USC-SIPI REPORT #208

Single Sensor Detection and Classification of Multiple Sources by Higher-Order Spectra

by

Mithat C. Dogan and Jerry M. Mendel

June 1992

Signal and Image Processing Institute
UNIVERSITY OF SOUTHERN CALIFORNIA
Department of Electrical Engineering-Systems
3740 McClintock Avenue, Room 404
Los Angeles, CA 90089-2564 U.S.A.

Abstract

We propose a method to detect the number of non-Gaussian sources using higher-order statistics. The distinguishing feature of our method is its ability to perform detection with single sensor data unlike existing covariance-based detection algorithms that require an array of sensors. In addition, our method is blind to Gaussian observation noise. After the detection procedure, we propose an algorithm for classification of sources employing a priori knowledge of their spectra. Simulation results indicate the performance of our algorithms.

Contents

List of Figures

List of Tables

Chapter 1

INTRODUCTION

In many problems in signal processing, observations can be modeled as a superposition of an unknown number of signals corrupted by additive noise. An important issue is to detect the number of sources that emit the waveforms and classify them using a priori information about their statistical characteristics.

Existing approaches to the *multiple* source detection problem employ *multichannel* data and utilize information-theoretic criteria for model selection, as introduced by Akaike (AIC) or by Schwartz and Rissanen (MDL) [6]. The number of signals is determined as the value for which one of these criteria is optimized. If, however, multichannel data is not available (e.g., when observations are limited to data from a *single* sensor), these approaches do not work. This is the problem we address in this report. In addition, after estimating the number of sources, it is important to classify them. We propose a classification method that utilizes a priori knowledge of the shape of the spectrum of the sources.

The report is organized as follows: in Chapter 2, we formulate our problem, state the limitation of second-order statistics and describe the relevant work in the area of higher-order spectral analysis.

Chapter 3 establishes an analogy with array processing that results in an algorithm for detection and classification of multiple sources. After describing the estimation of the fourth-order spectrum of the observed signals in Chapter 4, we present simulation experiments in Chapter 5, and summarize our results and conclusions in Chapter 6.

Chapter 2

FORMULATION OF THE PROBLEM

In this report, we address the problem of detecting the number of sources that emit non-Gaussian signals, where we have access to only the superposition of the waveforms, and this observation may be further corrupted by additive Gaussian noise of unknown covariance. Mathematically, we have the measurements

$$x(t) = \sum_{k=1}^{P} x_k(t) + n(t) \tag{2.1}$$

where $n(t)$ represents the Gaussian noise with spectrum $S_n(w)$ and $\{x_k(t)\}_{k=1}^{P}$ are the waveforms from sources, which in turn can be modeled as

$$x_k(t) = h_k(t) \star u_k(t) \quad k = 1, 2, \ldots, P \tag{2.2}$$

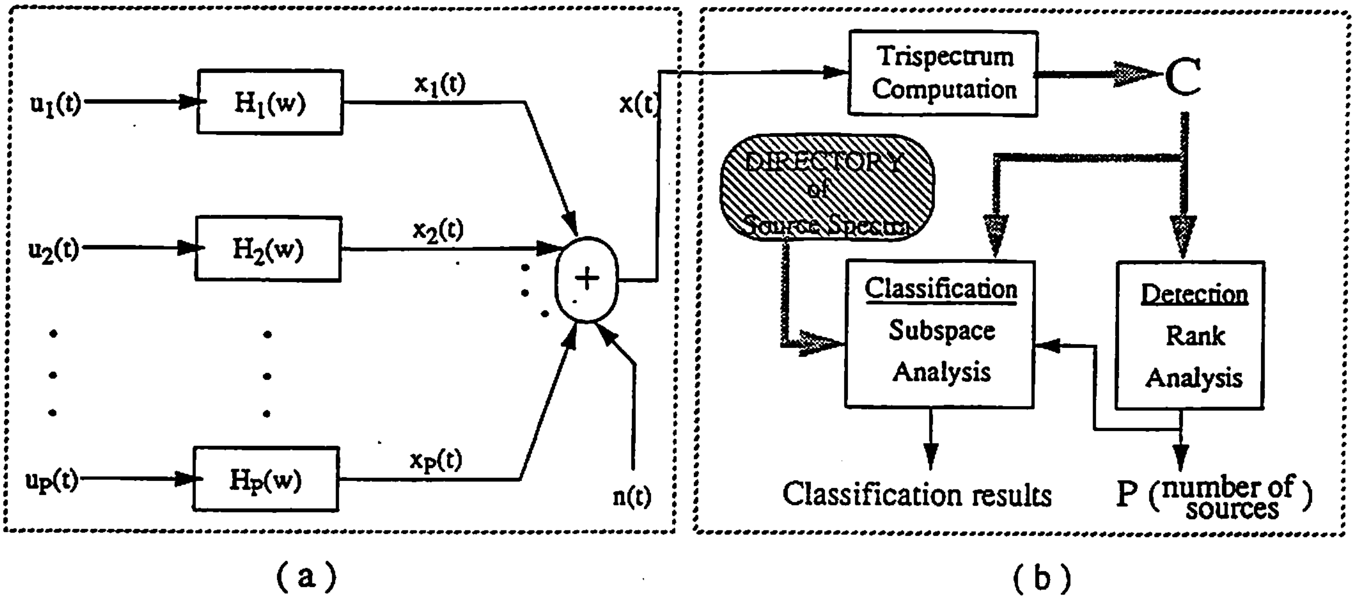


Figure 2.1: (a) Signal generation process, (b) Proposed system.

where $\{u_k(t)\}_{k=1}^{P}$ are real, stationary, white, non-Gaussian excitation sequences which are statistically independent among themselves, with variance σ_k^2 and fourth-order cumulant $\gamma_{4,k}$, and, the filter $h_k(t)$ models the waveform generation process of the kth source, with a frequency response $H_k(w)$. The signal model is illustrated in Figure 2.1.

The detection problem is to determine the number of sources, P, whereas, the classification problem is to sort the signals into specific categories based on some characteristics of the emitted waveforms. In an underwater military application, the detection problem can be to determine the number of submarines in a specific zone with only a single sensor. In this scenario, the classification problem will be to identify the submarines as friendly/hostile, or as conventional/nuclear. In the speaker verification problem of speech processing, it is desirable to identify the presence of the true speaker in the presence of noise, interference, or an imitating speaker.

For purposes of classification, there exist a need for *templates.* In this report, we consider the availability of spectrum shape information of sources, i.e., we have in our directory

$$S_k(w) = H_k(w)\, H_k^*(w) \tag{2.3}$$

for all sources that we want to classify. If $x(t)$ contains signals from sources with spectral shape that do not exist in our directory, the corresponding sources will be classified as "unknown".

Neither the detection nor the estimation problem can be solved by second-order statistics of the observations, since the spectrum of the observed signal $x(t)$, $S_x(w)$, can be expressed as

$$S_x(w) = \sum_{k=1}^{P} \sigma_k^2\, S_k(w) \; + S_n(w) \tag{2.4}$$

for all frequencies of interest.

To demonstrate the inadequacy of output spectral information more clearly, let us consider the following vector formulation of the problem: the received signal spectrum evaluated at discrete frequencies (FFT bins), and stacked in a vector $\mathbf{s}_x \triangleq [S_x(w_1), S_x(w_2), \ldots, S_x(w_M)]^T$, $w_k = 2\pi(k-1)/M$, can be represented as a superposition of *source spectrum* vectors

$$\mathbf{s}_x = \sum_{k=1}^{P} \sigma_k^2\, \mathbf{s}_k \; + \sigma_n^2\, \mathbf{s}_n \tag{2.5}$$

where $\mathbf{s}_k$ and $\mathbf{s}_n$ are defined analogously to $\mathbf{s}_x$. If all contributing sources (including noise) have their spectrum vectors in our directory, i.e., the spectrum vectors are known and linearly independent then it is possible to identify the sources, since (2.5) can be uniquely expressed as a linear combination of spectrum vectors. But with unknown contributions which may also have linearly independent spectrum vectors, it is not possible to express (2.5), as a linear combination of known spectrum vectors. Even when all the observed sources are registered in our directory, and the noise spectrum is of known shape (e.g., white) then the detection and classification task will require an exhaustive search procedure over our directory of spectral information, which will converge only after

$$\sum_{l=1}^{P+1} \binom{d}{l} \tag{2.6}$$

iterations where d denotes the total number of sources registered in the directory. The exhaustive search procedure stops when there is no improvement in approximating $\mathbf{s}_x$ with the spectra of $P+1$ sources. As a numerical example, in order to detect and classify 5 sources which are actually present in the data, and (which, therefore, implies the presence of P sources) which have already been registered in a 10 source directory, we need 847 iterations. Each iteration requires the construction of a projection matrix that spans the columns of selected source spectrum vectors; therefore, even in this simplified case of known noise spectrum and all registered sources, computations become very excessive. Clearly, the *observation dimensionality* provided by second-order statistics is *inadequate* to solve our problem.

Recently, higher-order statistics have been proposed to increase the *effective* dimensionality of an array [2]; however, the single sensor problem has received very little attention except for [4], which is an excellent paper in which the problem of separating the spectrum for the sum of two time-series is treated. That paper utilizes a particular submanifold of the trispectrum, $T(w_i, w_k, w_j)$, of the observed signal, for which $w_k = -w_i$. For the multiple sources case, this statistic can be expressed as[1]

$$T(w_i, -w_i, w_j) = \sum_{k=1}^{P} \gamma_{4,k} \cdot S_k(w_i)\, S_k(w_j) \quad 1 \le i, j \le M, \; w_i = 2\pi(i-1)/M. \tag{2.7}$$

Unfortunately, the approach in [4] does not handle measurement noise and is limited by the assumption that it requires one of the sources to have a null in its spectrum when the other source must have a finite value in its spectrum. The authors claim that their method can be extended to the case where there are more than two time-series, but this makes the assumption about the spectra of the sources even less reasonable.

In this report, we assume the presence of an unknown number of sources. The detection algorithm to be presented in the next chapter estimates the number of sources. For the classification problem, the information about the shape of the source spectra plays the role of a steering vector in an array processing scenario; this duality will be utilized to construct a subspace-based approach for the classification problem.

[1] For more background on trispectrum, refer to Chapter 4 and also [1].

Chapter 3

ANALOGY WITH ARRAY PROCESSING

In this chapter, we construct an analogy between our problem stated in the previous chapter and a narrowband array processing problem, in which the goal is to detect the number of far-field sources and estimate their directions-of-arrival (DOA).

Let us consider a narrowband array processing scenario, where there are M sensors and P far-field sources ($M > P$) with steering vectors $\mathbf{a}_k$. The measured Mx1 signal vector $\mathbf{y}(t)$ can be expressed as

$$\mathbf{y}(t) = \sum_{k=1}^{P} \mathbf{a}_k \, y_k(t) \; + \mathbf{n}(t) \tag{3.1}$$

where $\mathbf{n}(t)$ represents the effects of spatially white measurement noise with power σ^2, and the source waveforms, $y_k(t)$, are not fully correlated (coherent) among themselves [5]. Then the covariance matrix of measurements takes the form

$$\mathbf{R} = \mathbf{A}\,\mathbf{R}_s\,\mathbf{A}^H + \sigma^2\mathbf{I} \tag{3.2}$$

in which MxP matrix $\mathbf{A}$ is the steering matrix and $\mathbf{R}_s$ is the positive-definite covariance matrix of sources.

Now, let us return to our single channel problem. We can form an MxM trispectrum matrix $\mathbf{C}$, $(M > P)$, that is formed by the samples of the trispectrum of the received signal as

$$c_{ij} = T(w_i, -w_i, w_j) \quad 1 \leq i, j \leq M, \;\; w_i = 2\pi(i-1)/M. \tag{3.3}$$

Then following (2.7) and the definition of source spectrum vectors we obtain

$$\mathbf{C} = \sum_{k=1}^{P} \gamma_{4,k}\,\mathbf{s}_k\,\mathbf{s}_k^T \;=\; \mathbf{S}\,\Gamma\,\mathbf{S}^T \tag{3.4}$$

where $\mathbf{S}$, which we refer to as a *source spectrum matrix*, has columns which are the source spectrum vectors (hence, $\mathbf{S}^T = \mathbf{S}^H$), and the PxP diagonal matrix Γ consists of the fourth-order cumulants $\gamma_{4,k}$ of the sources (hence, $\gamma_{4,k}$ are real). The trispectrum matrix $\mathbf{C}$ has the following properties:

- It is real, since all of its terms are the product of real factors (2.7).
- It is symmetric, i.e., $\mathbf{C} = \mathbf{C}^T = \mathbf{C}^H$. This follows from (3.4).
- Hence with the properties just stated, $\mathbf{C}$ can be viewed as a pseudo-covariance matrix; but, it is *indefinite*, since fourth-order cumulants of driving sources are not necessarily positive.

The analogy between the array problem and the single sensor problem is summarized in Table 1.

Based on the analogy, we can utilize the detection and DOA algorithms already formulated for array processing for the single sensor detection and classification problem.

3.1 DETECTION OF THE NUMBER OF SOURCES

It is a common assumption in array processing to have the steering vectors (Mx1) corresponding to the sources illuminating the array be linearly independent. The equivalent assumption for the single sensor trispectral analysis scheme is that all the source spectrum vectors corresponding to the sources in the field must be linearly independent. Clearly, when this assumption is violated, solution becomes impossible even with the exhaustive search scheme described in Chapter 2; hence, in this report we assume that this is a valid assumption. Note that this is a very mild assumption, and is less restrictive than the one used in [4].

If the source spectrum vectors are linearly independent (there are P of them), then the rank of the matrix C must be P (since $M > P$). This follows from the above assumption and (3.4) and is motivated by the similar use of covariance matrix $\mathbf{R}$ in array processing [5]; hence, the number of sources can be detected by computing the rank of C.

Table 3.1: Analogy between two problems

Array Problem	Single Sensor Problem
Steering Vector: $\mathbf{a}_k$	Spectrum Vector: $\mathbf{s}_k$
Steering Matrix: A	Spectrum Matrix: S
Source Covariance Matrix: $\mathbf{R}_s$	Source Cumulant Matrix: $\mathbf{\Gamma}$
Noise Covariance Matrix: $\sigma^2\mathbf{I}$	Noise Cumulant Matrix: $\mathbf{0}$
Array Covariance Matrix: $\mathbf{R}$	Trispectrum Matrix: C

3.2 CLASSIFICATION OF SOURCES

Using the MUSIC algorithm, the DOA's of sources in array processing are determined by a search procedure [5]. A vector from the array manifold is selected and its distance from the so-called noise subspace is computed. If the vector is in the signal subspace then this implies that we have an arrival from a source with this particular steering vector.

The MUSIC algorithm can be used to classify the sources in the single sensor problem as follows:

(1) Compute the rank of $\mathbf{C}$ to reveal the number of sources P. This is the detection algorithm.

(2) Form the Mx$(M - P)$ matrix $\mathbf{E}_n$ containing the eigenvectors of $\mathbf{C}$, associated with its zero eigenvalues, as its columns.

(3) Pick a source spectrum vector $\mathbf{s} \in \{\mathbf{s}_j\}_{j=1}^{d}$ from our directory that contains spectral shape information about the kth source, and compute

$$f(k) = \frac{\mathbf{s}_k^T \mathbf{s}_k}{\mathbf{s}_k^T \mathbf{E}_n \mathbf{E}_n^T \mathbf{s}_k} \quad \text{for } k = 1, 2, \ldots, d. \tag{3.5}$$

The numerator is included to provide normalization. After $f(\cdot)$ is computed for all the sources in the directory, further normalization can be done to force this function to have a maximum of unity.

(4) The higher the value of $f(k)$, the higher the possibility of existence of the kth source in the field.

(5) If only K sources $(K < P)$ from our directory are classified to be in the field, then there must be $(P - K)$ sources of unknown spectra.

The proposed system is illustrated in Figure 2.1b.

Chapter 4

ESTIMATION OF FOURTH-ORDER SPECTRA

In any application, we do not have access to the *true* higher-order statistics, and therefore we estimate them. In this chapter, we propose two approaches for estimating the required trispectrum samples in (3.3).

4.1 FIRST APPROACH: TIME-DOMAIN AVERAGING

Trispectrum is the 3-dimensional Fourier transform of the fourth-order cumulant function of a stationary random process, $x(t)$, i.e.,

$$C_{4,x}(\tau_1,\tau_2,\tau_3) = E\{x(t)x(t+\tau_1)x(t+\tau_2)x(t+\tau_3)\} - E\{x(t)x(t+\tau_1)\}\,E\{x(t+\tau_2)x(t+\tau_3)\}$$

$$- E\{x(t)x(t+\tau_2)\}\,E\{x(t+\tau_1)x(t+\tau_3)\} - E\{x(t)x(t+\tau_3)\}\,E\{x(t+\tau_1)x(t+\tau_2)\} \quad (4.1)$$

and

$$T_x(w_1, w_2, w_3) = \sum_{\tau_1,\tau_2,\tau_3=-\infty}^{\infty} C_{4,x}(\tau_1, \tau_2, \tau_3) \exp\left(-j(w_1\tau_1 + w_2\tau_2 + w_3\tau_3)\right) \quad (4.2)$$

If x(t) is the observed output of a linear system (with impulse response h(t)) which is driven by independent, identically distributed (i.i.d.) non-Gaussian random variables $u(t)$, and is corrupted by additive Gaussian noise $n(t)$, i.e.,

$$x(t) = h(t) \star u(t) + n(t)\,, \quad (4.3)$$

then the fourth-order cumulants can be expressed as [3]

$$C_{4,x}(\tau_1, \tau_2, \tau_3) = \gamma_{4,u} \sum_{t=-\infty}^{\infty} h(t)h(t+\tau_1)h(t+\tau_2)h(t+\tau_3) \quad (4.4)$$

in which $\gamma_{4,u} = C_{4,u}(0,0,0) \neq 0$. The trispectrum can then be evaluated as[1]

$$T_x(w_1, w_2, w_3) = \gamma_{4,u} H(w_1)H(w_2)H(w_3)H(-(w_1 + w_2 + w_3)) \quad (4.5)$$

In the case of multiple independent inputs as in Figure 2.1a, superposition holds. The argument of the summation in (4.4) has in fact a finite support ($(\tau_1, \tau_2, \tau_3) \in \Omega$) for linear systems with finite-impulse response, therefore, it is sufficient to evaluate the Fourier transform over Ω, as follows:

[1]In this report, we are concerned with real signals and impulse responses. For real signals, the Fourier transform has the symmetry property, $H(-w) = H^*(w)$. With real signals, the source spectrum vectors have a redundant part: for even M, the first L, ($L = M/2 + 1$), samples of a spectrum vector are sufficient. Consequently, the results of Chapter 3 remain the same, by using the LxL principal submatrix of C, and replacing M by L.

(1) After observing $\{x(t)\}_{t=0}^{T-1}$, calculate the sample cumulants on a predetermined finite domain of support Ω, by time-domain averaging, i.e.,

$$\widehat{C}_{4,x}(\tau_1,\tau_2,\tau_3) = 1/T \sum_{t=0}^{T-1} x(t)x(t+\tau_1)x(t+\tau_2)x(t+\tau_3)$$

$$-1/T^2 \sum_{t_1,t_2=0}^{T-1} x(t_1)x(t_1+\tau_1)x(t_2)x(t_2+\tau_3-\tau_2) - 1/T^2 \sum_{t_1,t_2=0}^{T-1} x(t_1)x(t_1+\tau_2)x(t_2)x(t_2+\tau_3-\tau_1)$$

$$- 1/T^2 \sum_{t_1,t_2=0}^{T-1} x(t_1)x(t_1+\tau_3)x(t_2)x(t_2+\tau_2-\tau_1) \qquad (4.6)$$

in which the received signal is assumed to be zero outside the observation window.

- Compute the trispectrum samples $T_x(w_1,-w_1,w_2)$ as required in (3.3) by performing the following 2D-Fourier transform over Ω:

$$\widehat{T}_x(w_1,-w_1,w_2) = \sum_{(\tau_1,\tau_2,\tau_3)\in\Omega} \widehat{C}_{4,x}(\tau_1,\tau_2,\tau_3)\exp\left(-j(w_1(\tau_1-\tau_2)+w_2\tau_3\right) \qquad (4.7)$$

4.2 SECOND APPROACH: FREQUENCY-DOMAIN AVERAGING

The trispectrum samples can be estimated by first computing the Fourier transform of the observed signal $x(t)$, and then computing the higher-order statistics of the Fourier transform coefficients. This approach requires some background on the estimates of higher-order spectra [1].

Let $X_M(w)$ denote the M-point Fourier transform of the observed signal, $x(t)$, i.e.,

$$X_M(w) = \sum_{t=0}^{M-1} x(t)\exp(-jwt) \tag{4.8}$$

and let $\Delta_M(w)$ denote the Fourier transform of the constant function $x(t) = 1$

$$\Delta_M(w) = \sum_{t=0}^{M-1} 1\cdot\exp(-jwt) = \exp(jw(M-1)/2)\ \sin(wM/2)/\sin(w/2) \tag{4.9}$$

so that $\Delta_M(w) = M$ if $w \equiv 0 \,(\mathrm{mod}\ 2\pi)$, and $\Delta_M(w) = 0$ at $w = 2\pi n/M$, $n \in \mathcal{Z}$ but not a multiple of M.

Assumption ([1], pg. 160): Given the strictly stationary process $x(t)$, we assume

$$\sum_{\tau_1,\tau_2,\tau_3=-\infty}^{\infty} |\tau_l \cdot \mathrm{cum}_x(\tau_1,\tau_2,\tau_3)| < \infty \tag{4.10}$$

for $l = 1,2,3$.

If the domain of support (Ω) of the fourth-order cumulant is finite (as a result of finite-impulse response filtering of an i.i.d. input) the above assumption is always true. Based on this assumption on the cumulants of $x(t)$ (or, correspondingly, the generating impulse response) we have the following:

Theorem ([1], pg. 161): If (4.10) is valid, then

$$\mathrm{cum}(X(w_a), X(w_b), X(w_c), X(w_d)) = (2\pi)^3\Delta_M(w_a+w_b+w_c+w_d)\ T_x(w_a,w_b,w_c) + O(1) \tag{4.11}$$

where the error term $O(1)$ is uniform for (w_a, w_b, w_c, w_d).

As a consequence, when the window length M increases, we divide both sides by M and obtain the following important result:

$$\lim_{M\to\infty}(2\pi)^{-3}M^{-1}\text{cum}(X(w_a),X(w_b),X(w_c),X(w_d)) = \delta(w_a+w_b+w_c+w_d)\,T_x(w_a,w_b,w_c) \quad (4.12)$$

where $\delta(w)$ is the Kronecker comb function

$$\delta(w) = \begin{cases} 1, & \text{if } w = 2\pi l,\ l \in Z \\ 0, & \text{otherwise} \end{cases} \quad (4.13)$$

which can also be viewed as $\delta(w) = \lim_{M\to\infty}\Delta_M(w)/M$. This observation shows that as the Fourier analysis window length M increases, the cumulants of frequency samples tend to the trispectrum samples, i.e., trispectrum samples estimated in this way are asymptotically unbiased. To reduce the variance of the estimates, averaging over multiple windows are required.

To estimate the trispectrum, we first need to compute the cumulant term in the left side of (4.12). To accomplish this, we use the definition of fourth-order cumulants for zero-mean random variables [3] , that:

$$\text{cum}(X(w_a),X(w_b),X(w_c),X(-(w_a+w_b+w_c)) = E\{X(w_a)X(w_b)X(w_c)X(-(w_a+w_b+w_c))\}$$

$$-E\{X(w_a)X(w_b)\}\cdot E\{X(w_c)X(-(w_a+w_b+w_c))\}-E\{X(w_a)X(w_c)\}\cdot E\{X(w_b)X(-(w_a+w_b+w_c))\}$$

$$-\ E\{X(w_b)X(w_c)\}\cdot E\{X(w_a)X(-(w_a+w_b+w_c))\} \quad (4.14)$$

where [1]:

$$E\{X(w_a)X(w_b)\} = (2\pi)\,\Delta_M(w_a + w_b)S_x(w_a) + O(1) \tag{4.15}$$

in which $S_x(w)$ denotes the power spectrum of $x(t)$. Dividing both sides of (4.15) by M and taking the limit yields

$$\lim_{M\to\infty} E\{X(w_a)X(w_b)\} = (2\pi)\,\delta(w_a + w_b)S_x(w_a)\ ; \tag{4.16}$$

hence, power spectrum can be estimated through (4.16).

Substituting (4.14) and (4.16) into (4.12), we see that

$$\begin{aligned} T_x(w_a, w_b, w_c) = \lim_{M\to\infty}[\ & (2\pi)^{-3}M^{-1} \cdot E\{X(w_a)X(w_b)X(w_c)X(-(w_a + w_b + w_c))\} \\ & - (2\pi)^{-1}M\,(S_x(w_a)\cdot S_x(w_c)\cdot \delta(w_a + w_b) + S_x(w_a)\cdot S_x(w_b)\cdot \delta(w_a + w_c) \\ & + S_x(w_b)\cdot S_x(w_a)\cdot \delta(w_b + w_c))\ ] \end{aligned} \tag{4.17}$$

To obtain (3.3) from (4.17), we set $w_a = w_1$, $w_b = -w_1$, and $w_c = w_2$, in the latter, i.e.,

$$\begin{aligned} T_x(w_1, -w_1, w_2) = \lim_{M\to\infty}[\ & (2\pi)^{-3}M^{-1} \cdot E\{X(w_1)X(-w_1)X(w_2)X(-w_2)\} \\ & - (2\pi)^{-1}M\,(S_x(w_1)\cdot S_x(w_2) + S_x(w_1)\cdot S_x(-w_1)\cdot \delta(w_1 + w_2) \\ & + S_x(w_1)\cdot S_x(-w_1)\cdot \delta(w_1 - w_2))\ ] \end{aligned} \tag{4.18}$$

In summary, to compute $T_x(w_1, -w_1, w_2)$:

(1) After observing $\{x(t)\}_{t=0}^{T-1}$, separate the data into K windows of M samples $T = K \cdot M$.

(2) For each window, take the Fourier transform of the data, and use (4.15) and (4.18) to compute the required trispectrum samples for each window. In (4.18), $E\{X(w_1)X(-w_1)X(w_2)X(-w_2)\}$ is obtained by sample averaging over the K windows. Although infinite amount of data is not available, these equations yield accurate estimates with sufficiently large data lengths.

(3) Finally, average the trispectrum estimates from each window to get the final estimate.

In this report , we use the second approach in the simulation experiments. The primary reason for this selection is the latter approach does not require an a priori estimate of the region of support for the fourth-order cumulant function.

Chapter 5

SIMULATIONS

We tested our method with a directory of size four. The available spectral information is illustrated in Figs. 5.1 a-d. We assume the first source in our directory is present in the field. In addition, there exists an unknown source, with spectrum given in Figure 5.1 e. Ambient noise is Gaussian with spectrum as in Figure 5.1 f. Two targets are represented by non-Gaussian white processes driving the appropriate filters that give rise to the corresponding spectra. The SNR at the excitation level is 20 dB for both sources (refer to Figure 2.1 a).

Figure 5.1g illustrates the spectrum of the received signal estimated from $10,000$ segments of length 16. Clearly, this information is inadequate for deciding how many sources are present in the field; however, eigenanalysis of the trispectrum matrix, reveals two major eigenvectors; hence, signal subspace is two-dimensional, indicating the presence of two targets. Next, we plot the $f(k)$ $(k = 1, 2, 3, 4)$ values which indicate the possibility of the presence of the sources in the directory (see Fig. 5.1h). Clearly, only one source is present in the field, resulting in the conclusion that there must be a source of unknown spectrum.

We also performed 30 Monte-Carlo runs for this experiment, to illustrate the variations of

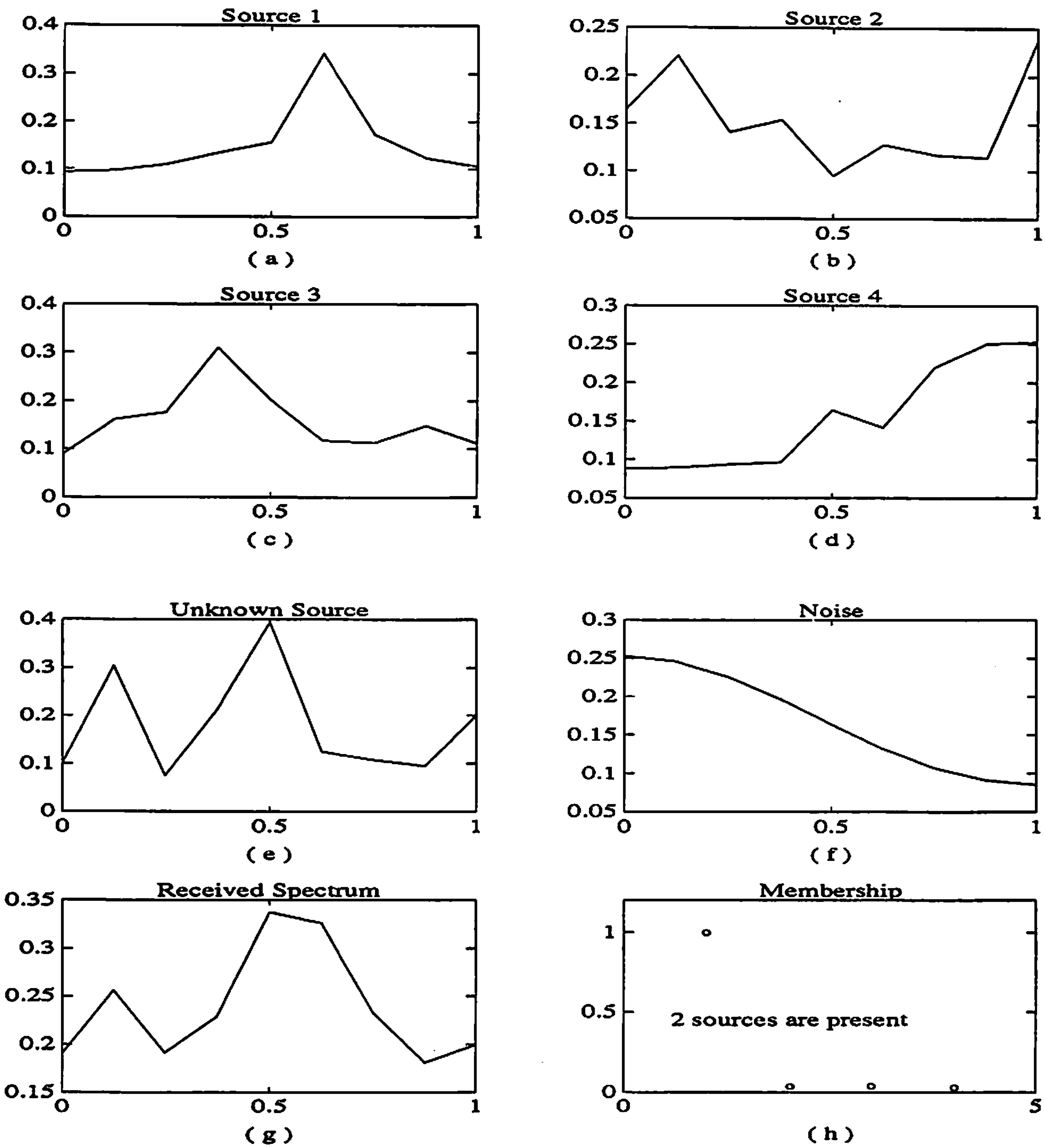
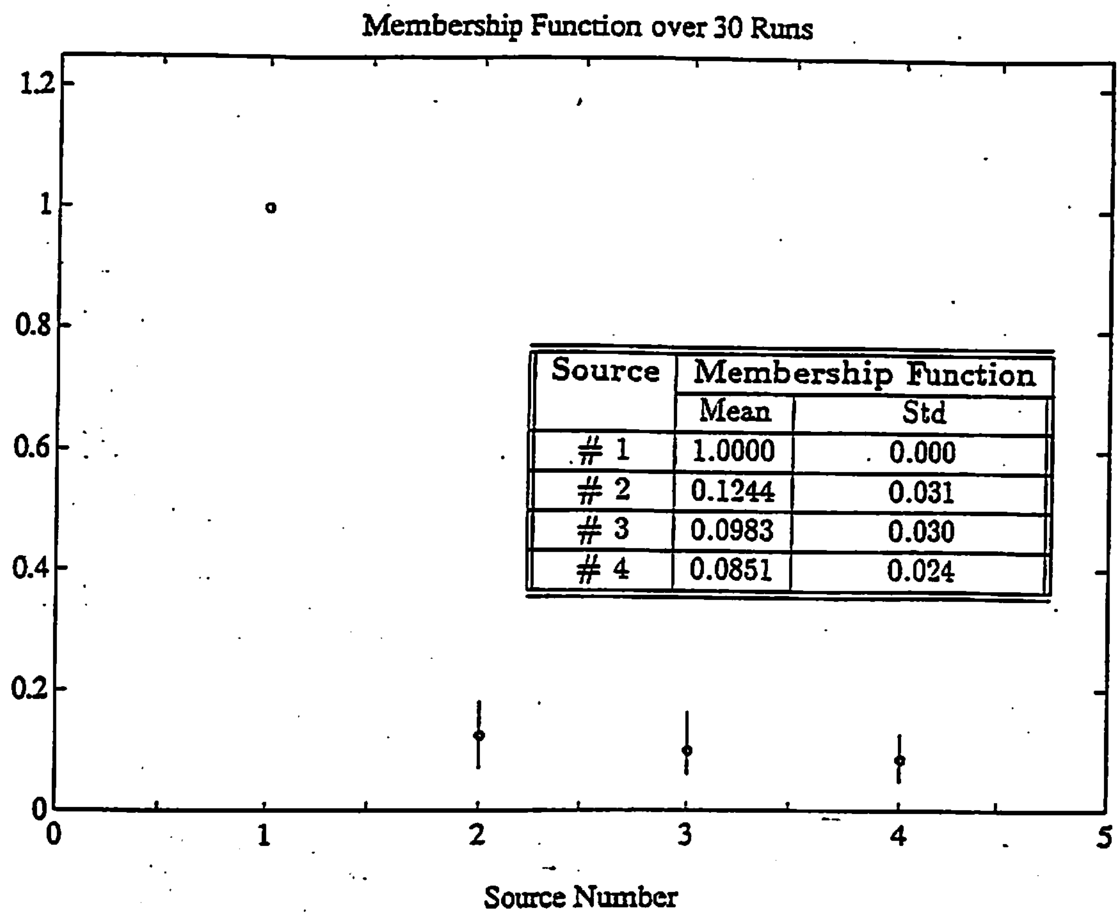


Figure 5.1: (a)-(d) Spectra of sources in our directory, (e) Spectrum of unknown target, (f) Spectrum of ambient noise, (g) Spectrum of the received signal, (h) Membership values for the candidate sources.



Source	Membership Function	
	Mean	Std
# 1	1.0000	0.000
# 2	0.1244	0.031
# 3	0.0983	0.030
# 4	0.0851	0.024

Figure 5.2: Results from 30 Monte-Carlo runs: variation of membership functions.

the membership functions depicted in Figure 5.1h. In each run, the membership functions are normalized by their maximum value over the sources. The first source always has the maximum value since it is actually the only one in the field. The results shown in Figure 5.2 indicate that the proposed algorithm correctly classifies the source in the directory.

Chapter 6

CONCLUSIONS

In this report, we established a framework for the analysis of single channel, multicomponent data with higher-order statistics. We showed that by using only second-order statistics, it is impossible to detect and classify multiple sources from single channel data; however, by using higher than second-order statistics it is possible to form an analogy with the direction-of-arrival estimation problem in array processing. This analogy enables us to further utilize existing ideas from array processing; blind estimation of source spectra and employing information theoretic criteria for detection, both of which will be described in future publications.

Bibliography

[1] D.R. Brillinger and M. Rosenblatt, "Asymptotic theory of the estimates of kth order spectra," in *Spectral Analysis of Time Series*, B. Harris, Ed. New-York: Wiley, pp.189-222, 1967.

[2] J.F. Cardoso, "Higher-order narrowband array processing," *International Signal Processing Workshop on Higher Order Statistics*, pp.121–130,Chamrousse-France, July 10-12, 1991.

[3] J.M. Mendel, "Tutorial on higher-order statistics (spectra) in signal processing and system theory: theoretical results and some applications," *in Proc. IEEE*, vol.79, no.3, pp.278–305, March 1991.

[4] F. Sakaguchi and H. Sakai, "A spectrum separation method for the sum of two non-Gaussian stationary time series using higher-order periodograms," *IEEE Journal of Oceanic Engineering*, vol.OE-12, no.1, pp.80–89, January 1987.

[5] R.O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas and Propagation*, vol.AP-34, no.3, pp.276–280, March 1986.

[6] M. Wax and I. Ziskind, "Detection of the number of signals by the MDL principle," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.ASSP-37, no.8, pp.1190–1196, August 1989.