\frac{2(\mu_{A,i}[k]^2 + \mu_{B,i}[k]^2)}{N\sigma^2}$ . When these quantities are averaged over

$I$  frames, the non-centrality parameters are averaged over their values for each frame as well to produce  $\lambda_k$ , and again the final powers in each bin  $k$  remain independent. Similar steps as in the RFI free case are applied to combine the  $k = 0$  and  $N/2$  bins in a manner that is consistent with the other bins. Multiplying the CDF's for each frequency bin to determine the CDF of the maximum, the resulting probability of detection for a threshold value of  $T_{\text{cross}}$  is

$$P_d = 1 - \prod_{k=1}^{N/2} CDF_{2I, \lambda_k} \left( \frac{T_{\text{cross}} 2I}{N T_{\text{sys}}} \right) \quad (18)$$

where  $CDF_{2I, \lambda}$  denotes the cumulative distribution function of a non-central chi-squared random variable with  $2I$  degrees of freedom and non-centrality parameter  $\lambda_k$ .

Since the RFI source frequency, phase, arrival frame, and arrival sample are assumed to be uniformly distributed, the probability of detection obtained with equation (18) is averaged over all the possible values of these parameters and associated  $\lambda_k$ 's.

#### REFERENCES

- [1] L. Li, E. G. Njoku, E. Im, P. S. Chang, and K. M. St.Germain, "A preliminary survey of radio-frequency interference over the US in Aqua AMSR-E data," *IEEE Trans. Geosc. Rem. Sens.*, vol. 42, pp. 380-390, 2004.
- [2] S. W. Ellingson and J. T. Johnson, "A polarimetric survey of radio frequency interference in C- and X-bands in the continental United States using WindSAT radiometry," *IEEE Trans. Geosc. Rem. Sens.*, vol. 44, pp. 540-548, 2006.
- [3] J. T. Johnson, A. J. Gasiewski, B. Güner, G. A. Hampson, S. W. Ellingson, R. Krishnamachari, N. Niamsuwan, E. McIntyre, M. Klein, and V. Y. Leuski, "Airborne radio-frequency interference studies at C-band using a digital receiver," *IEEE Trans. Geosc. Rem. Sens.*, vol. 44, pp. 1974-1985, 2006.
- [4] D. M. LeVine, "ESTAR experience with RFI at L-band and implications for future passive microwave remote sensing from space" *IEEE Geoscience and Remote Sensing Symposium*, conference proceedings, vol. 3, pp. 847-849, 2002.
- [5] A. Camps, I. Corbella, F. Torres, J. Bara and J. Capdevila, "RF interference analysis in aperture synthesis interferometric radiometers: application to L-band MIRAS instrument," *IEEE Trans. Geosc. Rem. Sens.*, vol. 38, pp. 942-950, 2000.
- [6] A. J. Gasiewski, M. Klein, A. Yevgrafov, and V. Leuski, "Interference mitigation in passive microwave radiometry," *IEEE Geoscience and Remote Sensing Symposium*, conference proceedings, vol. 3, pp. 1682-1684, 2002.
- [7] M. Basseville and I. V. Nikiforov, *Detection of abrupt changes: theory and applications*. Englewood Cliffs, NJ: Prentice-Hall, 1993.
- [8] P. A. Fridman and W. A. Baan, "RFI mitigation methods in radio astronomy," *Astronomy&Astrophysics*, vol. 378, pp. 327-344, 2001.
- [9] G. A. Hampson, S. W. Ellingson, and J. T. Johnson, "Design and demonstration of an interference suppressing microwave radiometer," *IEEE Aerospace Conference*, conference proceedings, vol.2, pp. 993-999, 2004
- [10] N. Niamsuwan, J. T. Johnson, and S. W. Ellingson, "Examination of a simple pulse blanking technique for RFI mitigation" *Radio Science*, vol. 40, June 2005.
- [11] B. Güner, J. T. Johnson, and N. Niamsuwan, "Time and frequency blanking for radio frequency interference mitigation in microwave radiometry," *IEEE Trans. Geosc. Rem. Sens.*, vol. 45, pp. 3672-3679, 2007.
- [12] C. S. Ruf, S. M. Gross, and S. Misra, "RFI detection and mitigation for microwave radiometry with an agile digital detector," *IEEE Trans. Geosc. Rem. Sens.*, vol. 44, pp. 694-706, 2006.
- [13] S. Misra, S. S. Kristensen, S. S. Sobjaerg, and N. Skou, "CoSMOS: Performance of kurtosis algorithm for radio frequency interference detection and mitigation," *IGARSS 2007*, conference proceedings, 2007.
- [14] R. D. De Roo, S. Misra, and C. S. Ruf, "Sensitivity of the kurtosis statistics as a detector of pulsed sinusoidal RFI," *IEEE Trans. Geosc. Rem. Sens.*, vol. 45, pp. 1938-1946, 2007.
- [15] J. R. Piepmeier, P. Mohammed, and J. Knuble, "A Double Detector for RFI Mitigation in Microwave Radiometers," *IEEE Trans. Geosc. Rem. Sens.*, vol. 46, pp. 458-465, 2008.
- [16] B. Güner, M. Frankford and J. T. Johnson, "Comparison of three omnibus normality tests for the detection of pulsed sinusoidal RFI sources," *URSI General Assembly*, conference proceedings, 2008.
- [17] B. Güner, M. T. Frankford and J. T. Johnson, "A study of the Shapiro-Wilk test for the detection of pulsed sinusoidal radio frequency interference," *IEEE Trans. Geosc. Rem. Sens.*, vol. 47, pp. 1745-1751, 2009.
- [18] S. Misra and C. Ruf, "Comparison of pulsed sinusoidal RFI detection algorithms using time and frequency sub-sampling" *IGARSS 2008*, conference proceedings, 2008.
- [19] S. Misra, P. N. Mohammed, B. Güner, C. Ruf, J. R. Peipmeier and J. T. Johnson, "Microwave radiometer radio frequency interference detection and mitigation algorithms: a comparative study," accepted by *IEEE Trans. Geosc. Rem. Sens.*, 2009.
- [20] J. T. Johnson and L. C. Potter, "A study of algorithms for detecting pulsed sinusoidal interference in microwave radiometry," *IEEE Trans. Geosc. Rem. Sens.*, vol. 47, pp. 628-636, 2009.
- [21] P. N. Mohammed, J. Knuble and J. Piepmeier, "Pulse and kurtosis detection of radio-frequency interference (RFI): an experimental comparison," *IGARSS 2008*, conference proceedings, 2008.
- [22] S. M. Kay, *Fundamentals of Statistical Signal Processing: Volume II, Detection Theory*, Upper Saddle Creek, NJ: Prentice Hall, 1998.
- [23] M. G. Kendall and A. Stuart, *The advanced theory of statistics*, Volume 1, Chapter 14, New York: Hafner, 1969.
- [24] M. Abramowitz and I. A. Stegun, *Handbook of mathematical functions*, Chapter 26, New York: Dover, 1965.