

---

---

# Utilizing Waveform Features for Adaptive Beamforming and Direction Finding with Narrowband Signals

Keith W. Forsythe

■ Extensive research has been done on the use of antenna arrays for direction finding and beamforming; this research focuses on the detailed behavior of specific techniques rather than on actual signal processing applications. In most applications, there is a fundamental signal feature that provides essential leverage for an effective processing approach. This article, which is structured around such features, presents a comprehensive framework for selecting an appropriate adaptive approach for processing cochannel narrowband signals. We address the roles of antenna calibration and prior waveform knowledge, and give examples of effective, practical direction-finding and beamforming procedures that cover a wide range of potential applications.

**D**IRECTION FINDING and adaptive beamforming for narrowband sources in cochannel interference are traditionally performed with adaptive antenna arrays by employing high-resolution spectral estimators such as multiple signal classification, or MUSIC [1–3]. These estimators resolve closely spaced emitters by eigenanalyzing an array covariance matrix, which, along with an antenna-calibration table, provides enough information to estimate angles of arrival and form adaptive beams. Antenna-calibration errors can cause high-resolution spectral estimators to behave poorly. Consequently, adaptive beamformers based on direction-finding estimates also can suffer performance losses.

Known properties of signal waveforms can often be exploited to form adaptive beams without relying on any antenna calibration. In many cases, the level of performance achieved is close to that of the best possible adaptive beamformer: one that maximizes signal-to-interference-plus-noise ratio (SINR). In ad-

dition, the combination of waveform feature exploitation and antenna calibration can lead to better direction-finding algorithms.

A simple model of a snapshot taken at the output of an array can be expressed by

$$\mathbf{z}(t) = \sum_{i=1}^S \mathbf{x}_T(u_i) a_i(t) + \mathbf{n}(t), \quad (1)$$

where  $\mathbf{z}(t)$  is the vector of  $M$  sensor outputs at time  $t$ ,  $S$  is the number of signals received at the array,  $\mathbf{x}_T(u_i)$  is the array-response vector associated with the  $i$ th signal arriving from direction  $u_i$ ,  $a_i(t)$  is the amplitude of the  $i$ th signal, and  $\mathbf{n}(t)$  is receiver noise plus background noise. The subscript  $T$  on the array-response vector  $\mathbf{x}_T$  indicates the “true” array response rather than an array response that matches an entry in the array’s calibration table  $\mathbf{x}(u)$ . The two differ by calibration errors, which are an important component of system performance. For simplicity, the

model does not involve polarization explicitly. We make the additional assumption that the direction parameters  $u_i$  are scalars.

The signal model represented by Equation 1 encompasses a variety of different approaches to adaptive-array processing for narrowband waveforms [4]. For the purpose of discussion, we classify adaptive-array processing techniques into two categories. They can either be calibration based or waveform based. Calibration-based techniques rely exclusively on accurate antenna calibration to form adaptive beams and provide direction finding. Waveform-based approaches additionally utilize information about the structure of the received signals' waveforms to form adaptive beams or to perform direction finding.

In some cases, the joint use of calibration and array knowledge is not robust to modeling errors (usually because of antenna-calibration errors as opposed to waveform uncertainties). Fortunately, many waveform-based procedures do not require calibration for adaptive beamforming. Thus waveform-based techniques can be used first for adaptive beamforming, then calibration can be employed for what is often called copy-based direction finding.

We discuss calibration-based and waveform-based approaches to adaptive beamforming and direction finding below. We also discuss copy-based direction finding. The discussions focus on various assumptions that can be made concerning the array response and signal parameters illustrated in Equation 1. We present examples of procedures that are built on these assumptions.

### Calibration-Based Adaptive-Array Processing

Most of the superresolution procedures treated in the literature are based on accurate antenna calibration. For completeness we present here a brief survey of calibration-based techniques. To simplify notation, our algorithm discussions assume perfect array calibration. Since calibration errors are an important source of direction-finding and copy errors, we treat the effects of calibration errors later.

The notation introduced in Equation 1 can be organized in a more useful manner for presenting the results in the following sections. Specifically, we can define the more compact notation

$$\begin{aligned} \mathbf{Z} &= [\mathbf{z}(t_1), \dots, \mathbf{z}(t_L)], \\ \mathbf{X} &= [\mathbf{x}(u_1), \dots, \mathbf{x}(u_S)], \\ \hat{\mathbf{R}} &= L^{-1} \sum_k \mathbf{z}(t_k) \mathbf{z}(t_k)^H = L^{-1} \mathbf{Z} \mathbf{Z}^H, \end{aligned}$$

and

$$\mathbf{A} = \begin{pmatrix} a_1(t_1) & \cdots & a_1(t_L) \\ \vdots & \ddots & \vdots \\ a_S(t_1) & \cdots & a_S(t_L) \end{pmatrix}.$$

Then we can rewrite Equation 1 as  $\mathbf{Z} = \mathbf{X} \mathbf{A} + \mathbf{N}$ , where  $\mathbf{N}$  is a matrix with independent, identically distributed columns representing receiver and background noise. For the remainder of this section, we assume that the receiver noise and background noise are spatially white with unit complex variance. In other words, the entries of  $\mathbf{N}$  are independent, identically distributed, complex circular Gaussian random variables of unit variance.

For the purpose of discussing calibration-based adaptive-array processing, we further categorize adaptive-beamforming and direction-finding procedures based on this equation by assumptions made about the signal amplitude matrix  $\mathbf{A}$ . These assumptions amount to a Gaussian or deterministic model for the signal amplitudes. In the deterministic case, the amplitudes (the entries of  $\mathbf{A}$ ) are assumed to be unknown unconstrained parameters that must be estimated in the process of beamforming or direction finding. Later, in the discussion of waveform-based adaptive-array processing, constraints are placed on the allowable signal amplitudes. In the Gaussian case, the signal amplitude matrix  $\mathbf{A}$  is assumed to have independent, identically distributed columns with common covariance

$$\mathbf{P} = L^{-1} \mathbf{E}[\mathbf{A} \mathbf{A}^H],$$

where  $\mathbf{E}[\ ]$  represents the expectation of  $\mathbf{A} \mathbf{A}^H$ . Both cases are treated in more detail below.

Central to the discussion of adaptive-array processing is the measurement of signal separations. Traditionally, signal separation is measured as a fraction of the width of a steered, nonadaptive beam formed by the array. Such definitions are often formulated on a

case-by-case basis for each array geometry. A more abstract definition of array beamwidth, presented in the sidebar entitled “Definition and Properties of Beamwidth,” is closer to the mathematical notion of separation and agrees as well with physical separations when the separations are small.

### Deterministic Signal Model

As discussed above, the deterministic signal model assumes that the baseband antenna-array data  $\mathbf{Z}$  are complex circular Gaussian random matrices with mean  $\mathbf{X}\mathbf{A}$  and identity (white) noise covariance

$$\mathbb{E}[\mathbf{N}(t)\mathbf{N}(t)^H] = \mathbf{I}_M.$$

The array covariance of the data  $\mathbf{Z}$  can be expressed as

$$\mathbb{E}[\hat{\mathbf{R}}] = \mathbf{X}\mathbf{P}\mathbf{X}^H + \mathbf{I}_M,$$

where  $\mathbf{P} = L^{-1}\mathbf{A}\mathbf{A}^H$ . Note that we will change the definition of  $\mathbf{P}$  to suit the signal model under consideration (deterministic or Gaussian). Thus the probability density function of the deterministic signal model can be written as

$$p_{DS}(\mathbf{Z} | \mathbf{X}, \mathbf{A}) = \pi^{-LM} e^{-\text{tr}[(\mathbf{Z}-\mathbf{X}\mathbf{A})(\mathbf{Z}-\mathbf{X}\mathbf{A})^H]}, \quad (2)$$

through which the maximum-likelihood estimates of the signal direction parameters  $u_i$  and the signal amplitude matrix  $\mathbf{A}$  can be formulated. Let

$$\mathbf{P}_{\mathbf{X}}(\mathbf{u}) = \mathbf{X}(\mathbf{u})[\mathbf{X}(\mathbf{u})^H \mathbf{X}(\mathbf{u})]^{-1} \mathbf{X}(\mathbf{u})^H$$

define the projection matrix, parameterized by the directions  $\mathbf{u} = (u_1, \dots, u_G)$ , onto the subspace spanned by the columns of  $\mathbf{X}$ . Then maximum-likelihood direction finding is performed by estimating the signal direction parameters  $u_i$  as expressed by solving

$$\arg \max_{\mathbf{u}} \text{tr}[\mathbf{P}_{\mathbf{X}}(\mathbf{u}) \hat{\mathbf{R}}], \quad (3)$$

where  $\text{tr}[\ ]$  denotes trace. Maximum-likelihood signal amplitude estimates of  $\mathbf{A}$  are provided by

$$\hat{\mathbf{A}} = \mathbf{W}^H \mathbf{Z},$$

where the matrix  $\mathbf{W}$  represents the outputs of a beamforming matrix

$$\mathbf{W} = \hat{\mathbf{X}}(\hat{\mathbf{X}}^H \hat{\mathbf{X}})^{-1}.$$

Each column of  $\mathbf{W}$  forms a beam steered at the signal represented by the corresponding column of  $\hat{\mathbf{X}}$ , which consists of the array-response estimates based on the signal direction parameters estimated by solving Equation 3. Each adaptive beam tries to place nulls on all but one signal (note that  $\mathbf{W}^H \hat{\mathbf{X}} = \mathbf{I}_G$ ).

### Gaussian Signal Model

The signal amplitudes can also be modeled as complex Gaussian random variables with zero mean and unknown covariance

$$\mathbf{P} = L^{-1} \mathbb{E}[\mathbf{A}\mathbf{A}^H].$$

In this case the array covariance becomes

$$\mathbf{R} = \mathbb{E}[\hat{\mathbf{R}}] = \mathbf{X}\mathbf{P}\mathbf{X}^H + \mathbf{I}_M.$$

The probability distribution function of the Gaussian signal model of the data becomes

$$p_G(\mathbf{Z} | \mathbf{X}, \mathbf{P}) = \pi^{-LM} |\mathbf{R}|^{-L} e^{-\text{tr}(\mathbf{R}^{-1} \mathbf{Z}\mathbf{Z}^H)}.$$

We define the real-valued function  $g(x)$ :

$$g(x) = \begin{cases} x - \ln x & x \geq 1 \\ 1 & 0 \leq x < 1 \end{cases}$$

For any hermitian matrix  $\Gamma$ , we can define  $g(\Gamma)$  as follows. Let  $\mathbf{U}\mathbf{\Lambda}\mathbf{U}^H = \Gamma$  express the eigenanalysis of  $\Gamma$  in terms of the unitary matrix  $\mathbf{U}$  whose columns are eigenvectors of  $\Gamma$ , and the diagonal matrix  $\mathbf{\Lambda}$  whose entries are the corresponding eigenvalues. Then we define

$$g(\Gamma) = \mathbf{U} \text{diag}[g(\lambda_1), \dots, g(\lambda_M)] \mathbf{U}^H,$$

where the  $\lambda_k$  are the diagonal entries of  $\mathbf{\Lambda}$ . We define

$$\hat{\mathbf{R}}_{\mathbf{X}} = (\mathbf{X}^H \mathbf{X})^{-1/2} \mathbf{X}^H \mathbf{R} \mathbf{X} (\mathbf{X}^H \mathbf{X})^{-1/2}.$$

Then maximum-likelihood direction finding is performed by finding

$$\arg \max_{\mathbf{u}} \text{tr}[g(\hat{\mathbf{R}}_{\mathbf{X}})].$$

Asymptotically,  $g(x) \approx x$  when  $x \gg 1$ . Thus for  $\mathbf{P} \gg \mathbf{I}_M$ ,

$$\text{tr}[g(\hat{\mathbf{R}}_{\mathbf{X}})] \approx \text{tr}(\hat{\mathbf{R}}_{\mathbf{X}}) = \text{tr}(\mathbf{P}_{\mathbf{X}} \hat{\mathbf{R}}),$$

## DEFINITION AND PROPERTIES OF BEAMWIDTH

THE ANGLE  $b$  between two array-response vectors is fundamentally important in the geometrical interpretation of adaptive-array signal processing. Mathematically,  $b$  is proportional to the geodesic distance between the array responses, interpreted as points in projective space. Specifically, two array responses  $\mathbf{v}_1$  and  $\mathbf{v}_2$  are  $b$  beamwidths apart whenever

$$\cos\left(\frac{\pi b}{2}\right) = \frac{|\mathbf{v}_1^H \mathbf{v}_2|}{\|\mathbf{v}_1\| \|\mathbf{v}_2\|}.$$

Note that  $b$  has been normalized so that  $b = 1$  corresponds to orthogonal array responses.

The abstract definition of beamwidth given above can be interpreted physically. Let the function  $\mathbf{v}(\theta, \phi)$  be the modeled array response parameterized by zenith angle (complement of elevation)  $\theta$  and azimuth angle  $\phi$ . A convenient coordinate system is provided by the unit direction-of-arrival vectors  $\mathbf{u}(\theta, \phi)$ , where

$$\mathbf{u}(\theta, \phi) = [\sin(\theta) \cos(\phi), \sin(\theta) \sin(\phi), \cos(\theta)]^T.$$

This vector points in the direction corresponding to  $(\theta, \phi)$ . For arrays with perfectly matched element patterns (a common modeling assumption), we can write the response of the  $k$ th element of the response vector  $\mathbf{v}$  as

$$v_k(\theta, \phi) = g[\mathbf{u}] e^{i(2\pi/\lambda) \mathbf{d}_k^T \mathbf{u}},$$

where  $\mathbf{d}_k$  is the three-dimensional location of the  $k$ th antenna element,  $g[\mathbf{u}(\theta, \phi)]$  is the antenna pattern common to all elements, and  $\lambda$  is the wavelength. Let

$$\mathbf{u}_\Delta = \mathbf{u}(\theta_1, \phi_1) - \mathbf{u}(\theta_2, \phi_2).$$

For small separation vectors  $\mathbf{u}_\Delta$ , we have the approximation

$$b^2 \approx \frac{16}{\lambda^2} \mathbf{u}_\Delta^T \mathbf{D} \mathbf{u}_\Delta$$

for  $b$  as defined above. The matrix  $\mathbf{D}$  is given by

$$\mathbf{D} = \frac{1}{M} \sum_{k=1}^M (\mathbf{d}_k - \mathbf{d}_{av})(\mathbf{d}_k - \mathbf{d}_{av})^T,$$

where  $\mathbf{d}_{av}$  is the average location vector. By extrapolating this expression out to  $b = 1$  (where the array responses  $\mathbf{v}(\theta_1, \phi_1)$  and  $\mathbf{v}(\theta_2, \phi_2)$  would be orthogonal if the approximation held), we have the definition of beamshape:

$$1 = \frac{16}{\lambda^2} \mathbf{u}_\Delta^T \mathbf{D} \mathbf{u}_\Delta. \quad (\text{A})$$

Two direction-of-arrival vectors differing by  $\mathbf{u}_\Delta$  are considered to be one beamwidth apart. This equation is motivated by a peak-to-null definition of beamwidth. As mentioned above, we choose the normalization of  $b$  so that  $b = 1$  corresponds to a null (orthogonality) in the pattern.

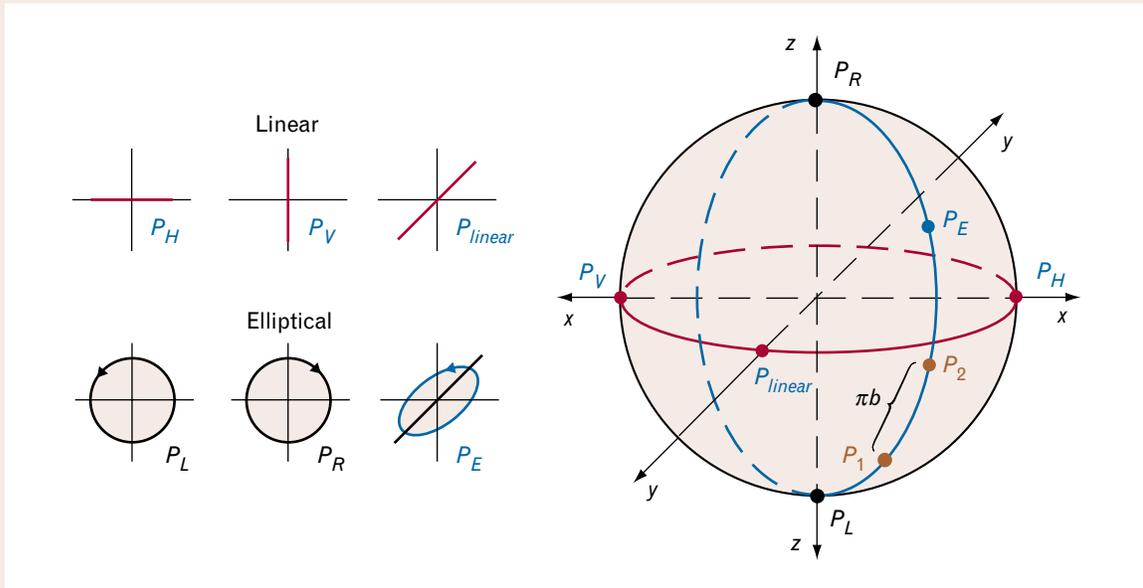
Equation A can be applied to arbitrary antenna-array geometries to yield a physical beam. For example, a completely filled (sampled arbitrarily finely) line array of length  $A$  has a beam of size  $\sqrt{3/4} \lambda/A$ . A filled circular array of diameter  $A$  has a circular beamshape of diameter  $\lambda/A$ . Both cases, and many others, indicate a beam size related to  $\lambda/A$ , where  $A$  is a characteristic size of the array. Similar beamshapes can be motivated by Cramér-Rao bounds on direction finding.

The physical geometry underlying beamwidths is particularly simple in the case of polarization. We consider a two-element array with polarization response to an incident signal  $P = E_R \hat{\mathbf{R}} + E_L \hat{\mathbf{L}}$ , given in terms of the right ( $E_R$ ) and left ( $E_L$ ) circular components (or any fixed orthogonal reference polarizations). The polarization ratio can be mapped to the unit sphere (Poincaré sphere) by

$$\frac{E_R}{E_L} = \tan(\theta/2) e^{i\phi},$$

where  $\theta$  and  $\phi$  are spherical coordinates. Then the separation in beamwidths between two incident polarizations  $P_1$  and  $P_2$  is given by

$$b(P_1, P_2) = \frac{\text{arc length}(P_1, P_2)}{\pi}.$$



**FIGURE A.** The Poincaré sphere as a geometric model of pure linear and elliptical polarization states. Each polarization state (elliptical, in general) corresponds to a point on the sphere. In particular, points on the equator correspond to linear polarizations. The beamwidth separation  $b$  between polarizations is expressed by the great-circle arc length between corresponding points on the sphere, divided by  $\pi$ . Orthogonal polarizations correspond to antipodal points, which are one beamwidth apart.

Figure A illustrates the intrinsic geometry for the case of a dual polarized element based on the Poincaré sphere. Linear polarizations are represented by points along the equator.

The ideal nulling capability of an array can be expressed in terms of the beamwidth separation. For example, Figure 1 in the main text shows the loss of SINR due to a single cochannel inter-

ferer at various beamwidth separations and interference levels.

Note that the interpretation of beamwidths in terms of arc length on the sphere holds for any two-element array.

at least when  $\mathbf{X}$  is near the correct value. Thus the direction-finding estimates under the Gaussian signal model approximate the estimates under the deterministic signal model when the signal power matrix  $\mathbf{P}$  is large. When the power matrix  $\mathbf{P}$  is small, the direction-finding procedures differ, especially when there is more than one signal in the model (i.e.,  $S > 1$ ). Gaussian direction-finding procedures are more accurate with low signal powers.

Note that the Gaussian signal model does not provide adaptive beamforming. In these parametric formulations of signal processing, beamforming is connected with signal amplitude estimates. The Gaussian signal model, however, does not involve parameters for the individual signal amplitudes.

The deterministic signal model provides a beamforming matrix  $\mathbf{W}$  that attempts to place a null in the estimated direction of each interferer. This procedure does not attempt to maximize SINR at the beamformer output and thus has suboptimal performance even if the direction-finding estimates are perfect. To maximize SINR for a signal with array response  $\mathbf{x}(u_i)$ , we form the weight vector  $\mathbf{w} = \hat{\mathbf{R}}^{-1}\mathbf{x}(u_i)$ . This beamformer can be formed under either signal model. Calibration errors have a significant impact on this type of beamformer, as we discuss below.

#### Other Approaches

There are a variety of other calibration-based approaches to direction finding and beamforming (too

many to mention here), some of which are motivated by the maximum-likelihood procedures discussed above. Given direction-finding estimates, beamforming is treated briefly in the section on adaptive beamforming and so will not be considered further in this section. Most techniques are based on the eigenanalysis of the array sample covariance

$$\hat{\mathbf{R}} = \hat{\mathbf{E}}_S \hat{\Lambda}_S \hat{\mathbf{E}}_S^H + \hat{\mathbf{E}}_N \hat{\Lambda}_N \hat{\mathbf{E}}_N^H,$$

where the columns of  $\hat{\mathbf{E}}_S$  and  $\hat{\mathbf{E}}_N$  are eigenvectors corresponding to the top  $S$  and bottom  $M - S$  eigenvalues of  $\hat{\mathbf{R}}$ , which are the entries of the diagonal matrices  $\hat{\Lambda}_S$  and  $\hat{\Lambda}_N$ , respectively. The column span of  $\hat{\mathbf{E}}_S$  is called the signal subspace, and the column span of  $\hat{\mathbf{E}}_N$  is called the noise subspace.

Subspace-fitting techniques are motivated by both the deterministic and Gaussian maximum-likelihood procedures. They solve

$$\arg \max_{\mathbf{u}} \text{tr}[\mathbf{P}_X \hat{\mathbf{E}}_S f(\hat{\Lambda}_S) \hat{\mathbf{E}}_S^H],$$

where  $f$  is one of a variety of scalar-valued functions. For example,

$$f(x) = \frac{(x - 1)^2}{x}$$

provides a technique that has the same asymptotic (in the number of samples) accuracy as Gaussian maximum likelihood (with the noise covariance  $\mathbf{I}_M$  replaced by  $\sigma^2 \mathbf{I}_M$  with unknown  $\sigma$ ) for Gaussian waveforms [5]. Choosing  $f(x) = 1$  provides a technique sometimes called multidimensional MUSIC. In fact, using this version of  $f$  and replacing the rank- $S$  projector  $\mathbf{P}_X$  with the rank-one projector  $\mathbf{x}(u) \mathbf{x}^H(u)$  (assuming  $\|\mathbf{x}\| = 1$ ) leads to the MUSIC statistic:

$$\begin{aligned} & \arg \max_{\mathbf{u}} \frac{\mathbf{x}^H(u) \hat{\mathbf{E}}_S \hat{\Lambda}_S \hat{\mathbf{E}}_S^H \mathbf{x}(u)}{\mathbf{x}^H(u) \mathbf{x}(u)} \\ & = \arg \min_{\mathbf{u}} \frac{\mathbf{x}^H(u) \hat{\mathbf{E}}_N \hat{\Lambda}_N \hat{\mathbf{E}}_N^H \mathbf{x}(u)}{\mathbf{x}^H(u) \mathbf{x}(u)}, \end{aligned}$$

since

$$\hat{\mathbf{E}}_S \hat{\Lambda}_S \hat{\mathbf{E}}_S^H + \hat{\mathbf{E}}_N \hat{\Lambda}_N \hat{\mathbf{E}}_N^H = \mathbf{I}_M.$$

For future reference, replacing  $\hat{\mathbf{E}}_S \hat{\Lambda}_S \hat{\mathbf{E}}_S^H$  with  $\hat{\mathbf{R}}$  results in a direction-finding procedure called beamsum, which, in effect, measures the power at the output of a steered (summing over element outputs) beam as a function of steering direction  $u$ . If sources are separated in angle sufficiently, direction-finding performance can be good. However, performance is often poor when sources are closely spaced (i.e., closely spaced in beamwidths; this can be typical if the array has high sidelobes).

When the array is uniform linear, the modeled array response can be written as

$$\mathbf{x}(u) = (x_1, \dots, x_M)^T,$$

with  $x_k = w^k$  and  $w = e^{2\pi i d u / \lambda}$ , and where  $d$  is the antenna-element spacing and  $u = \cos(\theta)$  is the direction parameter that corresponds to an angle of arrival  $\theta$  measured from the array axis. For this type of array, the function

$$\mathbf{x}^H(u) \hat{\mathbf{E}}_N \hat{\Lambda}_N \hat{\mathbf{E}}_N^H \mathbf{x}(u)$$

occurring in the MUSIC statistic can be written, after multiplying by  $w^J$  for some  $J$  (depending on the coordinates used to model the array), as a polynomial  $Q(w)$  in  $w$  restricted to the unit circle  $|w| = 1$ . Since the minima in this polynomial correspond to estimates of signal directions, it is natural to find the roots of  $Q(w)$  and use the angles (i.e., arguments) of the roots for signal direction estimates. Such a procedure is called root MUSIC [6, 7]. Other algorithms based on uniform linear arrays also have rooting versions. For MUSIC, rooting increases resolution without any reduction in estimation accuracy. Rooting can also be applied to planar arrays in order to estimate jointly two-component signal directions (e.g., azimuth and elevation). Techniques of this type are introduced in References 8 and 9. An additional advantage of rooting procedures, particularly for planar (or three-dimensional) arrays, is the ability to perform direction finding without searching over all spatial directions. Such a search is often the most costly aspect of an implementation.

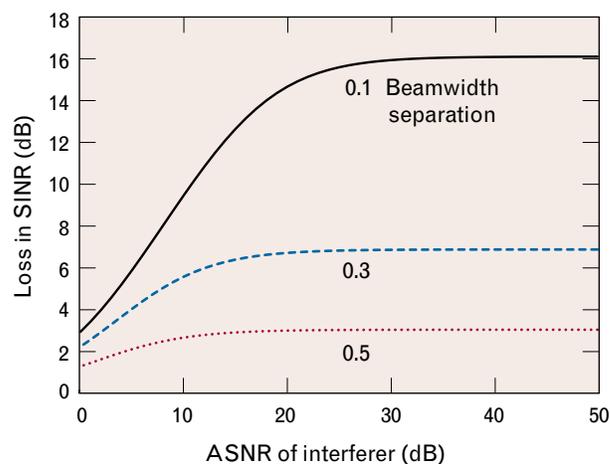
Another popular method, known as ESPRIT [10], is a technique for performing direction finding by comparing the data received simultaneously by two congruent translation-related subarrays of a given an-

tenna array. The direction estimates relate to the translation between the subarrays and thus consist of a single angle. ESPRIT provides simple angle estimates without a search over all spatial directions.

*Impact of Calibration Errors*

The direction-finding procedures described above have a predicted performance based on the evaluation of certain Cramér-Rao bounds or on behavior asymptotic in the number of samples. However, calibration errors tend to dominate performance at moderate to high signal levels. It is difficult to present a simple story of the effects of calibration errors in general. Precise statistical models of the errors are rarely available in practice and are needed to formulate performance bounds [11]. In the case of a single modeled signal ( $S = 1$ ), however, a simple result indicates the role of calibration errors, as discussed in the sidebar entitled “Single-Emitter Angle Estimation.”

*Direction-Finding Accuracy.* We can make some general comments concerning the effects of calibration errors on direction finding for multiple signals. First, techniques such as the deterministic-signal maximum likelihood or Gaussian-signal maximum likelihood are hypersensitive to calibration errors with strong signals. The same comments apply to most of the subspace-fitting procedures except those which place equal emphasis on all signal-subspace eigenvalues ( $f(x) \approx 1$  in the notation above). MUSIC is one of the more robust direction-finding procedures as well as one that is reasonable to implement. Rooting procedures or ESPRIT can offer implementation advantages over MUSIC (and, in the case of rooting, increased resolution), provided the arrays possess the appropriate structure. However, in some cases we can use arrays that do not have the requisite symmetry (because of calibration errors, for example). For the purposes of signal processing, we can obtain an effective symmetry condition either by preprocessing the data with a fixed beamforming matrix that forces its outputs toward the appropriate symmetry condition, or by using a Fourier-series model of antenna-element responses [4] and thus effectively replacing the physical array with a virtual array that possesses appropriate symmetry. The Fourier-series approach is typically better suited to handle array-response mismatches,



**FIGURE 1.** The presence of an interfering signal causes a loss in the signal-to-interference-plus-noise ratio (SINR) of a desired signal of interest at the output of a beamformer. With an ideal adaptive beamformer the loss can be minimized. The smallest possible loss of SINR is shown in terms of the strength of an interfering signal (measured as the array signal-to-noise ratio of the interferer, or ASNR) and the separation in beamwidths between the interfering signal and the signal of interest. The loss of SINR approaches a constant asymptote as the signal level of the interferer increases. In this asymptotic region, a null in the adaptive beam pattern moves ever closer to the interferer.

since an arbitrary number of basis functions can be used. Because ESPRIT cannot use the Fourier series approach, rooting procedures have an edge in many applications.

*Adaptive Beamforming.* Three types of SINR are useful for describing copy performance and scenario parameters. They are denoted array signal-to-noise ratio (ASNR), ideal array signal-to-interference-plus-noise ratio (ASINR), and achieved ASINR. All are SINRs at the output of different beamformers. ASNR, which indicates relative signal level, is the output of a beamformer focused on the signal in the absence of any interference except background noise. Ideal ASINR is measured at the output of the best possible beamformer: one that maximizes SINR in the given environment, including interference. Achieved ASINR is measured at the output of the beamformer determined by a copy algorithm.

In the absence of interference, performance is determined by the ASNR. With a cochannel interferer, the ideal beamformer offers a loss in SINR at the beam output, shown for example in Figure 1 as a

## SINGLE-EMITTER ANGLE ESTIMATION

THERE ARE A VARIETY of different models for uncertainties in antenna calibration. One of the simplest represents the true array response  $\mathbf{x}_T$  as a sum of an ideal array response (from the calibration table) and a vector of independent, identically distributed, complex circular Gaussian random variables representing the calibration errors. The error size is parameterized by the variance  $\epsilon^2$  of the complex Gaussian components. If  $\epsilon^2$  denotes the error variance per element for unity gain elements, then for small  $\epsilon$  the gain and phase errors are approximately  $(10\sqrt{2}/\ln 10)\epsilon$  dB and  $(180/\pi\sqrt{2})\epsilon$  degrees, respectively. Figure A shows the relationship between the variance of the element error and the size of gain/phase errors.

The accuracy of the direction finding estimate, in beamwidths, for a single angle variable  $\psi$  is

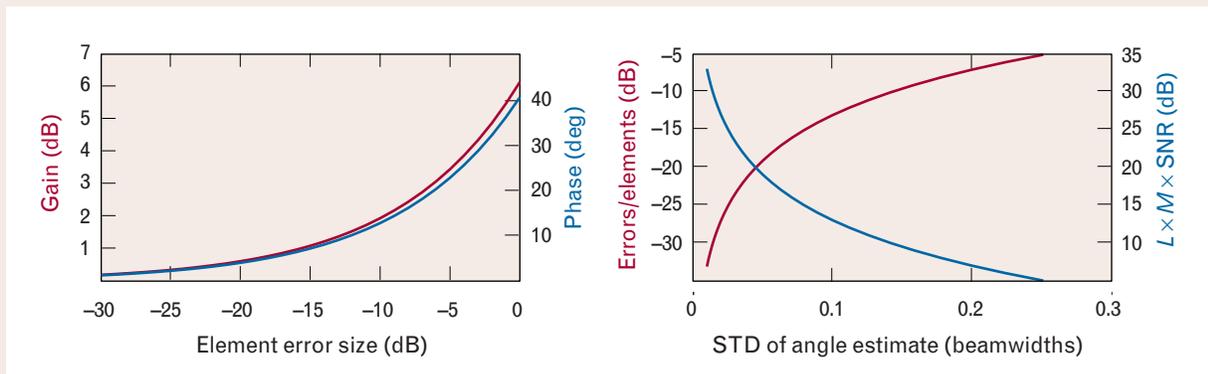
$$\text{STD}(\psi) = \sqrt{\frac{2}{\pi^2 M \rho}},$$

where  $\rho = \epsilon^{-2}$  with calibration errors (no noise), and  $\rho = L \times \text{SNR}$  for perfect calibration (SNR is the signal-to-noise ratio at each element output). There are  $M$  elements in the array. As an example, consider sixteen samples (i.e.,  $L = 16$ ) of a signal received by a four-element array with an element SNR of 10 dB. This yields a single-signal accuracy of about 0.02 beamwidths with perfect calibration. As indicated in Figure A, this level of performance is also achieved with perfect sampling (i.e., an arbitrarily large number of samples), given a calibration error

variance of  $-22$  dB. Thus, with even a modest number of samples, calibration errors can dominate accuracy even at small signal levels. Note also that the accuracy is a small fraction of a beamwidth, given practical levels ( $-22$  dB, corresponding to 0.5 dB and 3 degrees) of antenna calibration. The calibration levels required for good estimation accuracy with one signal can be inadequate for sub-beamwidth resolution with multiple signals. See Reference 1 for a treatment of multiple-signal direction finding in the presence of calibration errors.

### References

1. A. Kuruc, "Lower Bounds on Multiple-Source Direction Finding in the Presence of Direction-Dependent Antenna-Array-Calibration Errors," *Technical Report 799*, Lincoln Laboratory (24 Oct. 1989), DTIC #ADA-215825.



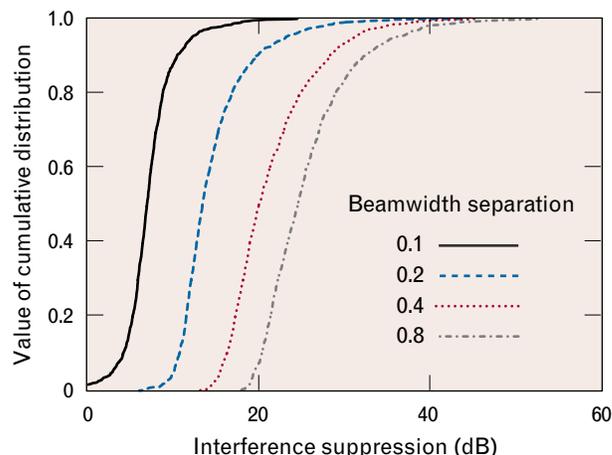
**FIGURE A.** Single-emitter angle estimation. The accuracy of angle estimates depends both on the quality of antenna calibration and on the level of the received signal. A simple model of antenna calibration expresses uncertainties in element gain as an additive error, sized relative to the element's nominal gain. This model can be interpreted in terms of gain and phase errors, as shown in the plot on the left. The plot on the right shows (in the red trace) the resulting standard deviation of the angle estimates due to calibration errors (in the absence of cochannel interference) as a fraction of an array beamwidth when the integrated signal-to-noise ratio (SNR) is asymptotically large. For comparison, the blue trace shows the accuracy of angle estimates with perfect calibration for various integrated (over time and space) SNRs.

function of the interferer ASNR and the beamwidth separation between the signal and interferer. The ASINR achieved by an algorithm suffers an additional loss, since signal parameters have to be estimated from the data. For many of the maximum-likelihood algorithms already discussed, this additional loss is only a few dB for a wide range of scenarios. Thus the ideal is a reasonable goal for algorithm design.

Adaptive beamforming also suffers from antenna-calibration errors. Null-forming procedures such as those based on the deterministic signal model suffer from inadequate null depth due to such errors. Figure 2 shows the performance of null-forming weights in the presence of independent element errors (see the sidebar entitled “Single-Emitter Angle Estimation”) in the estimates of the signal array responses. For various separations of the signal of interest and the interferer, the curves indicate the cumulative distribution functions of the achieved null depth. For example, with array-response errors of 0.08 beamwidths and a signal separation of 0.2 beamwidths (see the sidebar entitled “Definition and Properties of Beamwidth”), the median null depth is about 12 dB.

It is well known that calibration errors can also lead to poor performance of SINR-maximizing beamformers because of self-nulling [12]. Even perfect angle estimates do not prevent self-nulling. Figure 3 shows the achieved output SINR for selected values of the cumulative distribution function, given a perfect model of the array-response vector and an estimated covariance (using the sample covariance matrix as the estimate) based on one hundred samples. Even with a strong signal of interest, the median (0.5) output SINR is no more than about 10 dB.

Several remedies improve performance substantially. For null-forming approaches, we can place additional nulls in an attempt to broaden the nulls in the directions of interference. For SINR-maximizing techniques, one remedy [13] replaces  $\hat{\mathbf{R}}$  with a diagonally loaded version:  $\hat{\mathbf{R}} + \alpha \mathbf{I}_M$ . Varying the loading factor  $\alpha$  trades null depth for less self-nulling. Of course, there must be some scheme for choosing  $\alpha$ . Alternatively, we can replace the weight  $\mathbf{w}$  with the projection of  $\mathbf{w}$  into the signal subspace of the data [14] (the signal subspace is the span of the  $S$  largest



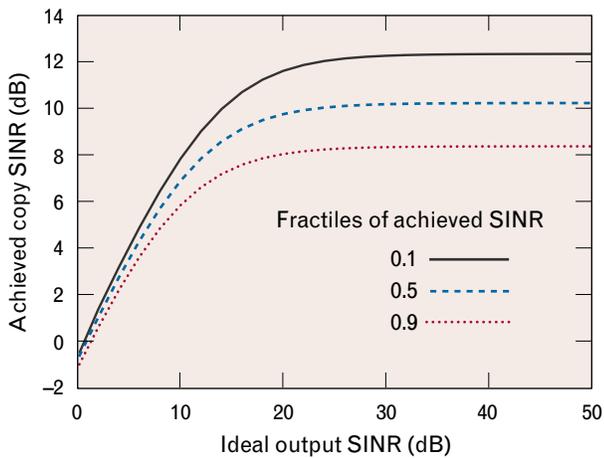
**FIGURE 2.** Interference suppression using null-forming weights. One variant of adaptive beamforming places nulls in the direction of the interferers. Errors in the estimated array response of the interferers result in misplaced nulls that reduce the achieved null depth. For estimation errors of 0.08 beamwidths, the cumulative distribution functions of the achieved null depths are shown for four separations of the signal of interest and the interferers.

eigenvectors of the sample covariance matrix  $\hat{\mathbf{R}}$ ). This remedy improves SINR substantially for the weaker signals.

Although the remedies mentioned above can improve performance substantially for practical antenna arrays, none offers the same level of beamforming performance that can be attained by exploiting signal waveforms.

### Waveform-Based Adaptive-Array Processing

Narrowband signals can be placed in a variety of categories for the purpose of waveform feature exploitation. These categories emphasize certain waveform features that can be used to form data-adaptive beams. Which feature is chosen depends on the waveform, the performance required, and the difficulty of implementation. Generally speaking, most waveforms fit into several categories, and some categories contain most waveforms. However, performance can vary substantially for the same waveform when different features are exploited. In many cases, we assume antenna-calibration information is unavailable for adaptive beamforming. Once signal parameters have been estimated, these estimates can be combined with calibration data to perform direction finding.



**FIGURE 3.** Statistics of output SINR due to finite sampling. Even if there is no separation, or mismatch, between the true array response and the estimated array response, the effects of finite sampling result in self-nulling. The SINR achieved at the output of a beamformer is always less than the SINR at the output of the ideal beamformer. With one hundred data samples and no mismatch, the 0.1, 0.5, and 0.9 fractiles of achieved SINR are shown as a function of the SINR at the output of the ideal beamformer.

The sections below entitled “Modeled Signals” and “Statistical Signals” discuss two approaches to waveform-based array processing. The signals can have parametric models based on a class of modulations, or they can be statistical in nature. In general, algorithms addressing both modeled and statistical signals play a role in array processing. Effective procedures can often be broken into two phases. The first phase provides an initial level of performance—possibly far from optimal—that is then improved in a second phase involving iterative refinement. Parametric models are important in the refinement phase, while statistical signal properties are adequate for initialization. Furthermore, the algorithms discussed in the section on statistical signals systematically handle all signals in the environment, while most of the modeled-signal approaches require some form of initial parameter estimates.

### Modeled Signals

We can treat modeled waveforms with maximum-likelihood techniques to yield a large number of effective waveform exploitation procedures. Typical exploited features include spectral lines, pulse leading

and trailing edges, constant envelopes, known data (training sequences), and modulation types (e.g., quadrature amplitude modulation). The maximum-likelihood techniques often provide near-optimal copy beams with small amounts of data (one hundred to one thousand samples). Furthermore, we do not have to use all known waveform information in order to approach optimal copy. For example, the constant-envelope property alone is sufficient for constant-envelope waveforms and provides some level of performance (although not optimal) for other types of waveforms.

The next section presents a concise description of an approach to model-based data-adaptive copy. This approach differs from well-known least-squares procedures described later in that the spatial covariance of the background noise (which may include undetected signals) is not assumed to be known (and hence is estimated as part of the maximum-likelihood procedure). This approach leads to a robust set of procedures that incorporate some form of data whitening, which enhances convergence rates.

*Blind Adaptation.* The signal model expressed by Equation 2 can be generalized to provide the basis for a more robust set of procedures for array processing. Consider the probability density function

$$p_{\text{BA}}(\mathbf{Z}|\mathbf{X}, \mathbf{A}, \mathbf{R}) = \pi^{-ML} |\mathbf{R}|^{-L} e^{-\text{tr}[(\mathbf{Z}-\mathbf{X}\mathbf{A})^H \mathbf{R}^{-1}(\mathbf{Z}-\mathbf{X}\mathbf{A})]}$$

Signals are modeled spatially in terms of the array-response vectors forming the columns of  $\mathbf{X}$  or in terms of the spatial distribution of power expressed by the covariance  $\mathbf{R}$ . Signals modeled through  $\mathbf{R}$  are treated as Gaussian sources. Signals modeled in  $\mathbf{X}$  have in addition amplitudes  $\mathbf{A}$  that are treated parametrically. Prior knowledge of these amplitudes is used to aid in their estimation. It is useful to think of the signals as represented by the parametric models given by  $\mathbf{X}$  and  $\mathbf{A}$ , while  $\mathbf{R}$  is used to represent an unknown colored (spatially) noise background that may itself be composed of additional signals or other types of noise (e.g., environmental or thermal). Although  $\mathbf{R}$  is estimated implicitly below, the estimates of  $\mathbf{X}$  and  $\mathbf{A}$  are the focus of the signal processing.

The most robust adaptive beamformers do not uti-

lize antenna calibration, even if this calibration is available. In the notation introduced above, this lack of calibration amounts to assuming that the array-response vectors represented by the columns of  $\mathbf{X}$  are unknown.

Maximizing  $p_{\text{BA}}(\mathbf{Z}|\mathbf{X}, \mathbf{A}, \mathbf{R})$  over the spatial covariance  $\mathbf{R}$  yields (see Reference 15, or Appendix A in Reference 16)

$$\left(\frac{\pi e}{L}\right)^{-LM} \left| (\mathbf{Z} - \mathbf{X}\mathbf{A})(\mathbf{Z} - \mathbf{X}\mathbf{A})^H \right|^{-L}.$$

Completing the square,

$$\begin{aligned} (\mathbf{Z} - \mathbf{X}\mathbf{A})(\mathbf{Z} - \mathbf{X}\mathbf{A})^H &= \\ [\mathbf{X} - \mathbf{Z}\mathbf{A}^H(\mathbf{A}\mathbf{A}^H)^{-1}](\mathbf{A}\mathbf{A}^H) & \\ \cdot [\mathbf{X} - \mathbf{Z}\mathbf{A}^H(\mathbf{A}\mathbf{A}^H)^{-1}]^H + \mathbf{Z}(\mathbf{I}_L - \mathbf{P}_\mathbf{A})\mathbf{Z}^H, & \end{aligned}$$

where  $\mathbf{P}_\mathbf{Q}$ , for any full-rank matrix  $\mathbf{Q}$  with at least as many columns as rows, is the projector onto the row space of  $\mathbf{Q}$ . Thus

$$\mathbf{P}_\mathbf{Q} = \mathbf{Q}^H(\mathbf{Q}\mathbf{Q}^H)^{-1}\mathbf{Q},$$

acting from the right on row vectors.

If  $\mathbf{X}$  can be chosen freely (in other words, if we choose to ignore calibration), its maximum-likelihood estimate is

$$\mathbf{X} = \mathbf{Z}\mathbf{A}^H(\mathbf{A}\mathbf{A}^H)^{-1}.$$

This is a least-squares estimate of  $\mathbf{X}$  based on the signal amplitude estimates provided by  $\mathbf{A}$  (see the section entitled “Least-Squares Procedures”). However, the estimate of  $\mathbf{A}$  is not based on the least-squares estimate obtained by fixing  $\mathbf{X}$ . If we choose  $\mathbf{X}$  as above, the maximum-likelihood decision for  $\mathbf{A}$  amounts to finding

$$\min_{\mathbf{A}} \left| \mathbf{Z}(\mathbf{I}_L - \mathbf{P}_\mathbf{A})\mathbf{Z}^H \right|.$$

The Frobenius relations for partitioned matrices can be used to show that

$$\left| \mathbf{I} - \mathbf{B}\mathbf{A} \right| = \begin{vmatrix} \mathbf{I} & \mathbf{A} \\ \mathbf{B} & \mathbf{I} \end{vmatrix} = \left| \mathbf{I} - \mathbf{A}\mathbf{B} \right|.$$

If we use this relation, then the expression

$$\begin{aligned} & \left| \mathbf{Z}(\mathbf{I}_L - \mathbf{P}_\mathbf{A})\mathbf{Z}^H \right| \\ &= \left| \mathbf{Z}\mathbf{Z}^H \right| \left| \mathbf{I}_N - (\mathbf{Z}\mathbf{Z}^H)^{-1}\mathbf{Z}\mathbf{P}_\mathbf{A}\mathbf{Z}^H \right| \\ &= \left| \mathbf{Z}\mathbf{Z}^H \right| \left| \mathbf{I}_L - \mathbf{P}_\mathbf{Z}\mathbf{P}_\mathbf{A} \right| \\ &= \left| \mathbf{Z}\mathbf{Z}^H \right| \left| \mathbf{I}_M - \mathbf{A}\mathbf{P}_\mathbf{Z}\mathbf{A}^H(\mathbf{A}\mathbf{A}^H)^{-1} \right| \quad (4) \\ &= \left| \mathbf{Z}\mathbf{Z}^H \right| \frac{\left| \mathbf{A}(\mathbf{I}_L - \mathbf{P}_\mathbf{Z})\mathbf{A}^H \right|}{\left| \mathbf{A}\mathbf{A}^H \right|} \end{aligned}$$

provides several equivalent maximum-likelihood statistics for parameter estimation.

When the signal amplitude matrix  $\mathbf{A}$  models a single signal ( $S = 1$ ), the maximum-likelihood parameter estimates are provided by

$$\arg \max_{\mathbf{A}} \frac{\mathbf{A}\mathbf{P}_\mathbf{Z}\mathbf{A}^H}{\mathbf{A}\mathbf{A}^H}. \quad (5)$$

Note that this is a constrained maximization, where the constraints involve modulation features. Even though this special case models only one signal, any interferers are treated as part of the unknown spatial covariance  $\mathbf{R}$ ; hence spatial nulling is performed.

There is a useful alternative formulation of Equation 5. By defining ( $\mathbf{W}$  is  $M \times 1$ )

$$\mu(\mathbf{W}, \mathbf{A}) = \frac{\left| \mathbf{W}^H \mathbf{Z}\mathbf{A}^H \right|^2}{(\mathbf{W}^H \mathbf{Z}\mathbf{Z}^H \mathbf{W})(\mathbf{A}\mathbf{A}^H)},$$

we have

$$\sup_{\mathbf{W}} \mu(\mathbf{W}, \mathbf{A}) = \frac{\mathbf{A}\mathbf{P}_\mathbf{Z}\mathbf{A}^H}{\mathbf{A}\mathbf{A}^H}.$$

The maximum is achieved with weights that are proportional to

$$\mathbf{W}_{\text{max}} = (\mathbf{Z}\mathbf{Z}^H)^{-1}\mathbf{Z}\mathbf{A}^H.$$

Applying the weight  $\mathbf{W}$  to the output of the array results in the time series

$$\mathbf{Y} = (y_1, \dots, y_L) = \mathbf{W}^H \mathbf{Z}.$$

The dependence of  $\mathbf{Y}$  on the weight  $\mathbf{W}$  has been suppressed from the notation. We can write

$$\mu \equiv \mu(\mathbf{W}, \mathbf{P}) = \frac{|\mathbf{Y}\mathbf{A}^H|^2}{\|\mathbf{A}\|^2 \|\mathbf{Y}\|^2}.$$

Thus

$$\begin{aligned} \left| \|\mathbf{A}\| - \frac{\mathbf{Y}\mathbf{A}^H}{\|\mathbf{A}\|} \right|^2 + \|\mathbf{Y}\|^2 - \frac{|\mathbf{Y}\mathbf{A}^H|^2}{\|\mathbf{A}\|^2} &= \|\mathbf{Y} - \mathbf{A}\|^2 \\ &= \left| \|\mathbf{Y}\| - \frac{\mathbf{Y}\mathbf{A}^H}{\|\mathbf{Y}\|} \right|^2 + \|\mathbf{A}\|^2 - \frac{|\mathbf{Y}\mathbf{A}^H|^2}{\|\mathbf{Y}\|^2}, \end{aligned}$$

so that

$$\begin{aligned} \frac{\|\mathbf{Y} - \mathbf{A}\|^2}{\|\mathbf{A}\|^2} &\geq 1 - \frac{|\mathbf{Y}\mathbf{A}^H|^2}{\|\mathbf{Y}\|^2 \|\mathbf{A}\|^2} \\ &= 1 - \mu \leq \frac{\|\mathbf{Y} - \mathbf{A}\|^2}{\|\mathbf{Y}\|^2}, \end{aligned} \quad (6)$$

with either equality (but not both) achievable by scaling  $\mathbf{W}$  appropriately.

An algorithm for finding the maximum-likelihood estimate of  $\mathbf{A}$  (and for finding  $\mathbf{W}$ ) can be based on minimizing either the left or right side of Equation 6. Consider, for example, minimizing the right side by alternately minimizing over the two variables  $\mathbf{Y}$  (really  $\mathbf{W}$ ) and  $\mathbf{A}$ . With initial values for  $\mathbf{Y}$  and  $\mathbf{A}$ , the (repeated) update steps are as follows:

1.  $\mathbf{A}_{\min} = \arg \min_{\mathbf{A}} \|\mathbf{Y} - \mathbf{A}\|^2$
2.  $\mathbf{W}_{\min} = \arg \min_{\mathbf{W}} \frac{\|\mathbf{Y} - \mathbf{A}_{\min}\|^2}{\|\mathbf{Y}\|^2}$
3.  $\mathbf{Y} = \mathbf{W}_{\min}^H \mathbf{Z}.$

We can perform the minimization in step two explicitly. Let  $\mathbf{P}_Z^\perp$  represent the projection onto the complement of the row space of  $\mathbf{Z}$ . We write

$$\begin{aligned} \frac{\|\mathbf{Y} - \mathbf{A}\|^2}{\|\mathbf{Y}\|^2} &= \frac{\|\mathbf{A}\mathbf{P}_Z^\perp\|^2 + \|\mathbf{Y} - \mathbf{A}\mathbf{P}_Z\|^2}{\|\mathbf{Y}\|^2} \\ &\geq \frac{\|\mathbf{A}\mathbf{P}_Z^\perp\|^2 + \|\mathbf{Y}\|^2 + \|\mathbf{A}\mathbf{P}_Z\|^2 - 2\|\mathbf{Y}\|\|\mathbf{A}\mathbf{P}_Z\|}{\|\mathbf{Y}\|^2} \end{aligned}$$

with equality when  $\mathbf{Y} \propto \mathbf{A}\mathbf{P}_Z$  with a real, positive proportionality factor. This is possible since  $\mathbf{Y}$  and  $\mathbf{A}\mathbf{P}_Z$  both lie in the row space of  $\mathbf{Z}$ . Let

$$\mathbf{W} = (\mathbf{Z}\mathbf{Z}^H)^{-1} \mathbf{Z}\mathbf{A}^H.$$

Note that

$$\mathbf{A}\mathbf{P}_Z = \mathbf{A}\mathbf{Z}^H (\mathbf{Z}\mathbf{Z}^H)^{-1} \mathbf{Z} = \mathbf{W}^H \mathbf{Z}.$$

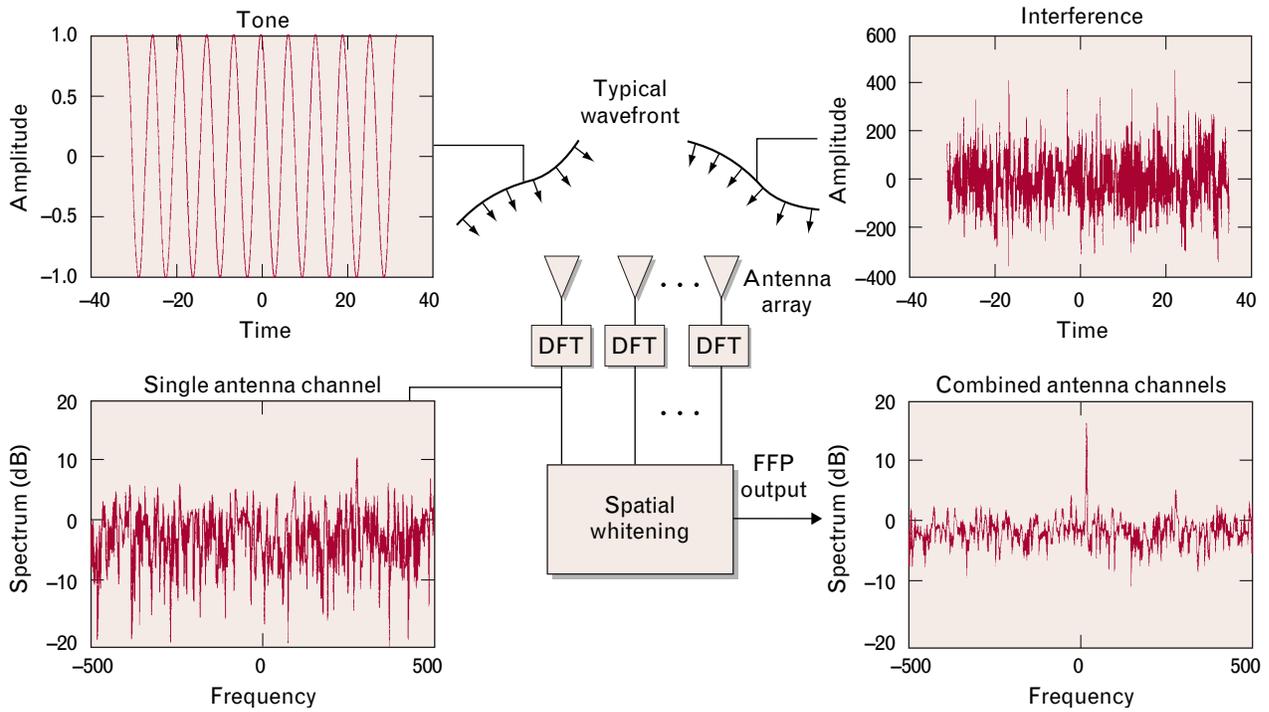
Furthermore, any  $\mathbf{W}$  that satisfies  $\mathbf{W}^H \mathbf{Z} \propto \mathbf{A}\mathbf{P}_Z$  is unique, up to a complex scalar, as long as the rows of  $\mathbf{Z}$  are independent, which holds with probability one whenever  $L \geq M$ . Thus  $\mathbf{W}_{\min}$  equals  $\mathbf{W}$  up to a scale factor that is easily evaluated, so that

$$\mathbf{W}_{\min} = \frac{\|\mathbf{A}\|^2}{\|\mathbf{A}\mathbf{P}_Z\|^2} (\mathbf{Z}\mathbf{Z}^H)^{-1} \mathbf{Z}\mathbf{A}^H.$$

In summary, an iterative solution for  $\mathbf{A}$  can be based on the steps

$$\begin{aligned} \mathbf{A}_{\min} &= \arg \min_{\mathbf{A}} \|\mathbf{Y} - \mathbf{A}\|^2 \\ \mathbf{W}_{\min} &= \frac{\|\mathbf{A}\|^2}{\|\mathbf{A}\mathbf{P}_Z\|^2} (\mathbf{Z}\mathbf{Z}^H)^{-1} \mathbf{Z}\mathbf{A}_{\min}^H \\ \mathbf{Y} &= \mathbf{W}_{\min}^H \mathbf{Z}. \end{aligned}$$

These three steps are repeated until a convergence condition (not discussed here) is satisfied. The first step can be regarded as demodulation. The second step is the decision-directed component of the algorithm. It feeds back demodulation decisions to update the copy weight  $\mathbf{W}$ . Note that the updated copy weight is proportional to the product of an inverse



**FIGURE 4.** Simulation of the detection of a tone, or spectral line, in the presence of a broadband interferer. Spectral lines in the waveform of the signal of interest, which are caused by residual carrier or data periodicities, can be detected by forming a discrete Fourier transform (DFT) of the data. Algorithms such as frequency feature processing (FFP) combine the coherent integration provided by a DFT with the spatial nulling provided by an antenna array to enhance substantially the detectability of the signal of interest.

sample covariance matrix  $(L^{-1}\mathbf{Z}\mathbf{Z}^H)^{-1}$  and an array-response estimate  $\mathbf{Z}\mathbf{A}_{\min}^H$ , which is formed by matched-filtering the antenna-element outputs with the estimated waveform. The third step applies the updated copy weight  $\mathbf{W}$  to the data  $\mathbf{Z}$ .

For a multisignal model, the maximum-likelihood statistic given in Equation 4 can be broken down into an iterative application of the one-signal procedure. This process is not discussed further here. It is worth mentioning that the multisignal approach combines spatial filtering (i.e., beamforming) with temporal filtering to enhance copy performance in the sense of providing better estimates of waveform parameters.

*Spectral Lines.* For spectral lines, which are informally called tones, we have

$$\mathbf{A} = (e^{i\theta}, e^{i2\theta}, \dots, e^{iL\theta}).$$

In this case we can rewrite Equation 5 in a more convenient form as

$$\max_k \mathbf{Z}^H(k) \hat{\mathbf{R}}^{-1} \mathbf{Z}(k), \quad (7)$$

where

$$\mathbf{Z} = [\mathbf{Z}(1), \dots, \mathbf{Z}(L)] = \mathbf{Z}\mathbf{U}_{DFT}$$

is the frequency-domain data ( $\mathbf{U}_{DFT}$  is the  $L \times L$  unitary matrix that performs row-wise discrete Fourier transforms, or DFTs) and the maximum-likelihood statistic looks for tones in spatially whitened DFT bins.

Figure 4 illustrates the detection of a tone in the presence of broadband noise. As shown in this example, a DFT at the output of a single channel does not provide enough integration to provide detection. In practice, signal duration or perhaps phase noise on the spectral line can limit integration times. Equation 7 provides the detection statistic shown in the bottom right panel of the figure. The tone is detected easily.

*Constant-Envelope Waveforms.* For constant-envelope waveforms,

$$\mathbf{A} = (e^{i\theta_1}, \dots, e^{i\theta_L}).$$

Demodulation amounts to estimating the phases  $\{\theta_k\}$ .

We observe that

$$\arg \min_{\mathbf{A}} \|\mathbf{Y} - \mathbf{A}\|^2 = (e^{i \arg y_1}, \dots, e^{i \arg y_L}).$$

Substituting the corresponding value of  $\mathbf{A}$  into the equivalent statistic in Equation 6 yields

$$\frac{\|\mathbf{Y} - \mathbf{A}\|^2}{\|\mathbf{A}\|^2} = L^{-1} \sum_{k=1}^L (|y_k| - 1)^2.$$

This is one member of a family of statistics suggested for constant-modulus signals in the literature. These statistics require the minimization of

$$\sum_k (|y_k|^p - 1)^q$$

for some  $p$  and  $q$ . The values  $p = 1$ ,  $q = 2$  provide a maximum-likelihood approach to adaptive beamforming for constant-modulus signals.

A variant of the constant-modulus technique, organized in a manner that provides beamforming for multiple cochannel constant-modulus signals, is called waveform improved nulling (WIN). The performance of WIN is discussed in a later section.

#### Maximum Likelihood Using Antenna Calibration.

The most general data-adaptive procedures use calibration data as well as waveform features to perform direction finding (see Reference 16 for a more general formulation). Such approaches can be more difficult to implement than some of the techniques discussed above, and may also suffer from sensitivity to calibration errors. Maximum-likelihood procedures are simplified by assuming that antenna calibration cannot be used. Waveform features alone can be used to provide good blind beamforming, which then can be used to aid direction finding (see the section entitled "Copy-Based Direction Finding").

In the applications below, the "component"  $Z_q$  of the data array  $\mathbf{Z}$  introduced in the appendix entitled "Maximum-Likelihood Parameter Estimation" is a block of samples, but the treatment in the appendix is much more general. With an  $M \times J$  matrix  $\mathbf{X}$ , a  $J \times S$  matrix  $\mathbf{A}$ , and an  $S \times L$  matrix  $\mathbf{T}$ , and the definition  $\mathbf{S} = \mathbf{Z}_q \mathbf{Z}_q^H$ , the notation introduced in the appendix allows us to write the maximum-likelihood solution (actually the generalized likely ratio test, or GLRT)

for Gaussian data with modeled mean  $E[\mathbf{Z}] = \mathbf{XAT}$  and covariance  $\mathbf{R}$  as

$$\begin{aligned} & \left( \frac{|\mathbf{X}^H \mathbf{S}^{-1} \mathbf{X}|}{|\mathbf{X}^H (\mathbf{Z} \mathbf{Z}^H)^{-1} \mathbf{X}|} \right)^L \\ &= \frac{\max_{\mathbf{R}, \mathbf{A}} \pi^{-ML} |\mathbf{R}|^{-L} e^{-\text{tr}[\mathbf{R}^{-1}(\mathbf{Z} - \mathbf{XAT})(\mathbf{Z} - \mathbf{XAT})^H]}}{\max_{\mathbf{R}} \pi^{-ML} |\mathbf{R}|^{-L} e^{-\text{tr}(\mathbf{R}^{-1} \mathbf{Z} \mathbf{Z}^H)}}, \end{aligned} \quad (8)$$

with the maximum-likelihood estimate of  $\mathbf{A}$  given by

$$(\mathbf{X}^H \mathbf{S}^{-1} \mathbf{X})^{-1} \mathbf{X}^H \mathbf{S}^{-1} \mathbf{Z} \mathbf{T}^H (\mathbf{T} \mathbf{T}^H)^{-1}. \quad (9)$$

Some special cases illustrate the scope of this machinery. These applications are, in most instances, fairly direct consequences of the expression above. No attempt is made to exhaust all possible applications. All of the applications discussed below assume  $J = 1$ ; in other words, there is only one signal of interest. This assumption doesn't mean that only one signal of interest can be detected, copied, or located. It means only that the signals are found one at a time, since the maximum-likelihood statistic is based on a single signal (but arbitrary Gaussian interference) model.

*Pulsed Signals.* The special case in which  $J = 1$  and the  $S \times L$  matrix  $\mathbf{T} = (\mathbf{I}_S \mathbf{0})$  corresponds to a pulsed emitter. In fact,  $\mathbf{A}$  is the row vector  $(a_1, \dots, a_S)$  in this case, and hence

$$\mathbf{AT} = (a_1, a_2, \dots, a_M, 0, \dots, 0),$$

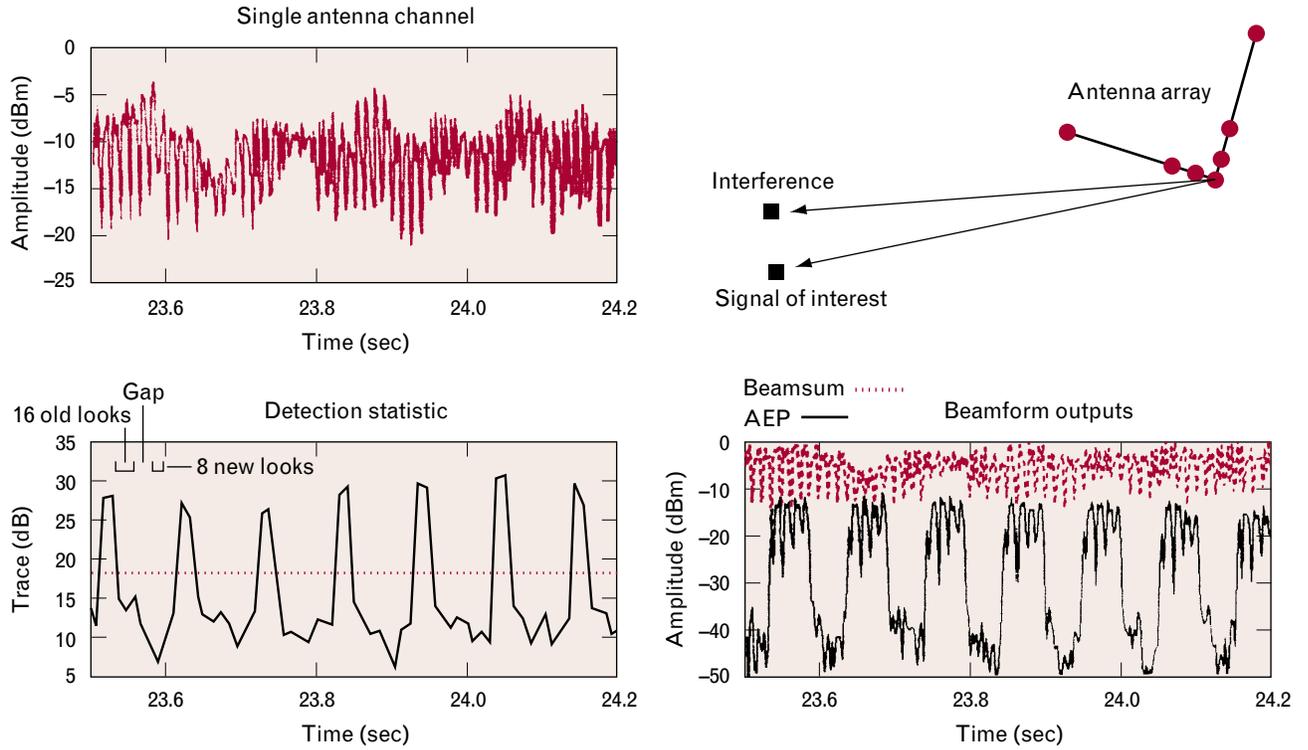
which corresponds to a signal that is "on" only during the first  $S$  samples. For the particular value of  $\mathbf{T}$ ,

$$\mathbf{Z} = (\mathbf{Z}_p \mathbf{Z}_q) \equiv (\mathbf{Z}_{\text{new}} \mathbf{Z}_{\text{old}})$$

expresses a partitioning of the observations  $\mathbf{Z}$  into what are called old samples, which don't contain the signal of interest, and new samples, which do. The GLRT, which is equivalent to the maximum-likelihood statistic, becomes

$$\max_u \frac{L}{L - S} \frac{\mathbf{X}^H(u) \hat{\mathbf{R}}_{\text{old}}^{-1} \mathbf{X}(u)}{\mathbf{X}^H(u) \hat{\mathbf{R}}^{-1} \mathbf{X}(u)}, \quad (10)$$

where



**FIGURE 5.** Exploitation of pulsed waveforms. Adaptive beamforming can be used to detect and copy intermittent signals buried in strong cochannel interference. A detection statistic and an adapted beam can be formed by comparing samples taken before (old looks) and after (new looks) a pulsed signal's leading edge. An example based on measured high-frequency data shows the ability of adaptive event processing (AEP) to detect and copy pulsed waveforms in strong cochannel interference.

$$\hat{\mathbf{R}}_{\text{old}} = \frac{1}{L - S} \mathbf{S}$$

$$\hat{\mathbf{R}}_{\text{new}} = \frac{1}{S} \mathbf{Z}_p \mathbf{Z}_p^H$$

$$\hat{\mathbf{R}} = \frac{1}{L} \mathbf{Z} \mathbf{Z}^H = \frac{L - S}{L} \hat{\mathbf{R}}_{\text{old}} + \frac{S}{L} \hat{\mathbf{R}}_{\text{new}},$$

and where  $\mathbf{X}(u)$  is the array-response vector ( $J = 1$ ) parameterized by the angle of arrival  $u$  (which can stand for both angular components). This expression provides a direction-finding statistic that can be used for pulsed waveforms.

When beamforming is desired, it is often better if we assume that the array-response vector  $\mathbf{X}$  is completely unknown; otherwise, calibration errors can result in a suboptimal beamformer, especially if the calibration errors are large. In this case the estimate of  $\mathbf{A}$  is given by

$$\mathbf{A} = (a_1, \dots, a_S) = \mathbf{W}^H \mathbf{Z}_{\text{new}},$$

with

$$\mathbf{W} = e_{\max}(\hat{\mathbf{R}}_{\text{old}}^{-1} \hat{\mathbf{R}}_{\text{new}}),$$

where  $e_{\max}(\cdot)$  denotes the principal eigenvector (and  $\lambda_{\max}(\cdot)$  denotes the corresponding eigenvalue) of its argument.  $\mathbf{W}$  is determined by Equation 9 upon solving for the (unconstrained)  $\mathbf{X}$  that maximizes Equation 10. We base detection on the GLRT equivalent (i.e., monotonically related) statistic

$$\lambda_{\max}(\hat{\mathbf{R}}_{\text{old}}^{-1} \hat{\mathbf{R}}_{\text{new}}),$$

or, more commonly, on a suboptimal approximation

$$\text{tr}[\hat{\mathbf{R}}_{\text{old}}^{-1} \hat{\mathbf{R}}_{\text{new}}].$$

Figure 5 shows the copy and detection of pulsed signals received at HF by a large-aperture ground-based array. The array geometry is L-shaped, as indicated in the figure. Detection occurs on pulse leading edges; a new copy weight (beamformer) is determined at each leading edge and provides the beamformer outputs shown.

*Dual Polarization.* Detection and direction finding can also be accomplished with polarized signals. We consider the case of pulsed waveforms as an example. Let  $\mathcal{X}(u)$  denote an  $M \times 2$  matrix parameterized by  $u$  whose columns span the two-dimensional subspace representing the array responses spanned by all incident polarizations with angle of arrival  $u$ . The matrix  $\mathcal{X}$  is a result of calibrating a diversely polarized array. Any incident polarization can be written as  $\mathcal{X}(u)\mathbf{q}$ , where  $\mathbf{q}$  is an unknown two-vector describing the polarization components of the signal in the basis provided by the columns of  $\mathcal{X}(u)$ . From above, the direction finding (or detection) statistic becomes

$$\begin{aligned} & \max_{\mathbf{q}} \frac{\mathbf{q}^H [\mathcal{X}^H(u) \hat{\mathbf{R}}_{\text{old}}^{-1} \mathcal{X}(u)] \mathbf{q}}{\mathbf{q}^H [\mathcal{X}^H(u) \hat{\mathbf{R}}^{-1} \mathcal{X}(u)] \mathbf{q}} \\ & = \lambda_{\max} \left\{ [\mathcal{X}^H(u) \hat{\mathbf{R}}^{-1} \mathcal{X}(u)]^{-1} [\mathcal{X}^H(u) \hat{\mathbf{R}}_{\text{old}}^{-1} \mathcal{X}(u)] \right\}. \end{aligned}$$

This expression is most useful as a direction-finding statistic, since the computation involved in using it for detection is excessive, given the potential performance improvement.

### Statistical Signals

An alternative to model-based exploitation involves statistical properties of the waveforms. For example, differences between the spectra of the signal of interest and the spectra of the interference can be leveraged to form adaptive beams. In the extreme case, when the spectra are disjoint, beamforming is easy. However, even when the spectra of the interference and the signal of interest are similar, some degree of adaptive beamforming is possible.

Another statistical approach uses higher-order moments to exploit the non-Gaussian character of typical signals. A number of procedures are available. Some fourth-order cumulant techniques offer the ability to form beams for multiple, statistically identical non-Gaussian emitters.

*Spectral Differences.* One class of algorithms exploits spectral differences between the signal of interest and the interference. The leverage is based on the constancy of the signals' array response (including the interferers' array response) over frequency (the narrowband assumption) in conjunction with the varia-

tion of the signals' power over the same band. If the signals are not spectrally identical over the averaging time of the algorithm, then some degree of signal separation is possible. Spectral differences are revealed by cutting the band into frequency cells and forming array covariances in each subband. The discussion below formulates the problem in more detail by using the notation and notions previously established. We have simplified the problem formulation to improve the exposition.

Recall that the  $M \times L$  data matrix  $\mathbf{Z}$  has mean

$$\mathbf{E}[\mathbf{Z}] = \mathbf{X}\mathbf{A}.$$

We assume  $M = S$  so that  $\mathbf{X}$  is  $S \times S$ . In other words, there are as many signals as sensors. When  $S < M$ , it is possible to reduce the data  $\mathbf{Z}$  to a signal subspace and exercise the algorithm there; this reduction is not discussed further. We assume that the covariance of any column of  $\mathbf{Z}$  can be written as  $\mathbf{I}_M$ . Let  $\mathbf{F}_k$  be an  $l \times L$  matrix with orthonormal rows. Filtering the data  $\mathbf{Z}$  in "cell" or "subband"  $k$  can be represented by  $\mathbf{Z}\mathbf{F}_k^H$ . This  $S \times l$  array can be interpreted as  $l$  outputs from  $S$  sensors in the  $k$ th subband. If we define  $\mathbf{A}_k = \mathbf{A}\mathbf{F}_k^H$ , then the mean of  $\mathbf{Z}\mathbf{F}_k^H$  is given by  $\mathbf{X}\mathbf{A}_k$  and the covariance of any column of  $\mathbf{Z}\mathbf{F}_k^H$  is  $\mathbf{I}_M$  (see Equation A in the appendix entitled "Maximum-Likelihood Parameter Estimation").

Up to now,  $\mathbf{A}$  (along with  $\mathbf{A}_k$ ) has been treated as a parameter matrix representing the signal amplitudes. For the remainder of this section, we consider  $\mathbf{A}$  (along with  $\mathbf{A}_k$ ) to be random with zero mean. Let  $\mathbf{P}_k = L^{-1}\mathbf{E}[\mathbf{A}_k\mathbf{A}_k^H]$  and  $\mathbf{P} = L^{-1}\mathbf{E}[\mathbf{A}\mathbf{A}^H]$ . We assume  $\mathbf{P}$  and  $\mathbf{P}_k$  are diagonal. In other words, the signals and their spectral components are uncorrelated. We define

$$\begin{aligned} \mathbf{R} &= L^{-1}\mathbf{E}[\mathbf{Z}\mathbf{Z}^H] = \mathbf{X}\mathbf{P}\mathbf{X}^H + \mathbf{I}_M \\ \mathbf{R}_k &= l^{-1}\mathbf{E}[\mathbf{Z}\mathbf{F}_k^H\mathbf{F}_k\mathbf{Z}^H] = \mathbf{X}\mathbf{P}_k\mathbf{X}^H + \mathbf{I}_M \\ \mathcal{R}_k &= (\mathbf{R} - \mathbf{I}_M)^{-1/2}\mathbf{R}_k(\mathbf{R} - \mathbf{I}_M)^{-1/2} \\ \mathbf{Q} &= (\mathbf{R} - \mathbf{I}_M)^{-1/2}\mathbf{I}_M(\mathbf{R} - \mathbf{I}_M)^{-1/2}. \end{aligned}$$

The expression for  $\mathbf{R}_k$  can be verified with the help of Appendix A. Then (recall that  $\mathbf{P}$  and  $\mathbf{P}_k$  are diagonal)

$$\mathcal{R}_k - \mathbf{Q} = \mathbf{U}(\mathbf{P}_k\mathbf{P}^{-1})\mathbf{U}^H$$

with common unitary matrix

$$\begin{aligned} \mathbf{U} &= (\mathbf{R} - \mathbf{I}_M)^{-1/2} \mathbf{X} \mathbf{P}^{1/2} \\ &= (\mathbf{X} \mathbf{P} \mathbf{X}^H)^{-1/2} \mathbf{X} \mathbf{P}^{1/2}. \end{aligned} \quad (11)$$

Since  $\mathbf{P}_k \mathbf{P}_k^{-1}$  is a diagonal matrix for each  $k$ , the above indicates that the  $\mathcal{R}_k - Q$  matrices are simultaneously diagonalizable by the unitary matrix  $\mathbf{U}$ . There are a variety of procedures for performing the simultaneous diagonalization; this step in the process is not considered further. Once  $\mathbf{U}$  is obtained, we can estimate  $\mathbf{X}$  by using Equation 11. We do not have to know  $\mathbf{P}$ , because it is diagonal and hence affects the estimate of  $\mathbf{X}$  by harmlessly scaling its columns. Direction finding and beamforming can be based on the estimate of  $\mathbf{X}$ .

An important special case occurs when  $k = 1$  or  $2$ . For the  $k = 2$  subband case, we redefine  $\mathbf{R} = \mathbf{R}_2$  and  $\mathbf{P} = \mathbf{P}_2$ . The above argument is still valid. The common eigenvectors  $\omega$  of  $\mathcal{R}_1 - Q$  and  $\mathcal{R}_2 - Q$  solve

$$(\mathcal{R}_1 - Q)\omega = \lambda(\mathcal{R}_2 - Q)\omega,$$

which is equivalent to the generalized eigenvalue problem

$$(\mathbf{R}_1 - \mathbf{I}_M)\mathbf{w} = \lambda(\mathbf{R}_2 - \mathbf{I}_M)\mathbf{w}.$$

The columns of  $\mathbf{X}$  become the eigenvectors of

$$(\mathbf{R}_1 - \mathbf{I}_M)(\mathbf{R}_2 - \mathbf{I}_M)^{-1},$$

which is often called a change matrix. Note that the eigenvalues of the change matrix are the diagonal entries of  $\mathbf{P}_1 \mathbf{P}_2^{-1}$ . If these entries are not distinct, we cannot separate signals in this fashion.

The beamforming achieved by using  $\mathbf{X}$  is occasionally close to optimal but not nearly as good as the near-optimal performance of maximum-likelihood techniques. However, in the strong signal case (i.e.,  $\mathbf{I}_M = 0$ ), a Cramér-Rao bound can be formulated that is often approached by the above technique.

Clearly, the level of performance that we can achieve depends significantly on the way the band is cut into spectral pieces, or subbands, to reveal spectral differences between the signals. In many cases, two subbands, as just discussed, offer an interesting level of performance, but more subbands have the potential to increase performance substantially.

Although spectral difference techniques have been discussed in terms of cutting the band into subbands, other time-frequency slices may be appropriate. For example, wavelet transforms can be used. In Figure 6, a broadband, spectrally flat signal and a frequency-shift-keyed signal are separated by exploiting differences in the waveform spectra, as discussed above.

*Cumulant Eigenanalysis.* Higher-order moment techniques can be used to provide coarse initial beamformers that are refined by the maximum-likelihood procedures described above. This section presents a brief introduction to cumulant eigenanalysis (CUE), which was introduced by J.F. Cardoso [17]. Cumulants are defined traditionally in terms of the moment generating function. Let  $\mathbf{z}$  be a random complex vector of length  $M$  and  $\mathbf{w}$  a fixed complex vector. The moment generating function of

$$\mathbf{z} = (z_1, \dots, z_M)^T$$

is given by

$$\phi_{\mathbf{z}}(\mathbf{w}) = E[e^{2\Re(\mathbf{w}^H \mathbf{z})}],$$

where  $\Re(\cdot)$  indicates the real part of the argument (when there is no confusion, the subscript  $\mathbf{z}$  on  $\phi$  is suppressed). An important property of moment generating functions is their factorization with independent random variables. Let  $\mathbf{x}$  and  $\mathbf{y}$  be independent complex random vectors with  $\mathbf{z} = \mathbf{x} + \mathbf{y}$ . Then

$$\phi_{\mathbf{z}}(\mathbf{w}) = \phi_{\mathbf{x}}(\mathbf{w}) \phi_{\mathbf{y}}(\mathbf{w}),$$

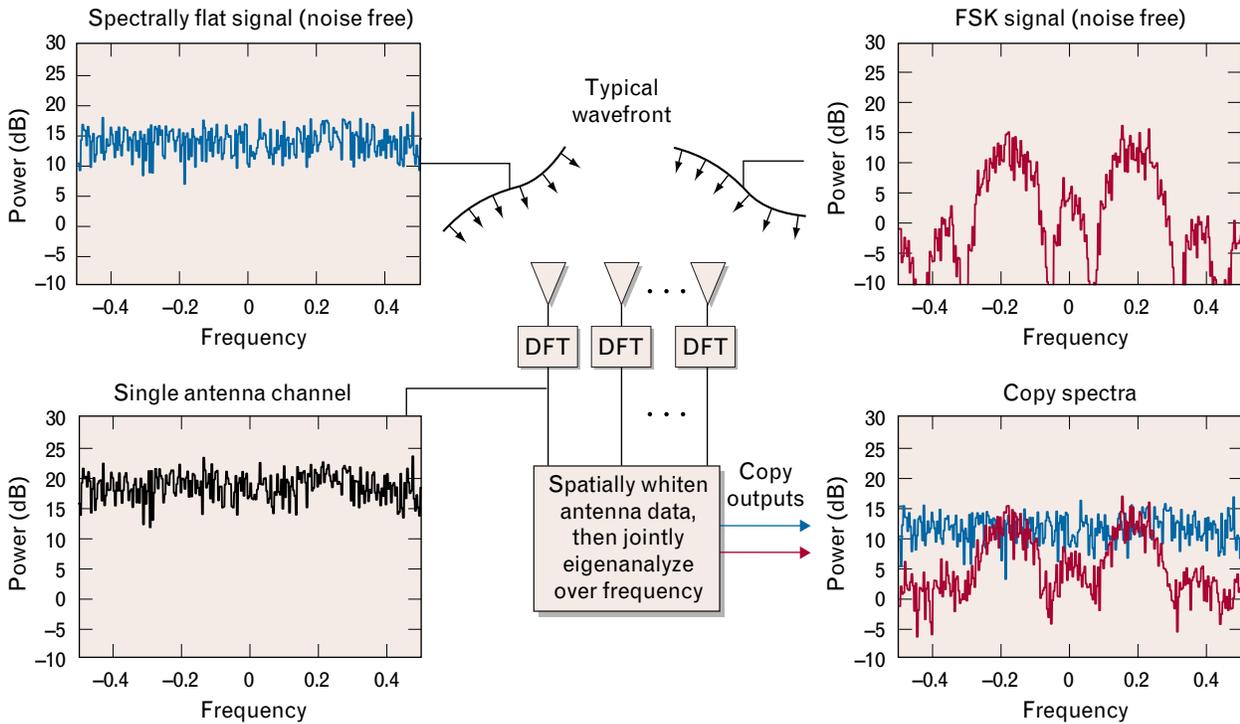
at least in a formal sense. Moments (really mixed moments) are defined by differentiating the moment-generating function. We define the complex differential operators

$$\frac{\partial}{\partial w} = \frac{1}{2} \left( \frac{\partial}{\partial u} - i \frac{\partial}{\partial v} \right) \text{ and } \frac{\partial}{\partial \bar{w}} = \frac{1}{2} \left( \frac{\partial}{\partial u} + i \frac{\partial}{\partial v} \right),$$

where  $w = u + iv$  is the expression of  $w$  in terms of real and imaginary parts. We can view  $w$  as a component of  $\mathbf{w} = (w_1, \dots, w_M)^T$ . The fourth-order moment tensor of  $\mathbf{z}$  has components

$$M_{ijkl} = \left. \frac{\partial^4 \phi(\mathbf{w})}{\partial \bar{w}_i \partial w_j \partial w_k \partial \bar{w}_l} \right|_{\mathbf{w}=0} = E[z_i \bar{z}_j \bar{z}_k z_l],$$

where  $E[\cdot]$  denotes expectation. Note that this defini-



**FIGURE 6.** Spectral difference processing (SDP). In some cases, cochannel signals can exhibit statistical differences that are observed in the spectral domain. These differences may be large, as shown in the top part of the figure for two signals of different types of modulation (spectrally flat and frequency-shift-keyed, or FSK), or the difference may be small (for signals with the same nominal modulation). The differing spectra facilitate the formation of adaptive beams for each signal. SDP, as with most of the techniques discussed in this article, does not rely on antenna calibration for beamforming.

tion makes sense as long as the mixed fourth-order moments exist.

Cumulants are defined by performing a similar differentiation on the log of the moment generating function. Thus the components of the second-order cumulant of a mean-zero random vector  $\mathbf{z}$  are

$$C_{ij}^{(2)} = \left. \frac{\partial^2 \ln \phi(\mathbf{w})}{\partial \bar{w}_i \partial w_j} \right|_{\mathbf{w}=0} = E[z_i \bar{z}_j].$$

This expression is simply the covariance matrix (and second-order moment) of the random vector  $\mathbf{z}$ . The superscript indicates the order of the cumulant. The fourth-order cumulant is expressed by

$$\begin{aligned} C_{ijkl}^{(4)} &= \left. \frac{\partial^4 \ln \phi(\mathbf{w})}{\partial \bar{w}_i \partial w_j \partial w_k \partial \bar{w}_l} \right|_{\mathbf{w}=0} \\ &= E[z_i \bar{z}_j \bar{z}_k z_l] - E[z_i \bar{z}_j] E[\bar{z}_k z_l] \\ &\quad - E[z_i \bar{z}_k] E[\bar{z}_j z_l] - E[z_i z_l] E[\bar{z}_j \bar{z}_k]. \end{aligned}$$

We can express the factoring of the moment generating function with independent summands  $\mathbf{z} = \mathbf{x} + \mathbf{y}$  as

$$\ln \phi_{\mathbf{x}}(\mathbf{w}) + \ln \phi_{\mathbf{y}}(\mathbf{w}) = \ln \phi_{\mathbf{z}}(\mathbf{w}),$$

which suggests that the cumulant ( $C = C^{(4)}$  here and in the following) is additive. In other words,

$$C(\mathbf{x} + \mathbf{y}) = C(\mathbf{x}) + C(\mathbf{y}),$$

with independent mean-zero summands  $\mathbf{x}$  and  $\mathbf{y}$ . This result can be shown directly from the definition and requires only the existence of all moments up to the fourth order. The notation expresses the cumulant as a tensor that is a function of the complex-valued random vector  $\mathbf{z}$ .  $C$  is additive, and  $C(\mathbf{z}) = 0$  for mean-zero complex Gaussian random vectors. The second relation is a direct consequence of the well-known complex Gaussian moment factoring theorem [18].

The analysis in the appendix entitled ‘‘Cumulant Eigenanalysis’’ demonstrates that the whitened data  $(\mathbf{R} - \mathbf{I}_M)^{-1/2} \mathbf{Z}$  can be used to estimate the matrix

$$(\mathbf{R} - \mathbf{I}_M)^{-1/2} \mathbf{X} \mathbf{P}^{1/2}$$

when  $\mathbf{P}$  is diagonal. We can use this matrix in turn to estimate the columns of  $\mathbf{X}$  up to complex scale factors, thus providing blind array-response estimates. The performance of CUE is discussed below.

### Copy-Based Direction Finding

Our initial discussion of direction finding in the section entitled “Calibration-Based Adaptive-Array Processing” revolved around techniques that do not utilize waveform features. In contrast, the discussion of direction finding in the previous section considered some techniques that utilize waveform features for joint beamforming and direction finding. An alternative approach to direction finding is based on forming adaptive beams first, without the use of array calibration, and then incorporating array calibration to perform direction finding. Such direction-finding procedures are called copy-based procedures. Waveform features can be used to form adaptive beams in the manner discussed in the previous section. In many cases, direction-finding performance can be enhanced significantly over that achieved by the calibration-based techniques discussed earlier by using the estimates of signal amplitudes provided by the kind of adaptive beamforming treated above. Some examples of copy-based direction finding are described below.

#### Least-Squares Procedures

The least-squares estimates of the array response  $\mathbf{X}$  and signal amplitudes  $\mathbf{A}$  are given by

$$\begin{aligned} \hat{\mathbf{A}} &= (\mathbf{X}^H \mathbf{R}^{-1} \mathbf{X})^{-1} \mathbf{X}^H \mathbf{R}^{-1} \mathbf{Z} \\ \hat{\mathbf{X}} &= \mathbf{Z} \mathbf{A}^H (\mathbf{A} \mathbf{A}^H)^{-1} \end{aligned}$$

(for the derivation of these expressions, see the appendix entitled “Least-Squares Procedures”). Note that the estimator  $\hat{\mathbf{X}}$  is independent of the spatial covariance  $\mathbf{R}$ ; the performance of the estimator depends strongly on  $\mathbf{R}$ , however, as seen below. When the maximum-likelihood techniques we discussed in the section entitled “Modeled Signals” are used to estimate  $\mathbf{A}$ , the estimate of  $\mathbf{X}$  is expressed by  $\hat{\mathbf{X}}$  above. But even more generally, if  $\mathbf{A}$  can be estimated by any procedure, we can form the least-squares estimate of

$\mathbf{X}$  given above. We can base practical direction-finding techniques on an estimate of a signal’s array response. Once the estimate is made, we can look it up in a calibration table by, for example, minimizing the angle between the estimated array-response vector and the calibration-table entries.

#### Copy Correlation

An alternative to the joint least-squares procedure just discussed relies on correlating the beamformer output for a particular signal with the output of a conventional adaptive beamformer of the form

$$\mathbf{W}(u) = \hat{\mathbf{R}}^{-1} \mathbf{X}(u).$$

Here  $\mathbf{X}(u)$  represents the modeled array response of a particular signal with an angle of arrival  $u$  and

$$\hat{\mathbf{R}} = \mathbf{L}^{-1} \mathbf{Z} \mathbf{Z}^H.$$

The motivation for this approach is that  $\mathbf{W}(u)$ , for the correct angle of arrival  $u$ , should be a good beamformer and hence its output should be a good approximation to the time series of the copied signal.

Let  $\mathbf{W}$  denote a column vector expressing a copy weight obtained, for example, from one of the copy algorithms discussed above. The correlation coefficient of the copy outputs from the two beamformers becomes

$$\begin{aligned} & \frac{|\mathbf{W}^H(u) \mathbf{Z}| (\mathbf{W}^H \mathbf{Z})^H|^2}{(\mathbf{W}^H \mathbf{Z} \mathbf{Z}^H \mathbf{W}) [\mathbf{X}(u)^H \hat{\mathbf{R}}^{-1} \mathbf{Z} \mathbf{Z}^H \hat{\mathbf{R}}^{-1} \mathbf{X}(u)]} \\ &= \frac{|\mathbf{W}^H \mathbf{X}(u)|^2}{(\mathbf{W}^H \hat{\mathbf{R}} \mathbf{W}) [\mathbf{X}(u)^H \hat{\mathbf{R}}^{-1} \mathbf{X}(u)]}. \end{aligned}$$

This statistic has the form of a product of a Capon spectral estimator [19] and the gain of the copy beamformer.

#### Advantage of Joint Direction Finding

To understand the benefits of waveform exploitation, we ask how well can the array response of a signal be estimated, given perfect knowledge of the waveform. This ideal example is obviously extreme, but it illustrates the power of feature-based direction finding.

We recall (in the appendix “Maximum-Likelihood

Parameter Estimation”) that the covariance  $\text{cov}(\mathbf{Z})$  of the data matrix is given by  $\mathbf{R} \otimes \bar{\Delta}$ . We can show (see Equation A in the appendix entitled “Maximum-Likelihood Parameter Estimation”) that

$$\text{cov}(\mathbf{AZB}) = (\mathbf{ARA}^H) \otimes \overline{(\mathbf{B}^H \Delta \mathbf{B})}$$

for arbitrary  $\mathbf{A}$  and  $\mathbf{B}$ . If  $\Delta = \mathbf{I}_L$ , then, since  $E[\mathbf{Z}] = \mathbf{X}\mathbf{A}$ , the least-squares estimate  $\hat{\mathbf{X}}$  of  $\mathbf{X}$  has the statistic

$$E[\hat{\mathbf{X}}] = \mathbf{X}$$

$$\text{cov}(\hat{\mathbf{X}}) = L^{-1} \mathbf{R} \otimes \bar{\mathbf{P}}^{-1},$$

where  $\mathbf{P} = \mathbf{A}\mathbf{A}^H/L$ . Thus the covariance of the estimate of the  $k$ th signal’s array response is given by  $L^{-1} \mathbf{P}^{kk} \mathbf{R}$ , where the superscript indicates an entry of the inverse matrix.

The “angle” between two array responses  $\mathbf{q}$  and  $\mathbf{r}$  is

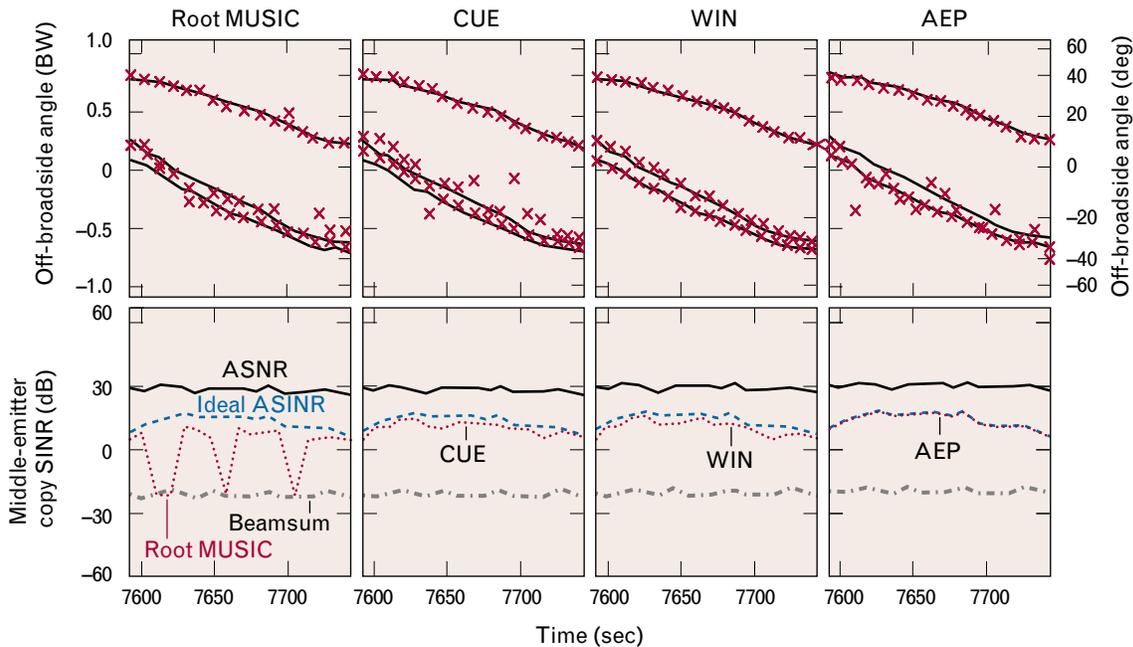
$b$  beamwidths (see the sidebar entitled “Definition and Properties of Beamwidths”), where

$$\frac{|\mathbf{q}^H \mathbf{r}|}{\|\mathbf{q}\| \|\mathbf{r}\|} = \cos\left(\frac{\pi b}{2}\right).$$

When  $b = 0$ , the array responses coincide; when  $b = 1$ , the array responses are orthogonal. For small estimation errors (a large number of samples), the mean-squared error (measured in square beamwidths) for estimating the array response of the  $k$ th signal (which is denoted  $\mathbf{X}(k)$ , the  $k$ th column of  $\mathbf{X}$ ; we also assume  $\|\mathbf{X}(k)\| = 1$ ) becomes approximately

$$(2/\pi)^2 L^{-1} \mathbf{P}^{kk} [\text{tr}(\mathbf{R}) - \mathbf{X}(k)^H \mathbf{R} \mathbf{X}(k)].$$

Two special cases are of interest. If  $\mathbf{R} = \mathbf{I}_M$  and  $\mathbf{P}$  is diagonal, the root-mean-squared (RMS) error in



**FIGURE 7.** Copy and direction-finding performance for four different signal processing algorithms. Root MUSIC, an established technique for direction finding, can also be used to provide beamforming (by a method mentioned briefly in the section entitled “Adaptive Beamforming”). The traditional beamsum performance is shown here as a reference. The other algorithms—cumulant eigenanalysis (CUE), waveform-improved nulling (WIN), and adaptive event processing (AEP)—are based on waveform features and are used primarily for adaptive beamforming. These three techniques can also provide direction finding by methods discussed in the text. The top row of plots shows the accuracy of line-of-bearing estimates for three cochannel signals observed by a test aircraft over several minutes. The red x-markers show the angle estimates and the black lines show the actual line of bearing for the three signals in this experiment. The bottom row of plots shows the SINR at the output of a beamformer pointed at the middle signal. The CUE and WIN techniques are significantly better at copy than the beamsum and root MUSIC techniques, and AEP is nearly indistinguishable from the ideal SINR.

beamwidths becomes

$$(2/\pi)\sqrt{\frac{N-1}{LP_{kk}}},$$

where  $P_{kk}$  denotes the average power of the  $k$ th signal. The case  $\mathbf{R} \neq \mathbf{I}_M$  corresponds to least-squares estimation of some, but not all, cochannel signals. For example, if  $\mathbf{R} = \mathbf{I}_M + p_I \mathbf{V}\mathbf{V}^H$  for a unit vector  $\mathbf{V}$ , which models a single interferer with ASNR  $p_I$ , then the RMS error becomes

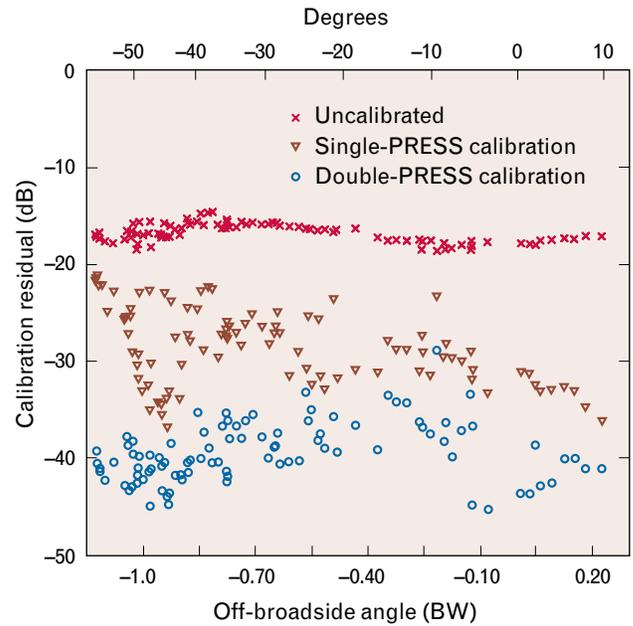
$$(2/\pi)\sqrt{\frac{N-1 + p_I \sin^2(\pi b/2)}{LP_{kk}}},$$

where  $b$  is the beamwidth separation of the signal and interferer. The RMS error is large if the interferer is much stronger than the signal. In effect, the estimation errors in the direction of the interferer's array response have a large variance. However, if the waveform of the interferer is exploited in addition to that of the signal of interest (i.e., the interferer is incorporated in the least-squares estimation), then the estimation error does not depend on the emitter separations and interference power. Direction-finding estimates for the two emitters are decoupled.

### Performance Examples

Figure 7 shows beamforming and direction-finding results for four different algorithms with measured data. The data and experiment design are discussed elsewhere [4]; here we limit the discussions to connections with the algorithms introduced earlier.

The data shown come from narrowband FM waveforms collected at VHF aboard an airborne platform; all processing was done off line. The environment consists of three cochannel signals of varying powers and geometries. Four antenna elements are used in a nonlinear array. Time along the flight path is indicated in the  $x$ -axes of the subplots. The marker symbols indicate direction-finding estimates and the solid lines indicate the corresponding true off-broadside angles of arrival, measured in degrees and in beamwidths. In this figure, the signal from the middle emitter is 20 dB below the signals from the outer emitters. The signal separations range from almost one beamwidth to less than a tenth of a beamwidth.

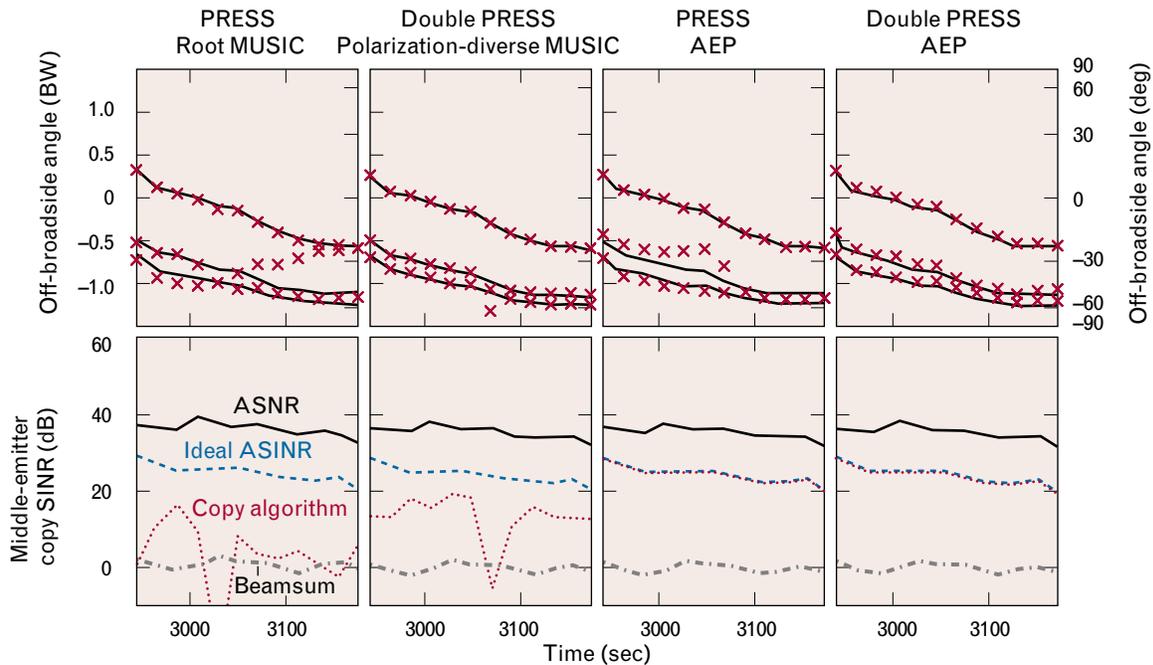


**FIGURE 8.** Calibration residuals. Good direction-finding performance requires accurate models of antenna patterns. The uncertainty of antenna-array calibration is measured by residuals that express the difference between measured and modeled array responses. Calibration based on vertical polarization models (pattern response equalization for spatial similarity, or PRESS) of the array response are used for the results shown in Figure 7. Although antenna elements and signal sources nominally have a vertical polarization response, the measured data are explained much more accurately, and with substantially lower residuals, by using dual-polarization models (double PRESS) for the array response.

Also shown in the figure is the SINR measured at the output of the copy beamformer for the middle signal. This SINR is compared with several other kinds of SINR that were introduced earlier.

The bottom row of plots in Figure 7 shows the ASNR of the middle signal as a function of time along the flight path. Also shown is the ideal ASINR and the ASINR achieved by the copy algorithm. For reference, the performance of a beamformer pointed at the middle signal is shown; since the middle signal is about 20 dB below the other signals, the beamformer output SINR is correspondingly low. Note that the ideal ASINR curve and the ASNR curve remain unchanged on all subplots.

The first column of panels in Figure 7 shows the performance of root MUSIC. This direction-finding algorithm has good accuracy and resolution as well as



**FIGURE 9.** Copy and direction-finding performance with improved antenna calibration. The improvements in antenna calibration provided by dual-polarization models (double PRESS) of array response leads to improvements in the accuracy of direction finding. Shown here is the direction-finding and beamforming performance for root MUSIC, polarization-diverse MUSIC, and AEP (with and without polarization diversity), using PRESS and double PRESS for array calibration. Double-PRESS calibration yields significant improvements in direction-finding performance in both cases, as shown in the top row of plots. The beamformer built from root MUSIC also shows improvement with double-PRESS calibration, as shown in the bottom row of plots. AEP beamforming remains unchanged because it does not utilize antenna calibration.

robustness to calibration errors. Direction-finding performance is good until the separation between signals is small, near the end of the flight path. Copy is based on the direction-finding estimates provided by root MUSIC. Typically, copy is about 5 dB worse than ideal unless the direction-finding estimates are poor, leading to dropouts. The remaining algorithms provide copy much closer to ideal. This is especially true for waveform improved nulling (WIN), a constant-envelope technique based on the blind-adaptive procedures discussed in the earlier section on modeled signals, and adaptive event processing (AEP), a maximum-likelihood technique for pulsed signals. Cumulant eigenanalysis (CUE) is a higher-order moment procedure discussed earlier. All three algorithms provide blind (i.e., without any calibration) copy.

Direction-finding estimates for two of the algorithms (CUE and WIN) are based on the copied data. AEP uses the maximum-likelihood statistics discussed previously to estimate directions for each emit-

ter. The outer emitters are located accurately, but resolving the middle signal is difficult. CUE uses the copy correlation procedure discussed earlier to perform direction finding. The performance on the middle signal appears slightly better with CUE, but the direction finding on the outer signals is worse. WIN uses a joint least-squares approach similar (but not identical) to the least-squares approach discussed earlier (see also the appendix entitled “Least-Squares Procedures”). In this case, all three signals are resolved and have good direction-finding estimates. The outer signals, which are about 20 dB stronger than the middle signal, must be copied well, since the least-squares procedure attempts to subtract them from the data in order to perform direction finding on the middle signal.

#### Antenna Calibration

All of the direction-finding techniques just discussed have to deal with imperfect antenna calibration.

Nominally, the signals used in the experiments have vertical transmit polarization. However, near-field reflections can rotate the received polarizations on the airframe, especially as the orientation of the airframe changes. If we take the polarization dependence of the antenna response into account, we can substantially reduce calibration residuals (the errors between modeled array response and true array response). Figure 8 shows calibration residuals as a function of off-broadside angle for a six-element linear array. The x-markers indicate the residuals present in the uncorrected patterns; the triangular markers show on average a 14-dB reduction in the residuals after calibration. This level of calibration is used for the results shown above in Figure 7.

By taking source polarization into account, we can reduce the calibration residuals another 11 dB on average, which leads to an improvement in direction-finding performance, as indicated in Figure 9. The figure shows before-and-after plots of copy and direction-finding performance for MUSIC and AEP for a slightly different set of data than that shown in Figure 7. Clearly, the improved calibration has a substantial impact on direction-finding performance. Both MUSIC and AEP resolve the signals and perform direction finding with good accuracy.

## Conclusions

Waveform features can be exploited to perform adaptive beamforming and direction finding for narrowband signals in cochannel interference. For many emitters, we can use a few generic signal features such as constant envelopes, periodicities, or leading-and-trailing edges of signal bursts to perform near-optimal adaptive beamforming without using any array calibration data. The beamformer outputs can be used to aid direction finding in several ways. The resulting copy-based direction-finding estimates can be significantly better than those provided by conventional high-resolution estimators such as MUSIC. It is still important, however, to achieve the best possible level of calibration in order to maximize direction-finding performance. Calibration for both senses of polarization can provide significant improvements in calibration accuracy that result in substantially better direction finding.

## REFERENCES

1. R. Schmidt, "Multiple Emitter Location and Signal Parameter Estimation," *Proc. RADC Spectrum Estimation Workshop, Rome, N. Y., 3-5 Oct. 1979*, pp. 243-258.
2. R.O. Schmidt, "A Signal Subspace Approach to Multiple Emitter Location and Spectral Estimation," Ph.D. thesis, Stanford University, Nov. 1981.
3. D.N. Simkins, "Multichannel Angle-of-Arrival Estimation," Ph.D. thesis, Stanford University, Nov. 1980.
4. L.L. Horowitz, "Airborne Signal Intercept for Wide-Area Battlefield Surveillance," in this issue.
5. B. Ottersten, M. Viberg, and T. Kailath, "Analysis of Subspace Fitting and ML Techniques for Parameter Estimation from Sensor Array Data," *IEEE Trans. Signal Process.* **40** (3), 1992, pp. 590-600.
6. J.E. Evans, J.R. Johnson, and D.F. Sun, "Application of Advanced Signal Processing Techniques to Angle of Arrival Estimation in ATC Navigation and Surveillance Systems," *Technical Report 582*, Lincoln Laboratory (23 June 1982), DTIC #ADA-118306.
7. A.J. Barabell, "Improving the Resolution Performance of Eigenstructure-Based Direction-Finding Algorithms," *ICASSP Proc.* **1**, Boston, 14-16 Apr. 1983, pp. 336-339.
8. G.F. Hatke and K.W. Forsythe, "A Class of Polynomial Rooting Algorithms for Joint Azimuth/Elevation Estimation Using Multidimensional Arrays," *28th Asilomar Conf. on Signals, Systems & Computers 1, Pacific Grove, Calif., 31 Oct.-2 Nov. 1994*, pp. 694-699.
9. Gary F. Hatke, "Superresolution Source Location with Planar Arrays," in this issue.
10. R. Roy, A. Paulraj, and T. Kailath, "ESPRIT—A Subspace Rotation Approach to Estimation of Parameters of Cisoids in Noise," *IEEE Trans. Acoust. Speech Signal Process.* **34** (5), 1986, pp. 1340-1342.
11. A. Kuruc, "Lower Bounds on Multiple-Source Direction Finding in the Presence of Direction-Dependent Antenna-Array-Calibration Errors," *Technical Report 799*, Lincoln Laboratory (24 Oct. 1989), DTIC #ADA-215825.
12. D.M. Boroson, "Sample Size Considerations for Adaptive Arrays," *IEEE Trans. Aerosp. Electron. Syst.* **16** (4), 1980, pp. 446-451.
13. W.D. White, "Artificial Noise in Adaptive Arrays," *IEEE Trans. Aerosp. Electron. Syst.* **14** (2), 1978, pp. 380-384.
14. R.O. Schmidt and R.E. Franks, "Multiple Source DF Signal Processing: An Experimental System," *IEEE Trans. Antennas Propag.* **34** (3), 1986, pp. 281-290.
15. N.R. Goodman, "Statistical Analysis Based on a Certain Multivariate Complex Gaussian Distribution," *Ann. Math. Stat.* **34** (1), 1963, pp. 152-177.
16. E.J. Kelley and K.W. Forsythe, "Adaptive Detection and Parameter Estimation for Multidimensional Signal Models," *Technical Report 848*, Lincoln Laboratory (19 Apr. 1989), DTIC #ADA-208971.
17. J.-F. Cardoso, "Blind Identification of Independent Components with Higher-Order Statistics," *IEEE Workshop on Higher-Order Spectral Analysis, Vail, Colo., 28-30 June 1989*, pp. 157-162.
18. I.S. Reed, "On a Moment Theorem for Complex Gaussian Processes," *IRE Trans. Info. Theory* **8** (3), 1962, pp. 194-195.
19. J. Capon, "High-Resolution Frequency-Wavenumber Spectrum Analysis," *Proc. IEEE* **57** (8), 1969, pp. 1408-1419.

## APPENDIX A: MAXIMUM-LIKELIHOOD PARAMETER ESTIMATION

LET  $\mathbf{Z}$  be an  $M \times L$  array of vector samples of the output of an  $M$ -element antenna array. Each column of  $\mathbf{Z}$  is a snapshot of the array output at a particular time. We assume below that these snapshots are independent, although the initial formulation is more general. The entries of  $\mathbf{Z}$  are taken to be jointly complex circular Gaussian with mean  $E[\mathbf{Z}] = \mathbf{XAT}$ .  $\mathbf{X}$  is an  $M \times J$  array whose columns represent the array responses of  $J$  emitters.  $\mathbf{T}$  is an  $S \times L$  array whose rows can be interpreted as time-domain-basis waveforms.  $\mathbf{A}$  is a  $J \times S$  array of signal amplitudes. In the case of interest here,  $S = 1$  and  $J = 1$  (the treatment is more general than the applications in the text). Then  $\mathbf{XAT}$  expresses the time history of a single coherent waveform. In the formulation of hypothesis-testing problems in Reference 16 and in the text, the  $\mathbf{A}$  array is assumed to be completely unknown. However, prior knowledge of  $\mathbf{X}$  and  $\mathbf{T}$  may vary considerably, leading to different forms of parameter estimators.

The covariance of  $\mathbf{Z}$  is given by

$$[\text{cov}(\mathbf{Z})]_{(i,j);(k,l)} = E[(Z_{ij} - \overline{Z_{ij}})(Z_{kl} - \overline{Z_{kl}})^*],$$

where the overbar designates expected value and the  $*$  denotes complex conjugate (see also the appendix entitled "Least-Squares Procedures"). It is often the case that the covariance has the special structure expressed by

$$\text{cov}(\mathbf{Z}) = \mathbf{R} \otimes \mathbf{C}^*,$$

where the Kronecker (tensor) product of matrices  $\mathbf{A}$  and  $\mathbf{B}$  is defined by

$$(\mathbf{A} \otimes \mathbf{B})_{(i,j);(k,l)} = A_{ik} B_{jl}.$$

The tensor-product structure of the covariance can be interpreted as a factoring into spatial ( $\mathbf{R}$ ) and temporal ( $\mathbf{C}$ ) factors. For the applications of interest here, we assume the factor  $\mathbf{C}$  is known. It can express, for example, the time-domain correlation imposed on a wideband waveform by the bandpass characteristic of a receiver.

Some additional notation and machinery is required. We consider data-independent matrices  $\mathbf{F}$  and  $\mathbf{D}$  that pre-multiply and post-multiply the data  $\mathbf{Z}$ . The matrix  $\mathbf{F}$  can represent spatial processing of the data, while the matrix  $\mathbf{D}$  can represent temporal processing. We can easily show that

$$E[\mathbf{FZD}] = \mathbf{FE}[\mathbf{Z}]\mathbf{D},$$

and that

$$\text{cov}[\mathbf{FZD}] = (\mathbf{FRF}^H) \otimes (\mathbf{D}^H \mathbf{C} \mathbf{D})^* \quad (\text{A})$$

(see Equation A1-44 in Reference 16). In the following, we assume that  $\mathbf{C} = \mathbf{I}_L$ , the  $L \times L$  identity matrix.

An orthonormal basis of the row space of  $\mathbf{T}$  can be chosen by defining

$$\mathbf{P} = (\mathbf{T}\mathbf{T}^H)^{-1/2}\mathbf{T}.$$

This is an  $S \times L$  array whose rows span the same subspace as the rows of  $\mathbf{T}$ . Furthermore, the rows of  $\mathbf{P}$  are orthonormal; that is,

$$\mathbf{P}\mathbf{P}^H = \mathbf{I}_M.$$

We choose a  $K \times L$  array  $\mathbf{Q}$  (where  $K = L - S$ ) whose rows are orthonormal and span the orthocomplement of the row space of  $\mathbf{P}$ . The  $\mathbf{P}$  and  $\mathbf{Q}$  arrays have the following properties:

$$\mathbf{I}_L = \mathbf{P}^H \mathbf{P} + \mathbf{Q}^H \mathbf{Q}$$

$$\mathbf{P}\mathbf{Q}^H = \mathbf{0}$$

$$\mathbf{P}\mathbf{P}^H = \mathbf{I}_S$$

$$\mathbf{Q}\mathbf{Q}^H = \mathbf{I}_K.$$

We post-multiply the data array  $\mathbf{Z}$  with a data-independent, unitary  $L \times L$  matrix  $\mathbf{U}_L^H$ , where

$$\mathbf{U}_L = \begin{pmatrix} \mathbf{P} \\ \mathbf{Q} \end{pmatrix}.$$

Then

$$\mathbf{Z}\mathbf{U}_L^H = (\mathbf{Z}\mathbf{P}^H \mathbf{Z}\mathbf{Q}^H) \equiv (\mathbf{Z}_p \mathbf{Z}_q)$$

is a left-right partition of the transformed data. Post-multiplication by  $\mathbf{U}_L^H$  constitutes time-domain processing of the data array  $\mathbf{Z}$  with a bank of orthogonal filters. The covariance machinery presented above tells us that

$$\text{cov}(\mathbf{Z}\mathbf{U}_L^H) = \mathbf{R} \otimes (\mathbf{U}_L \mathbf{I}_L \mathbf{U}_L^H)^* = \mathbf{R} \otimes \mathbf{I}_L.$$

In order to find the transformed mean of  $\mathbf{Z}$ , we write

$$\mathbf{X}\mathbf{A}\mathbf{T} = \mathbf{X}\mathbf{B}\mathbf{P},$$

where

$$\mathbf{B} = \mathbf{A}(\mathbf{T}\mathbf{T}^H)^{1/2}.$$

Then

$$\begin{aligned} \mathbb{E}[\mathbf{Z}_p] &= \mathbf{X}\mathbf{B}\mathbf{P}\mathbf{P}^H = \mathbf{X}\mathbf{B} \\ \mathbb{E}[\mathbf{Z}_q] &= \mathbf{X}\mathbf{B}\mathbf{P}\mathbf{Q}^H = \mathbf{0}. \end{aligned} \quad (\text{B})$$

Thus  $\mathbf{Z}_p$  and  $\mathbf{Z}_q$  are independent Gaussian arrays with independent columns and identical column (spatial) covariances. The  $\mathbf{Z}_q$  array has zero mean, while the  $\mathbf{Z}_p$  array has a nonzero mean containing the signal parameters. Although the coordinates just de-

scribed are based on the true signal parameters expressed by  $\mathbf{T}$ , these coordinates also make sense when  $\mathbf{T}$  is regarded parametrically, as is the case below. Of course, the signal means (Equation B) are no longer valid, but the covariance is unchanged.

Let  $\mathbf{S} = \mathbf{Z}_q\mathbf{Z}_q^H$  for use below. In addition, note that

$$\mathbf{Z}\mathbf{Z}^H = \mathbf{Z}_p\mathbf{Z}_p^H + \mathbf{S}.$$

From Reference 16 (see Equations 2-1, 2-2, 2-8, and 2-57), the generalized-likelihood-ratio test (GLRT), maximized over  $\mathbf{A}$ , can be written

$$\begin{aligned} & \left( \frac{|\mathbf{X}^H \mathbf{S}^{-1} \mathbf{X}|}{|\mathbf{X}^H (\mathbf{Z}\mathbf{Z}^H)^{-1} \mathbf{X}|} \right)^L \\ &= \frac{\max_{\mathbf{R}, \mathbf{A}} \pi^{-ML} |\mathbf{R}|^{-L} e^{-\text{tr}[\mathbf{R}^{-1}(\mathbf{Z}-\mathbf{X}\mathbf{A}\mathbf{T})(\mathbf{Z}-\mathbf{X}\mathbf{A}\mathbf{T})^H]}}{\max_{\mathbf{R}} \pi^{-ML} |\mathbf{R}|^{-L} e^{-\text{tr}(\mathbf{R}^{-1}\mathbf{Z}\mathbf{Z}^H)}}, \end{aligned}$$

with the maximum-likelihood estimate of  $\mathbf{A}$  given by

$$(\mathbf{X}^H \mathbf{S}^{-1} \mathbf{X})^{-1} \mathbf{X}^H \mathbf{S}^{-1} \mathbf{Z}\mathbf{T}^H (\mathbf{T}\mathbf{T}^H)^{-1}.$$

## APPENDIX B: CUMULANT EIGENANALYSIS

IN THE SPECIAL CASE that  $\mathbf{z} = a\mathbf{x}$  for a scalar  $a$  and a vector  $\mathbf{x} = (v_1, \dots, v_N)^T$ ,

$$C_{ijkl}(a\mathbf{x}) = C(a)v_i \bar{v}_j \bar{v}_k v_l,$$

where  $C(a)$  is the (scalar) fourth-order cumulant of the scalar  $a$ . When  $\mathbf{z}$  is composed of several independent signal components, the cumulant matrix is the sum of each signal's contribution.

Let  $\mathbf{X}$  denote the  $M \times S$  matrix of array responses as described in the main text of the article. Specifically, the  $c$ th column of  $\mathbf{X}$  is an  $M$ -vector that expresses the wavefront of the  $c$ th emitter. Let  $\mathbf{a} = (a_1, \dots, a_S)^T$  be a column vector of length  $S$  with random, mean-zero components  $a_c$  denoting the independent complex

amplitudes associated with each signal. Let  $\mathbf{b}$  be an  $M \times 1$  vector of independent additive mean-zero complex Gaussian noise. We assume

$$\mathbf{z} = \mathbf{X}\mathbf{a} + \mathbf{b}.$$

For use below, we define the tensor product of two vectors  $\mathbf{x}$  and  $\mathbf{y}$  as the vector whose  $\alpha$ th component is

$$(\mathbf{x} \otimes \mathbf{y})_\alpha = \mathbf{x}_i \mathbf{y}_j,$$

where  $\alpha = (ij)$ . The definition extends obviously for more than two factors; it is an "associative" product.

By using the additivity of the cumulant and the fact that the cumulant of a Gaussian vector is zero, we have

$$\begin{aligned} \mathbf{C}(\mathbf{z}) &= \sum_k \mathbf{C}(a_k)(\mathbf{X}_{\cdot k} \otimes \overline{\mathbf{X}_{\cdot k}} \otimes \overline{\mathbf{X}_{\cdot k}} \otimes \mathbf{X}_{\cdot k}) \\ &= \sum_k \mathbf{C}(a_k)(\mathbf{X}_{\cdot k} \otimes \overline{\mathbf{X}_{\cdot k}}) \otimes (\overline{\mathbf{X}_{\cdot k}} \otimes \mathbf{X}_{\cdot k}), \end{aligned}$$

where  $\mathbf{X}_{\cdot k}$  denotes the  $k$ th column of  $\mathbf{X}$ . Recall that the *Vec* of a rectangular array is the vector formed by stacking together in sequence the rows of the array. In particular, for column vectors  $\mathbf{x}$  and  $\mathbf{y}$ ,

$$\text{Vec}(\mathbf{x}\mathbf{y}^T) = \mathbf{x} \otimes \mathbf{y}.$$

Thus we can express the cumulant as

$$\mathbf{C}(\mathbf{z}) = \text{Vec} \left[ \sum_k \mathbf{C}(a_k)(\mathbf{X}_{\cdot k} \otimes \overline{\mathbf{X}_{\cdot k}})(\mathbf{X}_{\cdot k} \otimes \overline{\mathbf{X}_{\cdot k}})^H \right].$$

This notation merely formalizes the “matrix” viewpoint of  $\mathbf{C}$ . That is,  $\mathbf{C}$ , viewed as a matrix, has the structure shown within the *Vec*. In the particular case in which the columns of  $\mathbf{X}$  are orthonormal, this structure expresses an eigenanalysis of the “matrix”  $\mathbf{C}$  because

$$(\mathbf{X}_{\cdot j} \otimes \overline{\mathbf{X}_{\cdot j}})^H (\mathbf{X}_{\cdot k} \otimes \overline{\mathbf{X}_{\cdot k}}) = \delta_{jk}$$

since, for arbitrary vectors  $\mathbf{d}$ ,  $\mathbf{e}$ ,  $\mathbf{f}$ , and  $\mathbf{g}$ ,

$$(\mathbf{d} \otimes \mathbf{e})^H (\mathbf{f} \otimes \mathbf{g}) = (\mathbf{d}^H \mathbf{f})(\mathbf{e}^H \mathbf{g}).$$

Because the columns of  $\mathbf{X}$  are arbitrary array-response vectors, they are not orthonormal. However, a change of coordinates leads to an equivalent problem with orthonormal array-response vectors.

We assume  $\mathbf{P} = \text{E}[\mathbf{a}\mathbf{a}^H]$  is diagonal and  $S = M$ . Then whitening the data as described in the section of the article on spectral differences yields

$$(\mathbf{R} - \mathbf{I}_M)^{-1/2} \mathbf{z} = \mathbf{U}\mathbf{a}' + (\mathbf{R} - \mathbf{I}_M)^{-1/2} \mathbf{b},$$

where  $\mathbf{a}' = \mathbf{P}^{-1/2} \mathbf{a}$  are the normalized signal amplitudes and the matrix  $\mathbf{U} = (\mathbf{X}\mathbf{P}\mathbf{X}^H)^{-1/2} \mathbf{X}\mathbf{P}^{1/2}$  is unitary. Let  $\mathbf{a}' = (a'_1, \dots, a'_S)$ . The properties of  $\mathbf{C}$  mentioned above lead to

$$\begin{aligned} &\mathbf{C}[(\mathbf{R} - \mathbf{I}_M)^{-1/2} \mathbf{z}] \\ &= \text{Vec} \left[ \sum_k \mathbf{C}(a'_k)(\mathbf{U}_{\cdot k} \otimes \overline{\mathbf{U}_{\cdot k}})(\mathbf{U}_{\cdot k} \otimes \overline{\mathbf{U}_{\cdot k}})^H \right]. \end{aligned}$$

Because the columns of  $\mathbf{U}$  are orthonormal, eigenanalysis of  $\mathbf{C}$  yields the signal subspace (span of the top  $S$  eigenvectors)

$$\text{Span}\{(\mathbf{U}_{\cdot 1} \otimes \overline{\mathbf{U}_{\cdot 1}}), \dots, (\mathbf{U}_{\cdot S} \otimes \overline{\mathbf{U}_{\cdot S}})\}.$$

Note that  $\mathbf{C}$  is an  $S^2 \times S^2$  hermitian matrix but has only at most  $S$  nonzero eigenvalues. Any element of this subspace can be written in the form

$$\begin{aligned} &\sum_m \rho_m \mathbf{U}_{\cdot m} \otimes \overline{\mathbf{U}_{\cdot m}} \\ &= \text{Vec} \left[ \mathbf{U} \cdot \text{diag}(\rho_1, \dots, \rho_S) \cdot \mathbf{U}^H \right] \end{aligned} \quad (\text{A})$$

for, in general, arbitrary complex  $\rho_k$ . (The eigenspaces of  $\mathbf{C}$  have “real” structure, as can be shown from some symmetries of  $\mathbf{C}$ ; thus the  $\rho_k$  can be chosen as real.) Typically, the  $\rho_k$  are distinct. If this is the case, Equation A can be eigenanalyzed to recover the individual vectors  $\mathbf{U}_{\cdot s}$ . More generally, the matrices corresponding to the vectors given in Equation A can be diagonalized simultaneously by the unitary  $\mathbf{U}$ . Since

$$(\mathbf{R} - \mathbf{I}_M)^{1/2} \mathbf{U} = \mathbf{X}\mathbf{P}^{1/2},$$

Equation A determines the columns of  $\mathbf{X}$  up to unknown scale factors when  $\mathbf{P}$  is diagonal.

## APPENDIX C: LEAST-SQUARES PROCEDURES

For arbitrary matrices  $\mathbf{A}$  and  $\mathbf{B}$ , define the Kronecker product  $\mathbf{A} \otimes \mathbf{B}$  so that  $(\mathbf{A} \otimes \mathbf{B})_{(i,j)(k,l)} = A_{ik}B_{jl}$ . If  $\text{Vec}(\cdot)$  is defined as above (in the appendix entitled ‘‘Cumulant Eigenanalysis’’), unwrapping the rows of a matrix, then

$$(\mathbf{A} \otimes \mathbf{B})\text{Vec}(\mathbf{C}) = \text{Vec}(\mathbf{ACB}^T).$$

Recall that  $(\mathbf{A} \otimes \mathbf{B})^{-1} = \mathbf{A}^{-1} \otimes \mathbf{B}^{-1}$  and note that

$$\text{Vec}(\mathbf{A})^H \text{Vec}(\mathbf{B}) = \text{tr}(\mathbf{A}^H \mathbf{B}).$$

The covariance of an  $M \times L$  matrix  $\mathbf{Z}$  is defined as

$$\text{cov}(\mathbf{Z}) = \mathbb{E}[\text{Vec}(\mathbf{Z} - \mathbb{E}[\mathbf{Z}])\text{Vec}(\mathbf{Z} - \mathbb{E}[\mathbf{Z}])^H].$$

In some special cases, the covariance factors as a tensor product:  $\text{cov}(\mathbf{Z}) = \mathbf{R} \otimes \Delta^*$ . In particular, this factorization occurs with narrowband signals sampled at the Nyquist rate. In this case,  $\Delta = \mathbf{I}_L$ , expressing the fact that the columns of  $\mathbf{Z}$ , which correspond to snapshots of the array output, are uncorrelated. The matrix  $\mathbf{R}$  expresses the spatial correlation (between rows) in the data  $\mathbf{Z}$ .

If the data  $\mathbf{Z}$  has mean  $\mathbf{XA}$ , as above, then we can find (using the above relations) the least-squares estimate of the product  $\mathbf{XA}$  by solving

$$\begin{aligned} \min_{\mathbf{X}, \mathbf{A}} \text{Vec}(\mathbf{Z} - \mathbf{XA})^H \text{cov}(\mathbf{Z})^{-1} \text{Vec}(\mathbf{Z} - \mathbf{XA}) \\ = \min_{\mathbf{X}, \mathbf{A}} \text{tr}[(\mathbf{Z} - \mathbf{XA})^H \mathbf{R}^{-1}(\mathbf{Z} - \mathbf{XA})\Delta^{-1}]. \end{aligned}$$

By completing the square, we have

$$\begin{aligned} \text{tr}[(\mathbf{Z} - \mathbf{XA})^H \mathbf{R}^{-1}(\mathbf{Z} - \mathbf{XA})\Delta^{-1}] \\ = \text{tr}\{[(\mathbf{X} - \mathbf{Z}\Delta^{-1}\mathbf{A}^H(\mathbf{A}\Delta^{-1}\mathbf{A}^H)^{-1})(\mathbf{A}\Delta^{-1}\mathbf{A}^H) \\ \cdot (\mathbf{X} - \mathbf{Z}\Delta^{-1}\mathbf{A}^H(\mathbf{A}\Delta^{-1}\mathbf{A}^H)^{-1})^H + \mathbf{Z}\Delta^{-1}\mathbf{Z}^H \\ - \mathbf{Z}\Delta^{-1}\mathbf{A}^H(\mathbf{A}\Delta^{-1}\mathbf{A}^H)^{-1}\mathbf{A}\Delta^{-1}\mathbf{Z}^H]\mathbf{R}^{-1}\} \\ = \text{tr}\{[(\mathbf{A} - (\mathbf{X}^H\mathbf{R}^{-1}\mathbf{X})^{-1}\mathbf{X}^H\mathbf{R}^{-1}\mathbf{Z})^H(\mathbf{X}^H\mathbf{R}^{-1}\mathbf{X}) \\ \cdot (\mathbf{A} - (\mathbf{X}^H\mathbf{R}^{-1}\mathbf{X})^{-1}\mathbf{X}^H\mathbf{R}^{-1}\mathbf{Z}) + \mathbf{Z}^H\mathbf{R}^{-1}\mathbf{Z} \\ - \mathbf{Z}^H\mathbf{R}^{-1}\mathbf{X}(\mathbf{X}^H\mathbf{R}^{-1}\mathbf{X})^{-1}\mathbf{X}^H\mathbf{R}^{-1}\mathbf{Z}]\Delta^{-1}\}. \end{aligned}$$

Thus it is apparent that minima in  $\mathbf{X}$  and  $\mathbf{A}$  (fixing  $\mathbf{A}$  and  $\mathbf{X}$ , respectively) occur when

$$\begin{aligned} \hat{\mathbf{A}} &= (\mathbf{X}^H\mathbf{R}^{-1}\mathbf{X})^{-1}\mathbf{X}^H\mathbf{R}^{-1}\mathbf{Z} \\ \hat{\mathbf{X}} &= \mathbf{Z}\Delta^{-1}\mathbf{A}^H(\mathbf{A}\Delta^{-1}\mathbf{A}^H)^{-1}. \end{aligned}$$

Of course, the product  $\mathbf{XA}$  does not uniquely determine the factors  $\mathbf{X}$  and  $\mathbf{A}$  without additional information in the form of calibration tables (which provide structure for  $\mathbf{X}$ ) or waveform features (which provide structure for  $\mathbf{A}$ ). Therefore, a general estimation procedure based on the least-squares formulation must incorporate a nonlinear optimization step that includes calibration as well as waveform features.



**KEITH W. FORSYTHE** is a staff member in the Advanced Techniques group. He received an S.B. degree and an S.M. degree, both in mathematics, from MIT. In 1978 he joined Lincoln Laboratory, where he has worked in the areas of spread-spectrum communications, adaptive sensor array processing, and SAR imaging. His work on spread-spectrum systems includes electromagnetic modeling (geometric theory of diffraction) of antennas mounted on an airframe, error-correction coding, jam-resistant synchronization techniques, and digital matched-filter design and performance. In the area of adaptive sensor array processing he helped develop a number of signal processing algorithms that exploit waveform features to achieve levels of performance (beamforming, direction finding, geolocation, and other forms of parameter estimation) beyond those attainable by non-exploitive techniques. His work on SAR imaging involves techniques for resolution enhancement and interference rejection for foliage penetration systems.