Applications of Cumulants to Array Processing Part II: Non-Gaussian Noise Suppression

Mithat C. Doğan, *Member, IEEE*, and Jerry M. Mendel, *Fellow, IEEE*

***Abstract*—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. In this paper, we investigate the effects of non-Gaussian ambient noise on cumulant-based direction-finding systems using the interpretation for the information provided by cumulants for array processing applications described in the companion paper [4]. We first demonstrate the suppression of uncorrelated non-Gaussian noise that has spatially varying statistics. Then, we indicate methods to suppress spatially colored non-Gaussian noise using cumulants and an additional sensor whose measurement noise component is independent of the noise components of the original array measurements. We also indicate the noise suppression properties of the virtual-ESPRIT algorithm proposed in [4]. In addition, we propose a method that combines second- and fourth-order statistics together in order to suppress spatially colored non-Gaussian noise. Finally, we also illustrate how to suppress spatially colored non-Gaussian noise when the additional sensor measurement is not available. Simulations are presented to verify our results.**

I. Introduction

OVER the last decade, many algorithms have been developed that utilize higher order statistics, which address only Gaussian noise suppression, and, for the most part, repeat the work accomplished when using second-order statistics [7], [8]. A major objection to these algorithms is their sensitivity to additive non-Gaussian noise. In this paper, we address the effects of non-Gaussian noise on high-resolution direction-finding problems. We establish methods to suppress non-Gaussian noise by adding just one new sensor to an array. Our algorithms depend on a new interpretation of cumulants in array processing problems proposed in the companion paper [4]: one that focuses on the information embedded in the multiple arguments of cumulants.

In [4], we proposed a cumulant-based approach to increase the effective aperture of antenna arrays by substituting cumulants for the cross-correlations between actual and so-called virtual sensors to form a virtual covariance matrix that resembles the array covariance matrix, as if the measurements from the virtual sensors were available. We call this approach a virtual cross-correlation computer ($\mathbf{VC}^3$). We demonstrate in Section II that this approach can also be used to suppress additive uncorrelated non-Gaussian noise, whose statistics vary from sensor to sensor.

The $\mathbf{VC}^3$ can be used to calibrate arbitrary antenna arrays using only a sensor doublet [4]. The calibration problem is solved by computing the cross-correlations between the actual array and its virtual copy, using the cumulants of measurements and applying the ESPRIT algorithm [10]. We refer to this result as a virtual-ESPRIT algorithm (VESPA) [4]. Extensive simulations are provided in [4]; they compare the performances of VESPA and ESPRIT in direction-finding and signal recovery applications.

In array processing, it is commonly assumed that the measurements are corrupted by additive Gaussian noise that 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 array covariance matrix. If the ambient noise is spatially colored but its covariance matrix is known to within a scale factor, prewhitening can then 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-based methods. The higher order statistics based methods mentioned above assume the hypothetical case, when there is Gaussian noise present in the measurements.

In this paper, we investigate the possibility of combating the effects of *non-Gaussian noise* using cumulants. Using the geometric interpretation of cumulants explained in [4], we describe, in Section II, a way to suppress spatially independent non-Gaussian sensor noise whose statistics vary from sensor to sensor. The conditions necessary to suppress the effects of noise in more general situations are described in Section III, where it is shown that by using an additional sensor whose noise component is independent from the noise components of the measurements from the main array, it is possible to suppress non-Gaussian noise. Noise suppression properties of the virtual-ESPRIT algorithm are described in Section IV. Finally, we propose a hybrid method in Section V that combines second- and fourth-order statistics to achieve even better results. We also propose a way to suppress spatially

Manuscript received May 28, 1993; revised December 20, 1994. The associate editor coordinating the review of this paper and approving it for publication was Prof. John Goutsias.

M. C. Doğan was with the Signal and Image Processing Institute, Department of Electrical Engineering-Systems, University of Southern California, Los Angeles, CA 90089-2564. He is now with the TRW Avionics and Surveillance Group, Sunnyvale, CA 94088-3510 USA.

J. M. Mendel is with the Signal and Image Processing Institute, Department of Electrical Engineering-Systems, University of Southern California, Los Angeles, CA 90089-2564 USA.

IEEE Log Number 9411985.

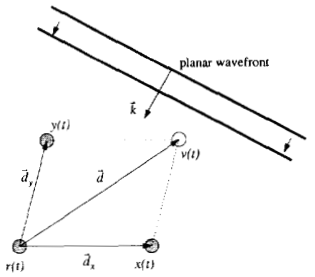


Fig. 1. Sensors that measure $r(t), x(t), y(t)$ are actual sensors, whereas $v(t)$ is a virtual-process measured by a virtual sensor.

colored non-Gaussian noise when an additional sensor measurement is not available. We demonstrate our theoretical results by simulations in Section VI. Final observations and future directions are given in Section VII.

II. Non-Gaussian Noise Suppression (Uncorrelated Noises)

Theorem 1: Consider an array of isotropic sensors that 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, then it is possible to do the following:

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.
2) Extend the aperture regardless of the sensor noise.

These can happen provided that the number of sources is less than the number of effective sensors.

Proof: Since the far-field sources are assumed to be independent, we can consider the presence of a single source without loss of generality. Consider Fig. 1, which illustrates an array of three sensors. 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, e.g., 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.

$$\frac{\sigma_s^2}{\gamma_{4,s}} \text{cum}(r^*(t), x(t), x^*(t), r(t)) = E\{s^*(t)s(t)\}. \quad (1)$$

To derive (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

$$\begin{aligned}&\text{cum}(r^*(t), x(t), x^*(t), r(t))\\ &= \underbrace{\text{cum}(s^*(t), s(t)e^{-j\vec{k}\cdot\vec{d}_x}, s^*(t)e^{j\vec{k}\cdot\vec{d}_x}, s(t))}_{\gamma_{4,s}}\\ &\quad + \text{cum}(n_r^*(t), n_x(t), n_x^*(t), n_r(t)). \quad (2)\end{aligned}$$

Since the noise components are independent of each other, the second term $\text{cum}(n_r^*(t), n_x(t), n_x^*(t), n_r(t))$ is equal to zero [7]. Scaling (2) by $\sigma_s^2/\gamma_{4,s}$ gives the equality in (1) because $E\{s^*(t)s(t)\} = \sigma_s^2$. The result in (1) is valid regardless of additive noise characteristics: It can be Gaussian or non-Gaussian.

In the hypothetical case, where there is no noise present, the second-order statistics of the measurements will be equal to those of the signal components only. Therefore, we can write

$$E\{r^*(t)r(t)\}\big|_{\text{no noise}} = E\{s^*(t)s(t)\} \quad (3)$$

Here, we emphasized the assumption of no noise condition for the equality. Since the right sides of (1) and (3) are identical, their left sides must be equal as well:

$$\begin{aligned}&\frac{\sigma_s^2}{\gamma_{4,s}} \text{cum}(r^*(t), x(t), x^*(t), r(t))\big|_{\text{with non-Gaussian noise}}\\ &= E\{r^*(t)r(t)\}\big|_{\text{no noise}}. \quad (4)\end{aligned}$$

The left-hand side of (4) (to within the scale factor $\sigma_s^2/\gamma_{4,s}$) is computed in the actual scenario where additive non-Gaussian noise can be present. We emphasized this in the equation because the right-hand side of (4) can only be computed in the hypothetical case where there is no measurement noise. 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.

The scale factor $\sigma_s^2/\gamma_{4,s}$ does not cause a problem because if all the required covariances are computed through cumulants, then the resulting covariance matrix will correspond to the case in which source powers are scaled by these unknown (but nonzero) factors, which preserves the signal subspace. We refer the reader to [4] for further discussion on scaling. The convention established in (4) will be used in the paper to indicate non-Gaussian noise suppression. For example (with proof similar to that of (4))

$$\begin{aligned}&\text{cum}(r^*(t), x(t), r^*(t), r(t))\big|_{\text{with non-Gaussian noise}}\\ &= \frac{\gamma_{4,s}}{\sigma_s^2} E\{r^*(t)x(t)\}\big|_{\text{no noise}}. \quad (5)\end{aligned}$$

In the case of P non-Gaussian sources with cumulants $\{\gamma_{4,k}\}_{k=1}^P$ and powers $\{\sigma_k^2\}_{k=1}^P$, if one constructs a matrix of covariances that are computed by using cumulants, through relations such as (4) and (5), and ignores 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 (σ_k^2's) are scaled by β_k's. This matrix takes the form $\mathbf{A}\mathbf{\Gamma}_{\mathbf{ss}}\mathbf{A}^H$, where $\mathbf{A}$ is the full-rank steering matrix, and $\mathbf{\Gamma}_{\mathbf{ss}}$ is a diagonal matrix that contains the fourth-order cumulants of sources and 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; therefore, the signal subspace can be identified as the eigenvectors of this cumulant matrix that have nonzero (but perhaps negative since scale factors may be negative) eigenvalues. Of course, the number of sources must be less than the number of (effective) sensors for any subspace method (e.g., MUSIC) to correctly determine the noise subspace and estimate the source bearings. This proves part 1 of Theorem 1.

Virtual aperture extension is the term coined in [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 Fig. 1, where $\vec{d} = \vec{d}_x + \vec{d}_y$). From our interpretation, this can be viewed as reaching a virtual location by adding two nonzero vectors (otherwise, we cannot reach a virtual location) that extend between actual array elements. A nonzero 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 from the other components; for example (see Fig. 1)

$$\begin{aligned}&\text{cum}(r^*(t), x(t), r^*(t), y(t))\big|_{\text{with non-Gaussian noise}}\\ &= \frac{\gamma_{4,s}}{\sigma_s^2} E\{r^*(t)v(t)\}\big|_{\text{no noise.}}\end{aligned} \tag{6}$$

The derivation of (6) can be done by observing similar results in [4]. $E\{r^*(t)v(t)\}$ is not computable since we do not have access to $v(t)$ (a virtual sensor); however, we have $r(t)$,$x(t)$, and $y(t)$, and the noise components in these channels are independent; hence, there is equality in (6). We have therefore shown that $E\{r^*(t)v(t)\}$ (virtual statistic) can be computed using cumulants by processing the measured signals $r(t)$,$x(t)$, and $y(t)$, even in the presence of non-Gaussian noise. Of course, the number of sources must be less than the number of (effective) sensors for any subspace method to correctly determine the noise subspace and estimate the source bearings. This proves the second part of Theorem 1. □

Comments:

- The convention established in (4) will be used throughout this paper. It is important to note that (4) is valid only for ensemble averages. With finite samples, the standard deviations of the two sides will be different. In addition, there may be a bias due to finite sample size.
- The geometric interpretation of (4) based on the results in [4] is as follows: Since fourth-order cumulants can be considered to be addition of two vectors extending from conjugated channels to unconjugated ones, for non-Gaussian noise suppression purposes, using the statistic $\text{cum}(r^*(t), x(t), x^*(t), r(t))$, we *move* from $r(t)$ to $x(t)$ (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 [1] for accomplishing non-Gaussian noise insensitivity by removing the diagonal elements of quadricovariance steering matrices.
- The primary limitation of the proposed approach comes from the assumption about the sensor-to-sensor independence of the non-Gaussian noise. In addition, we used the far-field assumption and independence of sources.

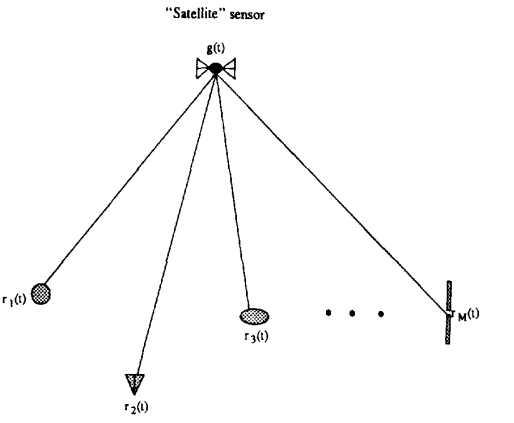


Fig. 2. Non-Gaussian noise suppression for correlated noises in the main array that measures $\{r_k(t)\}_{k=1}^M$. An arbitrary array of M sensors whose measurements are corrupted by colored non-Gaussian noise can still be used for direction finding if the noise in $g(t)$ is independent of the noise present in the rest of the array elements. Such a unit can be imposed in the field to correct the performance of existing systems that suffer from colored noise.

III. Non-Gaussian Noise Suppression (Correlated Noises)

Theorem 2: Consider an array of arbitrary sensors that is illuminated by linearly correlated non-Gaussian sources whose number is less than the number of sensors. 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 that 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 that the sources are independent. Later, we consider linearly correlated sources.

Consider Fig. 2, in which there is an array of M sensors that measure $\{r_k(t)\}_{k=1}^M$, and there is another sensor that measures $g(t)$ where

$$g(t) = \mathbf{g}^T \mathbf{s}(t) + n_g(t) \tag{7}$$

in which, g is a P-vector, which contains the response of the sensor that measures $g(t)$ to the far-field source wavefronts. We define g as the sensor response vector. The main array measurements take the form

$$\mathbf{r}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}_\mathbf{x}(t). \tag{8}$$

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

$$\begin{aligned}&\text{cum}(r_j^*(t), g(t), g^*(t), r_k(t))\big|_{\text{with non-Gaussian noise}}\\ &= \frac{\gamma_{4,s}|g_s|^2}{\sigma_s^2} E\{r_j^*(t) r_k(t)\}\big|_{\text{no noise}}\end{aligned} \tag{9}$$

where g_s is the response of the satellite sensor to the single source wavefront. The derivation of (9) is similar to that of (4).

In this way, the following noise-free array covariance matrix can be constructed using cumulants to replace the second-order statistics (using (9)), where we assume multiple independent sources and that superposition holds:

$$\mathbf{C} = \mathbf{A}\,\Gamma\mathbf{A}^H \tag{10}$$

where

- $\mathbf{A}$ steering matrix for the original M sensors (excluding the satellite sensor)
- Γ diagonal matrix whose kth diagonal entry is $\gamma_{4,k}|g_k|^2$
- g_k response of the satellite sensor to the kth source (the vector $\mathbf{g}$ is the collection of such responses).

Equation (10) follows from the fact that in the absence of noise the array covariance matrix takes the form $\mathbf{A}\,\mathbf{\Sigma_{ss}}\mathbf{A}^H$, where $\mathbf{\Sigma_{ss}}$ is a diagonal matrix that contains source powers. When cumulants are used to compute the cross-correlations, we obtain a covariance matrix for the case in which source powers are scaled by unknown constants, $\beta_k = \gamma_{4,k}|g_k|^2/\sigma_k^2$, which yields (10).

Any subspace method can be applied to $\mathbf{C}$ in (10), whose elements are computed using (9). Of course, the number of sources must be less than the number of sensors in the main array for any subspace method to correctly determine the noise subspace and estimate the source bearings. There is no need to know the response of the satellite sensor to the far-field sources (i.e., $\mathbf{g}$ in (7)); however, the elements of $\mathbf{g}$ must be nonzero in order to make Γ a nonsingular matrix. We just need the time series recorded by the satellite sensor to actually compute the left-hand side of (9).

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

$$\mathbf{s}(t) = \mathbf{Q}\mathbf{u}(t) \tag{11}$$

where $\mathbf{Q}$ is nonsingular (but arbitrary), and components of $\mathbf{u}(t)$ are independent. Then, the observation (7) and (8) change to

$$\begin{aligned} g(t) &= \mathbf{h}^T\mathbf{u}(t) + n_g(t) \\ \mathbf{r}(t) &= \mathbf{B}\mathbf{u}(t) + \mathbf{n_x}(t) \end{aligned} \tag{12}$$

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) and (8) with a generalized steering matrix $\mathbf{B}$ and a response vector $\mathbf{h}$. The cumulant matrix $\mathbf{C}$, which is computed as described in (9) for the independent sources scenario, now takes the form

$$\mathbf{C} = \mathbf{B}\,\tilde{\Gamma}\mathbf{B}^H \tag{13}$$

which was obtained by substituting $\mathbf{B}$ for $\mathbf{A}$ and $\mathbf{h}$ for $\mathbf{g}$ in (10). In (13), $\tilde{\Gamma}$ is defined as the diagonal matrix whose kth 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 $\mathbf{Q}\,\tilde{\Gamma}\,\mathbf{Q}^H$ is full rank so that $\mathbf{C}$, which is expressed as

$$\mathbf{C} = \mathbf{A}\,(\mathbf{Q}\,\tilde{\Gamma}\,\mathbf{Q}^H)\,\mathbf{A}^H \tag{14}$$

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 nonzero. This completes the proof of Theorem 2. □

Comments:

- This method can be interpreted as follows: Consider a totally different problem in which the sensors $\{r_k\}_{k=1}^M$ are viewed as mobile communication antennas that 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. From the results in [4], we know that communication between sensors means implementing cross-covariance. Here, we cannot 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 $E\{r_j^*(t)r_k(t)\}$, first move from $r_j(t)$ to the satellite sensor $g(t)$, and then, let the satellite distribute the message, i.e., move from $g(t)$ to $r_k(t)$.
- A similar technique was developed in [14] as a covariance-based approach; however, it requires a second array of sensors whose noise component is independent of the noise in the existing array. Consequently, [14] 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 (VESPA) [4] (which is described briefly in the next section) as compared with covariance ESPRIT. The reason for this difference is that cumulants, unlike cross-correlation, have an array of arguments. Based on this observation, we have also developed algorithms for single sensor detection/classification of multiple sources [5] and two-sensor multiple source nonparametric time-delay estimation [3].
- 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 [12] to the covariance matrix in (14) to estimate the parameters of *coherent sources* (i.e., when $\mathbf{Q}$ is singular). Simulations in Section VI investigate the coherent sources in non-Gaussian noise scenario. In addition, an approach that applies spatial smoothing to the generalized steering vectors estimated by VESPA is proposed in [6]. This approach can estimate more sources than sensors, even in the coherent sources case, and it can utilize sensors that are nonuniformly spaced with different responses.
- Virtual aperture extension using Theorem 2 is also possible: It requires fixing one of the four arguments of the cumulant to be $g(t)$. Consequently, this problem is similar to aperture extension using third-order cumulants since we now only have three free cumulant arguments with which to extend the aperture. This issue is described in detail in [3].

- It is possible to verify the independence of the noise contained in the satellite sensor's measurements from the noise components of the main array sensors by eigenanalysis of the cumulant matrix in (10). There are P (P is the number of far-field sources illuminating the array) signal subspace eigenvalues and $(M - P)$ null eigenvalues if the noise conditions are satisfied.
- In practice, it is possible to obtain the satellite sensor measurements without separating a sensor from the main array if one of the sensors is electromagnetically shielded against near-field noise of the main array, e.g., the main array and the satellite sensor are separated by a metallic wall on a platform. Another way is to preprocess the main array measurements to create a time series in which noise contribution is decreased but not totally eliminated and use this data as the satellite sensor measurement. The latter method is investigated in the simulations.

IV. Virtual-ESPRIT and Non-Gaussian Noise

The virtual-ESPRIT algorithm (VESPA) calibrates arbitrary arrays using a single doublet and fourth-order cumulants [4]. Consider an M element arbitrary array that measures $\{r_1(t), r_2(t), \ldots, r_M(t)\}$ (see Fig. 4 in [4]). Let the first two sensors be identical, and denote the vector between them by $\vec{\Delta}$. In addition, let $\{v_1(t), v_2(t), \ldots, v_M(t)\}$ denote the mesurements from a virtual array. The virtual array is a copy of the original array, which is displaced by $\vec{\Delta}$; therefore, from geometry, we have $v_1(t) = r_2(t)$ in the absence of noise. The actual sensors form the actual subarray, and the virtual sensors form the virtual subarray.

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. 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. Therefore, by using fourth-order cumulants, and assuming that only one doublet $\{r_1(t), r_2(t)\}$ is available, we can compute the cross-correlations between subarray elements, e.g.

$$\text{cum}(r_1^*(t), r_2(t), r_k^*(t), r_l(t)) = \frac{\gamma_{4,s}|a_1|^2}{\sigma_s^2} E\{r_k^*(t)v_l(t)\}. \tag{15}$$

Here, a_1 is defined as the response of the first sensor (that measures $r_1(t)$) to the signal wavefront. Use of fourth-order cumulants provides the necessary 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.

$$\text{cum}(r_1^*(t), r_1(t), r_k^*(t), r_l(t)) = \frac{\gamma_{4,s}|a_1|^2}{\sigma_s^2} E\{r_k^*(t)r_l(t)\} \tag{16}$$

In this way, (15) and (16) can be used to form the covariance matrix of ESPRIT [10]. For obvious reasons, we call the single pair of sensors that form the doublet "guiding sensors," and the resulting method "the virtual-ESPRIT algorithm." For additional details about VESPA, as well as simulation results, see [4]. Next, 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 Theorem 2 to VESPA (see Fig. 4 in [4]). Let us assume that the noise component of the second sensor measurement ($r_2(t)$) is independent of the rest of the sensors, and consider a single wavefront. Let a_s denote the response of this sensor to the single wavefront. We choose the second sensor as the "satellite" sensor, i.e., $g(t) = r_2(t)$. Since the responses of the first two sensors are identical, the response of the first sensor to the wavefront is also a_s.

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

$$\begin{aligned}&\text{cum}(r_k^*(t), g(t), g^*(t), r_l(t))\big|_{\text{non-Gaussian noise}} \\ &= \frac{\gamma_{4,s}|a_s|^2}{\sigma_s^2} E\{r_k^*(t)r_l(t)\}\big|_{\text{no noise}}\end{aligned} \tag{17}$$

where $1 \le k, l \le M$. This is in fact the idea presented for Theorem 2: Using the cumulant expression in (17), we add two vectors: one that extends from $r_k(t)$ to $g(t)$ another one that extends from $g(t)$ to $r_l(t)$. This vector addition results in the vector from $r_k(t)$ to $r_l(t)$, which is equivalent to the right side of (17). A slight modification of this idea can be used to compute the cross correlations between actual and virtual sensors as

$$\begin{aligned}&\text{cum}(r_1^*(t), g(t), r_k^*(t), r_l(t))\big|_{\text{non-Gaussian noise}} \\ &= \frac{\gamma_{4,s}|a_s|^2}{\sigma_s^2} E\{r_k^*(t)v_l(t)\}\big|_{\text{no noise}}\end{aligned} \tag{18}$$

although $v_l(t)$ is not physically available. The cross correlations between virtual sensors are identical to those between actual sensors (e.g., $E\{v_k^*(t)v_l(t)\} = E\{r_k^*(t)r_l(t)\}$); therefore, (17) can also be used to compute cross correlations between virtual sensors.

For multiple independent sources, the superposition property of cumulants (**[CP3]** in [4]) holds, and (17) and (18) still apply. This completes the proof of Theorem 3. □

V. Combining 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 results from Theorem 2 can be improved by using second-order statistics. We also propose a cumulant-based method to generate a vector that can be used to estimate source bearings when an additional sensor is not available.

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

$$\mathbf{d} \triangleq E\{\mathbf{x}(t)g^*(t)\} \tag{19}$$

where $\mathbf{x}(t)$ denotes the measurements of the main array, and $g(t)$ is the measurement of the satellite sensor (see Fig. 2 and (7)). 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 (only when the 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 (11)), 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

$$\mathbf{d} = E\{\mathbf{x}(t)g^*(t)\} = E\{\mathbf{A}\,\mathbf{Q}\mathbf{u}(t)\mathbf{u}^H(t)\mathbf{Q}^H\mathbf{g}^*\}$$
$$= \mathbf{A}\,\mathbf{Q}\,\mathbf{\Sigma}_{uu}\,\mathbf{Q}^H\,\mathbf{g}^* \triangleq \mathbf{A}\mathbf{z} \tag{20}$$

where $\mathbf{\Sigma}_{uu} \triangleq E\{\mathbf{u}(t)\mathbf{u}^H(t)\}$. If none of the components of $\mathbf{z}$ are zero, then $\mathbf{d}$ is a superposition of steering vectors from the sources and leads to an algorithm that we describe next to estimate the directions of arrival.

With finite samples, we estimate the noise-free vector $\mathbf{d}$ as

$$\hat{\mathbf{d}}_N = \frac{1}{N}\sum_{t=1}^{N}\mathbf{x}(t)g^*(t). \tag{21}$$

If the received signal vectors are independent and identically distributed with finite moments up to order eight, then $\hat{\mathbf{d}}_N$ is an asymptotically normal sequence of random vectors [2], i.e.

$$\sqrt{N}\,(\hat{\mathbf{d}}_N - \mathbf{d}) \xrightarrow{\mathcal{L}} \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}) \tag{22}$$

where the $(m,n)th$ entry of $\mathbf{\Sigma}$ can be expressed as

$$\Sigma_{m,n} \triangleq \lim_{N\to+\infty} N(E\{(\hat{d}_N(m) - d(m))(\hat{d}_N(n) - d(n))^*\}$$
$$= E\{|g(t)|^2 r_m(t) r_n^*(t)\} - d(m)d^*(n)\,, \tag{23}$$

which is derived in the Appendix.

Since the estimation errors are asymptotically Gaussian, we can use the following cost function to estimate the parameters of interest [16]:

$$\hat{\theta} = \arg\min_{\theta,\mathbf{Z}} J(\theta) \triangleq \arg\min_{\theta,\mathbf{Z}} \|\,\mathbf{\Sigma}^{-1/2}\,(\hat{\mathbf{d}}_N - \mathbf{A}(\theta)\,\mathbf{z}\,)\,\|^2. \tag{24}$$

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

$$\hat{\mathbf{z}} = (\mathbf{A}^H(\hat{\theta})\mathbf{\Sigma}^{-1}\mathbf{A}(\hat{\theta}))^{-1}\mathbf{A}^H(\hat{\theta})\mathbf{\Sigma}^{-1}\hat{\mathbf{d}}_N. \tag{25}$$

Substituting (25) into (24), we obtain

$$\hat{\theta} = \arg\min_{\theta}\ \|\mathbf{\Sigma}^{-1/2}\,[\,\mathbf{I} - \mathbf{A}(\theta)$$
$$(\mathbf{A}^H(\theta)\mathbf{\Sigma}^{-1}\mathbf{A}(\theta))^{-1}\,\mathbf{A}^H(\theta)\mathbf{\Sigma}^{-1}\,]\,\hat{\mathbf{d}}_N\|^2. \tag{26}$$

Since the true covariance matrix $\mathbf{\Sigma}$ is not available, we may estimate it from data using fourth-order statistics of the received signals (see (23)). Alternatively, we may take a suboptimal approach by letting $\mathbf{\Sigma} = \mathbf{I}$, in which case, the contents of $\hat{\theta}$ are the least-squares estimates of source bearings.

The direction estimation from (26) 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 method described in Theorem 2 for initialization.* The minimization in (26) 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 [16]. We refer the reader to [16] 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 VI indicate that the SFS method can decrease the variance of the estimates from the cumulant-based approach, which would now only be used for initialization purposes.

If a satellite sensor is not available, cumulants can still be used to obtain a cumulant vector in which the contribution of noise is significantly reduced; then, this vector can be used to estimate source bearings. Specifically, consider the following simple scenario in which there are only two identically distributed and independent non-Gaussian processes: $\{u_1(t), u_2(t)\}$ with variance σ^2 and fourth-order cumulant γ_4, i.e.

$$\mathbf{r}(t) = \mathbf{a}_1 u_1(t) + \mathbf{a}_2 u_2(t) + \mathbf{n}(t) \tag{27}$$

where $\mathbf{n}(t)$ is Gaussian noise, and the second non-Gaussian component $u_2(t)$ is also treated as noise. Suppose that we preprocess the measurements with a weight-vector $\mathbf{w}$ to obtain $g(t) = \mathbf{w}^H\mathbf{r}(t)$. The selection of $\mathbf{w}$ depends on the *a priori* information about the source bearings, i.e., $\mathbf{w}$ is selected in a way to not only pass the signal $u_1(t)$ relatively undistorted but to also considerably suppress the second non-Gaussian component $u_2(t)$. Let the beamformer be $\mathbf{w}^H\mathbf{a}_1 = g_1$ ($|g_1| \simeq 1$) for the desired non-Gaussian source, and $\mathbf{w}^H\mathbf{a}_2 = g_1\beta$, where $|\beta| < 1$ for the undesired non-Gaussian source.[1] Then, the cumulant vector

$$(\mathbf{c})_k \triangleq \mathrm{cum}(g^*(t), g(t), g^*(t), r_k(t)) \quad 1 \le k \le M \tag{28}$$

can be expressed as

$$\mathbf{c} = \gamma_4|g_1|^2 g_1^*\mathbf{a}_1 + \gamma_4|g_1\beta|^2 g_1^*\beta^*\mathbf{a}_2. \tag{29}$$

This implies that if the second source is scaled by a factor β during the beamforming step (that uses $\mathbf{w}$) to form the satellite sensor measurement, then its contribution to $\mathbf{c}$ is scaled by $|\beta|^2\beta^*$; therefore, the ratio of contributions of the non-Gaussian noise to non-Gaussian desired source in the expression for $\mathbf{c}$ will be $|\beta|^2\beta^*$, which is small enough to ignore, provided that $|\beta|$ is small enough. If we can ignore the non-Gaussian noise component in $\mathbf{c}$ due to this scaling, we can use $\mathbf{c}$ in place of the second-order statistics vector $\mathbf{d}$ for DOA estimation. Of course, we can use the generated satellite signal $g(t) = \mathbf{w}^H\mathbf{r}(t)$ by the beamformer, but this will only provide a ratio of $|\beta|$ in terms of source contributions, which may be unacceptable for direction-finding applications. For the multiple sources/multiple non-Gaussian noises scenario,

[1] β is neither chosen nor known by the user in (29). It is only the relative gain of the beamformer to the undesired component. The goal here is to show that if β is smaller than unity, then it is possible to suppress the effect of an undesired non-Gaussian component. The beamformer weight vector can be determined by using *a priori* information about the expected sector for the source of interest.

the superposition property of cumulants ([**CP3**] in [4]) applies to yield a similar result.

VI. Simulations

In this section, we provide simulations that demonstrate the paper's non-Gaussian noise-insensitive direction-finding methods. Our first simulation experiment illustrates virtual aperture extension in the presence of spatially nonstationary but independent non-Gaussian sensor noise. Our second simulation compares cumulant and covariance-based algorithms in the presence of correlated non-Gaussian noise. 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.

Experiment 1: Virtual Aperture Extension in Non-Gaussian Noise

In this experiment, we consider a two-element linear array illuminated by two equal-power, independent non-Gaussian sources from 70 and 110° measured from the array axis (90° is the broadside). We assume the sensors are isotropic and separated by a half wavelength. The signal model is as follows:

$$\begin{bmatrix} r_1(t) \\ r_2(t) \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ e^{j\pi\cos(70^\circ)} & e^{j\pi\cos(110^\circ)} \end{bmatrix} \begin{bmatrix} s_1(t) \\ s_2(t) \end{bmatrix} \beta + \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} n_1(t) \\ n_2(t) \end{bmatrix}. \tag{30}$$

The statistically independent signal components $\{s_1(t), s_2(t)\}$ are zero-mean and non-Gaussian with unity variance. The noise components are generated as a mixture of non-Gaussian and Gaussian components as follows:

$$n_k(t) = \frac{1}{\sqrt{2}} \left(n_{k,1}(t) + n_{k,2}(t)\right) \quad k = 1, 2, \tag{31}$$

where $n_{k,1}(t)$ is circular Gaussian, and $n_{k,2}(t)$ is non-Gaussian and represents the contribution of 4-QAM communication signals. The noise components $\{n_{1,1}(t), n_{1,2}(t), n_{2,1}(t), n_{2,2}(t)\}$ are zero-mean, have unity variance, and are statistically independent; therefore, $n_1(t)$ and $n_2(t)$ are statistically independent as well. The SNR is defined as $20\log_{10}(\beta)$.

Using cumulants, it is possible to extend the aperture to three sensors. Since the noise components in the actual sensor measurements $\{n_1(t), n_2(t)\}$ are independent, it is possible to apply Theorem 1 to construct a 3×3 matrix that is blind to the presence of non-Gaussian noise as data length grows to infinity. It is therefore possible to estimate the autocorrelation at a sensor uniquely and the cross-correlation between the two actual sensors in two different ways using cumulants (see [4] for additional discussions on this issue). The cumulant statistics are placed in a 3×3 matrix that plays the role of the array covariance matrix, as if there were three actual sensors. The MUSIC algorithm was then applied to this 3×3 matrix to estimate the bearings of two sources. If we are constrained to use only second-order statistics, we cannot identify the bearings of the sources since the number of actual sensors (two) is not larger than the number of sources (two).

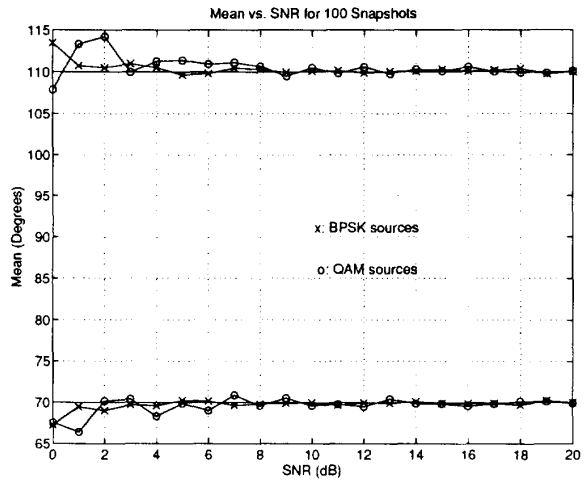


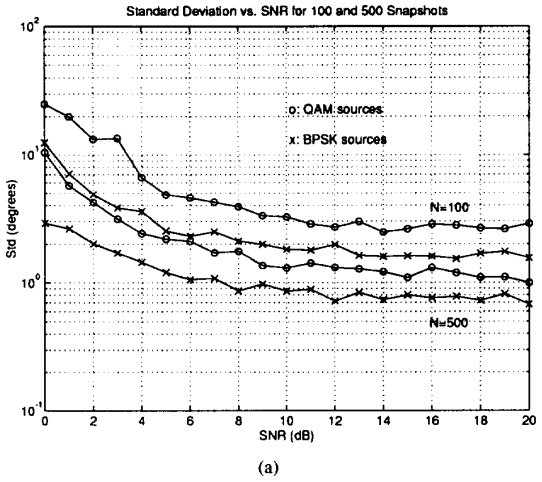
Fig. 3. Mean of estimates for BPSK and 4-QAM sources for 100 snapshots. True bearings are 70 and 110°. Each estimate is obtained from 100 independent realizations. Mean of estimates are satisfactory except in the low-SNR region.

We present the mean and standard deviation of the estimates for various data lengths and SNR levels for two different source distributions: For the first case, we let the sources be BPSK, and for the second case, we let the sources be 4-QAM. The first case has a greater cumulant-to-power ratio.[2] Hence, we expect the results for BPSK sources to be better that those for 4-QAM sources. Fig. 3 depicts the means obtained from 100 snapshots for BPSK and 4-QAM sources for varying SNR's. BPSK sources yielded slightly better results in terms of mean. Fig. 4 illustrates the standard deviation versus SNR for data lengths of 100, 500, 1000, and 2000 snapshots. Observe that the bearing estimates for BPSK sources have less estimation error than that of 4-QAM sources. At high SNR's, performance is limited by presence of cross terms in the 3×3 cumulant matrix between the independent waveforms in the cumulant expressions (these components should converge to zero in theory). At low-SNR's, performance is limited by the estimation errors due to the presence of high levels of noise that dominate the presence of cross terms.

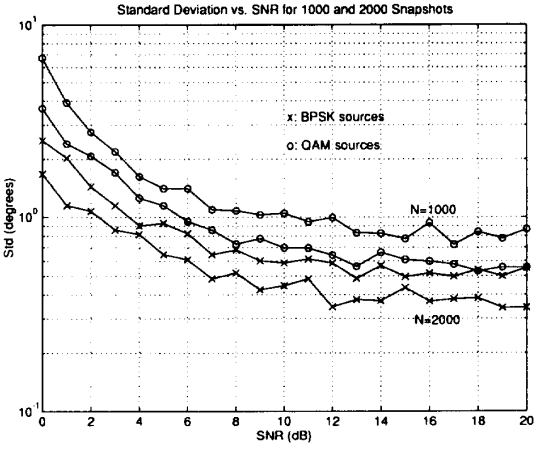
Experiment 2: Incoherent Sources in Non-Gaussian Noise

Here, we estimate the bearings of two far-field sources that illuminate a uniformly spaced linear array of five sensors from 85 and 95°. The bearings are measured from the axis of the linear array. Both sources broadcast BPSK waveforms

[2] A BPSK information sequence takes on values $\pm A$ (equally likely), and it has a variance of A^2 and a fourth-order cumulant $-2A^4$. A 4-QAM information sequence takes on values $\{Ae^{j\pi/4}, Ae^{j3\pi/4}, Ae^{-j3\pi/4}, Ae^{-j\pi/4}\}$ (equally likely). It has a variance of A^2 and a fourth-order cumulant $-A^4$.



(a)



(b)

Fig. 4. Standard deviation of estimates versus SNR for different data lengths and source distributions: (a) 100 and 500 snapshots; (b) 1000 and 2000 snapshots. Each estimate is obtained from 100 independent realizations. Standard deviations of bearing estimates for BPSK sources are always lower than those for 4-QAM sources. The results are given for the source from 110°.

of unity variance. The signal model for the experiment is as follows:

$$\begin{bmatrix} r_1(t) \\ r_2(t) \\ \vdots \\ r_5(t) \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ e^{j\pi\cos(85^\circ)} & e^{j\pi\cos(95^\circ)} \\ \vdots & \vdots \\ e^{j\pi 4\cos(85^\circ)} & e^{j\pi 4\cos(95^\circ)} \end{bmatrix} \begin{bmatrix} s_1(t) \\ s_2(t) \end{bmatrix} \beta + \mathbf{L} \begin{bmatrix} n_1(t) \\ n_2(t) \\ \vdots \\ n_5(t) \end{bmatrix}. \tag{32}$$

The signal components $\{s_1(t), s_2(t)\}$ are zero-mean, non-Gaussian with unity variance. The independent noise components $\{n_k(t)\}_{k=1}^{5}$ are generated as a mixture of circular Gaussian and 4-QAM non-Gaussian processes as in (31), and they have unity variance. The SNR is defined as $20\log_{10}(\beta)$.

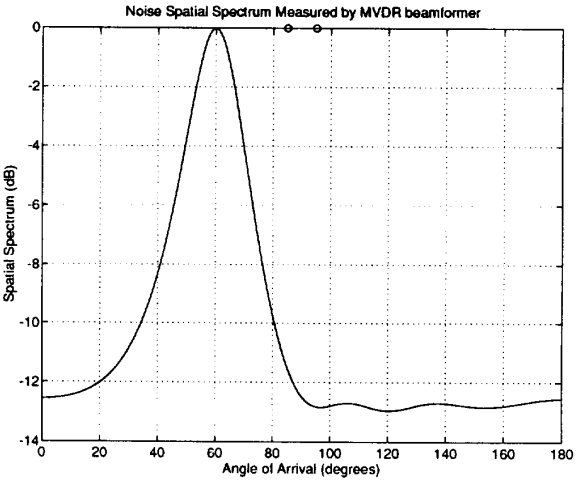


Fig. 5. Spatial spectrum of the noise as measured by an MVDR beamformer. The two circles indicate the source bearings.

The matrix $\mathbf{L}$ represents the spatial color of the noise in the linear array. It is obtained as the Cholesky decomposition of a noise covariance matrix that corresponds to major noise contribution from the range [50, 70°]. More specifically, we summed the rank-one matrices $\mathbf{a}(\theta)\mathbf{a}^H(\theta)$ in the range [50, 70°] at increments of 1° and scaled the result by α_1 so that it has a trace equal to 5 (the number of elements) to obtain $\alpha_1 \sum_{\theta=50^\circ}^{70^\circ} \mathbf{a}(\theta)\mathbf{a}^H(\theta)$, and then, to this, we added the identity matrix that represents the spatially white part of the measurement noise scaled by 0.2 so that the directional part is stronger in power than the white part. The resulting noise covariance matrix is

$$\mathbf{R_n} = \frac{1}{1.2}\left(\alpha_1 \sum_{\theta=50^\circ}^{70^\circ} \mathbf{a}(\theta)\mathbf{a}^H(\theta) + 0.2\mathbf{I} \right) \tag{33}$$

where the factor $1/1.2$ is included to make the trace of $\mathbf{R_n}$ be equal to 5. Then, we determined $\mathbf{L}$ so that $\mathbf{R_n} = \mathbf{L}\mathbf{L}^H$. The spatial spectrum of this measurement noise, as measured by an MVDR beamformer, is illustrated in Fig. 5, i.e., for each source bearing, we constructed an MVDR beamformer that passes the noise component from that particular direction undistorted and minimizes the contributions of noise from other directions.

In this experiment, we do not assume the availability of a satellite sensor. Instead, we process the measurements to obtain one such signal without an additional sensor. Using the *a priori* information that the sources are close to broadside (90°), we utilize the conventional beamformer with a look direction of 90° to obtain $g(t)$ as

$$g(t) = \mathbf{a}^H(90^\circ)\, \mathbf{r}(t). \tag{34}$$

Since the far-field sources are close to 90°, they pass through the conventional beamformer almost undistorted, whereas the noise, with its principal component in the range [50, 70°], is attenuated severely.

We construct the cumulant matrix

$$(\mathbf{C})_{m,n} = \text{cum}(g^*(t), g(t), r_m(t), r_n^*(t)) \tag{35}$$

which takes the form

$$\mathbf{C} = \sum_{k=1}^{2} \gamma_{4,k} |\mathbf{a}^H(90^o)\mathbf{a}(\theta_k)|^2 \mathbf{a}(\theta_k)\mathbf{a}^H(\theta_k) + \sum_{k=1}^{5} \gamma_{4,n_k} |\mathbf{a}^H(90^o)\mathbf{l}_k|^2 \mathbf{l}_k \mathbf{l}_k^H \tag{36}$$

where γ_{4,n_k} is the fourth-order cumulant of the noise component $n_k(t)$ in (32), and $\mathbf{l}_k$ is the kth column of $\mathbf{L}$. To derive (36), we first assume a single source so that $\mathbf{C} = \gamma_4 |g(\theta)|^2 \mathbf{a}(\theta)\mathbf{a}^H(\theta)$, where $g(\theta) \triangleq \mathbf{a}^H(90^o)\mathbf{a}(\theta_k)$. We consider the columns of $\mathbf{L}$ as steering vectors for the non-Gaussian noise sources and use the superposition property of cumulants to obtain the final result. Since the directional noise illuminates the array from directions that correspond to sidelobes of the conventional beamformer weight vector, we have $|\mathbf{a}^H(90^o)\mathbf{a}(\theta_k)|^2 >> |\mathbf{a}^H(90^o)\mathbf{l}_k|^2$, which implies noise suppression, i.e., we can ignore the second sum $\sum_{k=1}^{5}(\cdot)$ in (36). Algorithms such as MUSIC or ESPRIT can be used to process $\mathbf{C}$ to estimate source bearings.

We performed experiments to compare the proposed method that uses ESPRIT on the cumulant matrix $\mathbf{C}$ (defined in (35)) versus covariance-based ESPRIT that uses the array covariance matrix. We varied SNR (defined after (32)) from -10 to 20 dB in 1–dB steps and estimated the directions of arrival for data lengths of 500, 1000, 2000, and 5000 snapshots. For each experiment (with data length and SNR fixed), we performed 100 independent trials to estimate the directions of arrival. Since the bearing estimates from cumulant and covariance-based methods are expected to be biased due to the presence of colored non-Gaussian noise (the cumulant-based method does not use a true satellite sensor measurement whose noise component is statistically independent of the noise component in the main array measurements), we used the following performance criterion (RMSE) to compare the results:

$$\text{RMSE} \triangleq \sqrt{\frac{1}{N_e}\sum_{k=1}^{N_e}(\hat{\theta}_k - \theta_{\text{true}})^2}$$

in which N_e is the number of independent realizations, and $\hat{\theta}_k$ is the estimate provided from the kth realization.

The RMSE results are shown in Fig. 6 for the two sources. At low-SNR's, the performance of both methods are similar: The covariance ESPRIT is biased by the colored noise (see Fig. 7 for the mean of estimates), and the cumulant-based ESPRIT has high RMSE and bias due to high level of noise; however, the effects of colored noise decrease sharply for the cumulant-based approach, e.g., for 500 snapshots, the RMSE drops to an acceptable level of 1°, at an SNR of 0 dB, which is due still to estimation errors in cumulant estimates but not due to bias (see also Fig. 7). On the other hand, for the same data length, covariance ESPRIT achieves the same RMSE performance at an SNR of 11 dB. As the data length increases, the gap between cumulant and covariance-based results increases since the covariance-based ESPRIT is limited by the bias in covariance estimates (not the variance of estimates) and its performance does not vary with data length (see Fig. 6). For this reason, increasing data length does not effect the RMSE from covariance-based ESPRIT for SNR's below 10 dB since the algorithm is SNR limited in this range. For SNR's larger than 10 dB, bias gradually decreases, and estimation errors in finite-sample covariance estimates become visible (i.e., observe in Fig. 6 that the performance shifts down slightly with increasing data length); however, the performance of covariance-based ESPRIT is always worse than that of cumulant ESPRIT (even when the former uses 5000 snapshots and the latter uses 500 snapshots) for all SNR levels.

This experiment demonstrates that the proposed cumulant-based method can be applied to beamspace array processing

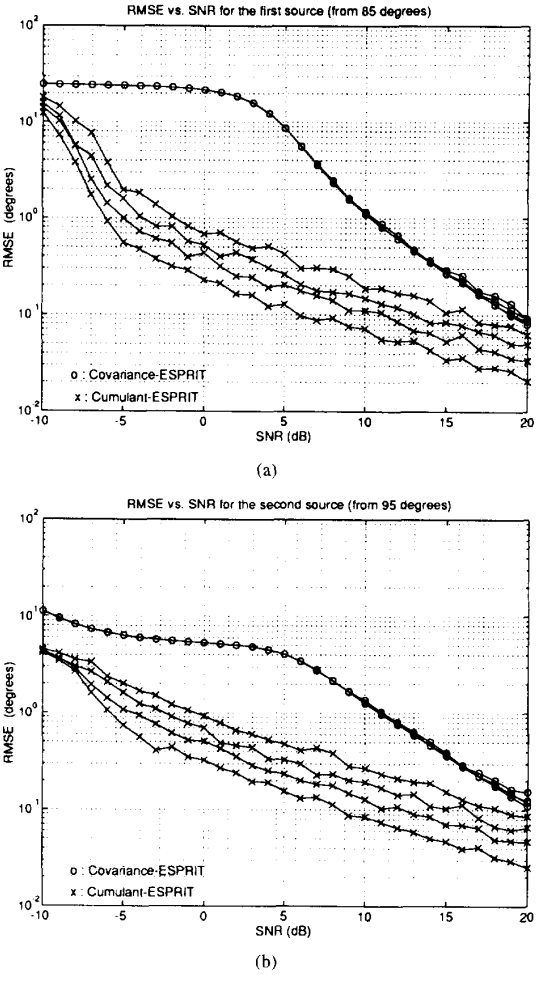


Fig. 6. Performance comparisons for covariance-ESPRIT and cumulant-based ESPRIT that uses a preprocessed sensor measurement for noise suppression: (a) RMSE for the first source; (b) RMSE for the second source. For each SNR level, we used four different data lengths of 500, 1000, 2000, and 5000 snapshots and performed each experiment 100 times to obtain the results. Performance of the cumulant-based approach improves with increasing data length for all SNR's, whereas only high-SNR results improve for the covariance ESPRIT.

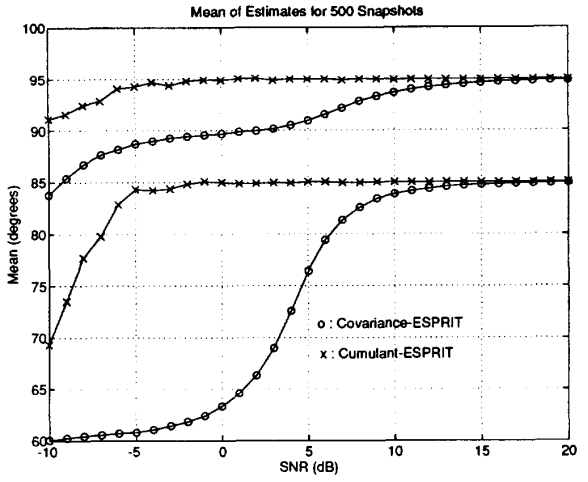


Fig. 7. Mean of estimates for 500 snapshots: Cumulant-based ESPRIT recovers from colored noise effects much faster than does the covariance-ESPRIT. The actual sources are located at 85 and 95°

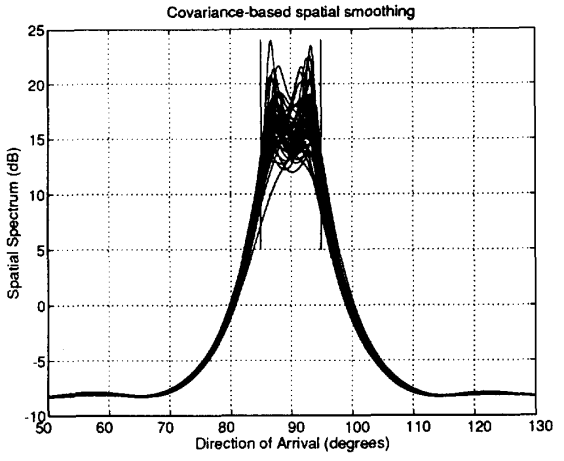


Fig. 8. Covariance-based MUSIC algorithm is unable to resolve the sources in general. Even when resolution is possible, the estimates are biased. The vertical lines indicate true source locations.

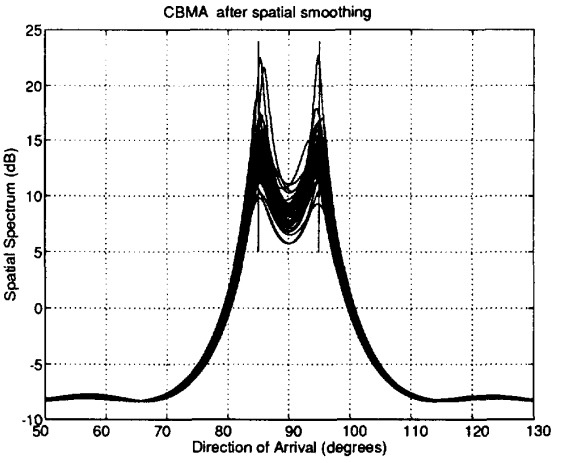


Fig. 9. Cumulant-based algorithm that uses an extra (satellite) sensor successfully resolves the sources and estimates the directions without bias. The vertical lines indicate true source locations.

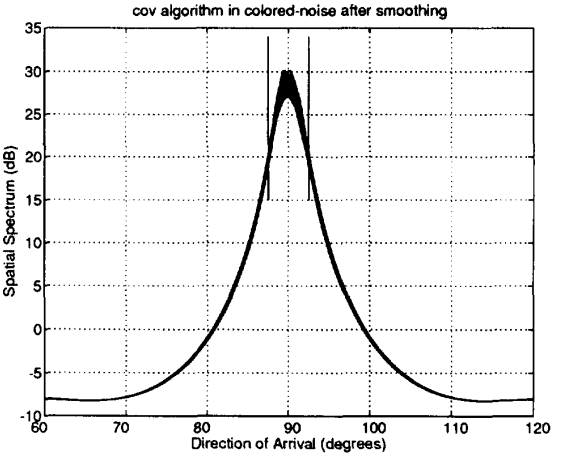


Fig. 10. Covariance-based MUSIC algorithm (cov) is unable to resolve the sources when the non-Gaussian noise is colored. The vertical lines indicate true source locations.

without reducing array dimensionality. This is important because reducing array dimensionality reduces the resolution of the array. Beamspace processing is known to partially correct colored noise effects by using *a priori* information about source locations in spatial preprocessing that decreases array dimensionality significantly. Our cumulant algorithm first uses preprocessing to create $g(t)$, which, in turn, is used to form an $M \times M$ cumulant matrix (M is the number of sensors) in which noise contributions are suppressed. By using cumulants, we achieve preprocessing-based noise suppression without reducing array dimensionality.

Experiment 3: Coherent Sources in Non-Gaussian Noise

In this experiment, we show how to incorporate a satellite sensor measurement into an array processing scenario that requires spatial smoothing, i.e., into a linear array. To compare the performance of the covariance- and cumulant-based algorithms for the case of coherent sources, we used a uniform linear array with eight sensors in order to improve resolution and enable spatial smoothing [12]. The uniform spacing between sensors is half wavelength.

First, we investigate the case of spatially white non-Gaussian noise. We consider a BPSK source that illuminates the array from 85°, and due to multipath, a perfectly coherent equal-power replica illuminates the array from 95°. We use an additional sensor in order to apply the method described in Theorem 2. This sensor is located 10 wavelengths to the left of the first element of the linear array. A covariance-based MUSIC algorithm cannot incorporate this measurement in the present case of coherent sources because the resulting array manifold does not possess a VanderMonde structure, which is required by the spatial smoothing algorithm [12] to decorrelate the sources.

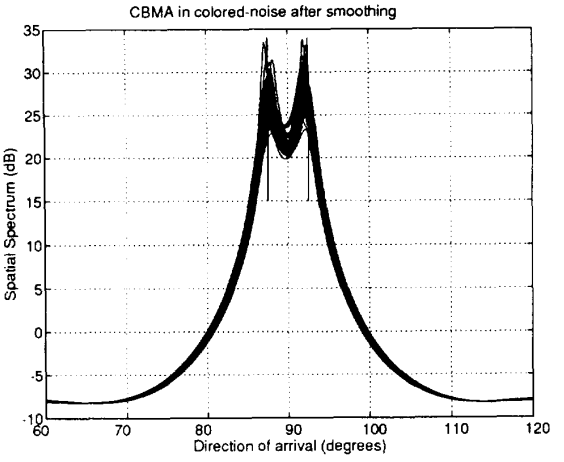


Fig. 11. Cumulant-based algorithm that uses an extra (satellite) sensor successfully resolves the sources in colored non-Gaussian noise. The estimates can be fine tuned by the SFS algorithm, which uses CBMA estimates for initialization. The vertical lines indicate true source locations.

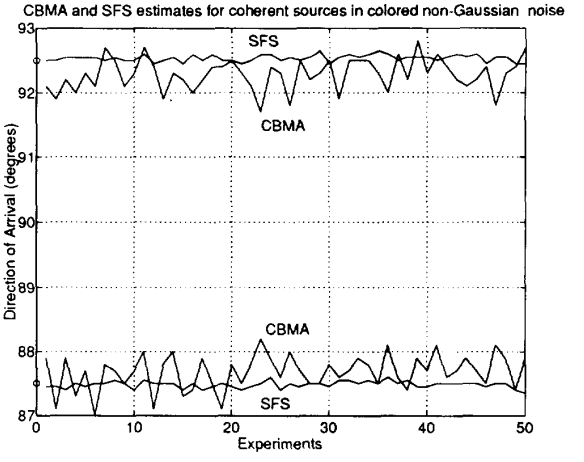


Fig. 12. SFS and CBMA performance comparison: SFS decreases the variation and the bias (due to finite number of samples) of CBMA estimates since it uses second-order statistics for estimation and fourth-order statistics for initialization. The mean of CBMA estimates are {87.680, 92.280 }, whereas the mean of SFS estimates are {87.482, 92.536}.

The noise components are assumed to originate from 4-QAM communications equipment, i.e., they are not a mixture of Gaussian and non-Gaussian components as we have considered in the previous experiments. This is done to introduce more problems to the cumulant-based approach: From cumulant properties, it is well known that cumulants can suppress additive Gaussian noise (in theory) but not non-Gaussian noise. We assume the noise power is identical (unity variance) at all sensors including the satellite sensor and that the signal power is equal to the noise power (0 dB). We used the MUSIC algorithm (cov) after one-level of spatial-smoothing to obtain a 7×7 matrix from the original 8×8 array covariance matrix [12] to investigate the performance of signal coherence on second-order statistics-based direction finding. Similarly, we employed the MUSIC algorithm after applying one level of spatial smoothing (which decreases the size of the matrix from 8×8 to 7×7) to the array cumulant matrix defined in (9) and (10) to investigate the performance of the cumulant-based MUSIC algorithm (CBMA). We used 1000 snapshots to estimate the required statistics and display spatial spectra in Fig. 8 for 50 independent realizations. 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, CBMA is able to resolve the sources, as illustrated in Fig. 9.

Next, we investigate the effects of colored non-Gaussian noise on the direction-finding methods. To make matters even worse, we assume the two coherent wavefronts are closer to each other: the bearings are now $\{87.5, 92.5°\}$. The noise covariance matrix for the main array takes the form $\mathbf{R_n} = \mathbf{a}(90^o)\mathbf{a}^H(90^o) + 0.01\mathbf{I}$, where $\mathbf{a}(90^o)$ is the steering vector that corresponds to 90°, i.e., $\mathbf{a}(90^o) = [1, 1, \ldots, 1]^T$. $\mathbf{R_n}$ represents an ambient noise structure whose major component illuminates the array from 90° (broadside) and shadows the presence of sources. The noise power at the satellite sensor remains at unity, and the signal power remains at 0 dB. Fig. 10 illustrates the results from the covariance-based MUSIC algorithm in this scenario: Sources are never resolved since the processor confuses the noise with a signal that arrives from the noise direction of 90°. On the other hand, the cumulant-based MUSIC algorithm (see Fig. 11) successfully resolves the two sources and suppresses the noise; however, CBMA estimates are slightly biased because the sample size (1000 snapshots) is not large enough to suppress the effects of the high-power noise source from 0°, which leaks into the spatial smoothing algorithm and pulls the estimates toward 0°. This observation is in accordance with the results of Xu and Buckley [15], who 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 V. We initialized the search required by SFS using the results of CBMA. We display the estimates provided by CBMA and the suboptimal SFS algorithm (in which $\boldsymbol{\Sigma}$ is replaced by $\mathbf{I}$ in (26)) for 50 trials in Fig. 12. Observe that suboptimal-SFS significantly reduces the variation and bias in the estimates.

Experiment 4: Virtual-ESPRIT and Non-Gaussian Noise

In this experiment, we used the eight-element linear array of Experiment 3 with the same noise correlation structure and strength. Two equal power, independent signals illuminate the array from $\{87.5, 92.5°\}$. Non-Gaussian noise suppression can be achieved by 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.
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.

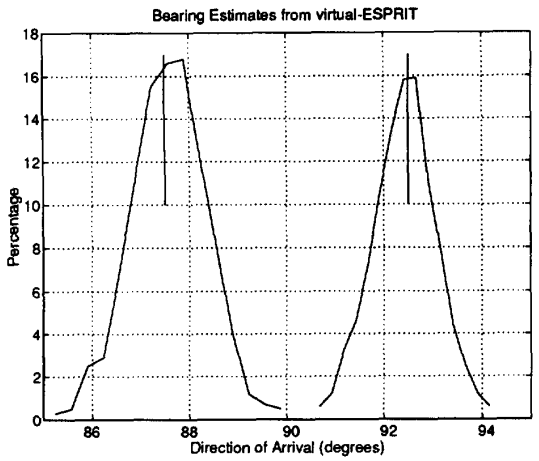


Fig. 13. Virtual-ESPRIT algorithm can estimate the source bearings in the presence of non-Gaussian noise. The graph indicates that the sources are resolved successfully. The mean of the estimates are {87.5852, 92.4206}.

The first method applies only when the response of one element in the main array is known. This is the major reason for using the second approach. In addition, the second method can be made insensitive to the noise correlation structure between the two guiding sensors if we only create a copy of the original array, i.e., an 8×8 cross-correlation matrix between the original array and its virtual copy, rather than a 10×10 matrix (see the description of VESPA in [4]). Consequently, we used two guiding sensors separated by $\lambda/2$. The guiding sensors are located on the axis of the main array: the first one of which is 10λ to the 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 0 dB with respect to noise at the guiding sensors.

We performed 1000 independent experiments to estimate source directions using virtual ESPRIT. The distribution of estimates is given in Fig. 13. VESPA resolves the sources as does the cumulant-based MUSIC algorithm (CBMA) described in Experiment 3. The bias in the estimates is less than that of the CBMA algorithm since sources are independent for this experiment.

VII. Conclusions

We have developed algorithms that are capable of suppressing the effects of non-Gaussian noise in array processing problems. We accomplished this by using the geometric interpretation of cumulants for array processing problems developed in the companion paper [4]. We first showed how to suppress statistically independent non-Gaussian noise that has different statistics at each sensor. Then, we generalized this method to suppress correlated non-Gaussian noise by using an additional sensor that is remotely located to the main array and whose noise component is uncorrelated with that of the main array. We also showed that it is possible to improve cumulant-based results by using second-order statistics.

Our simulations indicated that using second-order statistics and fourth-order cumulants together significantly improves the bias and standard deviation in the estimates over a cumulant-based algorithm. Note, however, that the performance of cumulant-based non-Gaussian noise suppression results are dependent on the degree of non-Gaussianity of the source waveforms, which indicates that cumulant-based algorithms may exhibit poor performance if the sources of interest are near-Gaussian. The performance of cumulant-based algorithms depend on data length and SNR results when the other source parameters are fixed (e.g., DOA's). In the simulations, it is observed that cumulant-based noise suppression improves only with increasing data length rather than with increasing SNR after the SNR exceeds a certain (scenario-dependent) level. Fig. 4 indicates this saturation effect for the two-source case. At high SNR's, the cumulant estimates are corrupted by cross-signal terms, which vanish in theory, and decay as data length grows to infinity. Use of second-order statistics (the SFS method) provides a way to eliminate the saturation effect because the distortion due to cross-signal terms decay at a faster rate for the estimates of second-order statistics.

Our overall conclusions are as follows:

- The richness of fourth-order cumulants over second-order statistics in terms of arguments, provides ways to increase the effective aperture of antenna arrays and reduce the adverse effects of additive correlated non-Gaussian noise.
- Combining second and higher order statistics may provide better results than the results obtained using only cumulants. The performance analysis of this approach is currently under investigation. In the Appendix, it is shown that estimation accuracy of the second-order statistics vector $\hat{\mathbf{d}}_N$ is independent of the higher-order statistics of the noise components in the main and satellite sensor measurements.
- Suppressing correlated measurement noise requires the use of an additional sensor that is located far-away from the main array so that its noise component is not correlated with the noise components of the main array; however, it is also possible to suppress non-Gaussian noise by using the received signals from the main array using cumulants and *a priori* information about the sources of interest without using an additional sensor.
- ESPRIT can be implemented in a practical manner using cumulants. Doing this saves hardware costs and achieves non-Gaussian as well as Gaussian noise suppression. This implementation is extended to the case of coherent sources in [6].

Appendix

Here, we derive the asymptotic covariance, which is given in (23), of the estimation error associated with $\hat{\mathbf{d}}_N$ as defined in (21). From (21) and $\mathbf{x}(t) = \text{col}[r_1(t), r_2(t), \ldots, r_M(t)]$

$$\hat{d}_N(m) = \frac{1}{N} \sum_{t=1}^{N} g^*(t) r_m(t) \tag{37}$$

and hence

$$\Sigma_{m,n} \triangleq \lim_{N\to+\infty} NE\{(\hat{d}_N(m) - d(m))(\hat{d}_N(n) - d(n))^*\}$$
$$= \lim_{N\to\infty} E\{\frac{1}{N}\sum_{t_1,t_2=1}^{N} g^*(t_1)r_m(t_1)g(t_2)r_n^*(t_2)\}$$
$$- E\{d^*(n)\sum_{t_1=1}^{N} g^*(t_1)r_m(t_1)\}$$
$$- E\{d(m)\sum_{t_1=1}^{N} g(t_2)r_n^*(t_2)\} + Nd(m)d^*(n) \quad (38)$$

where, using the definition of $\mathbf{d}$ in (19), we observe $d(m) = E\{g^*(t_1)r_m(t_1)\}$. Consequently, we are able to reexpress (38) as

$$\Sigma_{m,n} = \lim_{N\to\infty}\frac{1}{N}\sum_{t_1=t_2=t=1}^{N} E\{|g(t)|^2 r_m(t)r_n^*(t)\}$$
$$+ \frac{1}{N}\sum_{t_1\neq t_2=1}^{N} E\{g^*(t_1)r_m(t_1)\}E\{g(t_2)r_n^*(t_2)\}$$
$$- Nd(m)d^*(n). \quad (39)$$

The first summation has N identical terms, and the second summation has $(N^2 - N)$ identical terms. Using the definition of $d(m)$, we can express (39) as

$$\Sigma_{m,n} = \lim_{N\to\infty} E\{|g(t)|^2 r_m(t)r_n^*(t)\} - d(m)d^*(n)$$
$$= E\{|g(t)|^2 r_m(t)r_n^*(t)\} - d(m)d^*(n) \quad (40)$$

which is the result stated in (23).

To gain more insight into the asymptotic covariance matrix of the vector $\hat{\mathbf{d}}_N$, we first reexpress (40) in terms of cumulants as

$$\Sigma_{m,n} = \text{cum}(g^*(t), g(t), r_n^*(t), r_m(t))$$
$$+ E\{|g(t)|^2\}E\{r_m(t)r_n^*(t)\}$$
$$+ E\{g_1(t)r_m(t)\}E\{g_1^*(t)r_n^*(t)\}. \quad (41)$$

The last term vanishes for measurements that are circularly symmetric. Next, assume that there exist P far-field sources with powers $\{\sigma_k^2\}_{k=1}^P$, fourth-order cumulants $\{\gamma_{4,k}\}_{k=1}^P$, and steering vectors $\{\mathbf{a}_k\}_{k=1}^P$. If the noise covariance matrix for the main array is denoted as $\mathbf{R_n}$, the response of the satellite sensor to the kth source is g_k and the variance of noise in $g(t)$ is $\sigma_{n,v}^2$, then the matrix form of (41) is

$$\mathbf{\Sigma} = \sum_{k=1}^{P}\gamma_{4,k}|g_k|^2\mathbf{a}_k\mathbf{a}_k^H + (\sigma_{n,v}^2$$
$$+ \sum_{k=1}^{P}\sigma_k^2|g_k|^2)(\sum_{k=1}^{P}\sigma_k^2\mathbf{a}_k\mathbf{a}_k^H + \mathbf{R_n}) \quad (42)$$

which can be simplified to

$$\mathbf{\Sigma} = \sum_{k=1}^{P}(\gamma_{4,k}|g_k|^2 + \alpha\sigma_k^2)\mathbf{a}_k\mathbf{a}_k^H + \alpha\mathbf{R_n} \quad (43)$$

where $\alpha \triangleq (\sigma_{n,v}^2 + \sum_{k=1}^{P}\sigma_k^2|g_k|^2)$. Observe that $\mathbf{\Sigma}$ (covariance matrix of a second-order statistics vector) depends *only on* the second-order statistics of the noise components; it does not depend on the higher order statistics of the noise components. In other words, the degree of non-Gaussianity of the noise is irrelevant for the estimation accuracy of the estimated cross-correlation vector $\hat{\mathbf{d}}_N$.

ACKNOWLEDGMENT

This work was performed under a gift from Rockwell International Science Center, Thousand Oaks, CA. The final revision was performed while the first author was employed by TRW Avionics and Surveillance Group (ASG)—Sunnyvale (formerly ESL), CA. The first author thanks K. Anderson, J. Jenkins, M. Rands, and S. Stearns of TRW ASG—Sunnyvale for their help and support during this period.

The authors would like to thank the anonymous reviewers of this paper for their encouraging comments and important suggestions which significantly improved the paper.

REFERENCES

[1] J. F. Cardoso, "Higher-order narrowband array processing," *in Proc. Int. Conf. Higher Order Statistics*, Chamrousse, France, July 10-12, 1991, pp. 121–130.

[2] P. McCullagh, *Tensor Methods in Statistics*. New York: Chapman and Hall, 1987, Monographs on Statistics and Applied Probability.

[3] M. C. Doğan, "Cumulants and array processing," Ph.D. Dissertation, Signal Image Processing Inst., Univ. Southern Calif., Los Angeles, Dec. 1993.

[4] M. C. Doğan and J. M. Mendel, "Applications of cumulants for array processing, Part I: Aperture extension and array calibration," accepted for publication, *IEEE Trans. Signal Processing*.

[5] ——, "Single sensor detection and classification of multiple sources by higher order spectra," *Proc., Inst. Elec. Eng. F, Radar and Signal Processing*, Special issue on higher-order statistics, vol. 140, no. 6, pp. 350–355, Dec. 1993.

[6] E. Gönen, M. C. Doğan, and J. M. Mendel, "Applications of cumulants to array processing: direction-finding in a coherent signal environment," in *Proc. Twenty-Eighth Asilomar Conf. Signals, Syst. Comput.*, Monterey, CA, Oct. 1994.

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

[8] R. Pan and C. L. Nikias, "Harmonic decomposition methods in cumulant domains," in *Proc. ICASSP'88*, New-York, Apr. 1988, pp. 2356–2359.

[9] B. Porat and B. Friedlander, "Direction finding algorithms based on high-order statistics," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. SP-39, no. 9, pp. 2016–2024, Sept. 1991.

[10] R. Roy and T. Kailath, "ESPRIT—Estimation of signal parameters via rotational invariance techniques," *Opt. Eng.*, vol. 29, no. 4, pp. 296–313, Apr. 1990.

[11] S. Shamsunder and G. Giannakis, "Non-Gaussian source localization via exploitation of higher-order cyclostationarity," in *Proc. Sixth Signal Processing Workshop on Statistical Signal Array Processing* Victoria, Canada, Oct. 1992, pp. 193–196.

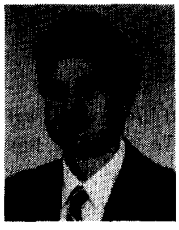
[12] T. Shan and T. Kailath, "Adaptive beamforming for coherent signals and interference," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-33, no. 3, pp. 527–536, June 1985.

[13] A. Swami and J. M. Mendel, "Cumulant-based approach to harmonic retrieval and related problems," *IEEE Trans. Signal Processing*, vol. 39, no. 5, pp. 1099–1109, May 1991.

[14] Q. Wu, K. M. Wong, and J. P. Reilly, "Maximum-likelihood estimation for array processing in unknown noise environments," *in Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, San Fransisco, Mar. 1992, pp. 241–244, vol. 5.

[15] 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, Oct. 1992.

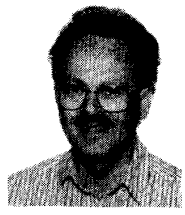
[16] I. Ziskind and M. Wax, "Maximum likelihood localization of multiple sources by alternating projection," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, no. 10, pp. 1553–1560, Oct. 1988.



Mithat C. Doğan (M'94) was born on March 1, 1967 in Ankara, Turkey. He received the B.S.E.E degree as a top-ranking senior from the Middle East Technical University (METU), Ankara, Turkey, in 1988, and the M.S.E.E. and Ph.D. degrees from the University of Southern California (USC), Los Angeles, in 1989 and 1994 respectively.

He joined TRW Avionics and Surveillance Group (ASG) (formerly ESL), Sunnyvale, CA, in 1994. His research interests are applications of higher order statistical techniques to communications signal analysis. Sensor array processing and blind equalization are his major topics of interest.

Dr. Doğan received several awards in recognition of academic excellence from the Scientific and Technical Research Council of Turkey (TÜBİTAK), including a NATO Science Fellowship for graduate studies in 1988. He is currently a honorary fellow of TÜBİTAK and a reviewer for several IEEE TRANSACTIONS.



Jerry M. Mendel (F'78) received the Ph.D. degree in electrical engineering from the Polytechnic Institute, Brooklyn, NY, in 1963.

He is currently Professor and Associate Chair of Electrical Engineering-Systems at the Unbiversity of Southern California, Los Angeles, where he has been since 1974. He has published more than 290 technical papers and is the author of five books, including *Lessons in Estimation Theory for Signal Processing, Communications, and Control* (Englewood Cliffs, NJ: Prentice-Hall, 1995) and *Maximum-Likelihood Deconvolution* (New York: Springer-Verlag, 1990). He is the editor of the book *A Prelude to Neural Networks: Adaptive and Learning Systems*. He is also the author of the IEEE Individual Learning Program *Kalman Filtering and Other Digital Estimation Techniques*.

Dr. Mendel is a Distinguished Member of the IEEE Control Systems Society, a member of the IEEE Signal Processing Society, the Society of Exploration Geophysicists, the European Association for Signal Processing, Tau Beta Pi, Pi Tau Sigma, and Sigma Xi. He was President of the IEEE Control Systems Society in 1986. He has received numerous awards including the 1983 Best Transactions Paper Award from the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, the 1992 Signal Processing Society Paper Award from the IEEE TRANSACTIONS ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING, and a 1984 Centennial Medal. He served as Editor of the IEEE Control Systems Society's IEEE TRANSACTIONS ON AUTOMATIC CONTROL and is presently an Associate Editor of the IEEE TRANSACTIONS ON FUZZY SYSTEMS.