# USC–SIPI REPORT #245

**Cumulants and Array Processing**

**by**

**Mithat C. Dogan**

**September 1993**

Signal and Image Processing Institute
**UNIVERSITY OF SOUTHERN CALIFORNIA**
Department of Electrical Engineering-Systems
3740 McClintock Avenue, Room 400
Los Angeles, CA 90089-2564 U.S.A.

# Dedication

This dissertation is dedicated to my parents,

In lieu of time
we might have spent together
and all of your efforts
towards my well-being

.... Nothing Compares to You ....

{♡, ♡}

# Acknowledgements

I would like to thank my advisor, Professor Jerry M. Mendel; my gratitude goes to him for his guidance, for sharing his thoughts and his time with me, and for convincing me how organization skills are as important as doing research. The most important distinction about his attitude towards Ph.D. research is to let students learn how to walk by themselves, and thanks to him, I think I am running freely... This work would not have been possible without his efforts, devotion, extensive planning and support. I hope to learn more from Dr. Mendel's experiences and wisdom after my graduation.

I am grateful to be a recipient of NATO Science Fellowship awarded by the Scientific and Technical Research Council of Turkey for graduate studies abroad. I am also thankful to my professors at the Middle East Technical University (METU) for providing an excellent undergraduate education.

I would like to thank my qualifying committee members, Dr. Solomon Golomb, Dr. Jay Kuo, Dr. Richard Leahy and Dr. Max Nikias. I would like to thank Dr. Golomb and Dr. Nikias again, for being on my dissertation committee. I would like to acknowledge interesting conversations with Dr. G.B. Giannakis of University of Virginia and Dr. A. Swami of $CS^3$.

Another important point that I appreciate is the excellent computer facilities of the Signal and Image Processing Institute. We owe Dr. Allan Weber a lot for solving our computer related problems in a timely manner.

Our secretarial staff helped me a lot in the paperwork necessary for payroll forms, conference expenditures, and even with some of the artwork necessary for our publications. I thank Ms. Mercedes Morente, Ms. Linda Varilla, and Ms. Mayitta Penoliar for all their efforts.

Living near USC was a really painful experience. If I did not have very good officemates, access to a very good bookstore and workout facility, it could have

been much worse. I thank Peter "tall-cool one" Chu, Hyun-Mun Kim, George "the great" Mouzouris, Chiu-Yeung "Mr. No" Ngo, and Wen Wei for their friendship. We shared "good times, bad times" and even last drops of coffee. As Dr. Mendel once said, our office looks more like the United Nations. I learned a lot from this diverse environment. I also enjoyed being with Brian Beard, Giri Mandyam, Julie "Fuzzy" Dickerson, Egemen Gönen, Mehmet Gürelli, Giri Mandyam, Greg Mitchell, Sabino Piazzolla and Saeed Rajput. Thanks a lot guys... especially for your tolerance to the kind of music that I constantly played. In this respect, I thank 95.5 KLOS radio for perfectly addressing my taste (an undefined term) in music.

Finally, I would like to thank my parents for all their love and support. I will always remember the night I started writing this dissertation; I felt my clock start ticking, indicating that I can see them very soon...

# CONTENTS

# LIST OF TABLES

# LIST OF FIGURES

# Abstract

Signal processing using an array of antennas is an attractive solution to problems of source detection and parameter estimation. Array processing algorithms are originally designed to extract important phase information about the propagating source wavefronts by cross-correlating sensor meausurements. If source processes are non-Gaussian, then higher-order statistics (cumulants) of received signals provide additional information about the parameters of interest and are insensitive to additive Gaussian observation noise. Our motivation of using cumulants in array processing applications is to recover *more* phase information than is possible by using only second-order statistics. We start the dissertation by a compact introduction to array processing models, spatial spectrum estimation techniques and properties of cumulants. Then, we show how fourth-order cumulants increase the phase information that can be extracted from the sensor measurements. We introduce the concept of *virtual* sensors, and explain how cumulants can be used to compute cross-correlations among actual and virtual sensors to increase the effective aperture of the array. Using the interpretation of cumulants, we address the joint array calibration and direction-finding problem using an arbitrary antenna array. We prove that using a single doublet, it is possible to estimate the directions, steering vectors, and the waveforms of the non-Gaussian sources using cumulants. We determine bounds on effective aperture extension (without knowing sensor locations) by using cumulants. The upper bound on existing cumulant-based methods can be exceeded when we use minimum-redundancy array design concepts together with cumulants. We propose designs for both linear and two-dimensional arrays. Cumulants have long been promoted in signal processing applications for their ability to suppress additive Gaussian observation noise. We show that it may be possible to suppress additive *non*-Gaussian noise if we place a sensor far-enough from the main array,

whose noise component can be non-Gaussian but independent from the noise components in the main array sensors. Finally, we address the problem of single sensor detection and classification of multiple linear non-Gaussian processes. This problem is solved by exploiting the fact that polyspectra possess an *array* of arguments, unlike spectrum. The dissertation concludes with future research directions and an extensive bibliography.

# Chapter 1

# Array Processing Fundamentals

Signal processing using an array of sensors is an attractive solution to problems of source detection and parameter estimation since an array offers ways of overcoming the directivity limitations of a single element [169]. The measured signals contain information about several parameters associated with the sources. For example, consider a single plane wave illuminating an array of two identical sensors. If the distance between the two sensors is known, then the time delay between the sensor outputs can be used to determine the direction-of-arrival (DOA) of the emitter.

The main problems of interest in array processing are: detection of far-field sources, estimation of their parameters, such as angles of arrival (azimuth and elevation), power levels, correlation structure of sources, and polarization properties, and, estimation of the source waveforms. All these problems fall into two (overlapping) categories: detection or estimation of signals and their parameters from multichannel information provided by an array of antennas. In some applications, perfect knowledge of the antenna responses over their field of view is required, whereas in other applications only partial information or some constraints on array geometry are sufficient.

In this chapter, we first describe passive antenna arrays, and the underlying assumptions on signal model and discuss the limitations of the assumptions. Then we describe complex representation of narrowband signals, since our major interest is on narrowband direction-finding. Finally, we present the generic data model associated with an array of passive sensors, its assumptions and limitations. We refer the reader to [58, Chapter1] for further reading.

## 1.1 Passive Sensor Arrays

In active sensing environments, a known waveform of finite duration is generated to illuminate a particular sector of the field of view. Processing is performed on the returned signals which are usually modified both in amplitude and in phase depending on the characteristics of a possible target and transmission medium. A radar is a perfect example of active remote sensing [117]. A passive array only receives signals as opposed to an active array which is equipped with a transmitter. Passive arrays are quite common in underwater acoustics [4, 17], and communication applications [2, 3, 8, 26, 29, 80, 156, 168, 177].

Consider the array configuration in Figure 1.1. In this structure, $M$ antennas are spaced at equal distance on a line. This array is commonly used in practical applications and it is called a uniform linear array (ULA). In the figure, there are $P$ sources (emitters) and the angles $\{\theta_1, \theta_2, \ldots, \theta_P\}$ (measured counter-clockwise) indicate the directions-of-arrival of the signals from the sources. The measured output of each sensor is assumed to be a *superposition* of the individual contributions from each emitter. We also assume that the emitters are located at distances that are much greater than the aperture of the array, so that their wavefronts can be approximated by *plane waves.* This assumption is called the *far-field* assumption. Another restriction is that the far-field sources must keep their locations during data collection. Under these assumptions, the wavefronts are approximately planar and the sensor outputs are delayed and scaled replicas of the emitter signals. Let $\{\bar{s}_1(t), \bar{s}_2(t), \ldots, \bar{s}_P(t)\}$ denote the source waveforms[1]; then, the measurement recorded by the $k$th sensor, $\bar{r}_k(t)$, is modeled as

$$\bar{r}_k(t) = \sum_{l=1}^{P} g_k(\theta_l)\bar{s}_l(t - \tau_k(\theta_l)) + \bar{n}_k(t) \tag{1.1}$$

where the response (also known as gain, or sensitivity) of the $k$th sensor to a wavefront from $\theta$ is represented by $g_k(\theta)$, and $\tau_k(\theta_l)$ represents the time-delay of the $l$th emitter signal at the $k$th sensor. The additive noise term, $\bar{n}_k(t)$ stands for everything that can not be explained by the modeling assumptions; e.g., measurement

[1] We use an overbar for real signals in order to avoid confusion with their complex representation which will be described in the next section.

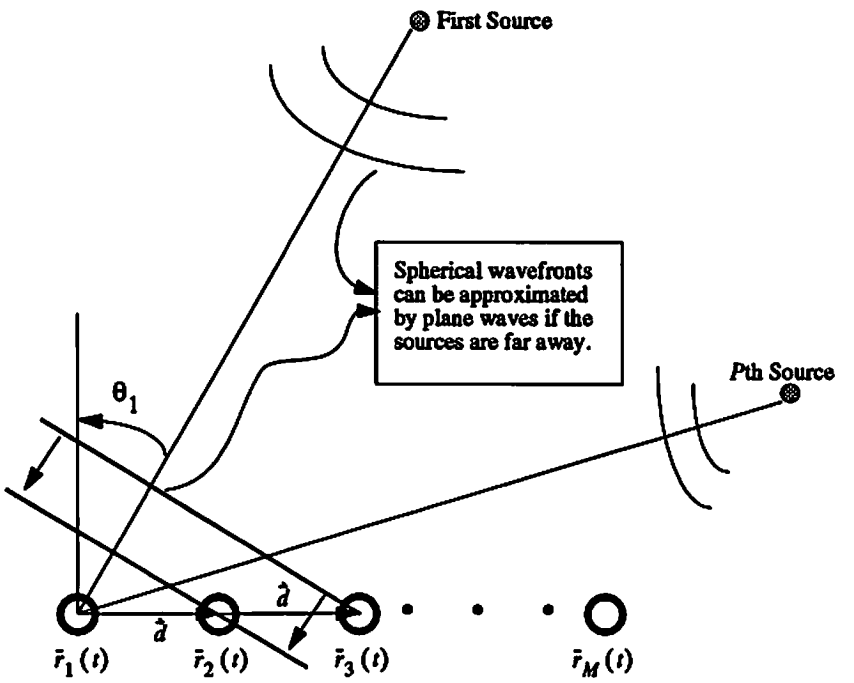



Figure 1.1: Uniformly spaced linear array of isotropic sensors illuminated by far-field sources. In this illustration, the antenna that measures $\bar{r}_1(t)$ is selected to be the reference sensor.

errors, ambient noise, deviations from the plane wave assumption, mutual coupling among sensors, etc. The response function $g(\theta)$ depends on the characteristics of the sensor; for example, an isotropic sensor has equal response in all directions, i.e., $g(\theta)$ is a constant. The time-delay function, $\tau(\theta)$, is a function of arrival angle $(\theta)$, array geometry and propagation speed of the wave (denoted as $c$) in the medium. For example, if two sensors are separated by a vector $\vec{d}$ in space, then the time-delay between the two locations can be quantified as

$$\tau(\theta) = \vec{\xi}(\theta) \cdot \vec{d} \tag{1.2}$$

where $\vec{\xi}(\theta)$ is defined as the *slowness vector*, with each of its components equal to the reciprocal of the speed in that dimension, i.e., $\|\vec{\xi}(\theta)\| = 1/c$.

The sensor outputs are collected in the $M$-vector $\bar{\mathbf{r}}(t)$, to form the array output, i.e.,

$$\bar{\mathbf{r}}(t) = \begin{bmatrix} \bar{r}_1(t) \\ \bar{r}_2(t) \\ \vdots \\ \bar{r}_M(t) \end{bmatrix} = \begin{bmatrix} \sum_{l=1}^{P} g_1(\theta_l)\bar{s}_l(t-\tau_1(\theta_l)) \\ \sum_{l=1}^{P} g_2(\theta_l)\bar{s}_l(t-\tau_2(\theta_l)) \\ \vdots \\ \sum_{l=1}^{P} g_M(\theta_l)\bar{s}_l(t-\tau_M(\theta_l)) \end{bmatrix} + \begin{bmatrix} \bar{n}_1(t) \\ \bar{n}_2(t) \\ \vdots \\ \bar{n}_M(t) \end{bmatrix} \tag{1.3}$$

A special case of the uniform linear array occurs when the sensor elements are assumed to have identical response functions (not necessarily isotropic). If there is only one emitter, and all the sensor gains are unity, the noise free measurements can be represented as

$$\begin{aligned} \bar{r}_1(t) &= \bar{s}_1(t) \\ \bar{r}_2(t) &= \bar{s}_1(t - \Delta \sin\theta/c) \\ &\vdots \\ \bar{r}_M(t) &= \bar{s}_1(t - (M-1)\Delta \sin\theta/c) \end{aligned} \tag{1.4}$$

which implies in the case of a uniform linear array with identical sensors, the response functions can be taken as $g_k(\theta) = 1$, and the time-delay functions are defined as $\tau_k(\theta) = (k-1)\Delta \sin\theta/c$. In (1.4), the reference point is selected as the location of the first sensor, and the time-delays are measured by considering the signal part of $r_1(t)$ as a reference. Clearly, there is a redundancy in (1.4): the signal waveform could be defined as an arbitrary (non-zero) constant times $\bar{s}_1(t)$, by letting the array response function at the corresponding direction be the reciprocal of the same constant.

Array observations (snapshots) are obtained by simultaneously measuring the sensor outputs at $N$ different time points. Although uniform sampling is not required, simultaneous sampling is necessary. Using the array snapshots, we can estimate the source parameters of interest. In communications, the source waveforms are of interest while the direction of the desired source is assumed to be known. By spatially filtering the snapshots (delaying and summing) it is possible to suppress interference and recover the message (source waveforms) of interest. In underwater applications, one is mainly interested in the locations of emitters. The source waveforms are of interest only if there is a need to classify the sources. In an underwater

communications problem, one is interested in both the message and the direction of sources. In order to estimate the parameters of sources, it is necessary to determine their number (detection) first.

## 1.2 Representation of Narrowband Signals

In array processing, a problem is classified as *narrowband* if the signal bandwidth is small compared to the inverse of the time required for the signal wavefront to propagate across the array. This implies that the complex envelopes of the signals (to be defined later in this section) do not change significantly over the time interval it takes the wavefront to propagate through the array aperture. For a discussion on bandwidth, we refer the reader to [126].

Consider a continuous-time signal $\bar{s}(t)$, whose energy is concentrated around a center frequency $w_c$. This type of signal is quite common in communications applications (except spread spectrum), where a low-frequency (slowly varying) message is modulated by a high-frequency carrier before transmission. The signal can be modeled as a sinusoid at $\omega_c$ with slowly time-varying amplitude and phase

$$\bar{s}(t) = A(t)\cos(\omega_c t + \phi(t)) \tag{1.5}$$

This expression can be rewritten as

$$\bar{s}(t) = \mathrm{Re}\{s(t)\} = \frac{1}{2}\left(s(t) + s^*(t)\right) \tag{1.6}$$

where

$$s(t) = \underbrace{A(t)\exp(j\phi(t))}_{\text{message}}\ \underbrace{\exp(j\omega_c t)}_{\text{carrier}} \tag{1.7}$$

The complex signal $s(t)$ is called the *analytic signal* representation of $\bar{s}(t)$, and the parts $A(t)\exp(j\phi(t))$ and $\exp(j\omega_c t)$ are the *complex envelope* and the *carrier* respectively. Similarly, the analytic signal can be obtained from the real signal as follows [57]:

$$s(t) = \bar{s}(t) + j\bar{s}_h(t) \tag{1.8}$$

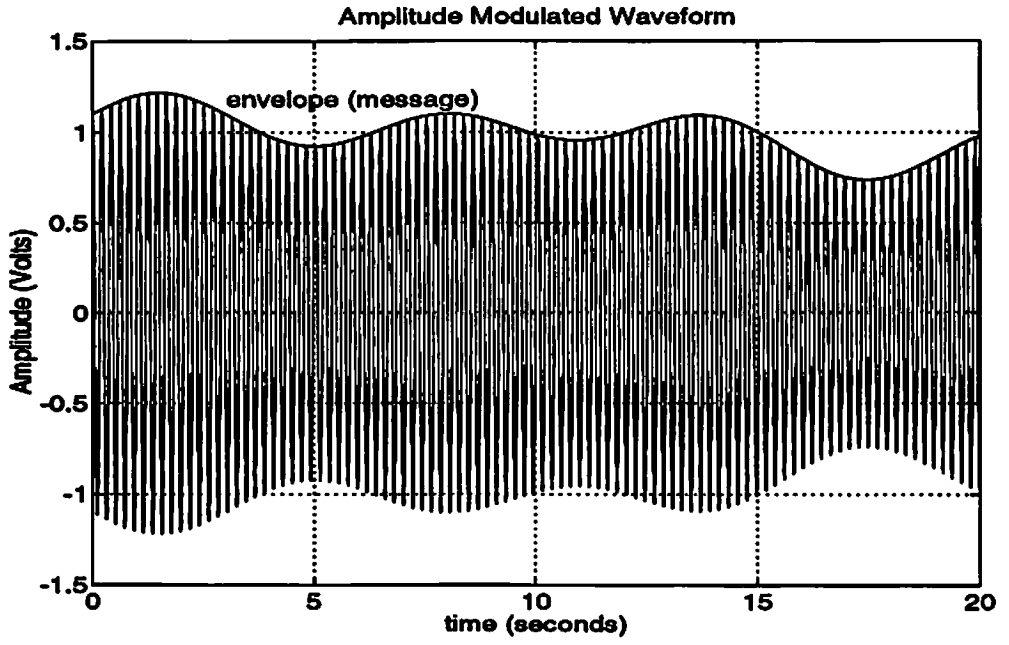



Figure 1.2: A narrowband amplitude modulated waveform.

where $\bar{s}_h(t)$ is the Hilbert transform of $\bar{s}(t)$. Since the highest frequency in the message is much less than the carrier frequency (i.e., the narrowband assumption), the Hilbert transform of the signal $\bar{s}(t)$ is [57]

$$\bar{s}_h(t) = A(t)\sin(\omega_c t + \phi(t)) \tag{1.9}$$

It is possible to approximately obtain the analytic signal (1.8), without using a Hilbert transformer as we shall show later.

Figure 1.2 illustrates an amplitude-modulated waveform ($A(t)$ is the envelope, as indicated by the solid line) which can be classified as a narrowband signal. Note that during a fraction of a second, the envelope (message) stays relatively unchanged, whereas the phase of the carrier changes considerably.

Suppose that $s(t)$ is subject to a small delay $\tau$, during which the message remains approximately unchanged ($A(t-\tau) \approx A(t)$ and $\phi(t-\tau) \approx \phi(t)$). Then, the effect of time delay on the analytic signal can be approximated as a phase shift since

$$s(t-\tau) \approx A(t)\exp(j\phi(t))\exp(j\omega_c(t-\tau)) = s(t)\exp(-j\omega_c\tau) \tag{1.10}$$

Equivalently, we have

$$\bar{s}(t-\tau) = \frac{1}{2}\left( s(t)\exp(-j\omega_c\tau) + s^*(t)\exp(j\omega_c\tau) \right) \tag{1.11}$$

Since it is necessary to include twice as many complex signals to represent the real waveforms, it is easier to deal with complex signals in practice. Using the fact that $\bar{s}(t)$ is narrowband, we have the approximations

$$\bar{s}(t-\frac{\pi}{2\omega_c}) = A(t-\frac{\pi}{2\omega_c})\cos(w_c t - \frac{\pi}{2} + \phi(t-\frac{\pi}{2\omega_c}) \approx A(t)\sin(\omega_c t + \phi(t)) \tag{1.12}$$

or, equivalently

$$\bar{s}(t-\frac{\pi}{2\omega_c}) \approx \bar{s}_h(t) \tag{1.13}$$

which follows from (1.9). Hence, the analytic signal can be obtained from the recevied signals, without using a Hilbert transformer:

$$s(t) \approx \bar{s}(t) + j\bar{s}(t-\frac{\pi}{2\omega_c}) \tag{1.14}$$

Due to the implementation in (1.14), $\mathrm{Re}\{s(t)\}$ is called the *in-phase* component of $s(t)$ and $\mathrm{Im}\{s(t)\}$ is called the *quadrature* component. The process of constructing $s(t)$ through (1.14) is known as *quadrature sampling.* In this thesis, we always assume the analytic signal representation in array processing problems.

The above derivation does not hold for broadband signals since the rate of change of the complex envelope of $\bar{s}(t)$ is comparable to the carrier frequency, making the approximation in (1.10) invalid. For example, if $\bar{s}_1(t)$ is a broadband signal, and if $\bar{s}_2(t) = \bar{s}_1(t-\tau)$, then the Fourier transforms are related as follows:

$$F_2(\omega) = F_1(\omega)\,\exp(-jw\tau) \tag{1.15}$$

where $F_2(\omega)$ is the Fourier Transform of $\bar{s}_2(t)$. We note that, although (1.15) is similar in form with (1.11) the term $\exp(-j\omega\tau)$ in (1.15) is a function of frequency, unlike the exponential term in (1.11) which depends only on the center frequency of the narrowband signal.

## 1.3 Problem Formulation

Array processing extensively uses results related to matrices, vectors, and their functions. For a successful presentation, we must establish a consistent notation. After that, we shall describe the effects of the narrowband assumption on the data model given in Section 1.1. We then briefly discuss the assumptions and limitations of the data model. Finally, we state the estimation problems associated with the introduced model.

### 1.3.1 Notation

Lower and upper case italic letters ($b, B$) are used to represent scalars, lower case bold face letters ($\mathbf{b}$) are used for vectors, and upper case bold face letters ($\mathbf{B}$) are used for matrices. The following matrix operators will be used frequently in this thesis:

| | | |
|---|---|---|
| $(\cdot)^*$ | ... | complex-conjugate |
| $(\cdot)^T$ | ... | transpose |
| $(\cdot)^H$ | ... | Hermitian transpose |
| unvec($\mathbf{b}, M, N$) | ... | $MN$ vector to $M$x$N$ matrix conversion |
| det($\mathbf{A}$) | ... | determinant of $\mathbf{A}$ |
| tr($\mathbf{A}$) | ... | trace of $\mathbf{A}$, $\text{tr}(\mathbf{AB}) = \text{tr}(\mathbf{BA})$ |
| $\|\mathbf{A}\|_F$ | ... | Frobenius norm, $\|\mathbf{A}\|_F^2 = \text{tr}(\mathbf{A}^H\mathbf{A})$ |
| $\mathbf{A}^\#$ | ... | left pseudo-inverse, $\mathbf{A}^\# = (\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H$ |
| $\mathbf{P}_\mathbf{A}$ | ... | projection matrix, $\mathbf{P}_\mathbf{A} = \mathbf{A}(\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H = \mathbf{A}\mathbf{A}^\#$ |
| $\mathbf{P}_\mathbf{A}^\perp$ | ... | complementary projection matrix, $\mathbf{P}_\mathbf{A}^\perp = \mathbf{I} - \mathbf{P}_\mathbf{A}$ |
| $\mathbf{A} \odot \mathbf{B}$ | ... | Schur (elementwise) product |
| $\mathbf{A} \otimes \mathbf{B}$ | ... | Kronecker product |

The books [53, 86, 135] provide a good background on linear algebra, which is essential to array signal processing.

### 1.3.2 Data Model

Recall the received signal model (1.3) for the passive sensor array presented in the first section. Assume that quadrature sampling is applied to $\bar{r}(t)$, i.e.,

$$\mathbf{r}(t) = \bar{\mathbf{r}}(t) + j\bar{\mathbf{r}}(t - \frac{\pi}{2\omega_c}) \tag{1.16}$$

Consider a second sensor located at a distance $\vec{d}_{12}$ from the first (reference) sensor, and assume that sensor responses are identical. For simplicity, assume that there is only one wavefront with wave propagation vector $\vec{k}_w(\theta)$. In the case of noiseless observations, we have the relation

$$\bar{r}_2(t) = \bar{r}_1(t - \vec{\xi}(\theta) \cdot \vec{d}_{12}) \tag{1.17}$$

For the analytic signals, we have

$$r_2(t) \approx r_1(t)\, \exp(-j\omega_c \vec{\xi}(\theta) \cdot \vec{d}_{12}) = r_1(t)\, \exp(-j\, \vec{k}(\theta) \cdot \vec{d}_{12}) \tag{1.18}$$

where $\vec{k}(\theta) \triangleq \omega_c \vec{\xi}(\theta)$ denotes the *wavenumber vector* (or, propagation vector) of the wavefront ($\|\vec{k}\| = 2\pi/\lambda$, where $\lambda$ is the wavelength). For the more general case, we can derive the following result

$$\mathbf{r}(t) = \begin{bmatrix} r_1(t) \\ r_2(t) \\ \vdots \\ r_M(t) \end{bmatrix} \approx \begin{bmatrix} \sum_{l=1}^{P} g_1(\theta_l) e^{-j\omega_c \tau_1(\theta_l)}\, s_l(t) \\ \sum_{l=1}^{P} g_2(\theta_l) e^{-j\omega_c \tau_2(\theta_l)}\, s_l(t) \\ \vdots \\ \sum_{l=1}^{P} g_M(\theta_l) e^{-j\omega_c \tau_M(\theta_l)}\, s_l(t) \end{bmatrix} + \begin{bmatrix} n_1(t) \\ n_2(t) \\ \vdots \\ n_M(t) \end{bmatrix} \tag{1.19}$$

where $s_l(t) = \bar{s}_l(t) + j\bar{s}_l(t - \frac{\pi}{2\omega_c})$, and $n_l(t) = \bar{n}_l(t) + j\bar{n}_l(t - \frac{\pi}{2\omega_c})$. Introducing the complex $M$-vector valued function $\mathbf{a}(\theta)$, to include the effects of sensor gains and time-delays, we have

$$\mathbf{a}(\theta) = [g_1(\theta)e^{-j\omega_c \tau_1(\theta)}, g_2(\theta)e^{-j\omega_c \tau_2(\theta)}, \ldots, g_M(\theta)e^{-j\omega_c \tau_M(\theta)}]^T \tag{1.20}$$

Additionally, noise vector $\mathbf{n}(t) = [n_1(t), n_2(t), \ldots, n_M(t)]^T$; hence, the observation equation (1.19) can be expressed as

$$\mathbf{r}(t) = \sum_{l=1}^{P} \mathbf{a}(\theta_l) s_l(t) + \mathbf{n}(t) \tag{1.21}$$

The vector $\mathbf{a}(\theta)$ is referred to as the *steering vector.* It represents the complex-envelope of the voltage induced on the array antennas when a unit amplitude wavefront illuminates the array from the direction represented by the angle $\theta$. For simplicity, we let the steering vector depend on a single parameter (angle) in (1.20). In general, the steering vector may depend on multiple parameters, e.g., azimuth and elevation angles, range, polarization, and center frequency. The *parameter space* $\Omega$ is defined as the set in which $\theta$ is allowed to vary. The collection of steering vectors over the parameter space is defined as the *array manifold*, $\mathcal{A}$

$$\mathcal{A} = \{ \mathbf{a}(\theta) \mid \theta \in \Omega \} \tag{1.22}$$

The definition of array manifold was introduced by Schmidt [118]. An important point to note is that array manifold is not a subspace, it is *only a set*, i.e., it is not necessarily closed under vector addition. This point is illustrated in Figure 1.3 for the case of two sources, and three sensors, in which we assumed the steering vectors are real. Direction-finding algorithms to be described in the next chapter rely on this (*no ambiguity* ) condition: the set of any $P$ ($P < M$) steering vectors (with distinct parameters) is linearly independent.

For simplicity, we consider the steering vectors to be described by a single parameter. In the case of $P$ sources, the total number of direction parameters is $P$, which can be put in a real $P$-vector $\theta$. We introduce the $M$x$P$ complex *steering matrix* $\mathbf{A}(\theta)$

$$\mathbf{A}(\theta) = [\mathbf{a}(\theta_1), \mathbf{a}(\theta_2), \ldots, \mathbf{a}(\theta_P)] \tag{1.23}$$

and the complex $P$-vector of source complex envelopes

$$\mathbf{s}(t) = [s_1(t), s_2(t), \ldots, s_P(t)]^T \tag{1.24}$$

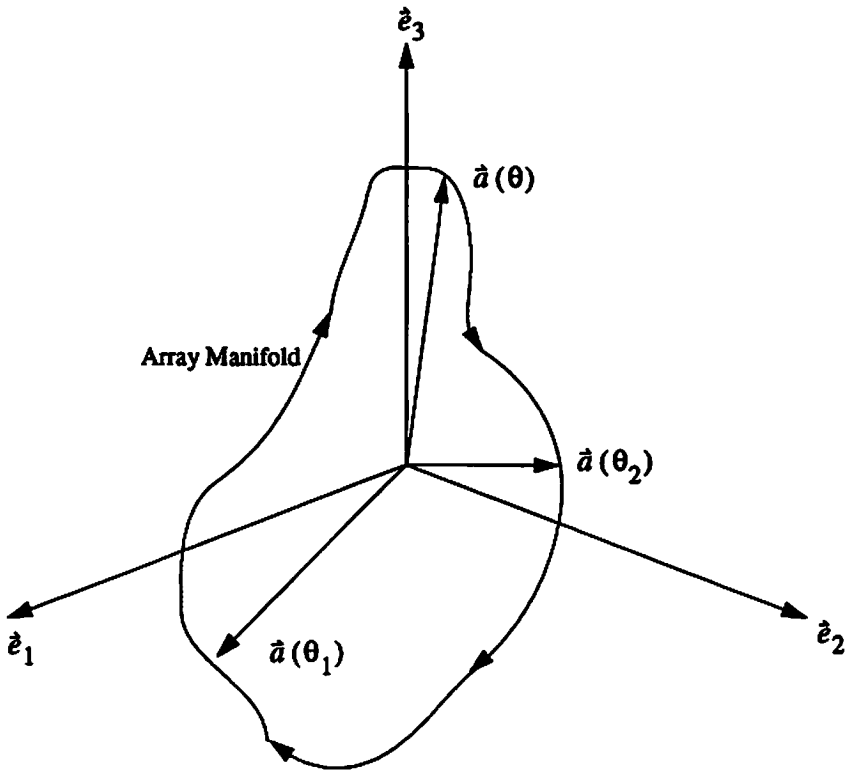



Figure 1.3: Two sources, three sensors scenario: $\vec{e}_1$ and $\vec{e}_2$ are the orthonormal basis vectors for the *subspace* (commonly known as signal-subspace) spanned by the steering vectors $\vec{a}(\theta_1)$ and $\vec{a}(\theta_2)$ of the far-field sources. The ring obtained by varying the argument in $\vec{a}(\theta)$, is the *array manifold.* Array manifold is *not a subspace*; it *only intersects* with the signal subspace at the locations which correspond to the source steering vectors. $\vec{e}_3$ is orthogonal to the signal subspace, and is called the *noise-subspace* (in this case, it is one-dimensional).

which allows us to write the *generic* array processing model in (1.21) as

$$\mathbf{r}(t) = \mathbf{A}(\theta)\,\mathbf{s}(t) \; + \mathbf{n}(t) \tag{1.25}$$

Due to the no-ambiguity assumption, $\mathbf{A}(\theta)$ has rank $P$. It is important to understand the geometry of (1.25) (which is illustrated in Figure 1.3 for a special case). The noiseless part of the $M$-vector $\mathbf{r}(t)$ is a linear combination of the columns of $\mathbf{A}(\theta)$, with coefficients from $\mathbf{s}(t)$. In the noiseless case, the observation vector is confined to the $P$ dimensional range space of $\mathbf{A}(\theta)$, which is a linear subspace of complex $M$ dimensional space. In the presence of noise, the observation process will no longer lie in the range space of $\mathbf{A}(\theta)$; however, the *signal-related* part will still be in the

$P$-dimensional signal subspace. The goal then is to somehow "undo" the effect of noise so as to analyze this subspace. Schmidt [118] proposed a way to accomplish this through the eigendecomposition of the covariance matrix of snapshots.

### 1.3.3 Assumptions and Extensions

To arrive at the generic equation (1.25), we made the following major assumptions: the wavefronts are planar over the aperture of the array, and the received signals are narrowband, which means the complex envelopes of the signals remain almost unchanged during the time it takes for the wavefronts to travel across the array.

Another important assumption in direction-finding applications is the *complete* knowledge of array manifold: for each angle $\theta$, the steering vector $\mathbf{a}(\theta)$ is assumed to be known. This can be achieved through the costly and time-consuming task known as *array calibration*: from each angle of interest, a transmitter in the far-field of the array transmits a signal which is high in power so that the effect of measurement noise can be ignored. The induced complex voltages on the array elements are measured, and stored in memory; hence, the functional description of the array manifold is not required; samples of the manifold are sufficient [118]. Estimation of the array manifold and related hardware issues are discussed in detail in [58, Chapter 2]. Array calibration becomes even more complicated when the parameter space is multidimensional, i.e., if one includes polarization, elevation angle [119], etc. Due to the complexity of calibration there is interest in algorithms which can estimate the direction parameters using only partial knowledge or constraints on the array manifold [108], and knowledge of signal properties [2, 116, 173].

Another assumption is that: for any set of $P$ distinct directions-of-arrival, the set of corresponding steering vectors are linearly independent. This assunption prevents ambiguities. For example, in the case of a uniform linear array, the steering vector takes the form (see Figure 1.1, and note the definition of $\theta$)

$$\mathbf{a}(\theta) = [1, \exp(-j2\pi\Delta \sin\theta/\lambda), \ldots, \exp(-j2\pi(M-1)\Delta \sin\theta/\lambda))]^T \qquad (1.26)$$

which follows[2] from (1.4) since $\omega_c/c = 2\pi/\lambda$. The uniform linear array clearly *has* an ambiguity: $\mathbf{a}(\theta) = \mathbf{a}(\pi - \theta)$. If one considers the parameter space of interest to be the set $\Omega = \{-\pi/2 < \theta < \pi/2\}$, and the element separation is less than $\lambda/2$, then the array becomes unambiguous in the azimuth angle (it is not possible to determine the elevation angle with a linear array).

The previous assumptions guarantee that the noise-free array measurements uniquely determine the DOA's if $P < M$, and the signals are not coherent (not scaled multiples of each other). For coherent signals, a smaller bound is necessary [10, 162].

The noise process $\mathbf{n}(t)$ is generally assumed to be a zero-mean, complex and circularly symmetric Gaussian stationary random process. Furthermore, it is very common to assume that the components of $\mathbf{n}(t)$ are independent of each other and have identical variances. These conditions imply

$$E\{\mathbf{n}(t)\} = \mathbf{0}, \quad E\{\mathbf{n}(t)\mathbf{n}(t)\} = \mathbf{0}, \quad E\{\mathbf{n}(t)\mathbf{n}^H(t)\} = \sigma^2\mathbf{I} \tag{1.27}$$

The signal waveforms $\mathbf{s}(t)$ are in general non-Gaussian random processes; however, for mathematical analysis purposes (e.g., developing Cramer-Rao bounds), signals are frequently considered to be independent Gaussian random processes. In this case, the signal covariance matrix $E\{\mathbf{s}(t)\mathbf{s}^H(t)\}$ is diagonal. In other applications, signal waveforms are modeled as deterministic processes. In this case, the only assumption is that the sum

$$\mathbf{R}_{ss} \triangleq \lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} \mathbf{s}(t)\mathbf{s}^H(t) \tag{1.28}$$

exists. $\mathbf{R}_{ss}$ is called the *signal covariance matrix*. In the random case, $\mathbf{R}_{ss} = E\{\mathbf{s}(t)\mathbf{s}^H(t)\}$. The signal waveforms are assumed to be independent of the noise process, and the covariance matrix of array outputs takes the form

$$\mathbf{R} \triangleq E\{\mathbf{r}(t)\mathbf{r}^H(t)\} = \mathbf{A}(\theta)\mathbf{R}_{ss}\mathbf{A}^H(\theta) + \sigma^2\mathbf{I} \tag{1.29}$$

[2]Due to the Vandermonde structure of the steering vector in (1.26), it is possible to develop an analogy with time-series analysis of sinusoids and devise spatial filters based on principles of FIR filter design [130].

There are applications in which the just-described model fails. For example, the received signals may deviate from the narrowband assumption. This problem is addressed in the frequency-domain by preprocessing data using the Discrete Fourier Transform (DFT) and transforming the broadband problem into a series of narrowband problems [161]. Alternative approaches have been developed to combine these individual narrowband problems together [62, 68, 69, 159].

Another restriction is on the noise among the sensors. Due to mutual-coupling among sensors and near-field noise, sensor to sensor independence assumption can easily be violated. Talham [145] evaluates the spatial spectra for special cases. If the noise is actually Gaussian, but the noise covariance matrix is modeled as a completely unknown Hermitian matrix, then the number of parameters to be estimated increases, and this problem is unsolvable in general since the stored array manifold does not have any information about these new components. For this reason, it is generally assumed that the noise covariance is *known* to within a scale factor $\sigma^2$, i.e., $\mathbf{R}_{nn} = \sigma^2\mathbf{\Sigma}_n$. Then, prewhitening is applied on the received signals as

$$\mathbf{r}(t) = \mathbf{\Sigma}_n^{-1/2}\mathbf{r}(t) = \mathbf{\Sigma}_n^{-1/2}\mathbf{A}(\theta)\mathbf{s}(t) + \mathbf{\Sigma}_n^{-1/2}\mathbf{n}(t) \tag{1.30}$$

which makes the undesired component white. During this operation (known as Mahalanobis transformation) the steering vectors change from $\mathbf{a}(\theta)$ to $\mathbf{\Sigma}_n^{-1/2}\mathbf{a}(\theta)$. The noise covariance structure ($\mathbf{\Sigma}_n$) can be measured if there are no signals present. If the Gaussian noise assumption is valid, then cumulants can be used to remove its effects provided that the signals of interest are non-Gaussian.

The planar wavefront approximation over the aperture is also invalid in some cases. This problem is treated in [129, 138], and also addressed in Chapter 5 of this thesis.

### 1.3.4 The Estimation Problem

Given the snapshots $\{\mathbf{r}(1), \mathbf{r}(2), \ldots, \mathbf{r}(N)\}$, modeled by the generic equation

$$\mathbf{r}(t) = \mathbf{A}(\theta)\mathbf{s}(t) + \mathbf{n}(t) \tag{1.31}$$

the estimation problem consists of the following three subproblems:

- Detect the number of signals, $P$.
- Estimate $\theta$, the vector of signal parameters.
- Recover signal waveforms (messages), $\mathbf{s}(t)$.

Algorithms that address these subproblems were originally developed using second-order statistics. In these algorithms, the observations are used to form the *sample* covariance matrix of the array,

$$\hat{\mathbf{R}}_N \triangleq \frac{1}{N}\sum_{t=1}^{N}\mathbf{r}(t)\mathbf{r}^H(t) \tag{1.32}$$

and the parameters of interest are estimated by processing the sample covariance matrix. Any parametric method for estimating the signal parameters requires an estimate of the number of parameters (detection problem). In this thesis, our emphasis is directed at the advantages offered by using higher-order statistics in the estimation problem. The detection problem is studied extensively in the book [149] and in the papers [158, 162, 179] for array processing applications.

# Chapter 2

# Spatial Spectrum Estimation Techniques

In this chapter, we review some of the parameter estimation algorithms that will be used in the thesis. The parameters of interest are the *directions* of far-field sources. The reader is referred to the books [55, 56, 58, 64, 147] and papers [1, 4, 10, 18, 19, 33, 39, 46, 47, 62, 63, 68, 69, 70, 77, 103, 108, 127, 128, 151, 118, 153, 157, 165, 159] for a more complete study of estimation techiques in array processing.

## 2.1 Beamforming Methods

Beamforming techniques are based on estimating a *spatial spectrum*, $P(\theta)$, which represents the energy in the measurements received from the direction represented by $\theta$. A peak in $P(\theta)$ indicates the presence of a source from $\theta$.

Beamforming methods do not fully exploit the data model, compared to the eigenstructure methods of the next section. There are advantages in this however: beamforming methods are robust to model errors and are computationally simple. The obvious disadvantage is the performance degradation when the model is actually accurate.

### 2.1.1 Conventional Beamformer

Conventional beamforming can be considered as a power maximization technique. Consider a single wavefront illuminating the array from an angle $\theta$. The noiseless measurements ($N$ snapshots) are modeled as

$$\mathbf{r}(t) = \mathbf{a}(\theta)s(t) \tag{2.1}$$

The output power of the linear combiner $y(t) = \mathbf{w}^H\mathbf{r}(t)$ is measured as the sample mean

$$\hat{P}(\mathbf{w}) = \frac{1}{N}\sum_{t=1}^{N}\|\mathbf{w}^H\mathbf{r}(t)\|^2 = \mathbf{w}^H\hat{\mathbf{R}}_N\mathbf{w} \tag{2.2}$$

where $\hat{\mathbf{R}}_N \triangleq \frac{1}{N}\sum_{t=1}^{N}\mathbf{r}(t)\mathbf{r}^H(t)$ is the sample covariance matrix of array measurements. The power function $\hat{P}(\mathbf{w})$ can be reexpressed as

$$\hat{P}(\mathbf{w}) = \|\mathbf{w}^H\mathbf{a}(\theta)\|^2\,\frac{1}{N}\sum_{t=1}^{N}\|s(t)\|^2 \tag{2.3}$$

Clearly, $\hat{P}(\mathbf{w})$ is maximized when $\mathbf{w} = \mathbf{a}(\theta)$. When these weights are used, the spatial spectrum estimate from the conventional beamformer is obtained from (2.2), as

$$\hat{P}(\theta) = \mathbf{a}^H(\theta)\hat{\mathbf{R}}_N\mathbf{a}(\theta) \tag{2.4}$$

If the norm of $\mathbf{a}(\theta)$ is not constant over the angles of interest, then a normalization is necessary. The spatial spectrum from the *conventional beamformer* can then be expressed as

$$\hat{P}_c(\theta) = \frac{\mathbf{a}^H(\theta)\hat{\mathbf{R}}_N\mathbf{a}(\theta)}{\mathbf{a}^H(\theta)\mathbf{a}(\theta)} \tag{2.5}$$

The spatial spectrum from the conventional beamformer can also be expressed as

$$\hat{P}_c(\theta) = \frac{\sum_{k=1}^{M}\lambda_k\,\|\mathbf{e}_k^H\mathbf{a}(\theta)\|^2}{\mathbf{a}^H(\theta)\mathbf{a}(\theta)} \tag{2.6}$$

where $\{\lambda_1,\ldots,\lambda_M\}$ and $\{\mathbf{e}_1,\ldots,\mathbf{e}_M\}$ denote the set of eigenvalues and eigenvectors of the sample covariance matrix respectively.

As seen from its derivation, the classical beamformer assumes a single source in no noise. There is no consideration of multiple sources or of colored noise; however, if the sources are well separated, so that their steering vectors are almost orthogonal, it is possible to resolve multiple sources. In this respect, the conventional beamformer resembles the classical Fourier analysis of time series: due to non-parametrization, resolution is limited by the window size (here, by aperture size). Although the resolution of multiple emitters by the conventional beamformer is not satisfactory, in the single emitter case the conventional beamformer *becomes* the deterministic maximum-likelihood processor.

### 2.1.2 Capon's Beamformer (MVDR)

Array processing techniques play an important role in enhancement of signals in the presence of interference. Capon's minimum-variance distortionless response (MVDR) beamformer [19] has been a starting point for both signal enhancement and high-resolution direction-of-arrival (DOA) estimation. This method is sometimes (misleadingly) called the "maximum-likelihood (ML) method" because of its similarity to the ML estimator of the amplitude of a sine wave of known frequency in Gaussian noise [97, Chapter 2]. MVDR does not maximize the likelihood, instead it is motivated as a beamformer with *disturbance rejection.*

Suppose we want to estimate the power of the source from an angle $\theta$. Therefore, we need to minimize the contribution of all sources from other directions. Then, we need to constrain the weight vector, $\mathbf{w}$, so that the signal from $\theta$ passes undistorted while minimizing the output power

$$\min_{\mathbf{w}} \hat{P}(\mathbf{w}) \quad \text{subject to} \quad \mathbf{w}^H \mathbf{a}(\theta) = 1 \tag{2.7}$$

Since the desired signal passes undistorted, minimizing the output power implies minimizing the contribution of all other signals. The solution to this constrained optimization problem can be solved by the method of Lagrange multipliers [7, 45],[147, Appendix A], and is

$$\mathbf{w} = \frac{\hat{\mathbf{R}}_N^{-1} \mathbf{a}(\theta)}{\mathbf{a}^H(\theta) \hat{\mathbf{R}}_N^{-1} \mathbf{a}(\theta)} \tag{2.8}$$

The spatial spectrum estimate from the *Capon's beamformer* can then be formed as

$$\hat{P}_{ca}(\theta) = \frac{1}{\mathbf{a}^H(\theta)\hat{\mathbf{R}}_N^{-1}\mathbf{a}(\theta)} \tag{2.9}$$

In terms of the eigendecomposition of the sample covariance matrix, the Capon's spatial spectrum estimate becomes

$$\hat{P}_{ca}(\theta) = \Big( \sum_{k=1}^{M} \frac{1}{\lambda_k} \|\mathbf{e}_k^H \mathbf{a}(\theta)\|^2 \Big)^{-1} \tag{2.10}$$

MVDR can be used to recover the waveform of source whose steering vector is known. Suppose we want to estimate the waveform from a source at $\theta_o$. Then,

$$y(t) = \mathbf{w}^H \mathbf{r}(t) \quad \text{where} \quad \mathbf{w} = \frac{\hat{\mathbf{R}}_N^{-1}\mathbf{a}(\theta_o)}{\mathbf{a}^H(\theta_o)\hat{\mathbf{R}}_N^{-1}\mathbf{a}(\theta_o)} \tag{2.11}$$

$y(t)$ is an estimate of the waveform of the source from $\theta_o$ in a mean-square sense. The equivalence of MVDR to other processors is determined in [15]. The performance and sensitivity of the MVDR beamformer is well-analyzed [27, 28, 31, 32, 37, 38, 50, 104, 176, 181]. The processor is very sensitive to the matrix inversion in (2.11) and errors in the steering vector ($\mathbf{a}(\theta_o)$) of the desired source, frequently resulting in the cancellation of the desired signal. To avoid this phenomenon, additional constraints are necessary, such as look-direction constraints in the vicinity of $\theta_o$ and derivative constraints on the spatial response of the array at $\theta_o$ [16]. If coherent sources accompany the desired signal, then it is necessary to put nulls on them; and this can be formulated as additional constraints. These constraints are linear in $\mathbf{w}$, hence they can be expressed as the matrix equation

$$\mathbf{C}^H\mathbf{w} = \mathbf{f} \tag{2.12}$$

The nonsingular matrix $\mathbf{C}$ is called the *constraint matrix.* $\mathbf{f}$ is a vector whose length and contents are determined by the number and type of constraints. The weight vector $\mathbf{w}$ is then found as the solution of

$$\min_{\mathbf{w}} \hat{P}(\mathbf{w}) \quad \text{subject to} \quad \mathbf{C}^H\mathbf{w} = \mathbf{f} \tag{2.13}$$

The solution can again be obtained by using Lagrange multipliers [45], and is

$$\mathbf{w} = \hat{\mathbf{R}}_N^{-1}\,\mathbf{C}\,(\,\mathbf{C}^H\,\hat{\mathbf{R}}_N^{-1}\,\mathbf{C}\,)^{-1}\,\mathbf{f} \tag{2.14}$$

The processor with the weight vector in (2.14) is called the *linearly-constrained minimum-variance* beamformer. Its adaptive implementation is discussed in [45, 54]. Bresler *et* al. [11] proposed a class of beamforming algorithms for coherent signal and interferences. Friedlander and Porat [43] investigated the performance of a null-steering signal recovery procedure based on DOA estimation.

## 2.2 Maximum-Likelihood Methods

The maximum-likelihood (ML) estimator requires the knowledge of conditional probability density function of the data given the unknown parameters that will be estimated. The ML estimator has very desirable properties that hold for sufficiently regular likelihood functions. *If* the ML estimates are *consistent*, then they are also asymptotically *efficient*; therefore, for sufficiently long data lengths, the ML estimator attains the Cramer-Rao bound on the estimation error variance for the parameters of interest.

Depending on the data model of array processing, ML technique has two different forms. The stochastic ML assumes the signals are Gaussian random processes, whereas the deterministic ML models the signals as deterministic (but unknown) quantities. We refer the reader to [9, 41, 105, 106, 129, 131, 132, 133, 157, 158, 160, 162, 164, 166, 172, 180] for the applications of ML estimation in array processing problems.

### 2.2.1 Stochastic ML Method

Stochastic ML method assumes that the signals are distributed as *jointly Gaussian* random processes. This brings considerable simplicity to develop the estimator. Since measurements are Gaussian, their second-order statistics (the covariance matrix, $\mathbf{R}$) completely describe their probability structure; from (1.29) we have

$$\mathbf{R} \triangleq \mathbf{A}(\theta)\mathbf{R}_{ss}\mathbf{A}^H(\theta) + \sigma_n^2\mathbf{I} \tag{2.15}$$

where $\mathbf{R}_{ss}$ is the covariance matrix of sources. The likelihood function of a single observation is

$$p(\mathbf{r}) = \frac{1}{\pi^M \det(\mathbf{R})} \exp(-\mathbf{r}^H \mathbf{R}^{-1} \mathbf{r}) \tag{2.16}$$

If snapshots are considered independent and identically distributed, then the set of measurements $\{\mathbf{r}(1), \ldots, \mathbf{r}(N)\}$ has the likelihood

$$L\{\mathbf{r}(1), \ldots, \mathbf{r}(N)\} = \prod_{t=1}^{N} \frac{1}{\pi^M \det(\mathbf{R})} \exp(-\mathbf{r}^H(t) \mathbf{R}^{-1} \mathbf{r}(t)) \tag{2.17}$$

By considering the log-likelihood, and after standard mathematical analysis, parameter estimation turns out to minimize the functional $l(\theta, \mathbf{R}_{ss}, \sigma_n^2)$, defined as

$$l(\theta, \mathbf{R}_{ss}, \sigma_n^2) \triangleq \log(\det(\mathbf{R})) + \mathrm{tr}(\mathbf{R}^{-1} \hat{\mathbf{R}}_N) \tag{2.18}$$

where $\hat{\mathbf{R}}_N$ is the sample covariance matrix. The unknown parameters are $P$ DOA's, noise variance $\sigma_n^2$, and the entries of the source covariance matrix $\mathbf{R}_{ss}$. The dimension of *optimization* can be *reduced* [5] to the number of sources:

$$\hat{\theta} = \arg\min_{\theta} \{ \det(\mathbf{P}_{\mathbf{A}} \hat{\mathbf{R}}_N \mathbf{P}_{\mathbf{A}} + \frac{1}{(M-P)} \mathrm{tr}(\mathbf{P}_{\mathbf{A}}^{\perp} \hat{\mathbf{R}}_N) \mathbf{P}_{\mathbf{A}}^{\perp}) \} \tag{2.19}$$

Although the dimension of optimization is reduced to $P$, it is still very costly to carry out the minimization in (2.19). There is no guarantee that the optimization will lead to a global minimum, due to the nonlinear character of the cost function in (2.19); it is possible that any iterative numerical approach can be trapped in a local minimum. This may be avoided if the optimization can be initialized with reliable estimates. In the case of global convergence, high-quality estimates are expected.

The likelihood function for the stochastic maximum-likelihood model is sufficiently regular, and the asymptotic covariance of the estimates asymptotically attain the *stochastic Cramer-Rao* bound (CRB) [87, 157, 158]. It is also known that stochastic modeling results in better detection algorithms than the deterministic model of the next section [157, 164].

### 2.2.2 Deterministic ML Method

In many applications concerning man-made signals, it is unrealistic to make the Gaussian signals assumption. In this case, considering *signals* as *deterministic processes* seems to be more reasonable if one insists on ML estimation.

In the *deterministic model*, the observations are Gaussian, given the values of unknown parameters. The first two statistics of $\mathbf{r}(t)$ are:

$$\mathbf{m}_\mathbf{r}(t) \triangleq E\{\mathbf{r}(t)\} = \mathbf{A}(\theta)\mathbf{s}(t) \tag{2.20}$$

and

$$E\{(\mathbf{r}(t) - \mathbf{m}_\mathbf{r}(t))(\mathbf{r}(t) - \mathbf{m}_\mathbf{r}(t))^H\} = \sigma_n^2\mathbf{I} \tag{2.21}$$

The unknown parameters are the angles $\theta$, the source waveforms $\{\mathbf{s}(1), \ldots, \mathbf{s}(N)\}$ and the noise variance $\sigma_n^2$.

The estimate for the noise variance takes the form [180]

$$\hat{\sigma}_s^2 = \frac{1}{(M-P)} \operatorname{tr}\left(\mathbf{P}_\mathbf{A}^\perp \hat{\mathbf{R}}_N\right) \tag{2.22}$$

whereas the direction and waveform estimates are obtained through the optimization

$$\{\hat{\theta}, \hat{\mathbf{s}}(t)\} = \arg\min_{\theta, \mathbf{s}(t)} \frac{1}{N} \sum_{t=1}^{N} \|\mathbf{r}(t) - \mathbf{A}(\theta)\mathbf{s}(t)\|_F^2 \tag{2.23}$$

The source parameters can be estimated as [180]

$$\hat{\mathbf{s}}(t) = \mathbf{A}^\#(\theta)\mathbf{r}(t) \tag{2.24}$$

and, therefore,

$$\hat{\theta} = \arg\min_\theta \operatorname{tr}\left(\mathbf{P}_\mathbf{A}^\perp(\theta)\,\hat{\mathbf{R}}_N\right) = \arg\max_\theta \operatorname{tr}\left(\mathbf{P}_\mathbf{A}(\theta)\,\hat{\mathbf{R}}_N\right) \tag{2.25}$$

Although the deterministic ML formulation leads to a cost function (2.25) which is *less complicated* than the stochastic counterpart (2.19), the optimization is still computationally costly. In both cases, the criterion functions possess local minima

since they are highly nonlinear in the unknown parameters; however, if there is only one emitter, then the ML cost function simplifies to

$$\hat{\theta} = \arg\max_{\theta} \mathrm{tr}\left( \mathbf{P_a}(\theta)\hat{\mathbf{R}}_N \right) = \arg\max_{\theta} \frac{\mathbf{a}^H(\theta)\hat{\mathbf{R}}_N\mathbf{a}(\theta)}{\mathbf{a}^H(\theta)\mathbf{a}(\theta)} = \arg\max_{\theta} P_c(\theta) \qquad (2.26)$$

(since $\mathbf{P_a} = \mathbf{a}\mathbf{a}^H/(\mathbf{a}^H\mathbf{a})$ ) which demonstrates that the deterministic ML beamformer is *equivalent* to the conventional beamformer.

It is important to note that number of the parameters in the deterministic ML method *grows* with the number of snapshots, since the source waveforms are treated as unknown (deterministic) parameters. As pointed out in [132], this makes the deterministic ML function *non-regular*. Although $\hat{\theta}_{ML}$ is consistent, unless the number of sensors ($M$) tends to infinity, it is *not efficient*. In [87], it is shown that the estimation error variance from deterministic ML is equal to or greater than than that of the stochastic ML method for large number of snapshots if the signals are uncorrelated. If the emitter signals are *uncorrelated*, then the asymptotic variances of the DOA estimates obtained from stochastic ML *depend only* on the *second moments* of the source signals [87], and the stochastic ML estimator is *always superior* to the deterministic ML estimator regardless of the actual probability density functions of the independent sources.

Although the ML approaches have certain optimality properties, they require the solution of nonlinear multidimensional optimization problems that introduce a *severe* computational load (stochastic ML is optimal for large $N$, whereas deterministic ML is optimal for large $M$). Due to these computational concerns, much array processing research is focused on suboptimal approaches.

## 2.3 Subspace Methods

The eigenstructure-based suboptimal methods (signal subspace methods) originated from Pisarenko's harmonic retrieval algorithm [98] which was developed to estimate pole locations from impulse response measurements. More powerful algorithms, such as the multiple signal classification (MUSIC) [118], or estimation of signal parameters via subspace rotation techniques (ESPRIT) [108], exploit the *geometrical structure*

of the estimation problem represented in terms of the sample covariance matrix. The array covariance matrix has the form

$$\mathbf{R} = \mathbf{A}\mathbf{R}_{ss}\mathbf{A}^H + \sigma_n^2\mathbf{I} \tag{2.27}$$

$\mathbf{R}$ is Hermitian and has the eigendecomposition

$$\mathbf{R} = \sum_{k=1}^{M} \lambda_k \mathbf{e}_k \mathbf{e}_k^H = \mathbf{E}\Lambda\mathbf{E}^H \tag{2.28}$$

where the real and positive eigenvalues of $\mathbf{R}$ are arranged in non-increasing order, and the eigenvectors are normalized so that they have unity norm. The following observation is the foundation for the signal subspace methods.

• Fact [118]: Assume the signal covariance matrix $\mathbf{R}_{ss}$ is full-rank $P$. Then, it is possible to partition the eigendecomposition of $\mathbf{R}$ as

$$\mathbf{E} = [\mathbf{E}_s, \mathbf{E}_n], \quad \Lambda = \mathrm{diag}(\Lambda_s, \Lambda_n), \quad \Lambda_s = \mathrm{diag}(\lambda_1, \ldots, \lambda_P), \quad \Lambda_n = \sigma_n^2\mathbf{I} \tag{2.29}$$

where $\mathbf{E}_s$ is $M$x$P$, and $\mathbf{E}_n$ is $M$x$(M-P)$. The first $P$ eigenvalues are all greater than $\sigma_n^2$ which has multiplicity of $M-P$. More importantly, the columns of $\mathbf{E}_s$ are the basis vectors for the subspace which is spanned by the columns of $\mathbf{A}$, i.e.,

$$\mathrm{span}(\mathbf{E}_s) = \mathrm{span}(\mathbf{A}) \tag{2.30}$$

where $\mathrm{span}(\mathbf{E}_s)$ denotes the range space of $\mathbf{E}_s$. Proof of this fact follows directly from the observation that the eigenvectors of $\mathbf{R}$ are also the eigenvectors of the positive semi-definite matrix $\mathbf{A}\mathbf{R}_{ss}\mathbf{A}^H$ [118].

### 2.3.1 Multiple Signal Classification (MUSIC)

If the emitter signals are not fully correlated (i.e., not coherent), then the signal covariance matrix $\mathbf{R}_{ss}$ has full-rank $P$. Then, since the range spaces of $\mathbf{E}_s$ and $\mathbf{A}$ are identical, by (2.30) and $\text{span}(\mathbf{E}_n) \perp \text{span}(\mathbf{E}_s)$, we have the important result

$$\text{span}(\mathbf{E}_n) \perp \text{span}(\mathbf{A}) \tag{2.31}$$

where $\text{span}(\mathbf{E}_n)$ is called as the *noise subspace* and the columns of $\mathbf{E}_n$ are called the *noise eigenvectors.* If the signals are not coherent, then steering vectors of the sources that contribute to the measurements are orthogonal to the noise subspace, i.e.,

$$\mathbf{E}_n^H \, \mathbf{a}(\theta_k) = \mathbf{0}, \qquad 1 \leq k \leq P \tag{2.32}$$

Due to the unambiguity assumption on the array manifold, (2.32) determines the source directions uniquely by a one dimensional search procedure. A consistent estimate of $\mathbf{E}_n$ is obtained from the sample covariance matrix

$$\hat{\mathbf{R}}_N = \hat{\mathbf{E}}_s \hat{\mathbf{\Lambda}}_s \hat{\mathbf{E}}_s^H + \hat{\mathbf{E}}_n \hat{\mathbf{\Lambda}}_n \hat{\mathbf{E}}_n^H \tag{2.33}$$

and then, considering the property in (2.32), the *normalized* MUSIC *spectrum* is calculated, as

$$P_{mu}(\theta) = \frac{\mathbf{a}^H(\theta)\mathbf{a}(\theta)}{\mathbf{a}^H(\theta)\,\mathbf{E}_n\,\mathbf{E}_n^H\,\mathbf{a}(\theta)} = \frac{\mathbf{a}^H(\theta)\mathbf{a}(\theta)}{\sum_{k=P+1}^{M} \| \, \mathbf{e}_k^H \, \mathbf{a}(\theta) \, \|^2} \tag{2.34}$$

The MUSIC estimates are the $P$ *maxima* of $P_{mu}(\theta)$. Note the similarity of (2.34) to the Capon beamformer (2.10). MUSIC estimator uses only the noise subspace eigenvectors, whereas the Capon beamformer uses all the eigenvectors, by scaling them with the reciprocal of the corresponding eigenvalues. If the norm of the steering vector is constant over the array manifold, and the signal to noise ratio tends to infinity, the Capon estimator has the *same* structure as the MUSIC estimator.

Generally, the MUSIC algorithm offers superior estimates as compared to beamforming methods: MUSIC requires a full eigendecomposition, which brings a slight

increase in computational load. Recent work [175] has addressed recursive eigendecomposition for subspace analysis.

MUSIC estimates are consistent, and they converge to true values as the number of snapshots grow to infinity [66, 118]. A *drawback* of the MUSIC algorithm is its *sensitivity* to model errors [44, 141]: MUSIC requires the *costly* procedure of array *calibration* for all parameters of interest, and is very sensitive to errors in the calibration procedure. Although the cost of array calibration can be decreased by the manifold interpolation technique of Schmidt [120], calibration remains to be a major problem. Friedlander [42] proposed a signal subspace method that uses the properties of the signal of interest for blind adaptive interference cancellation. The cost of calibration *increases* as the parameters that define the the array manifold increase, e.g., polarization [39]. An experimental implementation of the MUSIC algorithm is described in [119].

For uncorrelated sources, MUSIC can compete with the computationally more expensive deterministic ML method [42, 132, 133]. In the case of a uniform linear array, MUSIC resembles a root-finding procedure, and the resulting algorithm is called *root*-MUSIC. The resolution of root-MUSIC has been proved to be better than the MUSIC algorithm [101]. In the case of large arrays, *beam*forming and sub*space* (*beamspace* direction-finding) ideas can be combined to improve the estimation results [182].

When the sources are coherent, the rank of $\mathbf{R}_{ss}$ is less than the number of emitters. In the case of a linear array, spatial smoothing algorithms are proposed to restore the rank of the source covariance matrix [94, 96, 104, 123, 124, 144, 170, 171]. In the case of arbitrary arrays, the deterministic ML algorithm or the so-called multidimensional MUSIC algorithm [18] can be used, with an added expense in computational requirements. If the emitter waveforms are broadband, then DFT can be used to *separate* the wideband problem into several narrowband problems which can be properly averaged to cope with source coherency [62, 68, 69, 159].

### 2.3.2 Rotational Invariance Techniques (ESPRIT)

Due to the array calibration problems of the MUSIC algorithm, the ESPRIT algorithm (estimation of signal parameters via rotational invariance techniques) is

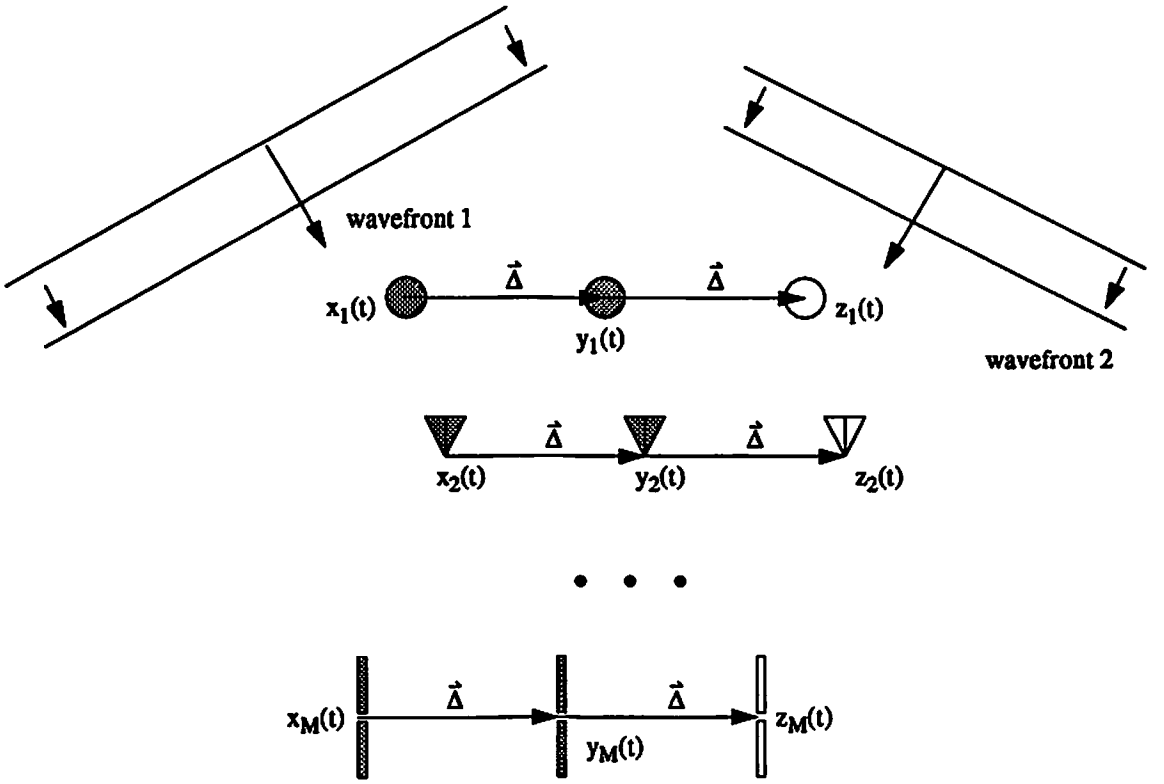



Figure 2.1: Array configuration for the ESPRIT algorithm. The filled sensors are sufficient for single DI-ESPRIT, but additional redundancy can be added using unfilled sensors for multiple invariance ESPRIT [140]. The sensors that have different shapes are presented to emphasize the fact that ESPRIT can be applied to any array consisting of sensors of arbitrary response.

introduced [108]. ESPRIT requires a specific type of *redundancy* in the array structure to eliminate the requirement for the knowledge of the array manifold. In this section, we describe the total-least squares version of the ESPRIT algorithm [108] (TLS-ESPRIT). Performance of TLS-ESPRIT algorithm is analyzed in [88, 134]. An experimental ESPRIT direction-finding system is described in [65].

Assume that there are $P$ narrowband far field sources illuminating an array of $M$ sensor doublets. The elements in each doublet have identical response patterns and are translationally separated by a known constant displacement vector $\vec{\Delta}$ ($\Delta \triangleq \|\vec{\Delta}\|_F$). The responses of each sensor in a subarray are quite arbitrary except for the requirement that the response of sensors are nonzero over the region of interest. Figure 2.1 illustrates the array set-up.

Received signals for the two subarrays $\mathcal{S}_x$ and $\mathcal{S}_y$ are expressed as

$$\begin{aligned} x_i(t) &= \sum_{k=1}^{P} s_k(t)a_i(\theta_k) + n_{x_i}(t) \\ y_i(t) &= \sum_{k=1}^{P} s_k(t)a_i(\theta_k)\exp(-j\omega_c\Delta\sin\theta_k/c) + n_{y_i}(t) \end{aligned} \tag{2.35}$$

where $1 \leq i \leq M$, $\theta_k$ is the direction-of-arrival of the $k$th source with respect to the direction of the translational displacement vector $\vec{\Delta}$, $\omega_c$ is the center frequency and $c$ is the propagation speed. The above equations can be expressed in vector form as

$$\begin{aligned} \mathbf{x}(t) &= \mathbf{A}\mathbf{s}(t) + \mathbf{n}_x(t) \\ \mathbf{y}(t) &= \mathbf{A}\Phi\mathbf{s}(t) + \mathbf{n}_y(t) \end{aligned} \tag{2.36}$$

in which $\mathbf{x}(t)$ and $\mathbf{y}(t)$ are $M$ dimensional vectors, $\mathbf{A}$ is the $M$x$P$ ($M > P$) full-rank steering matrix and $\Phi$ is a diagonal $P$x$P$ matrix of the phase delays between the doublet sensors for $P$ wavefronts, i.e.,

$$\Phi = \text{diag}\{\exp(-j\omega_c\Delta\sin\theta_1/c), \ldots, \exp(-j\omega_c\Delta\sin\theta_P/c)\} \tag{2.37}$$

The $P$x$P$ unitary matrix $\Phi$ relates the measurements from subarray $\mathcal{S}_x$ to those from subarray $\mathcal{S}_y$. Clearly, $\Phi$ is nonsingular.

If we concacanate observations from the two subarrays as

$$\mathbf{z}(t) = \begin{bmatrix} \mathbf{x}(t) \\ \mathbf{y}(t) \end{bmatrix}, \tag{2.38}$$

then the covariance matrix of $\mathbf{z}(t)$ is

$$\mathbf{R}_{zz} = E\{\mathbf{z}(t)\mathbf{z}^H(t)\} = \underbrace{\begin{bmatrix} \mathbf{A} \\ \mathbf{A}\Phi \end{bmatrix}}_{\triangleq\bar{\mathbf{A}}} \mathbf{R}_{ss} \begin{bmatrix} \mathbf{A} \\ \mathbf{A}\Phi \end{bmatrix}^H + \sigma^2\mathbf{I} \tag{2.39}$$

in which $\mathbf{R}_{ss}$ is the covariance matrix of $P$ non-coherent far-field sources. The eigenvectors of the covariance matrix $\mathbf{R}_{zz}$, corresponding to its $P$ largest eigenvalues will satisfy

$$\mathbf{E}_s \triangleq \begin{bmatrix} \mathbf{E}_x \\ \mathbf{E}_y \end{bmatrix} = \underbrace{\begin{bmatrix} \mathbf{A} \\ \mathbf{A}\,\mathbf{\Phi} \end{bmatrix}}_{\bar{\mathbf{A}}} \mathbf{T} \tag{2.40}$$

for a nonsingular $P$x$P$ matrix $\mathbf{T}$. Since $\mathbf{E}_x$ and $\mathbf{E}_y$ share a common column space, the rank of $[\mathbf{E}_x, \mathbf{E}_y]$ is $P$; therefore, the $2P$x$2P$ matrix $\mathbf{E}_{xy}$

$$\mathbf{E}_{xy} \triangleq \begin{bmatrix} \mathbf{E}_x^H \\ \mathbf{E}_y^H \end{bmatrix} \begin{bmatrix} \mathbf{E}_x & \mathbf{E}_y \end{bmatrix} \tag{2.41}$$

is a non-negative definite Hermitian matrix of rank $P$, so it has $P$ null eigenvalues; therefore, there is a rank $P$, $2P$x$P$ matrix[1] $\mathbf{F}$, such that

$$\mathbf{0} = [\mathbf{E}_x, \mathbf{E}_y]\mathbf{F} = \mathbf{E}_x\mathbf{F}_x + \mathbf{E}_y\mathbf{F}_y \tag{2.42}$$

This can be rewritten as

$$\mathbf{0} = \mathbf{A}\mathbf{T}\mathbf{F}_x + \mathbf{A}\mathbf{\Phi}\mathbf{T}\mathbf{F}_y \tag{2.43}$$

Since $\mathbf{A}$ is full-rank, this is equivalent to

$$\mathbf{T}\mathbf{F}_x + \mathbf{\Phi}\mathbf{T}\mathbf{F}_y = \mathbf{0} \tag{2.44}$$

or

$$\mathbf{T}^{-1}\mathbf{\Phi}\mathbf{T} = -\mathbf{F}_x\mathbf{F}_y^{-1} \tag{2.45}$$

which implies that the eigenvalues of $\mathbf{\Psi} \triangleq -\mathbf{F}_x\mathbf{F}_y^{-1}$ must be equal to the diagonal elements of $\mathbf{\Phi}$, and the columns of $\mathbf{T}^{-1}$ are the eigenvectors of $\mathbf{\Psi}$ ($\mathbf{E}_{\mathbf{\Psi}} = \mathbf{T}^{-1}$); hence, the DOA parameters can be obtained directly from an eigenanalysis of $\mathbf{\Psi}$. The remaining problem is to find such a matrix $\mathbf{F}$. Since, $\mathbf{F}$ must span the nullspace

[1]Such a matrix is not unique, as incorrectly claimed in [108], since $\mathbf{FQ}$ satisfies (2.42) with the same rank condition if $\mathbf{Q}$ is nonsingular; however, the arbitrary matrix $\mathbf{Q}$ does not effect the estimation procedure because it is canceled in (2.45).

of $[\mathbf{E}_x, \mathbf{E}_y]$, the obvious choice is to select $\mathbf{F}$ as the eigenvectors of $\mathbf{E}_{xy}$ with null eigenvalues.

In summary, the TLS-ESPRIT algorithm: forms $\mathbf{R}_{zz}$ in (2.39); performs an eigendecomposition of $\mathbf{R}_{zz}$ to obtain $\mathbf{E}_s$; determines $\mathbf{E}_x, \mathbf{E}_y$ in (2.40) to get $\mathbf{E}_{xy}$ in (2.41); computes $\mathbf{F}_x, \mathbf{F}_y$ from the null-space eigenvectors of $\mathbf{E}_{xy}$; and obtains DOA parameters by computing the eigenvalues in (2.45). In this way, a total-least squares (TLS) *cost function* is optimized and the resulting algorithm is called TLS-ESPRIT.

It is also possible to estimate the steering vectors (up to a scale factor) corresponding to the far-field sources whose bearings are already estimated. This can be accomplished by computing

$$\mathbf{E}_s \mathbf{E}_\Psi = \bar{\mathbf{A}} \mathbf{T} \mathbf{T}^{-1} = \bar{\mathbf{A}} \tag{2.46}$$

In the presence of noise, this estimate will not conform to the invariance structure imposed by the array. An improved estimate for the subarray steering vector can be computed as

$$\mathbf{A} = \frac{1}{2} (\mathbf{E}_x \mathbf{E}_\Psi + \mathbf{E}_y \mathbf{E}_\Psi \Phi^{-1}) \tag{2.47}$$

(note that $\Phi^{-1} = \Phi^*$). In this way it is possible to associate the steering vectors with the estimated bearings. After estimating the steering vectors, it is possible to construct a linearly-constrained minimum-variance beamformer to recover the emitter waveforms.

High-resolution algorithms such as multidimensional-MUSIC, deterministic ML, and TLS-ESPRIT can be analyzed within the framework known as *subspace fitting* [157, 158], which produces insight to their asymptotic relations. It is also shown in [157] that, by introducing certain weighting factors (which are determined from the received data) for the subspace fitting problem, it is possible to achieve the *stochastic* Cramer-Rao bound. Furthermore, weighted subspace fitting leads to better detection algorithms [158]. By using subspace-fitting arguments it is also possible to extend the ESPRIT algorithm to use multiple invariances [140, 143].

# Chapter 3

# Cumulants

In this chapter, we first review definitions and properties of *cumulants* and *polyspectra*. Rather than presenting an exhaustive overview of work that has been done in this field, we refer the reader to [78, 82] for a tutorial introduction, and proceed by illustrating properties of cumulants by an application to speech processing. Finally, we define cumulants for complex random processes, in a form applicable to the signal model of Chapter 1. For more detailed information on cumulants, reader is referred to the books [14, 84, 107, 183].

## 3.1 Definitions and Properties

Let $\{x_1, x_2, \ldots, x_n\}$ be a collection of random variables and $\{v_1, v_2, \ldots, v_n\}$ be a collection of deterministic variables. We can stack these variables in vectors $\mathbf{x} = [x_1, x_2, \ldots, x_n]^T$ and $\mathbf{v} = [v_1, v_2, \ldots, v_n]^T$. Then, the $n$th-order cumulant of the random variables is defined as the coefficient of $(v_1, v_2, \ldots, v_n)$ in the Mac-Laurin series expansion of the cumulant-generating function

$$K_{\mathbf{x}}(\mathbf{v}) = \log\left( E\left\{ \exp\left[ j\, \mathbf{v}^T \mathbf{x} \right] \right\} \right) \tag{3.1}$$

An alternate approach that defines the $n$th-order cumulant in terms of a weighted sum of joint moments of orders up to $n$ is provided in [78].

For zero-mean random variables that we frequently encounter in applications, the second-, third-, and fourth-order cumulants are expressed, as

$$\begin{aligned}
\text{cum}(x_1, x_2) &= E\{x_1 x_2\} \\
\text{cum}(x_1, x_2, x_3) &= E\{x_1 x_2 x_3\} \\
\text{cum}(x_1, x_2, x_3, x_4) &= E\{x_1 x_2 x_3 x_4\} - E\{x_1 x_2\} E\{x_3 x_4\} - \\
&\quad E\{x_1 x_3\} E\{x_2 x_4\} - E\{x_1 x_4\} E\{x_2 x_3\}
\end{aligned} \tag{3.2}$$

The expressions for non-zero mean random variables can be found in [85]. Throughout this thesis, we treat the zero-mean case.

There are several ways of collecting these random variables, e.g., in array processing, samples of sensor outputs separated in time, or, in system identification, samples from a random process. In the system *identification* context, if $x(t)$ is a random process, stationary up to order $n$, then the $n$th-order *cumulant function* of $x(t)$, $C_{n,x}(\tau_1, \tau_2, \ldots, \tau_{n-1})$, is defined as the $n$th-order cumulant of the random variables $\{x(t), x(t+\tau_1), \ldots, x(t+\tau_{n-1})\}$, i.e.,

$$C_{n,x}(\tau_1, \tau_2, \ldots, \tau_{n-1}) = \text{cum}(x(t), x(t+\tau_1), \ldots, x(t+\tau_{n-1})) \tag{3.3}$$

Due to the stationarity assumption, the $n$th-order cumulant of the random process $x(t)$ has $(n-1)$ degrees of freedom $\{\tau_1, \tau_2, \ldots, \tau_{n-1}\}$. Since the $n$th-order cumulant can be expressed as a sum of joint moments of the random variables of orders up to $n$, its existence is established if all absolute moments of orders $m \leq n$ exist and are bounded.

Note that for zero-mean processes the second- and third-order cumulants are identical to covariance and third-moment, respectively. The third- and higher-order cumulants of Gaussian processes are identically zero. This fact can be used for detection and characterization of deviations from non-Gaussianity [48]. In the next section, we will provide an experiment on speech processing.

• The following properties of cumulants will be used frequently in this thesis [78]:

**[CP1]** If $\{\alpha_i\}_{i=1}^n$ are constants and $\{x_i\}_{i=1}^n$ are random variables, then

$$\text{cum}(\alpha_1 x_1, \alpha_2 x_2, \ldots, \alpha_n x_n) = (\prod_{i=1}^{n} \alpha_i)\ \text{cum}(x_1, x_2, \ldots, x_n) \tag{3.4}$$

**[CP2]** If a subset of random variables $\{x_i\}_{i=1}^n$ are independent of the rest, then

$$\text{cum}(x_1, x_2, \ldots, x_n) = 0 \tag{3.5}$$

**[CP3]** Cumulants are additive in their arguments,

$$\text{cum}(x_1 + y_1, x_2, \ldots, x_n) = \text{cum}(x_1, x_2, \ldots, x_n) + \text{cum}(y_1, x_2, \ldots, x_n) \tag{3.6}$$

**[CP4]** If the random variables $\{x_i\}_{i=1}^n$ are independent of the random variables $\{y_i\}_{i=1}^n$, then

$$\text{cum}(x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n) = \text{cum}(x_1, x_2, \ldots, x_n) + \text{cum}(y_1, y_2, \ldots, y_n) \tag{3.7}$$

**[CP5]** The permutation of the random variables does not change the value of the cumulant.

**[CP6]** Cumulants suppress Gaussian noise of arbitrary covariance, i.e., if $\{z_i\}_{i=1}^n$ are Gaussian random variables independent of $\{x_i\}_{i=1}^n$ and $n > 2$, we have

$$\text{cum}(x_1 + z_1, x_2 + z_2, \ldots, x_n + z_n) = \text{cum}(x_1, x_2, \ldots, x_n) \tag{3.8}$$

**[CP7]** If $\alpha_o$ is a constant, then

$$\text{cum}(\alpha_o + x_1, x_2, \ldots, x_n) = \text{cum}(x_1, x_2, \ldots, x_n) \tag{3.9}$$

The $n$th-order polyspectrum is defined as the $(n-1)$-dimensional Fourier transform of the $n$th-order cumulant function. For discrete-time signal processing applications, this can be expressed as

$$S_{n,x}(\omega_1,\ldots,\omega_{n-1}) = \sum_{\tau_1,\ldots,\tau_{n-1}=-\infty}^{\infty} C_{n,x}(\tau_1,\ldots,\tau_{n-1}) \exp\left[-j\left(\sum_{i=1}^{n-1}\omega_i\tau_i\right)\right] \qquad (3.10)$$

Following the existence condition for Fourier transform, absolute summability is a sufficient condition for existence of the $n$th-order polyspectrum. Since cumulants are symmetric in their arguments so are their Fourier transforms. Exploiting symmetries provides computational advantages in implementation.

Cumulants are blind to phase shifts and scale factors. This originates from their definition. This observation is used in [23, 112, 150] for pattern recognition and reconstruction purposes. Third-order cumulants are blind to processes that have a symmetric probability density function; consequently, fourth-order cumulants must be used in such environments. Cumulants of independent and identically distributed (i.i.d.) random processes are delta functions, i.e., if $u(t)$ is such a process, then

$$C_{n,u}(\tau_1,\tau_2,\ldots,\tau_{n-1}) = \gamma_{n,u}\,\delta_{\tau_1,\tau_2,\ldots,\tau_{n-1}} \qquad (3.11)$$

where $\gamma_{n,u}$ is the $n$th-order cumulant of a single time sample from $u(t)$. It is important to note that joint moments do not possess this property. Suppose $x(t)$ is the output of the linear time-invariant system with impulse response $h(t)$ and driven by i.i.d. input $u(t)$. By using cumulant properties, the cumulant function of $x(t)$ can be expressed as

$$C_{n,x}(\tau_1,\tau_2,\ldots,\tau_{n-1}) = \gamma_{n,u}\sum_{t=-\infty}^{\infty} h(t)h(t+\tau_1)\cdots h(t+\tau_{n-1}) \qquad (3.12)$$

The polyspectra of $x(t)$ can be computed through its definition as

$$S_{n,x}(\omega_1,\omega_2,\ldots,\omega_{n-1}) = \gamma_{n,u}\,H(\omega_1)H(\omega_2)\cdots H(\omega_{n-1})H(-\sum_{k=1}^{n-1}\omega_k) \qquad (3.13)$$

where $H(\omega)$ is the frequency response of the filter. If the impulse response is real, then $H(-\omega) = H^*(\omega)$.

Cumulants of order higher than two are blind to Gaussian noise and can reveal phase characteristics of the system under consideration. On the other hand, covariance-based approaches are blind to phase information and sensitive to Gaussian noise. The use of these properties, as initiated in [74], make higher-order-statistics promising candidates to previously unsolvable problems. Several system identification methods are described in the tutorial [78].

In applications, we do not have access to true cumulants; we estimate them from the received data. The presence of additive Gaussian noise does effect the quality of the estimates, due to finite sample averaging in the estimation procedure. In order to get satisfactory results, longer data lengths are required for higher-order processing. An analysis of the asymptotical behavior of estimates of polyspectra can be found in [13].

## 3.2 An Application to Speech Processing

In this section, we illustrate the basic properties of higher-order-statistics by an application to speech processing. The aim in this experiment is to propose a robust method for speech state detection and pitch period estimation. This will be accomplished by observing the behavior of a cumulant-based adaptive predictor which processes the speech signal. Higher-order-statistics is proposed for discimination of speech states. Comparing the energy of the original speech signal with that of the prediction-error residual yields the decision method.

### 3.2.1 Speech Production Model

The state of speech signal belongs to three categories: voiced, unvoiced and silence. Silent periods can be detected easily by monitoring zero crossing rate and energy of the received signals [102]. For this reason, we shall concentrate on voiced/unvoiced classification of speech.

Unvoiced sounds are generated by forming a constriction at some point in the vocal tract and forcing air through the constriction at a high velocity to produce turbulence. This creates a broad spectrum noise source to excite the vocal tract. The energy concentration is shifted to the high-frequency end of the spectrum for

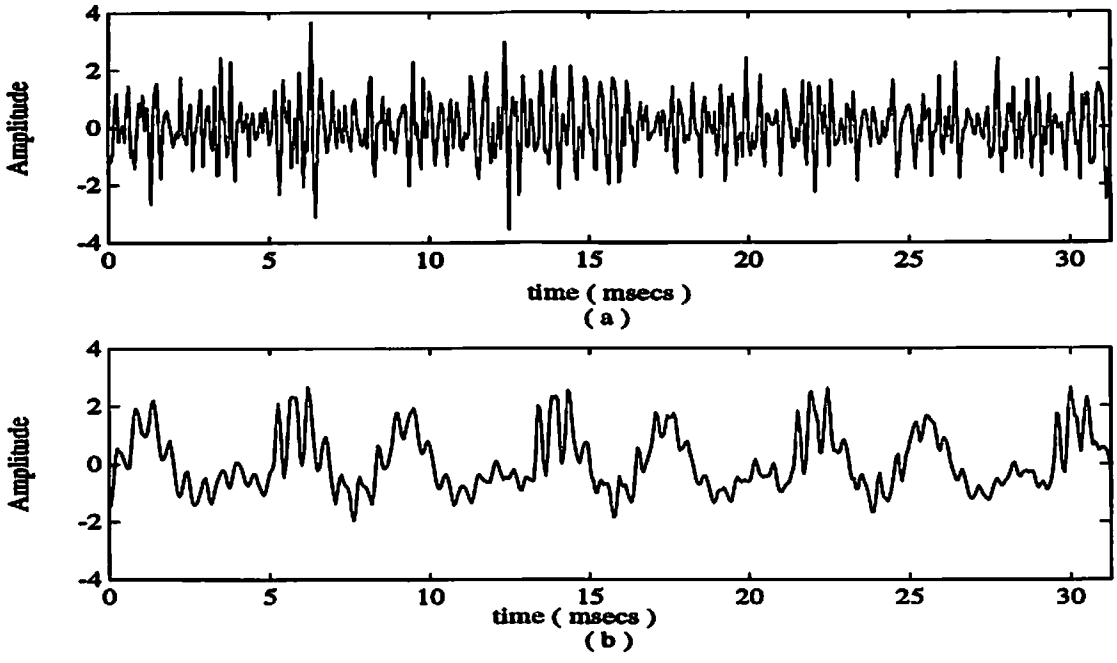



Figure 3.1: Typical speech signals: (a) Unvoiced speech, (b) Voiced speech.

unvoiced sounds, but the spectrum is relatively flat when compared with that of voiced speech. Due to large number of random effects involved in the production of unvoiced speech, Gaussian noise is a valid candidate as the excitation source. This assumption is validated by Wells [167]. In his work, the bispectrum is used to make V/UV decision. It has been found that bispectrum of English fricatives tend to zero, but for vowels the situation is just the opposite. A typical unvoiced segment of speech is shown in Fig. 3.1.a.

For voiced sounds, the vocal tract can be modelled as an all-pole linear system. The same model also holds for unvoiced sounds but the AR order is less. Correlation between adjacent samples is high for voiced sounds. On the other hand, unvoiced speech resembles white noise since its spectrum is relatively flat, yielding small correlation between adjacent samples.

The differences in the excitation and correlation properties for these two cases can be used to discriminate between them; however, with second-order statistics we can only use the correlation properties but can not utilize the information about the excitation model. This motivates the use of higher-order cumulants of speech signals.

### 3.2.2 Cumulants for Speech Analysis

In our work on speech processing we employ fourth-order cumulants since the third-order cumulants are blind to sources with symmetric probability density function. Although it is hard to propose an explicit density function for speech signals, a symmetric probability density function appears to be a reasonable one [102, pg.176].

Linear-prediction is a popular method to analyze speech signals. This method assumes an autoregressive (AR) model for speech production

$$s(n) = \sum_{k=1}^{p} a_k \, s(n-k) \; + u(n) \tag{3.14}$$

where $s(n)$ is the speech signal and $u(n)$ is the excitation sequence. In [78] it is shown that the AR parameters ( i.e., the $a_k$'s ) satisfy the following equations,

$$\sum_{k=0}^{p} a_k \, \mathrm{cum}(s(n), s(n+k_0), s(n+m-k), s(n), \ldots, s(n)) \; = \; 0 \tag{3.15}$$

where $m > 0$, $a_0 = 1$ and $k_0$ is a parameter whose selection is addressed in [78]. Concatenating (3.15), for $m = 1, 2, \ldots, p+M$, where $M \geq 0$, we obtain the so-called cumulant-based normal equations,

$$\mathbf{C}(k_0)\,\mathbf{a} = \mathbf{0} \tag{3.16}$$

where $\mathbf{C}(k_0)$ is a Toeplitz matrix and $\mathbf{a} = [1, a_1, a_2, \ldots, a_p]^T$. The AR parameters can be obtained by solving (3.16) if the matrix $\mathbf{C}(k_0)$ is full-rank. Note that, if the excitation sequence $u(n)$ is Gaussian and we employ higher than second-order cumulants, then the cumulant matrix $\mathbf{C}(k_0)$ becomes the all-zeros matrix, at least in theory. Due to finite-sample averaging, this condition will not hold but the identification problem becomes ill-posed since $\mathbf{C}(k_0)$ does not carry information about the speech production system. This observation is the starting point on the application of cumulants to speech processing.

### 3.2.3 Adaptive Filtering

In the previous section, we mentioned the distinctions between voiced and unvoiced sounds: correlation among adjacent samples and excitation models. In this section, we shall investigate methods that fully utilize this information.

Linear prediction (LP) methods are employed to accomplish our goal; however, we shall not use batch-type methods for reasons outlined in the introduction. Linear prediction can be based on second-or higher-order statistics, however the former is usually employed. Linear prediction is essentially identifying the inverse of a linear system driven by white noise; hence, it can be considered as a system identification problem. The system under consideration can be approximated by an AR model, so an FIR prediction filter will whiten the spectrum of the incoming signal. We shall investigate the differences between cumulant-and covariance-based adaptive prediction methods in this section.

#### Second-order statistics based adaptive filtering

*Correlation*-based adaptive prediction filters tend to minimize the prediction error power at the output of the filter. Since correlation among adjacent samples is high for voiced signals, we can remove a large proportion of energy from the original speech signal using prediction. On the other hand, in the case of unvoiced sounds, LP will not be that successful due to small correlation among samples. Therefore, a comparison of the input signal power with the power in the prediction residual may reveal the state of the speech signal.

Lattice prediction filters enable monitoring the variation of prediction error power with model order due to their specific structure. Autoregressive model-order-selection can be performed by selecting the tap which results in minimum prediction-error power. This leads to another discrimination between voiced and unvoiced sounds, since this order will be relatively lower for the unvoiced case.

#### Fourth-order statistics based adaptive filtering

We now investigate the behavior of a fourth-order cumulant-based adaptive filter. An adaptive algorithm for estimating the parameters of nonstationary AR processes,

excited by non-Gaussian signals is proposed in [136], and some modifications are suggested in [40]. The ideas for the covariance-based filter directly apply to this case with one important *exception*: the cumulant-based adaptive filter provides the solution to the cumulant-based normal equations, and this solution is not *the* one that minimizes the prediction-error power; however, one may argue that if the speech production system can be identified accurately, then the prediction error should be close to the minimum possible value.

With higher-order statistics, we have the diversity of using the excitation information: for voiced sounds, the excitation is non-Gaussian; hence, the speech production mechanism can be identified by cumulant-based AR equations. On the other hand, for unvoiced sounds the *excitation* can be characterized *as Gaussian*, making the identification problem *ill-posed*[1]. The cumulant-based adaptive filter *will not be able to identify the system* and, since there is no associated output-power minimization criterion, prediction-error power may *arbitrarily increase*. In this case, a cumulant-based filter may even amplify the speech signal making the power reduction by prediction comparison more clear than when using a covariance-based method.

To validate our ideas about covariance and cumulant-based adaptive prediction of speech signals, we summarize them in Figure 3.2. We also performed experiments using data from the TIMIT speech recognition database. The results verify our claims and are provided in the next section.

### 3.2.4 Experiment

To analyze the behavior of adaptive predictors in voiced and unvoiced speech states, we selected a 250 msec period of speech segment in which there are two transitions: voiced (0-75 msec), unvoiced (75-190 msec) and again voiced (190-250 msec). We used an order ten predictor for adaptive filtering of the speech waveform.

To make better comparisons concerning the energy of the original speech and prediction residuals, obtained via the two different filters, we illustrate the energy

[1] A cumulant-based filter provides the solution of cumulant-based normal equations in an adaptive fashion; however, this set of equations becomes trivial when the input to be analyzed is a Gaussian linear process, because higher than second-order cumulants of Gaussian processes are zero.

| | COVARIANCE | CUMULANT |
|---|---|---|
| VOICED | Mimimizes output power by whitening the spectrum.<br><br>Problems arise due to quasi-periodic excitation. | Solves cumulant-based normal equations.<br><br>No power minimization criterion. |
| UNVOICED | Shuts itself off due to small adjacent sample cross-correlation.<br><br>Output power is equal to input signal power. | System identification task is ill conditioned.<br><br>No output power minimization.<br><br>Filter runs uncontrolled.<br><br>Output power is larger than input power. |

Figure 3.2: Summary of the properties of adaptive predictors for different speech states.

estimates in Fig. 3.3. Energy is estimated by first squaring the signal and then performing low-pass filtering using a 15 point Hamming window. Figure 3.3 shows that, by comparing the prediction-residual power and the original-signal power, it is possible to make reliable V/UV decisions. With the cumulant-based method, even better results are obtained, because it amplifies the input data during unvoiced periods.

### 3.2.5 Remarks

In this section, we have described a speech analysis system which exploited the speech model and the properties of cumulants together. The lesson from this experiment is: in applying cumulants to a problem, first develop an insight to the conditions of the problem and then provide a clear interpretation.

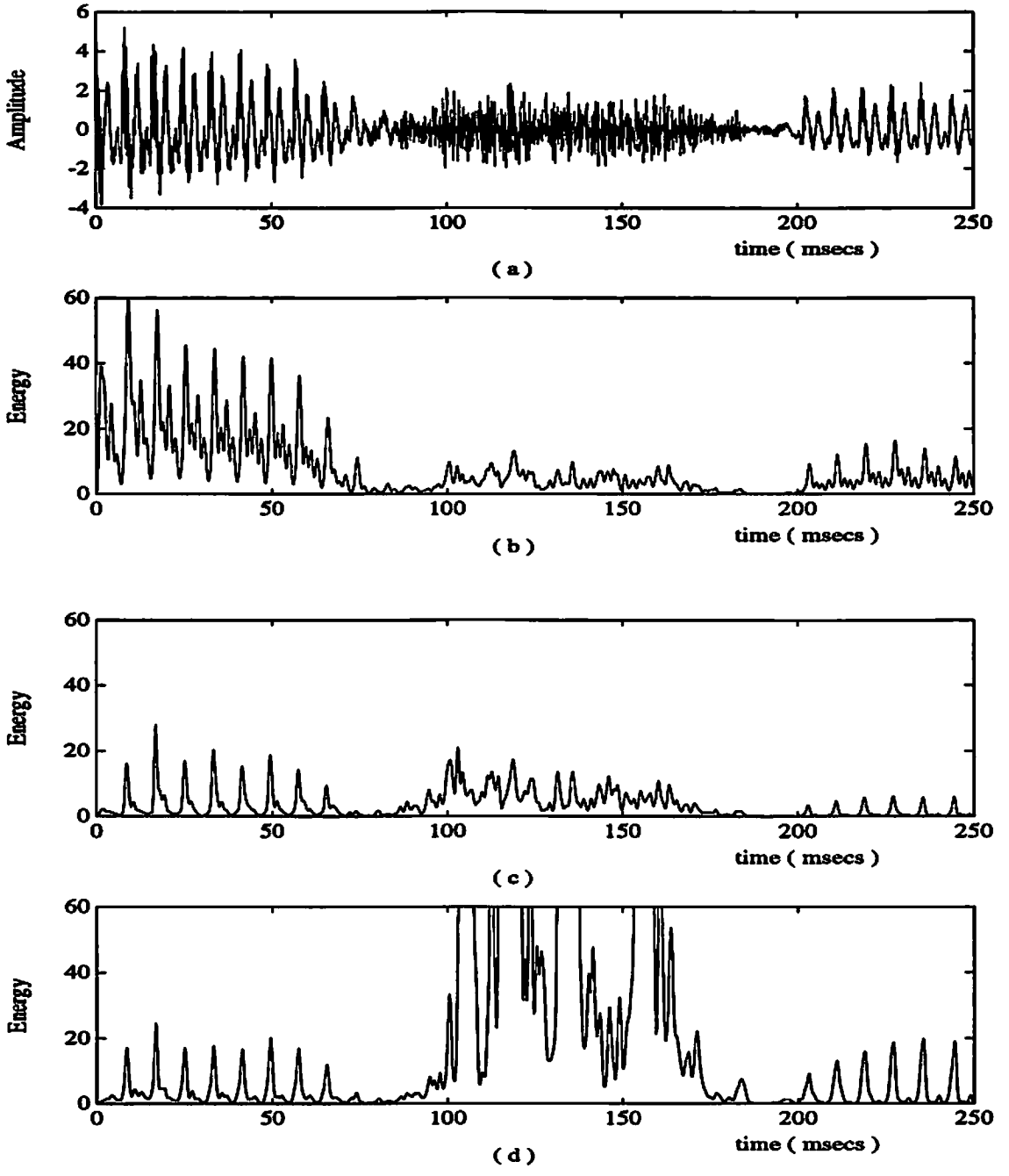



Figure 3.3: Energy estimates. (a) Original speech signal; (b) energy estimate of original speech signal; (c) energy estimate of prediction-error residual from covariance-based filter, (d) energy estimate of prediction-error residual from cumulant-based filter.

## 3.3 Cumulants of Complex Random Processes

In this thesis, we shall mainly consider direction-finding for narrowband signals. Hence, our signals of interest are non-Gaussian (possibly complex) narrowband random processes, e.g., communication signals. Measurement noise components, being the outcome of many unknown, independent factors can be assumed to be Gaussian random processes with unknown covariance. Not knowing the noise covariance reflects our lack of knowledge about the spatial propagation of noise. Spatial correlation of noise creates problems in covariance-based processing (e.g., biased direction-of-arrival (DOA) estimates), but with HOS this will not be an issue since higher-than second-order cumulants are blind to additive Gaussian noise. However, Gaussian noise suppression is only one benefit of using cumulants and it is not the major issue that we address in our work.

To handle symmetric probability density functions from the sources of interest (e.g., communication signals), we must use fourth-order cumulants of sensor outputs. Fourth-order (zero-lag) cumulants are defined in a *balanced* way as follows [137]:

$$\begin{aligned}\text{cum}\{r_i(t), r_j(t), r_k^*(t), r_l^*(t)\} \triangleq\ & E\{r_i(t)r_j(t)r_k^*(t)r_l^*(t)\} - E\{r_i(t)r_j(t)\}\, E\{r_k^*(t)r_l^*(t)\} \\ & - E\{r_i(t)r_k^*(t)\}\, E\{r_j(t)r_l^*(t)\} \ -\ E\{r_i(t)r_l^*(t)\}\, E\{r_j(t)r_k^*(t)\} \qquad (3.17) \\ & \text{for } (i,j,k,l) \in \{1,\ldots,M\}\end{aligned}$$

in which $\{r_k(t)\}_{k=1}^M$ denotes the received signal vector from an array of $M$ sensors. This definition is in-keeping with the definition of cross-covariance, which can be expressed as, $E\{r_i(t)r_j^*(t)\}$, and which has only two arguments. The *richness* of cumulants in terms of arguments will prove to be an interesting feature.

If the signal of interest also possesses a non-zero third-order cumulant, then we may also consider third-order cumulants. However, third-order cumulants can not be defined in a balanced way

$$\text{cum}\{r_i(t), r_j(t), r_k^*(t)\} \triangleq E\{r_i(t)r_j(t)r_k^*(t)\} \qquad (3.18)$$

since the number of arguments is odd.

# Chapter 4

# An Interpretation of Cumulants for Array Processing

In this chapter, we propose a novel interpretation for the use of fourth-order cumulants in narrowband array processing problems. We show that fourth-order cumulants of multichannel observations double the directional information available compared with second-order statistics. Based on our interpretation, we explain how cumulants can increase the effective aperture of antenna arrays. We give several examples to illustrate our interpretation. In addition, we extend our results to third order cumulants. Finally, we describe the role of our interpretation in the following chapters of the thesis.

## 4.1 Fourth-Order Cumulants

Conventional array processing techniques utilize only the second-order statistics of received signals. These methods require the computation of the sample covariance matrix as described in Chapter 2. Second-order statistics are sufficient whenever the signals can be completely characterized by knowledge of the first two moments, as in the Gaussian case; however, in real applications, far-field sources emit non-Gaussian signals, e.g., as in a communications scenario.

Whenever second-order statistics can not completely characterize all of the statistical properties of underlying signals, it is beneficial to consider information embedded in higher than second-order moments. Higher-order statistics (HOS) prove

to be rewarding alternatives to second-order statistics; there are many signal processing problems that are not solvable without access to HOS [78]. In this section, we provide an interpretation for the use of fourth-order cumulants in array processing problems based on the properties of higher-order statistics summarized in Chapter 3.

As we summarized earlier in Chapter 2, high-resolution direction-finding methods such as MUSIC and ESPRIT, which use second-order statistics of array measurements, have been developed for the model

$$\mathbf{r}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t) \tag{4.1}$$

where $\mathbf{A}$ is the full-rank steering matrix, $\mathbf{s}(t)$ denotes the source waveforms and $\mathbf{n}(t)$ is the noise contribution. If there are $M$ sensors and $P$ sources, then $\mathbf{r}(t)$ and $\mathbf{n}(t)$ are $M$-vectors, $\mathbf{s}(t)$ is a $P$-vector and $\mathbf{A}$ is an $MxP$ matrix. If the noise is spatially-white, then the covariance matrix of $\mathbf{r}(t)$ takes the form

$$\mathbf{R} \triangleq E\{\mathbf{r}(t)\mathbf{r}^H(t)\} = \mathbf{A}\,\mathbf{\Sigma}_{ss}\,\mathbf{A}^H + \sigma^2\mathbf{I} \tag{4.2}$$

where $\mathbf{\Sigma}_{ss} \triangleq E\{\mathbf{s}(t)\mathbf{s}^H(t)\}$. If there are $P$ sources, and they are all incoherent, then the source covariance matrix $\mathbf{R}$ is diagonal, and it can be reexpressed as

$$\mathbf{R} = \mathbf{A}\,\mathbf{\Sigma}_{ss}\,\mathbf{A}^H + \sigma^2\mathbf{I} = \sum_{k=1}^{P} \sigma_k^2 \mathbf{a}_k \mathbf{a}_k^H + \sigma^2\mathbf{I} \tag{4.3}$$

where $\sigma_k^2$ and $\mathbf{a}_k$ denote the power and the steering vector for the $k$th source respectively. If at another data collection time, the power of the $k$th source is scaled by $\beta_k$, then the array covariance matrix $\tilde{\mathbf{R}}$ for this scenario takes the form

$$\tilde{\mathbf{R}} = \sum_{k=1}^{P} \beta_k \sigma_k^2 \mathbf{a}_k \mathbf{a}_k^H + \sigma^2\mathbf{I} \tag{4.4}$$

High-resolution direction-finding methods use the structure of (4.3) to eliminate the noise component ($\sigma^2\mathbf{I}$), and then search for the vectors in the array manifold that lie in the range space of $\mathbf{A}\,\mathbf{\Sigma}_{ss}\,\mathbf{A}^H$. If the noise is spatially-colored, elimination of noise in the array covariance matrix is not possible unless one knows the noise covariance

matrix and whitens the received signals. If the noise is Gaussian, then cumulants can be used to suppress its effects.

In order to provide an interpretation for the use of second and higher-order statistics in array processing applications, we illustrate an array set-up in Figure 4.1. For convenience, the elements of the array (represented by circles) are assumed to be isotropic, and the narrowband sources that illuminate the array are assumed to be statistically independent. In this case, we can further assume the presence of a single source $s(t)$ (without loss of generality due to **[CP4]**), with propagation vector $\vec{k}$, where $\vec{k} = k_x\hat{a}_x + k_y\hat{a}_y$ ($\hat{a}_x$ and $\hat{a}_y$ denote the unit vectors along the x and y-axis, respectively), power $\sigma_s^2$ and fourth-order cumulant $\gamma_{4,s}$.

Suppose we compute the cross-correlation (ignoring noise effects) between real signal $r(t)$ and virtual signal $v(t)$, $E\{r^*(t)v(t)\}$, assuming the reference point to be the position of the sensor that records $r(t)$. We call such a cross-correlation a "virtual" cross-correlation. Because $r(t) = s(t)$, and $v(t) = s(t)\exp(-j\vec{k}\cdot\vec{d})$, it follows that the directional information provided by the correlation operation is embedded in the dot product, $\vec{k}\cdot\vec{d}$, i.e.,

$$E\{r^*(t)v(t)\} = \sigma_s^2 \exp(-j\ \vec{k}\cdot\vec{d}) \tag{4.5}$$

The source power $\sigma_s^2$ does not provide any directional information; hence, the information recovered by cross-correlation of two sensor outputs can be represented by the *vector* extending *from* the conjugated sensor, *to* the unconjugated one, i.e., cross-correlation is a *vector* in the geometrical sense. After giving such an interpretation for cross-correlation, the problem is how to interpret fourth-order cross-cumulants that have four-arguments. Consider the cumulant,

$$\text{cum}(r^*(t), x(t), r^*(t), y(t)) = \text{cum}(s^*(t), s(t)\exp(-j\vec{k}\cdot\vec{d}_x), s^*(t), s(t)\exp(-j\vec{k}\cdot\vec{d}_y)) \tag{4.6}$$

and use **[CP1]** to obtain:

$$\text{cum}(r^*(t), x(t), r^*(t), y(t)) = \gamma_{4,s}\exp(-j\vec{k}\cdot\vec{d}_x)\exp(-j\vec{k}\cdot\vec{d}_y) \tag{4.7}$$

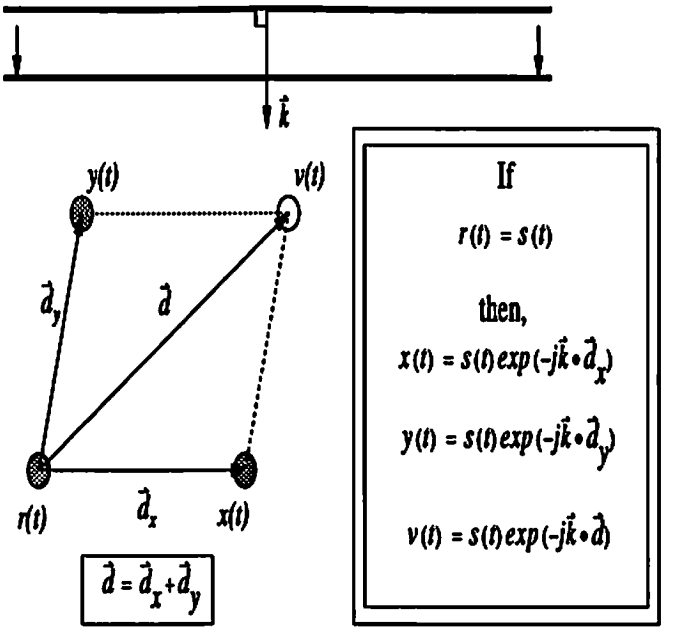



Figure 4.1: Mechanism of second-and higher-order cumulants for array processing: the sensors that measure $r(t), x(t), y(t)$ are actual sensors, whereas $v(t)$ is a virtual-process measured by a virtual-sensor.

Finally, by comparing (4.5) and (4.7), we observe the following:

$$E\{r^*(t)v(t)\} = \frac{\sigma_s^2}{\gamma_{4,s}}\text{cum}(r^*(t), x(t), r^*(t), y(t)) \tag{4.8}$$

Equation (4.8), which relates a fourth-order statistic to a second-order statistic, demonstrates that it is possible to *recover* directional information which is provided by channels $r(t)$ and $v(t)$ without actually using a real sensor to measure $v(t)$. We refer to (4.8) as a "virtual cross-correlation computation" ($\mathbf{VC^3}$)[1].

From the development of (4.8), we see fourth-order cumulants can be interpreted as *addition of two vectors* each extending *from* a conjugated channel *to* an unconjugated one. It can be shown that (4.8) holds for both: multiple independent sources (due to **[CP4]**) and in the presence of additive colored Gaussian noise (due to **[CP6]**). Furthermore, **[CP5]** which indicates that cumulants are invariant with respect to a permutation of the random variables is a restatement of the fact that addition of two vectors is a *commutative* operation.

[1]A patent for the $\mathbf{VC^3}$ has been filed for by the University of Southern California.

Our geometrical interpretation gives an idea why cumulants have a great potential for array processing: cumulants *virtually* increase the effective aperture of an array without any design procedure or configuration constraints (as in minimum redundancy arrays, which are constrained to be linear). To investigate the *virtual aperture extension* by using fourth-order cumulants we redraw the three element array of Figure 4.1 and indicate the *lattice* structure defined by the vectors $\vec{d}_x, \vec{d}_y, \vec{d}_z$ in Figure 4.2. We only have 3 sensors available: the ones labeled as $r(t), x(t)$ and $y(t)$. The intersections of the lines in the lattice determine the candidate locations for virtual-sensors. To implement a covariance-like subspace algorithm, we need to compute the cross-correlation of all sensor outputs, actual or virtual. In other words, we need to connect the sensors to be used, with a single vector. With fourth-order cumulants, we have the liberty of using two vectors for connection purposes. These connecting vectors must be selected from the set of vectors that define the lattice. In Figure 4.2 we have indicated a group of sensors (not a unique selection) that consist of four virtual and three actual sensors which can *communicate* by two *jumps* (vector additions). It is possible to form a 7x7 matrix in which we use cumulants to compute the required covariances (see Figure 4.3). Observe, from (4.8), that the computed cumulants are scaled versions of the cross-correlations. This is the case covered by (4.4), where $\beta_k = \gamma_{4,k}/\sigma_k^2$. The resulting matrix $\tilde{\mathbf{R}}$, will take the form of the covariance matrix of array measurements, in which source powers are changed by an unknown factor $\beta_k = \gamma_{4,k}/\sigma_k^2$, that results in

$$\tilde{\mathbf{R}} = \sum_{k=1}^{P} \gamma_{4,k}\, \mathbf{a}_k \mathbf{a}_k^H = \mathbf{A}\boldsymbol{\Gamma}\mathbf{A}^H \tag{4.9}$$

Note also that, because $\tilde{\mathbf{R}}$ is computed using cumulants, $\sigma^2 = 0$ (cumulants suppress additive Gaussian noise). In (4.9), the steering vectors ($\mathbf{a}_k$) are 7x1, i.e., they fully represent the delays encountered by the wavefronts as if we have 7 actual elements at the locations selected in Figure 4.2. The diagonal matrix $\boldsymbol{\Gamma}$ consists of the fourth-order cumulants (which can be negative) of sources. Therefore, the cumulant matrix $\tilde{\mathbf{R}}$ is not necessarily positive-definite, unlike the covariance matrix of sensor measurements. This difference does not pose a problem in the direction-finding processor, when the eigenvalues are sorted with respect to their magnitude.

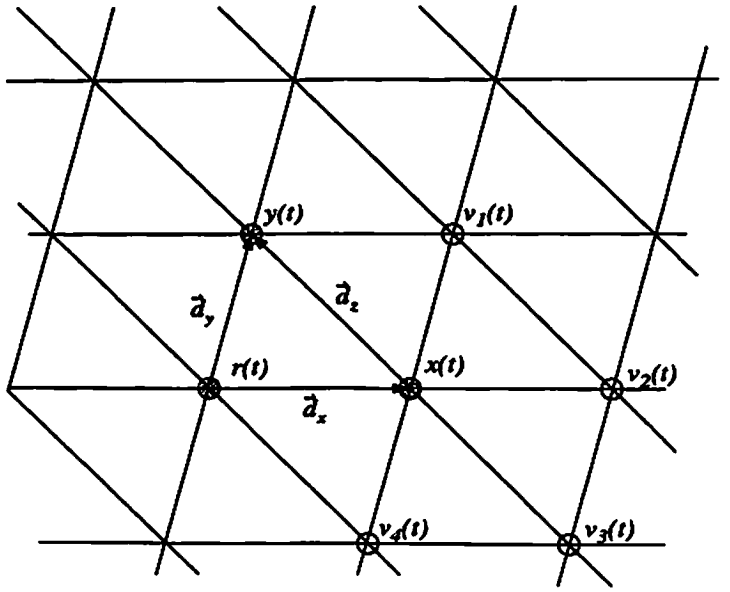



Figure 4.2: Virtual aperture extension by cumulants: lattice indicates the possible locations for virtual sensors. To be used in a DOA algorithm all sensors must be connected by at most two vectors. One such group (not unique) is illustrated in the figure.

For the example in Figure 4.2, a covariance-based algorithm can estimate the bearings of two sources, whereas the cumulant-based *virtual aperture extension* approach can estimate the parameters of six sources (one less than the number of elements that form the *effective* aperture). In addition, the cumulant based approach can *survive* in the presence of colored noise due to [CP6]. In the next section, we provide more examples with different array configurations.

Note that it is possible to compute cross-correlations among the actual sensors by cumulants, because cross-correlation is a "*vector*", and any vector can be expressed as the addition of the zero vector to itself. In other words, cross-correlation between two channels can be computed by using two of the arguments of the fourth-order cumulant as required by correlation, and then using the remaining two arguments by repeating one of the channels twice, e.g.,

$$E\{x^*(t)y(t)\} = \frac{\sigma_s^2}{\gamma_{4,s}}\text{cum}(x^*(t), y(t), \underbrace{x^*(t), x(t)}_{\text{repeat}}) \tag{4.10}$$

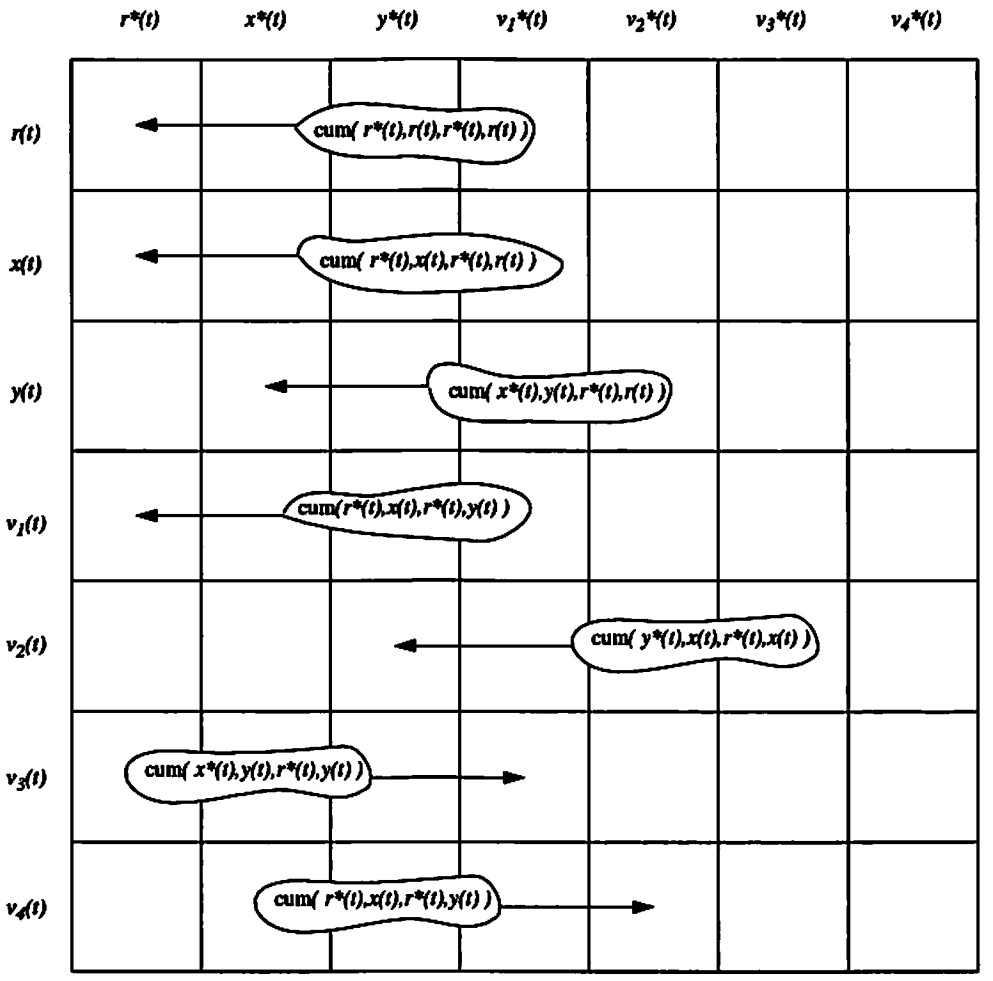



Figure 4.3: Construction of a matrix ($\tilde{\mathbf{R}}$) to be used in the direction-finding processor using cumulants.

The advantage of computing $E\{x^*(t)y(t)\}$ by (4.10) is that additive Gaussian noise can be suppressed by the cumulant calculation. If, $E\{x^*(t)y(t)\}$ were computed directly, it would be severely affected by additive Gaussian noise. Equation (4.10) is the approach taken in the initial attempts to incorporate higher-order statistics into direction-finding algorithms for Gaussian noise suppression purposes [24, 41, 83, 91]. In this section, we showed that cumulants have more important properties than just Gaussian noise suppression. Chapter 7 addresses the issue of non-Gaussian noise suppression.

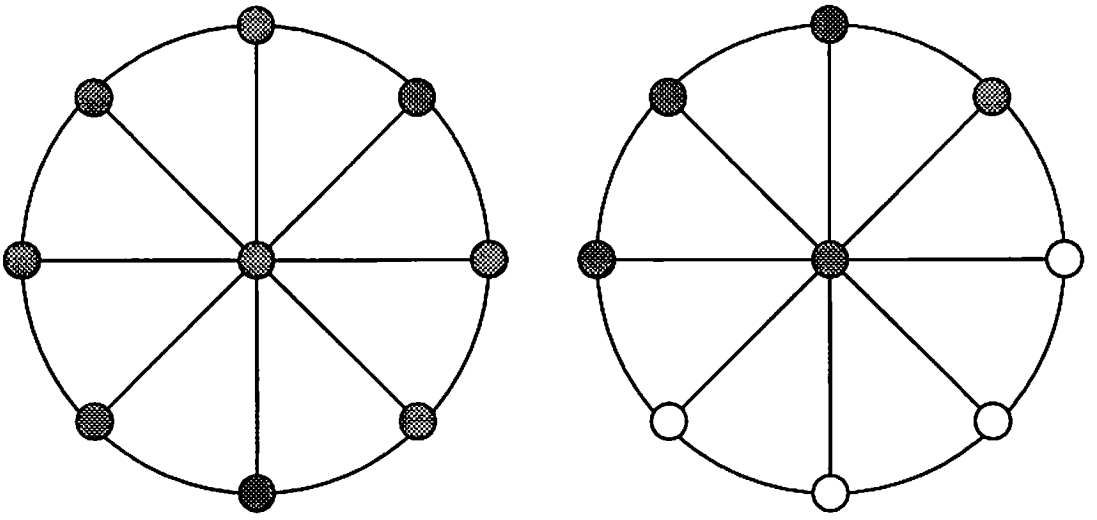



Figure 4.4: Circular arrays and element reduction by cumulants: (a) circular array for linear-prediction direction-finding, (b) the empty sensors become redundant with the use of cumulants.

## 4.2 Examples of Aperture Extension

We illustrate three different array geometries: circular, linear, and rectangular.

• Circular Array

In Figure 4.4a, we illustrate a circular array with a sensor at the center. Such a configuration is very suitable for linear-prediction direction-finding when the sensor at the center is used as the reference [46]. However, some sensors are redundant if we use cumulants; the empty sensors in Figure 4.4b indicate the virtual elements whose second-order statistics can be computed by higher-order statistics.

A more careful investigation (see Figure 4.5) of the circular array reveals that with the actual sensors of Figure 4.4b, it is possible to extend the aperture so that the effective aperture is even larger than that of the original array in Figure 4.4a. This result indicates that cumulants are very promising in the design of two-dimensional *minimum-redundancy* arrays. Chapter 6 investigates this issue.

• Linear Array

Consider the fully-redundant linear array of $N$ isotropic sensors in Figure 4.6a. In [122] it is proved that the aperture can be extended to $2N - 1$ elements. We now provide a very simple and geometric proof for this fact: the most distant point

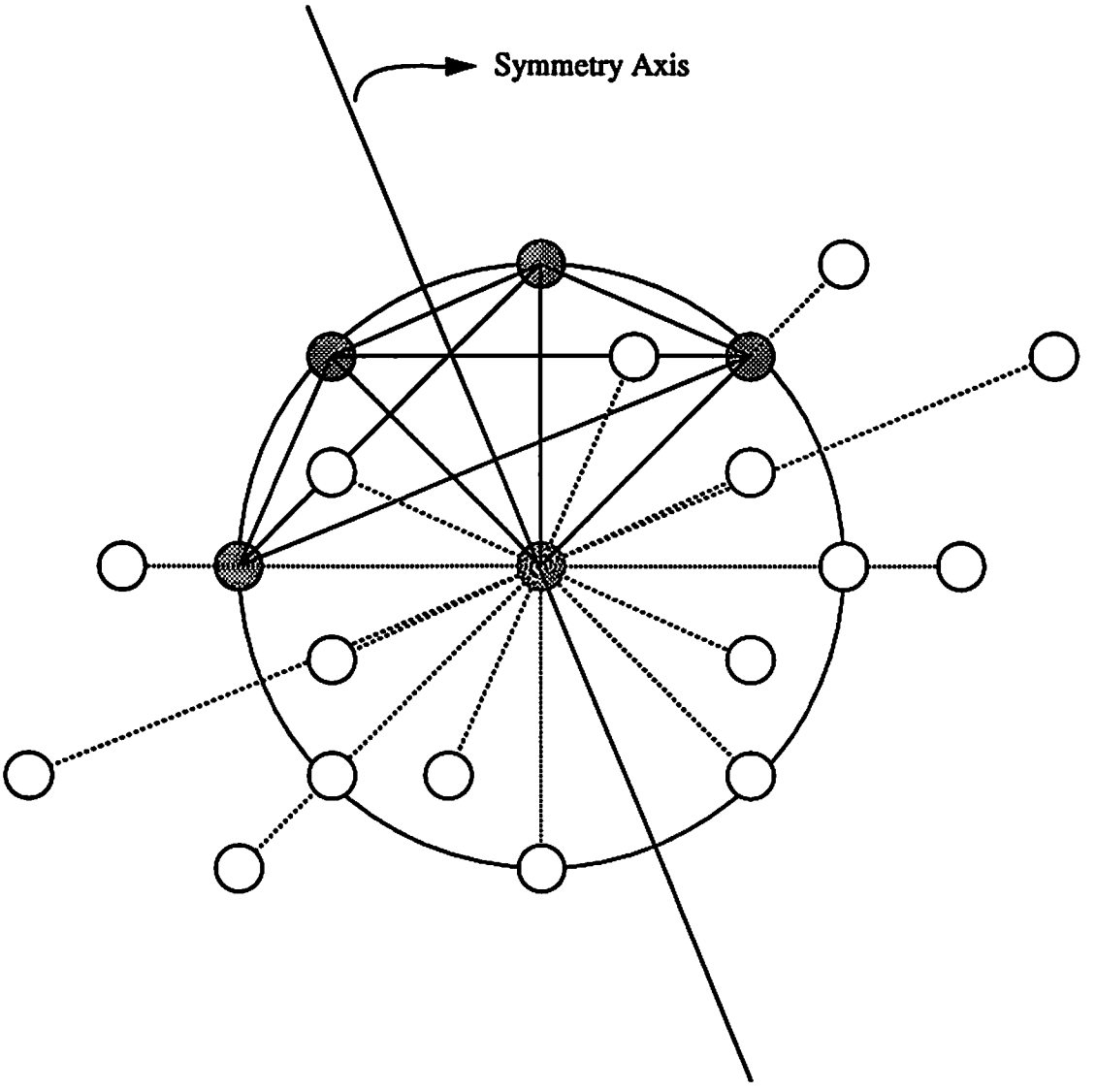



Figure 4.5: It is possible to generate 12 more virtual sensors by using cumulants. These new locations are obtained by picking one of the 6 possible vectors among the 4 actual sensors (4!/(2! 2!)=6) on the circle (excluding the sensor at the center), and attaching this vector to the sensor at the center. There are 2 ways to do this, hence we have 12 new locations. Since these 12 new locations *communicate* with the center sensor with only one vector, they communicate with the sensors of previous design (virtual sensors on the circle) with at most two jumps, implying that the aperture can be extended to 21 elements using cumulants.

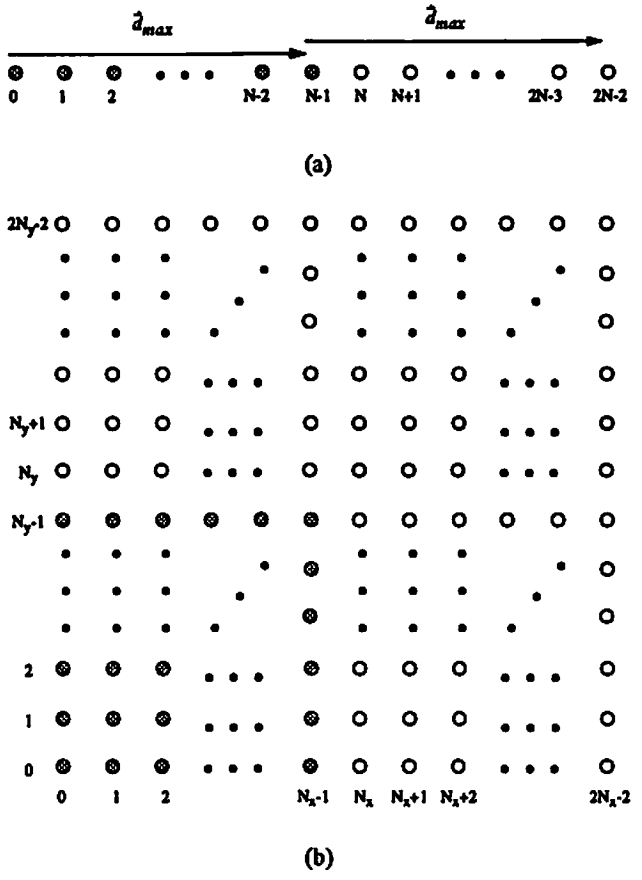



Figure 4.6: Linear and rectangular array aperture extension: (a) fully-redundant linear array ($N_x$ filled sensors), and virtual-sensors that can be generated by cumulants ($N_x - 1$ empty sensors), (b) fully-redundant rectangular array ($N_x \cdot N_y$ filled sensors), and virtual sensors that can be generated by cumulants. Effective aperture consists of $(2N_x - 1)(2N_y - 1)$ sensors.

from the origin that we can reach by two vector additions is $(2N - 2)$, (since the maximum length vector among the available sensors is $(N - 1)$ units long), i.e., the effective aperture consists of $2N - 1$ sensors! The proof is this simple because of the interpretation developed in this chapter. In addition, design of linear non-redundant arrays can be formulated based on our geometric interpretation. In Chapter 7, we provide a simulation, in which virtual aperture extension is accomplished for a two element array in the presence of spatially-white but spatially-nonstationary non-Gaussian noise.

• Rectangular Array

We can now analyze the more general case of the fully-redundant rectangular array

in Figure 4.6b. We have an array of $N_x \cdot N_y$ sensors. With similar reasoning to the linear case, the effective aperture consists of $(2N_x - 1)(2N_y - 1)$ sensors.

## 4.3 Third-Order Cumulants

Third-order cumulants are used less frequently in array processing applications than fourth-order cumulants [35, 41, 92, 139]. This is due to the non-symmetric nature of the definition of the third-order moments for complex random processes. Another important reason is that third-order cumulants of almost all man-made signals are identically zero (communication signals, telemetry signals); however, there are physical processes that have non-zero third-order statistics, such as signals associated with rotating machinery, radar returns, and voiced speech signals. For localization of these processes, it is advantageous to use third-order cumulants, since they are more computationally attractive than fourth-order cumulants, their data length requirements are less when compared with fourth-order statistics, and they can suppress symmetrically distributed ambient noise.

An important observation about third-order moments is that they are *not spatially stationary* even in the case of independent sources. This is a result of the asymmetry in the third-order moments with respect to conjugation. For example, refer to Figure 4.1, and consider

$$\text{cum}(r(t), r^*(t), r(t)) = \gamma_{3,s} \neq \text{cum}(x(t), x^*(t), x(t)) = \gamma_{3,s} \exp(-j\vec{k} \cdot \vec{d}_x) \quad (4.11)$$

This indicates that third-order cumulants have a *reference point problem*, unlike second and fourth-order statistics. They are floating in this sense. However, this can be cured very easily by allowing the third-order cumulant $\gamma_{3,s}$ to be an unknown complex quantity so as to include the multiplication by an unknown phase term due to the floating reference point. This is possible since the statistic of the source waveform does not provide information about the location of the source as pointed earlier.

Aperture extension (for identical sensors) starts by computing the autocorrelation as

$$E\{r(t)r^*(t)\} = E\{x(t)x^*(t)\} = E\{y(t)y^*(t)\} = \frac{\sigma_s^2}{\gamma_{3,s}} \text{cum}(r(t), r^*(t), r(t)) \quad (4.12)$$

Next, we must compute the cross-correlation between actual sensors, which can be done as follows (direct extension of the fourth-order cumulant results)

$$E\{r^*(t)x(t)\} = \frac{\sigma_s^2}{\gamma_{3,s}} \text{cum}(r^*(t), x(t), r(t)) \quad (4.13)$$

$$E\{r^*(t)y(t)\} = \frac{\sigma_s^2}{\gamma_{3,s}} \text{cum}(r^*(t), y(t), r(t)) \quad (4.14)$$

$$E\{x^*(t)y(t)\} = \frac{\sigma_s^2}{\gamma_{3,s}} \text{cum}(x^*(t), y(t), r(t)) \quad (4.15)$$

Observe that we always fix the last component of the cumulant in (4.13)-(4.15) to be $r(t)$. We call $r(t)$ the "common element". Now we shall complete the description of virtual aperture extension by demonstrating the way to compute the cross-correlations between real elements and and the lone virtual sensor (since we have only one virtual location, we can not calculate cross-correlations between two virtual sensor measurements), i.e.,

$$E\{r^*(t)v(t)\} = \frac{\sigma_s^2}{\gamma_{3,s}} \text{cum}(r^*(t), x(t), y(t)) \quad (4.16)$$

Equation (4.16) implies that for aperture extension purposes, the third-order cumulant can be considered as *addition* of *two vectors*, which are constrained to originate from the same sensor (the conjugated one). Therefore aperture extension is possible with third-order statistics, but it is not as powerful as using fourth-order cumulants. For example, in Figure 4.2, we can take the set $\{r(t), x(t), y(t), v_1(t)\}$ as an extended aperture with third-order cumulants, whereas with fourth-order cumulants the aperture can be extended to seven elements (see Figure 4.2). A third-order 4x4 cumulant matrix can be set-up as described in the fourth-order cumulants case (Figure 4.3).

In the circular array example (Figure 4.4), using third-order cumulants, it is possible to extend the aperture to nine elements (show as an exercise) as in Figure 4.4b;

however, this is less than the result obtained by using fourth-order cumulants (see Figure 4.5), since fourth-order cumulants can be viewed as unconstrained vector addition.

## 4.4 Conclusions

We have proposed a novel interpretation for describing the potential of cumulants in array processing applications. It has led to the virtual cross-correlation computer[2] ($VC^3$). The $VC^3$ can be used to: calibrate arbitrary arrays and jointly estimate the direction parameters of far-field sources, design minimum-redundancy arrays, and suppress undesired signals. Figure 4.7 summarizes these applications and their treatment in this thesis.

The most surprising result of this chapter is the explanation of how cumulants can increase the *effective aperture* of an arbitrary array. This fact is reasonable when one considers that forming the covariance matrix is a *data reduction* technique (rather than storing multichannel snapshots, we compute the sample covariance matrix). It is a well known fact of information theory that we always lose information by data reduction. Forming a cumulant matrix can be considered as an alternate data reduction technique, but it is much better than using only second-order statistics, since it is possible to recover more information from the cumulant matrix about the sources illuminating the array. The interpretation provided for cumulants can be extended for cyclostationarity-based array processing applications [116, 121, 174].

[2]A patent for the $VC^3$ has been filed for by the University of Southern California.

| APPLICATION | PROBLEM | SOLUTION | DESCRIPTION |
|---|---|---|---|
| Aperture Extension | To extend the effective aperture of an array without adding real sensors. | Compute all the required second-order statistics virtually, by fourth-order cumulants. | Chapter 4, 6. |
| Array Calibration | An identical copy of the array at a known displacement is required to compute cross-correlations. This requires extra hardware and precision. | Compute the required cross-correlations virtually by using a single doublet and fourth-order cumulants. | Chapter 5. |
| Antenna Array Design | Minimum redundancy array design is limited to linear arrays. Linear arrays have ambiguity problems. | Use the interpretation of cumulants to design 2-D minimum redundancy arrays. | Chapter 6. |
| Noise and Interference Suppression | To decorrelate the undesired components, another array which is far-away from the existing one is required. This means doubling the hardware. | A single sensor which is far-away from the existing array is sufficient. Compute the required statistics virtually, by using fourth-order cumulants. | Chapter 7. |

Figure 4.7: An overview of the possible applications of our interpretation.

# Chapter 5

# Array Calibration Issues: Virtual-ESPRIT Algorithm

We proposed a novel interpretation for the use of cumulants in narrowband array processing problems in Chapter 4. Based on this interpretation, we investigate the *amount* of partial information necessary to *jointly* calibrate an arbitrary array and estimate the directions of far-field sources in this chapter. We prove that the presence of a doublet and use of fourth-order cumulants is sufficient to accomplish this task. Our approach is computationally efficient and more general than covariance-based algorithms that have addressed this problem under constraints. A class of beamforming techniques is proposed to recover the source waveforms. All of the developed estimation procedures are based on cumulants, which bring insensitivity to the spatial correlation structure of additive Gaussian measurement noise. Simulations indicate excellent results by the proposed algorithms.

## 5.1 The Array Calibration Problem

During the last decade, there have been revolutionary advances in high-resolution direction-of-arrival estimation problems. Some of these algorithms were described in Chapter 2.

Among the algorithms proposed in the literature, the so-called subspace methods that are based on the eigendecomposition of the sample covariance matrix possess very appealing features: they have modest computational requirements when compared with the maximum-likelihood method [180]; there exists solid work on the

detection of sources problem [162]; and they provide asymptotically exact values for the parameters of interest.

The MUSIC algorithm [118] is most popular, due to its applicability to arrays of arbitrary orientation and response. In addition, it can estimate a multitude of parameters for each far-field source, such as azimuth, elevation angles, and polarization. This generality results in a major drawback: array response must be measured and stored for every possible combination of source parameters. This procedure, known as array calibration, is very undesirable, since it requires an enormous amount of memory to store the array manifold, especially in the case of multiple parameters. In addition, the MUSIC algorithm is very sensitive to calibration errors [44]. The direction-finding (DF) step of the MUSIC algorithm is also computationally expensive except for some specific array configurations. These problems limit the applicability of the MUSIC-like subspace algorithms.

In this chapter, we address the problem of *joint* array calibration and direction-of-arrival estimation (DOA) with arbitrary arrays. Our goal is to determine the minimal information necessary about the array structure to accomplish this task and develop an algorithm that utilizes this *sufficient* information. The problem resembles the blind equalization problem in data communications, where the data symbols are distorted by finite-memory channels. In blind equalization, the goal is to "open the eye" so that it is possible to jointly estimate the impulse response of the channel and recover the symbols. In the array problem, the aim is to "open the eye of the processor", so that it can "see" the far-field sources (DOA estimation), and "listen to" each of them (waveform recovery).

The blind equalization problem is known to be unsolvable for non-minimum phase systems if processing is limited to the power-spectrum. Higher-than second-order statistics (cumulants) have been shown to be invaluable for solving this problem, since it is possible to recover phase information [78] with cumulants. In array processing, we obtain phase information by cross-correlating channel measurements. It is this phase information that makes eigenstructure-based high-resolution spatial-spectrum estimation possible. In the array processing context, the *motivation* for using cumulants is to recover *more* phase terms than is possible by using only second-order statistics. This goal is accomplished in Chapter 4.

Given an arbitrary array, *joint* calibration and DOA estimation problems can be solved, if we have an identical copy of the array displaced in space, by using the ESPRIT algorithm [108]. In this way, the problems associated with the array calibration procedure are alleviated by incorporating a specific type of *redundancy* into the array configuration. The ESPRIT algorithm can blindly identify the steering vectors and DOA's of sources whose number is limited by the subarray size; hence, waveforms of the sources whose directions are identified can be estimated.

The special array geometry required by ESPRIT is not available in general, so it is difficult to calibrate arbitrary arrays in practice. Rockah and Schultheiss [105] did pioneering work on the conditions required for calibrating isotropic sensors of arbitrary arrays. They proved that if there are three non-colinear spectrally/temporally disjoint sources with unknown bearings, the calibration errors tend to zero as the signal-to-noise ratio (SNR) of the sources tends to infinity. In addition, if the direction of one sensor to the reference sensor is known, then DOA estimation can be done by taking the known direction vector as a reference. These results give us hope that not all of the redundancy required by ESPRIT is necessary, and, that one may do well by using a single doublet rather than having all the sensors occur in pairs.

In this aspect, the benefits of incorporating redundancy in an array structure is analogous to the benefits of channel coding for communication sytems which incorporates redundancy into the actual data. In the channel coding problem, the natural question is "How much redundancy is needed to transmit information with sufficient protection?" For high-rate transmission objectives, the task of the communication engineer is to minimize the required redundancy while maintaining the specifications for error correction. Similarly, the task of the antenna signal processing engineer is to minimize the redundancy required by the ESPRIT algorithm, while maintaining the capability of DOA estimation, array calibration and waveform estimation so that the algorithm can be applied to a wider class of array processing scenarios. In Section 5.2, we propose a solution to the redundancy minimization problem so that it will be possible to calibrate an arbitrary array, by using a single doublet and fourth-order cumulants. Our results are in agreement with the results of Rockah and Schultheiss, although their paper takes the approach of evaluating Cramer-Rao bounds for the unknown parameters and assuming Gaussian processes, whereas ours uses cumulants which are blind to Gaussian processes. The results of Rockah and

Schultheiss [105], about the conditions required for calibration and estimation, play the role of the Shannon Theorem for the capacity of a communication channel; they provide the conditions necessary for calibration but not the way to achieve it without further a-priori information. Our approach is totally different: we propose a method for calibration and then address its consistency.

Following the excellent work of Rockah and Schultheiss [105], work on direction-finding in the presence of sensor uncertainities continued [106], based mainly on the maximum-likelihood approach [166], or utilizing calibration sources whose bearings are known [76]-[89]. Problems with the existing methods arise from iterative nonlinear optimization procedures that require good nominal knowledge about array geometry, unrealistic constraints about sensor responses (e.g., all sensors are assumed to be isotropic), computational complexity of optimization techniques, local convergence problems and sensitivity to noise spatial correlation structure. The method that is proposed in this chapter is non-recursive, depends on the eigendecomposition of cumulant matrices obtained from the observed data, and does not involve any computationally intensive optimization or search procedure. In this way, local convergence problems and noise effects are avoided. In Section 5.3, we provide simulation experiments to illustrate the effectiveness of our method. We conclude the chapter with some final observations and extensions.

## 5.2 Joint Calibration and Parameter Estimation

In this section, we propose the use of $\mathbf{VC}^3$ for calibrating arrays which are illuminated by multiple incoherent far-field sources from unknown directions. The calibration problem can be summarized as estimation of the directions of far-field sources, with an array of unknown array manifold (i.e., sensor locations and responses). Contrary to [105], we allow sources to overlap in time and frequency, do not assume a nominal knowledge for the array geometry and spatial correlation of measurement noise. Clearly, there is need for *some* information about the array; here we investigate what that *sufficient* information is.

Given an arbitrary array, *joint* array calibration and source parameter estimation can easily be solved if we have an identical copy of the array displaced in space with a known displacement vector $\vec{\Delta}$ so that we can apply the ESPRIT algorithm of [108].

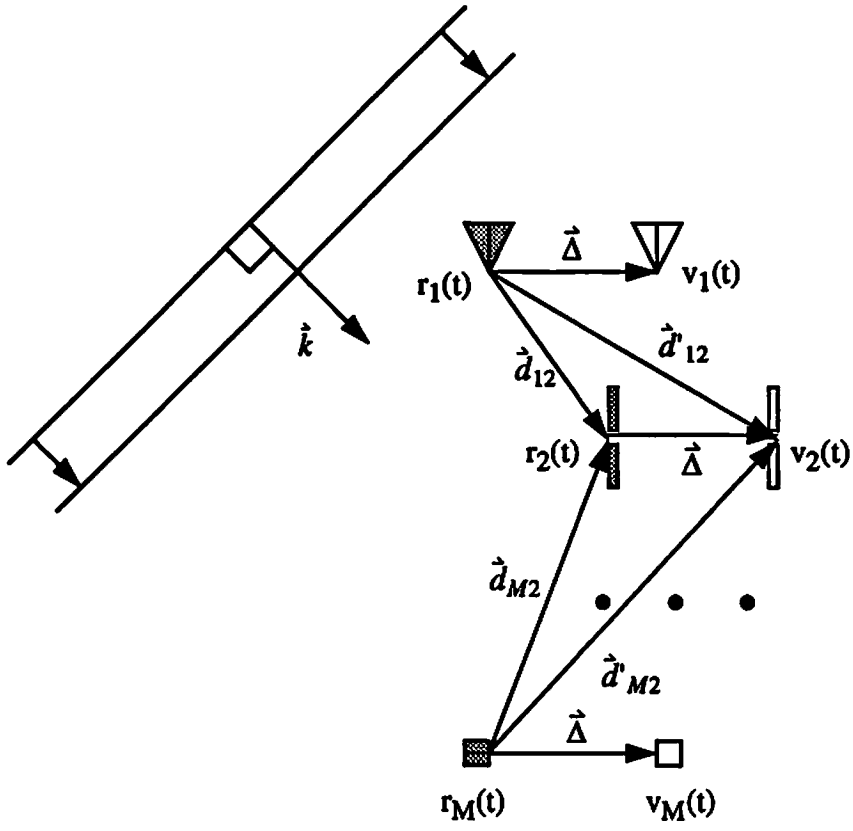



Figure 5.1: Virtual-ESPRIT algorithm: it is possible to reach any virtual sensor location from an actual sensor location by addition of two vectors between actual sensors when $v_1(t)$ is available. In other words, cross-correlation between actual and virtual sensor elements can be calculated by using cumulants, since cumulants can be interpreted as vector addition, whereas cross-covariance can be interpreted as a single vector.

The main questions we will answer in this section are: given an array of arbitrary geometry and sensor responses: (1) Is it necessary to have a full-copy of the array for calibration? (2) If not, how much redundancy is necessary? (3) How can such an algorithm be implemented? We provide the answers based on the results of Chapter 4.

Consider the arbitrary array and its image in Figure 5.1. In order to use the ESPRIT algorithm to jointly estimate the DOA parameters of multiple sources and

the associated steering vectors, we need to compute the cross-correlations between subarrays. For example (see (4.5) and Figure 4.1),

$$E\{r_1^*(t)v_M(t)\} = \sigma_s^2 a_1^* a_M \exp(-j\vec{k}\cdot\vec{d}_{1M})\exp(-j\vec{k}\cdot\vec{\Delta}) \tag{5.1}$$

where $a_M$ denotes the response of the $M$th sensor to the wavefront from the source. Unfortunately, we can not compute (5.1) since we do not have $v_M(t)$. Next, consider

$$E\{r_1^*(t)r_M(t)\} = \sigma_s^2 a_1^* a_M \exp(-j\vec{k}\cdot\vec{d}_{1M}) = E\{r_1^*(t)v_M(t)\}\exp(j\vec{k}\cdot\vec{\Delta}) \tag{5.2}$$

which is computable and is highly related to the correlation in (5.1). If we knew $e^{j\cdot\vec{k}\cdot\vec{\Delta}}$ then we could solve (5.2) for $E\{r_1^*(t)v_M(t)\}$; however, this is not possible since we do not know the propagation vector $\vec{k}$.

From Figure 5.1 we observe that all vectors joining two sensors in separate subarrays can be decomposed as the addition of two vectors, one in the actual subarray, and the other one being the displacement vector $\vec{\Delta}$, e.g., $\vec{d}'_{12} = \vec{d}_{12} + \vec{\Delta}$. All the computable correlations (5.2) lack the common term $\exp(-j\vec{k}\cdot\vec{\Delta})$. It is necessary to form a *bridge* between the subarrays to recover this phase term.

From our results in Chapter 4, we know that by using fourth-order cumulants, and assuming that only one doublet $\{r_1(t), v_1(t)\}$ is available, we can compute the cross-correlations between subarray elements, e.g.,

$$E\{r_k^*(t)v_l(t)\} = \frac{\sigma_s^2}{\gamma_{4,s}|a_1|^2}\,\mathrm{cum}(r_1^*(t), v_1(t), r_k^*(t), r_l(t)) \tag{5.3}$$

Use of fourth-order cumulants provides the *magic touch* necessary to compute cross-correlations between actual and virtual sensors so that the ESPRIT cross-correlation matrix can be generated. Similarly, the cross-correlation of actual sensors can be computed by cumulants, e.g.,

$$E\{r_k^*(t)r_l(t)\} = \frac{\sigma_s^2}{\gamma_{4,s}|a_1|^2}\,\mathrm{cum}(r_1^*(t), r_1(t), r_k^*(t), r_l(t)) \tag{5.4}$$

In this way, (5.3) and (5.4) can be used to form the covariance matrix required by ESPRIT [108] by using cumulants. For obvious reasons, we call the single pair

of sensors that form the doublet "*guiding* sensors", and the method the *virtual*-ESPRIT algorithm (VESPA). Note that the VESPA requires only a single doublet rather than a full-copy of the array, resulting in enormous hardware reductions. VESPA also alleviates the problems resulting from the perfect sampling synchronization requirements of the covariance-ESPRIT for the two subarrays. In VESPA, synchronization must be maintained only between the elements of the single doublet. A flowchart for VESPA is provided in Figure 5.2.

In applications, we do not have the true cumulants; they are replaced by *consistent* estimates which converge to true values as the data length grows to infinity. The results in [81] indicate that convergence to the true values is rapid at high SNR.

The requirement of a single doublet with known orientation $\vec{\Delta}$ is in fact the necessary requirement in [105]. Rockah and Schultheiss [105] derive this result assuming isotropic sensors and Gaussian processes that are disjoint in time or frequency. Our approach utilizes fourth-order cumulants which are blind to Gaussian processes, yet we obtain the same requirement about the necessary information for identifiability! In addition, the consistency results for both methods require high-SNR and long data lengths. On the other hand, the requirements of our approach are very mild when compared to those in [105] and its extensions: (1) We allow multiple sources sharing the same frequency band due to **[CP3]**, (2) Our approach is applicable to arbitrary arrays (the isotropic sensor arrays of [105] are a subclass); (3) Our approach is also applicable to nominally-linear arrays, whereas [105] fails theoretically in this case; (4) The cumulant-based approach does not require information about noise spatial correlation, unlike the covariance-based algorithm; (5) From an implementation point of view, our approach is non-iterative, and eliminates parameter search by using ESPRIT, unlike the method in [166] which requires iterations, and very good nominal knowledge about the array to start a multidimensional search; (6) In the presence of white observation noise, it is possible to initialize the maximum-likelihood approach proposed in [166] by the results of VESPA to get even better estimates; (7) In the presence of colored Gaussian noise, a trispectral maximum-likelihood approach can be designed along the lines of [41] initialized by VESPA to calibrate arbitrary arrays in a maximum-likelihood fashion without the knowledge of noise color; and, finally, (8) The next chapter sets the stage for removing the effects of non-Gaussian colored noise, which can be followed by a ML approach.

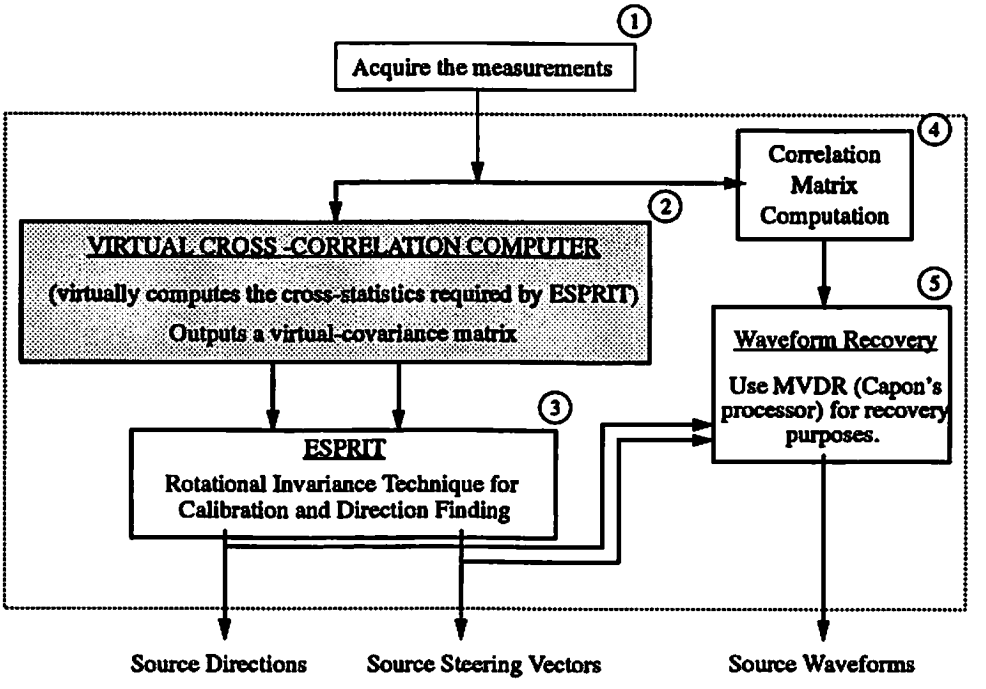



Figure 5.2: The role of virtual cross-correlation computer ($\mathbf{VC}^3$) in the joint array calibration and DOA estimation problem. $\mathbf{VC}^3$ provides ESPRIT the required covariance matrix without the need for a full copy of the sensor array. The matrices provided by $\mathbf{VC}^3$ are also insensitive to colored Gaussian noise.

To recover the waveforms associated with the far-field sources we first estimate the steering vectors by subspace rotation [108]. If the measured signals are represented as $\mathbf{r}(t) = [r_1(t), r_2(t), \ldots, r_M(t), v_1(t)]^T$, then we augment the steering matrix by the estimated response of the sensor that measures $v_1(t)$. Let $\mathbf{a}_1$ be the $(M+1)$x1 steering vector of the signal of interest (SOI) $s_1(t)$ with estimated bearing $\theta_1$, and let the augmented steering matrix be decomposed as $\mathbf{A} = [\mathbf{a}_1, \mathbf{A}_I]$. The augmented steering matrix $\mathbf{A}$ can be estimated from (2.47). We propose two approaches for signal-recovery:

• **Minimum-Variance Distortionless Response Beamformer (MVDR):**
This beamformer estimates the SOI waveform in the mean-square sense, i.e.,

$$\hat{s}_1(t) = \mathbf{w}_1^H \mathbf{r}(t) = (\mathbf{R}^{-1}\,\mathbf{a}_1)^H \mathbf{r}(t) \tag{5.5}$$

where $\mathbf{R} = E\{\mathbf{r}(t)\mathbf{r}^H(t)\}$. In (5.5), we ignored a scale factor since it does not effect the output SNR.

• **MVDR with Perfect Nulling (null-MVDR):**

This beamformer estimates the SOI waveform in the mean-square sense while putting perfect nulls on the interferers, i.e.,

$$\hat{s}_2(t) = \mathbf{w}_2^H \mathbf{r}(t) \tag{5.6}$$

where the weight vector $\mathbf{w}_2$ is the solution of the linearly-constrained minimum variance problem:

$$\mathbf{w}_2 = \min_{\mathbf{w}} \mathbf{w}^H \mathbf{R} \mathbf{w} \quad \text{subject to} \quad \left[ \mathbf{a}_1, \ \mathbf{A}_\mathrm{I} \right]^H \mathbf{w} = \mathbf{f} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{5.7}$$

which has the solution

$$\mathbf{w}_2 = \mathbf{R}^{-1} \mathbf{A} \left( \mathbf{A}^H \mathbf{R}^{-1} \mathbf{A} \right)^{-1} \mathbf{f} \tag{5.8}$$

Both of the above beamformers do not require knowledge of the measurement noise covariance matrix. Derivations of (5.5) and (5.8) can be found in [80].

## 5.3 Simulations

In this section, we present simulation experiments to demonstrate the performance of the joint parameter estimation and calibration algorithms proposed in this chapter.

We assume the presence of two far-field equal power BPSK broadcasts, with steering vectors:

$$\begin{aligned} \mathbf{a}_1 &= [1.0, 0.5e^{j\pi 0.3}, 1.2e^{j\pi 0.1}, 0.8e^{j\pi 0.7}, 1.0e^{j\pi 0.5}, 2.0e^{j\pi 1.2}, 0.0e^{j\pi 0.0}, 1.0e^{j\pi \sin(-2\pi/180)}]^T \\ \mathbf{a}_2 &= [1.0, 0.1e^{j\pi 0.8}, 0.5e^{j\pi 0.5}, 1.5e^{j\pi 0.1}, 0.1e^{j\pi 1.2}, 1.0e^{j\pi 0.2}, 0.6e^{j\pi 0.4}, 1.0e^{j\pi \sin(\ 2\pi/180)}]^T \end{aligned} \tag{5.9}$$

where $\theta_1 = -2^o$ and $\theta_2 = 2^o$. The first and last components of each steering vector correspond to the responses of the doublet sensors whose orientation is known. A

separation of half-wavelength is assumed between guiding sensors. Conventional covariance-ESPRIT can not be applied to this scenario, since it requires a full-copy of the array. To compare our cumulant-based virtual DOA finding algorithm (VESPA) with the conventional ESPRIT algorithm we assumed the existence of a second subarray for the latter (for the former this is not necessary and not available). In this case, the last element of the steering vectors (5.9) represents the response of the first element of the second subarray required for covariance-ESPRIT. We compare the performance of cumulant and covariance based algorithms in the presence of colored noise. We also present results for steering vector estimation and waveform recovery by the proposed two beamformers. Finally, we compare the waveform recovery performance of our approach and Capon's MVDR beamformer [19] which uses a perfectly known steering vector.

### 5.3.1 Direction-Finding

We created the following scenario to demonstrate noise suppression performance of VESPA: two sources from $\pm 2^o$ illuminate the array with an SNR of 10dB with respect to the noise level at the guiding sensors (noise components at the guiding sensors are uncorrelated and their power is unity). The remaining sensors have the correlation matrix $\mathbf{R}_n \triangleq \mathbf{L}\mathbf{L}^H$, where

$$\mathbf{L} \triangleq \begin{bmatrix} 10.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ 9.41e^{j2.05} & 8.41 & 0.00 & 0.00 & 0.00 & 0.00 \\ 6.08e^{-j1.62} & 3.05e^{j2.61} & 6.18 & 0.00 & 0.00 & 0.00 \\ 1.52e^{j1.11} & 0.76e^{-j0.94} & 0.29e^{j2.73} & 5.60 & 0.00 & 0.00 \\ 5.48e^{j2.41} & 2.75e^{j0.36} & 1.06e^{-j2.25} & 0.24e^{1.30} & 5.97 & 0.00 \\ 2.87e^{j2.27} & 1.44e^{j0.22} & 0.55e^{-j2.39} & 0.12e^{j1.16} & 0.41e^{-j0.14} & 5.66 \end{bmatrix}$$

This example shows us how to calibrate a poorly conditioned array by two better conditioned sensors. The success of the results can motivate the use of such a doublet as a carry-on unit to calibrate different antenna arrays. Table 5.1 shows DOA estimates as the sample size increases. As noticed, ESPRIT totally fails to estimate the parameters given the number of signals. On the other hand, VESPA performs better with increasing sample size. In this experiment, we tried ESPRIT by setting the number of signals to be 4 (overestimate). This setting provided

the best possible results (which are still not acceptable). Table 5.1 indicates total inconsistency for the covariance-ESPRIT.

To understand the reason why the standard deviations of covariance-ESPRIT are very large, we display the estimates from 200 Monte Carlo runs for 5000 snapshots in Figure 5.3. As clearly seen in the figure, covariance-ESPRIT breaks down many times. This is the result of the overparametrization and strict violation of the ESPRIT data model due to the presence of spatially-colored noise. Covariance-ESPRIT estimates are within $0.5^o$ of the true values for 88 of 200 runs, whereas VESPA yields estimates always within this range.

### 5.3.2 Waveform Recovery

In this experiment, we illustrate the difference between the MVDR and null-MVDR beamformers proposed earlier in the chapter. Null-MVDR puts exact nulls on the interferers to eliminate the effect of changes in the jamming strategy after weights are computed. This results in a slight increase in the output noise power since some degrees of freedom (equal to number of interferers) are lost in this way. The way in which these beamformers use directional information from VESPA is illustrated in Figure 5.2.

Suppose two sources (one BPSK and one CW) illuminate the array from $+2^o$ and $-2^o$ respectively, as described earlier. Both sources have an SNR of 10dB and noise is white. We are interested in the message of the BPSK source. After collecting 1000 snapshots, we estimate the steering vectors and construct the two beamformers. Figure 5.4 illustrates the results for the MVDR beamformer, which recovers the message almost perfectly. After weights are computed with the first 1000 shapshots, they are fixed and the beamformers continue processing the received signals; however, the CW jammer then increases its power to 60dB. With the MVDR weights already computed for the equal power sources, the message is totally destroyed; however, null-MVDR recovers the message since it puts a null on the jammer regardless of its power. It must be clearly understood that when the steering vector of the jammer can not be estimated reliably (short data lengths and very low SNR), it is better to use the MVDR, since the null-MVDR zero-forcing constraint will be useless.

Table 5.1: Results from 200 Monte-Carlo runs for spatially-colored noise

| Data Length | $\theta_1 = -2^o$ | | | | $\theta_2 = 2^o$ | | | |
|---|---|---|---|---|---|---|---|---|
| | VESPA | | ESPRIT | | VESPA | | ESPRIT | |
| | mean | std | mean | std | mean | std | mean | std |
| 50 | -1.7500 | 1.1232 | -4.6519 | 8.1553 | 1.2851 | 1.6423 | 5.0800 | 10.0122 |
| 100 | -1.8426 | 0.8788 | -4.3408 | 8.7436 | 1.3006 | 1.2729 | 3.5067 | 7.3308 |
| 200 | -1.9282 | 0.6016 | -4.1822 | 6.7576 | 1.6772 | 1.0005 | 3.5192 | 8.2243 |
| 500 | -1.9355 | 0.3403 | -3.4685 | 6.9108 | 1.8930 | 0.4597 | 2.7725 | 6.2439 |
| 750 | -1.9942 | 0.2717 | -3.5411 | 6.1400 | 1.9308 | 0.2692 | 2.9638 | 6.3929 |
| 1000 | -1.9937 | 0.2551 | -3.0133 | 5.6189 | 1.9402 | 0.2457 | 3.9767 | 8.0984 |
| 1500 | -1.9909 | 0.2088 | -3.3463 | 7.9549 | 1.9749 | 0.2056 | 3.4459 | 7.6065 |
| 2000 | -2.0075 | 0.1616 | -3.2354 | 6.2054 | 1.9920 | 0.1903 | 3.4155 | 7.4958 |
| 2500 | -1.9998 | 0.1622 | -3.2337 | 6.5632 | 2.0008 | 0.1593 | 3.2515 | 7.6784 |
| 3000 | -1.9933 | 0.1251 | -3.1169 | 6.3253 | 1.9953 | 0.1360 | 3.3567 | 6.9544 |
| 3500 | -2.0002 | 0.1224 | -4.5073 | 9.9299 | 1.9876 | 0.1430 | 2.6196 | 5.5402 |
| 4000 | -1.9990 | 0.1089 | -3.0528 | 6.2277 | 1.9952 | 0.1308 | 2.8882 | 6.5228 |
| 4500 | -2.0032 | 0.1120 | -2.4213 | 4.1917 | 2.0010 | 0.1130 | 3.3673 | 6.6588 |
| 5000 | -2.0032 | 0.1044 | -3.3839 | 7.3596 | 1.9994 | 0.1026 | 3.5328 | 8.2402 |
| 10000 | -2.0016 | 0.0738 | -2.6734 | 5.7416 | 2.0025 | 0.0794 | 3.3571 | 6.7082 |

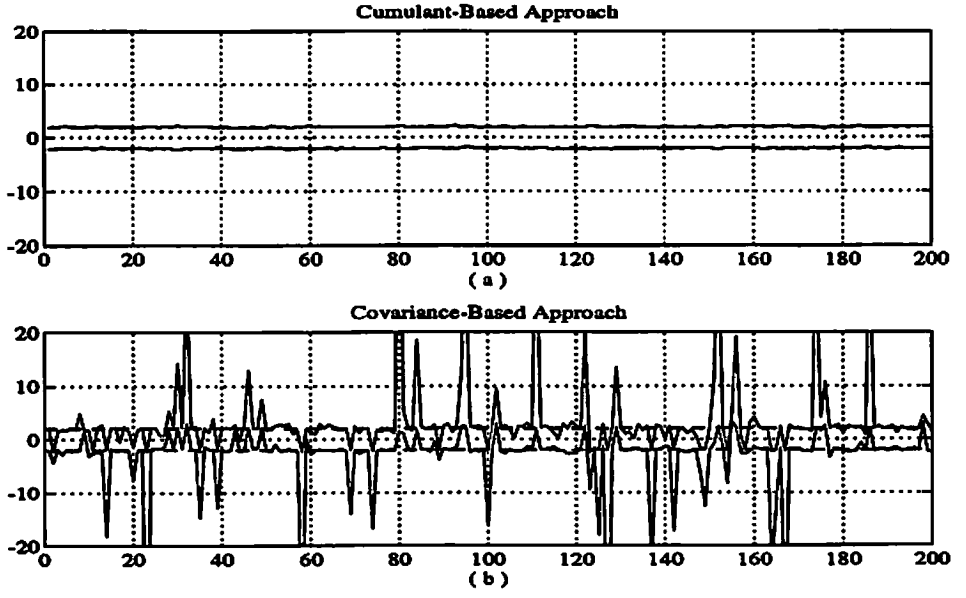



Figure 5.3: Cumulant (VESPA) and covariance-ESPRIT DOA estimates from 200 runs: (a) VESPA estimates are very close to the true values, (b) Covariance-ESPRIT with overparametrization breaks down many times.

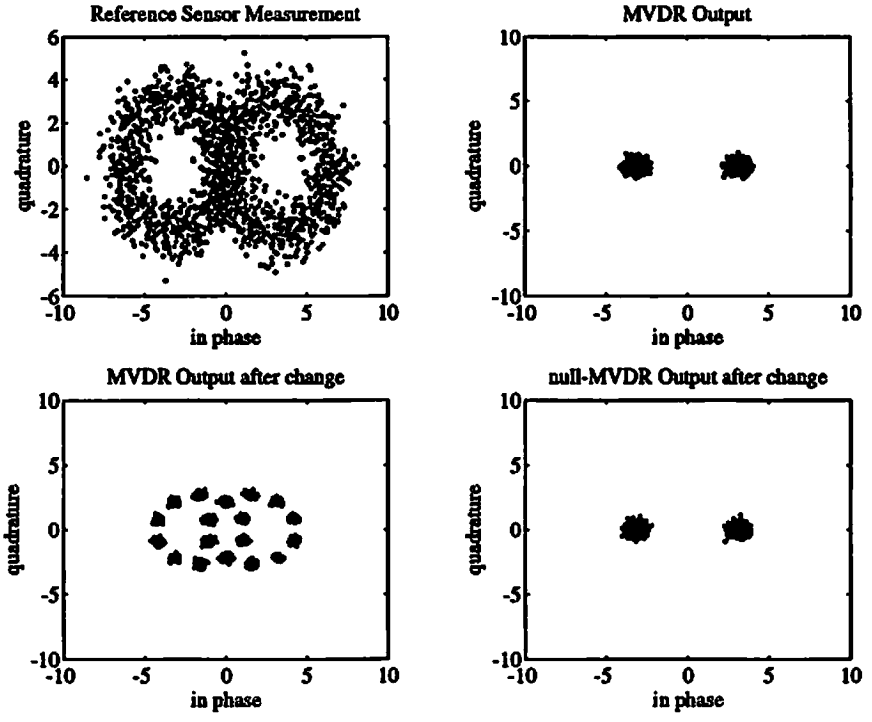



Figure 5.4: CW and BPSK sources illuminate the array: (top left) received signal at the reference sensor, (top right) MVDR output which nulls the CW component, (bottom left) CW jammer increases its power and penetrates through the MVDR processor destroying the BPSK message, (bottom right) null-MVDR nulls the interference with the precomputed weights and recovers the message.

### 5.3.3 Signal Enhancement Comparison

In this experiment, we compare the performance of the cumulant-based signal recovery techniques and Capon's MVDR beamformer with perfectly known steering vector in terms of the signal-to-interference-plus-noise (SINR) at the output of the processor.

Capon's MVDR beamformer is a very sensitive processor to mismatches about the steering vector of the source whose waveform is to be estimated. In this experiment, we assume that perfect knowledge of the steering vector is available to the Capon processor to implement (5.5). Even in this case, Capon's MVDR has problems due to the sample matrix inversion in (5.5).

There are two major techniques to address the mismatches created by sample matrix inversion: (1) Robust adaptive beamforming technique [31] which constrains the weight vector of the MVDR processor to lie in a hypersphere, and (2) Steering vector projection approach of [37] which, as its name implies, first projects the known steering vector on the signal subspace of the sample covariance matrix, and then computes (5.5). We decided to investigate the performance improvement offered by the second method since the hypersphere constraint in the former approach prevents the directional interference to be *sufficiently* removed from the output.

This experiment compares the performance of four beamformers: (1) Capon's MVDR which uses exact knowledge (MVDR1), (2) Capon's MVDR improved by the projection method (MVDR2), (3) MVDR beamformer that uses the *estimated* steering vector by the cumulant-based technique (VESPA) developed in this chapter (CUM1), and (4) Cumulant-based MVDR beamformer (VESPA) improved by the projection method (CUM2). We assume two BPSK sources of the previous experiments illuminate the array with the same directional parameters. Sensor noise is assumed to be white and of equal power at each sensor. We first let SNR be 20dB for both sources and designate the source from $-2^{o}$ as the signal of interest. We illustrate our results in Figure 5.5. Each data point in the figure is obtained by averaging results from 200 Monte-Carlo runs. The range from 50 to 1000 snapshots is sampled intensively. The SINR at the output of the Capon's beamformer with true statistics and steering vector is 29.6892 dB. We observe that cumulant-based approaches are very close to providing the maximum-possible output SINR if more than 500 snapshots are available. For short data lengths, MVDR1 is 15-20 dB away from the expected SINR. Projection helps MVDR2 to recover, but it can not outperform CUM1 for any data length! After 1000 snapshots, CUM1 and CUM2 performance is indistinguishable, i.e., we do not gain very much by projection since the estimates are *in harmony* with the sample covariance matrix. Figure 5.6 illustrates the case when SNR is increased to 30 dB. MVDR1 and MVDR2 can not *recover* with the available snapshots: even at 4000 snapshots, MVDR1 is 14 dB, and MVDR2 is 4dB away from expectations which are perfectly met by CUM1 and CUM2 only after 500 snapshots. Note that CUM2 offers little improvement over CUM1 whereas MVDR2 offers a large improvement over MVDR1. This is important since the projection method only applies when the signal subspace can be identified (white noise case).

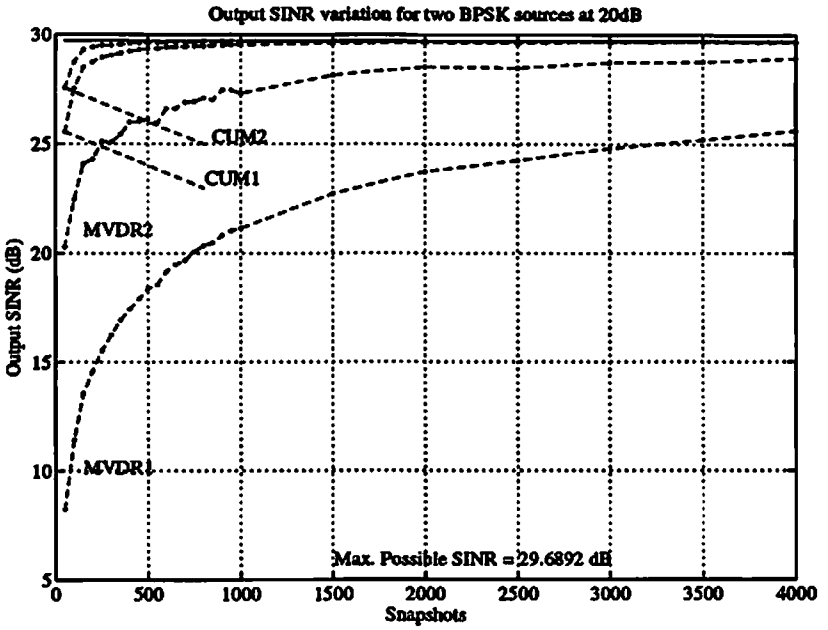



Figure 5.5: Performance comparison for the beamformers with two sources illuminating the array with SNR=20dB at the reference sensor.

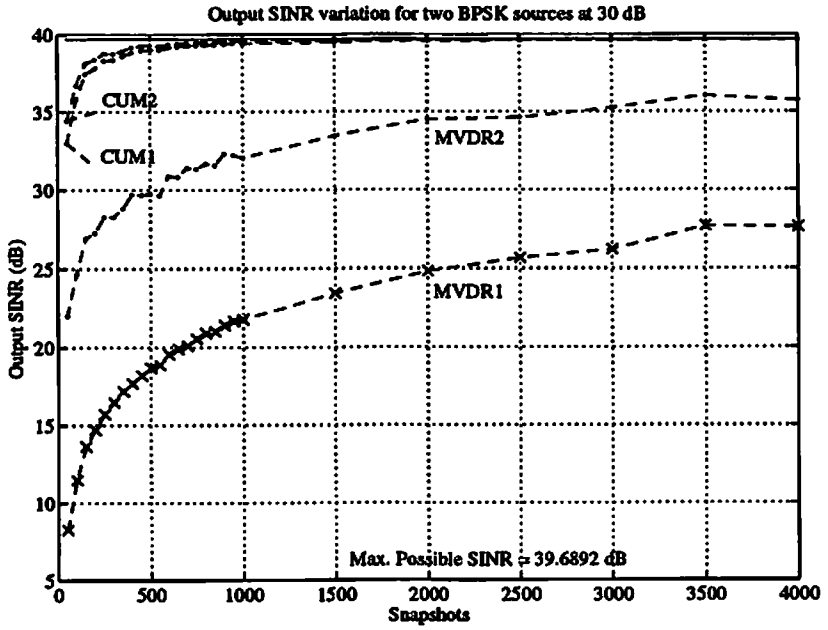



Figure 5.6: Performance comparison for the beamformers with two sources illuminating the array with SNR=30dB at the reference sensor. The crosses on MVDR1 performance curve indicate the sampling points.

The situation encountered here resembles the impedance matching problem in transmission line theory. The moral of that story is: "Always match your line with a load of equal impedance". Similarly the moral of our story is: "Always match your (inverted) sample covariance matrix with a steering vector of equal quality" (an estimate consistent with the received signals rather than a perfectly known one).

## 5.4 Conclusions

In this chapter, we applied virtual cross-correlation computation (Chapter 4) to the joint array calibration and direction-finding problem. We established several relationships between the proposed method and existing work.

The algorithms of this chapter (DOA estimation, array calibration and waveform recovery) are asymptotically insensitive to the spatial correlation structure of additive Gaussian sensor noise. Furthermore, our cumulant-based algorithms are *computationally simpler* than covariance-based counterparts, which require multidimensional search and suffer from local convergence problems [166]. The consistency requirements of cumulant and covariance-based approaches are found to be similar; but the latter are applicable to a very limited class of scenarios, i.e., isotropic sensor arrays [105], whereas the former can calibrate arbitrary arrays using only a doublet of unknown characteristics.

When compared with existing cumulant-based signal recovery algorithms [20, 25, 34, 148, 110], virtual-ESPRIT provides both source bearings in addition to source waveforms with minimal information on the array manifold. Extensions of virtual-ESPRIT are provided in the Appendix.

## 5.5 Appendix–Extensions

In this section, we extend the application of the virtual-ESPRIT algorithm to different scenarios. We start with the use of third-order cumulants for the joint array calibration and direction-finding problem, which in turn leads to a new class of algorithms suitable for fourth-order cumulants. We also extend the virtual-ESPRIT algorithm to wideband signals and near-field source localization problems.

### 5.5.1 Virtual-ESPRIT with Third-Order Cumulants

In Chapter 4, we showed how to compute cross-correlations between sensors by using third-order cumulants. In this section, we extend our results to compute the statistics required by the ESPRIT algorithm by using third-order cumulants and a doublet of guiding sensors.

Let $\{g_1(t), g_2(t)\}$ denote the guiding sensor measurements and consider the location of the first guiding sensor as the reference point. The cross-correlations among the rest of the array measurements $\{r_1(t), r_2(t), \ldots, r_M(t)\}$ can be computed by using our previous results (4.13)-(4.15) as:

$$E\{r_p^*(t)r_q(t)\} = \frac{\sigma_s^2}{\gamma_{3,s}}\text{cum}(r_p^*(t), r_q(t), g_1(t)) \quad 1 \leq p, q \leq M \tag{5.10}$$

Our goal is to compute the cross-correlation matrix between the actual sensor measurements and their virtual counterparts denoted as $\{v_1(t), v_2(t), \ldots, v_M(t)\}$. Before proceeding we note that

$$v_p(t) = r_p(t)\exp(-j\vec{k}\cdot\vec{\Delta}) \quad 1 \leq p \leq M \tag{5.11}$$

and similarly,

$$g_2(t) = g_1(t)\exp(-j\vec{k}\cdot\vec{\Delta}) \tag{5.12}$$

which implies that we can compute the required statistics by using cumulants as follows:

$$E\{r_p^*(t)v_q(t)\} = E\{r_p^*(t)r_q(t)\}\exp(-j\vec{k}\cdot\vec{\Delta}) = \frac{\sigma_s^2}{\gamma_{3,s}}\text{cum}(r_p^*(t), r_q(t), g_1(t))\exp(-j\vec{k}\cdot\vec{\Delta})$$

$$= \frac{\sigma_s^2}{\gamma_{3,s}}\text{cum}(r_p^*(t), r_q(t), g_1(t)\exp(-j\vec{k}\cdot\vec{\Delta})) = \frac{\sigma_s^2}{\gamma_{3,s}}\text{cum}(r_p^*(t), r_q(t), g_2(t)) \tag{5.13}$$

These expressions indicate that three arguments in a cumulant are sufficient to generate the statistics required by the ESPRIT algorithm. The interpretation of (5.10) and (5.13) is that if we change the location of the so-called common element ($g_1(t)$ in (5.10)) by $\vec{\Delta}$, then we obtain the cross-correlation matrix of the actual array measurements and their virtual counterparts which are displaced by $\vec{\Delta}$, due

to the spatial-nonstationarity of the third-order cumulants. This result indicates that one of the arguments of the fourth-order version of the virtual-ESPRIT may be redundant; and it can be used in alternate ways as we shall see in the next section.

### 5.5.2 Modifications of Virtual-ESPRIT

In the previous section, we indicated that three arguments in a cumulant expression are sufficient to implement the virtual-ESPRIT algorithm; hence, the question is "how can the fourth-argument be utilized when fourth-order cumulants are employed?". The first two parts of this section are devoted to this issue. The third part describes use of multiple guiding sensors.

Consider an array of $M$ elements that measures $\{r_1(t), r_2(t), \ldots, r_M(t)\}$, and the two guiding sensor measurements, $\{g_1(t), g_2(t)\}$ which are actually obtained from the first two-sensors (i.e., $g_1(t) = r_1(t)$ and $g_2(t) = r_2(t)$). Let $a(t)$ be a linear function of the array measurements (i.e., $a(t) = \mathbf{w}^H \mathbf{r}(t)$). Assume there is only one source $s(t)$, and the signal part of $a(t)$ is equal to $\beta s(t)$. The fourth-order cumulant obtained by fixing one of the arguments to be $a(t)$ is functionally identical to the third-order cumulant for direction-finding purposes, i.e.,

$$\text{cum}(a^*(t), g_1(t), r_p^*(t), r_q(t)) = \frac{\gamma_{4,s}\beta^*}{\gamma_{3,s}} \text{cum}(r_p^*(t), r_q(t), g_1(t)) \tag{5.14}$$

because we can view the constant term in (5.14) as a scaling of the new third-order cumulant of the source which does not affect the direction-finding performance. Covariances computed by using cumulants as in (5.14) are used to form the autocorrelation matrix required by the ESPRIT algorithm. Similarly,

$$\text{cum}(a^*(t), g_2(t), r_p^*(t), r_q(t)) = \frac{\gamma_{4,s}\beta^*}{\gamma_{3,s}} \text{cum}(r_p^*(t), r_q(t), g_2(t)) \tag{5.15}$$

can be used to compute the cross-correlation matrix required by the ESPRIT algorithm. The way to obtain $a(t)$ is important, and it will be discussed in the following.

#### 5.5.2.1 Increasing Dimensionality

The process $a(t)$ can be selected as one of the $M$ sensor measurements. Let

$$[\mathbf{C}_1(l)]_{p,q} = \text{cum}(a^*(t), g_1(t), r_p(t), r_q^*(t))\Big|_{a(t)=r_l(t)} \tag{5.16}$$

$$[\mathbf{C}_2(l)]_{p,q} = \text{cum}(a^*(t), g_2(t), r_p(t), r_q^*(t))\Big|_{a(t)=r_l(t)} \tag{5.17}$$

In this way, there are $M$ selections for $a(t)$. In the original development of virtual-ESPRIT, we let $a(t) = g_1(t)$, which in turn is only a small subset of all possible choices. Of course, the reason behind that selection was to illustrate the operation of the virtual-ESPRIT more clearly. With each selection of $a(t)$ in (5.16) and (5.17), we obtain a third-order virtual-ESPRIT problem. All $(M)$ of these problems can be combined to provide the direction estimates, if we construct the following $M^2$x$M$ matrices:

$$\mathbf{T}_1 = \begin{bmatrix} \mathbf{C}_1(1) \\ \mathbf{C}_1(2) \\ \vdots \\ \mathbf{C}_1(M) \end{bmatrix} \quad \mathbf{T}_2 = \begin{bmatrix} \mathbf{C}_2(1) \\ \mathbf{C}_2(2) \\ \vdots \\ \mathbf{C}_2(M) \end{bmatrix} \tag{5.18}$$

which take the form

$$\mathbf{T}_1 = (\mathbf{A}^* \otimes \mathbf{A})\,\mathbf{\Gamma}_{4,s}\,\mathbf{A}^H \quad \mathbf{T}_2 = (\mathbf{A}^* \otimes \mathbf{A})\,\mathbf{\Phi}\,\mathbf{\Gamma}_{4,s}\mathbf{A}^H \tag{5.19}$$

where $\mathbf{B} \triangleq (\mathbf{A}^* \otimes \mathbf{A})$ can be viewed as an effective steering matrix. $\mathbf{\Gamma}_{4,s}$ is a diagonal matrix that contains the fourth-order cumulants of the far-field sources. The derivation of (5.19) is possible by using (5.14) and (5.15). $\mathbf{\Phi}$ contains the directional information and was defined earlier in (2.37). We can use the ESPRIT algorithm to solve for the elements of $\mathbf{\Phi}$ which contain the direction information and the effective steering matrix $\mathbf{B}$, which consists of the effective steering vectors of sources. After the columns of $\mathbf{B}$ are determined ($\mathbf{b} = \mathbf{a}^* \otimes \mathbf{a}$), the $M^2$-vectors can be reconfigured in a Hermitian matrix which is rank one ($\text{unvec}(\mathbf{b}, M, M) = \mathbf{a}\mathbf{a}^H$). Then, the steering vectors can be determined by taking the principal components of these reconfigured matrices. This provides additional smoothing for the estimation of the steering vectors.

#### 5.5.2.2 Beamforming and Virtual-ESPRIT

In the previous section, we indicated a computationally demanding extension to the original virtual-ESPRIT. An alternate approach which can parallelize the computations is possible by the selection of $a(t)$ to be spatial-filtered array measurements, i.e, $a(t) = \mathbf{w}^H \mathbf{r}(t)$. The weight-vector can be determined (using the MVDR) to suppress all but one of the sources illuminating the array by using the initial estimates of the steering vectors provided by virtual-ESPRIT. Performing this procedure for all of the $P$ sources, we generate $P$ separate, theoretically rank-one virtual-ESPRIT problems in which the cross-correlation matrices between the subarrays can be computed as:

$$\text{cum}(a^*(t), g_2(t), r_p(t), r_q^*(t)) \tag{5.20}$$

Generating rank-one virtual-ESPRIT problems is motivated to alleviate the effects of finite-sample estimates of cumulants by suppressing the residual cross-terms between multiple signals which only decay to zero asymptotically. These problems can be solved in parallel for each source of interest.

This technique can be further improved by reducing the dimensionality of the main array by putting null-constraints on all but one of the sources. If $\mathbf{A}$ is an $M$x$(P-1)$ matrix that contains the estimates of steering vectors of the undesired sources, and $\mathbf{a}_1$ is the estimate of the steering vector for the desired source, then the modified array measurements can be obtained by the transformation

$$\tilde{\mathbf{r}}(t) = \mathbf{E}^H \mathbf{r}(t) \tag{5.21}$$

where the columns of $\mathbf{E}$ constitute an orthonormal basis for the left-nullspace of $\mathbf{A}$. This transformation can not suppress the desired signal, since $\mathbf{a}_1$ can not be represented as a linear combination of the columns of $\mathbf{A}$. Due to the tranformation in (5.21), $\tilde{\mathbf{r}}(t)$ is an $(M-P+1)$ vector. The virtual-ESPRIT subarray cross-correlation matrix can be computed as

$$\text{cum}(a^*(t), g_2(t), \tilde{r}_p(t), \tilde{r}_q^*(t)) \tag{5.22}$$

#### 5.5.2.3 Multiple Guiding Sensors

We now investigate how to use multiple guiding sensors in the virtual-ESPRIT algorithm. One approach is to lay the guiding sensors as a minimum-redundancy array so as to generate as many copies of the main array as possible. This is motivated by the 2-D MRA design procedure of Chapter 6. Figure 5.7 illustrates an example in which four guiding sensors are used to generate six virtual copies of an arbitrary array. The seven subbarrays can be grouped as two overlapping super-subarrays such that each super-subarray contains four subarrays (other grouping options are also possible). The ESPRIT algorithm can be applied with the effective subbarray size being equal to the super-subarray size. An alternative approach is to perform beamforming on the virtual covariance matrices as suggested in the previous section (on the received signals). Here we do not have the virtual signals, so we do the processing on the statistics.

In some applications, both azimuth and elevation information is necessary. Modifications of the ESPRIT algorithm (see the introduction of [143] for an extensive summary of algorithms) can be applied if the array contains displaced copies of a subarray for every dimension. We can compute the virtual statistics with $\mathbf{VC}^3$ if we have three guiding sensors. Figure 5.8 illustrates the guiding sensor configuration for this application. For example,

$$\text{cum}(g_2^*(t), g_3(t), r_k(t), r_l^*(t) = \frac{\gamma_{4,s}}{\sigma_s^2} E\{y_k(t) z_l^*(t)\} \tag{5.23}$$

### 5.5.3 Wideband Signals

In this section, we extend the virtual-ESPRIT algorithm to wideband signals. The spectrum of interest can be divided into $L$ narrowbands around the center frequency $w_c$. Let $r_k(t, \omega_j)$ denote the measurement from the $k$th sensor at frequency $\omega_j$, and

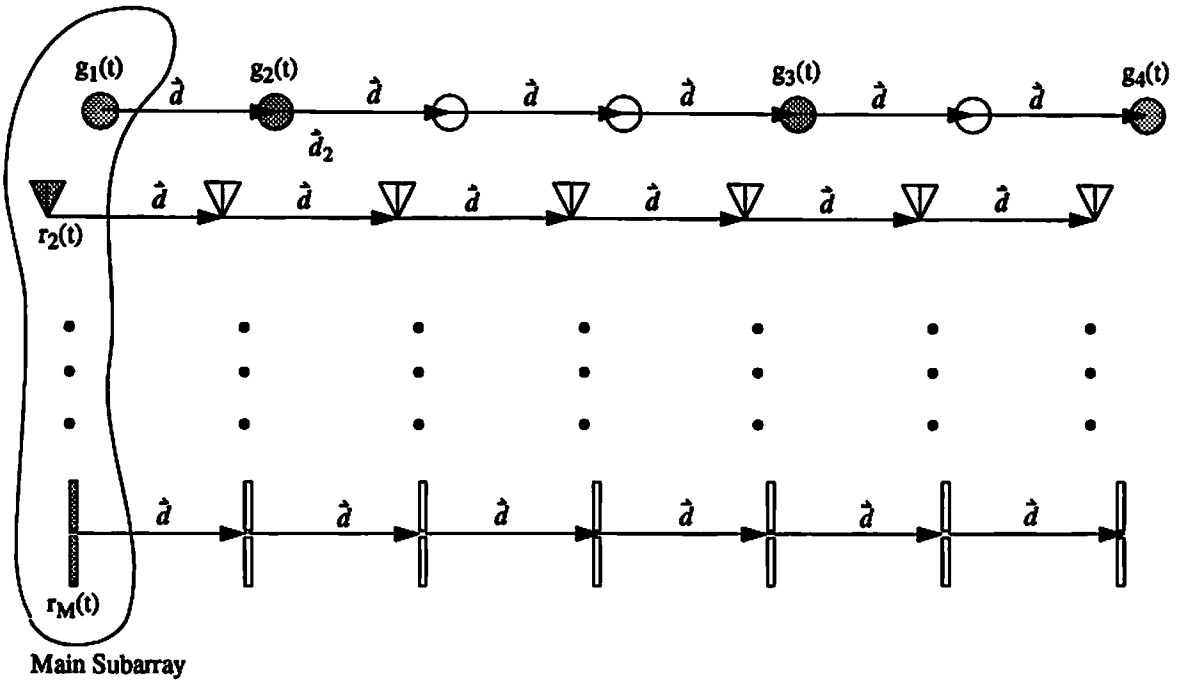



Figure 5.7: A minimum-redundancy array of four guiding sensors is used to create six copies of the main array. The filled sensors indicate the actual sensors.

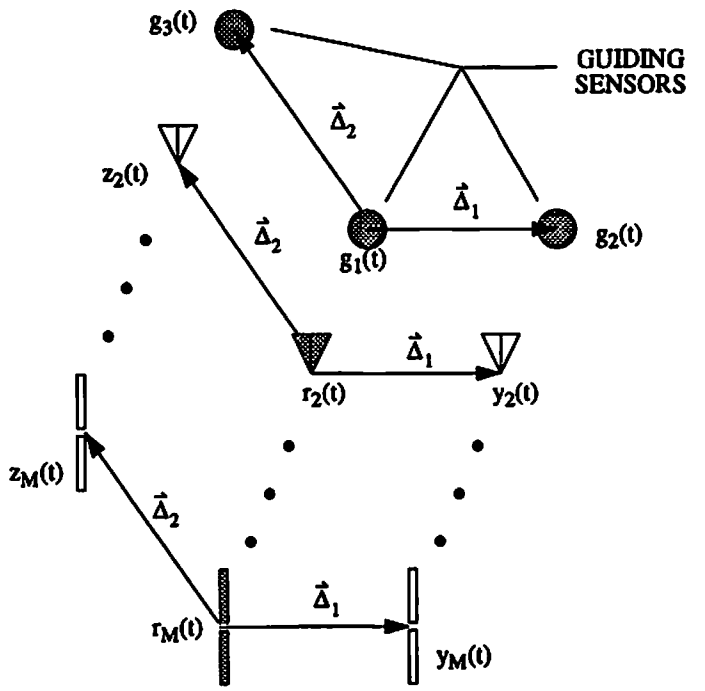



Figure 5.8: Azimuth elevation direction-finding is possible if we use $\mathbf{VC^3}$ and three guiding sensors.

$\{g_1(t,w_c), g_2(t,w_c)\}$ denote the guiding sensor measurements at the center frequency $\omega_c$. Let

$$[\mathbf{C}_1(\omega_j)]_{k,l} = \text{cum}(g_1^*(t,\omega_c), g_1(t,\omega_c), r_k(t,\omega_j), r_l^*(t,\omega_j)) = \frac{\gamma_{4,s}}{\sigma_s^2} E\{r_k(t,\omega_j), r_l^*(t,\omega_j)\} \tag{5.24}$$

and

$$[\mathbf{C}_2(\omega_j)]_{k,l} = \text{cum}(g_1^*(t,\omega_c), g_2(t,\omega_c), r_k(t,\omega_j), r_l^*(t,\omega_j)) = \frac{\gamma_{4,s}}{\sigma_s^2} E\{v_k(t,\omega_j), r_l^*(t,\omega_j)\} \tag{5.25}$$

where $v_k(t)$ is the process recorded by a virtual sensor, which is located at $\vec{d}$ away from the sensor that measures $r_k(t)$ ($\vec{d}$ is the vector from $g_1(t)$ to $g_2(t)$).

If $\mathbf{A}(\omega_j)$ denotes the steering matrix at frequency $\omega_j$, we have

$$\mathbf{C}_1(\omega_j) = \mathbf{A}(\omega_j)\mathbf{\Gamma}_{4,s}\mathbf{A}^H(\omega_j) \tag{5.26}$$

where $\mathbf{\Gamma}_{4,s} \triangleq \text{diag}(\gamma_{4,1}|g_{1,1}|^2, \ldots, \gamma_{4,P}|g_{1,P}|^2)$, in which $g_{1,k}$ represents the response of the first guiding sensor to the $k$th source at the center frequency. Similarly,

$$\mathbf{C}_2(\omega_j) = \mathbf{A}(\omega_j)\mathbf{\Phi}\mathbf{\Gamma}_{4,s}\mathbf{A}^H(\omega_j) \tag{5.27}$$

where $\mathbf{\Phi}$ is the diagonal matrix that contains the direction information as in the ESPRIT algorithm (see (2.37)). Using all the frequency bins ($L$ of them), we obtain two matrices

$$\mathbf{T}_1 = \begin{bmatrix} \mathbf{C}_1(\omega_1) \\ \mathbf{C}_1(\omega_2) \\ \vdots \\ \mathbf{C}_1(\omega_L) \end{bmatrix} = \underbrace{\begin{bmatrix} \mathbf{A}(\omega_1) \\ \mathbf{A}(\omega_2) \\ \vdots \\ \mathbf{A}(\omega_L) \end{bmatrix}}_{\triangleq \mathbf{B}} \mathbf{\Gamma}_{4,s}\mathbf{B}^H \tag{5.28}$$

$$\mathbf{T}_2 = \begin{bmatrix} \mathbf{C}_2(\omega_1) \\ \mathbf{C}_2(\omega_2) \\ \vdots \\ \mathbf{C}_2(\omega_L) \end{bmatrix} = \mathbf{B}\,\mathbf{\Phi}\,\mathbf{\Gamma}_{4,s}\,\mathbf{B}^H \tag{5.29}$$

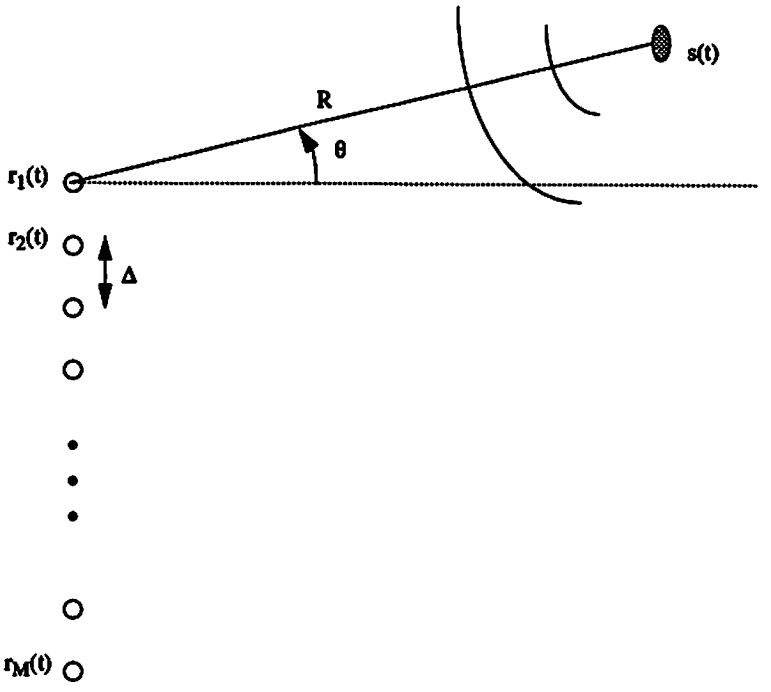



Figure 5.9: A uniform linear array in the near-field localization problem.

Now, we can use the ESPRIT algorithm to solve for the elements of $\mathbf{\Phi}$, which contain the direction information, and the effective steering matrix $\mathbf{B}$, which consists of the steering vectors for individual narrowbands over the spectral band of interest. The dimensionality of the cumulant matrices $\mathbf{T_1}$ and $\mathbf{T_2}$ can be increased along the lines of the previous section. Chapter 8 further investigates the use of higher-order statistics for wideband signals.

### 5.5.4 Near-Field Direction-Finding

In this section, we extend the virtual-ESPRIT algorithm to the case when the guiding sensors lie in the *Fresnel* region of the sources. This region is intermediate between the true near-field case where the wavefronts are spherical and the true far-field case where the wavefronts are very well approximated as planar.

Let the $i$th source be located at an unknown range $R_{i,1}$ from the first guiding sensor (reference point), with an unknown bearing $\theta_i$ with respect to the vector joining the two guiding sensors. Figure 5.9 illustrates an $M$ element uniform linear

array for the near-field problem, with interelement spacing $\Delta$. The distance from the $i$th source to the $k$th sensor is given by applying the law-of-cosines:

$$R_{i,k} = \sqrt{R_{i,1}^2 + (k-1)^2\Delta^2 - 2(k-1)\Delta R_{i,1}\sin\theta_i} \tag{5.30}$$

For sources in the far-field, $R_{i,1} \gg \Delta$, and $R_{i,k}$ is approximated by taking only the first two terms of the binomial expansion of (5.30):

$$R_{i,k} \simeq R_{i,1} - (k-1)\Delta\sin\theta_i \tag{5.31}$$

In many applications, the distances from the array to the emitters are on the order of only a few apertures, and the plane-wave approximation (5.31) is not valid. Retaining an additional term from the binomial expansion of (5.30) leads to an expansion that is quadratic in $k$:

$$R_{i,k} \simeq R_{i,1} - (k-1)\Delta\sin\theta_i + \frac{((k-1)\Delta\cos\theta_i)^2}{2R_{i,1}} \tag{5.32}$$

Equation (5.32) is called the *Fresnel* approximation; it models the spherical wavefronts as quadratic surfaces in the vicinity of the array; therefore, the received signal model for the linear array can be approximated as:

$$r_k(t) = \sum_{i=1}^{P} s_i(t)\alpha_i^{(k-1)}\,\beta_i^{(k-1)^2} + n_k(t) \tag{5.33}$$

where

$$\alpha_i(\theta_i) = \exp(-j(2\pi/\lambda)\Delta\sin\theta_i) \tag{5.34}$$

and

$$\beta_i(\theta_i, R_{i,1}) = \exp(j(2\pi/\lambda)\frac{(\Delta\cos\theta_i)^2}{2R_{i,1}}) \tag{5.35}$$

In this case, the steering vector for the uniform linear array takes the form:

$$\mathbf{a}(\theta_i, R_{i,1}) = [1, \alpha_i\beta_i, \alpha_i^2\beta_i^4, \ldots, \alpha_i^{(M-1)}\beta_i^{(M-1)^2}]^T \tag{5.36}$$

It is interesting to note that the array manifold for the uniform linear array does not conform directly to the displacement invariance structure required by the ESPRIT

algorithm since the Vandermonde structure of the steering vector is lost. Therefore, even for this simple array configuration a search over the array manifold is necessary. Starer and Nehorai [129] proposed a path following algorithm for localization of sources using a uniform linear array and parellelized computations. Swindlehurst and Kailath [138] proposed a spatial Wigner-Ville analysis (a heavy computational load) which is followed by the ESPRIT algorithm.

We now investigate how to calibrate arbitrary arrays with the virtual-ESPRIT algorithm using a guiding doublet that measures $\{g_1(t), g_2(t)\}$. If the location of the first guiding sensor is chosen as reference, the response of the doublet can be expressed as

$$[1, \alpha_i \beta_i] \tag{5.37}$$

which follows from viewing the guiding sensors as a uniform linear array with only two elements and using (5.36). If there is only one signal $s(t)$, then the autocovariance matrix required by ESPRIT can be computed as before by using cumulants.

$$\text{cum}(g_1^*(t), g_1(t), r_k(t), r_l^*(t)) = \frac{\gamma_4, s}{\sigma_s^2} E\{r_k(t), r_l^*(t)\} \tag{5.38}$$

A difference appears in the cross-correlation matrix required by ESPRIT, i.e.,

$$\text{cum}(g_1^*(t), g_2(t), r_k(t), r_l^*(t)) = \text{cum}(g_1^*(t), g_1(t), r_k(t), r_l^*(t)) \underbrace{e^{-j(2\pi/\lambda)\Delta \sin\theta_i}}_{\alpha} \beta \tag{5.39}$$

Due to near-field effects the additional term $\beta$ appears in (5.39) instead of $\alpha$ alone. This determines the generalized eigenvalue for the $i$th source provided by virtual-ESPRIT to be $\lambda_i = \alpha_i \beta_i$ instead of $\alpha_i$ alone; hence, the DOA can not be estimated as simply as in the far-field case; but, the steering vector estimates corresponding to the sources are not affected by this modification (since we assumed no modeling for them), which implies that signal recovery can be done without any difference from the far-field case. Using the definition of $\lambda_i$, the direction of the $i$th source can be found as the minimizer of:

$$\hat{\theta}_i = \arg\min_{R,\theta} \| \exp(-j(2\pi/\lambda)\Delta \sin\theta_i) \exp(j(2\pi/\lambda) \frac{(\Delta \cos\theta_i)^2}{2R_{i,1}}) - \lambda_i \| \tag{5.40}$$

The minimization is implemented for each source separately, and this parallelization provides fast throughput.

Our approach based on virtual-ESPRIT works even when the sources are very close in the near-field in which case the Fresnel approximation is not valid for the entire array but only for the guiding sensor part of the aperture that consists of only two elements.

# Chapter 6

# Minimum Redundancy Array Design

In this chapter, we first determine bounds on virtual aperture extension. We provide a summary on minimum redundancy arrays indicating design methods based on second-order statistics. Finally, we provide algorithms to design minimum-redundancy arrays based on cumulants.

## 6.1 Bounds on Aperture Extension

In this section, we derive lower and upper bounds for cumulant-based virtual-aperture extension as described in Chapter 4. The results are given for uncorrelated sources illuminating an array of identical sensors. The sensors can be at *arbitrary* locations. We are interested in identical sensor arrays because they play an important role in the design of minimum-redundancy antenna systems. The bounds are derived by letting the sensors have no area in the physical space, i.e., by representing them by dots; however, all the results apply to the real case, where the sensors consume a volume in three-dimensional space because the virtual aperture created by using actual sensor outputs and cumulants do not occupy any area either.

### 6.1.1 Lower Bound

Definition 1: Consider an antenna array of $M$ identical sensors, where the set of vectors $\{\mathbf{r}_k\}_{k=1}^{M}$ denote sensor locations. The *diameter* $d$ of such an array is defined as the maximum distance between any pair of the sensors, i.e.,

$$d \triangleq \max_{1 \le i,j \le M} \| \vec{r}_i - \vec{r}_j \|_F \tag{6.1}$$

Definition 2: A pair of sensors is called *limit-point sensors* if the distance between them is identical to the diameter ($d$) of the array. For a given array, there may be more than one pair of limit-point sensors.

Fact 1: For a given antenna array configuration, let $(\vec{r}_1, \vec{r}_2)$ denote the locations of a limit-point pair. The rest of the sensors must lie in the intersection of the following two spheres:

$$\begin{aligned} \mathbf{S}_1 &: \text{ centered at } \vec{r}_1, \text{ with radius } d, \\ \mathbf{S}_2 &: \text{ centered at } \vec{r}_2, \text{ with radius } d. \end{aligned} \tag{6.2}$$

Proof: If there is a single sensor that does not lie in the intersection of spheres defined above, then its distance to at least one of the limit-point sensors must be larger than $d$. But this implies the diameter of the array must be larger than $d$, which is a contradiction. Figure 6.1 gives an illustration. The intersection of the two spheres $\mathbf{S}_1, \mathbf{S}_2$ that contain all the actual sensors is defined as the *region of support.*

Theorem 1: Using the measurements from a limit-point pair, it is possible to extend the effective aperture from $M$ to at least $2M-1$, regardless of array geometry. The uniform linear array satisfies this lower bound.

Proof: Suppose that the limit-point sensor measurements are used as in the virtual-ESPRIT algorithm and let the vector from the first limit-point sensor to the second one be $\vec{d}$. Then, it is possible to generate a virtual array which is the shifted version of the actual array using $\mathbf{VC}^3$. The virtual sensors do not coincide with the actual sensors except for one of the limit-point sensor positions, due to Fact 1; hence, the effective aperture can be extended to $2M-1$ sensors. See Figure 6.2 for an illustration. The aperture extension result for the uniform linear array was also

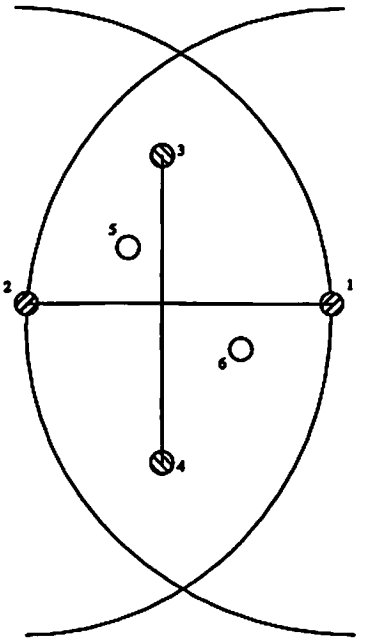


Figure 6.1: An antenna array of identical elements: the pairs (1,2) and (3,4) constitute limit-point sensors. All sensors must lie within the intersection of the two spheres. This intersection is defined as the *region of support.*

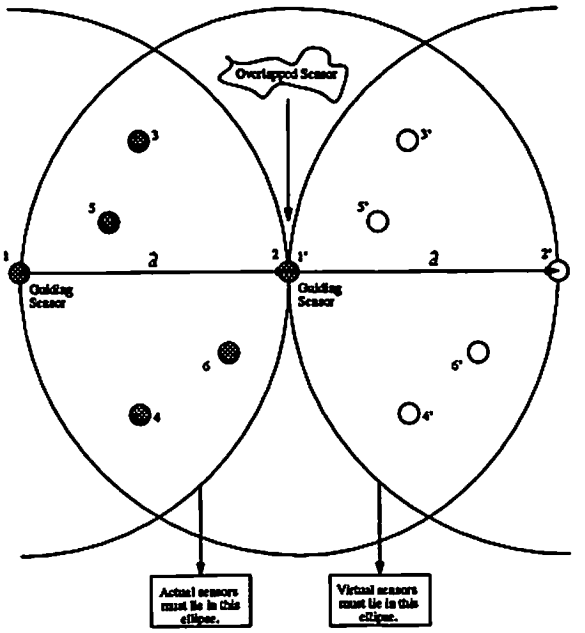



Figure 6.2: Aperture extension by using the limit-point sensors (sensors 1 and 2) as guiding sensors as in virtual-ESPRIT. Actual and virtual sensors must lie in distinct support regions due to Fact 1, and the two support regions intersect only at the guiding sensor 2. Each of the virtual sensors (e.g., 5') is obtained by shifting an actual sensor (e.g., 5) by the vector $\vec{d}$.

reported in [122] by using Carathedory's theorem. We displayed this result earlier in Section 4.2. Using the $\mathbf{VC}^3$ construction, it is very easy to obtain the result in [122].

Comment 1: Theorem 1 indicates that we can extend the aperture to at least $2M-1$ sensors when we know which sensors are the limit point sensors. This raises the issue of whether it is possible to accomplish this extension without this knowledge. We answer this question affirmatively in the next subsection, in the context of a direction-finding method proposed by Porat and Friedlander [99] (an equivalent method is described by Cardoso [21]).

### 6.1.2 Upper Bound

Consider an $M$ element array that measures the signals $\{r_1(t), r_2(t), \ldots, r_M(t)\}$, and consider cumulant matrices defined as

$$[\mathbf{C}(a,b)]_{k,l} = \text{cum}(\underbrace{r_a^*(t), r_b(t)}_{\text{guiding sensors}}, r_k(t), r_l^*(t)) \quad 1 \leq a,b,k,l \leq M. \tag{6.3}$$

Hence, $\mathbf{C}(a,b)$ is equivalent to the cross-correlation matrix of the actual array and its virtual copy which is shifted by the vector from $r_a(t)$ to $r_b(t)$. For example, in Figure 6.2, if we select $a = 1$ and $b = 2$, then, $\mathbf{C}(a,b)$ is cross-correlation matrix between the actual array and its virtual copy. Similarly, $\mathbf{C}(a,a)$ is equivalent to the autocorrelation matrix for direction-finding purposes.

Porat and Friedlander [99] propose to form the following $M^2 \text{x} M^2$ matrix:

$$\mathbf{C} = \begin{bmatrix} \mathbf{C}(1,1) & \mathbf{C}(1,2) & \ldots & \mathbf{C}(1,M) \\ \mathbf{C}(2,1) & \mathbf{C}(2,2) & \ldots & \mathbf{C}(2,M) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{C}(M,1) & \mathbf{C}(M,2) & \ldots & \mathbf{C}(M,M) \end{bmatrix} \tag{6.4}$$

If $\mathbf{a}_k$ denotes the steering vector of the $k$th source, then $\mathbf{C}$ can be decomposed as

$$\mathbf{C} = \sum_{k=1}^{P} \gamma_{4,k} (\mathbf{a}_k^* \otimes \mathbf{a}_k)(\mathbf{a}_k^* \otimes \mathbf{a}_k)^H \tag{6.5}$$

where $\gamma_{4,k}$ denotes the fourth-order cumulant of the $k$th source.

The following interpretation of the matrix $\mathbf{C}$ is very important: we fix one sensor (e.g., $r_1(t)$) to be one of the guiding sensors, and pick one of the remaining $M-1$ sensors (say $r_a(t)$) as the second guiding sensor, and extend the aperture as in virtual-ESPRIT using the matrix $\mathbf{C}(a,1)$ (which shifts the actual array by the vector from $r_a(t)$ to $r_1(t)$). Then, pick another sensor from the remaining $M-2$ sensors (say $r_b(t)$), and use it as the second guiding sensor by computing $\mathbf{C}(b,1)$. This explains the first block-column of $\mathbf{C}$. Of course, the cross-correlations between the two sets of virtual sensors must be computed, and it can be done by evaluating $\mathbf{C}(a,b) = \mathbf{C}^H(a,b)$. Figure 6.3 gives an illustration of this operation. For example, if we select sensor 1 as the reference sensor, and select sensor 4 as the second guiding sensor to compute $\mathbf{C}(4,1)$ and then select sensor 2 as the second sensor and compute $\mathbf{C}(2,1)$, these two sets of virtual sensors can be connected by the vector from sensor 4 to sensor 2, i.e., by computing $\mathbf{C}(4,2)$. This explains the remaining block-components of $\mathbf{C}$.

Theorem 2: The effective aperture provided by $\mathbf{C}$ is at least $2M-1$ elements. This result is valid regardless of the selection of the reference sensor among the $M$ actual sensors.

Proof: Let us pick an arbitrary sensor from the $M$ actual sensors and consider this sensor as the reference sensor (so it measures $r_1(t)$ in Figure 6.4). Then, extend the aperture by shifting the actual array by using the two vectors between the reference point and a pair of limit-point sensors $\{r_a(t), r_b(t)\}$. Virtual sensors obtained by this process only coincide at the reference point (where there exists an actual sensor), because actual sensors must lie in the original region of support, which forces the virtual sensors to lie in the shifted regions of support and in this construction, these two shifted regions intersect only at the reference point. Therefore, if we only count these sensors as the effective aperture (even without counting the actual sensors), then the lower bound becomes $2M-1$. The subtraction is due to the overlap at the reference point.

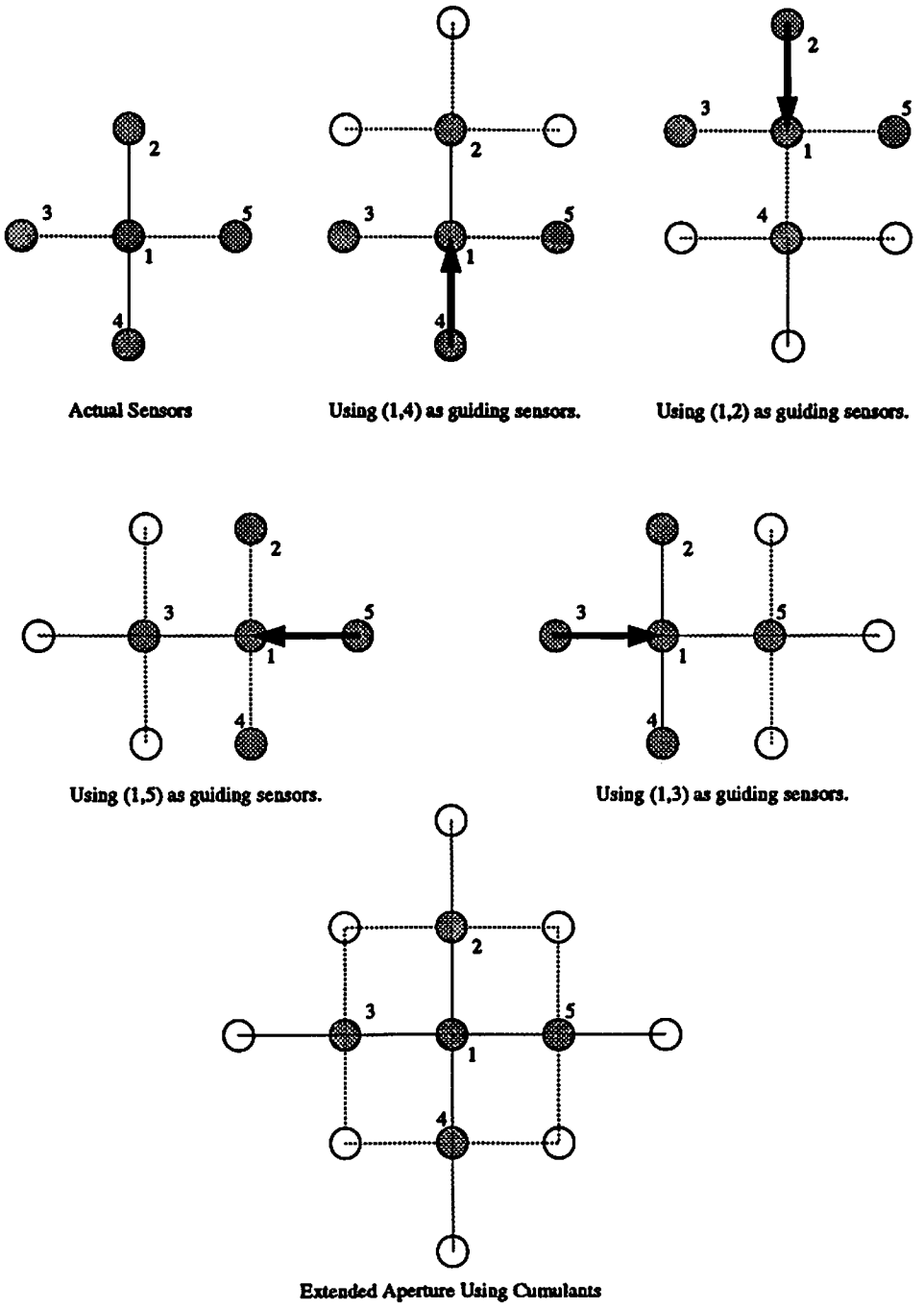



Figure 6.3: An interpretation of the matrix construction proposed in (6.4). We select sensor 1 as the reference sensor, and shift the actual array by the vectors from the other actual sensors to the reference sensor to obtain the effective aperture. Note that there exist virtual sensors at all actual sensor locations and this *redundancy* decreases the capacity to extend the effective aperture.

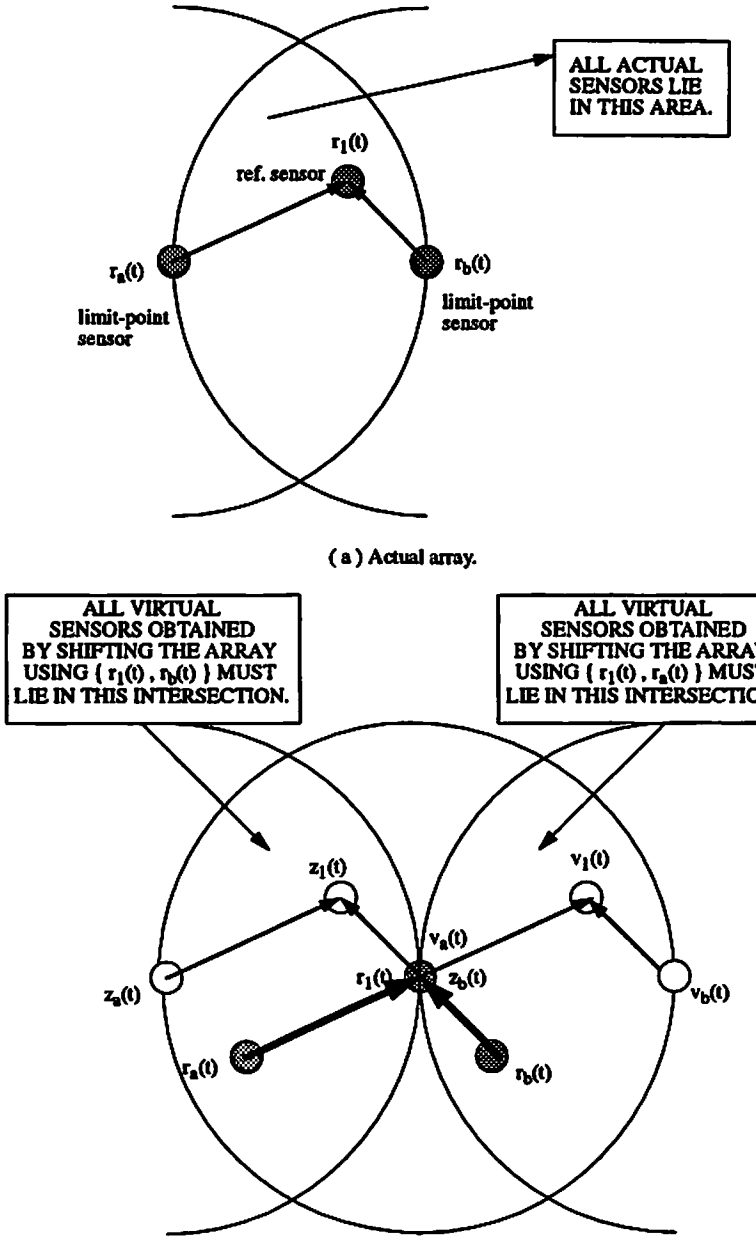



Figure 6.4: The lower bound on aperture extension: (a) Three elements of the $M$ element actual array are represented by filled circles. The rest of the elements must lie in the support region as pointed out in Fact 1. (b) By aperture extension, the support regions can only overlap at the reference sensor. Here, $v_x(t)$ $(1 \leq x \leq M)$ is the virtual sensor created by shifting the actual sensor $r_x(t)$ by the vector between $r_a(t)$ and $r_1(t)$. Similarly, $z_x(t)$ is the virtual sensor created by shifting the actual sensor $r_x(t)$ by the vector between $r_b(t)$ and $r_1(t)$.

Fact 2: The effective aperture provided by C does not depend on the choice of the reference sensor. This is obvious from the structure of the effective steering vectors, which take the form:

$$\mathbf{a}^* \otimes \mathbf{a} = [a_1^* a_1, a_1^* a_2, \ldots, a_1^* a_M, a_2^* a_1, a_2^* a_2, \ldots, a_2^* a_M, \ldots, a_M^* a_1, a_M^* a_2, \ldots, a_M^* a_M]^T \tag{6.6}$$

Suppose we change the labels on the first and second sensors, i.e., the second sensor is named as the first, etc. Then, the effective steering vector for this labeling takes the form:

$$\tilde{\mathbf{a}}^* \otimes \tilde{\mathbf{a}} = [a_2^* a_2, a_2^* a_1, \ldots, a_2^* a_M, a_1^* a_2, a_1^* a_1, \ldots, a_1^* a_M, \ldots, a_M^* a_2, a_M^* a_1, \ldots, a_M^* a_M]^T \tag{6.7}$$

The elements of the vector in (6.7) are a permutation of the elements of the effective steering vector in (6.6), which means a relabeling of the actual sensors result in a relabeling of the virtual sensors.

Comment 2: Fact 2 provides an alternate way to prove Theorem 2. Because the choice of the reference point does not affect the effective aperture, we can choose one of the limit-points as the reference sensor. When the other limit-point sensor is chosen as a guiding sensor to shift the array, the resulting virtual sensors that do not overlap with the actual ones will be $M-1$, as stated in Theorem 1, and displayed in Figure 6.2. Then, the effective aperture consists of at least $2M-1$ elements.

Theorem 3: The effective aperture provided by C can be at most $M^2 - M + 1$ elements.

Proof: After choosing the reference sensor (this choice does not affect the result, due to Fact 2), we shift the array as in virtual-ESPRIT by using the reference sensor as a guiding sensor and picking another sensor (say $r_a(t)$) from the remaining $M-1$ sensors as the other guiding sensor. With each such shift, there is at least one overlap between the virtual sensor locations and the actual ones (the reference sensor position coincides with the virtual sensor location created by the shift of $r_a(t)$, see Figure 6.3). Hence, each shift adds at most $M-1$ virtual sensors. We can make $M-1$ such shifts, that results in a total of $(M-1)^2 = M^2 - 2M + 1$ virtual sensors. Adding the number of actual sensors, $M$, to the number of virtual sensors, we obtain the upper bound for the number of effective sensors as $M^2 - M + 1$. The number

of the effective sensors clearly depends on the array geometry, but it is always lower bounded by $2M-1$. The result reported in Theorem 3 is also claimed in [21] without proof. The remaining question is whether there exist a class of arrays which always achieve the upper bound. We shall provide an affirmative answer in the context of two-dimensional array design later in this chapter, in Section 6.3.1.

## 6.2 Minimum Redundancy Arrays (MRA)

The structure of the array covariance matrix in the case of incoherent sources illuminating a uniformly spaced linear array of identical sensors has led to algorithms that can estimate the directions of more sources than sensors. This section is devoted to reviewing this concept within the framework developed in the previous chapter for the role of cumulants in array processing.

Consider a uniformly spaced antenna array of $M$ identical sensors which is illuminated by $P$ incoherent sources with waveforms $\{s_1(t),\ldots,s_P(t)\}$

$$x_k(t)=\sum_{l=1}^{P}s_k(t)\exp(-j(k-1)\sin(\theta_l))+n_k(t) \tag{6.8}$$

where the noise components $\{n_1(t),\ldots,n_M(t)\}$ are uncorrelated with the same variance $\sigma^2$. In (6.8) we assume the sensor separation is a half-wavelength to eliminate additional parameters. We also assume that noise components are independent of signals. After these assumptions, the cross-covariance between sensor outputs can be expressed as

$$E\{x_m(t)x_n^*(t)\}=\sum_{l=1}^{P}\sigma_k^2\exp(-j(m-n)\sin(\theta_l))+\sigma^2\delta(m-n) \tag{6.9}$$

where $\sigma_k^2$ is the power of the $k$th far-field source, and $\delta(m-n)$ is the Kronecker delta function which is unity if and only if $m=n$. Clearly, (6.9) indicates that the cross-covariance between two sensors can be interpreted as an integer which is the difference of their locations, i.e., $E\{x_m(t)x_n^*(t)\}=r_{m-n}$. This is also obvious from our previous results: covariance can be interpreted as a vector between two sensors, and when the sensors are constrained on a line, the sign and the magnitude of this

vector is sufficient to represent cross-covariance. In [95], this interpretation for cross-covariance is proved by using the Caratheodory theorem. We have extended this interpretation to fourth-order cumulants: cumulants are addition of two "covariance" integers.

From (6.9), the array covariance matrix takes the form

$$\mathbf{R} = \mathbf{A}\,\mathbf{\Sigma}_{ss}\,\mathbf{A}^H + \sigma^2\mathbf{I} \tag{6.10}$$

in which $\mathbf{A}$ is the $M$x$P$ steering matrix and its columns are the steering vectors. The diagonal matrix $\mathbf{\Sigma}_{ss}$ contains the source powers. The steering vectors take the form

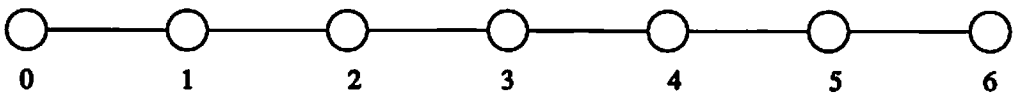
$$\mathbf{a}(\theta) = [1, \exp(-j\sin(\theta)), \ldots, \exp(-j(M-1)\sin(\theta))]^T \tag{6.11}$$

and as a result, the steering matrix possesses a Vandermonde structure. Furthermore, since signals are independent and because of the structure of the steering vectors, the array covariance matrix $\mathbf{R}$ is Toeplitz, i.e.,

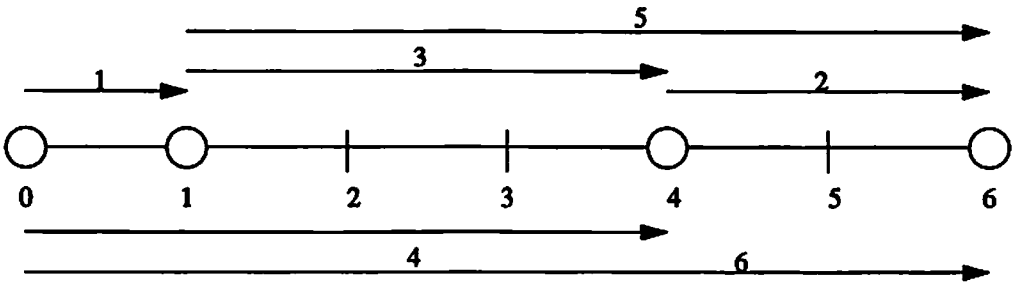
$$\mathbf{R} = \begin{bmatrix} r_0 & r_1^* & r_2^* & \cdots & r_M^* \\ r_1 & r_0 & r_1^* & \cdots & r_{M-1}^* \\ r_2 & r_1 & r_0 & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ r_M & r_{M-1} & r_{M-2} & \cdots & r_0 \end{bmatrix} \tag{6.12}$$

which implies that if we can compute the set of covariances $\{r_0, r_1, \ldots, r_M\}$, then we can reconstruct $\mathbf{R}$ due to the Toeplitz property. For an arbitrary scenario, a $K$x$K$ covariance matrix has $K(K+1)/2$ parameters due to Hermitian property. When the covariance matrix is constrained to be Toeplitz, the number of free parameters reduces to $K$. This highlights the redundancy in the uniform linear array: we do not need all the sensors in order to compute the set of necessary covariances ((M+1) of them). To eliminate redundancy, we must remove some of the sensors, such that we can still compute $\mathbf{R}$ from the remaining sensors as if the number of actual elements remained the same.

The problem of eliminating redundancy by exploiting the special signal model in (6.8) and the array covariance matrix (6.12) can be restated as the following




(a)




(b)

Figure 6.5: Elimination of redundant sensors in a uniform linear array where identical sensors are separated by $\lambda/2$: (a) Redundant array with seven sensors; (b) Nonredundant array obtained by removing 3 sensors; as shown, it is possible to generate all integers from zero to six, which means that the 7x7 covariance matrix of (a) can be constructed using only four sensors.

combinatorial problem [73]: "Represent $1, 2, \ldots, N$ by differences of $M$ $(M < N)$ non-negative integers drawn from non-negative natural numbers up to $N$." In this problem, $N$ corresponds to the number of sources that can be resolved. The mathematical description of this problem is as follows:

- Find a set of integers $\{d_1, d_2, \ldots, d_M\}$ such that every positive integer $k$ $(0 < k \leq N)$ can be represented in the form $k = d_i - d_j$ with the constraint

$$0 = d_1 < d_2 < \cdots < d_M = N \tag{6.13}$$

The set of integers that satisfy (6.13) is called the *restricted difference basis* with respect to $N$ [73]. The minimum number of integers in the restricted difference basis for given $N$ is denoted as $M_N$. In the array processing context, $M_N$ corresponds to the minimum number of elements to estimate the directions of $M$ sources, and

the difference basis indicates the sensor locations. Similarly, given $M$, $N_M$ denotes the length of the longest restricted difference basis. Figure 6.5 gives an example for the case with $N = 6$ in which the restricted difference basis $\{0, 1, 4, 6\}$ is used. Table 6.1 illustrates the optimal locations of the sensors for minimum redundancy up to $M = 17$ [93]. These locations were obtained by exhaustive search methods whose running time grows exponentially with the size of the array.

Finding the minimum number of elements to construct restricted sets and their locations has been of interest for a long time and results are published in different contexts [36, 59, 61, 71, 72, 73, 75, 79, 95, 146, 154]. The first important result is that given $M$ integers, it is possible to generate at most $\binom{M}{2} = M(M-1)/2$ integers that satisy (6.13) which are not necessarily unique; hence, an upper bound for $N_M$ is:

$$N_M \leq \frac{M(M-1)}{2} \tag{6.14}$$

Bracewell [6] showed that the strict inequality holds for $M > 4$. Leech [73] provided the asymptotic result

$$2.434\ldots \leq \lim_{N\to\infty} \frac{M_N^2}{N} \leq \frac{375}{112} = 3.348\ldots \tag{6.15}$$

More importantly Pearson *et* al. [93] proved the following result by providing a *constructive method* to generate a restricted difference basis:

• For any given $M > 3$, it is always possible to chose a restricted difference basis such that

$$M^2/N < 3 \tag{6.16}$$

which implies $M_N^2/N < 3$. Note that this is not just an asymptotic result. For $M = 3$, $M^2/N = 3$ (see Table 6.1). We refer the reader to [93] or [97, Chapter 2] for the proof of (6.16). It is important to note that asymptotically it is not possible to find configurations that use 10 percent fewer elements than the design procedure of [97, Chapter 2]. This follows from the lower bound (6.15) of Leech [73].

There are also improved search procedures for finding the optimal configuration of antennas. A recent approach [109] uses a numerical annealing method for the search and provides identical results with the exhaustive search method up to $M = 17$.

Table 6.1: Interelement spacing for optimal restricted difference basis determined by exhaustive search [93] ($M$ is the number of sensors and $N_M$ is the array length). Note that the first sensor is always located at the origin ($d_1 = 0$).

| $M$ | $N_M$ | Interelement Spacing | $M^2/N_M$ |
|---|---|---|---|
| 3 | 3 | •1•2• | 3.0 |
| 4 | 6 | •1•3•2• | 2.667 |
| 5 | 9 | •1•3•3•2• | 2.778 |
| 6 | 13 | •1•1•4•4•3• | 2.769 |
| 7 | 17 | •1•1•4•4•4•3• | 2.882 |
| 8 | 23 | •1•3•6•6•2•3•2• | 2.783 |
| 9 | 29 | •1•4•4•7•7•3•2•1• | 2.793 |
| 10 | 36 | •1•4•4•7•7•7•3•2•1• | 2.778 |
| 11 | 43 | •1•4•4•7•7•7•7•3•2•1• | 2.814 |
| 12 | 50 | •1•4•4•7•7•7•7•7•3•2•1• | 2.88 |
| 13 | 58 | •1•4•3•4•9•9•9•9•5•1•2•2• | 2.914 |
| 14 | 68 | •1•1•6•6•6•11•11•11•5•5•3•1•1• | 2.882 |
| 15 | 79 | •1•1•6•6•6•11•11•11•11•5•5•3•1•1• | 2.848 |
| 16 | 90 | •1•1•6•6•6•11•11•11•11•11•5•5•3•1•1• | 2.844 |
| 17 | 101 | •1•1•6•6•6•11•11•11•11•11•11•5•5•3•1•1• | 2.861 |

Additional results are also given for $M > 17$, but there is no proof that these designs are optimal.

The minimum redundancy array concept is only applicable to linear arrays, which can only provide the azimuth of the far-field sources. In many applications, both azimuth and elevation information is necessary; this requires at least a two-dimensional array configuration. Bracewell [6] has noted that there is no two-dimensional analog to the minimum redundancy arrays except for the case of a four-element "T" shaped array. Because of the lack of a solution, two-dimensional resolution can only be obtained by linear minimum redundancy arrays if the array is physically rotated. For example, in astronomy the *earth rotation synthesis* technique of [111] exploits the rotation of the earth with respect to the far-field source, which in turn changes the orientation of the linear array. The collection times are very large, and the performance depends on the actual direction of the source. In addition, during

the rotation time of the array, the source is assumed to be stationary. These assumptions will never hold in small-scale applications such as radar and sonar. In the next section, we show methods to design two-dimensional minimum redundancy arrays based on the interpretation of cumulants and already existing design procedures for linear minimum redundancy arrays (see Table 6.1). We also propose cumulant-based minimum redundancy linear arrays based on cumulants which can significantly outperform the existing covariance-based designs.

## 6.3 Cumulant-Based MRA Design

In this section, we first design two-dimensional minimum-redundancy arrays based on the existing one-dimensional covariance-based designs. We then concentrate on cumulant-based design of linear arrays.

### 6.3.1 Two-Dimensional Arrays

Suppose we wish to design a rectangular array so that both azimuth and elevation estimation is possible through the use of subspace rotation techniques (ESPRIT). Specifically, if we have $M_x M_y$ elements, then it is possible to generate a rectangular grid of length $M_x$ and width $M_y$ and put the sensors in the intersection points. For example, if we have 49 sensors, then a square array can be generated by using the intersections of the grids in Figure 6.6. To implement a direction-finding algorithm all the connections between the sensors must be completed by a single vector between actual sensors. Any vector in the rectangle can be decomposed into two parts: x-component and a y-component. If one uses fourth-order cumulants, then it is possible to add two vectors between actual sensors to obtain virtual cross-correlations; hence, the 2-D design problem reduces to designing two linear minimum-redundancy arrays which can cover the axes of the rectangle. This design was described in the previous section (the restricted difference basis). Figure 6.6 illustrates the design procedure with four actual sensors per axis. It is possible to extend the aperture to 49 sensors by using 7 actual elements. Note that this design breaks the bound of the existing linear minimum-redundacy arrays (see Bracewell's bound in (6.14), which upper bounds the number of resolvable sources to be 21).

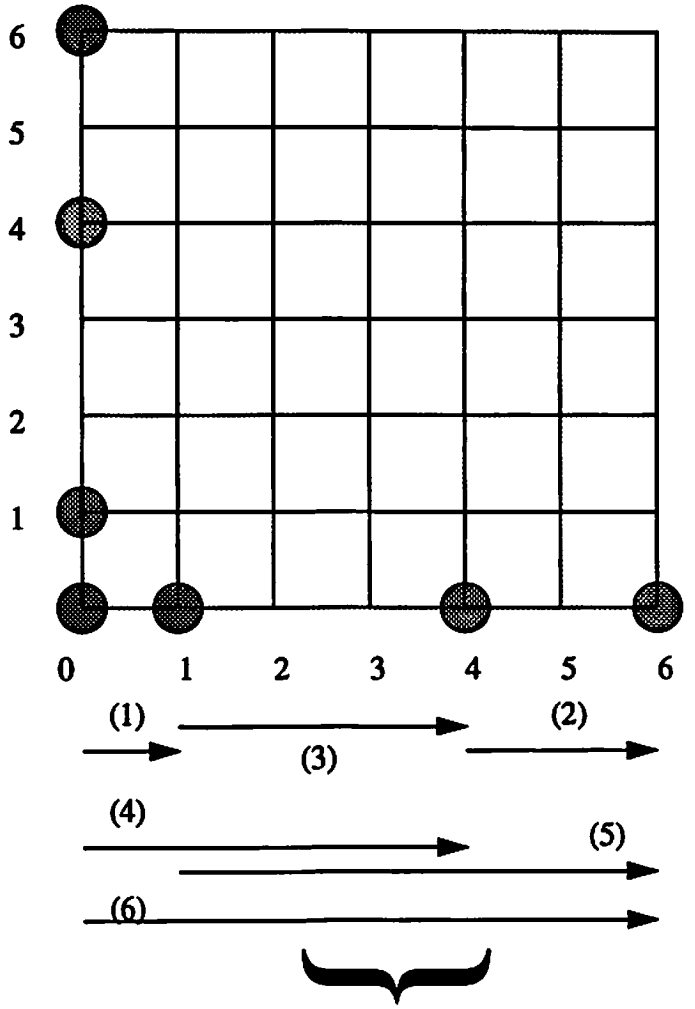



Figure 6.6: 2-D minimum redundancy array design: combining two 1-D arrays and using $VC^3$ allows the computation of cross-correlations among all the 49 grid points, using the data available just from the 7 real sensors (filled circles).

Our approach provides both two-dimensional resolution and more virtual sensors than covariance-based designs. Given a number of actual sensors, the best displacement is the one in which the extended aperture will be close to a square. The reason is that the perimeter of the rectangle is proportional to the number of actual elements, and the area of the rectangle is proportional to the number of sensors in the extended aperture. The minimum redundancy problem requires us to maximize the aperture (area) while the number of actual sensors (perimeter) is kept constant and the solution is the special rectangle; namely, the square. If the number of actual sensors $M$ is odd, then we can use $(M+1)/2$ sensors per axis to form a linear minimum-redundancy array, whose length is lower bounded by $(M+1)^2/12$, which comes from the constructive method of Pillai [97]. The total number of elements in the square aperture is the square of this number; hence, the effective aperture is lower bounded by $(M+1)^4/144$. Figure 6.7 gives a comparison of our two-D method and the optimal linear array. Our two-D method provides more virtual sensors than the conventional minimum-redundancy array based on second-order statistics. Next, we describe the general design steps for our approach (CUM-REC) given the number of sensors $(M)$:

1. If $M$ is even, then let $M_x = (M+2)/2$ and $M_y = M/2$. If $M$ is odd, then let $M_x = M_y = (M+1)/2$.

2. Decompose the two-dimensional design problem into two separate linear array design problems, with number of sensors equal to $M_x$ and $M_y$, respectively.

3. Use Table 6.1, or the extended results reported in [109] to find the solutions to the linear array design problems. Let $N[M_x]$ denote the effective number of sensors with $M_x$ actual sensors, and define $N[M_y]$ similarly.

4. Put the array with $M_x$ actual elements on the x-axis with the locations obtained from the previous step. Put the array with $M_y$ actual elements on the y-axis such that the first actual elements of both arrays coincide.

5. It is now possible to compute the cross-correlations between any pair of sensors located at the grid points of the rectangular structure of area $N[M_x]N[M_y]$ by using the $\mathbf{VC}^3$. An example was given for the $M = 7$ case in Figure 6.6.

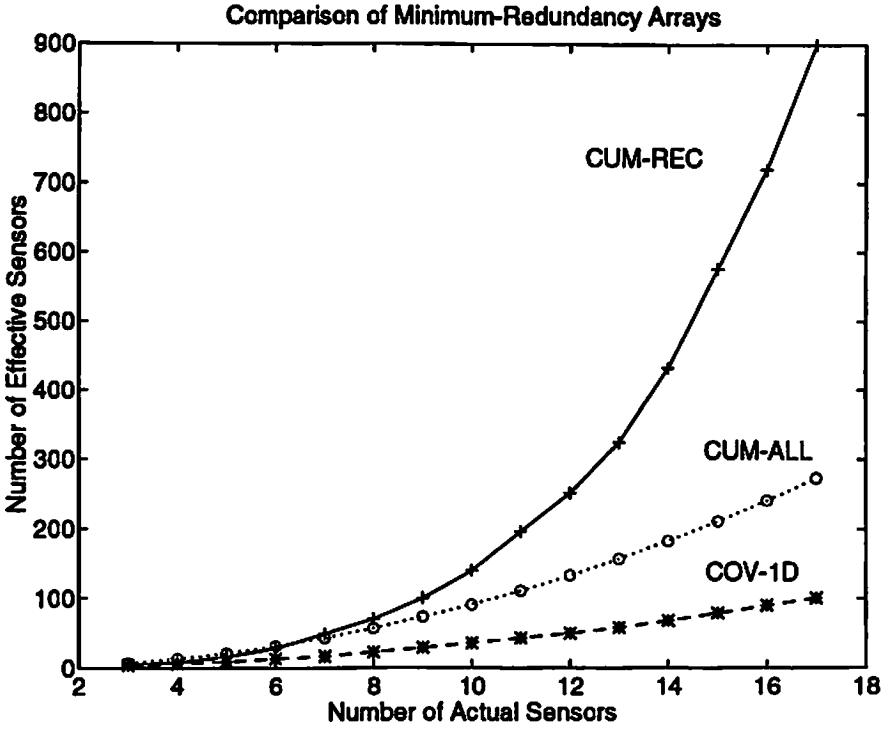



Figure 6.7: CUM-REC outperforms other designs and provides two-dimensional resolution. CUM-ALL also represents the upper bound of Section 6.1.2.

Based on the design procedure described above, we compare the covariance-based MRA and the cumulant-based two-dimensional MRA (CUM-REC) in Table 6.2 for a small number of actual sensors. In this table, CUM-ALL indicates the effective aperture generated by the sensors designed for CUM-REC, but whose measurements are processed by the algorithm in Section 6.1.2, by forming the matrix $\mathbf{C}$ in (6.4). An example with four actual sensors is given in Figure 6.8 for CUM-ALL. COV-1D is the effective sensors from the one-dimensional covariance-based MRA design. In Table 6.2, we observe that when the number of available sensors is less than 7, CUM-REC provides less sensors than CUM-ALL, and does not provide dramatic improvement over COV-1D. Particularly, CUM-REC results in an array whose effective number of sensors is less than the upper bound $(M^2 - M + 1)$ without the MRA design concept if $M < 7$, e.g., with six actual sensors the effective number of sensors from CUM-ALL and CUM-REC are 31 and 24 respectively. The reason

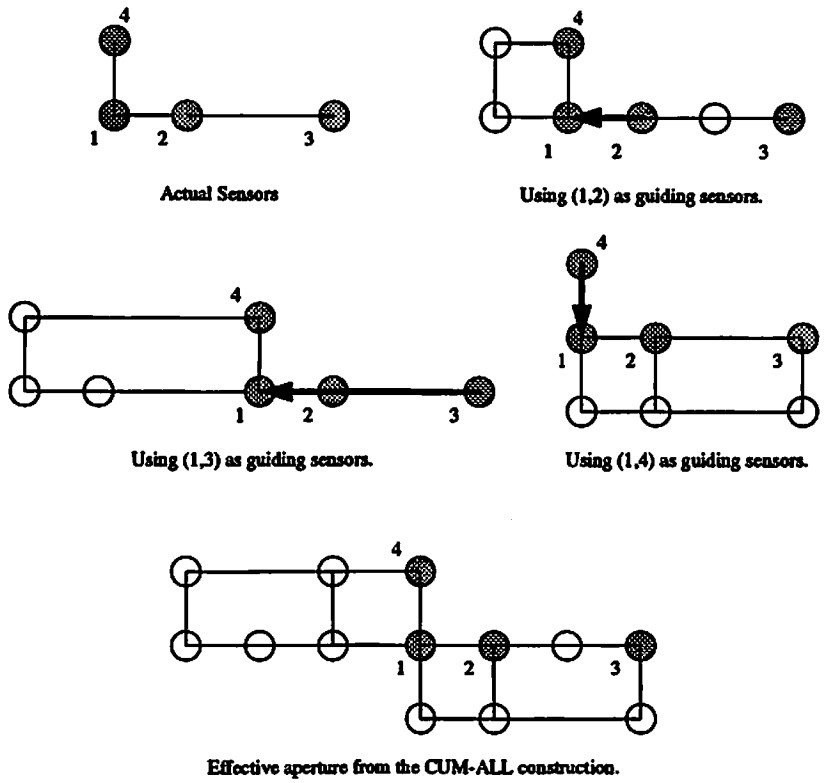



Figure 6.8: An example for aperture extension with CUM-ALL: there are four actual sensors: $M_x = 3$ and $M_y = 2$ from the CUM-REC procedure. Sensor 1 is selected as the reference sensor.

for this is that the partitioning of the available elements reduces the number of sensors per axis to less than four, which in turn causes an inefficiency by insisting on a rectangular shaped array. If we can guarantee that there exists a class of array configurations that meet the upper bound $M^2 - M + 1$, then, we can use the matrix C defined in (6.4) to increase the aperture for $M < 7$. We now show that Costas arrays achieve the upper bound.

Definition 3: For each positive integer $M$, construct an $M$x$M$ permutation matrix with the property that for all possible x-y shift combinations at most one pair of ones coincide with the unshifted matrix [30]. The resulting structure is called a *Costas array*. An alternate definition is to construct an $M$x$M$ permutation matrix with the property that the $\binom{M}{2}$ vectors connecting two 1's of the matrix are all distinct

Table 6.2: Comparison of aperture extension for small number of actual sensors.

| Actual Sensors | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|
| CUM-REC | 4 | 8 | 16 | 28 | 49 |
| CUM-ALL | 7 | 13 | 21 | 31 | 43 |
| COV-1D | 4 | 7 | 10 | 14 | 18 |

vectors [52]. An example is given for the case $M = 6$ in Figure 6.9.

Theorem 4: If the $M$ sensors are placed at locations determined by a Costas array of order $M$, then the effective aperture consists of sensors $M^2 - M + 1$ for the direction-finding method based on processing $\mathbf{C}$ in (6.4).

Proof: The virtual array is obtained by shifting the actual array by vectors between the actual sensors. In Theorem 3, we indicated that in each shift operation, one of the virtual sensors coincides with the reference sensor, regardless of the sensor that is chosen to shift the array. Therefore, by definition of the Costas arrays, there can be no other overlaps, i.e., each shift creates $M - 1$ more virtual sensors. There are $(M - 1)$ possible shifts, hence $(M - 1)^2$ virtual sensors that do not overlap with each other and the actual sensors. When we add to $(M - 1)^2$, the number of actual sensors $M$, we obtain $M^2 - M + 1$.

### 6.3.2 Linear Arrays

An initial attempt to design linear MRA's based on cumulants is to take an existing design from Table 6.1, and recall that cumulants are addition of two integers, and, can be used to double the aperture. Therefore, to obtain an effective aperture of $N$ units long, we need to look for a covariance-based design for an aperture $N/2$ units long; hence, replacing $N$ by $N/2$ in (6.16), we determine the upper bound for this simple cumulant-based procedure as

$$\frac{M^2}{N} < 3 \cdot \frac{1}{2} \tag{6.17}$$

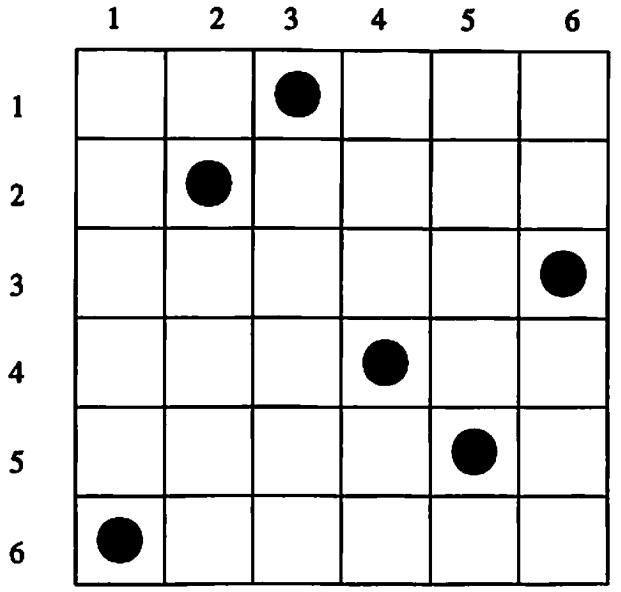



Figure 6.9: An example of a Costas array of order six.

Equation (6.17) implies that the effective aperture ($N$) is still proportional to the square of the number of actual sensors. Cardoso [21] also claimed that linear MRA's can be designed with an effective aperture of $M^2 - M + 1$ sensors, which is slightly better than (6.17). Cardoso did not prove whether or not the resulting effective aperture has "holes", i.e., whether it constitutes sensors configured as a uniform linear array with no missing sensors.

Of course these results are somewhat primitive because we already showed that it is possible to design two-dimensional MRA's (a harder problem than the linear array problem) whose effective aperture is proportional to the fourth power of the number of the actual sensors. However, for illustrative purposes we start with a simple method which generates similar primitive results without using existing covariance-based MRA results.

We know that cumulants are addition of two integers in the linear array problem; hence, if we have all the odd numbers, then we can generate even numbers by adding existing odd numbers. Let us place our sensors with this idea in mind: let the first sensor be the reference element; put the second sensor one unit to the right of the first one; put the third sensor three units to the right of the second one; put the

fourth-sensor five units to the right of the third sensor; etc. With six sensors, this placing method results in the following array:

$$\bullet 1 \bullet 3 \bullet 5 \bullet 7 \bullet 9 \bullet \tag{6.18}$$

with the sensor locations $\{0, 1, 4, 9, 16, 25\}$. This method actually puts the sensors at the squares of integers. This observation can be proved as follows: if $M$ actual sensors are available then the location of the last sensor is

$$\sum_{k=1}^{M-1} (2k-1) = (M-1)^2, \quad 1 \leq M \tag{6.19}$$

The next thing to show is that the effective aperture that can be created using cumulants has *no holes* in it, at least for the range $0 < k < (M-1)^2$ (end points are guaranteed to be sensor locations, by design). This can be proved by induction as follows: for $M = 1$ or $M = 2$, it can be proved by observing (6.18). Then, for any $M$, we observe that the last two elements of the $M$ element array are located at $\{(M-2)^2, (M-1)^2\}$, and these elements are separated by $2M-3$ units (just subtract the numbers). By construction (see (6.18)), we have the odd numbers $\{1, \ldots, 2M-3\}$ available between the actual sensors. Now, all the even (odd) numbers between $\{(M-2)^2, (M-1)^2\}$ can be generated by adding the appropriate number from the set of odd numbers to $(M-2)^2$ if $M$ is odd (even). Similarly, all the odd (even) numbers between $\{(M-2)^2, (M-1)^2\}$ can be obtained by subtracting the appropriate number from the odd numbers set from $(M-1)^2$ if $M$ is odd (even). For example, if $M = 6$ (even), the last two sensors are at 16 and 25. An odd number in this range can be generated by adding 16, an integer between the actual sensors (see (6.18)), e.g., 21=16+5. Similarly, for the even numbers, we subtract from 25, e.g., 22=25-3. This completes the proof that the effective aperture has no holes in it for $0 < k < (M-1)^2$.

Actually, this simple design can extend the aperture to more than $(M-1)^2$ sensors. For example, using a four element design with sensor locations $\{0, 1, 4, 9\}$, it is possible to create effective sensors at the locations $\{0-14, 16-18\}$, with a sensor missing at 15 (a total of 18 sensors). Our aim was not to find the exact

number of effective sensors for this design, but to show how to reach claims made at the beginning of this section.

We now propose a cumulant-based linear MRA design method (CUM-LIN) that uses results from covariance-based designs, and achieves an effective aperture proportional to the fourth-power of the number of actual sensors. The proposed method is actually a mapping of our two-dimensional array design method to a single dimension. The CUM-LIN method is described below:

1. Given $M$ actual sensors, divide them into two groups: the first group has $M_x$ sensors and the second group has $M_y$ sensors. If $M$ is even, then let $M_x = (M+2)/2$ and $M_y = M/2$. If $M$ is odd, then let $M_x = M_y = (M+1)/2$. This step is identical to the first step of the two-dimensional method, however, the ordering $M_y \leq M_x$ is important in the linear MRA design problem and is explained in the last item of this list.

2. Given $M_x$ actual sensors, design a one-dimensional MRA based on the results presented in Table 6.1. Let $N[M_x]$ denote the effective length of the array (one less than the number of sensors) using second-order statistics.

3. By adding two vectors between the $M_x$ actual sensors located in the previous step (using cumulants of measurements), it is possible to extend (double) the aperture to have a length of $2N[M_x]$. Therefore, if we put new sensors at multiples of $2N[M_x]+1$, then we can generate all integers using the $M_x$ element array of Step 2 and these new sensors. For example, if $k = l(2N[M_x]+1)+m$, where $0 \leq k$ and $0 \leq m \leq 2N[M_x]$, then $k$ can be represented by addition of two vectors between actual sensors; if $m \leq M_x$, then $m$ can be obtained as the difference between the locations of the $M_x$ element array of Step 2, if $M_x < m \leq 2N[M_x]$, then we rewrite $k = (l+1)(2N[M_x]+1)-(2N[M_x]+1-m)$, where the second term $(2N[M_x]+1-m)$ is not larger than $N[M_x]$ and hence can be obtained as the difference between the locations of the $M_x$ element array of Step 2.

4. The next step is to minimize the number of sensors which are separated by $2N[M_x]+1$ units, but maintain the maximum length possible. This is no different than a covariance-based MRA design problem with $M_y$ sensors, but

the separation between each sensor is a superunit, which is defined as $2N[M_x]+1$ units. The first element of this array coincides with the first element of the $M_x$ array. The design can be done by using Table 6.1. Let $N[M_y]$ be the length of the MRA array from Table 6.1.

5. Clearly, it is possible to obtain an array of length $N[M_y](2N[M_x]+1)$ by using cumulants and the actual sensors deployed as stated in the previous steps. We can also generate integers from $N[M_y](2N[M_x]+1)$ to $N[M_y](2N[M_x]+1)+N[M_x]$ by adding vectors from the $M_x$ element array to the last element of the $M_y$ element array which is located at $N[M_y](2N[M_x]+1)$. Finally, the difference between the last element of the $M_x$ element array (with location $N[M_x]$) and the second element of the $M_y$ element array (located at $2N[M_x]+1$) is $N[M_x]+1$, which can be added to the last element of the $M_y$ element array to obtain an effective aperture of length $L[M_x, M_y] = N[M_x](2N[M_y]+1)+N[M_y]+1$ without holes. The expression for $L([M_x, M_y]$ explains our selection for $M_y \leq M_x$ in Step 1.

An example is given below for the $M = 7$ case. We let $M_x = M_y = 4$ and find the locations of the $M_x$ array (from Table 6.1) as $\{0,1,4,6\}$, which indicates that $N[M_x] = 6$. Now, a superunit is 13 units (i.e., $2 \cdot 6 + 1$). We design the $M_y$ array by using Table 6.1, and multiplying the results by a superunit, i.e., locations are $13 \cdot \{0,1,4,6\} = \{0,13,52,78\}$; hence, the locations of all 7 sensors are $\{0,1,4,6,13,52,78\}$. Integers from 79 to 84 can be obtained by adding the differences between the $M_x$ array to 78. Finally, 85 can be obtained as $(78-0)+(13-6)$. The aperture is 85 units long, or it consists of 86 sensors. For comparison, CUM-REC provided 49 sensors, and COV-1D provided 18 sensors. In Table 6.3 we provide designs for $3 \leq M \leq 17$. In Table 6.4 we compare the length of the effective aperture with that of COV-1D and previously described two-dimensional designs. The results from Table 6.1 can be used to design CUM-LIN and CUM-REC arrays for up to $M = 34$. For $M > 34$ the results from [109] can be used.

From the results in Table 6.4, we observe that for large $M$, CUM-LIN provides twice the number of effective sensors that CUM-REC can provide. This is due to ease of designing linear arrays as compared to two-dimensional arrays. This observation can be proved as follows: let $M$ be odd, so that $M_x = M_y = (M+1)/2$. Then

Table 6.3: Cumulant-based linear MRA design (CUM-LIN) for $3 \leq M \leq 17$.

| Actual Sensors $M$ | $M_x$ ARRAY Sensor Locations | $M_y$ ARRAY Sensor Locations |
|---|---|---|
| 3 | $\{0,1\}$ | $\{0,3\}$ |
| 4 | $\{0,1,3\}$ | $\{0,7\}$ |
| 5 | $\{0,1,3\}$ | $\{0,7,21\}$ |
| 6 | $\{0,1,4,6\}$ | $\{0,13,39\}$ |
| 7 | $\{0,1,4,6\}$ | $\{0,13,52,78\}$ |
| 8 | $\{0,1,4,7,9\}$ | $\{0,19,76,114\}$ |
| 9 | $\{0,1,4,7,9\}$ | $\{0,19,76,133,171\}$ |
| 10 | $\{0,1,2,6,10,13\}$ | $\{0,27,108,189,243\}$ |
| 11 | $\{0,1,2,6,10,13\}$ | $\{0,27,54,162,270,351\}$ |
| 12 | $\{0,1,2,6,10,14,17\}$ | $\{0,35,70,210,350,455\}$ |
| 13 | $\{0,1,2,6,10,14,17\}$ | $\{0,35,70,210,350,490,595\}$ |
| 14 | $\{0,1,4,10,16,18,21,23\}$ | $\{0,47,94,282,470,658,799\}$ |
| 15 | $\{0,1,4,10,16,18,21,23\}$ | $\{0,47,188,470,752,846,987,1081\}$ |
| 16 | $\{0,1,5,9,16,23,26,28,29\}$ | $\{0,59,236,590,944,1062,1239,1357\}$ |
| 17 | $\{0,1,5,9,16,23,26,28,29\}$ | $\{0,59,295,531,944,1357,1534,1652,1711\}$ |

Table 6.4: Comparison of CUM-LIN, CUM-REC, CUM-ALL ($M^2 - M + 1$) and COV-1D designs for total number of effective sensors.

| $M$ | CUM-LIN | CUM-REC | CUM-ALL | COV-1D |
|---|---|---|---|---|
| 3 | 6 | 4 | 7 | 4 |
| 4 | 12 | 8 | 13 | 7 |
| 5 | 26 | 16 | 21 | 10 |
| 6 | 47 | 28 | 31 | 14 |
| 7 | 86 | 49 | 43 | 18 |
| 8 | 125 | 70 | 57 | 24 |
| 9 | 182 | 100 | 73 | 30 |
| 10 | 258 | 140 | 91 | 37 |
| 11 | 366 | 196 | 111 | 44 |
| 12 | 474 | 252 | 133 | 51 |
| 13 | 614 | 324 | 157 | 59 |
| 14 | 824 | 432 | 183 | 69 |
| 15 | 1106 | 576 | 211 | 80 |
| 16 | 1388 | 720 | 241 | 91 |
| 17 | 1742 | 900 | 273 | 102 |

Table 6.5: Sensor locations for CUM-SL method.

| $M$ | New Sensor Location | Effective Sensor Locations |
|---|---|---|
| 2 | $\{1\}$ | $\{0-2\}$ |
| 3 | $\{5\}$ | $\{0-6, 8-10\}$ |
| 4 | $\{13\}$ | $\{0-14, 16-18, 20, 21, 24-26\}$ |
| 5 | $\{28\}$ | $\{0-33, 35, 36, 38-43, 46, 50, 51, 54-56\}$ |
| 6 | $\{57\}$ | $\{0-62, 64, 65, 67-73, 75, 79-81, 83-86,$ $88, 96, 100, 101, 104, 108, 109, 112-114\}$ |

$N[M_x] = N[M_y]$ are lower bounded by $(M+1)^2/12$ (due to (6.16)), which implies the effective length $L[M_x, M_y]$ is lower bounded by $(M+1)^4/72$, for large $M$, which is twice the lower bound for CUM-REC.

We now present a final linear MRA design for small $M$ ($M < 7$) which does not employ covariance-based MRA methods and is competitive with CUM-LIN (actually slightly better) with the very important property that whenever a new sensor is available, the locations of the sensors from the previous design remain the same, so that calibration problems are not repeated every time the designer can afford a new sensor. This new design procedure (called CUM-SL) starts with the observation that given an existing linear array, cumulants can be used to double the aperture. Therefore, given a previous design, we put the new element as far as possible from the reference element with the constraint that all integers from 0 to the location of the new element can be produced by addition of two integers between the actual sensors. The procedure starts by putting the reference sensor at the origin ($M = 1$). When a second sensor is available, it can be put at 1. When the third sensor is available, it can be put 5 units to right of the reference sensor, since all the integers $\{0, 1, \ldots, 5\}$ can be generated from the set of integers between actual sensors (i.e., $\{0, 1, 4, 5\}$) by addition or subtraction of two elements. If we put the third element more than 5 units to right of the second sensor, holes appear, so we fix the location of the third element to be 5. The continuation of this search produces the results in Table 6.5, where effective sensor locations are the integers that can be obtained by adding/subtracting two integers between the actual sensors. CUM-SL is always better than CUM-LIN in this range, and by construction it guarantees an effective aperture without holes whose length is lower bounded by the location of the last

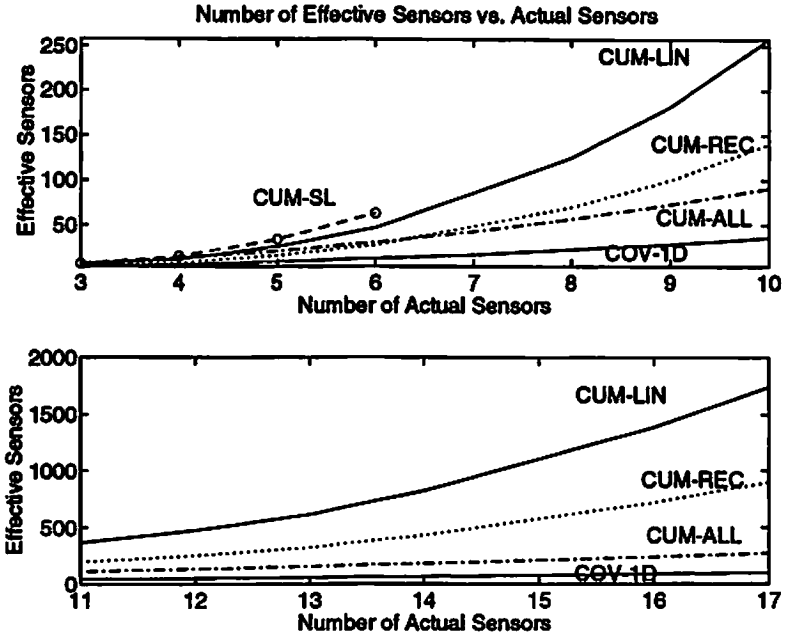



Figure 6.10: Comparison of minimum-redundancy arrays.

actual element. In addition, whenever a new sensor is available we do not have to alter the locations of the previously available sensors as in the other methods CUM-LIN, CUM-REC, COV-1D.

## 6.4 Conclusions

In this chapter, we determined lower and upper bounds on aperture extension by using the interpretation described in Chapter 4. We proved that cumulants can be used to at least double the effective aperture with an upper bound of $M^2 - M + 1$ sensors where $M$ is the number of actual sensors.

We showed ways to exceed the upper bound by designing arrays for cumulant-based processing. We started with two-dimensional MRA arrays and proposed the CUM-REC algorithm. We showed that the CUM-REC algorithm can significantly outperform the covariance-based optimum design (COV-1D), particularly when the number of actual sensors is large. After investigating the issues related to small number of sensors in the two-dimensional MRA design problem, we returned to the linear MRA design, and proposed the CUM-LIN method which is inspired by

the CUM-REC method. CUM-LIN provides twice as many effective sensors when compared to CUM-REC for large $M$. Finally, we described the CUM-SL method, which addresses the problem of linear MRA design in which availability of a new sensor does not affect the locations of the previous sensors, thus less calibration efforts are required. The effective sensors provided by all of the investigated methods are summarized in Figure 6.10.

# Chapter 7

# Non-Gaussian Noise Suppression

The main motivation of using higher-order-statistics in signal processing applications has been their insensitivity to additive colored Gaussian noise. The main objection to those methods is their possible vulnerability to non-Gaussian noise.

Our interpretation for the use of cumulants as presented in Chapter 4 assumed the *hypothetical* case, when there is additive Gaussian noise present in the measurements. In array processing, it is commonly assumed that the measurements are corrupted by additive Gaussian noise which is independent from sensor to sensor. In addition, measurement noise power is assumed to be identical for each sensor. Then, it is possible to separate signal and noise subspaces, and estimate the source directions and the noise power using the eigendecomposition of the source covariance matrix. If the ambient noise is spatially colored but its covariance matrix is known to within a scale factor, then prewhitening can be applied to the received signals, which in turn enables the separation of signal and noise subspaces. The problem of identifying the signal subspace is impossible to solve if one models the noise covariance matrix as a completely unknown Hermitian matrix; however, if the additive noise is Gaussian, then its covariance structure is not needed for the cumulant expressions which have been described in Chapters 4-5.

In this chapter, we investigate the possibility of combating the effects of *non-Gaussian noise* using cumulants. Using the geometric interpretation of cumulants, we describe a way to suppress spatially independent non-Gaussian sensor noise. Then the conditions necessary to suppress the effects of noise in more general situations are described. Noise suppression properties of the virtual-ESPRIT algorithm

(which can achieve joint array calibration and direction-finding with only a single doublet in the array) are also addressed. Finally, we propose a method, that combines second and fourth-order statistics, to achieve even better results. We demonstrate our theoretical results by simulations.

## 7.1 Weak Law of Non-Gaussian Noise Suppression

Theorem 1: Consider an array of isotropic sensors, which is illuminated by statistically independent non-Gaussian sources. Furthermore, assume that measurements are contaminated by additive non-Gaussian sensor noise, which is independent from sensor to sensor, and whose noise components can have varying power and kurtosis over the aperture. If one uses cumulants, it is possible to:

1. identify the signal subspace, although noise statistics vary from sensor to sensor; this implies the directions of far-field sources can be estimated using subspace techniques; and,

2. extend the aperture regardless of the sensor noise.

*Proof*: Since the far-field sources are assumed to be independent, we can consider the presence of a single source without loss of generality. Since noise components are independent from sensor to sensor, statistical expressions such as $E\{r^*(t)x(t)\}$ or $\text{cum}(r^*(t), x(t), r^*(t), r(t))$ are not affected by the noise. Noise affects the computation of variance at a sensor, i.e., if $r(t) = s(t) + n_r(t)$, then $E\{r^*(t)r(t)\} = \sigma_s^2 + \sigma_{n_r}^2 \neq \sigma_s^2$ whereas $E\{r^*(t)x(t)\} = \sigma_s^2 \exp(-j\vec{k} \cdot \vec{d}_x)$.

When noise power changes from sensor to sensor in an unknown way, it is not possible to remove its effects by an eigenanalysis of the sample covariance matrix, since the diagonal terms of the covariance matrix are corrupted by unknown (not necessarily identical) positive numbers; however, if one uses cumulants to compute correlations, then it is possible to exploit the sensor-to-sensor independence of noise, i.e.,

$$\frac{\sigma_s^2}{\gamma_{4,s}} \text{cum}(r^*(t), x(t), x^*(t), r(t)) \,|_{\text{with non-Gaussian noise}} =$$

$$= E\{s^*(t)s(t)\} = E\{r^*(t)r(t)\}\,|_{\text{no noise}} \tag{7.1}$$

The left-hand side of (7.1) (to within the scale factor $\sigma^2/\gamma_{4,s}$ is computed in the actual scenario where additive non-Gaussian noise is present. To derive (7.1), let $r(t) = s(t) + n_r(t)$ and $x(t) = s(t)\exp(-j\vec{k}\cdot\vec{d}_x) + n_x(t)$. Then, because the noise components $n_r(t)$ and $n_x(t)$ are independent of the signal component $s(t)$, it follows that

$$\text{cum}(r^*(t), x(t), x^*(t), r(t)) = \underbrace{\text{cum}(s^*(t), s(t)e^{-j\vec{k}\cdot\vec{d}_x}, s^*(t)e^{j\vec{k}\cdot\vec{d}_x}, s(t))}_{\gamma_{4,s}} +$$

$$\text{cum}(n_r^*(t), n_x(t), n_x^*(t), n_r(t)) \tag{7.2}$$

Since the noise components are independent of each other, the second term in the right-hand side of (7.2) is equal to zero due to [CP4]. Scaling (7.2) by $\sigma_s^2/\gamma_{4,s}$ gives the left equality in (7.1). If there is no noise, i.e., $n_r(t) = 0$, then $r(t) = s(t)$, which results in the right equality in (7.1).

The right-hand side of (7.1) can only be computed in the hypothetical case where there is no measurement noise, in which case $r(t) = s(t)$; however, when noise is present, $E\{r^*(t)r(t)\} \neq E\{s^*(t)s(t)\}$, but $E\{s^*(t)s(t)\}$ is still equal to $\text{cum}(r^*(t), x(t), r^*(t)r(t))$ to within the scale factor $\beta_s \triangleq \sigma_s^2/\gamma_{4,s}$ since the noise contributions in $r(t)$ and $x(t)$ are independent. This way of computing the variance at a sensor is illustrated in Figure 7.1. The scale factor $\sigma_s^2/\gamma_{4,s}$ does not cause a problem if all the required covariances are computed through cumulants. For example, (with proof similar to that of (7.1))

$$\text{cum}(r^*(t), x(t), r^*(t), r(t)) = \frac{\gamma_{4,s}}{\sigma_s^2} E\{r^*(t)x(t)\} \tag{7.3}$$

In the case of multiple sources, if one constructs a matrix of covariances which are computed by using cumulants through the equivalence relations (7.1) and (7.3) by ignoring the scale factors $\beta_k = \gamma_{4,k}/\sigma_k^2$, then the resulting matrix will be identical to the covariance matrix in which the source powers ($\sigma_k^2$'s) are scaled by $\beta_k$'s, and hence it takes the form (see (4.4)) $\mathbf{A}\boldsymbol{\Gamma}_{ss}\mathbf{A}^H$, where $\boldsymbol{\Gamma}_{ss}$ is the fourth-order cumulants of sources. This matrix has a rank that is equal to the number of sources, i.e., the noise subspace will be spanned by the eigenvectors that have zero eigenvalue;

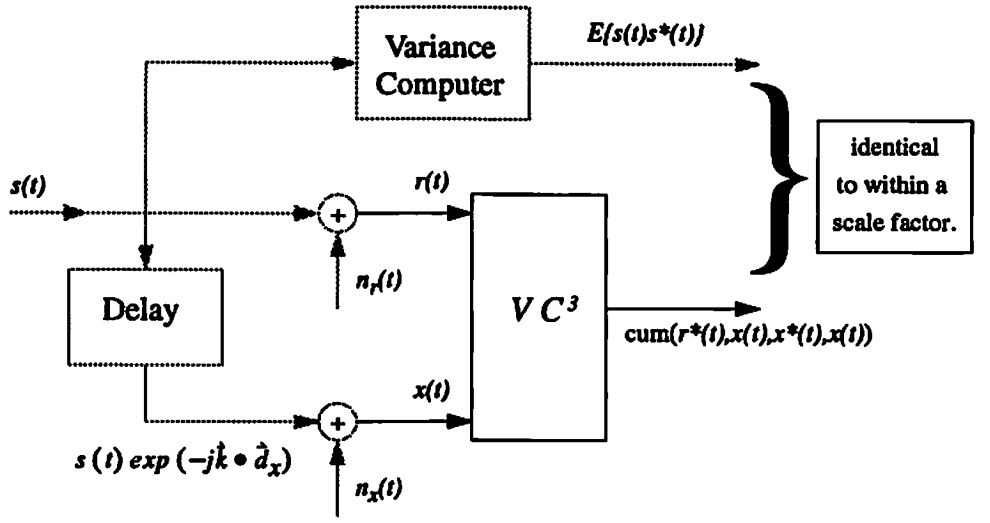



Figure 7.1: Weak law of noise suppression: The dashed lines indicate the unobservable signals and virtual processors that utilize them. The goal is to compute the variance of $s(t)$ which is present in $r(t)$, but is corrupted by noise. The virtual cross-correlation computer $\mathbf{VC^3}$ exploits the fact that the noise components $n_x(t)$ and $n_r(t)$ are independent, and computes the required second-order statistic to within a scale factor.

therefore, the signal subspace can be identified as the eigenvectors of this cumulant matrix that have non-zero (but, perhaps negative since scale factors may be negative) eigenvalues. This proves part 1 of Theorem 1.

Virtual aperture extension is the term coined in Chapter 4 to explain how cumulants increase the aperture of antenna arrays. Aperture extension is accomplished by using the cumulants of received signals to compute the cross-correlation between actual and virtual elements (e.g., see Figure 4.1, where $\vec{d} = \vec{d}_x + \vec{d}_y$). From our interpretation, this can be viewed as going to a virtual location by adding two non-zero vectors (otherwise we can not move to a virtual location) that extend between actual array elements. A non-zero vector implies that its tail and head do not coincide, i.e., in the cumulant expression to compute the virtual cross-correlation, at least one of the four components must be different than the other components; for example, (see Figure 4.2)

$$\text{cum}(r^*(t), x(t), r^*(t), x(t))\,|_{\text{with non-Gaussian noise}} = \frac{\gamma_{4,s}}{\sigma_s^2} E\{r^*(t) v_2(t)\}\,|_{\text{no noise}} \tag{7.4}$$

$E\{r^*(t)v_2(t)\}$ is not computable since we do not have $v_2(t)$ (virtual sensor); however, we have $r(t)$ and $x(t)$, and the noise in these two channels are independent; hence, equality in (7.4). We have therefore shown that $E\{r^*(t)v_2(t)\}$ (virtual statistic) can be computed using cumulants by processing the measured signals $r(t)$ and $x(t)$, even in the presence of non-Gaussian noise. This proves the second part of Theorem 1.

Comments:

• The convention established in (7.1) will be used throughout this chapter. It is important to note that (7.1) is valid for ensemble averages. With finite samples, the standard deviations of the two sides will be different.

• The geometric interpretation of (7.1) is: with cumulants, we *move* from one sensor to another one (which has non-Gaussian but independent noise), and come back to the starting point using the same path. This approach is in fact an interpretation of the technique proposed by Cardoso [21] for accomplishing non-Gaussian noise insensitivity by cutting off the diagonal elements of quadricovariance steering matrices.

• The limitation of the weak law comes from the assumption about the sensor-to-sensor independence of the non-Gaussian noise.

## 7.2 Strong-Law of Non-Gaussian Noise Suppression

**Theorem 2**: Consider an array of arbitrary sensors which is illuminated by *linearly-correlated* non-Gaussian sources. Assume that array measurements are contaminated by *non-Gaussian* sensor noise of arbitrary cross-statistics. Then, it is possible to identify the signal subspace to estimate the DOA parameters by subspace techniques if there is a single sensor whose measurements are contaminated by non-Gaussian noise which is independent of the noise component of other sensors. Furthermore, there is no need to store the spatial response of that sensor.

*Proof*: To begin, we assume the sources are independent. Later we consider linearly-correlated sources. Consider Figure 7.2, in which there is an array of $M$ sensors

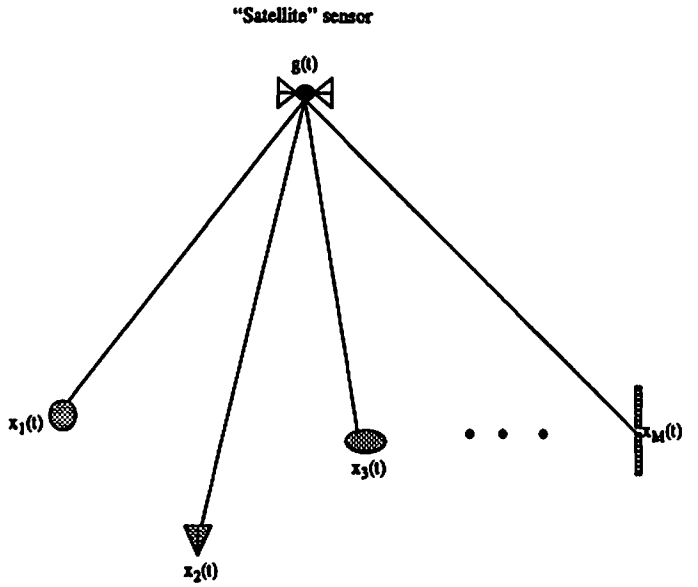



Figure 7.2: Strong-law of non-Gaussian noise suppression: an arbitrary array of M sensors whose measurements are corrupted by colored non-Gaussian noise can still be used for direction-finding if the noise in $g(t)$ is independent of the noise present in the rest of the array elements. Such a unit can be imposed in the field to correct the performance of existing systems which suffer from colored noise.

which measure $\{x_k(t)\}_{k=1}^{M}$ and there is another sensor that measures $g(t)$ where

$$g(t) = \mathbf{g}^T \mathbf{s}(t) + n_g(t) \tag{7.5}$$

whereas the main array measurements take the form

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}_\mathbf{x}(t) \tag{7.6}$$

We assume the noise component $n_g(t)$ in $g(t)$ is independent of the noise component $\mathbf{n}_\mathbf{x}(t)$ in the main array. The satellite sensor, $g(t)$, can be used to compute the second-order statistics by using cumulants (assume a single source for the moment), because:

$$\begin{aligned} \text{cum}(x_j^*(t), g(t), g^*(t), x_k(t)) \,|_{\text{with non-Gaussian noise}} \; = \\ = \frac{\gamma_{4,s}|g_s|^2}{\sigma_s^2} E\{x_j^*(t) x_k(t)\} \,|_{\text{no noise}} \end{aligned} \tag{7.7}$$

where $g_s$ is the response of the satellite sensor to the source wavefront. In this way, the following noise-free array covariance matrix can be constructed using cumulants to replace the second-order statistics (see (7.7)) as in Figure 4.3 (assuming multiple independent sources where superposition holds):

$$\mathbf{C} = \mathbf{A}\boldsymbol{\Gamma}\mathbf{A}^H \tag{7.8}$$

where $\mathbf{A}$ is the steering matrix for $M$ sensors (except the satellite sensor), $\boldsymbol{\Gamma}$ is a diagonal matrix whose $k$th diagonal entry is $\gamma_{4,k}|g_k|^2$, and $g_k$ is the response of the satellite sensor to the $k$th source (the vector $\mathbf{g}$ is the collection of such responses). Equation (7.8) follows from application of (7.7) to (4.4) and the fact that in the absence of noise the array covariance matrix takes the form $\mathbf{A}\,\boldsymbol{\Sigma}_{ss}\mathbf{A}^H$, where $\boldsymbol{\Sigma}_{ss}$ is a diagonal matrix that contains source powers (see (4.4)). Note that, $\beta_k = \gamma_{4,k}|g_k|^2/\sigma_k^2$.

Any subspace method can be applied to $\mathbf{C}$ in (7.8), whose elements are computed using (7.7). There is *no need* to know the response of the satellite sensor to the far-field sources (i.e., $\mathbf{g}$ in (7.5)); but, the elements of $\mathbf{g}$ must be nonzero in order to make $\boldsymbol{\Gamma}$ a nonsingular matrix. We just need the time series recorded by the satellite sensor to actually compute the left-hand side of (7.7).

Next, consider the source signals $\mathbf{s}(t)$ correlated in the following way:

$$\mathbf{s}(t) = \mathbf{Q}\mathbf{u}(t) \tag{7.9}$$

where $\mathbf{Q}$ is non-singular (but arbitrary), and components of $\mathbf{u}(t)$ are independent. Then the observation equations (7.5) and (7.6) change to

$$\begin{aligned} g(t) &= \mathbf{h}^T\mathbf{u}(t) + n_g(t) \\ \mathbf{x}(t) &= \mathbf{B}\mathbf{u}(t) + \mathbf{n}_\mathbf{x}(t) \end{aligned} \tag{7.10}$$

where $\mathbf{B} \triangleq \mathbf{A}\,\mathbf{Q}$ and $\mathbf{h}^T \triangleq \mathbf{g}^T\,\mathbf{Q}$. Since the components of $\mathbf{u}(t)$ are independent, they can be viewed as the actual source waveforms of (7.5) and (7.6) with an effective steering matrix $\mathbf{B}$, and a response vector $\mathbf{h}$. The cumulant matrix $\mathbf{C}$, computed as described in (7.7) for the independent sources scenario, now takes the form

$$\mathbf{C} = \mathbf{B}\tilde{\boldsymbol{\Gamma}}\mathbf{B}^H \tag{7.11}$$

which was obtained by substituting $\mathbf{B}$ for $\mathbf{A}$ and $\mathbf{h}$ for $\mathbf{g}$ in (7.8). In (7.11), $\tilde{\mathbf{\Gamma}}$ is defined as the diagonal matrix whose $k$th diagonal entry is $\tilde{\gamma}_{4,k}|h_k|^2$ and $\tilde{\gamma}_{4,k}$ is the fourth-order cumulant of $u_k(t)$. Note that $\mathbf{Q}\tilde{\mathbf{\Gamma}}\mathbf{Q}^H$ is full-rank, so that $\mathbf{C}$, expressed as

$$\mathbf{C} = \mathbf{A}(\mathbf{Q}\tilde{\mathbf{\Gamma}}\mathbf{Q}^H)\mathbf{A}^H , \tag{7.12}$$

maintains all the requirements for subspace algorithms like MUSIC and ESPRIT for direction-finding, even in the presence of correlated sources, correlated non-Gaussian noise, and arbitrary array characteristics. It is also important to note that we do not need to know the response of the satellite sensor to the waveforms (i.e., $\mathbf{g}$), as long as the components of $\mathbf{h}$ ($\mathbf{h}^T = \mathbf{g}^T\mathbf{Q}$) are non-zero. This completes the proof of Theorem 2.

Comments:

- This method can be interpreted as follows: consider a totally different problem in which the sensors $\{x_k\}_{k=1}^M$ are viewed as mobile communication antennas which suffer from interference effects, so that they can not communicate directly. It is necessary to use a *satellite* transponder ($g(t)$) to maintain communications among sensors. Remember, from the Chapter 4, that *communication* between sensors means *implementing cross-covariance.* Here we can not do that because of non-Gaussian sensor noise of arbitrary statistics; however, the satellite sensor, $g(t)$, can be used to make that communication possible: to implement $E\{x_j^*(t)x_k(t)\}$ first move from $x_j(t)$ to the satellite sensor $g(t)$, then let the satellite distribute the message; i.e., move from $g(t)$ to $x_k(t)$.
- A similar technique was developed in [172] as a covariance-based approach; however, it requires an additional array of sensors whose noise component is independent of the noise in the existing array. Consequently, [172] ends up doubling the number of sensors for direction-finding. We have accomplished noise reduction by using only one extra sensor. This gain on the number of required sensors is similar to the gain observed in the virtual-ESPRIT algorithm (Chapter 4) as compared with covariance-ESPRIT. The reason for this difference is that *cumulants, unlike cross-correlation, have an array of arguments.*
- If the original array is linear and consists of uniformly spaced sensors of identical

response, then it is possible to apply the spatial-smoothing algorithm of [123] to the covariance matrix in (7.12) to estimate the parameters of *coherent sources* (i.e., when **Q** is singular). Simulations in Section 7.5 investigate the coherent sources in non-Gaussian noise scenario.

• Virtual aperture extension using the strong law of noise suppression is possible: it requires fixing one of the four arguments of the cumulant to be $g(t)$. Consequently, this problem reduces to aperture extension using third-order cumulants, since we now have three free cumulant arguments with which to extend the aperture.

## 7.3 Virtual-ESPRIT and Non-Gaussian Noise

The virtual-ESPRIT algorithm (VESPA) calibrates arbitrary arrays using a single doublet and fourth-order cumulants (see Chapter 5). Here, we discuss the properties of VESPA in non-Gaussian noise.

Theorem 3: Assume independent non-Gaussian sources illuminate an array of arbitrary sensors whose measurements are corrupted by non-Gaussian noise of unknown statistics. Joint array calibration and direction-finding is possible, if we have a doublet and at least one of the doublet element's measurement noise component is independent of the noise components measured by the rest of the sensors.

*Proof*: We apply the strong-law of non-Gaussian noise suppression to VESPA. Consider $\{x_1(t), x_2(t), \ldots, x_M(t)\}$ to be the measurements from an arbitrary array. Let us assume that we have an identical copy of the first sensor (that measures $x_1(t)$) whose noise component is independent of the rest of the sensors. The measurement recorded by this sensor is denoted as $g(t)$. Let $a_s$ denote the response of the first sensor to the wavefront.

ESPRIT autocorrelation matrix for the measurements $\{x_k(t)\}_{k=1}^{M}$ can now be generated using cumulants as

$$\text{cum}(x_k^*(t), g(t), g^*(t), x_l(t))\,|_{\text{non-Gaussian noise}} = \frac{\gamma_{4,s}|a_s|^2}{\sigma_s^2} E\{x_k^*(t)x_l(t)\}\,|_{\text{no noise}} \tag{7.13}$$

where $1 \leq k, l \leq M$. This is in fact the idea presented for the strong-law of non-Gaussian noise suppression: we first move to the satellite sensor $g(t)$ and then come back. A slight modification of this idea can be used to virtually compute the ESPRIT cross-correlation matrix. Let $v_l(t)$ denote the virtual-process received by the virtual-twin of $x_l(t)$ where $1 \leq l \leq M$. We can compute

$$\text{cum}(x_k^*(t), g(t), x_1^*(t), x_l(t)) \,|_{\text{non-Gaussian noise}} = \frac{\gamma_{4,s}|a_s|^2}{\sigma_s^2} E\{x_k^*(t) v_l(t)\} \,|_{\text{no noise}} \tag{7.14}$$

although $v_l(t)$ is not physically available. This completes the proof of Theorem 3.

## 7.4 Combination of Second and Fourth-Order Statistics

We have shown several ways to use higher-order statistics to suppress non-Gaussian noise. In this section, we investigate possible use of second-order statistics along with fourth-order cumulants. We show that the strong law of noise suppression can be improved by using second-order statistics.

Consider the cross-correlation vector $\mathbf{d}$, defined as

$$\mathbf{d} \triangleq E\{\mathbf{x}(t) g^*(t)\} \tag{7.15}$$

where $\mathbf{x}(t)$ denotes the measurements of the main array and $g(t)$ is the measurement of the satellite sensor (see Figure 7.2). Since the noise component of $\mathbf{x}(t)$ is independent of the noise component in $g(t)$, $\mathbf{d}$ is free of the effects of noise (when ensemble average is considered). If $\mathbf{A}$ is the steering matrix for the main array, and the sources are linearly correlated ($\mathbf{s}(t) = \mathbf{Q}\mathbf{u}(t)$; see (7.9)), then, since the noise components $n_g(t)$ and $\mathbf{n}_\mathbf{x}(t)$ are independent, we can consider only the signal components of measurements to obtain

$$\mathbf{d} = E\{\mathbf{x}(t) g(t)\} = E\{\mathbf{A}\,\mathbf{Q}\mathbf{u}(t)\mathbf{u}^H(t)\mathbf{Q}^H \mathbf{g}^*\} = \mathbf{A}\,\mathbf{Q}\,\boldsymbol{\Sigma}_{uu}\,\mathbf{Q}^H\,\mathbf{g}^* = \mathbf{A}\mathbf{z} \tag{7.16}$$

where $\mathbf{\Sigma}_{uu} \triangleq E\{\mathbf{u}(t)\mathbf{u}^H(t)\}$. If none of the components of $\mathbf{z}$ are non-zero, then $\mathbf{d}$ is a superposition of steering vectors from the sources; hence, it is possible to estimate the directions-of-arrival by minimizing the cost function

$$\hat{\theta} = \arg\min_{\theta,\mathbf{z}} J(\theta) \triangleq \arg\min_{\theta,\mathbf{z}} \| \mathbf{d} - \mathbf{A}(\theta)\,\mathbf{z} \|_F^2 \tag{7.17}$$

It is possible to eliminate $\mathbf{z}$ in the optimization procedure, since given the optimal estimates for $\theta$, namely $\hat{\theta}$, $\hat{\mathbf{z}}$ can be estimated as

$$\hat{\mathbf{z}} = \mathbf{A}(\hat{\theta})(\mathbf{A}^H(\hat{\theta})\mathbf{A}(\hat{\theta}))^{-1}\mathbf{A}^H(\hat{\theta})\ \mathbf{d} \tag{7.18}$$

Substituting (7.18) into (7.17), we obtain

$$\hat{\theta} = \arg\min_{\theta} \|(\ \mathbf{I} - \mathbf{A}(\theta)\,(\mathbf{A}^H(\theta)\mathbf{A}(\theta))^{-1}\,\mathbf{A}^H(\theta)\ )\ \mathbf{d}\|_F^2 = \arg\min_{\theta} \|\mathbf{P}^{\perp}_{\mathbf{A}(\theta)}\,\mathbf{d}\|_F^2 \tag{7.19}$$

(where $\mathbf{P}_{\mathbf{X}}$ denotes the projection matrix on the range of $\mathbf{X}$) or, alternatively

$$\hat{\theta} = \arg\max_{\theta} \| \mathbf{P}_{\mathbf{A}(\theta)}\ \ \mathbf{d} \|_F^2 \tag{7.20}$$

The direction estimation from (7.20) requires a $P$ dimensional search procedure ($P$ is the number of sources). This search is quite complex unless we have good initial estimates. We use the estimates provided by the strong law of non-Gaussian noise suppression for initialization. The maximization in (7.20) can be performed by the alternating projection (AP) method, as suggested by Ziskind and Wax for the cost function associated with the deterministic maximum-likelihood method for direction-finding [180]. We refer the reader to [180] for the implementation of the AP algorithm.

Since this approach of suppressing non-Gaussian noise uses second-and fourth-order statistics together we call it the SFS method. Simulations presented in Section 7.5 indicate that the SFS method can decrease the variance of the estimates from the cumulant-based approach which is now only used for initialization.

## 7.5 Simulations

In this section, we provide simulations that demonstrate the performance of non-Gaussian noise-insensitive direction-finding methods. Our first simulation experiment illustrates virtual aperture extension in the presence of spatially non-stationary but independent non-Gaussian sensor noise. Our second simulation compares previous cumulant and covariance based algorithms with our strong law of noise suppression. Our third simulation investigates direction-finding for coherent sources in non-Gaussian colored noise; it also investigates the performance improvement obtained by using both second-and fourth-order statistics. Our final simulation illustrates the non-Gaussian noise suppression properties of VESPA.

### 7.5.1 Virtual Aperture Extension in Non-Gaussian Noise

In this experiment, we consider a two-element linear array illuminated by two equal-power, independent BPSK sources from $\pm 10^{\circ}$. We assume the sensors are isotropic. We define SNR as the ratio of the signal power to noise power at the first (reference) sensor. We let the non-Gaussian noise component at the second sensor be 3 dB above that of the reference sensor. The non-Gaussian noise is due to the near-field QAM communications equipment. We collect 1000 snapshots from the array.

Using cumulants, it is possible to extend the aperture to 3 sensors. If the noise components in the actual sensor measurements are independent, then it is possible to apply the weak-law of noise suppression to construct a 3x3 matrix which is asymptotically independent of the presence of non-Gaussian noise. If we are constrained to use only second-order statistics, we can not identify the bearings of the sources, since the number of actual sensors is not larger than the number of sources. Figure 7.3 illustrates the results obtained by using the ESPRIT algorithm (other direction-finding methods are applicable, but ESPRIT is chosen for its simplicity) on the 3x3 cumulant matrix. It is observed that cumulant-based results are unbiased, and they improve as the SNR increases.

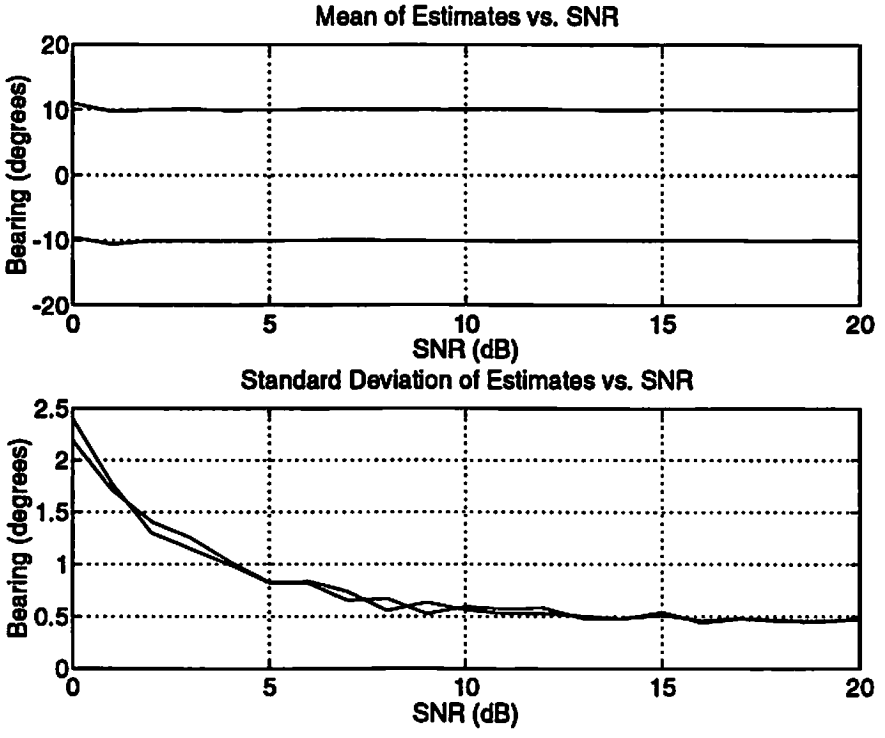



Figure 7.3: Properties of the estimates from virtual aperture extension in the presence of spatially-nonstationary but independent non-Gaussian noise. SNR is increased from 0dB to 20dB, and each result is obtained from 100 independent realizations.

### 7.5.2 Incoherent Sources in Non-Gaussian Noise

Here we estimate the DOA's of two far-field sources which illuminate a uniformly spaced linear array of five sensors from $-5^o$ and $5^o$. Both sources broadcast BPSK waveforms of unity variance. The noise covariance matrix has the following structure ($\mathbf{R}_n = \mathbf{L}\mathbf{L}^H$):

$$\mathbf{L} = \beta \begin{bmatrix} 1.4142e^{j0.0} & 0.0 & 0.0 & 0.0 & 0.0 \\ 1.4051e^{j0.0548} & 0.1606e^{j0.0} & 0.0 & 0.0 & 0.0 \\ 1.3987e^{j0.1097} & 0.1355e^{j0.0548} & 0.1588e^{j0.0} & 0.0 & 0.0 \\ 1.2585e^{j0.7928} & 0.2487e^{-j2.4036} & 0.3460e^{-j2.4584} & 0.4844e^{j0.0} & 0.0 \\ 1.1820e^{-j0.6831} & 0.4201e^{-j0.7380} & 0.4378e^{-j0.7928} & 0.4535e^{j1.6656} & 0.1702e^{j0.0} \end{bmatrix}$$

The SNR is defined as $-20\log_{10}\beta$. The signal power is 20 dB above the noise power at the satellite sensor. The noises at all sensors are 4-QAM type; they represent the

effect of near-field communications equipment on the direction-finding system. We tested three versions of the ESPRIT algorithm: (1) The original covariance-ESPRIT of [108] (denoted "cov"); (2) The cumulant-ESPRIT as suggested in [91] (denoted "PN") which is designed for suppressing colored Gaussian noise; and (3) Our strong-law of noise suppression method, which uses the satellite sensor to obtain the non-Gaussian noise free matrix in (7.8) (denoted "DM"). All versions use the total-least squares ESPRIT.

At a sample size of 1000 snapshots, we varied SNR from -10dB to 10dB and averaged the results from 200 experiments. Figure 7.4 depicts the mean values of estimates. Observe that below a certain SNR, cov and PN yield biased results; after a threshold SNR (around 3-4dB) their estimates get close to the true values; but, the DM algorithm provides unbiased results for all SNR levels.

The standard deviation plot in Figure 7.5 is more interesting. At low SNR cov yields the least variation, but the estimates are biased. Both PN and DM estimates have higher deviations, but DM estimates vary around the true values. As SNR increases, the variation in the DM estimate decreases monotonically; however, there is a *hill-climbing* effect in the variation of the estimates for cov and PN. This effect takes place around the threshold SNR (see Figure 7.4 for the mean of the estimate corresponding to the $-5^{\circ}$ source.). This is where, due to sufficiently high SNR, the effect of colored noise no longer is very significant, i.e., the leakage-to-noise subspace is less important due to the larger difference between noise and signal subspace singular values. However, SNR is still not high enough to prevent the breakdown of the cov and PN algorithms when the noise in the experiment is bad (outcome is worse than the ensemble average). In this case, both PN and cov break down many times, i.e., the estimates are far away from the true value. This increases the deviation of the estimates and results in the hill-climbing. Further SNR increase eliminates this phenomenon and standard deviations decrease; but DM always results in less variation for all SNR levels.

An important (and expected) observation is that the PN algorithm starts hill-climbing before cov. The reason for this is due to the role of cumulants in harmonic-retrieval problems [137]. For non-Gaussian noise, cumulants boost the SNR if it is

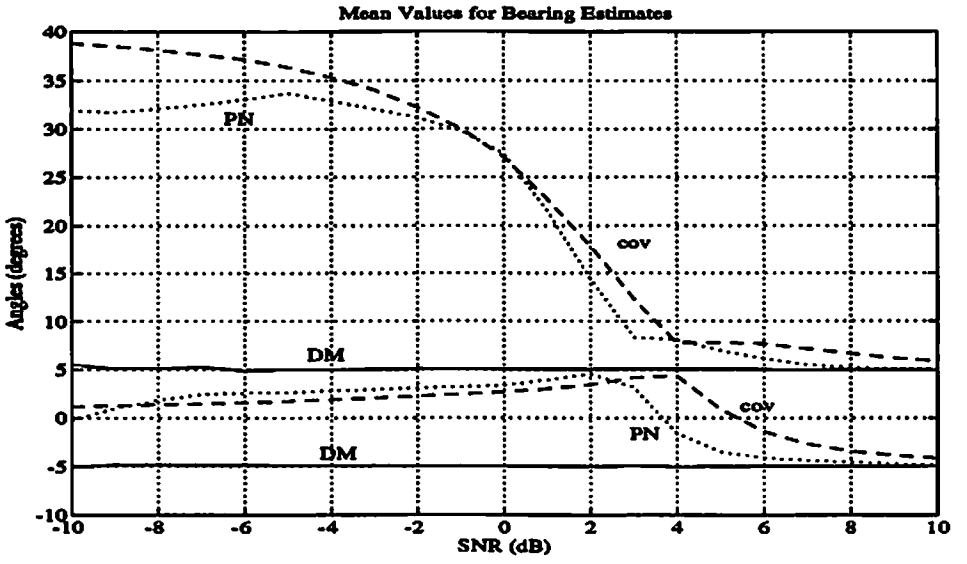



Figure 7.4: Performance comparison for the direction-finding algorithms with 1000 snapshots: the mean of the estimates.

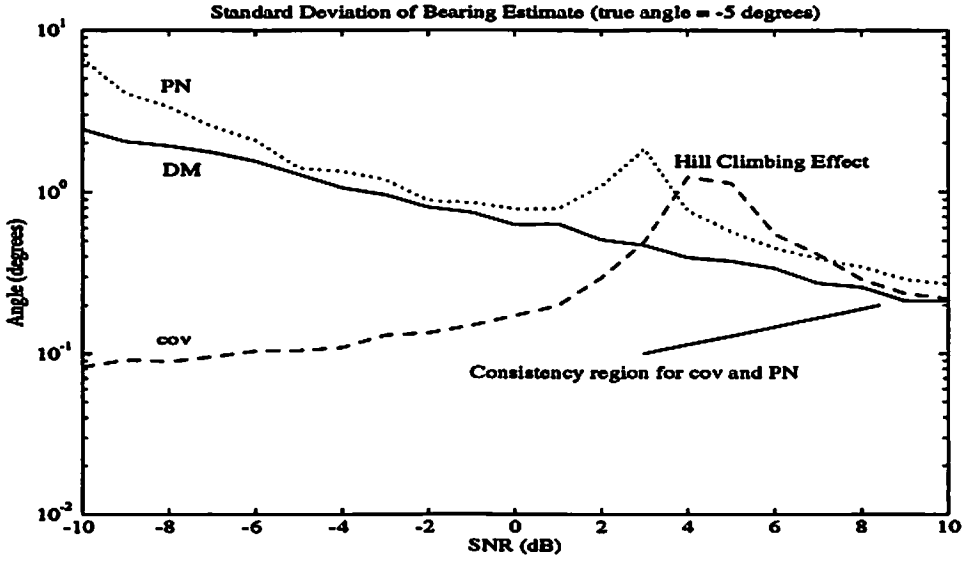



Figure 7.5: Performance comparison for the direction-finding algorithms with 1000 snapshots: the standard deviation of estimate for the source from $-5°$.

already above a threshold (below that threshold the opposite happens). For example, if $v(t) = \beta s(t)$, then $\sigma_v^2 = \beta^2 \sigma_s^2$ whereas $\gamma_{4,v} = \beta^4 \gamma_{4,s}$; hence

$$\frac{\gamma_{4,s}}{\gamma_{4,v}} = (\frac{\sigma_s^2}{\sigma_v^2})^2 \tag{7.21}$$

which implies that for two identically distributed signals the ratio of fourth-order cumulants is the square of the ratio of variances. This is the reason why PN achieves the top of the hill before cov: PN boosts the signal subspace singular values, so that the variation of noise singular values can be ignored, and hence bias in the bearing estimates approach zero sooner than for cov.

### 7.5.3 Coherent Sources in Non-Gaussian Noise

To compare the performance of the algorithms described in Experiment 2, for the case of coherent sources, we increased the number of main array sensors to 8 in order to improve resolution and enable spatial smoothing. First we investigate the case of spatially-white non-Gaussian noise. We consider a BPSK source illuminates the array from $5^o$, and due to multipath a perfectly coherent equal-power replica illuminates the array from $-5^o$. The extra sensor for the DM algorithm is located at 10 wavelengths away from the first (left) element of the linear array. The noise components are assumed to originate from QAM communications equipment. We assume the noise power is identical (unity) at all sensors including the satellite sensor. We assume the signal power is equal to the noise power (0 dB). We use MUSIC algorithm after one-level of spatial-smoothing [123] to analyze all of the methods: covariance-based (cov), cumulant-based (PN), and our approach which uses the satellite sensor (DM). We use 1000 snapshots to estimate the required statistics and display spatial spectra for 50 independent realizations. From Figure 7.6 we observe that in many realizations, cov is unable to resolve the sources in a satisfactory way. In addition, the estimates are biased, whenever cov can resolve the sources (indicated by the two vertical lines). In this case, both of the cumulant-based algorithms are able to resolve the sources, and their performances are similar. We illustrate the spatial spectrum from DM in Figure 7.7.

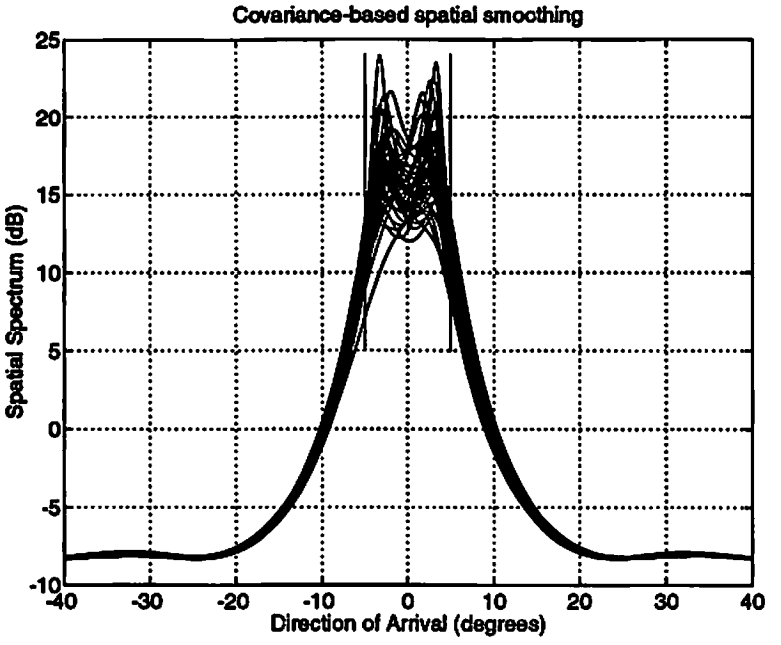



Figure 7.6: Covariance-based MUSIC algorithm is unable to resolve the sources in general. Even when resolution is possible, the estimates are biased.

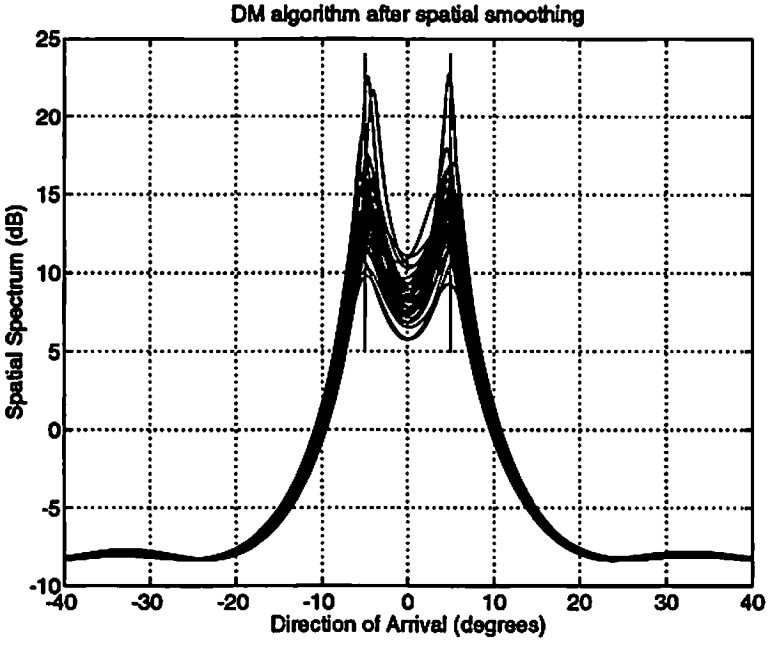



Figure 7.7: Cumulant-based algorithm which uses an extra (satellite) sensor successfully resolves the sources and estimates the directions without bias.

Next, we investigate the effects of colored non-Gaussian noise. To make matters even worse, we assume the two coherent wavefronts are closer to each other: the bearings are now $\{-2.5^o, 2.5^o\}$. We assumed the noise covariance matrix for the main array takes the form: $\mathbf{R}_{nn} = \mathbf{a}(0)\mathbf{a}^H(0) + 0.01\mathbf{I}$ where $\mathbf{a}(0)$ is the steering vector that corresponds to $0^o$, i.e., $\mathbf{a}(0) = [1, 1, \ldots, 1]$. $\mathbf{R}_{nn}$ represents an ambient noise structure whose major component illuminates the array from $0^o$ and shadows the presence of sources. The noise power at the satellite sensor remains at unity, and the signal power remains at 0dB. Figure 7.8 illustrates the PN algorithm in this scenario: sources are never resolved since the processor confuses the noise as a signal and indicates the arrival from the corresponding direction $0^o$. The covariance-based approach does not do any better; hence, we do not show its spatial spectrum. On the other hand, DM algorithm successfully resolves the two sources and suppresses the noise. However, DM estimates are slightly biased. The reason is that the sample size is not enough to suppress the effects of the high-power noise source from $0^o$, which leaks into the spatial smoothing algorithm and pulls the estimates towards $0^o$. This observation is in accordance with the results of Xu and Buckley [174], which indicate that as the correlation increases between closely separated sources, bias plays an increasingly important role.

Finally, we illustrate the improvement provided by the SFS algorithm of Section 7.4. We initialized the search required by SFS by the results of DM. We display the estimates provided by DM-MUSIC and SFS algorithms from 50 trials in Figure 7.10. SFS reduces the variation and bias in the estimates significantly.

### 7.5.4 Virtual-ESPRIT and Non-Gaussian Noise

In this experiment, we used the 8 element array of Experiment 3 with the same noise correlation structure and strength. Two equal power, independent signals illuminate the array from $\pm 2.5^o$. Non-Gaussian noise suppression can be achieved in VESPA in two ways: (1) Use one new sensor that is a copy of an existing sensor, but whose additive noise is independent of noises in the other sensors; or (2) Use a doublet located sufficiently far away from the original array so that the noise contribution of the doublet measurements are independent of the noises in the original array. The first method applies only when one of the responses of the main array elements is

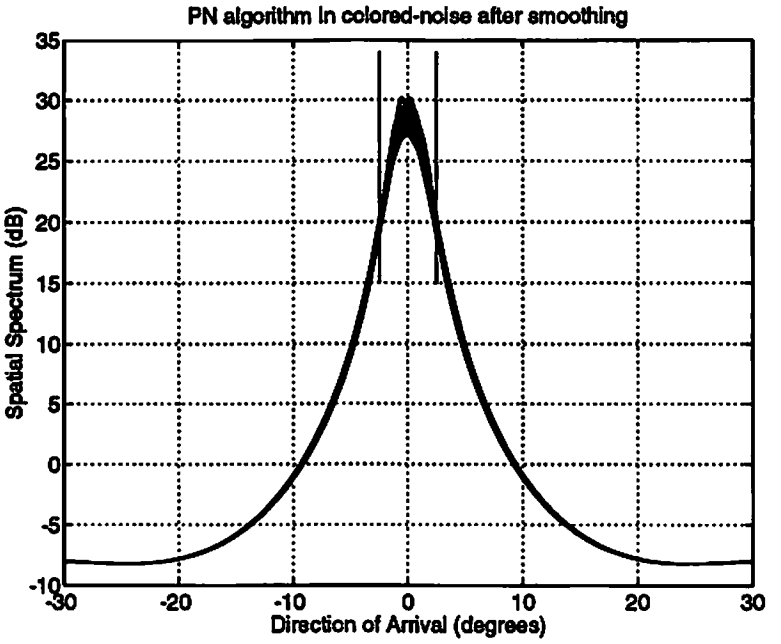



Figure 7.8: Cumulant-based MUSIC algorithm (PN) is unable to resolve the sources when the non-Gaussian noise is colored.

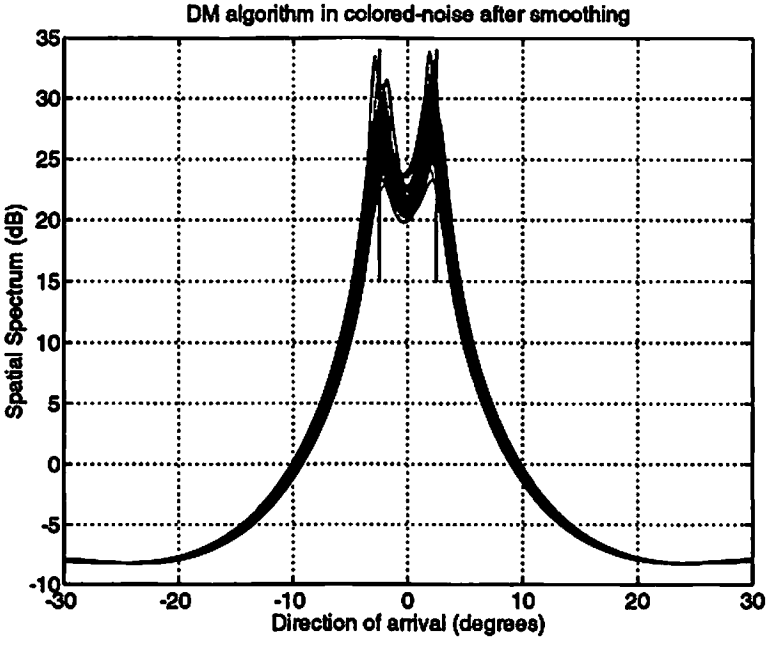



Figure 7.9: Cumulant-based algorithm which uses an extra (satellite) sensor successfully resolves the sources in colored non-Gaussian noise. The estimates can be fine-tuned by the SFS algorithm which uses DM estimates for initialization.

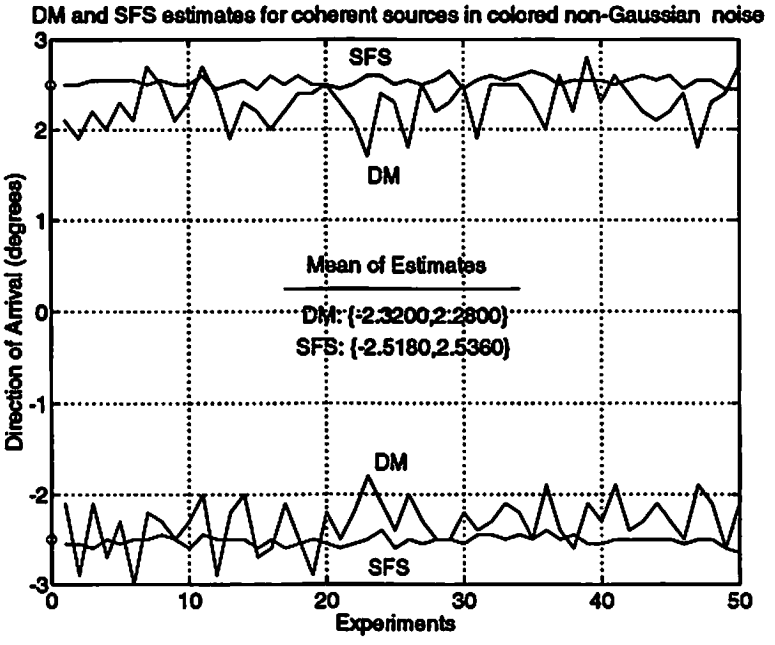



Figure 7.10: SFS and DM-MUSIC performance comparison: SFS decreases the variation and the bias (due to finite number of samples) of DM-MUSIC estimates, since it uses second-order statistics for estimation and fourth-order statistics for initialization.

known. This is the major reason for using the second approach. In addition, the second method is insensitive to noise correlation structure between the two guiding sensors if we only create a copy of the original array; i.e., an 8x8 cross-correlation matrix between the original array and its virtual copy rather than a 9x9 matrix. Specifically, we used two guiding sensors separated by $\lambda/2$. The guiding sensors are located on the axis of the main array, and the first one of which is $10\lambda$ left of the leftmost element of the main array. The noise power at the guiding sensors is identical to the noise level at the satellite sensor of the third experiment (unity power). The signals are at 0db with respect to noise at the guiding sensor.

We performed 1000 independent experiments to estimate source directions using virtual-ESPRIT. The distribution of estimates are given in Figure 7.11. VESPA resolves the sources as well as the DM-MUSIC algorithm described in Experiment 3. The biases in the estimates are less than that of the DM algorithm, since sources are independent for this experiment. In the next chapter we propose a combined beamforming and calibration method to improve the results from VESPA.

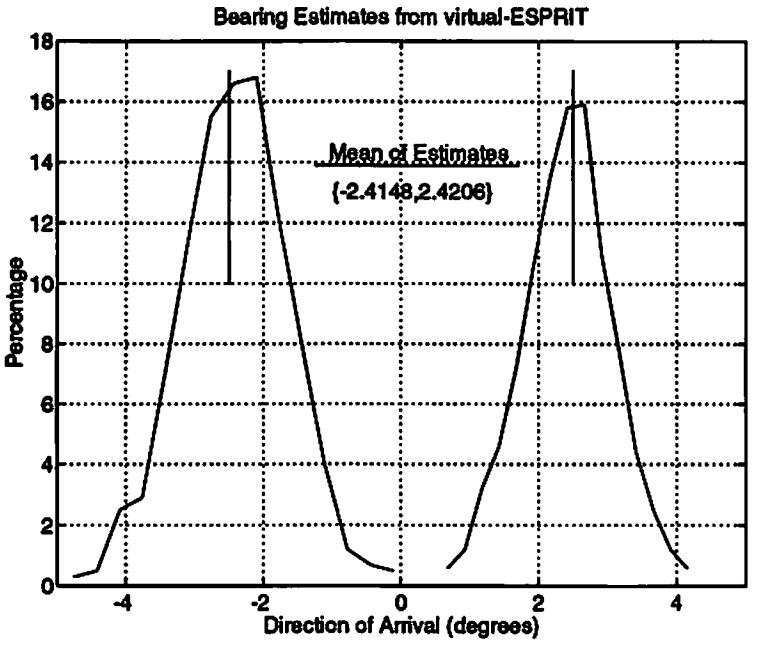



Figure 7.11: Virtual-ESPRIT algorithm can estimate the source bearings in the presence of non-Gaussian noise. The graph indicates that the sources are resolved successfully.

## 7.6 Conclusions

We have developed algorithms which are capable of suppressing the effects of non-Gaussian noise in array processing problems. We also compared the performance of existing cumulant-based work for suppressing non-Gaussian noise and explored the effects of SNR. We found out that cumulant-based algorithms tend to remove the effects of noise color above a threshold SNR by boosting the singular values of the signal subspace.

We also showed that it is possible to improve cumulant-based results by using second-order statistics. The vector $\mathbf{d} \triangleq E\{\mathbf{x}(t)g^*(t)\}$ is noise-free, and a least-squares optimization that uses $\mathbf{d}$ is proposed for DOA estimation. This algorithm can be initialized by the results of subspace algorithms. Our simulations indicated that this combined method significantly improves the bias and deviation in the estimates provided by cumulant-based algorithms. For even better asymptotic performance the covariance matrix of the cross-correlation vector $\mathbf{d}$ can be computed from data,

as suggested in [99], and a weighted least-squares procedure can be adopted at the expense of computational complexity.

Some remaining research issues related to non-Gaussian noise suppression are listed below:

• Use the interpretation for third-order cumulants for non-Gaussian noise suppression purposes. Next chapter addresses noise suppression in addition to array calibration and aperture extension using third-order statistics.

• Spatially-filtering main array measurements in a way to enhance the signal component from a location sector, and using this enhanced signal as the satellite sensor measurement, $g(t)$, to decrease the effects of non-Gaussian noise in direction-finding algorithms. For example, if the noise contribution can be decreased to -20 dB with respect to signal power ($\beta = 1/10$ in (7.21)) by spatial filtering of array measurements, then the contribution of noise in the cumulant matrix in the way it is constructed for the DM algorithm will have noise contribution scaled by 1/100. Slightly biased bearing estimates can be obtained from the DM algorithm. The results can be improved by processing the cumulant vector $\mathbf{c}$, defined as

$$(\mathbf{c})_k \triangleq \text{cum}(g^*(t), g(t), g^*(t), x_k(t)) \quad 1 \leq k \leq M \tag{7.22}$$

in which the $g(t)$ corresponds to spatially filtered array measurements. The noise contribution in $\mathbf{c}$ will be scaled by 1/1000 (due to the properties of cumulants the noise contribution to the cumulant vector $\mathbf{c}$ will be scaled by $\beta^3$). The fourth-order statistics vector $\mathbf{c}$ can then be processed by the alternating projection algorithm [180] to obtain better (relatively unbiased) results.

• Combine spatial filtering with direction-finding problem, and perform beamspace processing without decreasing the number of sensors by noting the fact that cumulants have an array of lags which can substitute for the reduction of the aperture in regular beamspace processing.

Finally, we note that third-order cumulants can be used for non-Gaussian noise suppression. We also indicate the noise suppression capability of virtual-ESPRIT based on third-order cumulants.

Consider the scenario in Section 7.2, in which an array of arbitrary sensors $\{x_1(t), x_2(t), \ldots, x_M(t)\}$ is illuminated by linearly correlated non-Gaussian sources, and there exists a satellite sensor measurement $g(t)$, i.e,

$$\mathbf{x}(t) = \mathbf{A}\underbrace{\mathbf{Q}\mathbf{u}(t)}_{\mathbf{s}(t)} + \mathbf{n}_{\mathbf{x}}(t) \tag{7.23}$$

and

$$g(t) = \underbrace{\mathbf{g}^T\mathbf{Q}}_{\mathbf{h}^T}\mathbf{u}(t) + n_g(t) \tag{7.24}$$

Due to the superposition property of cumulants, consider a single source, with third-order cumulant $\gamma_{3,s}$. Then, we have the following result for non-Gaussian noise suppression:

$$\text{cum}(x_j^*(t), x_k(t), g(t))\,|_{\text{with non-Gaussian noise}} = \frac{\gamma_{3,s}h_s}{\sigma_s^2} E\{x_j^*(t)x_k(t)\}\,|_{\text{no noise}} \tag{7.25}$$

where $h_s$ is the response of the satellite sensor to the wavefront. Equation (7.25) can be derived as in the fourth-order cumulant case which was presented in Chapter 7. The cumulant matrix $\mathbf{C}$, computed by replacing the required correlations by cumulants, takes the form:

$$\mathbf{C} = \mathbf{A}\,(\mathbf{Q}\tilde{\mathbf{\Gamma}}\,\mathbf{Q}^H)\,\mathbf{A}^H \tag{7.26}$$

where $\tilde{\mathbf{\Gamma}}$ is defined as the diagonal matrix whose $k$th diagonal entry is $\tilde{\gamma}_{3,k}h_k$, and $\tilde{\gamma}_{3k}$ is the third-order cumulant of $u_k(t)$. This third-order cumulant matrix $\mathbf{C}$ (as its fourth-order cumulant counterpart in (7.12)) satisfies all the requirements for subspace algorithms like MUSIC and ESPRIT for direction-finding, even in the presence of correlated sources, correlated non-Gaussian noise, and arbitrary array characteristics.

The virtual-ESPRIT algorithm based on third-order cumulants was described in Chapter 5. If the noise components of *both* of the guiding sensors are independent of the rest of the array, then these sensors can be thought of as two satellite sensors, indicating that the covariance matrices computed by using cumulants as in (5.10) and (5.13) are not affected by additive non-Gaussian noise. We note that the fourth-order

cumulants only require one of the guiding sensor noise components to be independent of the rest of the array.

The extra-argument offered by the fourth-order cumulant provides an advantage when the measurements of the satellite sensor are contaminated by non-Gaussian noise which is weak, but is correlated by the noise in the main array. To see this more clearly, if the response of the satellite sensor to the signal is unity, and to noise is $\beta < 1$, then the fourth-order cumulant of non-Gaussian noise will be scaled by $|\beta|^2$ whereas the third-order cumulant will be scaled by only $\beta$; hence, fourth-order cumulants can offer less biased results. Similar arguments were made in Chapter 7 where we compared the bias of second-and fourth-order statistics due to colored non-Gaussian noise by introducing the hill-climbing concept. On the other hand, if the source signals of interest are not symmetric, then third-order cumulants offers insensitivity to a broader class of disturbances since they suppress undesired signals with symmetric probability density functions.

# Chapter 8

# Single Sensor Detection and Classification

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) [162, and the references therein]. 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 are not applicable. This is the problem we address in this chapter.

It is shown that by using multiple frequency lags of the trispectrum of single sensor measurements, it is possible to form a trispectral matrix $\mathbf{C}$ that possesses the same structure as the array covariance matrix of narrowband multisensor measurements. Consequently, techniques that are applicable to narrowband array processing can be adapted for the analysis of single sensor data: rank of $\mathbf{C}$ reveals the number of sources, and a multiple signal characterization (MUSIC)-like method can be used for source classification using a directory of candidate source spectra. Simulations are included to illustrate the proposed methods. We conclude the chapter with several extensions of the proposed method.

## 8.1 Problem Formulation

In this chapter, 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

$$x(t) = \sum_{k=1}^{P} x_k(t) + n(t) \tag{8.1}$$

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

$$x_k(t) = h_k(t) \star u_k(t) \quad k = 1, 2, \ldots, P \tag{8.2}$$

where $\{u_k(t)\}_{k=1}^{P}$ are real, stationary, white, non-Gaussian excitation sequences which are statistically independent among themselves, with variance $\sigma_k^2$ and fourth-order cumulant $\gamma_{4,k}$, and, the filter $h_k(t)$ models the waveform generation process of the $k$th source, with a frequency response $H_k(w)$. The signal model is illustrated in Figure 8.1. The linear process model is used extensively in the cumulant-based time-delay estimation problem [152, and the references therein], which has significant underwater military applications. The algorithms to be proposed in this chapter for source detection and classification will assume the linear model.

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.

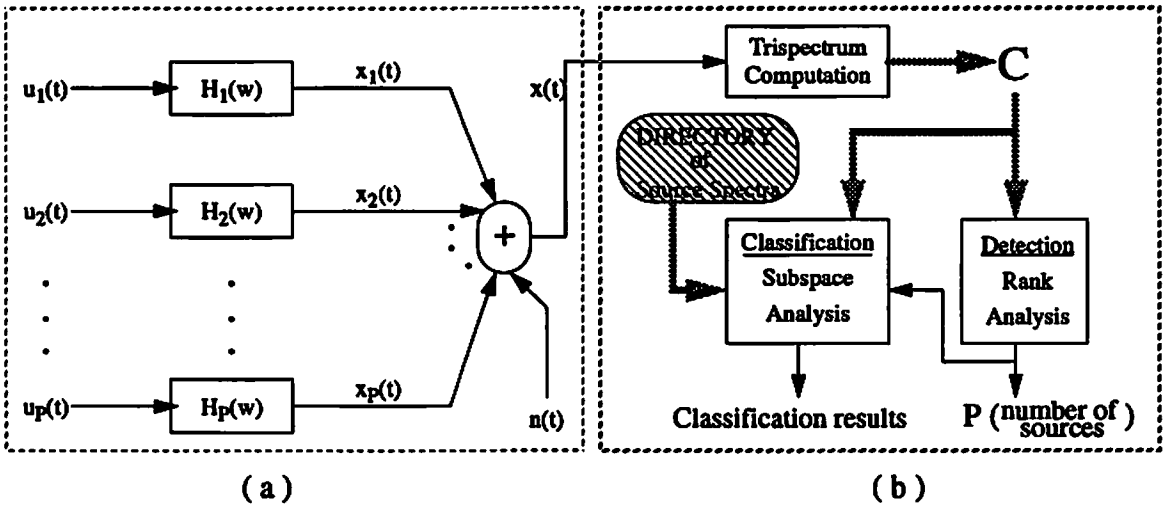



Figure 8.1: (a) Signal generation process, (b) Proposed system.

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

$$S_k(w) = H_k(w)\, H_k^*(w) \tag{8.3}$$

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 alone. To demonstrate this fact more clearly, we consider the vector formulation of the problem. We first define *source spectrum* vector for each source as

$$\mathbf{s}_k \triangleq [S_k(\omega_1), S_k(\omega_2), \ldots, S_k(\omega_M)]^T \quad 1 \leq k \leq P \tag{8.4}$$

in which[1] $\omega_k = 2\pi(k-1)/M$. Due to independence of the non-Gaussian sources and noise, the received signal spectrum ($S_x(\omega)$) is the superposition of the spectra of non-Gaussian sources and measurement noise. Therefore, the received signal spectrum vector (an $M$-vector) can be expressed as

$$\mathbf{s}_x = \sum_{k=1}^{P} \sigma_k^2 \mathbf{s}_k + \sigma_n^2 \mathbf{s}_n \tag{8.5}$$

If all contributing sources (including noise) have their spectrum vectors in our directory, (i.e., full knowledge of spectral shape of all possible sources) and the spectrum vectors are linearly independent, then it is possible to identify the sources, since (8.5) can be uniquely expressed as a linear combination of spectrum vectors. However, with unknown contributions (such as colored noise or a source with unregistered spectrum), which may also have linearly independent spectrum vectors, it is not possible to express (8.5) as a linear combination of known spectrum vectors. This implies that *second-order* statistics of the measurements is *not sufficient* to detect and classify the sources. 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 stop only after

$$\sum_{l=1}^{P+1} \binom{d}{l} \tag{8.6}$$

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 contribute to the received data, and which are 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

[1]For real signals, the Fourier transform has the symmetry property, $H(-w) = H^*(w)$, and 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 to completely characterize it. However, we shall continue to use $M$ for possible extensions to complex signals.

spectrum vectors; therefore, even in this simplified case of known noise spectrum and all registered sources, computations become very excessive.

Next, we consider the trispectrum of the single sensor data $T_x(\omega_1, \omega_2, \omega_3)$, which can be expressed in terms of the trispectra of individual source trispectra $T_k(\omega_1, \omega_2, \omega_3)$ $(1 \leq k \leq P)$ as

$$\begin{aligned} T_x(\omega_1, \omega_2, \omega_3) &= \sum_{k=1}^{P} \gamma_{4,k} T_k(\omega_1, \omega_2, \omega_3) \\ &= \sum_{k=1}^{P} \gamma_{4,k} H_k(\omega_1) H_k(\omega_2) H_k(\omega_3) H_k(-(\omega_1 + \omega_2 + \omega_3)) \end{aligned} \tag{8.7}$$

which can be proved by using the properties of cumulants and polyspectra [78]. The trispectrum of the additive measurement noise $n(t)$ is identically zero by the Gaussianity assumption.

In Chapter 4, we showed how cumulants can be used to increase the *effective* dimensionality of an antenna array; however, the single sensor problem has received very little attention except for [113], 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 of the observed signal, $T_x(\omega_i, \omega_k, \omega_j)$, for which $\omega_k = -\omega_i$. For the multiple sources case, using (8.7), it is possible to obtain

$$T(\omega_i, -\omega_i, \omega_j) = \sum_{k=1}^{P} \gamma_{4,k} \cdot S_k(w_i)\, S_k(w_j) \quad 1 \leq i, j \leq M, \; w_i = 2\pi(i-1)/M. \tag{8.8}$$

Unfortunately, the approach in [113] does not handle measurement noise (since it also employs the spectrum of received data) and is limited by the assumption that it requires one of the sources to have a null in its spectrum where 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.

There are two major approaches to estimate the trispectrum: (1) Direct Method, which utilizes the Fourier Transform samples of the received signal $x(t)$ (this approach is reported in [13] and further analyzed in [113] for the specific slice (submanifold) of (8.8)); (2) Indirect Method, which first computes the cumulants of the

received signal and then computes the Fourier Transform (this approach is recommended in [13] for the computation of polyspectra on submanifolds as in (8.8), and is further analyzed in [67]). We shall use the latter method in the simulations.

In this chapter, we assume the presence of an unknown number of sources. The detection algorithm to be presented in Section 8.3 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.

## 8.2 Analogy with Array Processing

In this section, we construct an analogy between our problem stated in the previous section 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 ($P < M$) with steering vectors $\mathbf{a}_k$. The measured signal $M$-vector $\mathbf{y}(t)$ can be expressed as

$$\mathbf{y}(t) = \sum_{k=1}^{P} \mathbf{a}_k \, y_k(t) \quad + \mathbf{n}(t) \tag{8.9}$$

where $\mathbf{n}(t)$ represents the effects of spatially white measurement noise with power $\sigma^2$, and the source waveforms, $y_k(t)$, are assumed to be not fully correlated (coherent) among themselves [118]. Then the covariance matrix of measurements takes the form

$$\mathbf{R} = \mathbf{A}\,\mathbf{R}_s\,\mathbf{A}^H + \sigma^2\mathbf{I} \tag{8.10}$$

in which $M$x$P$ 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 $M$x$M$ trispectrum matrix $\mathbf{C}$ ($P < M$) that contains samples of the trispectrum of the received signal $x(t)$, as

$$c_{ij} = T(w_i, -w_i, w_j) \quad 1 \le i, j \le M, \;\; w_i = 2\pi(i-1)/M. \tag{8.11}$$

Table 8.1: Analogy between two problems.

| Array Problem | Single Sensor Problem |
|---|---|
| Steering Vector: $\mathbf{a}_k$ | Spectrum Vector: $\mathbf{s}_k$ |
| Steering Matrix: $\mathbf{A}$ | Spectrum Matrix: $\mathbf{S}$ |
| 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: $\mathbf{C}$ |

Then using (8.8) and the definition of source spectrum vectors, we obtain

$$\mathbf{C} = \sum_{k=1}^{P} \gamma_{4,k}\, \mathbf{s}_k\, \mathbf{s}_k^T = \mathbf{S}\mathbf{\Gamma}\mathbf{S}^T \tag{8.12}$$

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 $P$x$P$ diagonal matrix $\mathbf{\Gamma}$ 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 (8.8).
- It is symmetric, i.e., $\mathbf{C} = \mathbf{C}^T = \mathbf{C}^H$. This follows from (8.12).
- Therefore, $\mathbf{C}$ 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 this analogy, we can utilize the detection and direction-finding algorithms already formulated for array processing for the single sensor detection and classification problem. Although the *observation dimensionality* provided by second-order statistics is *inadequate* to solve the problem in a realistic way, multiple lags of the trispectrum enable us to cast the single sensor problem as an array processing problem.

## 8.3 Detection and Classification of Sources

In this section, we propose methods for detection and classification of multiple sources from single sensor data. The proposed methods are based on the analogy developed in the previous section which enable us to adapt standard methods of array processing to single sensor data analysis.

### 8.3.1 Detection of the Number of Sources

It is a common assumption in array processing to have the steering vectors ($M$-vectors) corresponding to the sources illuminating the array be linearly independent. The equivalent assumption for the single sensor trispectral detection problem is that all the source spectrum vectors corresponding to the sources in the field must be linearly independent. For the classification problem, a stronger condition is necessary: in addition to the linear independence of the source spectrum vectors that belong to the contributing sources, no linear combination of the source spectrum vectors of sources present in the field is equal to a spectrum vector in the directory which belongs to a source that is not in the field. Clearly, when this assumption is violated, solution becomes impossible even with the exhaustive search scheme (that requires full information about the source and noise spectra) described in Section 8.1; hence, in this chapter 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 [113].

If the source spectrum vectors are linearly independent (there are $P$ of them), then the rank of the matrix $\mathbf{C}$ must be $P$ (since $P < M$). This follows from the above assumption and (8.12), and is motivated by the similar use of covariance matrix $\mathbf{R}$ in array processing [118]; hence, the number of sources can be detected by computing the rank of $\mathbf{C}$. In practice, $\mathbf{C}$ is always full-rank because of estimation errors due to finite-length data. A test is required to compute the effective rank of $\mathbf{C}$. In this chapter, we accomplish this heuristically; we sort the eigenvalues of $\mathbf{C}$ in a descending order ($\{\lambda_k\}_{k=1}^{M}$) and determine $P$ as the number such that the ratio $(\sum_{k=1}^{P} \lambda_k)/(\sum_{k=1}^{M} \lambda_k)$ first exceeds a threshold less than unity (we use 0.9).

In the proposed detection method, the sources show their effect on the trispectral matrix through their spectrum vectors. The components of the spectrum vectors

are real and positive. Due to this constraint, spectrum vectors must lie in the subset (not subspace) of the $M$ dimensional complex space where each component is real and positive; hence, the angle between any two such vectors is upper bounded when compared to the general unrestricted case. The angle between steering vectors plays an important role in the direction-finding problem; if the angle between steering vectors is small (i.e., sources are very close to each other), the sources may not be resolvable, and their number may be underestimated due to the values of the signal subspace eigenvalues approaching the noise subspace eigenvalues. Similarly, a small angle between spectrum vectors causes underestimation of the number of sources in single channel data; however, there is a way to circumvent this effect by changing (8.11) to

$$\tilde{c}_{ij}(p) = T_x(\omega_i, -\omega_i + 2\pi(p-1)/M, \omega_j) \quad 1 \leq i, j, p \leq M \tag{8.13}$$

Then, instead of (8.12), we have:

$$\tilde{\mathbf{C}}(p) = \sum_{k=1}^{P} \gamma_{4,k} \mathbf{b}_k(p) \mathbf{f}_k^T(p) = \mathbf{B}(p) \mathbf{\Gamma} \mathbf{F}^T(p) \tag{8.14}$$

where $(\mathbf{b}_k(p))_i = H_k(2\pi(i-1)/M) H_k^*(2\pi(i-1-p)/M)$ and $(\mathbf{f}_k(p))_j = H_k(2\pi(j-1)/M) H_k^*(2\pi(j+p)/M)$. Now, $\tilde{\mathbf{C}}(p)$ is not a pseudo-covariance matrix (it is not Hermitian) unless $p = 1$, and an SVD is required instead of an eigendecomposition. Note that the vectors in $\mathbf{B}(p)$ can be complex and they are not constrained to lie in a subset of $M$ dimensional complex space, so that the angles between the vectors can be larger. Here we only use $p = 1$ ($\mathbf{C}$=$\tilde{\mathbf{C}}(1)$) to reduce computation, since ($p = 1$) slice is necessary for classification from spectral information.

### 8.3.2 Classification of Sources

Using the MUSIC algorithm, the DOA's of sources in array processing are determined by a search procedure [118]. 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 $M$x$(M-P)$ matrix $\mathbf{E}_n$ containing the eigenvectors of $\mathbf{C}$, associated with its noise-subspace 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 $k$th source, and compute

$$f(k) = \frac{\mathbf{s}_k^T \mathbf{s}_k}{\mathbf{s}_k^T \mathbf{E}_n \mathbf{E}_n^T \mathbf{s}_k} \quad \text{for } k = 1, 2, \ldots, d. \tag{8.15}$$

The numerator is included to provide normalization. After $f(\cdot)$ is computed for all the sources in the directory, an optional 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 $k$th 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.

## 8.4 Simulations

We tested our method with a directory of size eight. The available spectral information about eight sources is illustrated in Figure 8.2. The last source of the directory is one of the two non-Gaussian sources in the field.

The signal generation process is described in Figure 8.3. The non-Gaussian processes are obtained by passing a Gaussian process through a cubic non-linearity. The measurement noise is colored due to the filter $h_n(l)$. The spectra of non-Gaussian sources, noise and the received signal are illustrated in Figure 8.4. We can observe that the received signal spectrum alone can not help to detect/classify the sources in the field (see the directory in Figure 8.2).

For each realization, we collected 16,384 samples of $x(l)$ and computed the fourth-order cumulant function $C_{4,x}(\tau_1, \tau_2, \tau_3)$ for all $|\tau_1|, |\tau_2|, |\tau_3| \leq 7$. We obtain the trispectrum matrix $\mathbf{C}$ by using Discrete Fourier Transform. We generated 60

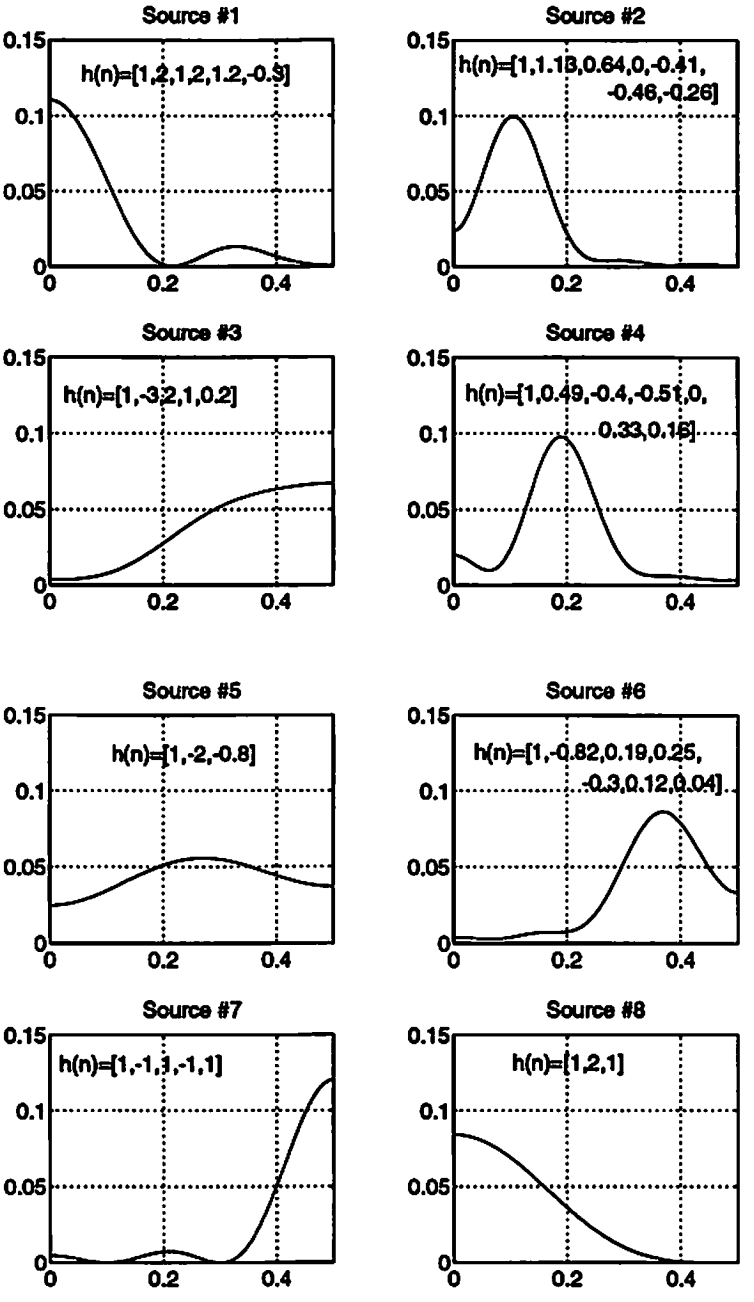



Figure 8.2: Spectral information on the eight sources of our directory. The impulse responses are also given for each source, although this is not required for detection or classification. The last registered source is actually in the field.

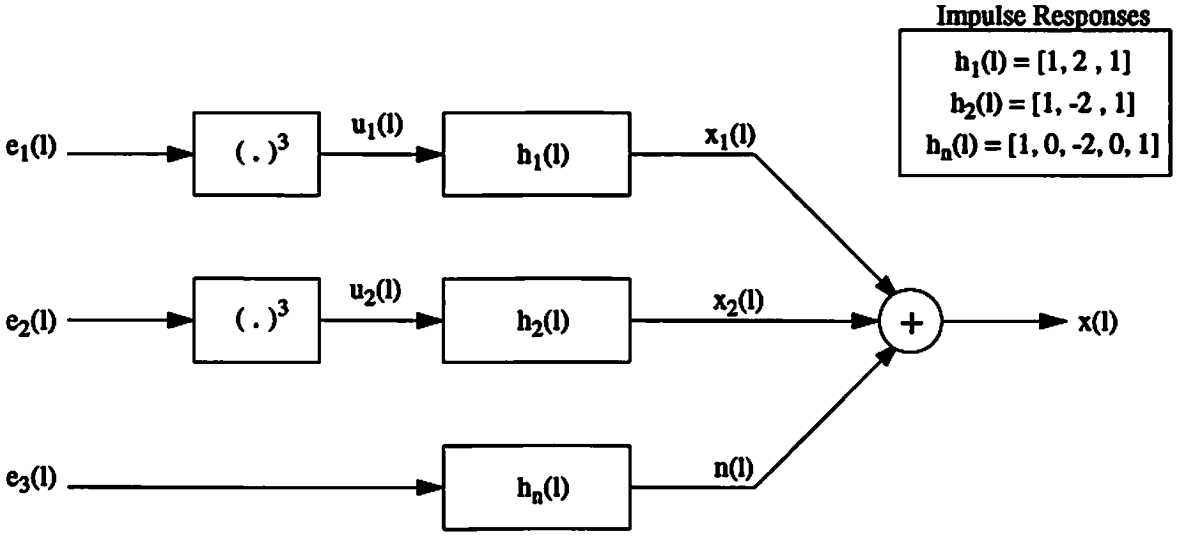



Figure 8.3: Data generation for the experiment. There are two non-Gaussian sources and measurements are corrupted by colored Gaussian noise. The source that contributes $x_1(t)$ to the measurements is the last registered source in the directory. The second non-Gaussian source and the spectral shape of the additive Gaussian noise are unknown to the processor.

Table 8.2: Statistics of the eigenvalues of the trispectrum matrix over 60 realizations.

| Eigenvalue Number | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| Mean | 0.5929 | 0.3509 | 0.0238 | 0.0130 | 0.0091 | 0.0059 | 0.0033 | 0.0012 |
| Standard Deviation | 0.0880 | 0.0792 | 0.0115 | 0.0044 | 0.0031 | 0.0018 | 0.0014 | 0.0011 |

independent realizations for the experiment. Table 8.4 presents the statistics of the eigenvalues of the trispectrum matrix indicating two principal eigenvalues, i.e., $P = 2$.

After determining the presence of two non-Gaussian sources, we use our directory of spectral information (Figure 8.2) to classify the sources. Table 8.4 illustrates the variation of the membership functions over 60 realizations for the eight sources. For each realization, we normalized the membership function to have a maximum of unity. We observe that there is one significant peak in the membership functions which corresponds to the last source of the directory (which is the first field source). The standard deviation for directory source 8 is zero, indicating that the membership

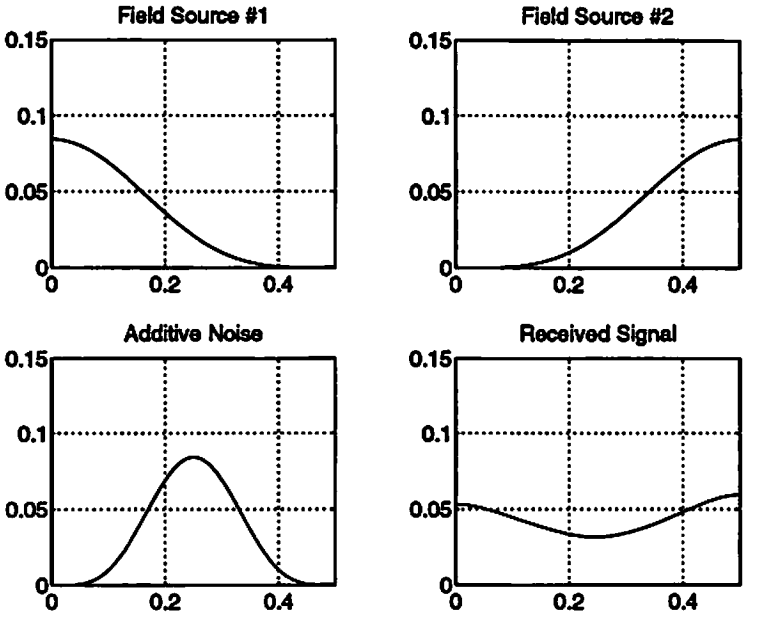



Figure 8.4: Spectra of the sources and noise in the field, and the spectrum of the received data estimated from a single realization.

function peaks at 8 for all realizations. The membership function for the third directory source is slightly larger than that of the directory sources because of the small angle between the spectrum vector of source 3 and that of the unknown source (see Figure 8.2 and Figure 8.4).

Table 8.3: Variation of membership functions over 60 realizations. The last source (the only one in the field from the directory) yields a significant peak when compared with the other sources of the directory.

| Source Number | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| Mean | 0.0537 | 0.0388 | 0.1464 | 0.0148 | 0.0493 | 0.0499 | 0.0367 | **1.0000** |
| Standard Deviation | 0.0541 | 0.0542 | 0.1589 | 0.0153 | 0.0532 | 0.0512 | 0.0323 | **0** |

## 8.5 Conclusions

We have established a framework for the analysis of single channel, multicomponent data using higher-order statistics. We have shown that by using only second-order statistics, it is impossible to detect and classify multiple sources from single channel data, and that, by using higher than second-order statistics, it is possible to form an analogy with the direction-of-arrival estimation problem in array processing, and accomplish single channel detection and classification for multiple sources.

Generally, cumulants are used in place of covariance information within popular algorithms such as ESPRIT, with the hope of suppressing the effects of Gaussian noise of arbitrary covariance. In this chapter, we suggest another viewpoint for the use of the extra information supplied by cumulants, namely: cumulants (or, equivalently, polyspectra) possess an *array of lags* unlike second-order statistics. When properly organized, signal processing tasks that are impossible to accomplish with just second-order statistics alone are possible with higher-order statistics. Furthermore, another important interpretation is: since cumulants possess an array of lags, it is possible to treat HOS information *using array processing techniques* to acquire more information than is available from cumulants.

The remaining topic for single sensor data analysis is to develop detection tests based on asymptotical properties of trispectrum estimates rather than heuristic ones proposed in this chapter. In the time-series analysis context, detection algorithms based on cumulants are described in [49, and the references therein].

It is possible to extend the developed framework to other areas such as time-delay estimation. This is carried out in the Appendix.

## 8.6 Appendix–Extensions

This section provides several extensions of the methods developed for single sensor data analysis.

### 8.6.1 Single Source Scenario

In the single source case, the trispectral matrix defined in Section 8.2 is rank one, and its principal eigenvector is the spectrum vector of the non-Gaussian source. It is possible to develop a non-parametric detection test based on the observation that the trispectral matrix is null under the no source hypothesis. When the source is present hypothesis is true, the principal eigenvector will be a smoothed estimate of the spectrum of the non-Gaussian signal over trispectral slices.

### 8.6.2 Blind Spectrum Estimation for Multiple Sources

If we resort to higher-than fourth-order cumulants, it is also possible to estimate the spectral shape of sources blindly by using subspace rotation direction-finding. Before illustrating how to do this by higher than fourth-order statistics, we first describe how fourth-order cumulants can solve the problem under some special conditions.

Consider two sets of data, collected at different times from the available single channel. We assume the excitation source powers (and hence cumulants) are scaled during the interval between the two collections. For the first set, we have

$$\mathbf{C}_1 = \mathbf{S}\boldsymbol{\Gamma}\mathbf{S}^T \tag{8.16}$$

In the second collection, excitation sequence is scaled (channels are excited with different power), which yields the trispectrum matrix:

$$\mathbf{C}_2 = \mathbf{S}\mathbf{D}\boldsymbol{\Gamma}\mathbf{S}^T \tag{8.17}$$

where $\mathbf{D}$ is a diagonal matrix that contains the scaling factors. We can consider the matrices in (8.16) and (8.17) as the covariance matrices required by the ESPRIT algorithm, and recover the spectral shapes of the sources.

The use of higher-than fourth-order cumulants enables us to generate the matrices in (8.16) and (8.17) without requiring two collection times. For example, if

$F_x(\omega_1, \omega_2, \omega_3, \omega_4)$ denotes the tetraspectrum (Fourier Transform of the fifth-order cumulant function) of $x(t)$, then the following submanifold

$$F_x(\omega_i, -\omega_i, \omega_j, -\omega_j, 0) = \sum_{k=1}^{P} \gamma_{5,k}\, H_k(\omega_i) H_k(-\omega_i) H_k(\omega_j) H_k(-\omega_j) H(0) \qquad (8.18)$$

can be expressed as

$$F_x(\omega_i, -\omega_i, \omega_j, -\omega_j, 0) = \sum_{k=1}^{P} \gamma_{5,k} H_k(0) S_k(\omega_i) S_k(\omega_j) \qquad (8.19)$$

An alternative expression is

$$F_x(\omega_i, -\omega_i, \omega_j, -\omega_j, 0) = \sum_{k=1}^{P} \frac{\gamma_{5,k} H_k(0)}{\gamma_{4,k}}\, \gamma_{4,k} S_k(\omega_i) S_k(\omega_j) \qquad (8.20)$$

We can construct an $M$x$M$ matrix $\mathbf{C}_2$ from the samples of the tetraspectrum slice on the manifold (8.20) by varying $\omega_i$ and $\omega_j$. Comparing (8.20) with (8.8), we observe that $\mathbf{C}_2$ can be decomposed as

$$\mathbf{C}_2 = \mathbf{S}\mathbf{D}\boldsymbol{\Gamma}\mathbf{S}^T \qquad (8.21)$$

where the trispectrum matrix was decomposed as $\mathbf{S}\boldsymbol{\Gamma}\mathbf{S}^T$. $\mathbf{D}$ is a diagonal matrix, $\mathbf{D} = \text{diag}(\frac{\gamma_{5,1} H_1(0)}{\gamma_{4,1}}, \ldots, \frac{\gamma_{5,P} H_P(0)}{\gamma_{4,P}})$. Now, the ESPRIT algorithm can be used to analyze $\mathbf{C}$ and $\mathbf{C}_2$, for estimating the spectral shapes contained in $\mathbf{S}$.

### 8.6.3 Time-Delay Estimation for Multiple Sources

Time-delay estimation using higher-order statistics is motivated to suppress the effects of additive spatially-colored Gaussian noise [60, 83, 114, 115, 152, 178]. In this section, we propose the use of cumulants in the case of multiple sources instead of a single non-Gaussian source. Let us assume that $\{r_1(t), r_2(t)\}$ are two sensor measurements satisfying

$$r_1(t) = \sum_{k=1}^{P} x_k(t) \; + n_1(t)$$

$$r_2(t) = \sum_{k=1}^{P} x_k(t - D_k) \; + n_2(t) \tag{8.22}$$

where $\{x_k(t)\}_{k=1}^{P}$ are non-Gaussian, stationary, linear processes, see (8.2). The additive Gaussian noise sources $\{n_1(t), n_2(t)\}$ may have unknown spectra.

Let us consider the following fourth-order cross-cumulant function for two zero-mean, stationary random processes, $a(t)$ and $b(t)$:

$$f_{a,b}(\tau_1, \tau_2, \tau_3) \triangleq \text{cum}(a(t), a(t+\tau_1), b(t+\tau_2), b(t+\tau_3)) \tag{8.23}$$

The corresponding cross-trispectrum is the 3-dimensional Fourier transform:

$$T_{a,b}(\omega_1, \omega_2, \omega_3) \triangleq \sum_{\tau_1,\tau_2,\tau_3=-\infty}^{\infty} f_{a,b}(\tau_1, \tau_2, \tau_3) \exp(-j(\omega_1\tau_1 + \omega_2\tau_2 + \omega_3\tau_3)) \tag{8.24}$$

Following the signal model in (8.22), the following holds:

$$T_{r_1,r_1}(\omega_1, \omega_2, \omega_3) = \sum_{k=1}^{P} \gamma_{4,k} \, H_k(\omega_1) H_k(\omega_2) H_k(\omega_3) H_k(-(\omega_1 + \omega_2 + \omega_3)) \tag{8.25}$$

$$T_{r_1,r_2}(w_1, w_2, w_3) = \sum_{k=1}^{P} \gamma_{4,k} \, H_k(\omega_1) H_k(\omega_2) H_k(\omega_3) H_k(-(\omega_1 + \omega_2 + \omega_3)) \, e^{-j(\omega_2+\omega_3) D_k} \tag{8.26}$$

In practical applications polyspectra are computed only at discrete points in the frequency domain which are located at multiples of $2\pi/M$, in which $M$ corresponds to the Fourier analysis window length. The frequency variable $w$ takes on the values in the range $\{2\pi(k-1)/M\}_{k=1}^{M}$. Consider the functions

$$g(\omega_1, \omega_2) \triangleq T_{r_1,r_1}(\omega_1, \omega_2, 2\pi/M - \omega_2) \tag{8.27}$$

$$h(\omega_1, \omega_2) \triangleq T_{r_1,r_2}(\omega_1, \omega_2, 2\pi/M - \omega_2) \tag{8.28}$$

which can be expressed as

$$g(\omega_1, \omega_2) = \sum_{k=1}^{P} \gamma_{k,4} H_k(\omega_1) H_k(\omega_2) H_k(2\pi/M - \omega_2) H_k(-(\omega_1 + 2\pi/M)) \tag{8.29}$$

$$h(\omega_1,\omega_2)=\sum_{k=1}^{P}\gamma_{k,4}H_k(\omega_1)H_k(\omega_2)H_k(2\pi/M-\omega_2)H_k(-(\omega_1+2\pi/M))\cdot e^{-j2\pi D_k/M} \tag{8.30}$$

Let us consider the vectors $\{\mathbf{s}_k,\mathbf{v}_k\}_{k=1}^{P}$ defined as follows:

$$\mathbf{s}_k(l)\triangleq H_k(\omega_1)H_k(-(\omega_1+2\pi/M))|_{\omega_1=2\pi(l-1)/M} \tag{8.31}$$

$$\mathbf{v}_k(l)\triangleq H_k(\omega_2)H_k(2\pi/M-\omega_2)|_{\omega_2=2\pi(l-1)/M} \tag{8.32}$$

Now we form two matrices, $\mathbf{G}$ and $\mathbf{H}$, where $g_{ij}\triangleq g(2\pi(i-1)/M,2\pi(j-1)/M)$ and $h_{ij}\triangleq h(2\pi(i-1)/M,2\pi(j-1)/M)$. Then, it is possible to express these matrices as

$$\mathbf{G}=\sum_{k=1}^{P}\gamma_{4,k}\mathbf{s}_k\mathbf{v}_k^T \tag{8.33}$$

$$\mathbf{H}=\sum_{k=1}^{P}\gamma_{4,k}\mathbf{s}_k\exp(-j\beta_k)\mathbf{v}_k^T \tag{8.34}$$

in which $\beta_k\triangleq 2\pi D_k/M$. In matrix form, if $\mathbf{S}\triangleq[\mathbf{s}_1,\mathbf{s}_2,\ldots,\mathbf{s}_P]$ and $\mathbf{V}\triangleq[\mathbf{v}_1,\mathbf{v}_2,\ldots,\mathbf{v}_P]$ with $\mathbf{\Gamma}\triangleq\mathrm{diag}(\gamma_{4,1},\ldots,\gamma_{4,P})$ and $\mathbf{D}\triangleq\mathrm{diag}(\exp(-j\beta_1),\ldots,\exp(-j\beta_P))$, we have

$$\mathbf{G}=\mathbf{S}\mathbf{\Gamma}\mathbf{V}^T \tag{8.35}$$

$$\mathbf{H}=\mathbf{S}\mathbf{D}\mathbf{\Gamma}\mathbf{V}^T \tag{8.36}$$

which implies

$$\begin{bmatrix}\mathbf{G}\\ \mathbf{H}\end{bmatrix}=\begin{bmatrix}\mathbf{S}\\ \mathbf{S}\mathbf{D}\end{bmatrix}\mathbf{\Gamma}\mathbf{V}^T \tag{8.37}$$

We observe that (8.37) is identical in form to the signal subspace provided by the ESPRIT data model (see (2.40)) provided that the matrices $\mathbf{S}$ and $\mathbf{V}$ are full-rank. In (2.40), $\mathbf{\Phi}$ contains the DOA parameters and plays the same role as $\mathbf{D}$ does in (8.37). Therefore, ESPRIT algorithm can be used to determine the components of the diagonal matrix that correspond to the time-delays encountered by the sources.

# Chapter 9

# Conclusions

In this thesis, we have shown that the use of higher-order statistics can provide solutions to important array processing problems. The multiple arguments of cumulants provide additional information about the phase of source wavefronts. In this respect, use of cumulants instead of correlation seems analogous to the use of antenna arrays rather than a single sensor; when one uses an array it is possible to overcome the sensitivity limitations of a single sensor, when one uses higher-order statistics, it is possible to eliminate the phase recovery limitations of cross-spectrum. In the single channel problem, spectrum can not provide phase information whereas polyspectra can.

Computation of cumulants (especially, fourth-order cumulants) is more expensive than computation of second-order statistics; however, the former can be parallelized. In array processing applications, the main computational burden that causes processing delay is the eigenanalysis of covariance and cumulant matrices. Provided that these matrices are of the same size, the major computational costs for their eigenanalyses are identical.

In Chapter 4, we provided an interpretation for the cumulants of narrowband sensor measurements. We showed that fourth-order cumulants *double* the phase information available from the sensor measurements; fourth-order cumulants are interpreted as interpreted as two "covariance vectors". We introduced the concept of virtual sensors, and explained how cumulants can be used to compute cross-correlations among actual and virtual sensors to increase the effective aperture of the array.

In Chapter 5, we addressed the joint array calibration and direction-finding problem using an arbitrary antenna array. In this problem, an identical copy of the array was needed so that the ESPRIT algorithm might be applicable. We considered extra sensors required by the ESPRIT algorithm as virtual sensors and proved that, using a single doublet, it is possible to compute all the cross-correlations required by the ESPRIT algorithm by means of fourth-order cumulants. The resulting algorithm was called as the *virtual*-ESPRIT algorithm (VESPA). It is also possible to estimate the directions, steering vectors, and the waveforms of the non-Gaussian sources using cumulants and a doublet. We provided extensions to indicate how VESPA can be applied in wideband scenarios, near-field sources, and using third-order cumulants.

In Chapter 6, we determined bounds on effective aperture extension by using cumulants. We proved that given an array of $M$ identical sensors, fourth-order cumulants can at least double the aperture $2M-1$ regardless of the location of the sensors, with an upper bound of $M^2-M+1$. These bounds are derived for existing cumulant-based algorithms, and therefore the upper bound was exceeded when we used minimum-redundancy array design concepts together with cumulants. We proposed designs for both linear and two-dimensional arrays. The effective number of sensors provided by our designs are proportional to $M^4$.

Cumulants have long been promoted in signal processing applications for their ability to suppress additive Gaussian observation noise. In Chapter 7, we showed that it is possible to suppress additive *non*-Gaussian noise if there exists a sensor far-enough from the main array whose noise component can be non-Gaussian, but is independent from the noise components in the main array sensors. We proved that our approach of computing cross-correlations using cumulants suppresses the effects of additive noise even when the sources of interest are correlated. We also indicated the noise suppression capabilities of the virtual-ESPRIT algorithm.

Finally, in Chapter 8, we addressed the problem of single sensor detection and classification of multiple linear non-Gaussian processes. This problem was solved by exploiting the fact that polyspectra possess an *array* of arguments, unlike spectrum. We showed how to form a trispectral matrix that has the same structure as the array covariance matrix in the narrowband array processing problem. Techniques that are well known in array processing were then applied for source detection and classification. This approach completed the "circle" in our array processing and

cumulants research, by showing how array processing techniques can be used to extract information from cumulants, instead of using cumulants to extract extra information for array processing problems. We indicated extensions of this approach for time-delay estimation and blind spectrum estimation problems.

Use of cumulants for array processing applications is an important research area, and several issues remain to be explored:

- Asymptotical performance analysis of the virtual-ESPRIT algorithm for direction-finding and signal recovery purposes. In [81], a performance analysis is carried out to evaluate the behavior of a cumulant-based direction-finding algorithm. In [22], it is proved that signal recovery by a cumulant-based method provides better results than an informed beamformer that uses the steering vector of the source of interest. This fact is also observed in [34]. Although algorithm development is easier with Kronecker products [12], performance analysis is simpler with tensors [22].

- Extension of the virtual-ESPRIT algorithm to correlated and/or coherent sources. Use of multiple guiding sensors seems to be a good starting point to address this problem.

- Development of algorithms for randomly perturbed arrays, for which sensor locations change randomly over the observation interval [125]. Investigation of the possibility of "virtually stationarizing" the array by a doublet which is stationary is an important issue.

- Development of adaptive virtual subspace rotation algorithms for tracking moving non-Gaussian sources without array calibration.

- Development of computationally efficient, cumulant-based algorithms for azimuth/elevation and polarization estimation. This was briefly discussed in Chapter 5.

- Extending virtual-ESPRIT algorithm to specific array geometries such as, rotationally-invariant arrays [142], and, arrays that satisfy original ESPRIT [108] requirements.

- Blind recovery of weak signals under strong jamming using multiple arguments of cumulants.

- Development and assessment of new methods for time-delay estimation using the framework developed for single sensor data analysis.

Although we showed how cumulants provide ways to handle important problems in array processing, we certainly do not conclude that second-order statistics should be abandoned. Actually, whenever applicable, use of second-order statistics can improve the estimation results, as indicated by the SFS method of Chapter 7.

# Reference List


[1] T.J. Abatzoglou, J.M. Mendel and G.A. Harada, "The constrained total least squares technique and its applications to harmonic superresolution," *IEEE Trans. Signal Processing*, vol.39, no.5, pp.1070–1087, May 1991.

[2] B.G. Agee, S.V. Schell and W.A. Gardner, "Spectral self-coherence restoral: a new approach to blind adaptive signal extraction using antenna arrays," *Proceedings of IEEE*, vol.78, no.4, pp.753–767, April 1990.

[3] K. Bahkru and D.J. Torrieri, "The maximin algorithm for adaptive arrays and frequency-hopping communications," *IEEE Trans. Antennas and Propagation*, vol.32, no.9, pp.919–927, September 1984.

[4] G. Bienvenu and L. Kopp, "Optimality of high resolution array processing using the eigensystem approach," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.31, no.5, pp.1235–1247, October 1983.

[5] J.F. Böhme, "Estimation of spectral parameters of correlated signals in wavefields," *Signal Processing*, vol.11, pp.329–337, 1986.

[6] R.N. Bracewell, "Radio astronomy techniques," in *Handbuch der Physik*, vol.54, pp.42-129, Berlin:Springer 1962.

[7] D. H. Brandwood, "A complex gradient operator and its application in adaptive array theory," *IEE Proceedings*, vol.130, *Pts. F and H*, no.1, pp.11-16, February 1983.

[8] D.G. Brennan, "Linear diversity combining techniques," *Proceedings of IRE*, vol.47, no.6, pp.1075–1102, June 1959.

[9] Y. Bresler and A. Macovski, "Exact maximum-likelihood parameter estimation of superimposed exponential signals in noise," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.34, no.5, pp.1081–1089, October 1986.

[10] Y. Bresler and A. Macovski, "On the number of signals resolvable by a uniform linear array," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.34, no.6, pp.1361–1375, December 1986.

[11] Y. Bresler, V.U. Reddy and T. Kailath, "Optimum beamforming for coherent signal and interferences," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.36, no.6, pp.833–843, June 1988.

[12] J.W. Brewer, "Kronecker products and matrix calculus in system theory," *IEEE Trans. Circuits Systems*, vol.25, no.9, pp.772–781, September 1978.

[13] D.R. Brillinger and M. Rosenblatt, "Asymptotic theory of estimates of $k$th-order spectra," in *Spectral Analysis of Time Series*, B. Harris, ed., New-York: John Wiley & Sons, pp.189–232, 1967.

[14] D.R. Brillinger. *Time-Series: Data Analysis and Theory.* Holt, Rinehart and Winston, Inc., New-York, 1975.

[15] L.W. Brooks and I.S. Reed, "The equivalence of the likelihood ratio processor, the maximum signal-to-noise ratio filter and the Wiener filter," *IEEE Trans. Aerospace and Electronic Systems*, vol.8, no.5, pp.690–691, September 1972.

[16] K.M. Buckley and L.J. Griffiths, "An adaptive generalized sidelobe canceler with derivative constraints," *IEEE Trans. Antennas and Propagation*, vol.34, no.3, pp.311–319, March 1986.

[17] W.S. Burdic. *Underwater Acoustic System Analysis.* Prentice-Hall, Englewood Cliffs, New-Jersey, 1984.

[18] J.A. Cadzow, "A high-resolution direction-of-arrival algorithm for narrowband coherent and incoherent sources," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.34, no.7, pp.965–979, July 1988.

[19] J. Capon, "High-resolution frequency-wavenumber spectral analysis," *Proceedings of IEEE*, vol.57, no.8, pp.1408–1418, August 1969.

[20] J.F. Cardoso, "Blind identification of independent components with higher-order statistics," *Proc. Vail Workshop on Higher-Order Spectral Analysis*, pp.157–162, June 1989.

[21] 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.

[22] J.F. Cardoso and A. Souloumiac, "An efficient techniques for the blind separation of complex sources," *Third International Workshop on Higher-Order Spectral Analysis*, pp.275–279, Lake Tahoe, USA, June 1993.

[23] V. Chandran and S.L. Elgar, "Pattern recognition using invariants defined from higher order spectra—one-dimensional inputs," *IEEE Trans. Signal Processing*, vol.41, no.1, pp.194–204, January 1993.

[24] H.H. Chiang and C.L. Nikias, "The ESPRIT algorithm with higher-order statistics," *Proc. Vail Workshop Higher-Order Spectral Analysis*, pp.163–168, June 1989.

[25] P. Comon, "Separation of stochastic processes," *Proc. Vail Workshop on Higher-Order Spectral Analysis*, pp.174–179, June 1989.

[26] R.T. Compton, "An adaptive array in a spread-spectrum communication system," *Proceedings of IEEE*, vol.66, no.3, pp.289–298, March 1978.

[27] R.T. Compton, "Pointing accuracy and dynamic range in a steered beam adaptive array," *IEEE Trans. Aerospace and Electronic Systems*, vol.16, no.3, pp.280–287, May 1980.

[28] R.T. Compton, "The effect of random steering error vectors in the Applebaum adaptive array," *IEEE Trans. Aerospace and Electronic Systems*, vol.18, no.5, pp.392–400, September 1982.

[29] R.T. Compton, *Adaptive Antennas: Concepts and Performance.* Prentice-Hall, New-Jersey, 1988.

[30] J.P. Costas, "A study of a class of detection waveforms having nearly ideal range-doppler ambiguity properties," *Proceedings of IEEE*, vol.72, no.8, pp.996–1009, August 1984.

[31] H. Cox, H.M. Zeskind and M.M. Owen, "Robust adaptive beamforming," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.35, no.10, pp.1365–1376, October 1987.

[32] H. Cox, H.M. Zeskind and M.M. Owen, "Effects of amplitude and phase errors on linear predictive array processors," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.36, no.1, January 1988.

[33] S. DeGraaf and D.H. Johnson, "Capability of array processing algorithms to estimate source bearings," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.33, no.6, pp.1368–1379, December 1985.

[34] M.C. Doğan and J.M. Mendel, "Cumulant-based blind optimum beamforming," *IEEE Trans. Aerospace and Electronic Systems*, April 1994.

[35] T.S. Durrani, A.R. Leyman and J.J. Soraghan, "New algorithms for array processing using higher-order statistics," *Proc. ICASSP-93*, vol.4, pp.500–503, 1993.

[36] P. Erdös and I.S. Gál, "On the representation of 1,2,$\cdots$,N by differences," *Nederl. Akad. Wetensch. Proc.*, vol.51, pp.1155-1148, 1948.

[37] D. Feldman and L.J. Griffiths, "A constraint projection approach for robust adaptive beamforming," *in Proc. IEEE Intl. Conf. Acoust., Speech, Signal Processing*, vol.2, pp.1381–1384, May 1991.

[38] D. Feldman and L.J. Griffiths, "A performance analysis of robust constrained beamforming with random errors," *Proc. ICAASP-93*, vol.4, pp.540–543, 1993.

[39] E.R. Ferrera and T.W. Parks, "Direction finding with an array of antennas having diverse polarizations," *IEEE Trans. Antennas and Propagation*, vol.31, no.2, pp.231–236, February 1983.

[40] J.A.R. Fonollosa, J. Vidal and E. Masgrau, "Adaptive system identification based on higher order statistics," *in Proc. IEEE Intl. Conf. Acoust., Speech, Signal Processing*, pp.3437–3440, 1991.

[41] P. Forster and C.L. Nikias, "Bearing estimation in the bispectrum domain," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.39, no.9, pp.1994–2006, September 1991.

[42] B. Friedlander, "A signal subspace method for adaptive interference cancellation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.36, no.12, pp.1835–1845, December 1988.

[43] B. Friedlander and B. Porat, "Performance analysis of a null-steering algorithm based on direction-of-arrival estimation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.37, no.4, pp.461–466, April 1989.

[44] B. Friedlander, "A sensitivity analysis of the MUSIC algorithm," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.38, no.10, pp.1740–1751, October 1990.

[45] O.L. Frost, "An algorithm for linearly constrained adaptive array processing," *Proceedings of IEEE*, vol.60, no.8, pp.926–935, August 1972.

[46] W. Gabriel, "Spectral analysis and adaptive array superresolution techniques," *Proceedings of IEEE*, vol.68, no.6, pp.654–666, June 1980.

[47] W. Gabriel, "Using spectral estimation techniques in adaptive processing antenna systems," *IEEE Trans. Antennas and Propagation*, vol.34, no.4, pp.291–300, March 1986.

[48] G.B. Giannakis and M. Tsatsanis, "HOS or SOS for parametric modelling," *Proc. IEEE Intl. Conf. on Acoust., Speech, Signal Processing*, vol.5, pp.3097–3100, May 1991.

[49] G.B. Giannakis and S. Shamsunder, "Information theoretic criteria for non-Gaussian ARMA order determination and parameter estimation," *Proc. ICASSP-93*, vol.4, pp.196–199, Minnesota, USA, April 1993.

[50] L.C. Godara, "Error analysis of optimal antenna array processors," *IEEE Trans. Aerospace and Electronic Systems*, vol.22, no.4, pp.395–409, July 1986.

[51] D.N. Godard, "Self-recovering equalization and carrier tracking in a two-dimensional data communication system," *IEEE Trans. Communications*, vol.28, no.11, pp.1867–1875, November 1980.

[52] S.W. Golomb, "Algebraic constructions for Costas arrays," *Journal of Combinatorial Theory*, Series A, vol.37, no.1, pp.13–21, July 1984.

[53] G.H. Golub and C.F. Van Loan, *Matrix Computations*. John Hopkins University Press, Baltimore, 1983.

[54] L.J. Griffiths and C.W. Jim, "An alternative approach to linearly constrained adaptive beamforming," *IEEE Trans. Antennas and Propagation*, vol.30, no.1, pp.27–34, January 1982.

[55] S. Haykin, (Editor). *Array Processing: Applications to Radar*. Dowden, Hutchinson & Ross, Inc., Stroudsberg, Pennsylvania, 1980.

[56] S. Haykin, (Editor). *Array Signal Processing*. Prentice-Hall, New-Jersey, 1984.

[57] S. Haykin. *Communication Systems*, Second Edition. John Wiley & Sons Inc., New-York, 1983.

[58] S. Haykin, (Editor). *Advances in Spectrum Estimation and Array Processing*. Volume 2. Prentice-Hall, New-Jersey, 1991.

[59] S. Haykin, J.P. Reilly, V. Kezys and E. Vertatschitsch, "Some aspects of array signal processing," *IEE Proceedings-F*, vol.139, no.1, pp.1–26, February 1992.

[60] M.J. Hinich and G.R. Wilson, "Time delay estimation using the cross-bispectrum," *IEEE Trans. Signal Processing*, vol.40, no.1, pp.106–113, January 1992.

[61] X.P. Huang, J.P. Reilly and K.M. Wong, "Optimal design of linear arrays of sensors," *Proc.ICASSP-91*, vol.2, pp.1405–1408, May 1991.

[62] H. Hung and M. Kaveh, "Focusing matrices for coherent signal subspace processing," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.36, no.8, pp.1272–1281, August 1988.

[63] D.H. Johnson and S. DeGraaf, "Improving the resolution of bearing in passive sonar arrays by eigenvalue analysis," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.30, no.4, pp.638–647, August 1982.

[64] D.H. Johnson and D.E. Dudgeon. *Array Signal Processing.* Prentice-Hall, New-Jersey, 1993.

[65] R. Johnson and G. Miner, "An operational system implementation of the ESPRIT DF algorithm," *IEEE Trans. Aerospace and Electronic Systems*, vol.27, no.1, pp.159-166, January 1991.

[66] M. Kaveh and A.J. Barabell, "The statistical performance of the MUSIC and the minimum-norm algorithms in resolving plane waves in noise," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.34, no.2, pp. 331–341, April 1986.

[67] P.T. Kim, "Consistent estimation of fourth-order cumulant spectral density," *Journal of Time Series Analysis*, vol.12, no.1, pp.63–71, January 1991.

[68] J. Krolik and D.N. Swingler, "Multiple broadband source location using steered covariance matrices," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.37, no.10, pp.1481–1494, October 1989.

[69] J. Krolik and D.N. Swingler, "Focused wideband array processing via spatial resampling," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.38, no.2, pp.356–360, February 1990.

[70] R. Kumaresan and D. Tufts, "Estimating the angles of arrival of multiple plane waves," *IEEE Trans. Aerospace and Electronic Systems*, vol.19, no.1, pp.134–139, January 1983.

[71] S.W. Lang, G.L. Duckworth and J.H. McClellan, "Array design for MEM and MLM array processing," *Proc.ICASSP-81*, vol.1, pp.145-148, March 1981.

[72] R.E. Leahy and B.D. Jeffs, "On the design of maximally sparse beamforming arrays," *IEEE Trans. Antennas and Propagation*, vol.39, no.8, pp.1178–1187, August 1991.

[73] J. Leech, "On the representation of 1,2,$\cdots$,n by differences," *J. London Math. Soc.*, vol.31, pp.160–169, 1956.

[74] K.S. Lii and M. Rosenblatt, "Deconvolution and estimation of transfer function phase and coefficients for non-Gaussian linear processes," *Annals of Statistics*, vol.10, no.4, pp.1195-1208, December 1982.

[75] D.A. Linebarger, I.H. Sudborough, I.G. Tollis, "Difference bases and sparse sensor arrays," *IEEE Trans. Information Theory*, vol.39, no.2, pp.716–721, March 1993.

[76] Y.T. Lo and S.L. Marple Jr., "Observability conditions multiple signal direction finding and array sensor localization," *IEEE Trans. Signal Processing*, vol.40, no.11, pp.2641–2650, November 1992.

[77] J. Marr, "A selected bibliography on adaptive antenna arrays," *IEEE Trans. Aerospace and Electronic Systems*, vol.22, no.6, pp.781–798, November 1986.

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

[79] A.T. Moffet, "Minimum-redundancy linear arrays," *IEEE Trans. Antennas and Propagation*, vol.16, no.2, pp.172–175, March 1968.

[80] R.A. Monzingo and T.W. Miller. *Introduction to Adaptive Arrays.* John-Wiley & Sons, Inc., New-York, 1980.

[81] E. Moulines and J.F. Cardoso, "Direction-finding algorithms using fourth-order statistics: asymptotic performance analysis," *Proc. ICASSP-92*, vol.2, pp.437–440, March 1992.

[82] C.L. Nikias and M.R. Raghuveer, "Bispectrum estimation: a digital signal processing framework," *Proceedings of IEEE*, vol.75, no.7, pp.869–891, July 1987.

[83] C.L. Nikias and R. Pan, "Time-delay estimation in unknown Gaussian spatially correlated noise," *IEEE Trans. Signal Processing*, vol.36, no.11, pp.1706–1714, November 1988.

[84] C.L. Nikias and A.P. Petropulu. *Higher-Order Spectra Analysis: A Nonlinear Signal Processing Framework.* Prentice-Hall, New-Jersey, 1993.

[85] C.Y. Ngo and J.M. Mendel, "A higher-moment formula for non-zero-mean AR processes," *Proc. IEEE Intl. Conf. on Acoust., Speech, Signal Processing*, vol.5, pp.3089–3096, May 1991.

[86] J.M. Ortega, *Matrix Theory: a second course.* Plenum Press, New-York, 1987.

[87] B. Ottersten and L. Ljung, "Asymptotic results for sensor array processing," *Proc. ICASSP-89*, vol.4, pp.2266–2269, May 1989.

[88] B. Ottersten, and T. Kailath, "Performance analysis of the total-least-squares ESPRIT algorithm," *IEEE Trans. Signal Processing*, vol.39, no.5, pp.2436–2449, May 1991.

[89] B. Ottersten, M. Viberg and B. Wahlberg, "Robust source localization based on array response modeling," *Proc. ICASSP-92*, vol.2, pp.441–444, March 1992.

[90] A. Paulraj and T. Kailath, "Eigenstructure methods for direction of arrival estimation in the presence of unknown noise fields," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.34, no.1, pp.13–20, February 1986.

[91] R. Pan and C.L. Nikias, "Harmonic decomposition methods in cumulant domains," *Proc. ICASSP'88*, pp. 2356–2359, New-York, NY, April 1988.

[92] H. Parthasarathy, S. Prasad and S.D. Joshi, "An ESPRIT-like algorithm for the estimation of quadratic phase coupling," Sixth SSAP Workshop on Statistical Signal and Array Processing, pp.189–192, Victoria, British Columbia, Canada, October 1992.

[93] D. Pearson, S. Pillai, Y. Lee, "An algorithm for near optimal placement of sensor elements," *IEEE Trans. Information Theory*, vol.36, no.6, pp.1280–1284, November 1990.

[94] S. Pei, C.C. Yeh and S.C. Chiu, "Modified spatial smoothing for coherent jammer suppression without signal cancellation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.36, no.3, pp.412–414, March 1988.

[95] S.U. Pillai, Y. Bar-Ness and F. Haber, "A new approach to array geometry for improved spatial spectrum estimation," *Proceedings of IEEE*, vol.73, no.10, pp.1522–1524, October 1985.

[96] U. Pillai and B. Kwon, "Forward/backward spatial smoothing schemes for coherent signal identification," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.37, no.1, pp.8–15, January 1989.

[97] S.U. Pillai. *Array Signal Processing.* Springer-Verlag, New-York, 1989.

[98] V.F. Pisarenko, "The retrieval of harmonics from a covariance function," *Geophysical Journal of Royal Astronomical Society*, vol.33, no.3, pp.347–366, September 1973.

[99] B. Porat and B. Friedlander, "Direction finding algorithms based on high-order statistics," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.39, no.9, pp.2016–2024, September 1991.

[100] S. Prasad, R.T. Williams, A.K. Mahalanabis and L.H. Sibul, "A transform-based covariance differencing approach for some classes of parameter estimation problems," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.36, no.5, pp.631–641, May 1988.

[101] B.D. Rao and K.V.S. Hari, "Performance analysis of root-MUSIC," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.37, no.12, pp.1939–1949, December 1989.

[102] L. Rabiner and R. Schafer. *Digital Processing of Speech Signals*, Prentice-Hall, New-Jersey, 1978.

[103] S.S. Reddi, "Multiple source location—a digital approach," *IEEE Trans. Aerospace and Electronic Systems*, vol.15, no.1, pp.95–105, January 1979.

[104] V. Reddy, A. Paulraj and T. Kailath, "Performance analysis of the optimum beamformer in the presence of correlated sources and its behavior under spatial smoothing," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.ASSP-35, no.7, pp.927–936, July 1987.

[105] Y. Rockah and P.M. Schultheiss, "Array shape calibration using sources in unknown locations—Part I: Far-field sources," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.35, no.3, pp.286–299, March 1987.

[106] Y. Rockah and P.M. Schultheiss, "Array shape calibration using sources in unknown locations—Part II: Near-field sources and estimator implementation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.35, no.6, pp.724–735, March 1987.

[107] M. Rosenblatt. *Stationary Sequences and Random Fields.* Birkhäuser Boston Inc., Boston, 1985.

[108] R. Roy and T. Kailath, "ESPRIT—Estimation of signal parameters via rotational invariance techniques," *Optical Engineering*, vol.29, no.4, pp.296–313, April 1990.

[109] C.S. Ruf, "Numerical annealing of low-redundancy linear arrays," *IEEE Trans. Antennas and Propagation*, vol.41, no.1, pp.85–90, January 1993.

[110] P. Ruiz and J.L. Lacoume, "Extraction of independent sources from correlated sources: a solution based on cumulants," *Proc. Vail Workshop on Higher-Order Spectral Analysis*, pp.146–151, June 1989.

[111] M. Rytle, B. Elsmore and A.C. Neville, "Observations of radio galaxies with the one mile telescope at Cambridge," *Nature*, vol.207, pp.1024–1027, September 1965.

[112] B. Sadler, "Shift and rotation invariant object reconstruction using the bispectrum," *Proc. Vail Workshop on Higher-Order Spectral Analysis*, pp.106–111, June 1989.

[113] 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.12, no.1, pp.80–89, January 1987.

[114] K. Sasaki, T. Sato and Y. Nakamura, "Holographic passive sonar," *IEEE Trans. Sonics Ultrasonics*, vol.24, no.3, pp.193–200, May 1977.

[115] T. Sato and K. Sasaki, "Bispectral Holography," *Journal of Acoustical Society of America*, vol.62, no.2, pp.404–408, August 1977.

[116] S.V. Schell and W.A. Gardner, "Progress on signal selective direction finding," *Fifth ASSP Workshop on Spectrum Estimation and Modelling*, pp.144–148, October 1990.

[117] D.C. Schleher. *MTI and Pulsed Doppler Radar.* Artech House, Boston, 1991.

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

[119] R.O. Schmidt and R.E. Franks, "Multiple source DF signal processing: an experimental system," *IEEE Trans. Antennas and Propagation*, vol.34, no.3, pp.281–290, March 1986.

[120] R.O. Schmidt, "Multilinear array manifold interpolation," *IEEE Transactions on Signal Processing*, vol.40, no.4, pp.857–866, April 1992.

[121] S. Shamsunder and G. Giannakis, "Non-Gaussian source localization via exploitation of higher-order cyclostationarity," *Sixth Signal Processing Workshop on Statistical Signal and Array Processing*, pp.193–196, Victoria, Canada, October 1992.

[122] S. Shamsunder and G. Giannakis, "Modeling of non-Gaussian array data using cumulants," *Signal Processing*, vol.30, no.3, pp.279–297, February 1993.

[123] T. Shan and T. Kailath, "Adaptive beamforming for coherent signals and interference," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.33, no.3, pp.527–536, June 1985.

[124] T. Shan, M. Wax and T. Kailath, "On spatial smoothing for direction-of-arrival estimation of coherent signals," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.33, no.4, pp.806–811, August 1985.

[125] P.M. Schultheiss and E. Ashok, "Localization with arrays subject to sensor motion," *Proc-ICASSP-83*, vol.1, pp.371–374, May 1983.

[126] D. Slepian, "On bandwidth," *Proceedings of IEEE*, vol.64, no.3, pp. 292–300, March 1976.

[127] Special Issue on Adaptive Arrays, *IEE Proceedings*, vol.130, *Pts. F and H*, no.1, February 1983.

[128] Special Issue on Adaptive Antenna Systems, *IEEE Trans. Antennas Propagation*, vol.34, no.3, March 1986.

[129] D. Starer and A. Nehorai, "Path following algorithm for passive localization of near-field sources," Report # 9008, Center of Systems Science, Electrical Engineering, Yale University, New Haven, Connecticut, 1990.

[130] B.D. Steinberg, *Principles of Aperture and Array System Design*, John Wiley & Sons, Inc., New-York, 1976.

[131] P. Stoica, R.L. Moses, B. Friedlander, and T. Soderstrom, "Maximum-likelihood estimation of the parameters of multiple sinusoids from noisy measurements," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.37, no.3, pp.378–392, March 1989.

[132] P. Stoica and A. Nehorai, "MUSIC, maximum-likelihood and Cramer-Rao bound," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.37, no.5, pp.720–741, May 1989.

[133] P. Stoica and A. Nehorai, "MUSIC, maximum-likelihood and Cramer-Rao bound: further results and comparisons," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.38, no.12, pp.2140–2150, December 1990.

[134] P. Stoica and T. Söderström, "Statistical analysis of MUSIC and subspace rotation estimates of sinusoidal frequencies," *IEEE Trans. Signal Processing*, vol.39, no.8, pp.1836–1847, August 1991.

[135] G. Strang. *Linear Algebra and Its Applications*, Second Edition. Academic Press, New-York, 1980.

[136] A. Swami and J.M. Mendel, "Adaptive system identification using cumulants," *in Proc. IEEE Intl. Conf. Acoust., Speech, Signal Processing*, pp.2248–2251, 1988.

[137] A. Swami and J.M. Mendel, "Cumulant-based approach to harmonic retrieval and related problems," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.39, no.5, pp.1099–1109, May 1991.

[138] A.L. Swindlehurst and T. Kailath, "Near field source parameter estimation using a spatial Wigner distribution approach," *Proc. 4th ASSP Workshop on Spectrum Estimation and Modeling*, pp.123–128, Minneapolis, August 1988.

[139] A.L. Swindlehurst and T. Kailath, "Detection and estimation using the third-moment matrix," *Proc. ICASSP-89*, vol.4, pp.2325–2328, 1989.

[140] A.L. Swindlehurst, B. Ottersten, R. Roy, and T. Kailath, "Multiple invariance ESPRIT," *IEEE Trans. Signal Processing*, vol.40, no.4, pp.867–881, April 1992.

[141] A.L. Swindlehurst and T. Kailath, "A performance analysis of subspace-based methods in the presence of model errors, part 1: the MUSIC algorithm," *IEEE Trans. Signal Processing*, vol.40, no.7, pp.1758–1774, July 1992.

[142] A.L. Swindlehurst, "DOA identifiability for rotationally invariant arrays," *IEEE Trans. Signal Processing* , vol.40, no.7, pp.1825–1828, July 1992.

[143] A.L. Swindlehurst and T. Kailath, "Azimuth/elevation direction finding using regular array geometries," *IEEE Trans. on Aerospace and Electronic Systems*, vol.AES-29, no.1, pp.145–156, January 1993.

[144] Y. Su, T. Shan and B. Widrow, "Parallel spatial processing: a cure for signal cancellation in adaptive arrays," *IEEE Trans. Antennas and Propagation*, vol.34, no.3, pp.347–355, March 1986.

[145] R.J. Talham, "Noise correlation functions for unisotropic noise fields," *J. Acoust. Soc. Amer.*, vol.69, no.1, pp.213–215, January 1981.

[146] H. Taylor and S.W. Golomb, *Rulers Part 1.* CSI Technical Report 85-05-01, University of Southern California, USA, 1985.

[147] C.W. Therrien. *Discrete Random Signals and Statistical Signal Processing.* Prentice-Hall, New Jersey, 1992.

[148] L. Tong, Y. Inouye and R. Liu, "Waveform preserving blind estimation of multiple independent sources," *IEEE Trans. Signal Processing*, vol.41, no.7, pp.2461–2470, July 1993.

[149] H.L. Van Trees. *Detection, Estimation, and Modulation Theory.* Volume 1. Wiley, New-York, 1968.

[150] M.K. Tsatsanis and G.B. Giannakis, "Translation, rotation and scaling invariant object and texture classification using polyspectra," *Proc. SPIE Int. Soc. Opt. Eng.*, vol.1348, pp.103–115, July 1990.

[151] D.W. Tufts and R. Kumaresan, "Estimation of frequencies of multiple sinusoids: making linear prediction perform like maximum-likelihood," *Proceedings of IEEE*, vol.70, no.9, pp.975–989, September 1982.

[152] J.K. Tugnait, "Time-delay estimation with unknown correlated Gaussian noise using fourth-order cumulants and cross-cumulants," *IEEE Trans. Signal Processing*, vol.39, no.6, pp.1258–1267, June 1991.

[153] B. Van Veen and K. Buckley, "Beamforming: a versatile approach to spatial filtering," *IEEE ASSP Magazine*, vol.5, no.2, pp.4–24, April 1988.

[154] E. Vertatschitsch and S. Haykin, "Impact of linear array geometry on direction of arrival estimation in a multipath environment," *IEEE Trans. Antennas and Propagation*, vol.39, no.5, pp.576-584, May 1991.

[155] G. Vezzosi, "Estimation of phase angles from the cross-spectral matrix," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.34, no.3, pp.405–422, June 1986.

[156] M. Viberg, "Sensor array processing using gated signals," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.37, no.3, pp.447–450, March 1989.

[157] M. Viberg and B. Ottersten, "Sensor array processing based on subspace fitting," *IEEE Transactions on Signal Processing*, vol.39, no.5, pp.1110–1121, May 1991.

[158] M. Viberg and B. Ottersten, and T. Kailath, "Detection and estimation in sensor arrays using weighted subspace fitting," *IEEE Trans. Signal Processing*, vol.39, no.11, pp.2436–2449, November 1991.

[159] H. Wang and M. Kaveh, "Coherent signal-subspace processing for the detection and estimation of angles of arrival of multiple wideband sources," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.33, no.4, pp.823–831, August 1985.

[160] M. Wax and T. Kailath, "Optimal localization of multiple sources by passive arrays," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.31, no.5, pp.1210–1218, October 1983.

[161] M. Wax, T.J. Shan, and T. Kailath, "Spatio-temporal spectral analysis by eigenstructure methods," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.32, no.4, pp.817–827, August 1984.

[162] M. Wax and I. Ziskind, "On unique localization of multiple sources by passive sensor arrays," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.37, no.7, pp.996–1000, July 1989.

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

[164] M. Wax, "Detection and localization of multiple sources via the stochastic signals model," *IEEE Trans. Signal Processing*, vol.39, no.11, pp.2450–2456, November 1991.

[165] A.J. Weiss, A.S. Willsky and B.C. Levy, "Eigenstructure approach for array processing with unknown intensity coefficients," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.36, no.10, pp.1613–1617, October 1988.

[166] A.J. Weiss and B. Friedlander, "Array shape calibration using sources in unknown locations–a maximum-likelihood approach," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.37, no.12, pp.1958–1966, December 1989.

[167] B. Wells, "Voiced/unvoiced decision based on the bispectrum," *in Proc. IEEE Intl. Conf. Acoust., Speech, Signal Processing*, pp. 1589–1592, 1985.

[168] B. Widrow, P.E. Mantey, L.J. Griffits, and B.B. Goode, "Adaptive antenna systems," *Proceedings of IEEE*, vol.55, no.12, pp.2143–2139, December 1967.

[169] B. Widrow, J.R. Glover, J.M. McCool, J. Kaunitz, C.S. Williams, R.H. Hearn, J.R. Zeidler, E.D. Dong, and R.C. Goodlin, "Adaptive noise cancelling: principles and applications," *Proceedings of IEEE*, vol.63, no.12, pp.1692–1716, December 1975.

[170] B. Widrow, K.M. Duvall, R.P. Gooch and W.C. Newman, "Signal cancellation phenomena in adaptive antennas: causes and cures," *IEEE Trans. Antennas and Propagation*, vol.30, no.3, pp.469–478, May 1982.

[171] R. Williams, S. Prasad, A.K. Mahalanabis and L.H. Sibul, "An improved spatial smoothing technique for bearing estimation in a multipath environment," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.36, no.4, pp.425–432, April 1988.

[172] Q. Wu, K.M. Wong, and J.P. Reilly, "Maximum-likelihood estimation for array processing in unknown noise environments," *in Proc. IEEE Intl. Conf. Acoust., Speech, Signal Processing*, pp.241–244, vol.5, San Fransisco, March 1992.

[173] G. Xu and T. Kailath, "Direction-of-arrival estimation via exploitation of cyclostationarity–a combination of temporal and spatial processing," *IEEE Trans. Signal Processing*, vol.40, no.7, pp.1775–1786, July 1992.

[174] X.L. Xu and K.M. Buckley, "Bias analysis of the MUSIC location estimator," *IEEE Trans. Signal Processing*, vol.40, no.10, pp.2559–2569, October 1992.

[175] K.B. Yu, "Recursive updating the eigenvalue decomposition of a covariance matrix," *IEEE Transactions on Signal Processing*, vol.39, no.5, pp.1136–1145, May 1991.

[176] C.L. Zahm, "Effects of errors in the direction of incidence on the performance of an adaptive array," *Proceedings of IEEE*, vol.60, no.8, pp.1008–1009, August 1972.

[177] C.L. Zahm, "Application of adaptive arrays to suppress strong jammers in the presence of weak signals," *IEEE Trans. Aerospace and Electronic Systems*, vol.9, no.2, pp.260–271, March 1973.

[178] W. Zhang and M. Raghuveer, "Nonparametric bispectrum-based time-delay estimators for multiple sensor data," *IEEE Trans. Signal Processing*, vol.39, no.3, pp.770–774, March 1991.

[179] L.C. Zhao, P.R. Krishnaiah and Z.D. Bai, "On detection of the number of signals in the presence of white noise," *Journal of Multivariate Analysis*, vol.20, no.1, pp.1–25, October 1986.

[180] I. Ziskind and M. Wax, "Maximum likelihood localization of multiple sources by alternating projection," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.36, no.10, pp. 1553–1560, October 1988.

[181] M.D. Zoltowski, "On the performance analysis of the MVDR beamformer in the presence of correlated interference," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.36, no.6, pp.945–947, June 1988.

[182] M.D. Zoltowski, G.M. Kautz, and S.D. Silverstein, "Beamspace root-MUSIC," *IEEE Trans. Signal Processing*, vol.41, no.1, pp.344–364, January 1993.

[183] I.G. Žurbenko. *The Spectral Analysis of Time Series.* Elsevier Science Publishers, Amsterdam, Netherlands, 1986.