IMPULSE DETECTION AND REJECTION METHODS FOR RADIO SYSTEMS

Harri Saarnisaari*
Cente for Wireless Communications
Oulu, Finland
and
Pertti Henttu
Centre for Wireless Communications
Oulu, Finland

ABSTRACT

Impulses that exist in the nature or are artificial may severely decrease the performance of any radio system including communication and interception receivers and radars. This paper investigates impulse detection in receivers that can have either a single antenna or an antenna array. The proposed detector uses either a backward or forward outlier search strategy and assumes zero mean signals and spatially and temporally white background noise. Signal need not to be weak, i.e., signal-to-noise ratio may be large and the methods still works. The backward method uses all the data in an initial step whereas the forward method uses smaller, (hopefully) clean set in an initial step. The forward method is expected to perform better. Simulations confirm this expectation. Indeed, the used forward method can detect outliers even if their contamination exceeds 90 % of all the samples.

INTRODUCTION

Outliers are observations that appear inconsistent with the remainder of the data, which follows the assumed distribution. Strong outliers prevent or mitigate capability to make correct decisions from the observed data. In radio systems outliers may be impulses from the nature such as those generated by lightning. They can also be man made such as signals from impulse radars or power lines [1]. Also, a fast sweeping tone signal, often used for jamming purposes, cause impulse like signal in a receiver. It is clear that impulses may continue one or few samples (short impulses) or may continue for several tens of samples (long impulses) in the receiver.

Several methods have been suggested to remove the impact of the outliers. Those can be divided into robust signal processing and diagnostic methods. In the diagnostic methods, outliers are first detected and then removed. The decisions are made from the rest of the data. In the robust approaches, decisions and outlier treatment is made at the same time.

Robust signal processing is mostly based on Huber's work [2], who developed the so called M-estimator, but also non-Gaussian noise models are used [3, 4, 5]. M-estimation based detection for direct sequence spread spectrum (DS-SS) systems is considered in [6, 7, 8]. Therein, it was observed that a small number of strong enough outliers significantly deteriorates the performance of detectors based on Gaussian statistics. The used robust detectors work well. In [8], it was shown that an asymptotic breakdown point of the single user detector is 50 % for single user case, if signal-to-noise ratio (SNR) per chip is large ($\gg$ 0 dB), and decreases if SNR decreases or number of users increases. The breakdown point determines maximum contamination rate at which the estimator or detector still does its job properly. In general, the breakdown point of the M-estimators is said to be not larger than 30 % [9].

The diagnostic methods are typically treated in the statistical literature, see, e.g., [10] and references therein. The diagnostic methods can be divided into two blocks. The first block uses sequential detection (backward search) and the second one sequential addition (forward search). In the backward search methods, the entire set of observations is initially considered and outliers are sequentially removed. In the forward search methods, a small suitably chosen set is initially used to set outlier detection parameters. The rest data is then treated sequentially. Typically, the initial set must be clean, free of outliers and this may be the disadvantage of the forward methods. It is said, that generally the forward search methods outperform the backward search methods [10].

In [11], a backward search method is presented. It is used to mitigate narrowband interference in single antenna DS-SS systems. The method uses assumption that observations are zero mean, independent Gaussian distributed variables and

*The research was supported by Nokia, Elektrobit, Finnish Navy, Air Force and Defence Forces Techical Research Centre, the National Technology Angency (Tekes), and Infotech Oulu Graduate School

0-7803-8140-8/03/$17.00 (C) 2003 IEEE.

computes the threshold according to this assumption, as will be discussed later. The algorithm is called the consecutive mean excision (CME) algorithm. In this paper this algorithm is applied for impulse detection and further developed in the sense that the forward version is presented. The receiver is also allowed to contain an antenna array. The CME algorithm is a special case of those used in the statistical literature in the sense that it assumes zero mean data while the general methods do not assume this.

After detection of the impulses, the corrupted samples has to be treated somehow. One method is to limit them, as is often done in the robust signal processing [12]. Another method is to reject (to zero) the corrupted samples. It is not investigated here what is the difference between limiting and rejecting to the overall system performance if impulses are present. This is an interesting future research topic. It is known that with the limiting the clean sample rejection rate can be rather high (10–15 %) and the overall system still performs well [13] in the impulse free case. Transform domain narrowband interference mitigation was investigated in [13], but the conclusions hold also for the impulse mitigation. In this paper we propose methods that reject only a few samples (1 % or less) in the impulse free case, but still offer a very good impulse detection capability. In the impulse free case the performance deterioration due to rejection is small since only a small number of samples is rejected. Especially this hold in DS-SS systems, where few chips can be zeroed without dramatic effects on performance. If symbols are zeroed instead of chips, then the channel coding has to take this into account.

The paper is organized as follows. First, the signal model is defined. Then the diagnostic outlier detection methods are reviewed. After that, the CME algorithm is given and expanded. Impulse detection capability is then verified by simulations. The results of the backward and forward CME algorithms are compared to those of the conventionally used one-shot impulse detector. Finally, conclusions are drawn.

SIGNAL MODEL

The baseband presentation of the received signal $x_i \in \mathbb{C}^n$ at time instant i in a receiver that contains an n-element antenna array is

$$x_i = a(\theta_s)s_i + a(\theta_i)o_i(\epsilon) + n_i, \quad i = 1, \dots, N \qquad (1)$$

where $a(\theta) \in \mathbb{C}^n$ is the steering vector, $s_i \in \mathbb{C}$ the desired signal, $o_i(\epsilon) \in \mathbb{C}$ an impulse that occurs with probability ϵ and $n_i \in \mathbb{C}^n$ presents system and background noise. N denotes the number of observed samples. Herein, a single signal model is selected for simplicity but as well a multiple signal model can be used.

The desired signal can be any communication or radar signal. Noise is modeled, as typically, to be spatially and temporally white, zero mean multivariate Gaussian process although the presented general outlier detectors can be used even this assumption is not valid. Impulses are either zero mean Gaussian variables with variance σ_o^2 or constant envelope signals. The latter are modeled as Gaussian variables which have a constant magnitude mean $\bar{o}$ and very small variance. The former impulse model leads to commonly in impulsive channels used mixed Gaussian model [6] whereas the latter models sweeping signals or radar pulses.

SNR is defined to be $|s_i|^2/\sigma^2$, i.e., it defines SNR at the input of a single antenna. Impulse-to-signal ratio (ISR) is defined either as $\sigma_o^2/|s_i|^2$ or $|\bar{o}|^2/|s_i|^2$. The former is used in the zero mean impulses case and the latter in the other case.

DIAGNOSTIC METHODS

The diagnostic methods often reject observations based on the Mahalanobis distance (MD) [9]. This means that x_i is decided to contain outliers if the Mahalanobis distance

$$\chi_i = \sqrt{(x_i - \bar{x})^{\mathrm{H}} C^{-1} (x_i - \bar{x})} \qquad (2)$$

is large. Here $\bar{x}$ and C denote the mean and covariance of x_i or their estimates.

There are two problems with χ_i [14]. First, outliers do not necessarily have large values for χ_i, since they may be affected to estimates of $\bar{x}$ and C. Second, all observations with large χ_i are not necessarily outliers, and valid observations will be removed. The first problem is known as masking because the presence of one outlier masks the appearance of another outlier. The second problem is known as swamping. To overcome these problems robust estimators of location $\bar{x}$ and covariance matrix C could be used. Several robust estimators have been suggested including the M-estimator, the least median of squares and the least trimmed squares [10].

The backward methods initially use all the observations $x_i \in \mathbb{C}^n, i = 1, \dots, N$ to determine the mean and covariance. After removing the first detected outliers new estimates of the mean and covariance are computed using the cleaner data. This process is repeated until outliers cannot be found anymore. In the forward methods, an initial clean set is selected and the set size is then increased until clean observations cannot be found. Some possibilities to select the initial set are discussed in [10, 14]. Basically, the initial robust estimates of the mean and covariance are used to rearrange observations in ascending order according to initial Mahalanobis distances χ_i [14]. Also other alternatives exist [15].

It is interesting to note that the Mahalanobis distance was used in univariate case ($n = 1$) already in 1935 [16]. Therein, the distribution of $(x_i - \bar{x})/\sqrt{C}$ is derived under assumption that x_i are Gaussian and that the location and covariance are replaced by their standard estimates. The resulting distribution is used to obtain a outlier detection threshold. This threshold depends on data size N. The outlier detection threshold can also be computed using knowledge that if the location and covariance are known, then χ_i^2 is central chi-square distributed with variance of individual elements as $1/2$ and $2n$ degrees of freedom (x_i are complex) [17]. 'Good' samples are rejected if outliers are not present. The number of rejected 'good' samples depends on the threshold. The threshold can be set in such a way that, e.g., 1 %, 5 % or 10 % of all the samples are rejected in the case that there do not exist any outliers. This is called the clean sample rejection rate.

CME

It is possible to use the above mentioned methods for the outlier detection. In this paper, however, a different approach is taken. The proposed approach takes into account the fact that many signals s_i are zero mean although signal-to-noise ratio (SNR) is high. This means that the mean cannot be estimated. It is therefore assumed that signal is zero mean. It is also assumed that noise is spatially and temporally white, such that only variance of the received signal has to be estimated. In the weak signal case zero mean assumption holds, at least approximately, also instantly.

Let $D_i = \|x_i\|^2$. Variable D_i is chi-square distributed with variance of individual elements as $\frac{1}{2}\sigma^2$ and $2n$ degrees of freedom. In the noise only case $\mathrm{E}\{D_i\} = n\sigma^2$. Probability that D_i exceeds $T\,\mathrm{E}\{D_i\} = Tn\sigma^2$ is [17]

$$P(D_i > T\,\mathrm{E}\{D_i\}) = e^{-Tn/2} \sum_{k=0}^{n-1} \frac{1}{k!}\left(\frac{T}{2}\right)^k, \qquad (3)$$

that does not depend on the variance. The outlier detection threshold can be determined using (3). Another information that is needed to understand the CME algorithm is the fact, that the average of squared samples converges to the true value, i.e., $\frac{1}{N}\sum_{i=1}^{N}\|x_i\|^2 \to \mathrm{E}\{\|x_i\|^2\} = \mathrm{E}\{D_i\}$. The CME algorithm reject samples that are larger than T times the mean $\frac{1}{N}\sum_{i=1}^{N}\|x_i\|^2$.

The CME algorithm finds outliers iteratively as follows:

1. Compute $z = \frac{1}{N}\sum_{i=1}^{N}\|x_i\|^2$.
2. Reject values that exceed Tz, keep rest.
3. Iterate steps 1-2 for the kept values until nothing is rejected or the maximum number of iterations is achieved.

Since the CME algorithm uses all the values in the starting point it is a backward search method. Especially in the one antenna case absolute values can be used instead of squares since this alleviates requirements in the dynamic range if the following approximation is used. In this case $\|x_i\|$ is Rayleigh distributed [17] and the outlier detection threshold should be appointed using it. The required square root, that is a computationally demanding operation, can be approximated in practical systems as $\sqrt{x} = \max\{\mathrm{Re}\,x, \mathrm{Im}\,x\} + 0.4\min\{\mathrm{Re}\,x, \mathrm{Im}\,x\}$ [18]. This approximation was proposed in [19], wherein the notch filter was implemented in a single chip.

The forward CME (FCME) is obtained by rearranging samples x_i according to $\|x_i\|^2$ in the ascending order. Let this set be denoted as the set $\{y_i\}_{i=1}^{N}$. Select the first (smallest in energy) m terms to form the 'clean' set. The size of the set can be one element, 5 % or maybe 10 % of set size. The larger the initial set is, the simpler the algorithm is. On the other hand, the larger the initial set is, the higher is the possibility that the initial set contains outliers and methods performance decreases.

The FCME algorithm finds outliers iteratively as follows:

1. Compute $z_m = \sum_{i=1}^{m}\|y_i\|^2$.
2. If y_{m+1} is smaller than the computed threshold Tz_m, increase m by one and goto 1. Otherwise, finish the algorithm.

The FCME algorithm requires sorting. Sorting algorithms are discussed in [20]. For example, Heapsort routine has complexity of order $N\log_2 N$ and its worst case performance is only 20 % worse than the average complexity. Quicksort is at best somewhat faster, but its worst case is of order N^2.

In order to compare the new methods to the older ones, the CME and FCME algorithms are compared to the conventionally used one-shot algorithm, where only the first iteration of the CME algorithm is used.

SIMULATIONS

Outlier detection capability is demonstrated by simulations. One and four antenna cases are considered. Number of observations N is 255. The signal is a complex signal formed by two 255-chip m-sequences. Cutting thresholds are selected according to 1 % clean sample rejection rate. One thousand repeats have been made in order to compute the impulse detection rate and the total detection rate. The latter describes how many samples exceed the threshold. The former tells how many corrupted samples exceed the threshold. The total detection rate is an important factor, since it is reasonable that the methods do not detect too many clean samples to contain impulses. Ten iterations are used in the CME in order to guarantee that it has been converged. 10 samples was used to form the clean set in the FCME algorithm, i.e., 4 % of all the samples.

In Fig. 1 the constant amplitude impulses case is studied in the single antenna case. The impulse detection rate of the methods is investigated with different SNR values as a function of ISR. The shown number of impulses is the highes one with methods attain 100 % detection rate at some ISR value. Results in the similar case for the random impulses are presented in Fig. 2.

It can be observed from Fig. 1 and 2 that the constant envelope impulses are easier to detect than the random ones. Especially the FCME algorithm is marvelous in the constant envelope impulses case. It can detect all the impulses even if their proportion is 90 % of all the samples providing that the impulses are strong enough. Strong enough means that impulse power is 12-15 dB above the 'noise' floor, where 'noise' floor is the level of the signal or noise, whichever is stronger. The CME algorithm can handle the constant envelope impulses if their proportion is 20 % or less and the one-shot method if the proportion is 10 % or less. In the random impulses case the one-shot method can not detect all the impulses at any proportion. The CME method can find all the impulses if their proportion is at maximum 40 % and the FCME method if the proportion is so large than 90 %. The required impulse-to-'noise' ratio is 25-30 dB that is 10 dB more than in the constant impulses case. It should, however, keep in mind that in the random impulses case all the impulses are not necessarily very strong such that they do not affect to the performance of a system even though

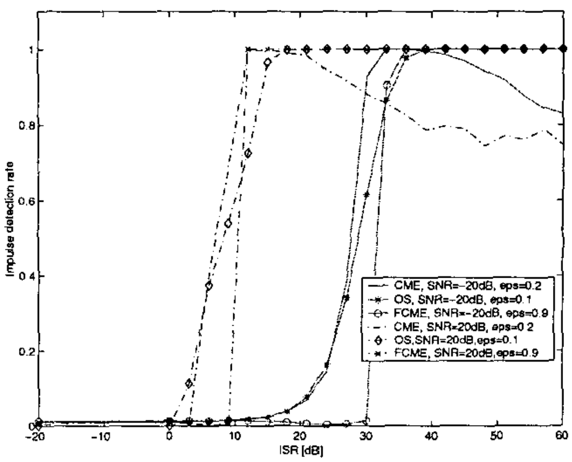

Figure 1. Impulse detection rate of the CME, one-shot (OS) and FCME algorithms as a function of ISR in the one antenna and constant impulses case for few SNR values and maximum number of impulses (eps) with the methods still obtain 100 % detection rate.

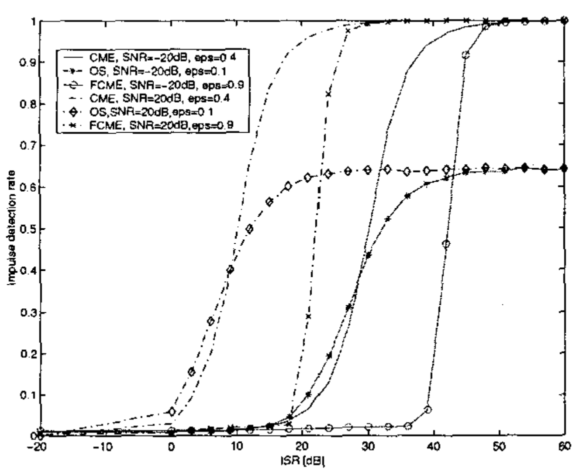

Figure 2. Impulse detection rate of the CME, one-shot (OS) and FCME algorithms as a function of ISR in the one antenna and zero mean impulses case for few SNR values and maximum number of impulses (eps) with the methods still obtain 100 % detection rate.

Similar set of results for the four antennae case are presented in Fig. 3.

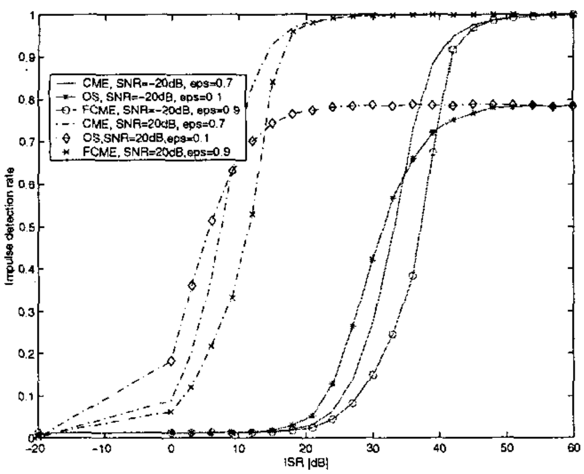

Figure 3. Impulse detection rate of the CME, one-shot (OS) and FCME algorithms as a function of ISR in the four antennae and zero mean impulses case for few SNR values and maximum number of impulses (eps) with the methods still obtain 100 % detection rate.

they are not detected and rejected. If the receiver contains a four element antenna array the impulse detection is somewhat easier than in the single antenna case, compare Fig. 2 and 3. It can also be seen from Fig. 1-3 that at low ISR value (-20 dB), the methods decide 1 % of samples to contain impulses that is the desired detection rate at the impulse free case.

The results in Fig. 4 shows the impulse detection and total sample detection rates as a function of the proportion of impulses. Random impulses case is considered in the one antenna case with SNR as -20 dB and ISR as 54 dB that guarantees the impulse detection rate be 100 % at the best situation. It can be seen that the CME methods do not detect too many samples to contain impulses that is an agreeable property.

From the above results it is clear that strong enough impulses are perfectly detected. But, the question is that can harmfull impulses detected? In DS-SS systems impulses are approximately harfull if *proportion of impulses* $\times$ ISR *-processing gain* > -6 dB. If the processing gain is 24 dB (as in the numerical examples), then at the proportion 10 %, the harmfull impulses should have ISR value 28 dB and at the proportion 50 % and 90 % they should have ISR value 21 dB and 18.5

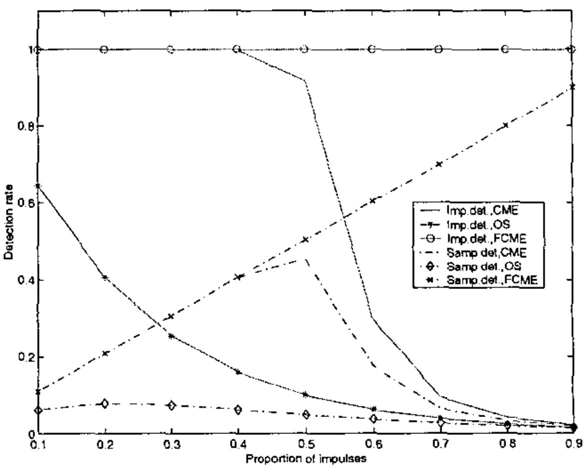

Figure 4. Impulse detection and total sample detection rates of the CME, one-shot (OS) and FCME algorithms as a function of the proportion of impulses in the one antenna case at SNR= -20 dB and ISR= 54 dB.

dB, respectively. Since the impulses with a constant envelope are perfectly detected if they exceed the 'noise' floor by 15 dB, the harmful impulses are perfectly detected by the FCME method if SNR is -3.5 dB or larger. At the lowest proportion, SNR has to be -13 dB or larger.

In the random impulses case harmful impulses can be perfectly detected if the signal is strong and the proportion is not too large. At the random impulses case the impulse power has to be 25 dB above the 'noise' floor for the perfect detection meaning that only 20 % impulse proportion is allowed in this particular example with SNR -3 dB or higher. This seems bad. In the random impulses case, it may happen that the strong impulses are detected whereas the weak ones are not such that the performance of the system is not necessarily reduced. Fig. 5 gives detection rate results for the FCME at different SNR and proportion of impulses. It can be concuded that if the proportion of impulses is 50 %, then 50 % harmful impulses (ISR 21 dB or larger) are detected if SNR is 0 dB or larger.

CONCLUSIONS

Impulse detection and rejection methods for radio systems was discussed. The CME methods where discussed more detailed. The FCME method can reject impulses even if their proportion is 90 % of all the samples! The FCME performed much better than the rest investigated methods that

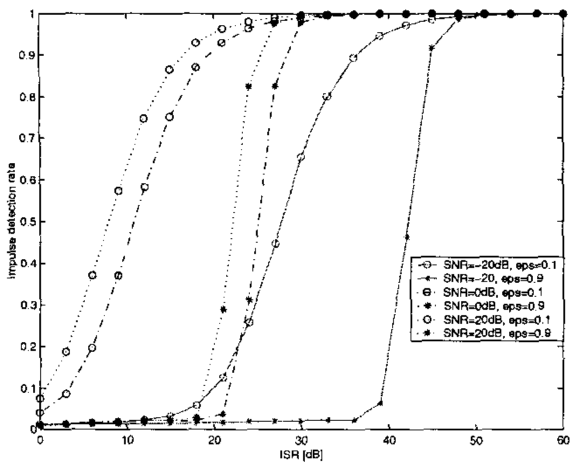

Figure 5. Impulse detection rate of the FCME algorithm as a function ISR for different SNR and proportion of impulses (eps).

where the CME and one-shot methods. The CME method performed better than the one-shot method.

REFERENCES

[1] D. Middleton, "Non-Gaussian noise models in signal processing for telecommunications: New methods and results for class A and B noise models," *IEEE Transactions on Information Theory*, vol. 45, no. 4, pp. 1129–1149, May 1999.

[2] P. J. Huber, "Robust estimation of a location parameter," *The Annals of Mathematical Statistics*, vol. 35, pp. 73–101, 1964.

[3] A. Spaulding and Middleton D., "Optimum reception in an impulsive interference environment-part I: Coherent detection," *IEEE Transactions on Communications*, vol. 25, no. 9, pp. 910–923, September 1977.

[4] A. Spaulding and Middleton D., "Optimum reception in an impulsive interference environment-part II: Incoherent detection," *IEEE Transactions on Communications*, vol. 25, no. 9, pp. 924–934, September 1977.

[5] P. Tsakalides and C. Nikias, "Maximum likelihood localization of sources in noise modeled as a stable process," *IEEE Transactions on Signal Processing*, vol. 43, no. 11, pp. 2700–2713, November 1995.

[6] K.-J. Wang and Y. Yao, "New nonlinear algorithms for narrowband interference suppression in CDMA spread-spectrum systems," *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 12, pp. 2148–2153, December 1999.

[7] X. Wang and V. Poor, "Robust multiuser detection in non-Gaussian channels," *IEEE Transactions on Signal Processing*, vol. 47, no. 2, pp. 289–305, February 1999.

[8] H. Delic and A. Hocanm, "Robust detection in DS-CDMA," *IEEE Transactions on Vehicular Technology*, vol. 51, no. 1, pp. 155–170, January 2002.

[9] Y.-Z. Liang and Kvalheim O. M., "Robust methods for multivariate analysis - a totorial review," *Chemometrics and Intelligent Laboratory Systems*, vol. 10, pp. 1–10, 1996.

[10] J. W. Wisnowski, D. C. Montgomery, and J. R. Simpson, "A comparative analysis of multiple outlier detection procedures in the linear regression model," *Computational Statistics & Data Analysis*, vol. 36, pp. 351–382, 2001.

[11] P. Henttu and S. Aromaa, "Consecutive mean excision algorithm," in *Proceedings of the IEEE International Symposium on Spread Spectrum Techniques and Applications*, Prague, Czech Republic, 2002, vol. 2/3, pp. 450–454.

[12] S. Kassam and H. V. Poor, "Robust techniques for signal processing: A survey," *Proceedings of the IEEE*, vol. 73, no. 3, pp. 433–481, March 1985.

[13] J. A. Young and J. S. Lehnert, "Analysis of DFT-based frequency excision algorithms for direct-sequence spread-spectrum communications," *IEEE Transactions on Communications*, vol. 46, no. 8, pp. 1076–1087, August 1998.

[14] A. S. Hadi, "Identifying multiple outliers in multivariate data," *Journal of the Royal Statistical Society–Series B*, vol. 54, no. 3, pp. 761–771, 1991.

[15] A. S. Kosinski, "A procedure for the detection of multivariate outliers," *Computational Statistics & Data Analysis*, vol. 29, 1999.

[16] W. R. Thompson, "On a criterion for the rejection of observations and the distribution of the ratio of deviation to sample standard deviation," *The Annals of Mathematical Statistics*, vol. 6, no. 4, pp. 214–219, December 1935.

[17] J. G. Proakis, *Digital Communications*, McGraw-Hill, New York, third edition, 1995.

[18] M. E. Frerking, *Digital Signal Processing in Communication Systems*, Van Nostrand Reinhold, New York, 1994.

[19] P. T. Capozza, B. J. Holland, T. M. Hopkinson, and R. L. Landrau, "A single-chip narrow-band frequency-domain excisor for a Global Positioning System (GPS) receiver," *IEEE Journal of Solid-State Circuits*, vol. 35, no. 3, pp. 401–411, March 2000.

[20] W. Press, W. Vetterling, S. Teukolsky, and B. Flannery, *Numerical Recipies in C*, Cambridge University Press, New York, USA, second edition, 1992.