Australian Telecommunications Research Institute &
School of Electrical and Computer Engineering

Goodness-of-Fit and Detection Problems in Impulsive Interference

by

Christopher Liong Brown

This thesis is presented as part of the requirements for the award of the Degree of Doctor of Philosophy of the Curtin University of Technology

2000

To my parents
and to my grandmother, Barbara

ABSTRACT

After defining the structure to a signal detection scheme, this dissertation describes and addresses some of the unresolved issues associated with its use when the interference encountered is impulsive. The alpha-stable (αS) family of distributions is used as a model of this interference due to its physical interpretation and its general form. Despite its attractive features, difficulties arise in using this distribution due to, amongst other things, the lack of a general closed form expression for its probability density function. Relevant to the detection scheme used, this affects parameter estimation, signal detector design and goodness-of-fit tests. Significant contributions are made in the latter through the introduction of characteristic function based test that uses the parametric bootstrap. A modification of this test is then made to define a test of the level of impulsive behaviour – again the parametric bootstrap is employed to maintain levels of significance for this and another test based on testing the αS parameter values. The performance of these tests is examined under simulated and two sources of real, impulsive data, namely human heart rate variability and fluctuations in stock prices. Once the appropriateness of the model assumption has been verified, the final, signal detection process may take place. Detectors based on the locally optimum criterion and approximations to it are described and compared to their rank-based counterparts. Results are presented that suggest compelling arguments based on performance and computational complexity for the consideration of rank-based techniques.

Keywords : Impulsive behaviour, alpha-stable distribution, stable laws, Gaussianity testing, parameter estimation, goodness-of-fit, parametric bootstrap, signal detection, locally optimum detectors, rank-based detectors.

Table of Contents

List of Figures 7

List of Tables 9

Acknowledgments 11

Publications 13

Abbreviations 16

Chapter 1. Introduction 17
- 1.1. Aims and Objectives 18
- 1.2. Contributions 19
- 1.3. Thesis Scope and Overview 20

Chapter 2. Impulsive and Alpha-Stable Processes 21
- 2.1. Definition 23
- 2.2. Properties 25
- 2.3. Parameter estimation 26
  - 2.3.1. Regression method proposed by Koutrouvelis 27
  - 2.3.2. Kernel characteristic function estimator 28
  - 2.3.3. Truncation of the linear regression 28
  - 2.3.4. Results and Discussion 29
  - 2.3.5. Conclusions 33
- 2.4. Summary 33

Chapter 3. Tests for Gaussianity 34
- 3.1. Review 34
  - 3.1.1. Chi-square (χ^2) tests 35
  - 3.1.2. Cumulant and Moment based Gaussianity tests 35
  - 3.1.3. Empirical distribution function tests 36
- 3.2. Characteristic Function Based Techniques 37
  - 3.2.1. Kernel Characteristic Function Estimator 39
  - 3.2.2. Test statistics 39
  - 3.2.3. Choice of t values 41
- 3.3. Results and Discussion 42
- 3.4. Multi-dimensional Gaussianity Testing 44
  - 3.4.1. Results and discussion 45
  - 3.4.2. Conclusions 48

TABLE OF CONTENTS—*Continued*

3.5. Summary . . . 48

CHAPTER 4. ALPHA-STABLE GOODNESS-OF-FIT TESTS . . . **49**
4.1. Testing for the αS Distribution . . . 49
4.1.1. Method . . . 50
4.1.2. Simulation Results and Discussion . . . 53
4.1.3. Conclusions . . . 56
4.2. Testing the Level of Impulsive Behaviour . . . 56
4.2.1. Testing the α parameter . . . 57
4.2.2. Characteristic function based test . . . 59
4.2.3. Simulation Results and Discussion . . . 62
4.2.4. Conclusions . . . 65
4.3. Summary . . . 66

CHAPTER 5. TESTING REAL SOURCES OF IMPULSIVE BEHAVIOUR . . . **68**
5.1. Heart Rate Variability . . . 68
5.1.1. Alpha-Stable modelling of HRV . . . 69
5.1.2. Results and Discussion . . . 70
5.2. Stock Prices . . . 72
5.3. Summary . . . 75
5.4. Acknowledgment . . . 75

CHAPTER 6. SIGNAL DETECTION IN IMPULSIVE INTERFERENCE . . . **76**
6.1. Optimal and Suboptimal Detectors . . . 77
6.1.1. Optimum Tests . . . 77
6.1.2. Locally Optimum Detectors . . . 78
6.1.3. Locally Suboptimum Detectors . . . 80
6.1.4. Cauchy Detector . . . 81
6.1.5. Other Detectors for SαS interference . . . 81
6.2. Rank-Based Detectors . . . 82
6.2.1. Locally Optimum Rank Detectors . . . 82
6.2.2. Score functions . . . 84
6.2.3. Locally Suboptimum Rank Detectors . . . 85
6.2.4. Signed Quadratic Rank-Based Detector . . . 86
6.3. Distribution of Test Statistics . . . 87
6.3.1. LO and LSO Detectors with Nonlinear Score Functions . . . 87
6.3.2. Matched Filter . . . 89
6.3.3. LO and LOR Detectors . . . 90
6.3.4. Quadratic detectors . . . 91
6.3.5. Cauchy detector . . . 92

Table of Contents—*Continued*

6.4. Performance Analysis 92
6.5. Conclusions 99
6.6. Summary 99

Chapter 7. Conclusions and Future Directions 100

Appendix A. Characteristic Functions and Their Estimation . . . 103
A.1. Empirical Characteristic Function 103
A.2. Kernel Characteristic Function Estimator 104

Appendix B. Hypothesis Testing with the Parametric Bootstrap 106

References 110

List of Figures

FIGURE 1.1. A signal detection scheme. 17

FIGURE 2.1. Characteristic functions of some symmetric αS distributions. . . 24
FIGURE 2.2. Pdfs of some standard symmetric αS distributions. 25
FIGURE 2.3. Normal probability plot of some standard symmetric αS distributions. 25

FIGURE 3.1. The cfs of the Gaussian, Uniform and Laplace distributions. . . 37
FIGURE 3.2. The ecfs of 10 realisations of $n = 64$ observations of iid standard Gaussian data. 38
FIGURE 3.3. Realisations of the ecfs of Gaussian, Uniform and χ^2_2 distributed data and the true Gaussian cf evaluated at $t = 1.8$, plotted on the complex plane. 40

FIGURE 4.1. Block diagram of the αS goodness-of-fit test method. 51
FIGURE 4.2. The cfs of the t_2 and $\alpha\mathrm{S}(\alpha = 1.4443, \beta = 0, c = 0.8668, \delta = 0)$ distributions. 55
FIGURE 4.3. Rejection rates (in %) for the $\alpha\mathrm{S}(\alpha = 1.4443,\ \beta = 0,\ c = 0.8668,\ \delta = 0)$ versus t_2 distribution goodness-of-fit test. 55
FIGURE 4.4. Block diagram of the parametric bootstrap test for α. 59
FIGURE 4.5. Block diagram of the characteristic function based test using the parametric bootstrap. 67

FIGURE 5.1. Segment of HRV signal from patient with ventricular tachycardias. 69

FIGURE 6.1. Structure of the LO Detector. 79
FIGURE 6.2. Locally Optimum score functions for various SαS distributions. 79
FIGURE 6.3. LSO nonlinear score functions when $\alpha = 1.6$. 80
FIGURE 6.4. Structure of the LOR Detector. 83
FIGURE 6.5. Locally Optimum Rank score functions for various SαS distributions. 84
FIGURE 6.6. Rank score functions used for LOR and LSOR detectors when $\alpha = 1.6$ and sample size of 100. 86
FIGURE 6.7. Receiver Operating Characteristics for some linear correlation detectors when $\alpha = 1.6$. 95
FIGURE 6.8. Receiver Operating Characteristics for some linear rank-based detectors when $\alpha = 1.6$. 95
FIGURE 6.9. Detection power functions for the LO, LSO-power, LOR and LSOR-tr detectors for sample sizes of 100, 5% level of significance and SαS interference with $\alpha = 1.6$. 97

LIST OF FIGURES—*Continued*

FIGURE A.1. The variance of the real and imaginary components of the ecf of 64 observations from a Gaussian process. 104
FIGURE A.2. The effect of multiplicative smoothing on the ecf of one realisation of a Gaussian process with $n = 64$ observations. 105

List of Tables

Table 2.1. MSE ($\times 10^{-3}$) of estimates of α using tabulated regression region, $\beta = 0$. 30
Table 2.2. MSE ($\times 10^{-3}$) of estimates of α using a variance threshold regression region, $\beta = 0$. 31
Table 2.3. MSE ($\times 10^{-2}$) of estimates of β using tabulated regression region. 32
Table 2.4. MSE ($\times 10^{-2}$) of estimates of β using a variance threshold regression region. 32
Table 2.5. MSE ($\times 10^{-2}$) of estimates of β for 200 observations, $\alpha = 1.5$. . . 32

Table 3.1. Rejection rate (in %) for the KCFE based test when testing at the 5% level. 43
Table 3.2. Rejection rate (in %) for the KCFE based test when testing at the 5% level, comparing the statistics based on the peak absolute difference, mean difference and MSE. 44
Table 3.3. Rejection rate (in %) for the KCFE, Shapiro Wilk W test and Mardia and Foster's S_W^2 test for bivariate Gaussianity at the 5% level of significance for $n = 20$. 46
Table 3.4. Rejection rate (in %) for the KCFE, Shapiro Wilk W test and Mardia and Foster's S_W^2 test for bivariate Gaussianity at the 5% level of significance for $n = 50$. 46

Table 4.1. Rejection rates of the αS goodness-of-fit test for various alpha-stable(α, β) distributions (in %) using sample sizes of 400, 5% nominal level and 500 replications. 54
Table 4.2. Rejection rates of the αS goodness-of-fit test for various non-stable distributions (in %) using sample sizes of 400, 5% nominal level and 500 replications. 54
Table 4.3. Rejection rates (in %) of tests on the level of impulsive behaviour of standard SαS distributed data with $\beta = 0$ where $\alpha_0 = 2$, estimated from 1000 replications. 62
Table 4.4. Rejection rates (in %) of tests on the level of impulsive behaviour of standard SαS distributed data with $\beta = 0$ where $\alpha_0 = 1.6$, estimated from 1000 replications. 63
Table 4.5. Rejection rates (in %) of tests on the level of impulsive behaviour of standard SαS distributed data with $\beta = 0.7$ where $\alpha_0 = 1.6$, estimated from 500 replications. 64
Table 4.6. Rejection rates (in %) of tests on the level of impulsive behaviour of standard SαS distributed data with $\beta = -0.3$ where $\alpha_0 = 1.6$, estimated from 500 replications. 65

LIST OF TABLES—*Continued*

TABLE 5.1. Rejection rate (in %) of goodness-of-fit tests for the αS distribution and the Gaussian distribution at the 5% level on five HRV data files. . . 70
TABLE 5.2. Conventional analysis of five HRV data files and estimated αS parameters. 71
TABLE 5.3. Sample sizes, p values and estimated αS parameter values for the logarithmic price changes of some stocks, when testing the goodness-of-fit of the αS distribution. 73
TABLE 5.4. Estimated α and p values for the tests of impulsive behviour. Four p values are shown for each company, two each for the $T_{n,\alpha}$ and $T_{n,\phi}$ tests and two each for $\alpha_0 = 1.8$ and $\alpha_0 = 2$. 74

TABLE 6.1. Detection and false alarm (bracketed) rates in % for linear correlation detectors of signals in SαS interference, $\theta = 0.5$. 93
TABLE 6.2. Detection and false alarm (bracketed) rates in % for linear rank-based detectors of signals in SαS interference, $\theta = 0.5$. 93
TABLE 6.3. Detection and false alarm (bracketed) rates in % for quadratic correlation detectors of signals in SαS interference, $\theta = 0.5$. 94
TABLE 6.4. Detection and false alarm (bracketed) rates in % for quadratic rank-based (Adichie) detectors of signals in SαS interference, $\theta = 0.5$. . . 94
TABLE 6.5. Detection and false alarm (bracketed) rates in % for a selection of detectors of signals in SαS ($\alpha = 1.6$) interference where the detector is designed for another α, $\theta = 0.5$ 98

TABLE B.1. The parametric bootstrap procedure for testing the hypothesis $\mathsf{H} : \boldsymbol{\xi} = \boldsymbol{\xi}_0$ against $\mathsf{K} : \boldsymbol{\xi} \neq \boldsymbol{\xi}_0$. 108
TABLE B.2. Beran's conditions to prove the asymptotically correctness of a bootstrap test and that its power can be consistently estimated [7]. . . . 109

Acknowledgments

How do you do it? How do you thank all the people who have helped you emerge from an ordeal/experience such as this, with a bit of sanity (though not much, possibly on a par with the amount of hair I have left) and with three letters after your name? It's a futile exercise and one that cannot possibly be completed through the vessel of a few pages in a thesis that just as few will ever read. Despite this, I'll continue writing this in an effort to convey at least some of the gratitude I feel towards some of the people who have been a part of my life over the past four and a half years.

My whole PhD experience has been "interesting" to say the least. Many highs and many lows have been encountered along the way. I think it was Dr Jonathon Ralston who told me in those critical early days to "low-pass filter everything". This turned out to be invaluable advice that I have always tried to adhere to, with varying degrees of success. Jonno was also the one who prophetically left me a box of paracetamol as a parting gift when he left for the "real world", knowing full well what lay in front of me. Ta, Jonno.

Special thanks to Dr D. Robert Iskander, a former co-supervisor, then colleague, then supervisor again. He's always been there, in body or in spirit, and has helped guide me and numerous others in the SPRC/CIP/CSP labs. Robert, I will always be grateful for the support and advice (technical and otherwise) you gave to me.

Thanks also to other friends and colleagues at QUT: Dr Geoff Roberts, Dr Ken Turner, Dr Norbert Härle, Joanne Smith, John "the grand wazoo" Williams, Pedro, CP, Maurice the Moose, Murray (IF23 4 eva!), Phil, Greg, Sid Meiers, Tanya, Dolores, Cherie and, in particular, Netty.

To all the gang at Chalmers University, especially Prof. Mats Viberg, Thomas, Anders, Håkan, Tony, Jonas and Sokol, thank you for making those few months of my northern sojourn some of the most enjoyable of my studies.

To those in the most isolated city on the Earth, whose company made me feel anything but alone: Kip, Iain "Mr Miri" Murray, Todd and all those members of the CSP lab – Ramon (ma-a-a-a-te), Matt (peace, love and mung beans), Mark, Dr Hwa-Tung Ong, Herr Doktor Per Pelin, Le Doktor Anisse Taleb and Dr Braham Barkat – thank you one and all.

Of course, it should go without saying that special thanks goes out to my supervisor, Prof. Abdelhak Zoubir. From his first lecture, Abdelhak has always encouraged and inspired me. All those who have had the privilege of working with him will understand this. I can only hope that the enthusiasm he has for signal processing and the pursuit of learning will stay with me.

And finally, all-embracing thanks and love to my family and friends (especially Eryn, Dave, Julia, Andrew, Loz, Mark and Carla, as well as the members of the "John Williams gang" listed above) who were there, not only from 9 to 5, but at all hours of the day and night for many, many years. Your unquestioning love and friendship always reminded me of what should always be our true priorities in this life.

To the 99% of you that will only read these two pages of this thesis, I leave you with something sent to me, and to which I hope very few will relate to

> *Stress is when you wake up screaming and you realize you haven't fallen asleep yet.*

CB 16 July, 2000 (as promised, *before* Sydney 2000)

Publications

The following publications have been produced during the period of PhD candidacy.

Internationally Refereed Journal Articles

1. C. L. Brown and A. M. Zoubir. A nonparametric approach to signal detection in impulsive interference. *IEEE Transactions on Signal Processing*, 48(9), September 2000. (to appear).

2. C. L. Brown and A. M. Zoubir. Testing for impulsive behaviour: A bootstrap approach. *Digital Signal Processing*. (in press).

Internationally Refereed Conference Papers

1. C. L. Brown and A. M. Zoubir. A new approach to testing Gaussianity using the characteristic function. In *Proceedings of the First International Conference on Information, Communications and Signal Processing, ICICS'97*, pages 1198–1202, Singapore, September 1997.

2. D. R. Iskander, C. L. Brown, and B. Boashash. A rank based approach to testing departure from Gaussianity of stationary signals. In *Proceedings of the First International Conference on Information, Communications and Signal Processing, ICICS'97*, pages 1194–1197, Singapore, September 1997.

3. K. J. Turner, F. A. Faruqi, and C. L. Brown. Tracking nonlinear systems using higher order moments. In *Proceedings of the IEEE Signal Processing Workshop on Higher-Order Statistics, HOS'97*, pages 52–56, Banff, Canada, July 1997.

4. A. M. Zoubir, C. L. Brown, and B. Boashash. Testing multivariate Gaussianity with the characteristic function. In *Proceedings of the IEEE Signal Processing*

Workshop on Higher-Order Statistics, HOS'97, pages 438–442, Banff, Canada, July 1997.

5. C. L. Brown and A. M. Zoubir. On the estimation of the parameters of alpha-stable distributions using a truncated linear regression. In *Proceedings of the 9th IEEE Signal Processing Workshop on Statistical Signal and Array Processing, SSAP'98*, Portland, USA, September 1998.

6. C. L. Brown, A. M. Zoubir, and B. Boashash. On the performance of an adaptation of Adichie's rank tests for signal detection: and its relationship to the matched filter. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP'98*, pages 1505–1508, Seattle, USA, May 1998.

7. C. L. Brown and S. Saliu. Testing of alpha-stable distributions with the characteristic function. In *Proceedings of the IEEE Signal Processing Workshop on Higher-Order Statistics, HOS'99*, pages 224–227, Caesarea, Israel, June 1999.

8. S. Saliu, N. Edvardsson, D. Poçi, and C. Brown. Alpha-stable modeling of arrhythmias on heart rate variability signals. In *Proceedings of the Third International Workshop on Biosignal Interpretation, BSI99*, Chicago, USA, June 1999.

9. A. M. Zoubir and C. L. Brown. Testing for impulsive interference: A bootstrap approach. In *Proceedings of the 1999 Workshop on Defense Applications of Signal Processing (DASP99)*, LaSalle, USA, August 1999.

10. C. L. Brown and A. M. Zoubir. Locally suboptimal and rank-based known signal detection in correlated alpha-stable interference. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP'00*, Istanbul, Turkey, June 2000.

Nationally Refereed Conference Papers

1. C. L. Brown and A. M. Zoubir. Nonparametric detection in the frequency domain using the Conditional Wilcoxon test. In *Proceedings of ISSPA 96*, volume 1, pages 105–108, Gold Coast, Australia, August 1996.

2. C. L. Brown. Using a rank test and the matched filter for signal detection in linear models with non-Gaussian interference. In *Proceedings of the Second Australian Workshop on Signal Processing and its Applications, WoSPA'97*, pages 139–142, Brisbane, Australia, December 1997.

Abbreviations

αS	alpha-stable
AR	autoregressive
cf	characteristic function
cdf	cumulative distribution function
ECG	electrocardiogram
ecf	empirical characteristic function
edf	empirical distribution function
HRV	Heart Rate Variability
iid	independent and identically distributed
KCFE	kernel characteristic function estimator
LO	locally optimum
LOR	locally optimum rank
LSO	locally suboptimum
LSOR	locally suboptimum rank
MF	matched filter
MLE	maximum likelihood estimate
MSE	mean squared error
PB	premature beats
pdf	probability density function
RNG	random number generator
ROC	Receiver Operating Characteristic
SαS	symmetric alpha stable
SPB	supraventricular premature beats
UMP	uniformly most powerful
VPB	ventricular premature beats
ZMNL	zero-memory nonlinearities

Chapter 1
INTRODUCTION

A complete system to efficiently determine the presence or absence of a signal within a finite observation sample containing interference is far more complicated than just a signal detection scheme. With the multitude of detection techniques available now, the choice of detector has become a significant issue.

Detectors have been designed for many of the common assumptions made on real data. While different assumptions lead to different detector structures, there is also often a trade-off between high performance and the generality of the necessary assumptions or pre-conditions. In order to achieve the best overall performance for a set of data, it is therefore necessary to test the data to determine which detector assumptions hold and will yield high performance.

A signal detection scheme may be considered to follow the general structure [39] in Figure 1.1. Using this scheme as a template, this thesis addresses some of the issues that remain when dealing with data that contains impulsive interference.

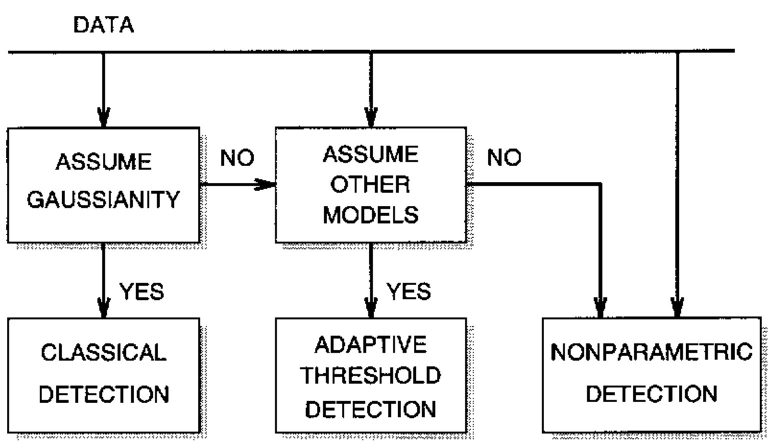


FIGURE 1.1. A signal detection scheme.

Impulsive interference is an interesting, naturally occurring phenomenon that can prove particularly troublesome to some traditional statistical techniques. Relatively infrequent, though large in magnitude, the spikes that are the characteristics of impulsive behaviour may have a disproportionate influence on methods that are not designed with their presence in mind. Consequently, standard techniques may suffer significant degradation in performance and the development of specialised approaches is warranted.

In the following chapters, a number of problems in impulsive interference are investigated, with special emphasis on goodness-of-fit testing and signal detection. These are two important components of the scheme in Figure 1.1 that remain underdeveloped. In particular, emphasis will be placed on one particular statistical model for impulsive behaviour.

The alpha-stable distribution (αS) has been used to model the statistical characteristics of impulsive processes and has received significant attention from the signal processing community in recent times. While numerous references exist in the literature to signal processing techniques that can be used in the presence of αS distributed interference, there remain many unsolved theoretical and practical problems. Improvements to current parameter estimation procedures may still be possible since no maximum likelihood estimate (MLE) has been found. Statistical tests to determine the appropriateness of the αS distribution are also lacking, as are practical signal detection schemes. Contributions in this area are expected to add to the viability and performance of testing and detection schemes for αS interference, as well as, more generally, impulsive interference.

1.1 Aims and Objectives

The aim of this dissertation is to develop the remaining steps for a signal detection scheme as described in Figure 1.1 when the interference is impulsive and may be modelled by an αS distribution. Where some of the elements of this scheme have already been addressed in the literature, they will be investigated and, where appropriate, modifications and improvements suggested. However, the main objectives revolve around the least developed aspects of the scheme, specifically

1. Develop goodness-of-fit techniques that can be used to justify the assumption that an αS distribution is an appropriate model for the interference encountered

2. Design of procedures for the detection of a known signal in additive interference that do not suffer from excessive computational burden

3. Performance study of the proposed techniques using simulated and real data where possible.

1.2 Contributions

The significant contributions made in this dissertation include

1. A characteristic function based goodness-of-fit technique for testing that observations are from an αS distribution, against the open alternative that they are not.

2. Tests of the level of impulsive behaviour of αS processes using the characteristic function and parameter estimates.

3. Application of these tests to human heart rate variability and stock price data.

4. Detectors derived from approximations to the local optimality criterion and corresponding rank-based detectors that achieve similar performance with reduced computational complexity and a constant false alarm rate.

1.3 Thesis Scope and Overview

Chapter 2. To begin, an introduction to impulsive processes, and in particular the αS distribution is presented. This will establish some definitions and the properties that motivate the investigation of the problem and will be used later on.

Chapter 3. The first component of Figure 1.1, namely Gaussianity testing is necessary. Particular emphasis is placed on some contributions in characteristic function based Gaussianity testing.

Chapter 4. These characteristic function based techniques are generalised and adapted in the introduction of a goodness-of-fit test for the alpha-stable distribution. Also in this chapter are two tests on the level of impulsive behaviour. These three tests form part of the contribution to testing the "Assume other models" component of the diagram mentioned above.

Chapter 5. Partly for validation of these tests and partly for demonstration of real life impulsive behaviour, the tests are applied to some impulsive real data sources.

Chapter 6. After completing the justification stages of Figure 1.1, the problem of signal detection in the presence of alpha-stable interference is investigated. A number of parametric and nonparametric detectors are proposed and evaluated.

Chapter 7. Conclusions are drawn and areas for future research are identified.

Chapter A. The characteristic function and some estimators of it are defined.

Chapter B. The parametric bootstrap, used extensively in Chapter 4 is described.

Chapter 2

Impulsive and Alpha-Stable Processes

Impulsive behaviour is a naturally occurring characteristic of many signals. Observations that appear far from the mean are sometimes regarded as outliers, a result of a freak occurrence, a measurement error or resulting from another process with another statistical distribution. As a result, they may be removed from the sample, censored, or limited to a lower level. This is often a necessity due to the disproportional effect an outlier may have on conventional statistical techniques.

However, rather than being disposable anomalies, it may be that these outliers carry useful information about the generating process, and therefore should be included in the statistical analysis. In this case, a unifying model should be used that appropriately characterises both the more probable observations located around the mean or median, as well as the less likely, but equally (or sometimes, more) important, outliers.

Statistical models that incorporate impulsive behaviour have probability density functions (pdfs) with tails that are heavier than the Gaussian distribution, indicating the higher likelihood of observations occurring a significant distance from the median. Such impulsive behaviour has been observed in low-frequency atmospheric noise, fluorescent lighting systems, combustion engine ignition, urban and indoor radio channels, underwater acoustic channels, economic stock prices, certain biomedical signals and in computer network traffic [2, 8, 9, 10, 36, 41, 42, 43, 47, 58, 61, 64, 89, 90]. Depending on the circumstances producing the impulsive process, the appropriateness of particular models is generally argued on the statistical physical or an empirical level.

Statistical physical models can be shown to be appropriate through detailed analysis of the physics or mechanics of the generating process. Perhaps the most widely known models in statistical signal processing that incorporate impulsive behaviour are those proposed by Middleton [60, 61] for electromagnetic interference. Physical interpretation of most parameters of these models can be made. Unfortunately, these models tend to be highly complicated as necessitated by the complexity of the generating mechanism. Because of their specialised nature, they may also be difficult or not suitable to apply to any other impulsive sources.

In some situations the theoretical justification of a model may be unavailable due to the prohibitively high degree of complexity in the generation process and may only be mathematically tractable after making broad assumptions or approximations. For example, it would be very difficult to complete a theoretical justification of a model for many econometric time series due to the plethora of input variables from diverse sources, including human psychological factors. It may be that even the widespread knowledge of such a model may alter the inputs to it and render it inappropriate.

In cases such as these, and in others, a statistical model may be deemed appropriate on the basis of empirical fitting to data – that is, as an empirical model. Overall, the complexity-performance tradeoff may mean simpler approximate models may find use above statistical physical models, and where the simplifying assumptions are valid, there may be a negligible difference in accuracy. Despite the lack of a direct physical interpretation of the model parameters, it is certainly true that intuitive relationships between statistical parameters and physical parameters may be evident. This is illustrated by example in Chapter 5.

The alpha stable (αS) distribution has been prominently promoted as an impulsive model. While a physical interpretation of the distribution does exist [36] and the Generalised Central Limit Theorem and the Stability Property give the distribution importance in the field of mathematical studies, debate still exists on its use as a model for real-world impulsive behaviour. The simplest argument against its use being that the infinite variance of the distribution (in all non-Gaussian cases) disqualifies it as a model for sources known to possess finite power.

Regardless of this, it is felt that as a broad family of distributions with four parameters, two for shape and one each for location and scale, in the very least the αS distribution may find use as an empirical model for many varied impulsive sources and, thus, statistical techniques must be developed for its justification and use.

2.1 Definition

The αS distribution can be defined in terms of its characteristic function (cf)

$$\phi(t) = \exp\{j\,\delta\,t - |\,c\,t\,|^{\alpha}\,[1 - j\beta\,\mathrm{sgn}(t)\,\omega(t,\alpha)]\} \tag{2.1}$$

where

$$\omega(t,\alpha) \;=\; \begin{cases} \tan(\alpha\pi/2)\,, & \alpha \neq 1 \\ -(2/\pi)\ln|t| & \alpha = 1 \end{cases}$$

and $\mathrm{sgn}(t)$ is the signum function [70, 30], or equivalently

$$\phi(t) = \begin{cases} \exp\{-|\,c\,t\,|^{\alpha} + j\,\delta\,t + j\,|\,c\,t\,|^{\alpha}\,\beta\,\mathrm{sgn}(t)\;\tan(\alpha\pi/2)\}\;, & \alpha \neq 1 \\ \exp\{-c\,|\,t\,| + j\,\delta\,t - j\,c\,\beta\,(2/\pi)\,t\,\ln|t|\}\;, & \alpha = 1\;. \end{cases}$$

See Appendix A for a definition and description of the cf. The four parameters of the distribution are

- α – the characteristic exponent, $0 < \alpha \leq 2$
- β – the skewness parameter, $-1 \leq \beta \leq 1$
- c – the scale parameter, $0 \leq c$, although sometimes the dispersion $\gamma = c^{\alpha}$ is used to indicate scale.
- δ – the location parameter, $-\infty < \delta < \infty$.

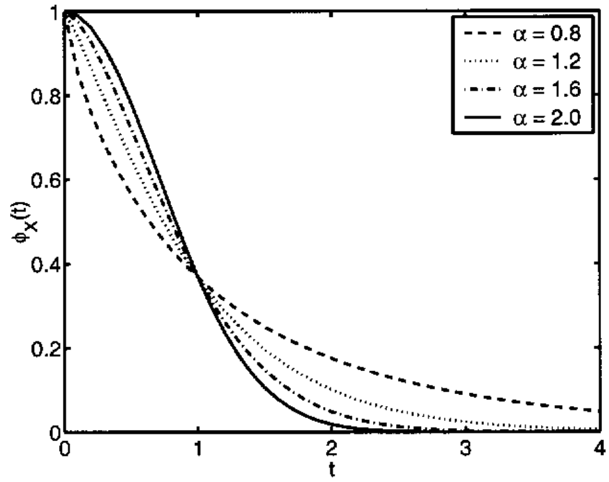


FIGURE 2.1. Characteristic functions of some symmetric αS distributions.

For a discussion on the sign of β when $\alpha = 1$, please refer to [30]. The cfs for the symmetric ($\beta = 0, c = 1, \delta = 0$) αS distribution are shown for several values of α in Figure 2.1.

Closed form expressions for the probability density function (pdf) of the distribution only exist for a small number of special cases: the Gaussian ($\alpha = 2$), Cauchy ($\alpha = 1, \beta = 0$) and Lévy ($\alpha = \frac{1}{2}, \beta = 1$) distributions. The only α-stable distribution to have finite variance is the Gaussian one.

The characteristic exponent, α, controls the impulsive behaviour of the process. When $\alpha = 2$, its maximum value, the αS distribution is equivalent to the Gaussian distribution. As α decreases, the tails of the distribution become heavier and the process becomes more impulsive, that is, the more outliers occur in an observed series. In Figure 2.2, numerical approximations to the pdfs of several standardised symmetric αS distributions are shown for varying α. The change in the weight of the tails is clearly evident.

An alternative illustration of the effect of α on the tails of the distribution is shown in Figure 2.3. Presented are the normal probability, or q-q, plots for standard symmetric αS distribution when $\alpha = 1.6$, $\alpha = 1.8$ and $\alpha = 2$. The Gaussian dis-

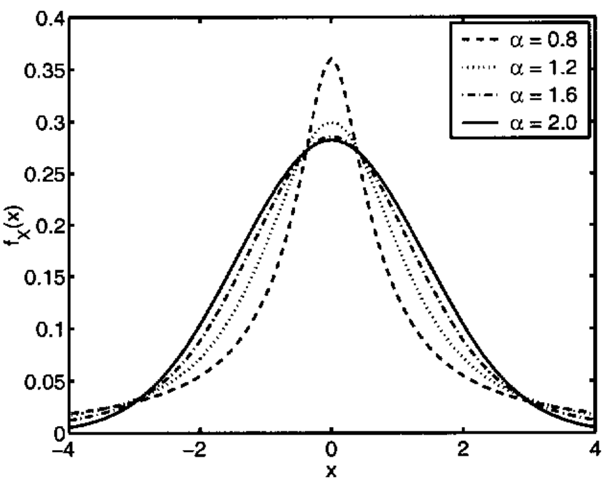


FIGURE 2.2. Pdfs of some standard symmetric αS distributions.

tribution ($\alpha = 2$) appears as a straight line on this type of plot. The non-Gaussian αS distributions look very Gaussian around the median, that is their plots are quite linear, yet the difference in the tails is quite pronounced.

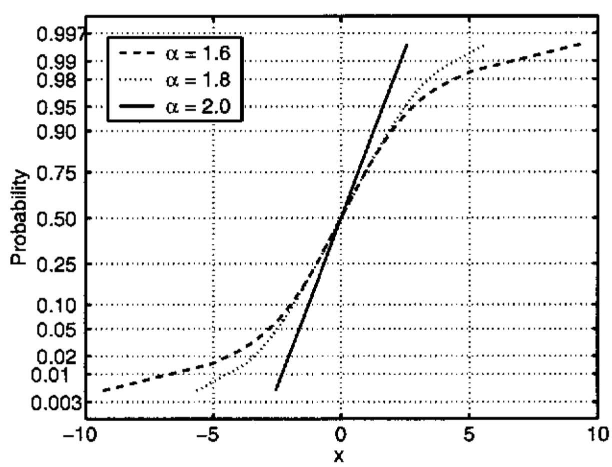


FIGURE 2.3. Normal probability plot of some standard symmetric αS distributions.

2.2 Properties

At the start of this chapter, the Generalised Central Limit Theorem and the Stability Property were cited as two properties of the αS distribution that gave the distribution mathematical importance. They are described here.

Theorem 1 (Stability Property). *If, for arbitrary positive numbers a_1 and a_2, there is a positive number a and a real number b such that*

$$a_1X_1 + a_2X_2 \stackrel{d}{=} aX + b$$

when X_1, X_2 and X are independent and identically distributed (iid), then X must have an α stable distribution [70].

Using a generalisation, the following theorem can be shown.

Theorem 2 (Generalised Central Limit Theorem). *The limit of the normalised sum of iid random variables is a stable distribution. That is, if $X_1, X_2, \ldots, X_n$ are iid, then as $n \to \infty$ the limit in distribution of*

$$S_n = (X_1 + X_2 + \ldots + X_n)/a_n - b_n$$

is α stable [64].

The Central Limit Theorem, as defined in [59] imposes the additional condition that $X_1, X_2, \ldots, X_n$ have finite variance – the resulting limiting distribution being Gaussian rather than the more general αS distribution.

Given the significance of these theorems, it is not surprising that a physical model can be defined that results in the generation of αS random processes. Indeed, it can be shown [64] that under certain assumptions, the received signal from independent, Poisson temporally and spatially distributed sources is a symmetric αS distribution. This same filtered-impulse mechanism, using different, more detailed assumptions, was used to develop the statistical physical Middleton Class B model [60].

2.3 Parameter estimation

The absence of closed form expressions for the density of general αS distributions may be seen as an impediment to the estimation of their parameters, especially for techniques such as maximum likelihood based approaches that rely on this expression. However, it does provide motivation for using the cf domain, where such expressions do exist. While numerical methods exist that mimic these traditional, optimal techniques, simulation results have shown that a cf based estimation procedure introduced by Koutrouvelis [50] provides better results than these, as well as other cf based methods.

2.3.1 Regression method proposed by Koutrouvelis

By manipulating the cf of the αS distribution it was shown that estimators of the parameters could be found through a regression technique. Due to the consistency of the empirical characteristic function (ecf) as an estimator of the true cf (see Section A.1), the parameter estimators were found to be consistent and asymptotically unbiased. The ecf of a vector of n iid observations, $\boldsymbol{X} = [X_1, X_2, \dots, X_n]^T$, is defined in equation (A.1) and is denoted by $\hat{\phi}^e_{\mathbf{X}}(t)$. Presently, expressions for the asymptotic distributions of the parameter estimators are unavailable.

To estimate α and c, the form of the cf of an αS random variable, $\phi(t)$, is manipulated to obtain the following relationship

$$\ln(-\ln \mid \phi(t) \mid^2) = \ln(2c^\alpha) + \alpha \ln \mid t \mid . \tag{2.2}$$

From this, it is possible to estimate the α and c parameters by regressing $\ln(-\ln \mid \phi(t) \mid^2)$ on $\ln \mid t \mid$. Similarly, the following expression was derived for estimating β and δ

$$\arctan\left\{\Im\left(\phi(t)\right)/\Re\left(\phi(t)\right)\right\} = \delta t - \beta c^\alpha \tan(\pi\alpha/2)\operatorname{sgn}(t) \mid t \mid^\alpha \tag{2.3}$$

where $\Im\left(\phi(t)\right)$ and $\Re\left(\phi(t)\right)$ are the imaginary and real components of $\phi(t)$, respectively.

By using the estimates for α and c evaluated using (2.2), estimates for β and δ may be found by regressing on t and $\operatorname{sgn}(t) \mid t \mid^\alpha$. Note, however, that integer multiples of 2π may have to be added to (2.3) to yield a continuous function spanning beyond the principal domain. The reader is referred to [50] for full details.

2.3.2 Kernel characteristic function estimator

Any technique in the cf domain will be dependent on the accuracy of the estimate of the cf used. From Section A.1 it can be seen that the ecf, while being a natural choice and an unbiased technique for estimating the cf of a process from a finite number of iid observations, suffers from variance that approaches a constant value of $1/2n$, where n is the number of observations available.

It has been shown in [93] that an alternative cf estimator, the kernel characteristic function estimator (KCFE), is more suited to Gaussianity testing in the cf domain. Similarly, it was proposed in [15] that the properties of the KCFE may lead to better parameter estimation performance for αS distributions. See Section A.2 for details on the KCFE.

Here the fixed KCFE (described in Section A.2) is considered, whereby a fixed kernel function is multiplied by the ecf. This process is equivalent to convolutive smoothing as occurs in kernel density estimation [75] and has the effect of reducing the variation in the tail of the estimator, at the expense of the introduction of a bias.

It has been suggested [31] that if the density being estimated has heavy tails, as αS distributions do, then the smoothing kernel should also have heavy tails. Consequently, kernels based on the cf of Student's-t distribution are investigated here.

2.3.3 Truncation of the linear regression

It is significant to note that the regression in the cf domain should only be performed on a truncated region of the t space. This is due to the non-vanishing variance properties of cf estimators. In the technique described by Koutrouvelis, a linear spacing of points in this region is proposed. To define these regions, the optimal numbers of such points are tabulated for seven or eight values of α and for sample sizes of 200, 800 and 1600. Through simulations, it was established that the tabulated values minimised the mean squared error of the estimates. If the actual number of observations or the true α differ from those provided in the tables, interpolation or extrapolation may be required.

An alternative introduced here and in [15] is to define the region of regression in terms of the variance of the ecf. The variances of the real and imaginary components of the ecf are given in (A.2) and (A.3). It is proposed that the region of regression span the values of t where the variance of the ecf is below a threshold. This may be specified as a proportion of the asymptotic variance of the ecf, that is $1/2n$ as $|t| \to \infty$.

An obvious advantage is that the regression region may be dynamically calculated, using the particular operating conditions encountered, rather than by approximation or interpolation on a table of values. It also avoids the possibility of including regions where the variance of the ecf is larger than the asymptotic value – as can occur when $\alpha < 1$. In this case, it can be shown that the variance of the ecf, as a function of t, does not rise monotonically to the asymptotic value. In fact, it exceeds this value, before approaching $1/2n$ as $|t| \to \infty$.

The spacing of points in the region of regression is to be the same as used in [50], that is, linear with fixed spacing of $\pi/25$ and $\pi/50$ for (2.2) and (2.3) respectively. It was suggested that the estimation procedure is relatively insensitive to the spacing of the points, but rather it is more dependent on the span of the chosen points.

2.3.4 Results and Discussion

For the sake of simplicity, data has been generated from standardised distributions, that is, $c = 1$ and $\delta = 0$. This does not affect the generality of these techniques. Rather, it has been done to allow more detailed analysis of results into the estimation of α and β, which, being shape parameters, are of the most interest. Each set of operating conditions was simulated 1000 times, and the mean squared error (MSE) of the estimates was computed as a measure of the performance of the technique.

The smoothing kernel used for the estimates using the KCFE was a t_3 kernel. Other more Gaussian-like kernels were tried, such as a t kernel with higher degrees of freedom, however the t_3 kernel was found to perform best under the largest number of conditions.

Estimation of α. The results shown in Tables 2.1 and 2.2 show the MSE in the α estimate using the tabulated regression region and the more recently proposed, variance threshold region [15]. Each table shows results for different values of α and using both the ecf and the KCFE. Although all distributions were symmetric, β does not affect the estimation of α. This should be evident from (2.2).

	cf			α			
n	estimator	0.8	1.0	1.2	1.4	1.6	1.8
50	ecf	31.83	38.36	48.73	53.24	48.54	33.42
	KCFE-t_3	61.44	52.57	47.84	41.20	29.26	19.31
200	ecf	8.13	9.74	11.20	12.59	13.11	10.04
	KCFE-t_3	24.47	19.96	16.28	14.60	10.34	6.38
400	ecf	3.86	4.79	5.58	6.22	6.72	5.01
	KCFE-t_3	14.71	12.66	10.14	7.93	6.35	3.48

TABLE 2.1. MSE ($\times 10^{-3}$) of estimates of α using tabulated regression region, $\beta = 0$.

It can be seen that the use of the KCFE, using a t_3 kernel, provides increased accuracy in the estimation of α for higher (true) values of α. For 50 observations, this technique provided superior results when $\alpha \geq 1.2$, while for 200 and 400 observations, it was superior when $\alpha \geq 1.6$.

The results using a threshold to determine the regression region in Table 2.2, can also be seen to compare favourably with those using the tabulated values in

n	cf estimator	0.8	1.0	1.2	1.4	1.6	1.8
50	ecf	37.19	41.07	51.35	56.37	53.40	34.42
	KCFE-t_3	70.92	53.44	50.29	45.81	33.38	18.69
200	ecf	7.48	9.47	10.79	13.06	13.94	10.75
	KCFE-t_3	30.76	21.27	16.46	15.45	11.66	6.83
400	ecf	3.29	4.47	5.28	6.33	7.31	5.38
	KCFE-t_3	20.31	13.77	10.25	8.28	7.22	3.78

TABLE 2.2. MSE ($\times 10^{-3}$) of estimates of α using a variance threshold regression region, $\beta = 0$.

Table 2.1. This threshold was set at 98% of the asymptotic variance of the ecf. While little difference in MSE can be observed between Tables 2.1 and 2.2, this confirms that the variance threshold method of setting regression regions is worthy of investigation since the tabulated regions were found, by simulation, to be optimal, that is, they minimised the MSE of the estimators.

Estimation of β. Although an estimate of α is used in (2.3) for the estimation of β, it can be noted from Tables 2.3 and 2.4 that the relative performance of the β parameter estimation procedures are relatively immune to changes in the value of α. Results for both tables were found for $\beta = 0$. While the MSEs varied with α, it can be seen that the improvement offered by the KCFE over the ecf based estimation technique appears to be similar for all the values of α tested.

In Table 2.5 the performance under different β is investigated when $\alpha = 1.5$. While the threshold in Table 2.4 was again set at 98%, in Table 2.5 three truncated regression regions have been used: the tabulated values and thresholds set at 80% and 98% of the asymptotic variance. Here it can be seen that a lower threshold of 80% improves performance for higher values of β.

n	cf estimator	α 0.8	1.0	1.2	1.4	1.6	1.8
50	ecf	15.03	16.64	20.22	27.59	39.14	55.58
	KCFE-t_3	11.32	14.09	15.55	20.90	25.39	35.18
200	ecf	3.70	3.77	4.52	6.33	10.97	28.18
	KCFE-t_3	2.94	2.97	3.90	4.96	7.30	15.16
400	ecf	1.73	2.11	2.23	3.27	4.52	14.68
	KCFE-t_3	1.43	1.67	1.83	2.79	3.59	8.75

TABLE 2.3. MSE ($\times 10^{-2}$) of estimates of β using tabulated regression region.

n	cf estimator	α 0.8	1.0	1.2	1.4	1.6	1.8
50	ecf	38.81	33.04	32.00	37.00	48.79	64.09
	KCFE-t_3	7.48	9.53	12.66	18.89	29.11	41.42
200	ecf	9.61	7.69	7.05	9.24	15.44	38.67
	KCFE-t_3	2.31	2.73	3.49	5.07	8.89	20.85
400	ecf	3.71	3.53	3.32	4.54	6.70	20.37
	KCFE-t_3	1.23	1.77	1.91	2.83	4.22	11.96

TABLE 2.4. MSE ($\times 10^{-2}$) of estimates of β using a variance threshold regression region.

regr. region	cf estimator	β 0	0.2	0.4	0.6	0.8
tabulated	ecf	7.70	7.55	7.63	6.02	4.39
	KCFE-t_3	5.90	5.76	5.98	5.64	5.17
80% of asymp. var	ecf	9.06	8.59	8.56	6.66	4.67
	KCFE-t_3	7.45	7.34	7.28	6.65	5.91
98% of asymp. var	ecf	10.52	10.82	10.18	9.03	6.84
	KCFE-t_3	5.93	6.31	6.37	6.60	6.52

TABLE 2.5. MSE ($\times 10^{-2}$) of estimates of β for 200 observations, $\alpha = 1.5$.

2.3.5 Conclusions

Using the KCFE and a t_3 kernel, improvements in the estimation procedure described in [50] for α were obtained for higher (true) values of α and in the estimation of β for smaller (true) values of $|\beta|$.

The results using calculated regression regions rather than tabulated values are more difficult to interpret and statements about their superior or inferior performance are difficult to make. However, generally, it has been shown that for estimating α, the performance of a 98% threshold is comparable to the tabulated regression regions. For the estimation of β, a smaller threshold (80%) was found to be better for the more skewed distributions, though still inferior in performance to the tabulated values.

It should not be expected that the variance threshold regression region will outperform the tabulated regression region since the tabulated values were found by extensive simulations to be optimal. However, while the variance threshold technique can be applied to any sample size and parameter values, the tabulated values are only available for a small number of settings.

2.4 Summary

An overview of the αS distribution as a model for impulsive processes has been presented, describing its definition, properties and parameter estimation techniques. Improvements to the latter, namely the use of the KCFE and of an automatic method of finding the regression region have been introduced to overcome some of the restrictions placed by current methods.

Chapter 3

TESTS FOR GAUSSIANITY

While it may be argued that strictly speaking "Normality is a myth, there never was and never will be a Normal Distribution" [27], the importance of the Gaussianity distribution is still unquestioned. Its unique properties make it the most important and most used in statistical signal processing.

The assumption that the distribution of a random process is indeed Gaussian may greatly simplify its analysis and lead to elegant, optimal solutions to signal processing problems. It is because of the need to justify this assumption that Gaussianity testing has received such attention and is of such importance.

In this chapter a brief summary of Gaussianity tests is presented, followed by a more detailed investigation of the merits and deficiencies in characteristic function (cf) based Gaussianity tests. Extension to the multivariate case is also performed and discussed.

3.1 Review

Due to the multitude of alternative distributions, there is no single "best" Gaussianity test. Each test utilises a particular property or feature of the distribution, and hence may be more sensitive to certain alternatives than other tests that use different properties. The aim of an omnibus test is to provide acceptably high rejection rates against all reasonable non-Gaussian distributions.

Many tests exist that test for Gaussianity of a stationary process where the data can be assumed to be independent and identically distributed (iid). These include Pearson's χ^2 test, the Shapiro-Wilk test [74] and the D'Agostino test [67]. Others, including bispectrum and trispectrum [34, 21, 80], entropy [40] and moment and cumulant based tests can be used on correlated data. Some of these tests are described below.

3.1.1 Chi-square (χ^2) tests

Chi-square (χ^2) tests have had a long history in Gaussianity testing and were used by Karl Pearson in 1900. Since then they have been widely adapted and used. After grouping the observations into bins or cells, the number of observations falling into each cell is compared to the number expected under the Gaussian distribution.

As information is discarded when observations are grouped, it is generally felt that these tests should only be used only when complete random samples are unavailable, for example when the data is censored or truncated [20, 62]. The effect of grouping is especially problematic for small sample sizes.

3.1.2 Cumulant and Moment based Gaussianity tests

The standardised third and fourth order moments of a Gaussian distribution, the skewness and kurtosis respectively, are [11]

$$\sqrt{b_1} = \frac{\mu_3}{(\mu_2)^{3/2}} = 0 \tag{3.1}$$

$$b_2 = \frac{\mu_4}{(\mu_2)^2} = 3 \tag{3.2}$$

where μ_k indicates the kth order central moment. Another class of tests compares the moments (or cumulants) of the observed process to that of a Gaussian random variable [20].

While any process that has $\sqrt{b_1} \neq 0$ or $b_2 \neq 3$ is certainly not Gaussian, it does *not* necessary imply that if $\sqrt{b_1} = 0$ and $b_2 = 3$ that the process is Gaussian. In other words, the values of the higher order standardised moments in (3.1) and (3.2) are necessary but not sufficient for the Gaussian distribution. It would be necessary to test *all* higher order moments, rather than just third and fourth orders. However, in practice, this is not feasible.

3.1.3 Empirical distribution function tests

An intuitively obvious way to test for the goodness-of-fit to the Gaussian distribution is to use the difference between the empirical distribution function (edf), $F_n(x)$ and the cumulative distribution function (cdf), $F(x)$. A large number of tests exist that use different measures of this difference function. The most widely known is the Kolmogorov-Smirnov test which is based on the statistic [79]

$$\begin{aligned} D &= \max_x |F_n(x) - F(x)| \\ &= \max\left(D^+, D^-\right) \quad , \end{aligned}$$

where $D^+ = \sup_X\{F_n(x) - F(x)\}$ and $D^- = \sup_X\{F(x) - F_n(x)\}$ are the largest positive and negative differences between the edf and cdf. This statistic is within the class of supremum statistics. A variation is the Kuiper statistic [79]

$$\begin{aligned} V &= \max_x\{F_n(x) - F(x)\} + \max_x\{F(x) - F_n(x)\} \\ &= D^+ + D^- \quad . \end{aligned}$$

Another class of tests integrate a weighted function of the square of the difference function to yield quadratic statistics of the form

$$Q = n\int_{-\infty}^{\infty} \{F_n(x) - F(x)\}^2\,\psi(x)dF(x) \quad .$$

The Cramér-von Mises,

$$\psi(x) = 1 \quad ,$$

and Anderson-Darling,

$$\psi(x) = \left[\{F(x)\}\{1 - F(x)\}\right]^{-1} \quad ,$$

statistics are perhaps the best known of this class of tests.

These difference measures result in tests which are sensitive to different alternatives, for example, while the Kolmogorov-Smirnov statistic can detect changes around the median well, its performance in the tails of the distribution is poor. By contrast, the weighting function of the Anderson-Darling statistic allows more powerful detection of deviations in tails – these points are confirmed by simulations in Chapter 4.

3.2 Characteristic Function Based Techniques

The characteristic function (cf) has been investigated by Koutrouvelis [49] and Epps [24], and more recently in [93], as the basis for developing Gaussianity tests. Some of these techniques are summarised here and related alternatives are proposed. Definitions and estimators of the cf are described in Appendix A.

The cfs for the Gaussian, Uniform and Laplace distributions are shown in Figure 3.1. All have been standardised to have zero mean and unit variance. Since their pdfs are real and symmetrically distributed about the origin, their cfs are too.

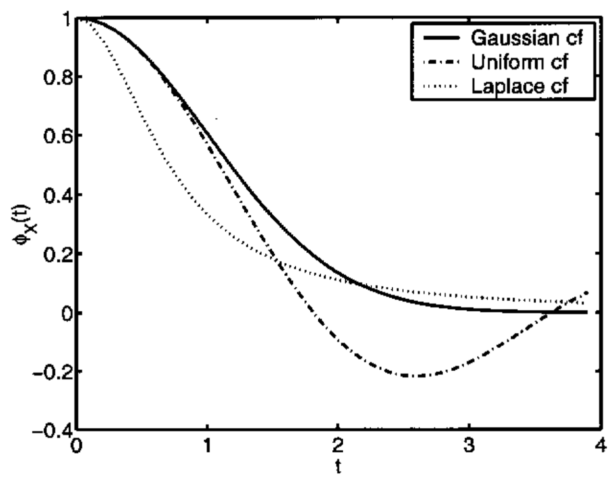
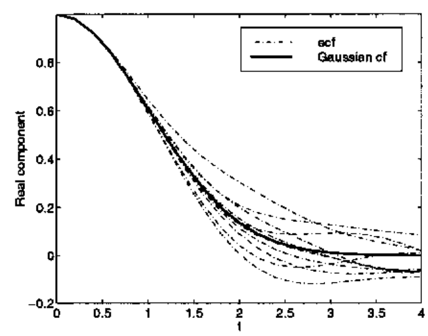


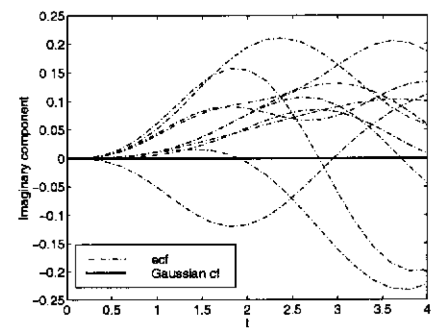
FIGURE 3.1. The cfs of the Gaussian, Uniform and Laplace distributions.

Like the pdf, the cf completely characterises a random variable. Therefore, by analogy with empirical distribution function tests, differences between an estimate of the cf based on the observations and the Gaussian cf may be used as the basis of a Gaussianity test. The relationship between the cf and pdf will be used again later as the motivation behind operations in the cf domain.

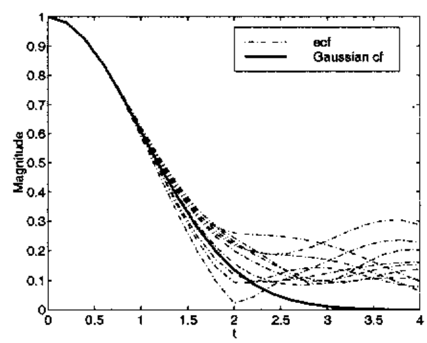
In Figure 3.2 are shown the ecfs calculated from 10 independent realisations of Gaussian data, each with $n = 64$ observations, along with the true Gaussian cf. The fluctuations between the realisations at higher values of t are clearly evident. A poorly chosen test statistic that uses the ecf may be unable to distinguish between systematic changes to the cf of the generating process and these fluctuations that are present even under the null hypothesis of Gaussianity, H.



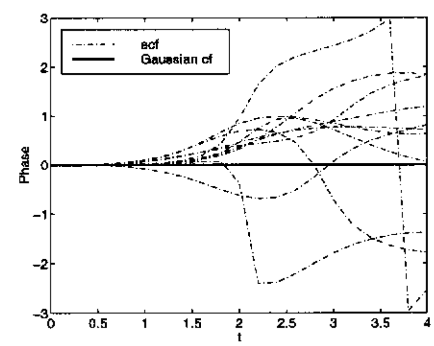
Real component.



Imaginary component.



Magnitude.



Phase.

FIGURE 3.2. The ecfs of 10 realisations of $n = 64$ observations of iid standard Gaussian data.

3.2.1 Kernel Characteristic Function Estimator

Despite the fluctuations in the ecf demonstrated above, until recently, cf based Gaussianity tests have concentrated on the use of the ecf as a cf estimator. In [93] the problems with using the ecf as a characteristic function estimator in goodness-of-fit tests, namely its non-vanishing variance, were highlighted. It was proposed that the fixed kernel characteristic function estimator (KCFE), as described in Section A.2, be used in place of the ecf since its variance can be controlled by the kernel function.

3.2.2 Test statistics

Finding a meaningful single measure of the difference between the estimated cf and the Gaussian cf is very important to the performance of cf based Gaussianity tests. The cf of every distribution is unique, and therefore the difference between it and the Gaussian cf will also be unique. The chosen statistic, along with the chosen set of t values at which to evaluate the statistic, must take into account a wide range of these distinct cfs in order to produce an omnibus Gaussianity test.

The real and imaginary components of the characteristic function contain different information about the distribution of the variable. A variable symmetrically distributed around the origin has a real cf. Any imaginary component in the estimate of the cf of such a variable will be due to finite sampling only and not to the distribution of the variable.

Illustrating this, Figure 3.3 shows 20 realisations of the ecfs of Gaussian ($\times$), Uniform ($\circ$) and χ_2^2 ($*$) distributed variables, standardised to have zero mean and unit variance, evaluated at $t = 1.8$ and plotted on the complex plane. Both the Uniform and Gaussian ecf values are scattered around the real line. Also, the Gaussian ecf values are centred near their true value, shown by a $+$ symbol, at $(0.1979, 0)$. The χ_2^2 ecf values, being from an asymmetric distribution, appear to have significant imaginary components.

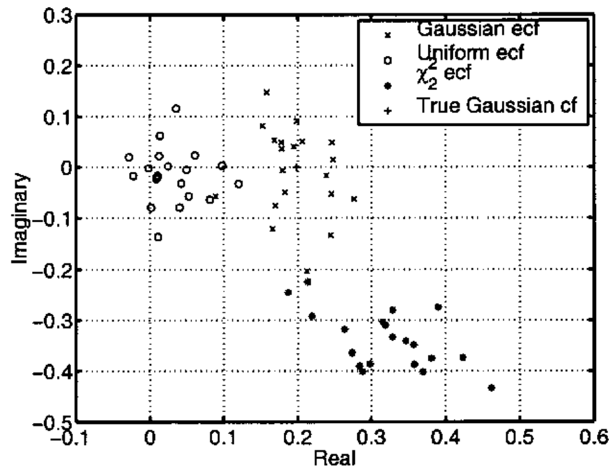


FIGURE 3.3. Realisations of the ecfs of Gaussian, Uniform and χ_2^2 distributed data and the true Gaussian cf evaluated at $t = 1.8$, plotted on the complex plane.

Consequently, using a test statistic based on only the real component of the cfs will be powerful against symmetric alternatives, but less powerful against asymmetric alternatives which have complex characteristic functions. One such statistic that was found to perform well [93] is

$$Q_{\mathbf{X}}^{\Re} = \max_t |\Re(\hat{\varepsilon}_{\mathbf{X}}(t;\varphi))| \tag{3.3}$$

where

$$\begin{aligned} \hat{\varepsilon}_{\mathbf{X}}(t;\varphi) &= \hat{\phi}_{\mathbf{X}}(t;\varphi) - \phi_0(t)\varphi_{\mathbf{X}}(t) \\ &= \left[\hat{\phi}_{\mathbf{X}}^{e}(t) - \phi_0(t)\right]\varphi_{\mathbf{X}}(t) \end{aligned} \tag{3.4}$$

is the product of a kernel function and the difference between the ecf and the cf under H, $\phi_0(t)$, that is, the Gaussian cf. The kernel smoothing can also be viewed as a weighting function, that is, more significance is given to differences between the cfs in the region where the variance of the estimator is low. An alternative statistic based on the imaginary component, such as

$$Q_{\mathbf{X}}^{\Im} = \max_t |\Im(\hat{\varepsilon}_{\mathbf{X}}(t;\varphi))| \tag{3.5}$$

will have little or no power against symmetric alternatives.

When the symmetry of the variable's distribution cannot be assumed, using the information in both the real and imaginary parts of the cf is necessary. This motivates the use of another statistic based on the magnitude of the smoothed difference function,

$$Q_{\boldsymbol{X}} = \max_t \mid \hat{\varepsilon}_{\boldsymbol{X}}(t;\varphi) \mid . \tag{3.6}$$

By incorporating the imaginary component of the cf, the $Q_{\boldsymbol{X}}$ based test will have slightly less power than the $Q_{\boldsymbol{X}}^{\Re}$ based test for symmetric alternatives, since an additional source of variation is introduced without the addition of any distributional information. However, against asymmetric alternatives, the imaginary component *is* adding information to the statistic, hence the $Q_{\boldsymbol{X}}$ based test is expected to be more powerful than the $Q_{\boldsymbol{X}}^{\Re}$ based test.

Several alternative test statistics present themselves for investigation, the most prominent of which are those based on the integrated squared error of the smoothed difference function, however, in [93] the statistics in (3.3) and (3.6) were recommended.

As yet, the distribution of the $Q_{\boldsymbol{X}}^{\Re}$ and $Q_{\boldsymbol{X}}$ statistics is unknown. Threshold setting is performed empirically by evaluating the statistics a large number of times under H.

3.2.3 Choice of t values

As mentioned in Section 3.2.2, the choice of values of t at which to evaluate the test statistic can be crucial to the power of the test against a particular alternative. The cfs of some non-Gaussian distributions may be very close to the Gaussian cf at *some* values of t, while vastly different at others. The values used must be able to provide good results against a large number of alternatives.

If increased sensitivity against a particular distribution were required, it would be possible to determine at what points the cfs of the distribution and the Gaussian cf differ the most and evaluate them only at these points. This would be particularly effective for test statistics that measure, for example, the mean squared error of the difference function. Evaluation at a small number of points also makes the test easier to implement and quicker to run.

Since the test statistic of choice here is a supremum statistic, the difference function should be evaluated at a fine spacing and should cover the range of values where the kernel function has significantly large values.

3.3 Results and Discussion

Monte-Carlo simulations were run to evaluate the performance of the proposed method and are shown in Table 3.1. Critical points or thresholds for the tests were obtained empirically from the distributions of the test statistics when the input data was known to be Gaussian.

Each test was run 1000 times for each distribution at the 5% level of significance, with t taking values between 0 and 5, in steps of 0.05. Samples contained $n = 64$ data points and were standardised by their respective means and variances. Standardisation was necessary as the cf is dependent on mean and variance. Omitting this procedure would result in testing for departure from a standard Gaussian distribution, $N(0,1)$, rather than any Gaussian distribution. Some of the non-Gaussian distributions considered are U: Uniform, L: Laplace, K: K-distribution, LN: Log-Normal.

Now comparing between the different test statistics in Table 3.1, as expected, it can be seen that the tests based on statistics of the real component of the differences, $Q^{\Re}_{\mathbf{X}}$, achieve higher power than the tests based on statistics of the magnitude of the differences, $Q_{\mathbf{X}}$ in the cases of $U(0,1)$, $\beta(4,4)$ and Laplacian distributions – all of

Distribution	Q_X	$Q_X^{\Re}$	$Q_X^{\Im}$
$N(0,1)$	5.0	5.0	5.0
$U(0,1)$	93.5	96.4	7.6
χ_2^2	99.6	23.6	99.9
χ_8^2	58.8	5.3	67.8
Laplace	30.7	31.8	11.8
$K(1,1)$	84.4	7.6	88.4
$LN(0,1)$	100.0	75.6	100.0
$\beta(4,4)$	16.0	22.5	3.3

TABLE 3.1. Rejection rate (in %) for the KCFE based test when testing at the 5% level.

which are symmetric.

In the $\beta(4,4)$ case no test is able to achieve high power against this *Gaussian-like* distribution. Against the other, asymmetric distributions, the $Q_X^{\Re}$ based test performs significantly worse than the omnibus Q_X based test.

The $Q_X^{\Im}$ based test also appears to be operating as expected and described in Section 3.2.2. That is, it is ineffective against the symmetric alternatives; $U(0,1)$, Laplace and $\beta(4,4)$, but effective against the asymmetric alternatives.

The simulations were also run for tests that used statistics not based on the peak or maximum difference between functions, but using both the mean and MSE of the difference functions. These results are shown in Table 3.2.

From this table, it can be seen that the peak absolute difference provides the greatest power for both methods. This confirms the choice of test statistic described in Section 3.2.2 and used in [93].

When symmetry of the data can be assumed, the $Q_X^{\Re}$ based test provides superior power. If no knowledge about the distribution of the data is available, the Q_X based test will generally have higher power. Of course if the alternative is known to be asymmetric, simple tests for symmetry would be considered.

It should be noted that these methods are not restricted to testing for Gaussianity,

Distribution	Q_X	mean	MSE
$N(0,1)$	5.0	4.9	5.0
$U(0,1)$	80.0	57.8	74.6
χ_2^2	99.9	100.0	100.0
χ_8^2	67.3	53.3	65.8
Laplace	63.8	55.6	66.1
$K(1,1)$	89.4	81.2	89.0
$LN(0,1)$	100.0	100.0	100.0
$\beta(4,4)$	8.5	6.6	8.6

TABLE 3.2. Rejection rate (in %) for the KCFE based test when testing at the 5% level, comparing the statistics based on the peak absolute difference, mean difference and MSE.

and may be used to test for any known distribution. As they rely on determining how closely the ecf, derived from the data, matches the cf of the distribution being tested, only the analytic form of the desired cf is required. This feature will be utilised in the following chapter.

3.4 Multi-dimensional Gaussianity Testing

Tests for Gaussianity using the characteristic function (cf) have been primarily developed for independent and identically distributed (iid) data only. It is necessary to address the issue of testing correlated data for Gaussianity using the cf in order to apply it to the many real sources of data where correlation is known to exist. Here, some multidimensional tests for Gaussianity are formulated [96].

Other tests for multivariate Gaussianity have generally lagged similar univariate tests [63] due to an increase in complexity and have required large numbers of samples (in the thousands) [19]. Unlike some other techniques, Gaussianity tests based on the cf are readily extendable to the multivariate case and achieve high power, while not being limited to large samples.

Consider a vector valued random variable

$$\boldsymbol{X} = [\boldsymbol{X}_1, \boldsymbol{X}_2, \dots, \boldsymbol{X}_m]$$

where $\boldsymbol{X}_k = [X_{1,k}, X_{2,k}, \dots, X_{n,k}]^T$ is the vector of n observations on the kth variable for $k = 1, 2, \dots, m$. Assume correlation exists between the $\boldsymbol{X}_1, \boldsymbol{X}_2, \dots, \boldsymbol{X}_m$. The multivariate ecf may be given by

$$\hat{\phi}^e_{\mathbf{X}}(t_1, t_2, \dots, t_m) = \frac{1}{n}\sum_{i=1}^{n} \exp\left(j\sum_{k=1}^{m} t_k X_{i,k}\right)$$

The smoothed difference between the multidimensional ecf and the cf under H is

$$\begin{aligned}&\hat{\varepsilon}_{\mathbf{X}}(t_1, t_2, \dots, t_m; \varphi_{\mathbf{X}}) = \\ &\qquad \hat{\varepsilon}_{\mathbf{X}}(t_1, t_2, \dots, t_m)\varphi_{\mathbf{X}}(t_1, t_2, \dots, t_m)\end{aligned}$$

where

$$\begin{aligned}\hat{\varepsilon}_{\mathbf{X}}(t_1, t_2, \dots, t_m) &= \hat{\phi}^e_{\mathbf{X}}(t_1, t_2, \dots, t_m) \\ &\quad -\phi_{\mathbf{X}}(t_1, t_2, \dots, t_m)\end{aligned}$$

and $\varphi_{\mathbf{X}}(t_1, t_2, \dots, t_m)$ is a multidimensional kernel function.

The multidimensional kernel function related to the univariate kernel function in (A.5) is

$$\varphi(t_1, t_2, \dots, t_m) = \exp\left(\frac{-\sigma_\varphi^2 \sum_{k=1}^{m} t_k^2}{2}\right) \quad . \tag{3.7}$$

3.4.1 Results and discussion

The power of the test for the bivariate case is investigated through Monte Carlo simulations. The results are presented in Table 3.3, for $n = 20$ data points, and Table 3.4, when 50 data points are used. The power of the proposed test is compared to bivariate tests based on an adaptation of the Shapiro Wilk W test and Mardia and Foster's S_W^2 test [63]. The bivariate chi-square distributions, $\chi_2^2(\nu_1, \nu_2; \nu)$, are the joint distributions of W_1 and W_2 where $W_i = V_i + V$. Each V_i is an independent χ^2 variate with ν_i degrees of freedom and V is a χ^2 variate with ν degrees of freedom, independent of all the V_i.

Distribution	Q_X	W	S_W^2
$\chi_2^2(2,2,1)$	63.3	75.6	49.4
$\chi_2^2(2,2,2)$	51.8	59.2	42.4
$\chi_2^2(2,2,3)$	43.1	48.1	35.4
t_3	54.8	47.6	48.6
Cauchy	97.6	96.9	94.9
$0.8N_2(0,\Sigma) + 0.2N_2(0,9\Sigma)$	60.9	44.3	57.0
$0.9N_2(0,\Sigma) + 0.1N_2(0,16\Sigma)$	59.5	52.3	61.9
Log Normal	89.0	87.6	70.2

TABLE 3.3. Rejection rate (in %) for the KCFE, Shapiro Wilk W test and Mardia and Foster's S_W^2 test for bivariate Gaussianity at the 5% level of significance for $n = 20$.

Distribution	Q_X	W	S_W^2
$\chi_2^2(2,2,1)$	96.2	99.7	95.3
$\chi_2^2(2,2,2)$	90.2	97.3	88.6
$\chi_2^2(2,2,3)$	81.7	91.2	83.0
t_3	90.7	75.5	86.3
Cauchy	100	100	100
$0.8N_2(0,\Sigma) + 0.2N_2(0,9\Sigma)$	92.5	78.1	93.6
$0.9N_2(0,\Sigma) + 0.1N_2(0,16\Sigma)$	89.2	85.8	93.1
Log Normal	99.8	100	99.4
Average	92.5	91.0	92.4

TABLE 3.4. Rejection rate (in %) for the KCFE, Shapiro Wilk W test and Mardia and Foster's S_W^2 test for bivariate Gaussianity at the 5% level of significance for $n = 50$.

For this investigation, 5000 replications of each test were performed. The level of significance was set at 5% and thresholds were derived empirically using bivariate

Gaussian data with marginal means of zero and covariance matrix $M = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}$.

All data was standardised by marginal means and variances, then the correlation between the two variables, $\hat{M}$, was estimated. The bivariate Gaussian cf, with correlation matrix $\hat{M}$, was then evaluated. The difference between this cf and the magnitude of the ecf was then smoothed by the Gaussian kernel function described in (3.7). Taking the magnitude of the mean of this function produced the test statistic $Q_{\boldsymbol{X}}$ which was compared to the empirically derived threshold.

The results in Tables 3.3 and 3.4 show an increase in average power by the smoothed difference cf based test, $Q_{\boldsymbol{X}}$, against both the W and S_W^2 tests. Each test has *some* distributions against which it is more powerful, however, none of the tests performs best for *all* distributions.

An important feature of the performance of the $Q_{\boldsymbol{X}}$ test is that of the results shown, only once was it the worst performing of the three tests – this was for the $\chi_2^2(2, 2, 3)$ distribution when $n = 50$. Both the W and the S_W^2 tests had their relative strengths and weaknesses, however, the $Q_{\boldsymbol{X}}$ test performed consistently well. For example, the W test was the best performing for all the bivariate χ^2 distributions, while the S_W^2 test generally achieved higher rejection rates for the other distributions.

The difference in relative performance of the $Q_{\boldsymbol{X}}$ test did not appear to be affected by a change in the number of data points used, n.

Several other features of cf based tests make them appealing compared to the other tests available and mentioned. Cf based tests can easily be adapted to test for any known, fixed distribution, and their sensitivity to certain alternatives can be adjusted through the choice of test statistic and t values. However, the problem of dimensionality may mean practical implementation of multivariate testing may be difficult – the dimensionality of the t-space occupied by the cf increases computation and storage requirements.

3.4.2 Conclusions

The KCFE based Gaussianity test has been described and extended to the multivariate case, allowing it to be used on correlated data. An empirical power study and comparison to other tests revealed encouraging results. Further studies into the performance of this test are certainly possible, as well as detailed investigation into optimisation of some aspects of the test, namely, the choice of a test statistic and the range of t values at which to evaluate the cfs. The distribution of the test statistic under H is difficult to find and is currently unknown. For the purposes of testing, empirically derived thresholds were used here.

3.5 Summary

A brief summary of current Gaussianity testing techniques was presented, with special emphasis on characteristic function (cf) based techniques. While there are several reasons for emphasising these tests in their own right, including the development here of multidimensional Gaussianity testing, another reason for the inclusion of cf based testing will become apparent in the following chapter.

Gaussianity testing is the important first step in Figure 1.1. If it can be assumed that the interference process is Gaussian, optimal solutions to the signal detection, and other signal processing problems, already exist in the literature. The remainder of this dissertation will deal with problems when Gaussianity cannot be assumed.

Chapter 4

ALPHA-STABLE GOODNESS-OF-FIT TESTS

The presence of *and* degree of impulsive behaviour in observed processes is often important in optimising performance. For example in detection problems, it is known that the optimal detector for additive Gaussian interference, the matched filter, will perform poorly when the interference is alpha-stable (αS) distributed with $\alpha < 2$. If tests reveal impulsive behaviour, significant improvements in detector performance may be achieved by using a detector designed for that particular type and level of impulsive interference or by using a robust, nonparametric detector that can operate in the presence of impulsive interference [18, 17]. Additionally, it may be argued that if the degree of impulsive behaviour is low, then the additional burden of dealing with complicated models such as the αS distribution may not be justified, and other distributions, such as the Gaussian or Generalised Gaussian may be sufficient. For these reasons, tests are described in this chapter to determine both the presence and the degree of impulsive behaviour in αS processes.

In Section 4.1, a goodness-of-fit test is described for determining the appropriateness of the assumption that observations come from an αS distribution, a broad family of distributions that have been used to model some impulsive sources. Following this, in Section 4.2, another test is presented for testing the level of impulsive behaviour within the αS family.

4.1 Testing for the αS Distribution

The goodness-of-fit test described here is a test for the entire family of stable distributions, against all other distributions, and not just a test for the non-Gaussian stable distributions. That is, the Gaussian distribution is included in the null hypothesis.

The test was first proposed in [13] and was used in [69], but is expanded here.

4.1.1 Method

Differences between the characteristic functions (cf) of random variables have been used extensively to test for changes in distributions. This has already been discussed in the context of Gaussianity goodness-of-fit testing, where the technique originated – refer back to Chapter 3. Much of the motivation for investigating cf based techniques has been its Fourier relationship with the more commonly employed, and more intuitive, probability density function. Tests and procedures developed for the latter can often be transformed to equivalent operations for the former.

Aside from this useful relationship, it has also frequently been noted that one of the major advantages of cf based goodness-of-fit testing is that the cf always exists and is smooth and well behaved, unlike the pdf. However, it still retains the advantages of pdf-based goodness-of-fit testing, mainly that it can be adapted to test for almost any distribution, as long as the cf is specified.

Noting the presence of general closed-form expressions for the αS cf but not the pdf, this lead to proposals for its application to testing for the αS distribution, in [51] and, more recently, in [13].

While, as shown in Section 3.4 and [96], cf based goodness of fit techniques can be easily extended to multivariate distributions, here emphasis is on univariate distributions, and more specifically, independent and identically distributed random variables.

Put simply, this test estimates the significance of the difference between the cf of the generating process, estimated nonparametrically from the data, and the cf of the "best fitting" αS process. The method is shown in block diagram form in Figure 4.1. A detailed explanation follows.

It is assumed that the data is independent and identically distributed (iid) with unknown, but fixed parameters. The hypothesis, H, is that the data is αS distributed

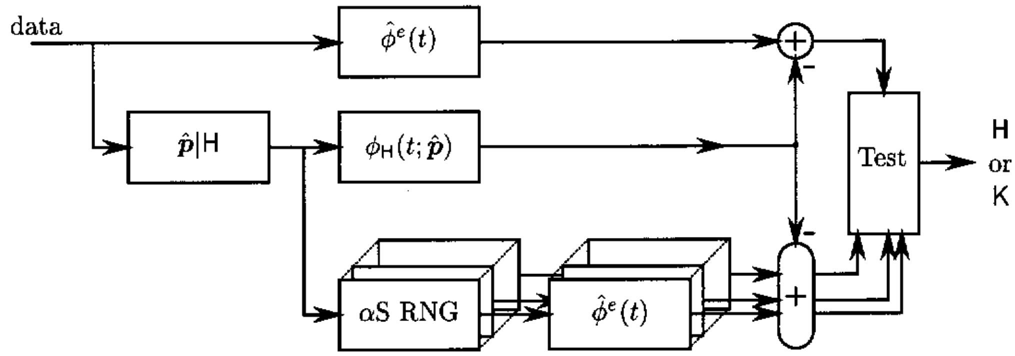


FIGURE 4.1. Block diagram of the αS goodness-of-fit test method.

against the alternative, K, that it is not. This alternative includes *all* non-stable distributions. The characteristic function (cf) of an αS process was given previously in (2.1) and the parameter vector will be denoted here by $\boldsymbol{p} = [\alpha\, \beta\, c\, \delta]^T$.

1. Estimate the cf of the process from the observations using the empirical characteristic function (ecf), $\hat{\phi}^e(t)$.

2. Estimate the αS parameters of the observations under the assumption that H holds, $\hat{\boldsymbol{p}}|\mathsf{H}$. This can be done through a variety of techniques, however, as discussed in Section 2.3, the cf based parameter estimation procedure proposed in [50] appears to have good consistency properties compared to other methods and is used here.

3. Evaluate the cf of the αS distribution having the estimated parameter values, $\phi_{\mathsf{H}}(t;\hat{\boldsymbol{p}})$.

4. Find the difference function $\phi_{\mathsf{H}}(t;\hat{\boldsymbol{p}}) - \hat{\phi}^e(t)$.

5. Test the significance of this difference.

By using the estimated parameter values $\hat{\boldsymbol{p}}|\mathsf{H}$ and the form of the αS cf, the cf generated in step 3 is a *parametric* estimate of the cf of the generating process, assuming H holds. Therefore, by subtracting the ecf, an unbiased *nonparametric* cf estimator, the difference function will give an indication of the validity of H. If the observed data is αS distributed, then the difference between these functions will be due to parameter estimation errors and variation in the ecf due to the finite sample size. However, under the alternative, K, there will be a systematic difference, which, for reasonable alternatives, should be larger than the random variation described above.

Drawing on the findings in [93] and Chapter 3 for cf based Gaussianity tests, the following test statistics are considered:

$$T_a = \max_t \left| \phi_{\mathsf{H}}(t; \hat{\boldsymbol{p}}) - \hat{\phi}^e(t) \right| \tag{4.1}$$

$$T_r = \max_t \left| \Re\, \phi_{\mathsf{H}}(t; \hat{\boldsymbol{p}}) - \Re\, \hat{\phi}^e(t) \right| \quad . \tag{4.2}$$

It was concluded in Section 3.2.2 that the statistic using the greatest magnitude difference provided the best omnibus test, here this is the T_a statistic. However, since the cf of a symmetric distribution is real, it is expected the T_r statistic will be superior when symmetry can be assumed. This is significant in the study of αS distributions due to the emphasis, in some areas, towards symmetric members of the αS family.

As the distributions of these test statistics is complicated and unknown, the parametric bootstrap [23, 94] is used to determine the significance of the statistics. To do this, multiple realisations of αS random data are generated using a random number generator (RNG) with parameters $\hat{\boldsymbol{p}}|\mathsf{H}$ and the corresponding ecfs and test statistics are calculated.

By using the regression-type procedure in the cf domain to estimate the parameters of the αS distribution in step 2, the cf of the generated αS bootstrap resamples is "fitted" to the data's ecf. That is, by some (unspecified) criterion, the cf of the resamples is the αS cf that best matches or approximates the cf of the data. Therefore, the distribution of these bootstrap statistics approximate the distribution of the test statistics under H and can be used to set thresholds. More details on the parametric bootstrap as a method to estimate distributions of test statistics can be found in Appendix B.

4.1.2 Simulation Results and Discussion

To evaluate the power of this test against various alternatives, the following simulations were performed for data lengths of 400 and nominal false alarm rate of 5%. These values were chosen as this allows the resulting rejection rates to be distinguishable – if too few or too many observations are used, all distributions yield similarly low or high rejection rates. The cfs where evaluated between $t = -10$ and $t = 10$ in steps of 0.1 and the number of bootstrap resamples was 500. As well as T_a and T_r, the Kolmogorov-Smirnov, T_{KS}, and Anderson-Darling, T_{AD}, tests, as defined in Section 3.1.3, are included for comparative purposes. However, due to the difficulty in numerically approximating the pdf of an asymmetric αS distribution, the T_{KS} and T_{AD} tests only test for the symmetric αS distribution, therefore, results for asymmetric distributions should be viewed with caution.

From Table 4.1 it can be seen that the false alarm rate has been maintained for the T_a and T_r tests. High detection rates are achieved for most of the non-αS distributions in Table 4.2. Exceptions to this are the t distributions. The t_8 distribution has enough degrees of freedom to appear almost Gaussian, and hence is often accepted by the tests.

Similarly, for lower degrees of freedom the difference between the cf of the t_2 distribution and the cf of the $\alpha\text{S}(\alpha = 1.4443, \beta = 0, c = 0.8668, \delta = 0)$ distribution is very small, peaking at only 0.0307, as shown in Figure 4.2. In addition, most of this difference is in the high variance region of the ecf and hence may be masked by the variability of the ecf. Only for very long data lengths will the distributions be distinguishable.

To illustrate this, a variant on the proposed test was run. Two tests were run

Distribution	T_a	T_r	T_{KS}	T_{AD}
αS(1.2,0)	6.0	5.0	68.2	0.6
αS(1.2,-0.3)	5.6	4.8	96.0	17.8
αS(1.7,0)	4.8	5.8	7.0	5.0
αS(1.7,-0.3)	4.6	4.6	19.6	17.4
αS(1.7,0.7)	5.2	5.6	66.4	69.2
αS(1.9,0)	5.0	4.6	0.8	3.8
αS(2,0)	3.8	4.4	0.4	4.6

TABLE 4.1. Rejection rates of the αS goodness-of-fit test for various alpha-stable(α, β) distributions (in %) using sample sizes of 400, 5% nominal level and 500 replications.

Distribution	T_a	T_r	T_{KS}	T_{AD}
$U(0,1)$	100.0	100.0	21.2	100.0
χ_2^2	100.0	99.8	100.0	100.0
χ_8^2	90.0	73.8	100.0	100.0
Laplace	34.2	43.8	24.6	9.6
Log Normal	76.2	77.2	100.0	93.8
t_2	4.8	5.2	39.2	4.0
t_3	5.8	5.0	12.6	5.8
t_8	3.2	5.4	1.8	4.8

TABLE 4.2. Rejection rates of the αS goodness-of-fit test for various non-stable distributions (in %) using sample sizes of 400, 5% nominal level and 500 replications.

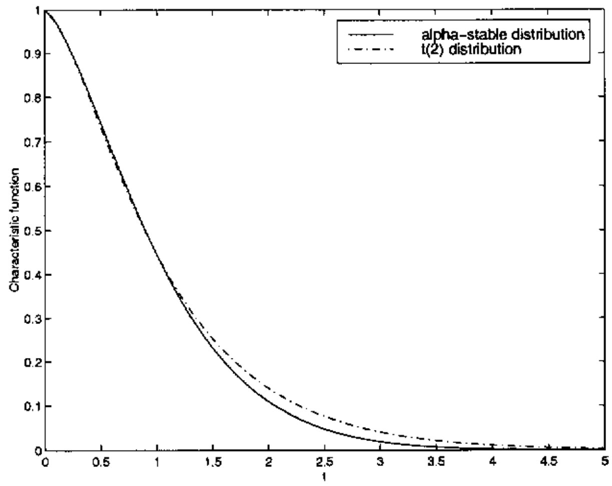


FIGURE 4.2. The cfs of the t_2 and $\alpha S(\alpha = 1.4443, \beta = 0, c = 0.8668, \delta = 0)$ distributions.

that tested that the observations are $\alpha S(\alpha = 1.4443,\ \beta = 0,\ c = 0.8668,\ \delta = 0)$ and t_2 distributed, respectively. The rejection rates indicate the approximate likelihood that the t_2 fit is better than the αS. As shown in Figure 4.3, it is difficult to distinguish between the distributions unless large amounts of data are available.

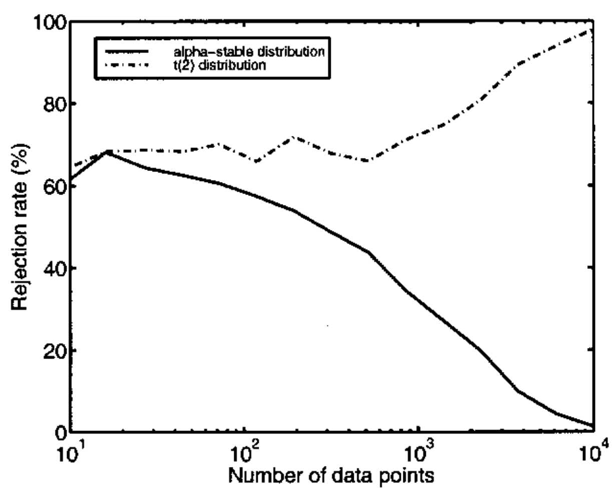


FIGURE 4.3. Rejection rates (in %) for the $\alpha S(\alpha = 1.4443,\ \beta = 0,\ c = 0.8668,\ \delta = 0)$ versus t_2 distribution goodness-of-fit test.

4.1.3 Conclusions

A characteristic function based test has been developed for testing the appropriateness of the assumption that data is αS distributed. The false alarm rate is able to be maintained through the use of a parametric bootstrap procedure. High rejection rates have been achieved against a number of alternative distributions. Although low rates were observed against some distributions, most notably the t_2, it is noted that the difference between these distributions and an αS distribution with appropriately chosen parameter values, is small.

Later, in Chapter 5, this test is applied to some real sources of impulsive data where it has been proposed that the αS distribution may be appropriate for modelling.

4.2 Testing the Level of Impulsive Behaviour

As noted in the introduction to this chapter, acknowledging the mere presence of impulsive behaviour may not be enough – it may be that its level is also important. This is associated with the likelihood of observing an outlier or "spike". For the αS distribution, this is determined by the characteristic exponent, α. In Section 2.1 and Figure 2.3 it was shown that even non-Gaussian ($\alpha < 2$) αS distributions appear very Gaussian around the median, it is really only in the tails that differences are observable.

Two techniques are presented that test the level of impulsive behaviour of αS processes based on the work in [95, 14]: testing the parameter α directly and a characteristic function (cf) based technique. The parametric bootstrap is used in both cases to estimate the distribution of the test statistics and in the setting of critical values. A description of the use of the bootstrap for hypothesis testing relevant to the tests to be described in this section, is provided in Appendix B.

These are not parameter estimation procedures, rather, they test their values. The procedures here are applicable to the testing of all four parameters of the αS distribution, however here the focus will be on testing α only, including the special case of testing for the Gaussian distribution ($\alpha = 2$) against non-Gaussian, stable distributions.

4.2.1 Testing the α parameter

Given the definition and properties of the αS distribution presented in Chapter 2, testing for the level of impulsive behaviour of αS distributed data may be considered in terms of testing the parameter α. Hence adopting the notation of Appendix B, the hypothesis and alternative formulation used in (B.3) and (B.4) may be rewritten in terms of the sole parameter of interest, that is $\boldsymbol{\xi} = \alpha$,

$$\mathsf{H} : \alpha = \alpha_0 \tag{4.3}$$

versus the two-sided alternative

$$\mathsf{K} : \alpha \neq \alpha_0 \quad . \tag{4.4}$$

An obvious test statistic is

$$T_{n,\alpha} = \frac{|\hat{\alpha} - \alpha_0|}{\hat{\sigma}_{\hat{\alpha}}} \tag{4.5}$$

where $\hat{\alpha}$ is an estimator of α derived from the observations and $\hat{\sigma}_{\hat{\alpha}}$ is an estimator of its standard deviation. Under H, $T_{n,\alpha}$ will be small, only deviating from zero due to finite sample characteristics and the variability of the estimator. This statistic would be modified to

$$T_{n,\alpha} = \frac{\hat{\alpha} - \alpha_0}{\hat{\sigma}_{\hat{\alpha}}}$$

if the alternative, K,

$$\mathsf{K} : \alpha < \alpha_0 \tag{4.6}$$

is one-sided. While the emphasis here will be on the two-sided alternative, in some applications a one-sided alternative may be appropriate, for example if testing to see if the level of impulsive behaviour exceeds a certain level.

Fitting this test statistic to the parametric bootstrap procedure in Table B.1, the parameter of interest, $\boldsymbol{\xi}$, is α and the nuisance parameter is $\boldsymbol{\vartheta} = [\beta\ c\ \delta]^T$.

In [50] it was shown that the estimator of α depends on the true values of α, β and c. Since β and c are elements of $\boldsymbol{\vartheta}$, the nuisance parameter will affect the test statistic and justifies the need to re-estimate the distribution of $T_{n,\alpha}$ for each $\hat{\boldsymbol{\vartheta}}$.

The conditions listed in Table B.2 from [7] are required to prove the asymptotic correctness of the test and that the power function may be approximated by a bootstrap procedure. They concern the convergence of the parameter estimates and of the distribution of the test statistic [71]. Given the consistency and asymptotic unbiasedness of the αS parameter estimators (as shown in [50]) and the simple form of the test statistic, it is assumed that these conditions are met.

Determining critical values for this test is made more complicated by the presence of nuisance parameters, β, c and δ, and in particular by their effect on the estimation of the parameter of interest, α.

While δ does not affect the estimation of α or of c, it does affect the estimator of β. In the estimation of β, no extra computational burden is added by the estimation of δ. Therefore, δ is included in the test structure since it adds generality to the procedure should it be adapted for the case when β is a parameter of interest, rather than a nuisance parameter.

The proposed parametric bootstrap based procedure for testing the α parameter is presented in block diagram form in Figure 4.4. Thick arrows indicate the bootstrap resamples seeded by the αS random number generator.

The distribution of $T_{n,\alpha}$ in (4.5) is approximated by the bootstrap statistics

$$T^*_{n,\alpha} = \frac{|\hat{\alpha}^* - \alpha_0|}{\hat{\sigma}^*_{\hat{\alpha}}}$$

where $\hat{\alpha}^*$ are parameter estimates from the resampled data, $\boldsymbol{X}^*$, and $\hat{\sigma}^*_{\hat{\alpha}}$ are estimates of its standard deviation.

The estimation of $\hat{\sigma}_{\hat{\alpha}}$ is performed by generating data using the estimated param-

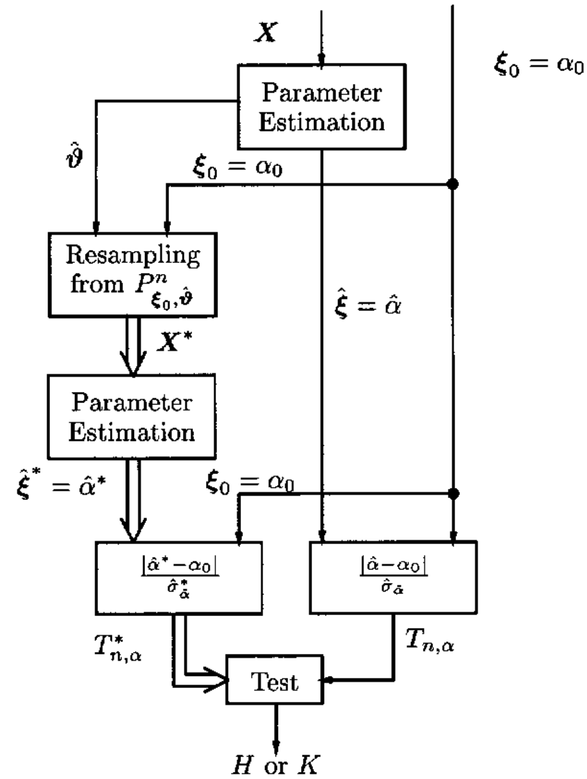


FIGURE 4.4. Block diagram of the parametric bootstrap test for α.

eters of $\boldsymbol{X}$, that is $\boldsymbol{X}^*$, and estimating the standard deviation of the α estimates from these resamples, $\hat{\alpha}^*$. A similar procedure is followed for each $T^*_{n,\alpha,b} = \frac{|\hat{\alpha}^*_b - \alpha_0|}{\hat{\sigma}^*_{\hat{\alpha},b}}$ by estimating the standard deviation of estimates of α derived from resamples generated using the estimated parameters of $\boldsymbol{X}^*_b$. Both estimation procedures have been omitted from Figure 4.4 as it adds unnecessary complexity to the diagram.

With the empirical distribution of $T^*_{n,\alpha}$ approximating the true distribution of $T_{n,\alpha}$, appropriate critical values can be chosen determined by the test's desired level of significance.

4.2.2 Characteristic function based test

Characteristic function based goodness-of-fit tests have already been described in Chapter 3 and Section 4.1. While there are similarities between these two applications, namely that they determine if a particular statistic model is valid for an observed process, in this section a cf based distance test is formulated for testing *within* the αS model. Consequently, a variation on the previously described tests is necessary.

Characteristic function estimators. In the cf based tests previously discussed, as well as in the references cited, statistics were calculated from the difference between a nonparametric estimate of the cf of $\boldsymbol{X}$ and the cf under the null hypothesis. Generally, this nonparametric estimator was the empirical cf (ecf), although the use of the kernel characteristic function estimator (KCFE) was also considered. The use of nonparametric estimators is appropriate for general goodness-of-fit tests where, under the alternative, the distribution of $\boldsymbol{X}$ may be *any* valid distribution.

Here the hypothesis and alternative specify different parameter values for within the αS distribution. That is, even under the alternative, K, $\boldsymbol{X}$ is αS distributed. To exploit this knowledge, a *parametric* cf estimator is used. This estimator uses the estimated parameter values and the known form of the cf of αS distributions and is denoted $\phi(t; \hat{\boldsymbol{\xi}}, \hat{\boldsymbol{\vartheta}})$. This is compared to the cf of the distribution under H, again using the estimated nuisance parameters, $\phi(t; \boldsymbol{\xi}_0, \hat{\boldsymbol{\vartheta}})$.

It is reasonable to expect that if the initial assumption is true that the observations *are* from an αS process and a suitably accurate parameter estimation procedure is available, then a parametric estimator of its cf will be better than a nonparametric estimator.

Test statistics. In [93] it was found that the peak of the absolute difference between two cfs provided a good measure of the distance between the two distributions. There are a large number of measures of the difference between the two functions that may be used. Each will have strengths and weaknesses against different types of alternative distributions. However, it was found that the peak absolute difference provides an omnibus measure that should yield high rejection rates for all reasonable alternatives.

Drawing on this, the test statistic to be used is defined as

$$T_{n,\phi} = \max_t \left| \phi(t; \hat{\boldsymbol{\xi}}, \hat{\boldsymbol{\vartheta}}) - \phi(t; \boldsymbol{\xi}_0, \hat{\boldsymbol{\vartheta}}) \right| \quad . \tag{4.7}$$

The parameter vectors, $\boldsymbol{\xi} = \alpha$ and $\boldsymbol{\vartheta} = [\beta \ c \ \delta]^T$, contain all four parameters of the distribution. This highlights an advantage of this test statistic over the statistic defined in (4.5), namely, it incorporates all the distributional parameters into the test. This allows its use in a broader range of problems including the testing of multiple parameters, compared to $T_{n,\alpha}$ that just tests the value of one parameter, namely, α. The expected cost of this generality is a slight drop in performance when only testing α.

Finding critical values. The distribution of $T_{n,\phi}$ is complicated and unknown. In [93], Monte Carlo simulations were performed to estimate the distribution of the test statistic under the simple null hypothesis of Gaussianity. From this, empirically derived thresholds were found.

As stated in Appendix B, due to the presence of nuisance parameters, no such "once off" Monte Carlo simulation would be appropriate here – it cannot be assumed that the distribution of the test statistics is independent of the nuisance parameters. Not only do they affect the estimation of the parameter of interest, as was the case with testing α in Section 4.2.1, they also affect the form of the cf. Again the parametric bootstrap is used to determine critical values. The distribution of $T_{n,\phi}$ is approximated by the distribution of the bootstrap statistics

$$T^*_{n,\phi} = \max_t \left| \phi(t; \hat{\boldsymbol{\xi}}^*, \hat{\boldsymbol{\vartheta}}) - \phi(t; \boldsymbol{\xi}_0, \hat{\boldsymbol{\vartheta}}) \right| \quad .$$

The procedure is presented in block diagram form in Figure 4.5. Its general structure is similar to that in Figure 4.4 and follows the general procedure described in Table B.1.

4.2.3 Simulation Results and Discussion

A simulation study was undertaken to determine the performance of the two tests. Here $\alpha_0 = 2$ and $\alpha_0 = 1.6$ are considered. In the former, the hypothesis is that the population is Gaussian distributed against any non-Gaussian αS distributed. This is probably the most important case to consider as it tests if the observations have bounded or infinite variance. While two-sided tests have been used here, the minor changes for one-sided tests are trivial.

Observations were generated from an αS random number generator with varying α, but with $c = 1$ and $\delta = 0$. Of most significance is the *symmetric* αS case (SαS), $\beta = 0$, however, results are also shown for $\beta = 0.7$ and $\beta = -0.3$. In many applications, such as some communications systems [64], it is common to consider only the SαS distribution when modelling naturally occurring impulsive interference. The nominal significance level for all tests was set at 5%, the sample size was 400 and the number of bootstrap resamples was 500.

Rejection rates for the tests when $\beta = 0$ have been estimated for a number of values of α and are presented in Table 4.3 for $\alpha_0 = 2$. As well as the $T_{n,\alpha}$ and $T_{n,\phi}$ tests, the Kolmogorov-Smirnov test, T_{KS}, the Anderson-Darling test, T_{AD}, and two of the cf based Gaussianity tests described in (3.6) and (3.3), $Q_{\mathbf{X}}$ and $Q_{\mathbf{X}}^{\Re}$ respectively, have also been added.

α	$T_{n,\alpha}$ (one-sided)	$T_{n,\alpha}$ (two-sided)	$T_{n,\phi}$	T_{KS}	T_{AD}	$Q_{\mathbf{X}}$	$Q_{\mathbf{X}}^{\Re}$
1.7	99.8	99.8	99.8	7.4	95.0	97.0	98.0
1.8	95.8	95.8	95.8	2.8	69.4	78.1	83.0
1.9	67.8	67.8	67.9	1.4	23.7	34.1	39.3
1.95	34.4	34.4	34.4	0.5	9.8	16.6	19.1
2	4.8	4.8	4.6	0.5	5.1	5.4	4.4

TABLE 4.3. Rejection rates (in %) of tests on the level of impulsive behaviour of standard SαS distributed data with $\beta = 0$ where $\alpha_0 = 2$, estimated from 1000 replications.

Table 4.4 shows results for $\alpha_0 = 1.6$. Here the two-sided and one-sided alternatives differ, and consequently two $T_{n,\alpha}$ tests have been included, although only one-sided T_{KS} and T_{AD} tests. For obvious reasons, the cf based Gaussianity tests have been removed for this hypothesis.

α	$T_{n,\alpha}$ (one-sided)	$T_{n,\alpha}$ (two-sided)	$T_{n,\phi}$	T_{KS}	T_{AD}
1.2	99.8	99.8	99.8	68.0	77.4
1.4	79.2	68.9	72.5	34.6	28.2
1.6	4.7	4.7	4.5	12.2	5.1
1.8	0.0	70.8	66.7	2.6	0.7
1.9	0.0	98.0	97.5	0.8	0.1
2	0.0	100.0	100.0	0.1	0.0

TABLE 4.4. Rejection rates (in %) of tests on the level of impulsive behaviour of standard SαS distributed data with $\beta = 0$ where $\alpha_0 = 1.6$, estimated from 1000 replications.

Very little difference can be detected between the $T_{n,\alpha}$ and $T_{n,\phi}$ tests, especially when $\alpha_0 = 2$. Remembering that the $T_{n,\phi}$ test implicitly tests the two-sided alternative, the results in Table 4.4 also show that the $T_{n,\phi}$ test is comparable to the two-sided $T_{n,\alpha}$ test. The one-sided $T_{n,\alpha}$ was more powerful when $\alpha < \alpha_0$ but obviously should not be used when $\alpha > \alpha_0$. It is also significant to note that the levels of significance have been maintained at the nominal 5% level for these tests – validating the use of the parametric bootstrap for threshold setting.

It is significant to note the difference in power between the Q_X and $T_{n,\phi}$ tests, especially for near alternatives, that is, when the actual, true parameter value α is close to α_0. The main difference between these tests is the use of a parametric estimator of the cf of the observed process for $T_{n,\phi}$, as opposed to the nonparametric estimator for Q_X. Here, the assumption about the distribution of the process is utilised. Of course if this assumption does not hold, we would expect the reverse situation to be true.

For far alternatives, the difference between parametric and nonparametric estimates of the cf of the generating process becomes less significant compared to their underlying differences with the cf under H. The $Q_{\mathbf{X}}^{\Re}$ test performs slightly better than the $Q_{\mathbf{X}}$ test since only symmetric distributions are considered.

A more disappointing, though not completely unexpected, result is the poor performance of the Kolmogorov-Smirnov test. This test is not sensitive to deviations in the tails of distributions, concentrating more on deviations near the median. A quick inspection of the normal probability plot in Figure 2.3 reminds us that αS distributed random variables appear very similar and Gaussian-like around the median and that their level of impulsive behaviour only becomes evident when inspecting their tails. Thus the Kolmogorov-Smirnov test is poorly suited to this problem, unlike the Anderson-Darling test that was specifically included in this simulation study since it is known to be a good test for deviations in the tail. While its rejection rates were better than the Kolmogorov-Smirnov, they were still noticeably less than those of $T_{n,\alpha}$ and $T_{n,\phi}$.

Tables 4.5 and 4.6, again test $\alpha_0 = 1.6$, however, the generated data was from an asymmetric αS distribution with $\beta = 0.7$ and $\beta = -0.3$ respectively. As discussed in Section 4.1.2, there is greater difficulty in numerically approximating asymmetric αS pdfs for these tests, therefore the Kolmogorov-Smirnov and Anderson-Darling tests have been omitted, as have the cf based Gaussianity tests, for obvious reasons.

α	$T_{n,\alpha}$ (one-sided)	$T_{n,\alpha}$ (two-sided)	$T_{n,\phi}$
1.2	99.8	99.6	99.8
1.4	77.4	66.8	70.2
1.6	4.2	5.0	4.6
1.8	0.0	73.4	68.2

TABLE 4.5. Rejection rates (in %) of tests on the level of impulsive behaviour of standard SαS distributed data with $\beta = 0.7$ where $\alpha_0 = 1.6$, estimated from 500 replications.

α	$T_{n,\alpha}$ (one-sided)	$T_{n,\alpha}$ (two-sided)	$T_{n,\phi}$
1.2	100.0	100.0	100.0
1.4	82.2	71.2	73.4
1.6	3.8	6.2	5.0
1.8	0.0	72.6	68.2

TABLE 4.6. Rejection rates (in %) of tests on the level of impulsive behaviour of standard SαS distributed data with $\beta = -0.3$ where $\alpha_0 = 1.6$, estimated from 500 replications.

Although individual rejection rates corresponding to a particular α may change slightly, the bootstrap tests still perform acceptably and their relative performances are similar. These results vindicate the use of estimated nuisance parameters in the estimation of critical values as they indicate the level is maintained in spite of the effect β has on the estimation of α.

4.2.4 Conclusions

Two tests have been presented for testing the parameter values of an αS distribution. While the emphasis has been on testing impulsive behaviour through the α parameter, the techniques are general enough to be applied to any of the distributional parameters.

The parametric bootstrap procedures implemented have been shown to allow the appropriate setting of critical values for the tests and resulted in them maintaining the nominal level of significance. Although testing the α parameter directly presented the flexibility in being able to test both alternatives, $\alpha < \alpha_0$ and $\alpha \neq \alpha_0$ compared to the cf based procedure, it is to be noted that only minor differences in performance were noted. It should also be noted that the latter procedure has a higher degree of generality in being able to be easily adapted to test more than one parameter.

The cf based test has also been shown to be a more powerful variation to existing cf based goodness-of-fit tests through the use of knowledge that the distribution of the observed process is within a particular parametric family. Simulation results show that the two presented tests reject the alternatives tested with rates varying depending on the degree of impulsive behaviour and the number of observations available.

4.3 Summary

Tests for the goodness-of-fit of an α stable (αS) distribution, using the characteristic function (cf), and for their level of impulsive behaviour, using the cf and parameter estimates, have been presented. The parametric bootstrap has been used to ensure levels of significance are kept. These contributions serve to answer the "Assume other models" component of the signal detection scheme in Figure 1.1 when considering αS interference.

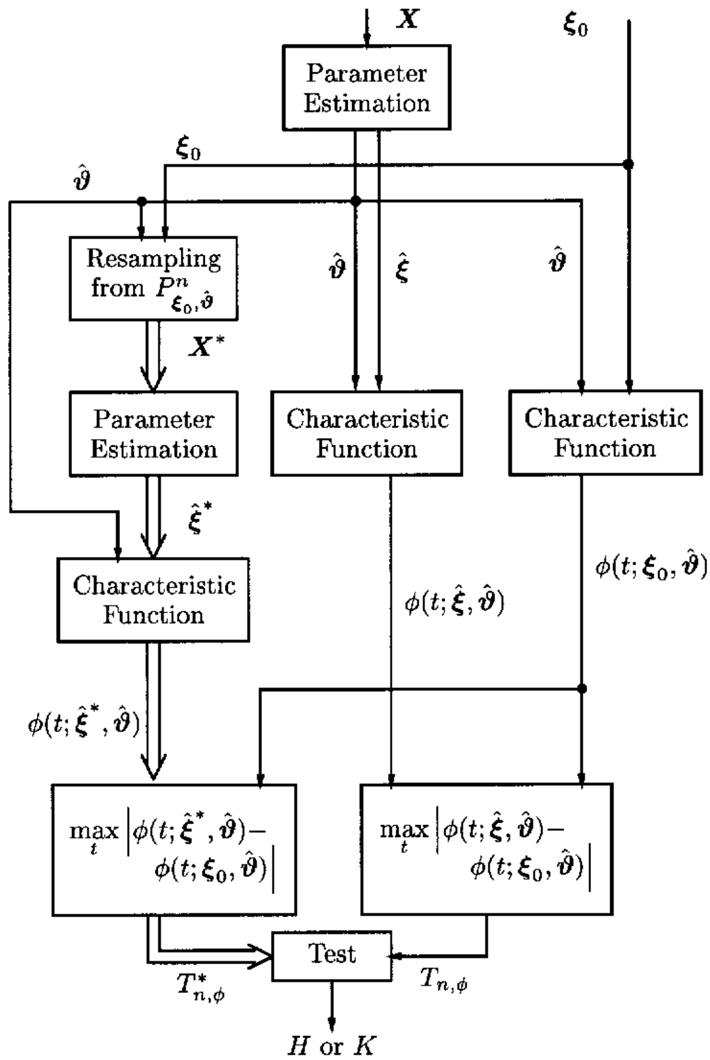


FIGURE 4.5. Block diagram of the characteristic function based test using the parametric bootstrap.

Chapter 5
Testing Real Sources of Impulsive Behaviour

The development of the tests described in the previous chapter was initially motivated to validate the modelling of heart rate variability (HRV) signals by αS processes [13, 69]. However, modelling by αS processes has been suggested for a number of impulsive sources. In this chapter, the goodness-of-fit tests from Chapter 4 are applied to naturally occurring HRV (a "man-made" series in the literal sense), as well as to stock price data, a process that is inherently governed by the conscious, and sub-conscious, actions of people.

5.1 Heart Rate Variability

Analysis of a patient's HRV may be useful in identifying a number of pathological heart conditions [45]. Figure 5.1 shows an HRV data segment. Observable are a number of premature beats (PB) appearing as spikes or outliers. The number and type of PBs may be significant in identifying and monitoring pathological heart conditions [35]. Current clinical evaluation of arrhythmias (abnormal heart rhythms) only identifies and counts the PB events on an ECG (electrocardiogram) signal.

This could be acceptable when few PB events are present, however in the case of extrasystolic arrhythmias, where there is a large number of PB, the signal should be considered in its entirety and an integrated analysis approach would be needed. The features of the αS distribution encouraged investigation of its use in [69] to further develop methods of clinical analysis of HRV signals.

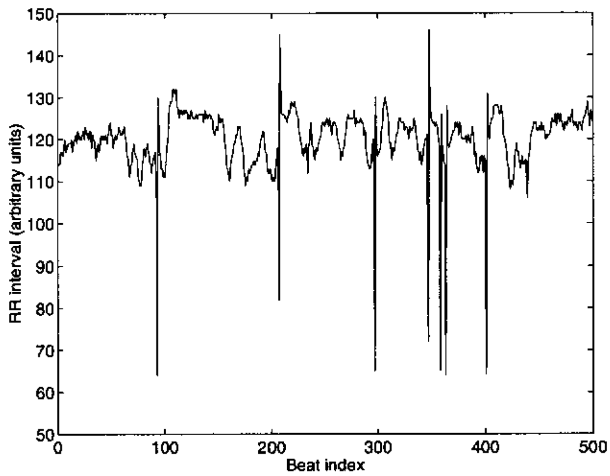


FIGURE 5.1. Segment of HRV signal from patient with ventricular tachycardias.

5.1.1 Alpha-Stable modelling of HRV

While it is not being proposed that the αS distribution is an exact model for HRV with arrhythmia, it is suggested that as an impulsive model with four parameters, the αS distribution may be broad enough to be a suitable empirical model that may allow a unified, statistical approach to the task of HRV analysis. In this context, the αS parameters do not have a direct physical interpretation, however, their effect on the distribution is well known and can be associated with HRV events as follows

- α– as a measure of the level of impulsive behaviour or the frequency of spikes, this parameter will be related to the amplitude of PBs and how frequently they occur

- β– the degree of symmetry characterises the mixture of PB types, for example, ventricular premature beats (VPB) and supraventricular premature beats (SPB).

- c– the dispersion may be a better measure of spread for impulsive observations than the standard deviation

- δ– the mean may not be a good measure of central tendency due to the impulsive behaviour, δ may give a better indication of the underlying heart rate.

5.1.2 Results and Discussion

A detailed explanation and discussion of the use of αS modelling of HRV signals is given in [69] and a summary of the findings is given here.

HRV signals from 24-hour Holter ECG tapes, chosen randomly from a patient group with arrhythmias, were divided into segments of 500 samples with 50% overlap. Before analysis, the data was prewhitened by an inverse αS autoregressive (AR) model of order 11 (Section 7.12 of [70] describes ARMA processes with stable innovations, as well as their invertibility), which generalizes the (implicit) Gaussian AR(11) model used predominantly in HRV studies [82]. The test was applied to the prewhitened segments. Results of applying the αS tests, T_a and T_r, defined in equations (4.1) and (4.2), as well as the Gaussianity tests, $Q_{\mathbf{X}}$ and $Q_{\mathbf{X}}^{\Re}$, defined in (3.6) and (3.3), are shown in Table 5.1.

Goodness-of-fit test	File 1	2	3	4	5
T_a	5.62	7.55	20.3	9.11	5.47
T_r	6.37	7.81	13.3	7.03	5.73
$Q_{\mathbf{X}}$	94.0	17.9	98.7	93.5	83.8
$Q_{\mathbf{X}}^{\Re}$	92.5	17.1	98.8	92.7	84.1

TABLE 5.1. Rejection rate (in %) of goodness-of-fit tests for the αS distribution and the Gaussian distribution at the 5% level on five HRV data files.

Significantly, the Gaussianity hypothesis could be confidently rejected in a number of cases, but not the αS hypothesis. This would suggest that the αS model is more appropriate than the Gaussian model. In the absence of a better alternative, the αS model can be used with some degree of confidence.

The mean of the αS parameter estimates, as well as their standard deviation (in brackets) are shown in Table 5.2 along with the number of ventricular and supraventricular premature beats, VPB and SPB respectively, occurring in each data file. The relationship hypothesised previously between αS parameter estimates and the type and frequency of PB appears to hold to some degree, the lower values of α indicate more impulsive behaviour and a higher PB count.

	File				
Feature	1	2	3	4	5
VPB	153	6	4967	239	130
SPB	180	10	93	230	1026
α	1.645	1.943	1.406	1.554	1.615
	(0.203)	(0.063)	(0.186)	(0.232)	(0.238)
β	0.401	0.191	0.040	0.06	0.282
	(0.325)	(0.787)	(0.534)	(0.09)	(0.283)
γ	2.53	1.733	4.476	2.62	3.89
	(1.525)	(1.349)	(1.495)	(2.529)	(1.291)
δ	0.459	0.023	1.33	1.72	0.775
	(0.897)	(0.075)	(27.69)	(0.445)	(7.125)

TABLE 5.2. Conventional analysis of five HRV data files and estimated αS parameters.

It is difficult to confirm the relationship between β and the proportion of VPBs (asymmetric) to SPBs (symmetric) since the variability of the estimates is quite high. It was noted in [57] that the effect of β on the distribution is reduced as $\alpha \to 2$. Inspection of the cf of the αS distribution in (2.1) confirms that the factor $\tan(\alpha\pi/2) \to 0$, thus reducing the weighting of the β parameter.

Given the smooth parametric transition from the existing implicitly Gaussian models ($\alpha = 2$) to the more general, proposed, αS models, as well as on the weight of the mentioned results, a clinical study to evaluate the physiological significance of αS modelling of HRV is being undertaken.

5.2 Stock Prices

It was hypothesised by Mandelbrot [58] that the logarithm of stock price changes follows an αS distribution. It is not surprising, therefore, that stock market data have been used in a number of prominent contributions to αS parameter estimation, specifically by Fama and Roll, Leitch and Paulson and Koutrouvelis in [25, 57, 50].

If the price of a stock at the end of the month is p_i, and the dividend paid is d_i, then the adjusted rate of return is

$$r_i = \frac{p_i - p_{i-1} - d_i}{p_{i-1}}$$

and the logarithmic price change used is

$$x_i = \log(r_i + 1) \quad .$$

The author was unable to obtain the data sets used by any of the previously mentioned authors, however historical monthly data for a number of stocks was obtained from `http://chart.yahoo.com/t` covering more recent time periods. The oldest records were for IBM and started in January 1962. The stocks chosen were generally from the lists used in the references above, although some other randomly chosen stocks were also included. The one criterion applied to choosing stocks was that sufficiently large data was available, and hence, the stocks tended to be larger, well established companies.

The stocks are listed in Table 5.3, along with the significance or p value associated with the cf based αS goodness-of-fit test described in Section 4.1. Also tabulated are the sample sizes and the estimated αS parameters.

From the p values shown, it would appear that the fit of the αS distribution to this type of data is very good. None of the sets would reject the null hypothesis that the data is αS distributed at any reasonable level.

Being unable to dismiss this hypothesis, it is now possible to test for the level of impulsive behaviour using the $T_{n,\alpha}$ and $T_{n,\phi}$ tests. Shown in Table 5.4 are the p

Company	Samples	p	$\hat{\alpha}$	$\hat{\beta}$	$\hat{c}$	$\hat{\delta}$
Am. Home Prod.	213	0.60	1.862	0.519	0.039	-0.014
American Exp.	270	0.90	1.907	0.942	0.053	-0.014
Boeing	357	0.32	1.866	0.296	0.061	-0.010
BHP	149	0.78	1.944	-1.000	0.053	-0.011
Colgate-Palm.	273	0.47	1.827	1.000	0.043	-0.010
General Motors	357	0.68	1.907	-0.545	0.046	-0.010
Goodyear	357	0.31	1.824	-0.206	0.051	-0.009
IBM	453	0.55	1.929	-0.084	0.045	-0.007
Intel	159	0.37	1.942	1.000	0.079	-0.025
J. P. Morgan	357	0.66	1.927	0.308	0.048	-0.010
Litton Ind.	214	0.60	1.891	0.287	0.048	-0.006
3M	357	0.42	1.814	0.180	0.038	-0.008
PepsiCo.	263	0.51	1.822	0.343	0.042	-0.015
Pfizer	213	0.47	1.962	1.000	0.051	-0.018
Vintage Petr.	110	0.58	1.670	0.167	0.077	-0.005

TABLE 5.3. Sample sizes, p values and estimated αS parameter values for the logarithmic price changes of some stocks, when testing the goodness-of-fit of the αS distribution.

values of the two tests for two different hypotheses, $\alpha_0 = 1.8$ and $\alpha_0 = 2$. While the $\alpha_0 = 2$ is obviously testing for Gaussianity, the $\alpha_0 = 1.8$ hypothesis was included due to the assertion in [65] that post-war monthly stock returns would have α close to this value.

		$\alpha_0 = 1.8$		$\alpha_0 = 2$	
Company	$\hat{\alpha}$	$T_{n,\alpha}$	$T_{n,\phi}$	$T_{n,\alpha}$	$T_{n,\phi}$
Vintage Petr.	1.670	0.16	0.59	0.00	0.06
3M	1.814	0.56	0.87	0.00	0.00
PepsiCo.	1.822	0.58	0.82	0.00	0.01
Goodyear	1.824	0.61	0.79	0.00	0.01
Colgate-Palm.	1.827	0.63	0.72	0.00	0.00
Am. Home Prod.	1.862	0.67	0.56	0.03	0.03
Boeing	1.866	0.83	0.51	0.01	0.06
Litton Ind.	1.891	0.83	0.46	0.04	0.09
American Exp.	1.907	0.87	0.21	0.05	0.02
General Motors	1.907	0.92	0.26	0.02	0.03
J. P. Morgan	1.927	0.94	0.16	0.06	0.13
IBM	1.929	0.98	0.14	0.04	0.11
Intel	1.942	0.89	0.20	0.21	0.09
BHP	1.944	0.87	0.21	0.21	0.15
Pfizer	1.962	0.97	0.10	0.27	0.20

TABLE 5.4. Estimated α and p values for the tests of impulsive behviour. Four p values are shown for each company, two each for the $T_{n,\alpha}$ and $T_{n,\phi}$ tests and two each for $\alpha_0 = 1.8$ and $\alpha_0 = 2$.

While broad conclusions from this table on stock price data are difficult to make considering the range of stocks, it is encouraging to note that generally the stocks with larger α achieved higher significance levels than the other stocks when testing $\alpha_0 = 2$ and for the one-sided $T_{n,\alpha}$ test when $\alpha_0 = 1.8$, while they had lower significance levels compared to other stocks for the two-sided $T_{n,\phi}$ test when $\alpha_0 = 1.8$. Conversely, the Vintage Petroleum stock, $\hat{\alpha} = 1.67$, had lower significance levels for $\alpha_0 = 2$ and for $T_{n,\alpha}$ when $\alpha_0 = 1.8$.

This apparent relationship between $\hat{\alpha}$ and the p values is to be expected, though, with the presence of nuisance parameters, it is not surprising that some stocks with lower values of $\hat{\alpha}$ achieved higher p values.

5.3 Summary

The αS goodness-of-fit test and tests of the level of impulsive behaviour have been applied to heart rate variability and stock price variations. The encouraging results give weight to the argument that the αS distribution may be used to model these real data sets.

5.4 Acknowledgment

The author would like to acknowledge the contributions of Dr Sokol Saliu of Chalmers University – Lindholmen and Dr N. Edvardsson and Dr D. Poçi of the Division of Cardiology at Sahlgrenska Hospital, all located in Göteborg, Sweden, in the HRV-related part of this work.

Chapter 6

SIGNAL DETECTION IN IMPULSIVE INTERFERENCE

Signal detection is an extensively studied and analysed topic. Detectors for a known signal in additive interference of unknown power have been derived for a number of interference distributions – the most famous and widely used being the matched filter (MF) for Gaussian interference.

Sources of impulsive interference, particularly those that follow a stable distribution, has presented some interesting problems for detection. Many conventional statistical and optimal detection techniques, such as likelihood ratio procedures, require a closed form expression for the pdf of the interference. As discussed in Section 2.1, there is no general expression for the pdf of the α-stable (αS) distribution. For this reason, it is necessary to investigate alternative strategies for the detection of signals in αS interference.

Consider the model

$$\boldsymbol{X} = \theta\, \boldsymbol{s} + \boldsymbol{W} \tag{6.1}$$

where $\boldsymbol{X} = [X_1, X_2, \dots, X_n]^T$ is the model for the real-valued observations, θ is a positive, real, unknown parameter, $\boldsymbol{s} = [s_1, \dots, s_n]^T$ is the known, deterministic signal to be detected and $\boldsymbol{W} = [W_1, W_2 \dots, W_n]^T$ is a stationary, iid, symmetric α-stable (SαS) interference process. To determine the presence of $\boldsymbol{s}$ in $\boldsymbol{X}$, the tested hypothesis is

$$\mathsf{H} : \theta = 0 \quad \text{against} \quad \mathsf{K} : \theta > 0 \tag{6.2}$$

The focus here is on the case where the SαS distribution provides an adequate model for the impulsive interference, $\boldsymbol{W}$. This model assumption should be validated from the physics of the process or empirically through the fitting of observations, as was described in Chapter 4 and performed in Chapter 5.

In the following sections, signal detection in SαS interference using correlation and rank-based detectors is discussed. Following that, simulation results are presented, compared and discussed, and finally conclusions are drawn. Some of this work was introduced in [17].

6.1 Optimal and Suboptimal Detectors

Since numerous references are available for detection theory, for example [56, 68, 88], only a brief introduction to conventional and classical detection will be presented here and is only to serve to justify the alternative procedures proposed later in this section and chapter. Of main interest here are generalised correlation detectors [44], which include locally optimum (LO) and locally suboptimum (LSO) detectors, although the Cauchy detector and some other alternatives are discussed later.

6.1.1 Optimum Tests

Before comparing tests it is necessary to define what is "best". The Neyman-Pearson criterion is used here to define optimality, that is, the optimum test maximises the probability of detection within the class of tests that maintain a certain, fixed false alarm rate (the probability of an error of Type I). If K is simple,

$$\mathsf{H} : \theta = 0 \quad \text{against} \quad \mathsf{K} : \theta = \theta_0$$

then by the Neyman-Pearson lemma, the likelihood ratio test is optimal for detection in the model (6.1). The test has the form

$$T(\boldsymbol{X}) = \sum_{i=1}^{n} \frac{f_W(X_i - \theta_0 s_i)}{f_W(X_i)} \underset{\mathsf{H}}{\overset{\mathsf{K}}{\gtrless}} \kappa \tag{6.3}$$

for some κ and where $f_W(w)$ is the pdf of the additive interference, W.

When K is composite, as in (6.2), a uniformly most powerful (UMP) test must have the highest probability of detection ("most powerful") under all alternatives ("uniformly"). This is a very strong and difficult condition to prove. Common alternatives when no UMP test can be found include generalised likelihood ratio tests, where maximum likelihood estimates of parameters are used in (6.3), and locally optimum tests.

6.1.2 Locally Optimum Detectors

Locally optimum (LO) [76, 77] tests attain the best detection performance amongst the class of detectors of the same size for weak signal conditions, that is, they maximise the slope of the detector power function at $\theta = 0$. Considering weak signal performance is an important criterion in comparing detectors since, typically, many detection schemes can be designed for high detection rates at high signal strengths.

A UMP test has maximal power for all parameter values, including the weak-signal case. Therefore a UMP, if it exists, will also be LO. While the converse relationship cannot be inferred in the general case, local optimality is clearly a highly desirable feature.

A LO detector for the detection of a signal in additive noise uses the test statistic [44]

$$T_{LO}(\boldsymbol{X}) = \sum_{i=1}^{n} s_i \, g_{LO}(X_i) \tag{6.4}$$

where the nonlinear score function is

$$g_{LO}(x) = -\frac{f_W'(x)}{f_W(x)} \tag{6.5}$$

and $f_W'(x)$ is the first derivative of $f_W(x)$, see Figure 6.1. It is simple to show that when the interference is Gaussian, $g_{LO}(x) = x$, producing the well known matched filter.

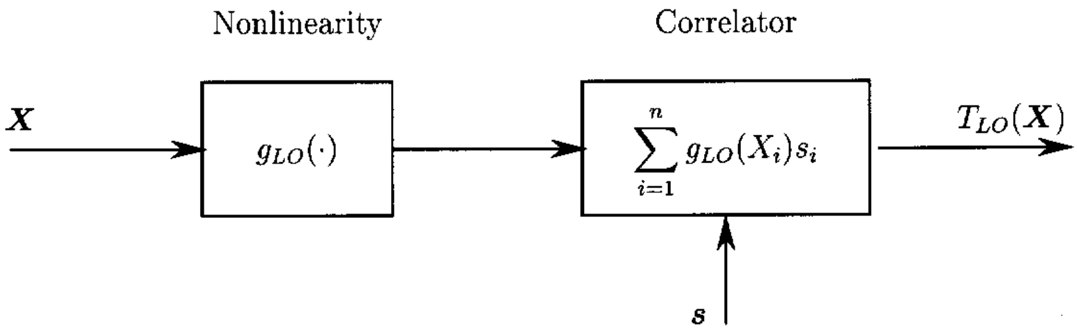


FIGURE 6.1. Structure of the LO Detector.

Due to the absence of closed form expressions for $f_W(x)$ when it is αS, the non-linear score function, $g_{LO}(x)$, cannot be found exactly. In [64], g_{LO} was found numerically for some SαS cases and plotted for a number of values of α. These results are reproduced in Figure 6.2.

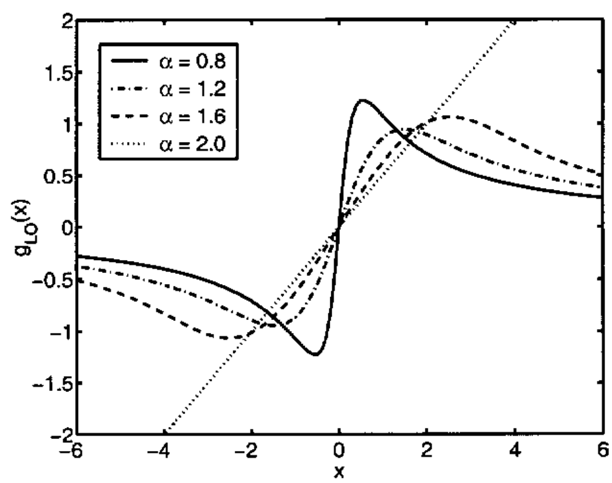


FIGURE 6.2. Locally Optimum score functions for various SαS distributions.

While these numerical approximations can be made to be extremely accurate, practical implementation requires coarser approximation, especially in applications where α is not constant. These approximations invalidate the "locally optimum" feature of the detector, and yield locally suboptimum (LSO) detectors.

6.1.3 Locally Suboptimum Detectors

Approximations to the computationally complex g_{LO} have been suggested in [3, 53, 87]. While the hole-puncher nonlinearity has been proposed as a simple approximation, it is seen as providing poor approximation to g_{LO} for larger values of x. Here two other nonlinear functions are considered that have the same apex as the hole-puncher, but decay at different rates. Since only the SαS case is considered, the nonlinearities are odd functions, and they are shown in Figure 6.3 for $x \geq 0$ only.

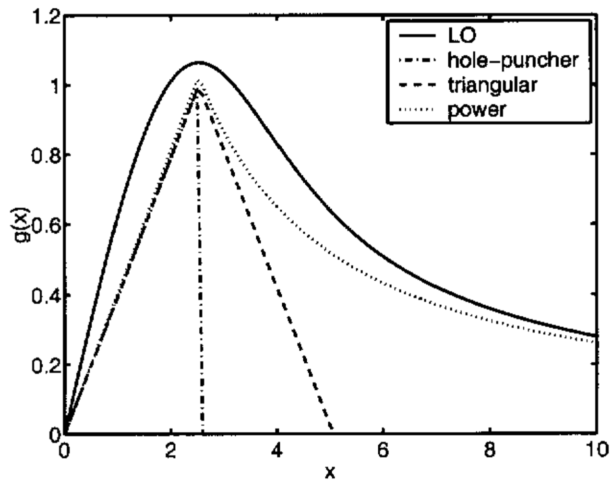


FIGURE 6.3. LSO nonlinear score functions when $\alpha = 1.6$.

The triangular nonlinear function decays to zero at the same linear rate with which the hole-puncher rose, while the LSO-power nonlinearity decays at the same asymptotic rate as g_{LO}, [64]

$$g_{LO}(x) \to \frac{\alpha + 1}{x} \quad , \text{ as } x \to \pm\infty \quad . \tag{6.6}$$

The apex of the LSO nonlinearities is found by numerically locating the peak of the g_{LO} nonlinearity, and are therefore dependent on α. All of these approximation functions have low computational complexity as each is specified in terms of a single apex point and follow simple algebraic expressions. However, it is expected that the introduced functions, especially LSO-power, will provide better approximations to g_{LO} than the hole-puncher. This should be apparent even upon simple visual inspection of Figure 6.3.

6.1.4 Cauchy Detector

Besides the Gaussian case, one of the few special cases where a closed form expression for the pdf of an αS random variable exists is the Cauchy distribution ($\alpha = 1, \beta = 0$),

$$f(x) = \frac{1}{\pi\gamma} \frac{1}{1 + \left(\frac{x-\delta}{\gamma}\right)^2} \quad .$$

Using the Maximum Likelihood criterion, the optimal detector for testing the hypothesis $\mathsf{H} : \theta = 0$ against $\mathsf{K} : \theta = \theta_0$ when the interference is Cauchy distributed is given by

$$\begin{aligned} T_C(\boldsymbol{X}) &= \sum_{i=1}^{n} \log \frac{f(X_i - \theta_0 s_i)}{f(X_i)} \\ &= \sum_{i=1}^{n} \log \frac{1 + \left(\frac{X_i - \delta}{\gamma}\right)^2}{1 + \left(\frac{X_i - \delta - \theta_0 s_i}{\gamma}\right)^2} \\ &= \sum_{i=1}^{n} \log \frac{\gamma^2 + (X_i - \delta)^2}{\gamma^2 + (X_i - \delta - \theta_0 s_i)^2} \end{aligned} \quad (6.7)$$

Although this detector is optimal only for the Cauchy distribution, it has been shown [17, 64, 87] that, unlike the matched filter, it maintains a high level of performance for $\alpha \neq 1$.

6.1.5 Other Detectors for SαS interference

Some detection techniques and receivers for SαS interference have been proposed using other estimators or approximations to the impulsive density or using limiters and other nonlinearities, while others have exploited Fractional Lower-Order Moments [73] or characteristics of the αS distribution. For example, in [84], the detection of impulsive transients in a background of Gaussian noise is considered by using the form of the pdf of the sum of two stable distributions with different parameters. While [3, 37, 38] consider detection in a mixture of Gaussian and stable noise.

Other nonlinearities and the Cauchy detector are analysed in [72, 85, 87, 86].

Zero-memory nonlinearities (ZMNL) have many similarities with LSO score functions. The ZMNL used in [81] is similar to the score functions presented here, in particular with the LSO-power nonlinearity. Specifically, both are linear at low-amplitude, while using different analytical functions to describe the high-amplitude decay. The primary difference between them is that the ZMNL does not incorporate the characteristic exponent, α, and has an exponential tail, rather than the algebraic rate of decay of the LSO-power nonlinearity.

Rather than impose a distributional family on the noise, [91] uses a nonparametric estimate of the impulsive density, while [52, 54] approximate the αS density as a scale mixture of Gaussian processes.

6.2 Rank-Based Detectors

Detection using the ranks of observations, rather than just their magnitude, has a long history. Though sometimes overlooked in favour of conventional methods, it has long been accepted that by using a weak set of assumptions, rank-based tests can achieve robust performance while often only suffering slight losses in efficiency against parametric tests. The literature on rank-based or nonparametric tests is extensive and well established, see for example [12, 28, 29, 55, 66, 83].

6.2.1 Locally Optimum Rank Detectors

Using the methodology that developed LO detectors, the corresponding locally optimum *rank* detectors (LOR) when f_W is symmetrically distributed can be shown to use the following test statistic [33, 78]

$$T_{LOR}(\boldsymbol{X}) = \sum_{i=1}^{n} s_i \, \mathsf{sgn}(X_i) \, g_{LOR}(r_i) \tag{6.8}$$

where

$$g_{LOR}(i) = \mathsf{E}_{\mathsf{H}} \left[g_{LO} \left(|X|_{(i)} \right) \right] \quad , \tag{6.9}$$

r_i is the rank of $|X_i|$ in the set $[\ |X_1|, |X_2|, \ldots, |X_n|\ ]$ and $|X|_{(i)}$ is the set's ith smallest member. $\mathsf{E}_{\mathsf{H}}[\cdot]$ denotes the expectation operation under the hypothesis H. The similarity between (6.8) and (6.4) should be evident – both correlate the signal to be detected with a function of the observations. The structure of the LOR detector is shown in Figure 6.4.

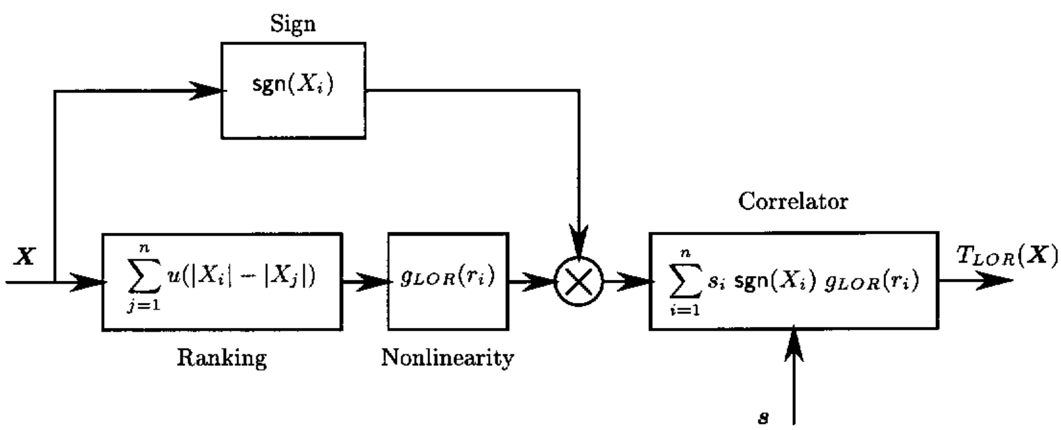


FIGURE 6.4. Structure of the LOR Detector.

While, by definition, the LO detector is the optimum detector for vanishing signal strength, it can be shown that, asymptotically as $n \to \infty$, the LOR and LO detectors become equivalent. Furthermore, the advantages in using rank based detectors are well documented [28, 29].

Rank based techniques require weak assumptions about the distributional properties of the interference – in this case only the symmetry of f_W is assumed. The ranking operation replaces distribution-dependent magnitude information with a discrete rank, determined by the relative position of the data sample amongst the other observations. Assuming the weak assumptions are met, this enables them to maintain a constant false alarm rate, independent of the interference distribution and is of particular interest when the exact form of the distribution is unknown.

As for LO detectors, for each symmetric distribution there exists a rank score function that will yield a LOR detector. However, unlike LO detectors, the same rank-based detector should still maintain its false alarm rate and achieve high (though not locally optimal) detection power against *any* symmetric distribution.

Ranking has additional benefits when outliers or spikes are encountered. By discarding magnitude information, ranking removes the very feature that makes some other tests overly sensitive to outliers. As a result, rank tests do not require the assumption of finite variance.

These properties make rank-based detection attractive when impulsive, and especially SαS, interference is encountered.

6.2.2 Score functions

The LOR nonlinearities, g_{LOR}, for a number of SαS distributions have been approximated numerically and are shown in Figure 6.5. While g_{LOR} is dependent on the LO

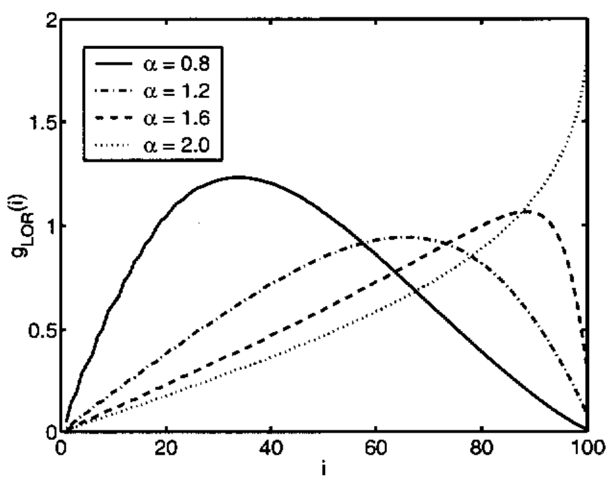


FIGURE 6.5. Locally Optimum Rank score functions for various SαS distributions.

score function, g_{LO}, there are several compelling arguments for implementing a LOR detector over a LO detector:

- The "usual" advantages of rank-based methods over parametric detectors apply, including a constant false alarm rate and higher detection rates when deviations in the distributional assumptions or errors in parameter estimates occur.
- While neither g_{LO} nor g_{LOR} have closed form expressions, for a fixed distribution and sample size, n, g_{LOR} need only be evaluated at n points. The results may be found *once* to a high degree of accuracy off-line and stored for on-line detection. By contrast, due to the slow rate of decay of g_{LO} and its infinite span, there may be great difficulty in providing sufficiently accurate and compact records of g_{LO} for on-line detection.
- If off-line estimation of g_{LOR} is not possible, for example, if α varies significantly, then on-line approximation of g_{LOR} is assisted by noting the smooth nature of the function and that its limiting values, $g_{LOR}(1)$ and $g_{LOR}(n)$, approach zero rapidly (see Figure 6.5). By contrast, while $g_{LO}(0) = 0$, for large $|x|$ there is a very slow decay of $g_{LO}(x)$ to zero, as shown in (6.6). An exception to this behaviour is when $\alpha = 2$.

6.2.3 Locally Suboptimum Rank Detectors

While it has already been noted that the g_{LOR} function may be evaluated off-line to sufficient accuracy for a fixed sample size and α, in circumstances where this cannot be assumed and a useful approximation to the LOR nonlinearities is required, a triangular function may suffice. In Figure 6.6 are shown a triangular score function (LSOR-tr) along with the LOR and linear score functions for $\alpha = 1.6$. As was the case with some of the LSO nonlinearities, the apex of the LSOR-tr score function was the peak of the LOR score function, evaluated numerically. From inspection of

Figure 6.5 it can be seen that a triangular function could be fitted to all of the g_{LOR} shown, with the exception of the Gaussian case.

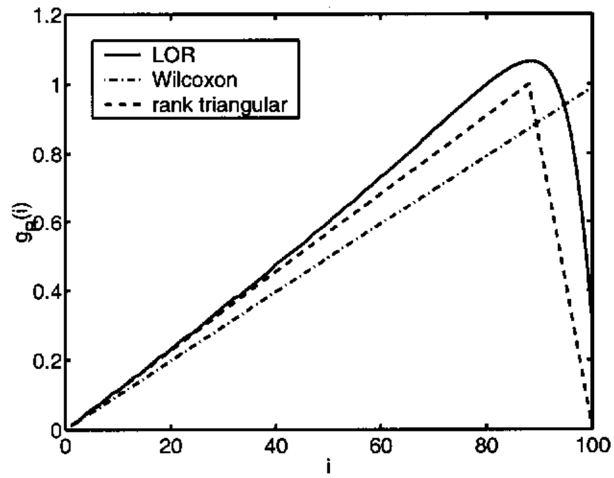


FIGURE 6.6. Rank score functions used for LOR and LSOR detectors when $\alpha = 1.6$ and sample size of 100.

The linear score function is referred to as the Wilcoxon score function. It is interesting to note that the Wilcoxon function is the locally optimal score function for the logistic distribution and not the Gaussian distribution. Following the naming convention introduced earlier, the rank score function approximations to g_{LOR} produce locally suboptimum rank (LSOR) detectors.

6.2.4 Signed Quadratic Rank-Based Detector

In [1], Adichie proposed a signed quadratic rank detector using the statistic

$$M^{+} = \frac{\left(\sum_{i=1}^{n} s_i \operatorname{sgn}(X_i) \, \xi \left(\frac{n+1+r_i}{2n+2} \right) \right)^2}{\left(\sum_{i=1}^{n} s_i^2 \right) \int \xi^2(u) du} . \tag{6.10}$$

The statistic is proportional to the square of the correlation between the signal, s, and a function of the ranks of the magnitude of the observations, $\xi(\cdot)$ being similar to the rank score functions defined previously.

Under certain regularity conditions it can be shown that, as n increases, asymptotically, M^+ has a χ_1^2 distribution. Hence, an appropriate threshold can be set for a detector of specified level. This distributional approximation is sufficiently accurate for the operating conditions used in this paper. Although tedious, the exact distribution may be calculated, however since it is dependent on $\boldsymbol{s}$, tabulated values are not available.

The relationship between the linear correlation rank detector in (6.8) and this quadratic rank detector should be evident [18]. The score functions developed and described previously should be able to be used in the quadratic statistics to achieve similar effect. The M^+ test is a quadratic rank-based detector with an additional scaling factor.

6.3 Distribution of Test Statistics

Although the LO, LSO, LOR and LSOR detectors all use different score functions, the similarity in their structure as correlators means the distribution of their test statistics are very similar. As will be discussed later, an exception to this is the matched filter due to its linear score function.

6.3.1 LO and LSO Detectors with Nonlinear Score Functions

Recall that the linear correlation statistics have the form

$$T(\boldsymbol{X}) = \sum_{i=1}^{n} s_i \, g(X_i)$$

that is, the test statistic is the sum of independent random variables, assuming the X_i are iid and s_i is some bounded, known sequence. Under H, the summed variables have similar distributions, differing only in the non-constant scale, s_i. If $g(x)$ has finite variance, the Central Limit Theorem may be invoked, meaning $T(\boldsymbol{X})$ is asymptotically Gaussian.

Since the score functions considered here are anti-symmetric, and the random variables, X, are SαS, then under H, $\mathsf{E}[g(X)] = 0$. Consequently,

$$\begin{aligned} \mathsf{E}[T(\boldsymbol{X})] &= 0 && (6.11) \\ \mathsf{var}[T(\boldsymbol{X})] &= \mathsf{var}[g(X)] \times \sum_{i=1}^{n} s_i^2 \quad . && (6.12) \end{aligned}$$

To determine if $g(X)$ has finite variance, an asymptotic expansion of the pdf of a standardised SαS random variable when $\alpha < 2$ and as $|x| \to \infty$, as given by [64] is used

$$\begin{aligned} f(x) &= \sum_{k=1}^{K} \frac{b_k}{|x|^{\alpha k+1}} + O\left(|x|^{-\alpha(K+1)-1}\right) \\ \text{where} \quad b_k &= -\frac{1}{\pi}\frac{(-1)^k}{k!}\Gamma(\alpha k+1)\sin\left(\frac{k\alpha\pi}{2}\right) \quad . \end{aligned}$$

This series may be approximated by its first term as this is the term with the slowest rate of decay for $|x| \to \infty$,

$$f(x) \approx \frac{b_1}{|x|^{\alpha+1}} \quad . \qquad (6.13)$$

Referring to Figure 6.3 and (6.6), of the *nonlinear* score functions considered, the LO and LSO-power nonlinearities decay at the slowest rate. Therefore to determine if

$$\mathsf{var}[g_{LO}(X)] = \int_{-\infty}^{\infty} g_{LO}^2(x)\ f(x)\ dx < \infty \quad ,$$

consider the integral

$$I = \int g_{LO}^2(x) f(x) dx$$

and its approximation using its highest order term as $x \to \infty$

$$\begin{aligned} I &\approx \int \frac{(\alpha+1)^2}{x^2}\frac{b_1}{x^{1+\alpha}}dx \\ &= (\alpha+1)^2\, b_1 \int x^{-3-\alpha}dx \\ &= -\frac{(\alpha+1)^2}{\alpha+2}\, b_1\, x^{-2-\alpha} \quad . \end{aligned}$$

The highest order term of I does not diverge to $\pm\infty$ as $x \to \pm\infty$, and therefore, neither will I. Furthermore, both $g_{LO}(x)$ and $f(x)$ are bounded functions, therefore it can be concluded that evaluation of the integral, I, between $-\infty$ and $+\infty$, that is, the variance of $g_{LO}(X)$, is finite.

Any further terms taken from the asymptotic expansion of the pdf will have faster rates of decay, as will any of the nonlinear score functions considered here, therefore it can be taken that the variance of the nonlinear score functions is finite and the corresponding test statistics are asymptotically Gaussian with mean and variance given in (6.11) and (6.12).

6.3.2 Matched Filter

While the above is true of the detectors considered with nonlinear score functions, the matched filter's test statistic

$$T_{MF}(\boldsymbol{X}) = \sum_{i=1}^{n} X_i s_i \quad .$$

is the sum of n scaled independent SαS random variables. Assuming $\alpha \neq 2$, they do not have finite variance and the Central Limit Theorem cannot be invoked. Therefore, rather than an asymptotically Gaussian test statistic, $T_{MF}(\boldsymbol{X})$ is an SαS random variable.

Consider the case

$$Z = aX_1 + bX_2$$

where X_1 and X_2 are iid SαS random variables ($\beta = \delta = 0$) with $\alpha_1 = \alpha_2 = \alpha$ and $c_1 = c_2 = c > 0$, then by the Stability Property, Theorem 1, Z is also SαS with $\alpha_Z = \alpha$. Therefore its cf is

$$\phi_Z(t) = \mathsf{E}\left[\exp(jZt)\right] = \exp\left\{-|c_Z t|^{\alpha}\right\}$$

and by substitution

$$\begin{aligned}
\phi_Z(t) &= \mathsf{E}\left[\exp(j(aX_1 + bX_2)t)\right] \\
&= \mathsf{E}\left[\exp(jaX_1)\right]\mathsf{E}\left[\exp(bX_2)t)\right] \\
&= \phi_{X_1}(at)\phi_{X_2}(bt) \\
&= \exp\left(-|c_1 at|^{\alpha_1} - |c_2 bt|^{\alpha_2}\right) \\
&= \exp\left(-|ct|^{\alpha}(|a|^{\alpha} + |b|^{\alpha})\right) \qquad\qquad c_1, c_2 = c \\
&= \exp\left(-\left| c\left(|a|^{\alpha} + |b|^{\alpha}\right)^{1/\alpha} t \right|^{\alpha}\right) \quad .
\end{aligned}$$

Therefore the scale parameter of Z is

$$c_Z = c\left(|a|^{\alpha} + |b|^{\alpha}\right)^{1/\alpha} \tag{6.14}$$

or in terms of the dispersion $\gamma = c^{\alpha}$,

$$\gamma_Z = \gamma\left(|a|^{\alpha} + |b|^{\alpha}\right) \quad .$$

By generalising this result, it can be shown that under H, the $T_{MF}(\boldsymbol{X})$ statistic is SαS with the same characteristic exponent, α, as X, centred around 0 and with dispersion

$$\gamma_{T_{MF}(\boldsymbol{X})} = \gamma_X \sum_{i=1}^{n} |s_i|^{\alpha_X} \quad .$$

6.3.3 LO and LOR Detectors

Now consider the general form of a linear rank-based detector statistic

$$T_R(\boldsymbol{X}) = \sum_{i=1}^{n} s_i \ \mathsf{sgn}(X_i)\ g_R(r_i)$$

where r_i is the rank of $|X_i|$. While $\mathsf{sgn}(X_i)$ is independent of $\mathsf{sgn}(X_j), i \neq j$, the same cannot be said of the ranks. If the possibility of ties is neglected, each r_i is an integer between 1 and n and each rank integer occurs only once, that is $r_i \neq r_j$ if $i \neq j$.

Then it is clearly seen that the ranking procedure introduces dependence between the terms. However, as $n \to \infty$, this dependence becomes negligible, and therefore *asymptotically* these terms are independent. Again, the Central Limit Theorem can be used to assert the asymptotic Gaussianity of $T_R(\boldsymbol{X})$.

If $|g_R(r_i)| < \infty$ for all $i = 1, \ldots, n$ then

$$\begin{aligned} \mathsf{E}\left[g_R(r)\right] &= \frac{1}{n}\sum_{i=1}^{n} g_R(i) \quad \text{and} \\ \mathsf{E}\left[g_R^2(r)\right] &= \frac{1}{n}\sum_{i=1}^{n} g_R^2(i) \end{aligned}$$

will be finite for finite n. For the signum function the following trivial results can be found when X is symmetrically distributed

$$\begin{aligned} \mathsf{E}\left[\mathsf{sgn}(X)\right] &= 0 \\ \mathsf{E}\left[\mathsf{sgn}^2(X)\right] &= 1 \quad . \end{aligned}$$

The distribution of the test statistic $T_R(\boldsymbol{X})$ under H is independent of the distribution of X, provided it is symmetric. While its *exact* distribution may be calculated for any g_R and $\boldsymbol{s}$, in practice, this is tedious and a suitably accurate approximation can be made using the Gaussian distribution with

$$\begin{aligned} \mathsf{E}\left[T_R(\boldsymbol{X})\right] &= 0 \\ \mathsf{var}\left[T_R(\boldsymbol{X})\right] &= \frac{1}{n}\sum_{i=1}^{n} g_R^2(i) \sum_{j=1}^{n} s^2(j) \quad . \end{aligned} \tag{6.15}$$

6.3.4 Quadratic detectors

When considering the score functions that yielded asymptotically Gaussian test statistics, the signed quadratic rank detector in (6.10) can be seen to be the square of the test statistic, normalised by its variance, see (6.12) and (6.15). It is well known that the square of a standard Gaussian random variable is χ_1^2 distributed.

6.3.5 Cauchy detector

The test statistic in (6.7) is, again, the sum of independent random variables with finite variance and thus is asymptotically Gaussian for large n. Its expected value and variance can be evaluated numerically for particular αS distributions.

6.4 Performance Analysis

For the simulation results in this section, interference was generated from an SαS distribution with unit dispersion and centred around zero. The signal to be detected, $\mathbf{s}$, was a sinusoid with unit amplitude. Observation samples were of length 100, the nominal false alarm rate was 5% and 1000 Monte Carlo simulations were performed to estimate detection rates. These operating conditions were chosen to ensure detection rates were neither too high nor too low since this would make comparison between detectors difficult.

The results for the LO and LSO detectors described previously, as well as the Cauchy detector are shown in Table 6.1. Also included are the quadratic forms of the correlation detectors in Table 6.3. They have been included to allow a comparison with the quadratic rank tests. Additionally, the Receiver Operating Characteristic (ROC) curves for a number of detectors are shown in Figure 6.7 and Figure 6.8.

A number of observations can be made

- All of the tests appear to maintain or closely maintain their false alarm rate. This encourages the continued use of the asymptotic distributions of the test statistics derived earlier.
- As expected, there is little difference in performance between the LO and LOR detectors. However, for practical implementation, if off-line evaluation of g_{LO} and g_{LOR} were made, the storage requirements of an accurate approximation to

	LO & LSO detectors					
α	LO	MF	hp	triangular	LSO-power	Cauchy
1	76.7	4.6	48.1	67.4	80.1	80.8
	(5.9)	(4.7)	(4.9)	(4.5)	(4.4)	(5.4)
1.2	73.3	7.2	54.3	71.0	74.3	75.0
	(4.9)	(4.9)	(5.5)	(4.8)	(6.7)	(4.1)
1.4	73.1	11.4	55.8	72.1	73.6	71.0
	(6.7)	(5.9)	(4.5)	(4.1)	(5.0)	(4.8)
1.6	74.2	31.7	65.8	76.5	74.4	68.3
	(5.4)	(3.5)	(5.7)	(4.7)	(5.9)	(5.0)
1.8	77.1	56.9	71.8	75.6	77.9	64.5
	(5.8)	(4.8)	(5.1)	(6.0)	(4.8)	(3.9)
2	81.5	79.4	79.1	80.4	81.8	64.3
	(5.3)	(5.5)	(4.7)	(5.2)	(4.8)	(4.4)

TABLE 6.1. Detection and false alarm (bracketed) rates in % for linear correlation detectors of signals in SαS interference, $\theta = 0.5$.

	LOR & LSOR detectors		
α	LOR	Wilcoxon	LSOR-tr
1	75.4	58.4	75.7
	(3.8)	(6.9)	(6.2)
1.2	75.3	65.3	71.9
	(4.8)	(6.6)	(4.8)
1.4	71.2	69.2	70.3
	(5.2)	(4.5)	(6.7)
1.6	74.5	70.9	73.1
	(6.0)	(6.2)	(4.5)
1.8	77.0	77.3	77.1
	(3.6)	(4.1)	(6.0)
2	79.4	79.5	78.7
	(4.1)	(5.8)	(4.3)

TABLE 6.2. Detection and false alarm (bracketed) rates in % for linear rank-based detectors of signals in SαS interference, $\theta = 0.5$.

	LO & LSO detectors					
α	LO	MF	hp	triangular	LSO-power	Cauchy
1	65.6	5.3	32.9	57.3	64.0	80.8
	(5.2)	(4.4)	(4.4)	(5.1)	(5.6)	(5.4)
1.2	62.1	4.5	38.3	57.7	63.3	75.0
	(5.1)	(5.2)	(5.9)	(4.1)	(4.6)	(4.1)
1.4	61.8	6.3	45.6	59.9	61.7	71.0
	(5.4)	(4.8)	(3.5)	(4.9)	(5.0)	(4.8)
1.6	63.1	12.0	52.7	63.6	64.2	68.3
	(5.4)	(5.3)	(4.9)	(5.7)	(6.0)	(5.0)
1.8	66.4	39.0	59.8	67.8	66.4	64.5
	(4.8)	(4.5)	(5.1)	(5.4)	(4.4)	(3.9)
2	70.9	68.9	69.4	70.9	71.8	64.3
	(5.2)	(5.0)	(3.8)	(5.3)	(4.2)	(4.4)

TABLE 6.3. Detection and false alarm (bracketed) rates in % for quadratic correlation detectors of signals in SαS interference, $\theta = 0.5$.

	LOR & LSOR detectors		
α	LOR	Wilcoxon	LSOR-tr
1	65.8	45.2	65.1
	(4.3)	(5.4)	(4.9)
1.2	62.1	48.0	60.2
	(5.7)	(4.2)	(5.5)
1.4	62.7	56.4	62.2
	(4.3)	(6.2)	(4.6)
1.6	59.4	62.6	60.8
	(5.8)	(5.0)	(5.0)
1.8	63.4	61.9	62.4
	(4.4)	(5.3)	(5.0)
2	68.5	66.9	66.8
	(5.4)	(5.3)	(4.4)

TABLE 6.4. Detection and false alarm (bracketed) rates in % for quadratic rank-based (Adichie) detectors of signals in SαS interference, $\theta = 0.5$.

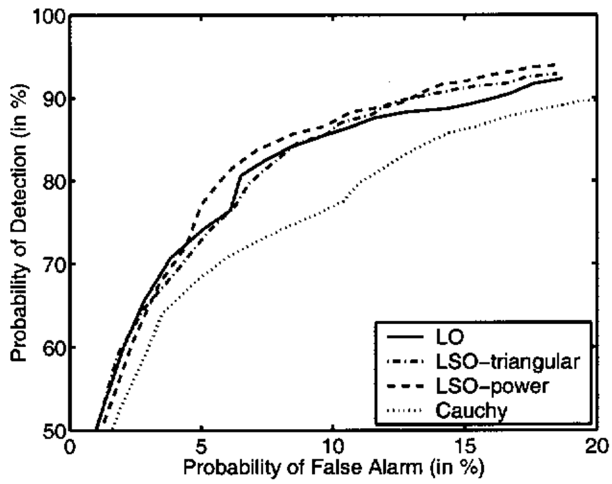


FIGURE 6.7. Receiver Operating Characteristics for some linear correlation detectors when $\alpha = 1.6$.

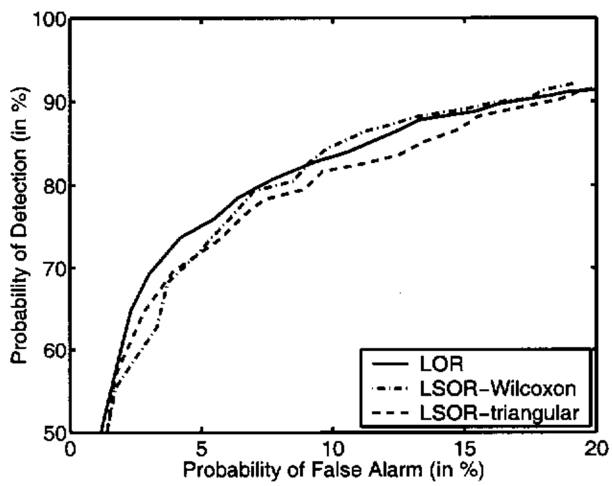


FIGURE 6.8. Receiver Operating Characteristics for some linear rank-based detectors when $\alpha = 1.6$.

g_{LO} would be very large since it extends from $-\infty$ to ∞, while the g_{LOR} can be stored *without further approximations* in an array of length n.

- The matched filter, achieves its highest detection rate for the Gaussian case $\alpha = 2$, where it is optimal, however it deteriorates rapidly as α decreases.

- The Cauchy detector is optimal when $\alpha = 1$ and its high detection rates, compared to other detectors, extends around $\alpha = 1$. However, it does decrease compared to the other detectors when α departs significantly from 1. It is important to note that this detector tested the simple alternative $\mathsf{K} : \theta = 0.5$, rather than the more general case of $\theta > 0$ that the other detectors used.

- The LSO-power nonlinearity appears to yield the best LSO detector of the nonlinearities tested here for non-Gaussian SαS distributions, in fact, little difference is noticeable compared to the LO detector.

- Similarly, the LSOR detector using the triangular nonlinearity, LSOR-tr, appears to follow the performance of the LOR and LO detectors very closely across the entire range of α values tested.

- Variations in the relative performance of the detectors may be attributable to the usual variation in estimating probabilities of detection and false alarm through a finite number of Monte Carlo simulations. It should also be remembered that the optimality of the LO detector is only for vanishing signal strength. These may be the reasons why it can occasionally be seen that LSO detectors achieve higher detection than the LO detector.

- The linear detectors appear to outperform the quadratic detectors. The quadratic detectors implicitly test the two-sided alternative, $\theta \neq 0$, while the linear detectors test the one-sided alternative, $\theta > 0$.

While these observations were based on results for $\theta = 0.5$ under K, since these detectors were designed to meet or approximate the locally optimum criterion, it is significant to note that similar results could be obtained for other values of θ. This is confirmed by the detection power graph in Figure 6.9 which shows that while the detection rate of the best performing detectors varies with signal strength, θ, their performance relative to the other detectors is similar. The power graphs for other detectors show similar results.

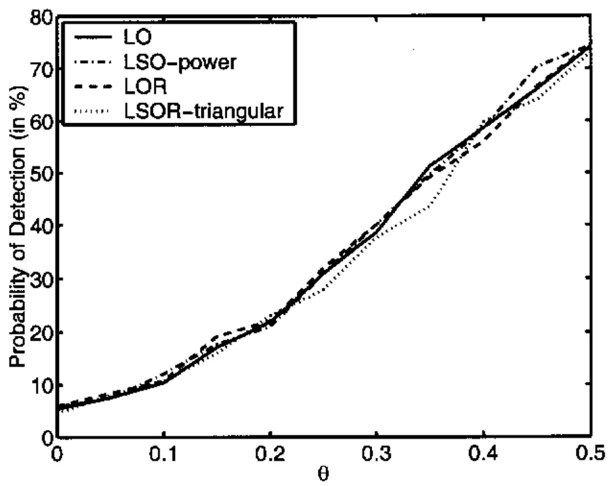


FIGURE 6.9. Detection power functions for the LO, LSO-power, LOR and LSOR-tr detectors for sample sizes of 100, 5% level of significance and SαS interference with $\alpha = 1.6$.

The detectors used in the results above assumed true knowledge of the parameter values to determine the shape of their nonlinearities and in the setting of thresholds. In Table 6.5 the detection and false alarm rates are shown for a selection of detectors when there is an error in the estimation of the parameter α. The interference was generated using $\alpha = 1.6$, however, the detectors were designed for $1 \leq \alpha \leq 2$.

As was discussed previously, the rank tests should maintain a constant false alarm rate when the interference is symmetrically distributed, while no such statement can be made about the LSO detectors. This is confirmed by the results shown, although the LO detectors' false alarm rates is close to the nominal level of 5%.

	Selection of detectors			
α	LO	LSO-power	LOR	LSOR-tr
1	69.8	69.3	59.3	57.8
	(5.6)	(7.4)	(5.2)	(5.0)
1.2	73.3	75.3	69.2	64.8
	(5.8)	(4.8)	(5.3)	(3.1)
1.4	75.8	74.9	73.1	70.7
	(5.2)	(4.9)	(5.7)	(4.2)
1.6	74.7	74.1	74.9	72.9
	(5.2)	(5.9)	(4.1)	(6.2)
1.8	74.9	75.0	72.6	74.0
	(4.5)	(5.7)	(4.1)	(5.8)
2	58.2	70.3	65.9	70.5
	(29.2)	(19.0)	(4.4)	(4.9)

TABLE 6.5. Detection and false alarm (bracketed) rates in % for a selection of detectors of signals in SαS ($\alpha = 1.6$) interference where the detector is designed for another α, $\theta = 0.5$

6.5 Conclusions

While the matched filter and Cauchy detectors are optimal when $\alpha = 2$ and 1 respectively, their performance deteriorated for other values of α. The LOR detector has been seen to achieve similar performance to the LO detector. It also has inherent computation and storage advantages for on-line detection. Additionally, the LSOR detector using the triangular rank score function, LSOR-tr, and the LSO detector using the LSO-power nonlinearity have been successfully introduced, showing similar performance to the LO and LOR detectors across all values of α tested. This has been achieved while maintaining computational simplicity. This is especially true of the LSOR-tr detector which, as a rank-based test, can also maintain a constant false alarm rate and high detection rates when parameter estimation errors occur.

6.6 Summary

The use of detectors approximating the locally optimum criterion has been investigated for the detection of known signals in impulsive interference modelled by an SαS process. The form of the LOR score functions for this interference has been provided. A number of nonlinear score functions for both correlation (LSO) detectors, as well as rank-based correlation detectors (LSOR) have been developed and their performance compared to the LO, LOR, matched filter and Cauchy detectors.

Only known signal detection was considered here in order to illustrate that the deterioration in performance when rank-based detectors are used may be small and may be outweighed by gains in computational simplicity and constant false alarm rates. The extension to the unknown signal case can be performed through parameter estimation and the adoption of a generalised approach. This is considered beyond the scope of this thesis. Rank-based detectors similar to those detailed here have also been considered for correlated interference [16, 4, 46, 48], multiplicative interference [5, 6] and random signals [78].

Chapter 7
Conclusions and Future Directions

Impulsive behaviour can be observed in many naturally occurring and man-made processes. Much more attention has been paid in recent engineering literature to dealing with it. Many models for this behaviour have been presented, including the alpha-stable (αS) family. Even in the absence of strong evidence to support the αS distribution as a statistical-physical model of the process, the properties and broad nature of this distribution may often warrant its investigation as an empirical model.

Based on the signal detection scheme in Figure 1.1, this dissertation discussed and addressed several of the unique problems surrounding the use of αS processes to model impulsive behaviour. Specifically

1. Investigation of αS parameter estimation procedures was performed and resulted in a proposal to dynamically determine the regression region for the extensively used Koutrouvelis [50] estimation procedure.

2. Tests for Gaussianity using the kernel characteristic function estimator (KCFE) were investigated and extended to the multivariate case – contributing to the body of literature answering the "Assume Gaussianity" block of Figure 1.1..

3. Drawing on characteristic function (cf) based Gaussianity tests, a goodness-of-fit test for the αS distribution was described. The parametric bootstrap procedure was utilised to estimate the distribution of the test statistic.

4. A variation of the above test was formulated that tested the level of impulsive behaviour of αS processes through changes in the cf. This test was compared to testing the parameters of the αS distribution directly. In both cases, the

parametric bootstrap was used to determine critical values. This and the previous item can be used to determine if it is possible to "Assume other models", namely the αS model.

5. Rank based tests were designed to meet or approximate the locally optimum criterion and were compared to existing locally optimum and some novel suboptimum detectors of known signals in symmetric αS interference. Results and analysis show strong arguments for the use of rank tests in this scenario based on detection performance and computational complexity. While these detectors have been designed to achieve local optimality or suboptimality for symmetric αS interference, the rank-based detectors may be readily applied when other symmetric interference is encountered, particularly if it is impulsive. Thus they contribute to "Nonparametric detection".

The contents of this dissertation have covered several topics, although most is concentrated on goodness-of-fit testing and rank-based signal detection. It is noted that extensions to the work are possible in the goodness-of-fit testing

- using the KCFE or other cf estimators in place of the empirical cf to reduce fluctuations due to limited sample sizes
- determining appropriate test statistics to increase sensitivity to particular classes of alternatives – as has been done in empirical distribution function based goodness-of-fit tests
- other test statistics may have distributions that can be found exactly or approximately, thus removing the need for the computationally expensive parametric bootstrap procedures

and detection

- the investigation of different detection scenarios are possible, including the presence correlated or multiplicative interference and unknown or random signals

- validation of the presented procedures on real data.

Appendix A

Characteristic Functions and Their Estimation

The characteristic function (cf) of a random variable, X, is

$$\phi_X(t) = \mathsf{E}\left[e^{jtX}\right]$$

where $\mathsf{E}[\cdot]$ is the expectation operator. By replacing $\mathsf{E}[\cdot]$, an alternative expression, and interpretation, of the cf is

$$\phi_X(t) = \int_{-\infty}^{\infty} f_X(x) e^{jxt} dx \quad .$$

That is, the cf is the Fourier transform of the pdf, the only difference being a reversal of the sign of t. The cf has the advantage that it always exists, has a smooth form and is well behaved.

A.1 Empirical Characteristic Function

The sample or empirical characteristic function (ecf) is a classical, unbiased cf estimator. The ecf of independent and identically distributed (iid) observations, $\boldsymbol{X} = [X_1, X_2, \ldots, X_n]^T$, is

$$\hat{\phi}^e_{\mathbf{X}}(t) = \frac{1}{n} \sum_{i=1}^{n} e^{jtX_i} \quad . \tag{A.1}$$

Its expected value and variance are [26]

$$\begin{aligned} \mathsf{E}\left[\hat{\phi}^e_{\mathbf{X}}(t)\right] &= \phi_X(t) \\ \mathsf{var}\left[\Re\left(\hat{\phi}^e_{\mathbf{X}}(t)\right)\right] &= \frac{1}{2n}\left\{1 + \Re\left(\phi_X(2t)\right) - 2\left(\Re\left(\phi_X(t)\right)\right)^2\right\} &\text{(A.2)} \\ \mathsf{var}\left[\Im\left(\hat{\phi}^e_{\mathbf{X}}(t)\right)\right] &= \frac{1}{2n}\left\{1 - \Re\left(\phi_X(2t)\right) - 2\left(\Im\left(\phi_X(t)\right)\right)^2\right\} \quad . &\text{(A.3)} \end{aligned}$$

The ecf is an unbiased cf estimator, however it has high variance that does not decay to zero with large t. Since $\phi_X(t) \to 0$ as $|t| \to \infty$, equations (A.2) and (A.3) approach $\frac{1}{2n}$ as $|t| \to \infty$, see Figure A.1.

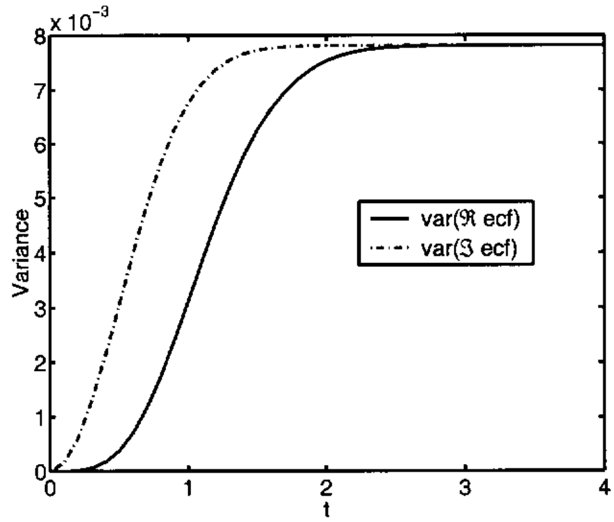


FIGURE A.1. The variance of the real and imaginary components of the ecf of 64 observations from a Gaussian process.

A.2 Kernel Characteristic Function Estimator

The development of the KCFE was motivated by the analogy with the theory of kernel density estimation of pdfs [75]. While kernel density estimation involves convolving a kernel function with the zero-width histogram, due to the Fourier relationship between the pdf and cf, the KCFE, $\hat{\phi}_{\mathbf{X}}(t;\varphi)$, is produced by multiplying the ecf by a kernel, $\varphi_{\mathbf{X}}(t)$, that is

$$\hat{\phi}_{\mathbf{X}}(t;\varphi) = \hat{\phi}^e_{\mathbf{X}}(t)\varphi_{\mathbf{X}}(t) \quad . \tag{A.4}$$

An appropriate kernel function should satisfy

- $\varphi_{\mathbf{X}}(0) = 1$
- $\varphi_{\mathbf{X}}(t) \leq \varphi_{\mathbf{X}}(0)$, for $-\infty < t < \infty$

- $\varphi_{\mathbf{X}}(t) \to 0$, as $|t| \to \infty$.

One choice, as used in [93], is the Gaussian function itself

$$\begin{aligned} \varphi(t) &= e^{-\sigma_\varphi^2 t^2/2} \\ \sigma_\varphi &= 0.97\sigma n^{-0.2} \end{aligned} \tag{A.5}$$

where σ^2 is the variance of the observed process.

By multiplying the ecf by a kernel in (A.4) the magnitude of the variations in the large t region of the KCFE are reduced, see Figure A.2. A bias is introduced since the expected value of the KCFE is also reduced by this operation

$$\begin{aligned} \mathsf{E}\left[\hat{\phi}_{\mathbf{X}}(t;\varphi)\right] &= \mathsf{E}\left[\hat{\phi}^e_{\mathbf{X}}(t)\ \varphi_{\mathbf{X}}(t)\right] \\ &= \phi_X(t)\ \varphi_{\mathbf{X}}(t) \quad . \end{aligned}$$

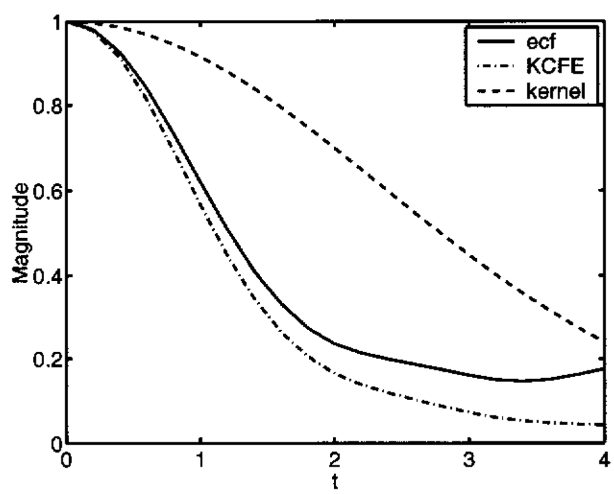


FIGURE A.2. The effect of multiplicative smoothing on the ecf of one realisation of a Gaussian process with $n = 64$ observations.

Appendix B

Hypothesis Testing with the Parametric Bootstrap

The bootstrap is a versatile computation statistical tool for estimating the sample distribution of statistics. Since being introduced by Efron [22], it has been applied in a significant number of problems when difficulties arise in applying standard methods. Examples include when large sample methods are inapplicable due to small sample sizes, or, as in the considered cases here, when complexity of the problem makes analytical solutions unfeasible.

With the dramatic increases in computational power available to researchers and engineers in recent years, the bootstrap has found application in a range of problems which would otherwise be too difficult with traditional statistical analysis.

A well designed bootstrap procedure can replace intensive mathematical analysis with computational load. Random resampling of the data is performed and the statistic of interest is recalculated. When repeated a large number of times, the sample distribution of the recalculated statistics will approximate the true distribution. In this way, the bootstrap procedure repeats the "experiment", just as a researcher would.

Here, the statistics of interest are test statistics $T_n = T_n(\boldsymbol{X})$ calculated from a finite set of random data, $\boldsymbol{X} = [X_1, X_2, \dots, X_n]^T$. The data has a joint distribution $P^n_{\boldsymbol{\xi},\boldsymbol{\vartheta}}$ and is parameterised by the unknown parameter vectors $\boldsymbol{\xi} \in \Xi$ and $\boldsymbol{\vartheta} \in \Theta$. The form of T_n will depend on the hypothesis to be tested.

It may be that T_n will depend on only *some* of the parameters of the distribution of $\boldsymbol{X}$. These parameters of interest shall be deemed to be the elements of the parameter vector $\boldsymbol{\xi}$, while the other, nuisance, parameters are contained in $\boldsymbol{\vartheta}$.

The null and alternative hypotheses can be expressed in terms of the distribution $P^n_{\boldsymbol{\xi},\boldsymbol{\vartheta}}$ and the parameter values

$$\mathsf{H} : \boldsymbol{X} \sim P^n_{\boldsymbol{\xi}_0,\boldsymbol{\vartheta}}, \quad \boldsymbol{\xi}_0 \in \Xi, \quad \boldsymbol{\vartheta} \in \Theta \tag{B.1}$$

versus

$$\mathsf{K} : \boldsymbol{X} \sim P^n_{\boldsymbol{\xi},\boldsymbol{\vartheta}}, \quad \boldsymbol{\xi} \in \Xi, \quad \boldsymbol{\xi} \neq \boldsymbol{\xi}_0, \quad \boldsymbol{\vartheta} \in \Theta \tag{B.2}$$

or alternatively, by assuming $\boldsymbol{X} \sim P^n_{\boldsymbol{\xi},\boldsymbol{\vartheta}}$, $\boldsymbol{\xi} \in \Xi$, $\boldsymbol{\vartheta} \in \Theta$ the hypothesis and alternative become

$$\mathsf{H} : \boldsymbol{\xi} = \boldsymbol{\xi}_0 \tag{B.3}$$

versus

$$\mathsf{K} : \boldsymbol{\xi} \neq \boldsymbol{\xi}_0 \quad . \tag{B.4}$$

By comparing T_n to an appropriately chosen critical value, $d_n(\gamma, \boldsymbol{\xi}_0, \hat{\boldsymbol{\vartheta}})$, the outcome of a test denoted by the test function $\varphi_n(\boldsymbol{X})$, is determined by

$$\varphi_n(\boldsymbol{X}) = \begin{cases} 1\,, & \text{if } T_n > d_n(\gamma, \boldsymbol{\xi}_0, \hat{\boldsymbol{\vartheta}}) \\ 0\,, & \text{otherwise} \end{cases} \tag{B.5}$$

$d_n(\gamma, \boldsymbol{\xi}_0, \hat{\boldsymbol{\vartheta}})$ is chosen such that the test has a level, or probability of Type I error, ζ.

When there is difficulty in finding the distribution of T_n, and in the consequent setting of $d_n(\gamma, \boldsymbol{\xi}_0, \hat{\boldsymbol{\vartheta}})$, the parametric bootstrap tests may be used. The bootstrap procedure appropriate to the conditions and hypotheses described previously is outlined in Table B.1.

When designing powerful bootstrap tests, it is critical to determine the best form of resampling to use in order to best approximate the distribution of T_n. The resampling scheme used will depend on the hypothesis to be tested. In the procedure described above, note that resampling was performed by a random number generator using the distribution $P^n_{\boldsymbol{\xi}_0,\hat{\boldsymbol{\vartheta}}}$.

TABLE B.1. The parametric bootstrap procedure for testing the hypothesis $\mathsf{H} : \boldsymbol{\xi} = \boldsymbol{\xi}_0$ against $\mathsf{K} : \boldsymbol{\xi} \neq \boldsymbol{\xi}_0$.

Step 1. *Parameter estimation :* Estimate the parameters $\boldsymbol{\xi}$ and $\boldsymbol{\vartheta}$ from the sample $\boldsymbol{X}$ and denote them $\hat{\boldsymbol{\xi}}$ and $\hat{\boldsymbol{\vartheta}}$, respectively.

Step 2. *Calculate test statistic :* Evaluate $T_n = T_n(\boldsymbol{X})$.

Step 3. *Resampling :* Using a pseudo-random number generator, generate B independent resamples $\boldsymbol{X}_b^* = [X_{1b}^*, \ldots, X_{nb}^*]^T$, $b = 1, \ldots, B$ from the distribution $P^n_{\boldsymbol{\xi}_0, \hat{\boldsymbol{\vartheta}}}$.

Step 4. *Calculate bootstrap statistics :* Evaluate $T_{n,b}^* = T_n\left(\boldsymbol{X}_b^*\right)$ for $b = 1, \ldots, B$.

Step 5. *Ranking :* Rank the collection $T_{n,1}^*, T_{n,2}^*, \ldots, T_{n,B}^*$ into increasing order to obtain $T_{n,(1)}^* \leq T_{n,(2)}^* \leq \cdots \leq T_{n,(B)}^*$.

Step 6. *Critical value :* Set $d_n(\gamma, \boldsymbol{\xi}_0, \hat{\boldsymbol{\vartheta}}) = T_{n,(C)}$, where $C = \lceil \zeta(B+1) \rceil$.

Step 7. *Test :* Compare T_n to $d_n(\gamma, \boldsymbol{\xi}_0, \hat{\boldsymbol{\vartheta}})$ in (B.5).

In the absence of nuisance parameters, a fixed, simple H will *completely* specify the distribution of $\boldsymbol{X}$ under the null and consequently, the distribution of T_n is fixed, though maybe unknown. When this is the case, it may be possible to approximate the distribution of T_n under H through a "once off" Monte Carlo simulation.

However, when nuisance parameters are present, they may affect the distribution of T_n and make it necessary to estimate this distribution for each distinct $\hat{\boldsymbol{\vartheta}}$. In the special cases where it can be proven that T_n is independent of $\boldsymbol{\vartheta}$, it may be possible to revert to a Monte Carlo simulation to determine critical values, since, again, the distribution of T_n is fixed under H.

Beran [7] has proven two important properties of bootstrap tests, which supports the approach just described. Under some assumptions, described in Table B.2, on the convergence of the parameter estimates $\hat{\boldsymbol{\vartheta}}$ and of the distribution of the test statistic T_n, it can be shown that the test is asymptotically correct, and its power can be consistently estimated by a related bootstrap power estimator.

TABLE B.2. Beran's conditions to prove the asymptotically correctness of a bootstrap test and that its power can be consistently estimated [7].

BERAN'S CONDITIONS: *Suppose* Ξ *is* $\mathbb{R}^k$ *and the following requirements are met*

1. $P^{\infty}_{\boldsymbol{\xi},\boldsymbol{\vartheta}}\left[\lim_{n\to\infty}\hat{\boldsymbol{\vartheta}} = \boldsymbol{\vartheta}\right] = 1$
2. $\lim_{n\to\infty}\sup_{\boldsymbol{\xi}} P^{n}_{\boldsymbol{\xi},\boldsymbol{\vartheta}}[m_2(\hat{\boldsymbol{\vartheta}},\boldsymbol{\vartheta}) > \varepsilon] = 0$ *for every positive* ε, *where* m_2 *is a metric on* Θ.
3. *If* $\{(\boldsymbol{\xi}_n,\boldsymbol{\vartheta}_n) \in \Omega; n \geq 1\}$ *is any sequence such that* $\lim_{n\to\infty}\sqrt{n}(\boldsymbol{\xi}_n - \boldsymbol{\xi}_0) = \boldsymbol{h}$ *for some* $\boldsymbol{h} \in \mathbb{R}^k$ *and* $\lim_{n\to\infty}\boldsymbol{\vartheta}_n = \boldsymbol{\vartheta}$, *then* $K_{n,T}(\boldsymbol{\xi}_n,\boldsymbol{\vartheta}_n) \Rightarrow K_T^{(\boldsymbol{h})}(\boldsymbol{\xi}_0,\boldsymbol{\vartheta})$, *a limit distribution which is continuous and does not depend upon the particular sequence* $\{(\boldsymbol{\xi}_n,\boldsymbol{\vartheta}_n)\}$ *chosen. Moreover,* $K_T^{(0)}(\boldsymbol{\xi}_0,\boldsymbol{\vartheta})$ *has a strictly monotone survival function.*
4. *If* $\{(\boldsymbol{\xi}_n,\boldsymbol{\vartheta}_n) \in \Omega, n \geq 1\}$ *is any sequence such that* $\lim_{n\to\infty}\sqrt{n}|\boldsymbol{\xi}_n - \boldsymbol{\xi}_0| = \infty$ *and* $\lim_{n\to\infty}\boldsymbol{\vartheta}_n = \boldsymbol{\vartheta}$ *then* $\lim_{n\to\infty}K_{n,T}(x;\boldsymbol{\xi}_n,\boldsymbol{\vartheta}_n) = 1$ *for every finite real* x.

These results present a very powerful argument and validation for bootstrap hypothesis tests. The principle of hypothesis testing using the bootstrap is discussed in more detail in [32, 92].

References

[1] J. N. Adichie. Rank tests in linear models. In P. R. Krishnaiah and P. K. Sen, editors, *Nonparametric Methods*, volume 4 of *Handbook of Statistics*, chapter 11, pages 229–257. Elsevier Science Publishers, Amsterdam, The Netherlands, 1984.

[2] R. J. Adler, R. E. Feldman, and M. S. Taqqu, editors. *A Practical Guide to Heavy Tails: Statistical Techniques and Applications*. Birkhäuser, 1998.

[3] S. Ambike, J. Ilow, and D. Hatzinakos. Detection of binary transmission in a mixture of Gaussian noise and impulsive noise modeled as an alpha-stable process. *IEEE Signal Processing Letters*, 1(3):55–57, March 1994.

[4] J. Bae, S. I. Park, and I. Song. A known-signal detector based on ranks in weakly dependent noise. *Signal Processing*, 54(3):309–314, November 1996.

[5] J. Bae and I. Song. Rank-based detection of weak random signals in a multiplicative noise model. *Signal Processing*, 63(2):121–131, December 1997.

[6] J. Bae, I. Song, H. Morikawa, and T. Aoyama. Nonparametric detection of known signals based on ranks in multiplicative noise. *Signal Processing*, 60(2):255–261, July 1997.

[7] R. Beran. Simulated power functions. *The Annals of Statistics*, 14(1):151–173, 1986.

[8] K. L. Blackard, T. S. Rappaport, and C. W. Bostian. Measurements and models of radio frequency impulsive noise for indoor wireless communications. *IEEE Journal on Selected Areas in Communications*, 11(7):991–1001, September 1993.

[9] T.K. Blankenship, D.M. Kriztman, and T.S. Rappaport. Measurements and simulation of radio frequency impulsive noise in hospitals and clinics. In *Proceedings of the IEEE 47th Vehicular Technology Conference. Technology in Motion*, volume 3, pages 1942–1946, 1997.

[10] M. Bouvet and S. C. Schwartz. Comparison of adaptive and robust receivers for signal detection in ambient underwater noise. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 37(5):621–626, May 1987.

[11] K. O. Bowman and L. R. Shenton. Moment $(\sqrt{b_1}, b_2)$ techniques. In R. B. D'Agostino and M. A. Stephens, editors, *Goodness-of-Fit Techniques*, pages 279–329. Marcel Dekker, 1986.

[12] J. V. Bradley. *Distribution-Free Statistical Tests*. Prentice-Hall, 1968.

[13] C. L. Brown and S. Saliu. Testing of alpha-stable distributions with the characteristic function. In *Proceedings of the IEEE Signal Processing Workshop on Higher-Order Statistics, HOS'99*, pages 224–227, Caesarea, Israel, June 1999.

[14] C. L. Brown and A. M. Zoubir. Testing for impulsive behaviour: A bootstrap approach. *Digital Signal Processing.* (in press).

[15] C. L. Brown and A. M. Zoubir. On the estimation of the parameters of alpha-stable distributions using a truncated linear regression. In *Proceedings of the 9th IEEE Signal Processing Workshop on Statistical Signal and Array Processing, SSAP'98*, Portland, USA, September 1998.

[16] C. L. Brown and A. M. Zoubir. Locally suboptimal and rank-based known signal detection in correlated alpha-stable interference. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP'00*, Istanbul, Turkey, June 2000.

[17] C. L. Brown and A. M. Zoubir. A nonparametric approach to signal detection in impulsive interference. *IEEE Transactions on Signal Processing*, 48(9), September 2000. (to appear).

[18] C. L. Brown, A. M. Zoubir, and B. Boashash. On the performance of an adaptation of Adichie's rank tests for signal detection: and its relationship to the matched filter. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP'98*, pages 1505–1508, Seattle, USA, May 1998.

[19] P. Comon and L. Deruaz. Normality tests for coloured samples. In *IEEE-ATHOS Workshop on Higher-Order Statistics*, pages 217–221, Begur, Spain, 12–14 June 1995.

[20] R. B. D'Agostino. Tests for the normal distribution. In R. B. D'Agostino and M. A. Stephens, editors, *Goodness-of-Fit Techniques*, pages 367–420. Marcel Dekker, 1986.

[21] J. W. Dalle Molle and M. J. Hinich. Trispectal analysis of stationary random time series. *Journal of the Acoustical Society of America*, 97:2963–2978, 1995.

[22] B. Efron. Bootstrap methods: another look at the jackknife. *The Annals of Statistics*, 7(1):1–26, 1979.

[23] B. Efron and R. Tibshirani. *An Introduction to the Bootstrap.* Chapman and Hall, 1993.

[24] T. W. Epps. Testing that a stationary time series is Gaussian. *The Annals of Statistics*, 15:1683–1698, 1987.

[25] E. Fama and R. Roll. Parameter estimates for symmetric stable distributions. *Journal of the American Statistical Association*, 66:331–338, June 1971.

[26] A. Feuerverger and A. Mureika. The Empirical Characteristic Function and its Applications. *The Ann. of Statist.*, 5(1):88–97, 1977.

[27] R. C. Geary. Testing for normality. *Biometrika*, 34:209–242, 1947.

[28] J. D. Gibson and J. L. Melsa. *Introduction to Nonparametric Detection with Applications.* Academic Press, 1975.

[29] J. Hájek, Z. Sidak, and P. K. Sen. *Theory of rank tests.* Academic Press, 1999.

[30] P. Hall. A comedy of errors: The canonical form for a stable characteristic function. *Bulletin of the London Mathematical Society*, 13:23–7, May 1980.

[31] P. Hall. On Kullback-Leiber loss and density estimation. *The Annals of Statistics*, 15:1491–1519, 1987.

[32] P. Hall. *The Bootstrap and Edgeworth Expansion.* Springer-Verlag, New York, 1992.

[33] T. P. Hettmansperger. *Statistical Inference Based on Ranks.* John Wiley, New York, 1984.

[34] M. J. Hinich. Testing for Gaussianity and Linearity of a Stationary Time Series. *J. Time Series Anal.*, 3:169–176, 1982.

[35] I. Holmes, K. H. Kobo, Cody R. J., and P. Kligfield. Arrhythmias in ischemic and nonischemic dilated cardiomyopathy: prediction of mortality by ambulatory electrocardiology. *American J Cardiolog*, 55:146–151, 1985.

[36] J. Ilow and D. Hatzinakos. Analytic alpha-stable noise modeling in a Poisson field of interferers or scatterers. *IEEE Transactions on Signal Processing*, 46(6):1601–1611, June 1998.

[37] J. Ilow and D. Hatzinakos. Applications of the empirical characteristic function to estimation and detection problems. *Signal Processing*, 65:199–219, March 1998.

[38] J. Ilow, D. Hatzinakos, and A. N. Venetsanopoulos. Detection of narrowband signals in alpha-stable noise environments. In *Proceedings of MILCOM '94*, pages 287–290, Fort Monmouth, NJ, USA, October 1994.

[39] D. R. Iskander. *The Generalised Bessel Function K Distribution and its Application to the Detection of Signals in the Presence of Non-Gaussian Interference.* PhD thesis, Queensland University of Technology, Brisbane, Australia, February 1997.

[40] D. R. Iskander and A. M. Zoubir. Testing Gaussianity of Multivariate Data Using Entropy. In *Fourth International Symposium on Signal Processing and Its Applications*, pages 109–112, Gold Coast, Australia, August 1996.

[41] R. Kapoor, A. Banerjee, G. A. Tsihrintzis, and N. Nandhakumar. UWB radar detection of targets in foliage using alpha-stable clutter models. *IEEE Transactions on Aerospace and Electronic Systems*, 35(3):819–834, July 1999.

[42] A. Karasaridis and D. Hatzinakos. On the modeling of network traffic and fast simulation of rare events using α-stable self-similar processes. In *Proceedings of the IEEE Signal Processing Workshop on Higher-Order Statistics, HOS'97*, pages 268–272, Banff, Canada, July 1997.

[43] A. Karasaridis and D. Hatzinakos. Bandwidth allocation bounds for α-stable self-similar internet traffic models. In *Proceedings of the IEEE Signal Processing Workshop on Higher-Order Statistics, HOS'99*, pages 214–218, Caesarea, Israel, June 1999.

[44] S. A. Kassam. *Signal detection in non-Gaussian noise.* Springer texts in electrical engineering. Springer-Verlag, New York, 1988.

[45] U. Kaul et al. Prognostic implications of complex ventricular ectopy in patients with and without structural heart disease. A study based on programmed electrical stimulation. *Int J Cardiol*, 14(1):79–89, January 1987.

[46] K. S. Kim, S. Y. Kim, I. Song, and S. R. Park. Locally optimum detector for correlated random signals in a weakly dependent noise model. *Signal Processing*, 74(3):317–322, 1999.

[47] K. S. Kim, S. I. Park, I. Song, and J. Bae. Performance of ds/ssma systems using tcm under impulsive noise environment. *Signal Processing*, 64(2):225–230, 1998.

[48] T. Kim, J. S. Yun, I. Song, and Y. J. Na. Comparison of known signal detection schemes under a weakly dependent noise model. *IEE Proceedings-Vision, Image and Signal Processing*, 141(5):303–310, October 1994.

[49] I. A. Koutrouvelis. A goodness-of-fit test of simple hypotheses based on the empirical characteristic function. *Biometrika*, 67(1):238–40, 1980.

[50] I. A. Koutrouvelis. Regression-type estimation of parameters of stable laws. *Journal of the American Statistical Association*, 75(372):918–928, December 1980.

[51] I. A. Koutrouvelis and J. Kellermeier. A goodness-of-fit test based on the empirical characteristic function when parameters must be estimated. *Journal of the Royal Statistical Society*, 43(2):173–176, 1981.

[52] E. E. Kuruoğlu, W. J. Fitzgerald, and P. J. W. Rayner. Near optimal detection of signals in impulsive noise modeled with a symmetric α-stable distribution. *IEEE Communications Letters*, 2(10):282–284, October 1998.

[53] E. E. Kuruoğlu, C. Molina, and W. J. Fitzgerald. Approximation of α-stable probability densities using finite Gaussian mixtures. In *Proceedings of EUSIPCO'98*, Rhodes, Greece, September 1998.

[54] E. E. Kuruoğlu, P. J. W. Rayner, and W. J. Fitzgerald. A near optimal receiver for detection in α-stable distributed noise. In *SSAP98*, pages 419–422, Portland, USA, September 1998.

[55] E. L. Lehmann. *Nonparametrics: Statistical Methods Based on Ranks.* Holden-Day, San Francisco, 1975.

[56] E. L. Lehmann. *Testing Statistical Hypotheses.* Wiley, 1986.

[57] R. A. Leitch and A. S. Paulson. Estimation of stable law parameters: Stock price behavior application. *Journal of the American Statistical Association*, 70(351):690–697, September 1975.

[58] B. Mandelbrot. The variation of speculative prices. *Journal of Business*, 36:394–419, October 1963.

[59] E. B. Manoukian. *Modern concepts and theorems of mathematical statistics.* Springer-Verlag, 1986.

[60] D. Middleton. Statistical-physical models of electromagnetic interference. *IEEE Transactions on Electromagnetic Compatibility*, EMC-19(3):106–127, August 1977.

[61] D. Middleton. Non-Gaussian noise models in signal processing for telecommunications: New methods and results for class a and class b noise models. *IEEE Transactions on Information Theory*, 45(4):1129–1149, May 1999.

[62] D. S. Moore. Tests of the chi-squared type. In R. B. D'Agostino and M. A. Stephens, editors, *Goodness-of-Fit Techniques*, pages 63–96. Marcel Dekker, 1986.

[63] G. S. Mudholkar, D. K. Srivastava, and C. T. Lin. Some p-variate adaptations of the Shapiro-Wilk test of normality. *Communications in Statistics, Theory and Methods*, 24(4):953–985, 1995.

[64] C. L. Nikias and M. Shao. *Signal Processing with Alpha-Stable Distributions and Applications.* John Wiley & Sons, 1995.

[65] R. R. Officer. The distribution of stock returns. *Journal of the American Statistical Association*, 67:807–812, 1972.

[66] P. Papantoni-Kazakos and D. Kazakos. *Nonparametric Methods in Communications.* Marcel Dekker, 1977.

[67] E. S. Pearson, R. B. D'Agostino, and K. O. Bowman. Test for departure from normality. *Biometrika*, 64:231–246, 1977.

[68] H. V. Poor. *An Introduction to Signal Detection and Estimation.* Springer-Verlag, 2 edition, 1988.

[69] S. Saliu, N. Edvardsson, D. Poçi, and C. Brown. Alpha-stable modeling of arrhythmias on heart rate variability signals. In *Proceedings of the Third International Workshop on Biosignal Interpretation, BSI99*, Chicago, USA, June 1999.

[70] G. Samorodnitsky and S. Taqqu. *Stable NonGaussian Random Processes, Stochastic Models with Infinite Variance.* Chapman Hall, 1994.

[71] J. Shao and D. Tu. *The Jackknife and Bootstrap.* Springer, 1995.

[72] M. Shao and C. L. Nikias. Signal detection in impulsive noise based on stable distributions. In *Conference Record of The Twenty-Seventh Asilomar Conference on Signals, Systems and Computers*, pages 218–222, Pacific Grove, CA, USA, November 1993.

[73] M. Shao and C. L. Nikias. Signal processing with fractional lower order moments: Stable processes and their applications. *Proceedings of the IEEE*, 81(7):986–1010, July 1993.

[74] S. S. Shapiro, M. B. Wilk, and H. J. Chen. Comparative study of various tests of normality. *J. Amer. Statis. As.*, 63:1343–1372, 1968.

[75] B. W. Silverman. *Density estimation for statistics and data analysis.* Chapman and Hall, 1986.

[76] I. Song and S. A. Kassam. Locally optimum detection of signals in a generalized observation model: The known signal case. *IEEE Transactions on Information Theory*, 36(3):502–515, May 1990.

[77] I. Song and S. A. Kassam. Locally optimum detection of signals in a generalized observation model: The random signal case. *IEEE Transactions on Information Theory*, 36(3):516–530, May 1990.

[78] I. Song and S. A. Kassam. Locally optimum rank detection of correlated random signals in additive noise. *IEEE Transactions on Information Theory*, 38:1311–1322, July 1992.

[79] M. A. Stephens. Tests based on edf statistics. In R. B. D'Agostino and M. A. Stephens, editors, *Goodness-of-Fit Techniques*, pages 97–194. Marcel Dekker, 1986.

[80] T. Subba Rao and M. M. Gabr. Testing for Linearity of a Stationary Time Series. *J. Time Series Anal.*, 1:145–158, 1980.

[81] A. Swami and B. Sadler. TDE, DOA, and related parameter estimation problems in impulsive noise. In *Proceedings of the IEEE Signal Processing Workshop on Higher-Order Statistics, HOS'97*, pages 273–277, Banff, Canada, July 1997.

[82] Task Force of European Society of Cardiology and the North American Society of Pacing and Electrophysiology. Heart rate variability, standard of measurement, physiological interpretation, and clinical use. *Circulation*, 93:1043–1065, 1996.

[83] J. B. Thomas. Nonparametric detection. *Proceedings of the IEEE*, 58:623–631, 1970.

[84] G.A. Tsihrintzis and C.L. Nikias. Detection and classification of signals in impulsive noise modeled as an alpha-stable process. In *Conference Record of The Twenty-Seventh Asilomar Conference on Signals, Systems and Computers*, pages 707–710, Pacific Grove, CA, USA, November 1993.

[85] G.A. Tsihrintzis and C.L. Nikias. Performance of optimum and suboptimum receivers in the presence of impulsive noise modeled as an alpha-stable process. In *Proceedings of MILCOM '93*, pages 658–662, Boston, MA, USA, October 1993.

[86] G.A. Tsihrintzis and C.L. Nikias. Signal detection in incompletely characterized impulsive noise modeled as a stable process. In *Proceedings of MILCOM '94*, pages 271–275, Fort Monmouth, NJ, USA, October 1994.

[87] G.A. Tsihrintzis and C.L. Nikias. Performance of optimum and suboptimum receivers in the presence of impulsive noise modeled as an alpha-stable process. *IEEE Transactions on Communications*, 43(2-4):904–914, March 1995.

[88] H. L. Van Trees. *Detection, Estimation, and Modulation Theory. Part I.* John Wiley and Sons, New York, 1968.

[89] X. Wang and H. V. Poor. Robust multiuser detection in non-Gaussian channels. *IEEE Transactions on Signal Processing*, 47(2):289–305, February 1999.

[90] E. J. Wegman, S. G. Schwartz, and J. B. Thomas, editors. *Topics in Non-Gaussian Signal Processing.* Academic Press, New York, 1989.

[91] S. M. Zabin and G. A. Wright. Nonparametric density estimation and detection in impulsive interference channels – Part II: Detectors. *IEEE Transactions on Communications*, 42:1698–1711, 1994.

[92] A. M. Zoubir. Signal detection using the bootstrap. *Signal Processing*, (in press).

[93] A. M. Zoubir and M. J. Arnold. Testing Gaussianity with the characteristic function. *Signal Processing*, 53:245–255, September 1996.

[94] A. M. Zoubir and B. Boashash. The bootstrap and its application in signal processing. *IEEE Signal Processing Magazine*, 15(1):56–76, January 1998.

[95] A. M. Zoubir and C. L. Brown. Testing for impulsive interference: A bootstrap approach. In *Proceedings of the 1999 Workshop on Defense Applications of Signal Processing (DASP99)*, LaSalle, USA, August 1999.

[96] A. M. Zoubir, C. L. Brown, and B. Boashash. Testing multivariate Gaussianity with the characteristic function. In *Proceedings of the IEEE Signal Processing Workshop on Higher-Order Statistics, HOS'97*, pages 438–442, Banff, Canada, July 1997.