J-Divergence Detection Currency before and after Conventional and Adaptive Beamforming

D. A. Abraham

Technical Report
APL-UW TR 2501
January 2025

Applied Physics Laboratory
1013 NE 40th Street

University of Washington
Seattle, Washington 98105-6698

Contract N00024-21-D-6400 under task order N00024-22-F-8714

Acknowledgements

This report was sponsored by the Office of Naval Research, Code 32, Undersea Signal Processing, through Naval Sea Systems Command contract N00024-21-D-6400 under task order N00024-22-F-8714.

Abstract

Beamforming algorithms are typically designed to maximize their output signal-to-interference-and-noise power ratio (SINR), under the assumption that doing so will optimize the probability of detection (P_d), given a design probability of false alarm (P_f), in the ensuing detection algorithm. An alternative performance metric, the J-divergence detection currency (JDC), is employed here to represent performance before and after beamforming. Building on early use of the J-divergence in modeling array processing performance, the basic analysis is extended to account for correlated multipath signals and shaded conventional beamforming. The reduction in performance observed in practical adaptive beamforming algorithms that must estimate the array covariance matrix (ACM) or its eigen-structure is then assessed for processors having a beta-distributed SINR loss factor, representing a number of popular processors. Simple approximations to the JDC in this scenario that are accurate at low SINR, as well as more involved ones for higher SINR, are presented along with the tools required to evaluate them. The analysis presented in this report allows assessing the potential gain in performance from combining multipath signals and the losses incurred by ACM estimation in a metric that is easily combined across multiple measurements, is more closely related to the (P_d, P_f) detection metrics than SINR, and can be evaluated throughout the signal and information processing chain.

Contents

1 Introduction

The signal and information processing chain (SIPC) of most remote sensing systems is formed by a sequence of operations, with each module designed and optimized using its own performance metrics. For example, beamforming algorithms are typically designed to maximize their output signal-to-noise-and-interference power ratio (SINR) and detection algorithms operating on beam responses are usually formulated to maximize the probability of detection (P_d) given a fixed probability of false alarm (P_f). When each module performs well, the resulting cascade typically provides good "end-to-end" performance, even though the metrics of the individual components differ.

This report represents a step toward evaluating the performance of a sonar system throughout the SIPC using the same performance metric, the J-divergence detection currency (JDC). As described in Sect. 1.1, the JDC [units: dB] is the logarithmic form of the J-divergence [1], which is representative of the distance between two probability distributions. The J-divergence was applied in [2] as an alternative to SINR for assessing array processors and in [3] as a simple-to-use replacement for the (P_d, P_f) operating point of matched filters and energy detectors. In [4], it was extended to account for a heavy-tailed noise background and incorporated the thresholding operation commonly found in systems that combine multiple measurements (e.g., across waveforms, frequency bands, or arrays).

The focus of this report is on evaluating JDC before and after beamforming. Building on the results of [2], the JDC inherent in the array data is obtained for correlated multipath signals and compared to the JDC after conventional and clairvoyant (i.e., perfect) minimum-variance-distortionless-response (MVDR) adaptive beamforming in Sect. 3. The losses arising from estimation of the array covariance matrix (ACM) or its eigen-structure, which is required in any practical MVDR-based adaptive beamformer, are then assessed in Sects. 4 & 5 using the beta-distributed SINR loss model of [5] for the sample-matrix-inversion (SMI) MVDR processor. Simple approximations to the JDC achieved by the SMI-MVDR processor are presented, as are the tools required to more accurately evaluate JDC, including results for an energy detector and the case of combining correlated signals arriving at the array on multiple paths.

Pertinent background material is provided in terms of a brief introduction to J-divergence and JDC (Sect. 1.1) and then a review of beamforming and statistical modeling of array data (Sect. 2). In Sect. 6, background material and tools related to the scaled-F distribution are presented, owing to their importance in obtaining JDC for practical adaptive beamformers. Finally, MATLAB® code implementing the key results can be found in the appendix.

1.1 J-divergence and detection currency

The J-divergence [1] is a measure of the distance between two probability density functions (PDFs). When it is used to represent detection performance, the PDFs describe the data or a decision statistic under the noise-only (H_0) and signal-present (H_1) hypotheses. If $f_0(t)$ and $f_1(t)$ are the PDFs under these two hypotheses, the J-divergence between them is

$$J = \int [f_1(t) - f_0(t)] \log\left[\frac{f_1(t)}{f_0(t)}\right] \, \mathrm{d}t. \tag{1}$$

An advantage J-divergence has in comparison to the traditional (P_d, P_f) detection metrics is that it can be evaluated at different points in the SIPC. For example, in this report it is first evaluated from

the array data and then after beamforming. For a single beam response (complex or instantaneous intensity) containing a bandpass Gaussian-fluctuating signal in Gaussian noise, the J-divergence is

$$J = \frac{s^2}{1+s} \quad \text{[unitless]}, \tag{2}$$

where s is the linear-quantity beam-output SINR [3]. This mapping from SINR to J-divergence illustrates that detection performance is quadratically related to the linear-quantity SINR when $s \ll 1$ and linearly related when $s \gg 1$. Given the wide range of values that might be observed, conversion to a logarithmic quantity, termed the *J-divergence detection currency* (JDC) via[1]

$$\text{JDC} = 5 \log_{10} J \quad \text{[units: dB]} \tag{3}$$

can help relate how much J-divergence is required for different levels of performance. As described in [3], low-, medium-, and high-quality operating points are achieved by, respectively, 5, 8, and 10 dB of detection currency. For example, 5 dB of JDC is roughly equivalent to $P_d = 0.5$ and P_f between 10^{-3} and 10^{-4}, which often represents a minimum detectable level (MDL) operating point.

A property of the J-divergence making it an attractive choice when modeling systems that combine multiple measurements is that the total J-divergence across M independent measurements is simply the sum of the M individual J-divergences. This result carries with it the caveat that the measurements are combined optimally (e.g., through a log-likelihood ratio). As an example, an energy detector with a time-bandwidth product of M would have J-divergence of $Ms^2/(1+s)$ if all of the bins have SINR equal to s and the aforementioned Gaussian statistical models. Sub-optimal processing, however, results in a reduction in detection currency. This highlights the *data processing inequality* where processing cannot increase J-divergence, but only maintain it or degrade it.

In this report, the detection performance achieved by an optimal beamformer is assessed in Sect. 3, first from the array data and then from the beam responses. The loss in JDC incurred by the practical necessity of using an estimated covariance matrix in an adaptive beamformer is then evaluated in Sects. 4 & 5 from a statistical characterization of the beam response. This demonstrates how performance can be assessed at different points in the SIPC in the same currency, providing a clear comparison of the loss incurred by each component.

2 Modeling the array data and beamforming

As with most performance metrics, evaluation of J-divergence requires definition of a detector or decision statistic and statistical models of the signal and noise. In this report, as is common in the analysis of array signal processing for sonar applications, the bandpass signal and noise are assumed to follow Gaussian distributions. The detection processing consists of narrowband beamforming and formation of an instantaneous intensity from the squared modulus of the complex beam response. Although most of the analysis is carried out for a single intensity sample, an energy detector formed by integrating a number of intensities is also considered.

In support of the derivations of J-divergence before and after array processing to be presented in Sect. 3, this section contains a brief introduction to statistical modeling of array data and beamforming in the sonar application. The reader is referred to textbooks such as [6, 7] for more detail in this area.

[1] From (2), J can be viewed as a ratio of squared intensities and is therefore converted to decibels according to $10 \log_{10} \sqrt{J} = 5 \log_{10} J$ as shown in (3).

2.1 Statistical models of the array data

In narrowband beamforming of an array with N sensors, a single "snapshot" of the array data consists of a dimension-N vector $\mathbf{x}$, where each element contains a specific frequency bin of a discrete Fourier transform (DFT) formed from one coherent processing interval (CPI) of acoustic time-series measurements from each sensor. The units of $\mathbf{x}$ are field-quantity units, such as volts or micropascals. To account for different scalings, such variables will be described generically as having field-quantity units or, after squaring, power-quantity units. Under the Gaussian models for the bandpass signal and noise, the detection hypotheses can be stated as

$$\begin{aligned} H_0&\text{: } \mathbf{x} \sim \mathcal{CN}(\mathbf{0}, \mathbf{Q}) \\ H_1&\text{: } \mathbf{x} \sim \mathcal{CN}(\mathbf{0}, \mathbf{R}) \,, \end{aligned} \tag{4}$$

where the notation $\mathbf{x} \sim \mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ indicates $\mathbf{x}$ follows a complex-Gaussian distribution with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. Note that the signal enters through the array covariance matrix (ACM) and not the mean. When a signal is present the ACM, $\mathbf{R} = \mathbf{P} + \mathbf{Q}$, is assumed to comprise a signal component ($\mathbf{P}$) and to simplify to the noise-and-interference ACM ($\mathbf{Q}$) when the signal power is zero.

The probability density function (PDF) of the zero-mean, multivariate, complex Gaussian distribution with covariance matrix $\mathbf{R}$ is

$$f(\mathbf{x}) = \frac{1}{\pi^N |\mathbf{R}|} e^{-\mathbf{x}^H \mathbf{R}^{-1} \mathbf{x}} \text{ for } \mathbf{x} \in \mathbb{C}^N \tag{5}$$

from [8, Sect. 5.6.9], where $|\mathbf{R}|$ is the determinant of $\mathbf{R}$. Note that the superscripts ^{T} and ^{H} represent the transpose and conjugate-transpose operations, respectively. Using this PDF in the two hypotheses in (4) leads to a log-likelihood-ratio (LLR)

$$l(\mathbf{x}) = \log\left[\frac{f_1(\mathbf{x})}{f_0(\mathbf{x})}\right] = \mathbf{x}^H \left(\mathbf{Q}^{-1} - \mathbf{R}^{-1}\right) \mathbf{x} + \log\left(|\mathbf{Q}\mathbf{R}^{-1}|\right). \tag{6}$$

This will be used in Sect. 3.1 to derive the J-divergence of the array data.

Although convention dictates that random variables should be represented by upper-case letters and their observed values by lower-case ones, this conflicts with the notation of lower and upper case bold letters for, respectively, vectors and matrices. The latter is followed in this report to aid in interpreting dimensions in matrix-vector equations. Whether or not a variable is a random quantity must be discerned from context.

2.1.1 Signal and noise models

When the object of interest (OOI) is in the far field of the array in a basic ocean model [8, pg. 96], the signal is accurately represented at the array as a plane-wave arrival. This implies that the same signal is observed at each sensor in the array, subject to a set of time delays ($\tau_n(\theta)$ [units: s] for sensor n) relative to when the signal arrives at the array origin. The time delays only depend on the geometry of the array and θ [units: rad], the direction from which the signal arrives at the array. In general θ is two-dimensional, with angles for azimuth and elevation. However, in the line-array examples used in this report, it simplifies to a scalar conical angle.

The DFT operation converts the delays into phases, which are represented in vector form as

$$\mathbf{d}(\theta) = \begin{bmatrix} e^{-j2\pi f \tau_1(\theta)} & \cdots & e^{-j2\pi f \tau_N(\theta)} \end{bmatrix}^T, \tag{7}$$

where f [units: Hz] is the center frequency of the DFT bin under consideration. In this formulation, $\mathbf{d}(\theta)$ acts as a transfer function mapping the complex amplitude of the signal observed at the array origin to that observed at each sensor. If the signal for a given CPI arrives from angle θ_s [units: rad] and has a zero-mean complex-Gaussian-distributed amplitude a_s [units: field], the array data can be represented by

$$\mathbf{x} = a_\mathrm{s}\mathbf{d}_\mathrm{s} + \mathbf{v} \quad \text{[units: field]}, \tag{8}$$

where $\mathbf{d}_\mathrm{s} = \mathbf{d}(\theta_\mathrm{s})$ and the interference and noise is in the vector $\mathbf{v} \sim \mathcal{CN}(\mathbf{0}, \mathbf{Q})$ [units: field], which is assumed to be independent of the signal. The ACM when signal is present is obtained by taking the expected value of the vector outer product of the data,

$$\mathbf{R} = E_1\left[\mathbf{x}\mathbf{x}^H\right] = \lambda_\mathrm{s}\mathbf{d}_\mathrm{s}\mathbf{d}_\mathrm{s}^H + \mathbf{Q}, \tag{9}$$

where $\lambda_\mathrm{s} = E_1[|a_\mathrm{s}|^2]$ [units: power] is the signal power observed at a single sensor. Note that the subscript on the expectation operator indicates the hypothesis under which the expectation is taken.

In underwater acoustic remote sensing, most scenarios involve a signal propagating from the object of interest (OOI) to a sensor array along multiple paths. The single-path model in (9) is easily extended to the case of N_s paths, leading to

$$\mathbf{R} = \mathbf{D}\mathbf{\Lambda}_\mathrm{s}\mathbf{D}^H + \mathbf{Q}, \tag{10}$$

where

$$\mathbf{D} = \begin{bmatrix} \mathbf{d}(\theta_1) & \cdots & \mathbf{d}(\theta_{N_\mathrm{s}}) \end{bmatrix} \tag{11}$$

is formed by the N_s array transfer-function vectors representing the different paths and $\mathbf{\Lambda}_\mathrm{s}$ [units: power] is an N_s-by-N_s covariance matrix representing the signal powers and coupling across the paths. In most cases, $N_\mathrm{s} \ll N$, so the signal component only occupies a low-rank subspace within the N-dimensional array data. The inter-path covariance matrix $\mathbf{\Lambda}_\mathrm{s}$ can vary from a diagonal matrix if there is no correlation between the paths to being a rank-one vector outer product if the signals are fully correlated.

The ACM when no signal is present ($\mathbf{Q}$) is typically assumed to consist of isotropic ambient noise and some number of interferences having compact angular support. An interference arriving from a specific angle can be modeled in the same manner as the signal (e.g., with a covariance proportional to $\mathbf{d}_i\mathbf{d}_i^H$, where $\mathbf{d}_i$ is the array transfer-function vector for the angle of arrival of the interference). To represent angular spreading, the rank-one matrix $\mathbf{d}_i\mathbf{d}_i^H$ can be tapered through a Hadamard (element-wise) product with a Toeplitz matrix containing ϱ^i on the ith sub- and super-diagonals and $\varrho^0 = 1$ on the diagonal, where $\varrho \in [0, 1]$. The spreading increases as the tapering becomes more severe, which occurs when ϱ is decreased from one toward zero.

Using [9, pg. 38, Table 3.1], the sensor-to-sensor correlation of isotropic noise produces an ACM with

$$\{\mathbf{Q}\}_{n,m} = \lambda_\mathrm{n}\mathrm{sinc}\left(\frac{2f}{c_w}\|\vec{x}_n - \vec{x}_m\|_2\right) \tag{12}$$

in the (n, m) element, where λ_n [units: power] is the sensor-level average noise power, $\mathrm{sinc}(x) = \sin(\pi x)/(\pi x)$, f [units: Hz] is the frequency, c_w [units: m/s] is the speed of sound in water, $\vec{x}_n$ [units: m] is the position of the nth sensor, and $\|\vec{x}_n - \vec{x}_m\|_2$ [units: m] is the distance between the nth and mth sensors. For the uniformly spaced line array used in the examples found in this report, this simplifies to $\mathbf{Q} = \lambda_\mathrm{n}\mathbf{I}$ at the design frequency of the array (i.e., the sensors are spaced every $d = \lambda/2 = c_w/(2f)$ [units: m] so $\{\mathbf{Q}\}_{n,m} = \lambda_\mathrm{n}\mathrm{sinc}(n-m) = \lambda_\mathrm{n}\delta[n-m]$). Noise having a scaled-identity covariance matrix is referred to as being spatially white.

2.2 Beam responses and beam-output SINR

The beamforming operation can be viewed as a spatial matched filter, where the processing correlates the array data with the model for signals coming from a given direction. A conventional beamformer (CBF) uses the transfer-function vector in (7) to form the complex beam response

$$Z = \mathbf{d}^H(\theta)\mathbf{x} \quad \text{[units: field]}, \tag{13}$$

which emphasizes signals arriving from the angle θ while suppressing other directions. It accomplishes this by undoing the changes in phase the signal on each sensor was subjected to, which is the frequency-domain equivalent of time-aligning the temporal signals so they add constructively. Signals arriving from other angles will have different delays and be suppressed. Beams responses are formed or steered in directions spanning the angles from which signals of interest might arrive, typically spaced so their beampatterns[2] overlap 3-dB down from their peak value. From this usage, the vector $\mathbf{d}(\theta)$ is commonly known as the array steering vector.

2.2.1 Shaded conventional beamforming

When shading is applied to a CBF to suppress sidelobes in the beampattern (at the expense of a wider mainlobe), the beam response is

$$Z = \mathbf{d}^H(\theta)\mathbf{A}\mathbf{x} \quad \text{[units: field]}, \tag{14}$$

where $\mathbf{A}$ is a diagonal matrix containing the weights applied to the sensors in the array. More generally, the complex beam response can be described as applying a filter vector $\mathbf{w}$ to the array data,

$$Z = \mathbf{w}^H\mathbf{x} \quad \text{[units: field]}. \tag{15}$$

Given $\mathbf{w}$, the linearity of (15) dictates that Z is also zero-mean and complex-Gaussian distributed,

$$Z \sim \mathcal{CN}(0, \mathbf{w}^H\mathbf{R}\mathbf{w}) \quad \text{[units: field]}. \tag{16}$$

Its variance, $\mathbf{w}^H\mathbf{R}\mathbf{w}$ [units: power], is the same as its power because it has zero mean.

In some applications, the complex beam response is combined over multiple frequencies to reconstitute a time-domain beam response. In other applications, the instantaneous intensity formed by squaring the modulus of Z is used as a decision statistic or combined across several beam responses

[2]The beampattern is the response of a beamformer steered to a fixed direction when observing a unit-strength test signal arriving from a hypothetical angle that is varied. It is similar to the transfer function of a filter, where the beam response is equivalent to the output of the filter.

to form one. Noting that Z is complex-Gaussian distributed, its squared modulus is seen to be the sum of the squares of two independent and identically distributed real Gaussian random variables. This leads to

$$Y = |Z|^2 = |\mathbf{w}^H\mathbf{x}|^2 \sim \text{Expon}\{\lambda_y\} \quad \text{[units: power]} \tag{17}$$

following an exponential distribution with mean $\lambda_y = \mathbf{w}^H\mathbf{R}\mathbf{w}$ [units: power]. When a CBF with shading is pointed at the signal (i.e., $\mathbf{w} = \mathbf{A}\mathbf{d}(\theta_\text{s}) = \mathbf{A}\mathbf{d}_\text{s}$), the average beam response is

$$\lambda_y = \mathbf{w}^H\mathbf{R}\mathbf{w} = \lambda_\text{s}(\mathbf{a}^T\mathbf{1})^2 + \mathbf{d}_\text{s}^H\mathbf{A}\mathbf{Q}\mathbf{A}\mathbf{d}_\text{s} \quad \text{[units: power]} \tag{18}$$

where $\mathbf{a} = \mathbf{A}\mathbf{1}$ is a vector of the shading weights and $\mathbf{1}$ is a vector of ones.

The beam-output SINR for this processor is

$$s^\text{cbf} = \frac{\lambda_\text{s}(\mathbf{a}^T\mathbf{1})^2}{\mathbf{d}_\text{s}^H\mathbf{A}\mathbf{Q}\mathbf{A}\mathbf{d}_\text{s}} \quad \text{[unitless]}. \tag{19}$$

In spatially white noise, $\mathbf{Q} = \lambda_\text{n}\mathbf{I}$, where λ_n [units: power] is the sensor-level noise power. The beam-output SINR is then the product of the sensor-level SINR ($\lambda_\text{s}/\lambda_\text{n}$) and the array gain,

$$s^\text{cbf} = \frac{\lambda_\text{s}}{\lambda_\text{n}} \cdot \frac{(\mathbf{a}^T\mathbf{1})^2}{\mathbf{a}^T\mathbf{a}} \quad \text{[unitless]}. \tag{20}$$

It is straightforward to show that this is maximized by uniform shading (i.e., $\mathbf{a} = \mathbf{1}$).

When the signal arrives on only one path, the J-divergence after conventional beamforming is obtained by using the beam-output SINR from (19) or (20) in (2) and then (3) to form detection currency. Although the uniformly shaded CBF is the optimal processor in spatially white noise, the background often contains spatially isolated interferences, such as noise from surface ships. In these scenarios, an adaptive beamformer can reduce the impact of the interferences and improve the beam-output SINR.

2.2.2 Clairvoyant minimum-variance-distortion-response (MVDR) adaptive beamforming

Many approaches to adaptive beamforming employ what is known as a distortionless-response constraint. For a beamformer using a weight vector $\mathbf{w}$ pointed to a direction having steering vector $\mathbf{d}_\text{s}$, this requires

$$\mathbf{w}^H\mathbf{d}_\text{s} = 1, \tag{21}$$

so a signal arriving from the main response axis (MRA) of the beam passes through the beamformer with unit gain. In the shaded CBF, this results in requiring the shading weights to satisfy $\mathbf{a}^T\mathbf{1} = 1$, which leads to the signal component of the average beam response in (18) to equal λ_s, which is the sensor-level signal power. When using this scaling, the SINR gain of the beamformer manifests as a reduction in the beam-output interference-and-noise power relative to that observed on a single sensor.

Subject to the distortionless-response constraint, criteria such as minimizing power or variance or maximizing SINR can be used to show that the optimal weight vector is

$$\mathbf{w} = \frac{\mathbf{Q}^{-1}\mathbf{d}_\text{s}}{\mathbf{d}_\text{s}^H\mathbf{Q}^{-1}\mathbf{d}_\text{s}}. \tag{22}$$

This is typically known as the minimum-variance distortionless-response (MVDR) adaptive beamformer. When the noise ACM ($\mathbf{Q}$) is assumed to be known perfectly, it is called a *clairvoyant*-MVDR processor. In practice $\mathbf{Q}$ must be estimated, the impact of which will be covered in Sects. 4 & 5. In some scenarios it is only possible to estimate the signal-plus-noise ACM ($\mathbf{R}$). Although this can be used to construct an adaptive beamformer, there are potential issues with signal cancellation and these are not addressed in this report.

It is straightforward to derive the clairvoyant-MVDR beamformer from the variance of Z in (16),

$$\mathrm{Var}\{Z\} = \mathbf{w}^H\mathbf{R}\mathbf{w} = \lambda_\mathrm{s}|\mathbf{w}^H\mathbf{d}_\mathrm{s}|^2 + \mathbf{w}^H\mathbf{Q}\mathbf{w} \quad \text{[units: power]}, \tag{23}$$

using a Lagrange multiplier to enforce the distortionless-response constraint. When using the clairvoyant-MVDR beamformer, the variance simplifies to

$$\mathbf{w}^H\mathbf{R}\mathbf{w} = \lambda_\mathrm{s} + \frac{1}{\mathbf{d}_\mathrm{s}^H\mathbf{Q}^{-1}\mathbf{d}_\mathrm{s}} \quad \text{[units: power]}. \tag{24}$$

Noting that (23) and (24) also represent the average beam response (i.e., $E[Y]$), the beam-output SINR is seen to be

$$s^\mathbf{w} = \frac{\lambda_\mathrm{s}|\mathbf{w}^H\mathbf{d}_\mathrm{s}|^2}{\mathbf{w}^H\mathbf{Q}\mathbf{w}} \quad \text{[unitless]} \tag{25}$$

for a beamforming vector $\mathbf{w}$ and

$$s^\mathrm{mvdr} = \lambda_\mathrm{s}\mathbf{d}_\mathrm{s}^H\mathbf{Q}^{-1}\mathbf{d}_\mathrm{s} \quad \text{[unitless]} \tag{26}$$

for the optimal processor.

As noted for the CBF processor, these beam-output SINRs can be used in (2) to obtain J-divergence and then (3) for detection currency. As might be expected, the clairvoyant-MVDR adaptive beamformer will always have equal or better detection currency than the CBF because its SINR is always equal or better.

The potential improvement in SINR is captured by the array gain improvement (AGI), which is

$$\mathrm{AGI} = \frac{s^\mathrm{mvdr}/\lambda_\mathrm{s}}{s^\mathrm{cbf}/\lambda_\mathrm{s}} = \frac{\mathbf{d}_\mathrm{s}^H\mathbf{Q}^{-1}\mathbf{d}_\mathrm{s}}{\mathbf{d}_\mathrm{s}^H\mathbf{d}_\mathrm{s}} \cdot \frac{\mathbf{d}_\mathrm{s}^H\mathbf{Q}\mathbf{d}_\mathrm{s}}{\mathbf{d}_\mathrm{s}^H\mathbf{d}_\mathrm{s}} \quad \text{[unitless]} \tag{27}$$

when the CBF employs uniform shading. Based on the derivation of the clairvoyant-MVDR adaptive beamformer, this is always greater than or equal to one. Using the eigen-decomposition of $\mathbf{Q} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H$, where $\mathbf{U}$ is a matrix of eigenvectors and $\mathbf{\Lambda}$ is a diagonal matrix of eigenvalues, the AGI can be written as

$$\mathrm{AGI} = \frac{(\mathbf{u}^H\mathbf{\Lambda}^{-1}\mathbf{u})(\mathbf{u}^H\mathbf{\Lambda}\mathbf{u})}{(\mathbf{u}^H\mathbf{u})^2} \quad \text{[unitless]} \tag{28}$$

where

$$\mathbf{u} = \mathbf{U}\mathbf{d}_\mathrm{s} \tag{29}$$

is a projection of the steering vector onto the eigenvector matrix. From this form, it can be seen that AGI will be one when the steering vector isolates any single eigenvector. If it isolates two

eigenvectors, then it is only greater than one if the corresponding eigenvalues are not equal. Constant eigenvalues, where the AGI is expected to be close to one, are representative of spatially flat noise at the angles represented by their corresponding eigenvectors. A strong spatially isolated interference would dominate a single eigenvector, implying the AGI when steered directly to the interference is close to one, but above one when near enough to be affected by it. From the perspective of detection performance, the use of adaptive beamforming requires balancing the potential gain through AGI with the losses incurred by the need to estimate the array covariance matrix. Adaptive beamformers are also more sensitive to mismatch in the array modeling. For example, the steering vectors can be misaligned when the sensor positions are not known exactly, which can lead to signal suppression in an adaptive beamformer. Losses arising from this type of mismatch are not considered in this report.

3 Detection currency in the array data

The basic result for the linear-quantity J-divergence of a signal impinging on an array from a single direction is simply (3)

$$J = \frac{s^2}{1+s} \quad \text{[unitless]}, \tag{30}$$

where s is the beam-output SINR of the clairvoyant-MVDR beamformer defined in (26). This result was derived in [2, eq. 30].[3] Note that the results presented here are a factor of two greater than those found in [2] owing to their treatment of the occurrence of H_0 or H_1 as an equally likely random event (with probability 1/2). The hypotheses here are treated as deterministic events.

This result, as derived in [2] from the array data, is identical to that obtained when working with the distributions representing the (complex or intensity) beam response for a clairvoyant-MVDR beamformer, which implies it is an optimal array processor with respect to detection. The effect of using an estimated covariance matrix is considered in Sects. 4 and 5. The focus of this section is on the case of a signal arriving at the array along a set of paths, with the above result arising as a special case of the more general result.

3.1 Derivation for multiple correlated signal paths

Consider the detection hypotheses described in (4), where the array snapshot ($\mathbf{x}$) is modeled as a dimension-N, zero-mean, complex-Gaussian random vector with a covariance matrix $\mathbf{Q}$ under H_0 and $\mathbf{R}$ under H_1. From the integral definition of the J-divergence in (1), it can be seen that it is also equal to the difference in the average LLR when the expectation is taken under H_1 and H_0. Using (6) and letting $\mathrm{tr}\{\mathbf{R}\}$ be the trace of $\mathbf{R}$, it is straightforward to show that the average LLR under H_1 is

$$E_1[l(\mathbf{x})] = \mathrm{tr}\big\{\big(\mathbf{Q}^{-1} - \mathbf{R}^{-1}\big)\, E_1\big[\mathbf{x}\mathbf{x}^H\big]\big\} + \log\big(|\mathbf{Q}\mathbf{R}^{-1}|\big) \tag{31}$$

$$= \mathrm{tr}\big\{\mathbf{Q}^{-1}\mathbf{R}\big\} - N + \log\big(|\mathbf{Q}\mathbf{R}^{-1}|\big) \tag{32}$$

and that it is

$$E_0[l(\mathbf{x})] = N - \mathrm{tr}\big\{\mathbf{R}^{-1}\mathbf{Q}\big\} + \log\big(|\mathbf{Q}\mathbf{R}^{-1}|\big) \tag{33}$$

[3] The second line of [2, eq. 30] should have $(L\eta_\theta^2)^2$ in the numerator.

under H_0. The J-divergence is then the difference,

$$J = E_1[l(\mathbf{x})] - E_0[l(\mathbf{x})] = \mathrm{tr}\{\mathbf{Q}^{-1}\mathbf{R} + \mathbf{R}^{-1}\mathbf{Q}\} - 2N \quad \text{[unitless]}. \tag{34}$$

With the exception of the aforementioned factor-of-two difference, this is identical to [2, eq. 27].

When a signal of interest arrives at the array along multiple paths, the array covariance matrix (ACM) is modeled as described in (10). Using the matrix inversion lemma, the inverse of $\mathbf{R}$ for this low-rank signal model is

$$\mathbf{R}^{-1} = \mathbf{Q}^{-1} - \mathbf{Q}^{-1}\mathbf{D}\big(\mathbf{D}^H\mathbf{Q}^{-1}\mathbf{D} + \mathbf{\Lambda}_\mathrm{s}^{-1}\big)^{-1}\mathbf{D}^H\mathbf{Q}^{-1}. \tag{35}$$

Inserting (35) into (34) produces the J-divergence

$$J = \mathrm{tr}\Big\{2\mathbf{I} + \mathbf{Q}^{-1}\mathbf{D}\mathbf{\Lambda}_\mathrm{s}\mathbf{D}^H - \mathbf{Q}^{-1}\mathbf{D}\big(\mathbf{D}^H\mathbf{Q}^{-1}\mathbf{D} + \mathbf{\Lambda}_\mathrm{s}^{-1}\big)^{-1}\mathbf{D}^H\Big\} - 2N \tag{36}$$

$$= \mathrm{tr}\Big\{\mathbf{D}^H\mathbf{Q}^{-1}\mathbf{D}\mathbf{\Lambda}_\mathrm{s}\Big[\mathbf{I} - \big(\mathbf{D}^H\mathbf{Q}^{-1}\mathbf{D}\mathbf{\Lambda}_\mathrm{s} + \mathbf{I}\big)^{-1}\Big]\Big\} \tag{37}$$

$$= \mathrm{tr}\Big\{\mathbf{\Lambda}_\mathrm{n}^{-1}\mathbf{\Lambda}_\mathrm{s}\Big[\mathbf{I} - \big(\mathbf{\Lambda}_\mathrm{n}^{-1}\mathbf{\Lambda}_\mathrm{s} + \mathbf{I}\big)^{-1}\Big]\Big\} \quad \text{[unitless]}, \tag{38}$$

where

$$\mathbf{\Lambda}_\mathrm{n}^{-1} = \mathbf{D}^H\mathbf{Q}^{-1}\mathbf{D} \quad \text{[units: 1/power]} \tag{39}$$

has diagonal elements equal to one over the average beam-output noise power pointing in the directions of the source signals and off-diagonals related to the coupling of these directions relative to the interference and noise background. The product $\mathbf{\Lambda}_\mathrm{n}^{-1}\mathbf{\Lambda}_\mathrm{s}$ [unitless] can be viewed as collectively characterizing the beam-output SINR and coupling across the directions from which the signal of interest arrives.

If $\mathbf{\Lambda}_\mathrm{n}^{-1}\mathbf{\Lambda}_\mathrm{s}$ is diagonalizable, using its eigen-decomposition in (38) results in

$$J = \sum_{i=1}^{N_\mathrm{s}} \frac{\nu_i^2}{1+\nu_i} \quad \text{[unitless]}, \tag{40}$$

where ν_i [unitless] is the ith eigenvalue. This form of the J-divergence allows for a computationally efficient evaluation compared with using the array matrices of the general result found in (34). It also permits consideration of some typical scenarios with respect to how correlated the signal and beam noise are from one path to the next.

3.1.1 Widely separated paths

If the paths on which a signal arrives are far enough apart in angle, the beam responses are likely to be only weakly correlated under the noise-only hypothesis, which implies $\mathbf{\Lambda}_\mathrm{n}$ can be approximated by a diagonal matrix. If the signal arrivals are also assumed to be only weakly correlated, then the J-divergence simplifies to

$$J = \sum_{i=1}^{N_\mathrm{s}} \frac{s_i^2}{1+s_i} \quad \text{[unitless]}, \tag{41}$$

where $s_i = \lambda_{\mathrm{s},i}/\lambda_{\mathrm{n},i}$ [unitless] is the clairvoyant-MVDR beam-output SINR from the ith path ($\lambda_{\mathrm{s},i}$ and $\lambda_{\mathrm{n},i} = 1/\mathbf{d}_i^H\mathbf{Q}^{-1}\mathbf{d}_i$ [units: power] are the ith diagonal elements of $\mathbf{\Lambda}_\mathrm{s}$ and $\mathbf{\Lambda}_\mathrm{n}$, respectively). This is the same J-divergence attained when N_s independent intensity measurements obtained from clairvoyant-MVDR beam responses are combined using an optimal detector.

3.1.2 Perfectly correlated signal in beams with uncorrelated noise

If the signals arriving at the array are perfectly correlated, but have been subjected to different propagation loss, the path-to-path covariance matrix can be described as

$$\mathbf{\Lambda}_\mathrm{s} = \mathbf{a}_\mathrm{s}\mathbf{a}_\mathrm{s}^H \quad \text{[units: power]}, \tag{42}$$

where $\mathbf{a}_\mathrm{s}$ [units: field] is a vector containing the transfer functions mapping the source signal to that observed along the different paths multiplied by the square root of the signal power. The squared modulus of the ith element is the power in the ith path ($\lambda_{\mathrm{s},i}$ [units: power]). If the noise beam responses in these directions are assumed to be only weakly correlated, then $\mathbf{\Lambda}_\mathrm{n}$ is a diagonal matrix of the clairvoyant-MVDR beam-output noise powers ($\lambda_{\mathrm{n},i}$ [units: power]). Using these in (38) results in

$$J = \frac{(\mathbf{a}_\mathrm{s}^H\mathbf{\Lambda}_\mathrm{n}^{-1}\mathbf{a}_\mathrm{s})^2}{1+\mathbf{a}_\mathrm{s}^H\mathbf{\Lambda}_\mathrm{n}^{-1}\mathbf{a}_\mathrm{s}} = \frac{\left(\sum_{i=1}^{N_\mathrm{s}} s_i\right)^2}{1+\sum_{i=1}^{N_\mathrm{s}} s_i} \quad \text{[unitless]}, \tag{43}$$

where $s_i = \lambda_{\mathrm{s},i}/\lambda_{\mathrm{n},i}$ [unitless] is the clairvoyant-MVDR beam-output SINR. This is the same as a single intensity sample with SINR equal to $s_\mathrm{tot} = \mathbf{a}_\mathrm{s}^H\mathbf{\Lambda}_\mathrm{n}^{-1}\mathbf{a}_\mathrm{s} = \sum_{i=1}^{N_\mathrm{s}} s_i$ [unitless]. In terms of the eigen-decomposition-result of (40), $\mathbf{\Lambda}_\mathrm{n}^{-1}\mathbf{\Lambda}_\mathrm{s}$ has rank one and therefore only one non-zero eigenvalue, which is equal to s_tot.

The result shown in (43) is similar to a coherent combination of the paths, which can be achieved if the signals are fully correlated. The result found in (41) is inherently an incoherent combination, which explains why (43) always equals or exceeds (41) (this can be seen by noting that the function $J = s^2/(1+s)$ is effectively quadratic in s when $s \ll 1$ and linear if $s \gg 1$). This suggests that there will be little to gain over incoherent combination when the SINRs of fully correlated signals are large, whereas some improvement might be obtained if a number of weak paths can be combined coherently. Partial signal correlation across the paths will produce a J-divergence somewhere between these two extremes.

3.1.3 Partially correlated signals from the same direction

Suppose multiple partially correlated signals arrive at the array from the same direction. In this scenario,

$$\mathbf{D} = \mathbf{d}_\mathrm{s}\mathbf{1}^T, \tag{44}$$

where $\mathbf{d}_\mathrm{s}$ is the steering vector pointing toward the source and $\mathbf{1}$ is a vector of ones. Using this in (39) produces

$$\mathbf{\Lambda}_\mathrm{n}^{-1} = (\mathbf{d}_\mathrm{s}^H\mathbf{Q}^{-1}\mathbf{d}_\mathrm{s})\mathbf{1}\mathbf{1}^T \quad \text{[units: 1/power]}. \tag{45}$$

Although the right side of (45) is not invertible, it can still be used in (38) to obtain the J-divergence through the product

$$\mathbf{\Lambda}_\mathrm{n}^{-1}\mathbf{\Lambda}_\mathrm{s} = (\mathbf{d}_\mathrm{s}^H\mathbf{Q}^{-1}\mathbf{d}_\mathrm{s})\mathbf{1}(\mathbf{\Lambda}_\mathrm{s}\mathbf{1})^T \quad \text{[unitless]}, \tag{46}$$

which is a vector outer product whose only non-zero eigenvalue is

$$\nu_1 = (\mathbf{d}_\mathrm{s}^H\mathbf{Q}^{-1}\mathbf{d}_\mathrm{s})(\mathbf{1}^T\mathbf{\Lambda}_\mathrm{s}\mathbf{1}) \quad \text{[unitless]}. \tag{47}$$

From (40), the J-divergence is simply $J = \nu_1^2/(1+\nu_1)$.

As might be expected, the signals combine in the water to produce a sum with power equal to $\mathbf{1}^T\mathbf{\Lambda}_\mathrm{s}\mathbf{1}$, so (47) is the beam-output SINR when pointing at the source. This can also be seen by describing the signal component as

$$\mathbf{x}_\mathrm{s} = \sum_{i=1}^{N_\mathrm{s}} a_{\mathrm{s},i}\,\mathbf{d}_i = \mathbf{D}\mathbf{a}_\mathrm{s} \quad \text{[units: field]}, \tag{48}$$

where $\mathbf{a}_\mathrm{s} = [a_{\mathrm{s},1}\cdots a_{\mathrm{s},N_\mathrm{s}}]^T$ [units: field] is a dimension-N_s, zero-mean, complex-Gaussian random vector with covariance matrix $\mathbf{\Lambda}_\mathrm{s}$. Using (44) in (48) leads to

$$\mathbf{x}_\mathrm{s} = \sum_{i=1}^{N_\mathrm{s}} a_{\mathrm{s},i}\mathbf{d}_i = (\mathbf{1}^T\mathbf{a}_\mathrm{s})\mathbf{d}_\mathrm{s} = a_\mathrm{s}\mathbf{d}_\mathrm{s} \quad \text{[units: field]}, \tag{49}$$

where $a_\mathrm{s} = \mathbf{1}^T\mathbf{a}_\mathrm{s}$ [units: field] is a zero-mean, complex-Gaussian random variable with variance $\mathbf{1}^T\mathbf{\Lambda}_\mathrm{s}\mathbf{1}$. Positive correlation among the signal components (indicated by positive off-diagonal terms in $\mathbf{\Lambda}_\mathrm{s}$) enhances the overall power, whereas negative correlation degrades it. This can be seen by considering a two-component signal with amplitude covariance matrix

$$\mathbf{\Lambda}_\mathrm{s} = \lambda_\mathrm{s}\begin{bmatrix} 1 & \varrho \\ \varrho^* & 1 \end{bmatrix} \quad \text{[units: power]}, \tag{50}$$

which leads to

$$\mathbf{1}^T\mathbf{\Lambda}_\mathrm{s}\mathbf{1} = 2(1+\mathrm{Real}\{\varrho\})\,\lambda_\mathrm{s} \quad \text{[units: power]}. \tag{51}$$

For fully correlated signals coming from the same direction, the combined signal power is $4\lambda_s$ and the J-divergence is

$$J = \frac{(4s)^2}{1+4s} \quad \text{[unitless]}, \tag{52}$$

where $s = \lambda_\mathrm{s}\mathbf{d}_\mathrm{s}^H\mathbf{Q}^{-1}\mathbf{d}_\mathrm{s}$ is the clairvoyant-MVDR beam-output SINR for a single component as in (26). The difference between the $4s$ used here and the $2s$ that would arise in (43) for two signal components is that the signals in the latter case are subjected to noise arriving from two different directions.

3.2 Detection currency after beamforming

In order to determine if there is any loss in detection currency from the beamforming operation, the beam data need to be characterized. This is done here under the assumption that the beamforming vector is deterministic, which covers conventional beamforming and a clairvoyant adaptive beamformer. In order to focus on the signal components, only the beams pointing to the signal arrivals are considered. Let the columns of

$$\mathbf{W} = \begin{bmatrix}\mathbf{w}_1 & \cdots & \mathbf{w}_{N_\mathrm{s}}\end{bmatrix} \tag{53}$$

contain these beamforming vectors. The complex beam responses are then

$$\mathbf{z} = \mathbf{W}^H\mathbf{x} \quad \text{[units: field]}. \tag{54}$$

Under this data-reducing transformation, the hypotheses stated in (4) become

$$\begin{aligned} H_0&\text{: } \mathbf{z} \sim \mathcal{CN}(\mathbf{0}, \mathbf{Q}_z) \\ H_1&\text{: } \mathbf{z} \sim \mathcal{CN}(\mathbf{0}, \mathbf{R}_z)\,, \end{aligned} \tag{55}$$

where the covariance matrices under H_0 and H_1 are, respectively,

$$\mathbf{Q}_z = \mathbf{W}^H \mathbf{Q} \mathbf{W} \quad \text{[units: power]} \tag{56}$$

and

$$\mathbf{R}_z = \mathbf{W}^H \mathbf{D} \mathbf{\Lambda}_\mathrm{s} \mathbf{D}^H \mathbf{W} + \mathbf{Q}_z \quad \text{[units: power]}. \tag{57}$$

When $\mathbf{Q}_z^{-1}$ exists, the J-divergence after forming beams pointing to the N_s signal paths can be shown to be

$$J_\mathrm{bf} = \mathrm{tr}\left\{ \tilde{\mathbf{\Lambda}}_\mathrm{n}^{-1} \mathbf{\Lambda}_\mathrm{s} \left[\mathbf{I} - \left(\tilde{\mathbf{\Lambda}}_\mathrm{n}^{-1} \mathbf{\Lambda}_\mathrm{s} + \mathbf{I} \right)^{-1} \right] \right\} \quad \text{[unitless]} \tag{58}$$

through use of (34) and the derivation leading to (38), where

$$\tilde{\mathbf{\Lambda}}_\mathrm{n}^{-1} = \mathbf{D}^H \mathbf{W} \left(\mathbf{W}^H \mathbf{Q}^{-1} \mathbf{W} \right)^{-1} \mathbf{W}^H \mathbf{D} \quad \text{[units: 1/power]} \tag{59}$$

plays the same role as $\mathbf{\Lambda}_\mathrm{n}^{-1}$ in (39).

3.2.1 Clairvoyant-MVDR adaptive beamforming

For the clairvoyant-MVDR adaptive beamformer described in (22), the matrix $\mathbf{W}$ can be described by

$$\mathbf{W} = \mathbf{Q}^{-1} \mathbf{D} \mathbf{U} \tag{60}$$

where the scalar terms in the denominator of (22) are placed on the diagonal of

$$\mathbf{U} = \mathrm{diag}\left\{ 1/\mathbf{d}_1^H \mathbf{Q}^{-1} \mathbf{d}_1, \ldots, 1/\mathbf{d}_{N_\mathrm{s}}^H \mathbf{Q}^{-1} \mathbf{d}_{N_\mathrm{s}} \right\}. \tag{61}$$

Inserting (60) in (59) leads to

$$\tilde{\mathbf{\Lambda}}_\mathrm{n}^{-1} = \mathbf{\Lambda}_\mathrm{n}^{-1}, \tag{62}$$

which implies the complex beam responses of a clairvoyant-MVDR adaptive beamformer pointing to the signal paths have the same total J-divergence as the array data—there is no loss in detection currency. Owing to the J-divergence's data processing inequality (processing can only maintain or degrade it), this result indicates that the complex beam responses in the directions of the signal arrivals are sufficient statistics for detection.

3.2.2 Conventional beamforming with shading

For the shaded conventional beamformer described in (14), the N_s complex beam responses are formed using

$$\mathbf{W} = \mathbf{AD} \tag{63}$$

where $\mathbf{A}$ is a diagonal matrix containing the shading applied to each sensor. Using (63) in (59) produces

$$\tilde{\mathbf{\Lambda}}_\mathrm{n}^{-1} = \mathbf{D}^H\mathbf{AD}\big(\mathbf{D}^H\mathbf{AQAD}\big)^{-1}\mathbf{D}^H\mathbf{AD} \quad \text{[units: 1/power]}. \tag{64}$$

When the beams (accounting for shading) are linearly independent (i.e., $\mathbf{D}^H\mathbf{AD}$ has full rank),

$$\tilde{\mathbf{\Lambda}}_\mathrm{n} = (\mathbf{D}^H\mathbf{AD})^{-1}\big(\mathbf{D}^H\mathbf{AQAD}\big)\,(\mathbf{D}^H\mathbf{AD})^{-1} \quad \text{[units: power]}, \tag{65}$$

which illustrates that $\tilde{\mathbf{\Lambda}}_\mathrm{n}$ is a normalized form of the noise-only covariance matrix of the CBF beams pointing to the signal paths. The special cases considered in Sects. 3.1.1–3.1.3 for the array data apply here when the clairvoyant-MVDR beam-output SINR is replaced by the CBF beam-output SINR from (19). The loss in detection currency is therefore commensurate with the reduction in SINR of the CBF relative to the clairvoyant-MVDR adaptive beamformer for the particular scenario under consideration.

3.3 Example: Correlated multipath signals

A uniformly spaced line array (ULA) is used to illustrate how the location of the signal components relative to each other and any interferences can impact performance. The average beam responses of the CBF and MVDR beamformers for the example are shown in Fig. 1 for an array containing $N = 32$ sensors at the design frequency, where the sensor spacing is one half of a wavelength. The nth element of the array steering vector is

$$\{\mathbf{d}(\theta)\}_n = e^{j\pi n\gamma\cos\theta} \;\text{ for }\; n = 0, 1, \ldots, N-1, \tag{66}$$

where θ [units: rad] is the angle from the forward end-fire of the line array and γ is the ratio of the operating frequency to the design frequency of the array (here it is set to $\gamma = 1$). Note that the results in this section are presented in units of $u = \cos\theta$ [unitless] or $Nu/2$ [units: #], with the latter representing a beam index for a CBF processor with uniform shading and 3-dB-down overlapping beampatterns. The average CBF beam response shown in Fig. 1 employed Taylor shading,[4] so it has a slightly larger beamwidth than for uniform shading and a loss of ≈ 0.7 dB in array gain against white noise.

The scenario considered consists of a primary signal component impinging on the array at a constant angle, a secondary signal component arriving from an angle that will be varied, and an interference with a sensor-level interference-to-noise ratio (INR) of 0 dB. The primary signal path is broadside to the ULA at $u = \cos\theta = 0$ and the interference is placed less than two beamwidths away at $u = 0.1$. In Fig. 1, the secondary path is shown at $u = 0.3$. The signal power was chosen so the J-divergence detection currency (JDC) for the primary signal path by itself was 5 dB. This required a beam-output SINR of ≈ 10.4 dB and resulted in a sensor-level SNR of -4.5 dB for each path. The secondary signal path was assumed to have the same power as the primary path.

[4]The Taylor window was designed so the first $\bar{n} = 4$ sidelobes are held 30 dB down from the mainlobe peak.

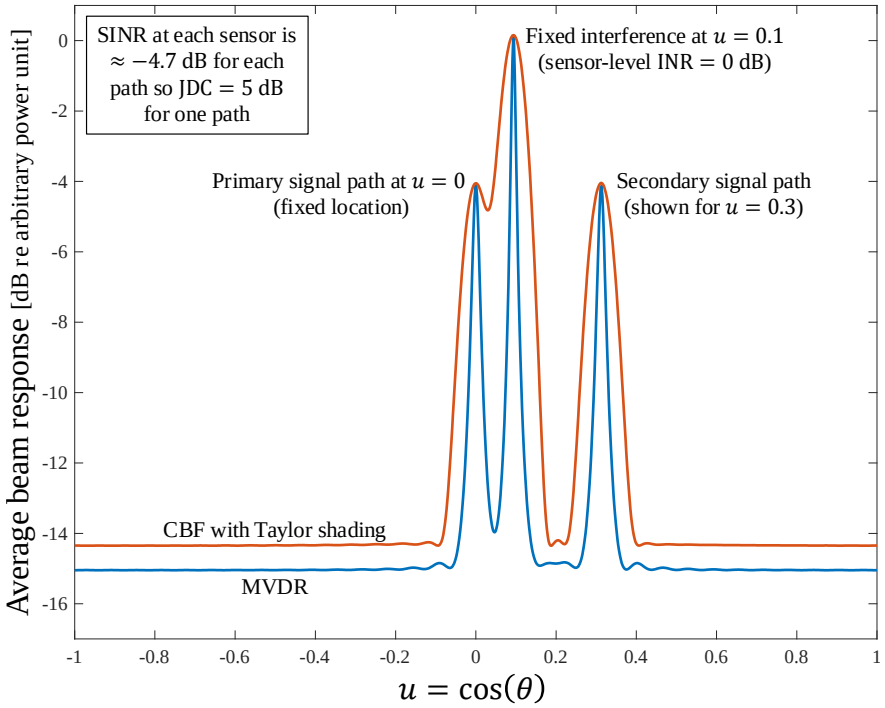

Figure 1: Average beam responses for the line-array example containing a single fixed interference, a primary signal path at broadside and a secondary signal path that is shown here at $u = 0.3$.

The array-level JDC is shown in Fig. 2 as a function of the position of the secondary signal path. Note that the abscissa is restricted to $u \in [-0.25, 0.375]$, which is roughly from four beams aft of the primary signal path to four beams forward of the interference. Most of the cases are seen to lie between the JDC attained by a single path (5 dB) and the maximum achieved for two fully correlated paths arriving from the same direction (8.15 dB). The exception arises when the paths arrive at the same angle and are negatively correlated, which leads to a destructive combination. However, even a correlation coefficient of $\varrho = -0.9$ has minimal impact when the paths are separated by half of a beamwidth or more. In these areas, the fully correlated paths ($\varrho = 1$) have a slightly higher nominal performance (the dotted line) compared to the others, which are closer to an incoherent combination (dash-dot line). Mathematically, the difference is between $J = (2s)^2/(1 + 2s)$ via (43) for the fully correlated signals arriving in independent beams and $J = 2s^2/(1 + s)$ from (41) for the incoherent combination.

As might be expected, the effect of the interference is to suppress the secondary signal path when it is coincident in angle, causing JDC to revert to the performance achieved with only the primary path. The results are intuitive and easily explained by the correlation between the signal paths and their locations relative to each other and the interference. A traditional (P_d, P_f) analysis of this scenario would require definition and analysis of the detector combining the beam responses, which is not straightforward when there is correlation between the signal paths. This illustrates one of the advantages of using JDC as a performance metric.

The JDC achieved when applying Taylor-shaded conventional beamforming in the directions of the two signal paths is shown for the $\varrho = 0$ case (dashed gold line). In regions not impacted by the interference, the CBF JDC is about one third of a decibel below the optimal JDC, owing to

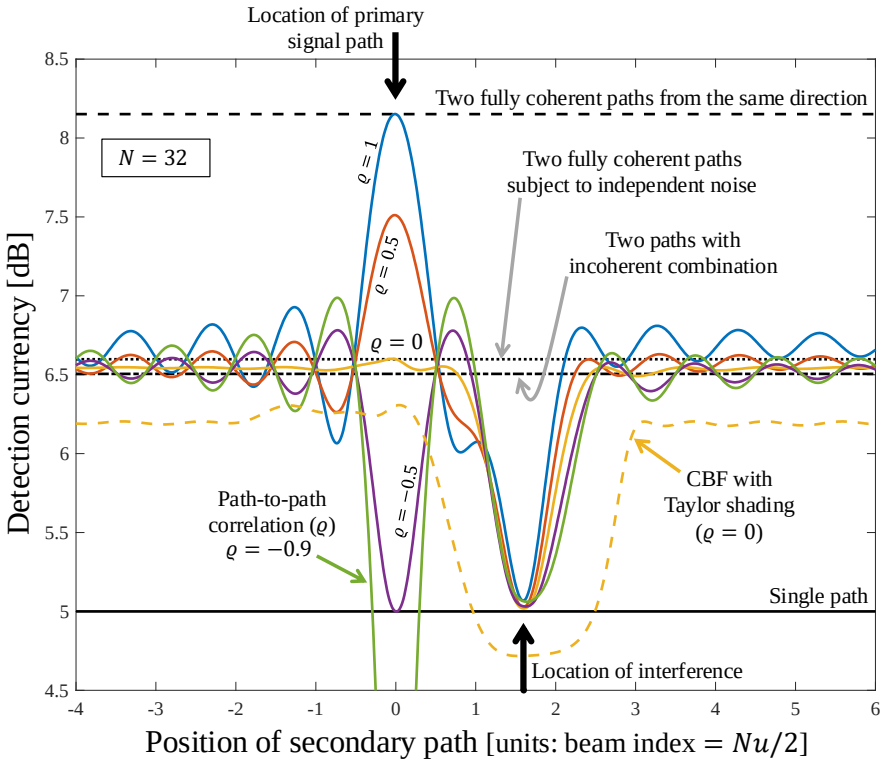

Figure 2: Detection currency in the array data as a function of the position of the secondary signal path.

the inherent loss in array gain from shading the array. However, when the interference is near the secondary signal path, but not too close, the loss is greater. Although not considered here, the performance of the CBF processor would be further exacerbated if the primary path were affected by an interference. These cases present an opportunity for improved performance through the use of adaptive beamforming, as long as the losses arising from mismatch and estimation errors are less than those encountered by the shaded CBF processor.

Using JDC as the performance metric also allows a straightforward assessment of the detection performance attainable when combining the beam responses across the multipath arrivals. As noted in Sect. 1.1, an MDL operating point requires 5 dB of JDC. This example was designed so the primary path by itself achieved MDL. Adding a second independent observation with the same SINR would result in a $5 \log_{10} 2 = 1.5$ dB increase in JDC, which is seen to be achieved in Fig. 2 for correlated signals (even negatively correlated ones) when the secondary path is distant from both the interference and the primary signal path. When it is near the primary path and positively correlated, the JDC can increase to over 8 dB, which was defined as a medium-quality operating point. Thus, incorporating an additional path of equal strength will in most scenarios stabilize an MDL operating point, moving it partway to a medium-quality one. More rarely, it might bring it fully to a medium-quality operating point. These same qualitative assessments could have been shown through an analysis of P_d given a fixed P_f. However, the effort required to perform that analysis is generally inhibitive and if the value of P_f is not representative of that observed in a system, only the trends in P_d would be informative.

4 Modeling detection currency with an estimated array covariance matrix (ACM)

In practice, the array covariance matrix (ACM) is unknown and must be estimated from data. The classic results of Reed, Mallett, and Brennan (RMB) in [5] are used here to illustrate the impact of the estimation losses on J-divergence detection currency. In this analysis, the signal is assumed to arrive from a single direction, so it can be represented as a rank-one component in the ACM,

$$\mathbf{R} = \lambda_{\mathrm{s}} \mathbf{d}_{\mathrm{s}} \mathbf{d}_{\mathrm{s}}^H + \mathbf{Q} \quad \text{[units: power]}, \tag{67}$$

where λ_{s} [units: power] is the sensor-level signal power, $\mathbf{d}_{\mathrm{s}}$ is the steering vector pointing toward the signal source, and $\mathbf{Q}$ [units: power] is the ACM of the noise and interference background. It is also assumed that the array data used to estimate $\mathbf{Q}$ are statistically independent of the data to which the beamforming filter vector is applied.

The focus in this section is on the theory supporting the evaluation and approximation of JDC for an adaptive beamformer using an estimated ACM. The tools presented here are applied in Sect. 5 to examine the loss in JDC arising from the estimation error.

4.1 SINR loss and a statistical model for the beam response

The squared modulus of the complex beam response pointed in the direction of the signal is

$$Y = |\mathbf{w}^H \mathbf{x}|^2 \quad \text{[units: power]}, \tag{68}$$

where $\mathbf{x}$ [units: field] is an N-by-1 vector of array data for a single frequency and $\mathbf{w}$ is the MVDR beamforming vector formed according to (22), but using an estimated ACM ($\hat{\mathbf{Q}}$),

$$\mathbf{w} = \frac{\hat{\mathbf{Q}}^{-1} \mathbf{d}_{\mathrm{s}}}{\mathbf{d}_{\mathrm{s}}^H \hat{\mathbf{Q}}^{-1} \mathbf{d}_{\mathrm{s}}}. \tag{69}$$

Conditioned on $\hat{\mathbf{Q}}$, Y is exponentially distributed with mean

$$\lambda_Y = \mathbf{w}^H \mathbf{R} \mathbf{w} = \mathbf{w}^H \mathbf{Q} \mathbf{w} + \lambda_{\mathrm{s}} \tag{70}$$

$$= \frac{1}{\rho \mathbf{d}_{\mathrm{s}}^H \mathbf{Q}^{-1} \mathbf{d}_{\mathrm{s}}} + \lambda_{\mathrm{s}} \quad \text{[units: power]}, \tag{71}$$

where

$$\rho = \frac{\left(\mathbf{d}_{\mathrm{s}}^H \hat{\mathbf{Q}}^{-1} \mathbf{d}_{\mathrm{s}} \right)^2}{(\mathbf{d}_{\mathrm{s}}^H \mathbf{Q}^{-1} \mathbf{d}_{\mathrm{s}})(\mathbf{d}_{\mathrm{s}}^H \hat{\mathbf{Q}}^{-1} \mathbf{Q} \hat{\mathbf{Q}}^{-1} \mathbf{d}_{\mathrm{s}})} \quad \text{[unitless]} \tag{72}$$

represents the loss in the signal-to-interference-and-noise power ratio (i.e., the SINR loss) arising from mismatch in the ACM estimate. Note that $\rho \in [0, 1]$ and $\rho = 1$ when $\hat{\mathbf{Q}} = \mathbf{Q}$. As will be described in Sect. 4.1.1, the classic result from [5] dictates that ρ follows a beta distribution when $\hat{\mathbf{Q}}$ is formed from a sample covariance matrix.

Forming the SINR from the conditional average in (71) (i.e., retaining the conditioning on $\hat{\mathbf{Q}}$) results in

$$\mathrm{SINR} = \rho \lambda_{\mathrm{s}} \mathbf{d}_{\mathrm{s}}^H \mathbf{Q}^{-1} \mathbf{d}_{\mathrm{s}} = \rho s \quad \text{[unitless]}, \tag{73}$$

where s [unitless] is the same as in (26), which assumed perfect knowledge of the ACM. This clearly illustrates how ρ represents the loss in SINR relative to a clairvoyant-MVDR beamformer. In a statistical analysis of the beam response, it is convenient to normalize Y by the noise power achieved in the clairvoyant-MVDR beamformer,

$$T = \frac{Y}{1/\mathbf{d}_\mathrm{s}^H \mathbf{Q}^{-1} \mathbf{d}_\mathrm{s}} \sim \mathrm{Expon}\{\rho^{-1} + s\} \quad \text{[unitless]}, \tag{74}$$

given ρ. This provides a relatively simple statistical model where the mean of the normalized beam response, conditioned on the ACM estimate, is simply $\rho^{-1} + s$.

When the same beamforming vector is applied to multiple snapshots and the beam responses are combined incoherently, as in an energy detector, the decision statistic becomes

$$Y = \sum_{m=1}^{M} |\mathbf{w}^H \mathbf{x}_m|^2 \quad \text{[units: power]}. \tag{75}$$

The normalized decision statistic (T), conditioned on ρ, is then gamma distributed with shape parameter M and scale parameter $\rho^{-1} + s$. Note that s is the SINR in a single snapshot. Although the primary focus of this report is on testing the instantaneous intensity formed from a single snapshot, the energy-detector case is considered in Sect. 5.3.

4.1.1 SINR loss for various ACM estimators

The most basic estimate of the ACM is the "sample" covariance matrix,

$$\hat{\mathbf{Q}} = \frac{1}{K} \sum_{k=1}^{K} \mathbf{z}_k \mathbf{z}_k^H \quad \text{[units: power]}, \tag{76}$$

formed by averaging the vector-outer-product of K independent snapshots of signal-free background data, $\mathbf{z}_k \sim \mathcal{CN}(\mathbf{0}, \mathbf{Q})$ [units: field] for $k = 1,\ldots,K$. As previously noted, the data used to estimate the ACM are assumed to be signal-free and statistically independent of the test data to which it is applied in an adaptive beamformer. Using the sample covariance matrix in the MVDR beamforming vector in (69) is known as a sample-matrix-inversion (SMI) MVDR processor.

The RMB results in [5] dictate that the SINR loss (ρ) for the SMI-MVDR processor is beta distributed,

$$\rho \sim \mathrm{Beta}(\alpha, \beta)\,, \tag{77}$$

with shape parameters

$$\alpha = K - N + 2 \;\text{ and }\; \beta = N - 1. \tag{78}$$

In order for $\hat{\mathbf{Q}}$ to be invertible, $K \geq N$, which implies $\alpha \geq 2$.

From [8, Sect. 5.6.5], the PDF of the beta distribution with shape parameters α and β is

$$f_\rho(\rho) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \rho^{\alpha-1} (1-\rho)^{\beta-1} \;\text{ for }\; \rho \in [0, 1]. \tag{79}$$

The analysis in [5] led to a recommendation that $K = 2N$ or higher in order for the average SINR loss,

$$E[\rho] = \frac{\alpha}{\alpha+\beta} = \frac{K-N+2}{K+1}, \tag{80}$$

to be less than 3 dB. This requirement can be difficult to meet when N is large, owing to limits on the temporal stability of the noise and interference background. Limiting the adaptivity of the beamformer resolves this problem, but comes at the expense of a more complicated analysis. Fortunately, the beta distribution is still useful in representing the SINR loss for several reduced-adaptivity processors, where N is (essentially) replaced by the number of spatially distinct interferences plus one to account for the distortionless-response constraint on the signal. For example, the results of [10] for principal component inverse (PCI), [11] for diagonal loading, and [12] for dominant-mode-rejection (DMR) dictate that

$$\alpha = K - D + 1 \text{ and } \beta = D, \tag{81}$$

where D is the rank of the interference subspace.[5] Because D is generally much smaller than N, the mean in (80) with $N = D + 1$ illustrates that the bulk of the distribution on ρ is closer to one for a fixed value of K. This implies there is less SINR loss when using a reduced-adaptivity processor compared with the sample-matrix ACM estimate.

As seen in the following sections, the beam response when ρ is beta distributed can be approximated using a generalized Pareto distribution (GPD). It is reasonable to expect this to hold for most processors, which provides a means for evaluating the loss in detection currency arising from ACM estimation. The GPD parameters are obtained by matching the moments of the normalized decision statistic (T), which depends on the first two inverse moments of ρ. It is straightforward to estimate these through a Monte-Carlo analysis, use them in (88) and (90) to obtain the moments of T, obtain the GPD parameters through (104), assess the fit of the GPD model under H_0 and H_1, and then evaluate the JDC by converting (113) to decibels.

4.2 Moments and scintillation index of the beam response

Recall that $T|\rho \sim \text{Expon}\{\rho^{-1} + s\}$ when signal is present, with the conditional mean simplifying to $1/\rho$ when there is only noise. This characterization provides a simple method for obtaining the moments of T under the two hypotheses, which are necessary to approximate its distribution by the GPD and to evaluate the scintillation index. Because the majority of the quantities related to T and ρ (e.g., moments or moment ratios) are unitless and dominate the remaining analysis, units are only noted when necessary in the remainder of this report.

Under H_0, the moments of T are simply obtained from consecutive application of the moment definitions for the exponential and beta distributions while exploiting the conditioning,

$$E_0[T^k] = E_\rho[E_0[T|\rho]] = E_\rho\left[\rho^{-k}\right]\Gamma(k+1) \tag{82}$$

$$= \frac{\Gamma(\alpha+\beta)\Gamma(\alpha-k)\Gamma(k+1)}{\Gamma(\alpha+\beta-k)\Gamma(\alpha)}, \tag{83}$$

[5] As described in [10–12], the result shown in (81) carries with it certain assumptions for each of the processors.

where the kth inverse moment of ρ is

$$E\left[\frac{1}{\rho^k}\right] = \frac{\Gamma(\alpha+\beta)\Gamma(\alpha-k)}{\Gamma(\alpha+\beta-k)\Gamma(\alpha)} = \frac{(\alpha+\beta-1)\cdots(\alpha+\beta-k)}{(\alpha-1)\cdots(\alpha-k)} \tag{84}$$

$$= E\left[\frac{1}{\rho^{k-1}}\right]\frac{(\alpha+\beta-k)}{(\alpha-k)} \text{ for } k<\alpha. \tag{85}$$

In Sect. 4.1.1 it was noted that $\alpha \geq 2$, so the first two moments of T under H_0,

$$\mu_1' = \frac{\alpha+\beta-1}{\alpha-1} \text{ and } \mu_2' = \frac{2(\alpha+\beta-1)(\alpha+\beta-2)}{(\alpha-1)(\alpha-2)}, \tag{86}$$

are always available.

When signal is present, the conditional moment produces a polynomial in $1/\rho$ that must be expanded,

$$E_1[T^k] = E_\rho[E_1[T|\rho]] = E_\rho\left[\left(\frac{1}{\rho}+s\right)^k\right]\Gamma(k+1). \tag{87}$$

This results in

$$\mu_1'(s) = E\left[\frac{1}{\rho}\right] + s = \frac{\alpha+\beta-1}{\alpha-1} + s \tag{88}$$

for the first moment and

$$\mu_2'(s) = 2\big(E\big[\rho^{-2}\big] + 2sE\big[\rho^{-1}\big] + s^2\big) \tag{89}$$

$$= \frac{2(\alpha+\beta-1)(\alpha+\beta-2)}{(\alpha-1)(\alpha-2)} + \frac{4s(\alpha+\beta-1)}{(\alpha-1)} + 2s^2 \tag{90}$$

for the second.

The scintillation index (SI), which is the ratio of the intensity variance to its squared average, provides an assessment of the heaviness of the PDF tails relative to the exponential distribution. Using (88) and (90), SI for the beam response obtained using the SMI-MVDR adaptive beamformer is

$$\text{SI} = \frac{\mu_2'(s) - [\mu_1'(s)]^2}{[\mu_1'(s)]^2} = 1 + \frac{2\beta(\alpha+\beta-1)}{(\alpha-2)[(\alpha-1)s+\alpha+\beta-1]^2} \quad \text{[unitless]}. \tag{91}$$

Substituting the beta-distribution parameters from the RMB results described in (78) leads to

$$\text{SI} = 1 + \frac{2K(N-1)}{(K-N)[K+(K-N+1)s]^2} \tag{92}$$

$$\approx 1 + \frac{2c^2}{N(1-c)[1+(1-c)s]^2} \text{ for } N \gg 1, \tag{93}$$

where $c = N/K \leq 1$ is the factor by which N is smaller than K. The approximation in (93) additionally assumes N is large. The scintillation index of a conventional or clairvoyant-MVDR beamformer is equal to one, representing an exponentially distributed intensity. Using an estimated ACM increases tail heaviness (as expected) and is reflected by SI exceeding one. The results shown

here illustrate that increasing s, decreasing c (i.e., increasing K relative to N), or increasing N while holding c fixed help drive down the tails of the beam-response distribution when the ACM is estimated by a sample covariance matrix. This also implies that the beam responses will have heavier tails when N is smaller (while keeping c fixed) and the SINR is not overly large.

It can also be seen from (92) that SI is infinite when $K = N$. Although the average beam intensity response in this scenario is finite ($E_1[T] = N + s$), the distribution is so heavy tailed that the variance is infinite. In most applications, a shaded CBF processor will provide better detection performance than an SMI-MVDR processor with $K = N$.

4.3 Integral and special-function representations of the beam-response distributions

The beam-response PDF is obtained by removing the conditioning on the exponential distribution described in (74),

$$f_1(t) = E_\rho\left[\frac{\rho}{1+\rho s} e^{-\rho t/(1+\rho s)}\right] \tag{94}$$

$$= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 \frac{\rho^\alpha (1-\rho)^{\beta-1}}{1+\rho s} e^{-\rho t/(1+\rho s)}\, \mathrm{d}\rho. \tag{95}$$

When no signal is present (i.e., under H_0), this simplifies to

$$f_0(t) = E_\rho\big[\rho e^{-\rho t}\big] = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 \rho^\alpha (1-\rho)^{\beta-1} e^{-\rho t}\, \mathrm{d}\rho \tag{96}$$

$$= \frac{\alpha}{\alpha+\beta} M_\rho(-t; \alpha+1, \beta) = \frac{\alpha}{\alpha+\beta} {}_1F_1(\alpha+1, \alpha+\beta+1; -t)\,. \tag{97}$$

The integral in (96) is seen to be proportional to the moment generating function (MGF) of a beta distribution with shape parameters $\alpha+1$ and β (i.e., $E[e^{x\rho}] = M_\rho(x; \alpha, \beta)$ if $\rho \sim \mathrm{Beta}(\alpha, \beta)$), which in turn can be obtained from a confluent hypergeometric function (i.e., ${}_1F_1(a, c; x)$, which is also known as a Kummer function [13, pg. 487, eq. 47:3:6]). This is a special case of the more general result found in [14, eq. 17]. Although it is possible to describe the PDF under H_1 in (95) as a confluent hypergeometric series of two variables (via [15, pg. 351, eq. 3.385]), the integrals in (95) & (96) are straightforward to evaluate numerically, with the J-divergence subsequently obtained through a trapezoidal-rule approximation (e.g., see the procedure recommended in [4, Sect. 5.2] and the MATLAB® code in App. A.3). Numerical routines for evaluating the Kummer function in (97) may (or may not) yield a simpler approach.

The cumulative distribution functions (CDFs) are similarly straightforward to describe in integral form:

$$F_1(t) = E_\rho\Big[1 - e^{-\rho t/(1+\rho s)}\Big] = 1 - \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 \rho^{\alpha-1} (1-\rho)^{\beta-1} e^{-\rho t/(1+\rho s)}\, \mathrm{d}\rho \tag{98}$$

and

$$F_0(t) = E_\rho\big[1 - e^{-\rho t}\big] = 1 - \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 \rho^{\alpha-1} (1-\rho)^{\beta-1} e^{-\rho t}\, \mathrm{d}\rho = 1 - {}_1F_1(\alpha, \alpha+\beta; -t)\,. \tag{99}$$

When evaluating $P_f = 1 - F_0(t)$, where t is the decision threshold, the expansion in [13, 47:6:2] for large negative arguments,

$$F_0(t) = 1 - \frac{\Gamma(\alpha+\beta)}{\Gamma(\beta)t^\alpha}\sum_{j=0}^{\infty}\frac{c_j}{t^j}, \tag{100}$$

is useful. The coefficients in the summation start with $c_0 = 1$ and can be obtained iteratively through

$$c_j = (\alpha+j-1)(j-\beta)\frac{c_{j-1}}{j}. \tag{101}$$

This form does require the argument t to be somewhat larger than both α and β to converge.

An interesting point can be discerned from (100) by examining the zero*th*-order approximation to $P_f = 1 - F_0(t)$ (represented by the term outside the summation) and in particular the t^α in the denominator. This implies that the tails of the beam-response distribution tend to a power-law decay, as opposed to the exponential decay predicted for the clairvoyant-MVDR beamformer. This will be exploited in Sect. 4.4 when approximations are developed to simplify evaluation of J-divergence.

4.4 Generalized-Pareto-distribution approximations to the beam response

Although evaluating J-divergence numerically is straightforward using the integral representations of the beam-response distributions shown in Sect. 4.3, it requires effort equivalent to a two-dimensional integral. Fortunately, reasonably accurate approximations are available, providing a simpler and quicker approach. The asymptotic (i.e., large argument) form of the CDF in (100) implies the tails tend to a power-law decay, which suggests using the generalized Pareto distribution (GPD) under H_0. Noting that the GPD model includes an exponentially distributed intensity, the proximity of the scintillation index to one when a high-SINR signal is present or when N and K are appropriately large enough (see Sect. 4.2) supports the argument under H_1. This leaves small values of N and K that might not yield a good GPD fit. From [8, Sect. 5.6.20], the GPD has PDF

$$f(t) = \frac{1}{\lambda(1+\gamma t/\lambda)^{\gamma^{-1}+1}} \text{ for } t \geq 0, \tag{102}$$

and CDF

$$F(t) = \frac{1}{(1+\gamma t/\lambda)^{\gamma^{-1}}} \text{ for } t \geq 0, \tag{103}$$

where γ is the shape parameter and λ is the scale parameter. It has mean $\lambda(1-\gamma)$ for $\gamma < 1$ and variance $\lambda^2/[(1-\gamma)^2(1-2\gamma)]$ for $\gamma < 0.5$.

Applying the method of moments to estimate the GPD parameters results in

$$\gamma = \frac{\mu_2' - 2[\mu_1']^2}{2(\mu_2' - [\mu_1']^2)} \text{ and } \lambda = \frac{\mu_1'\mu_2'}{2(\mu_2' - [\mu_1']^2)}. \tag{104}$$

Substituting the moments of T under H_0 from (86) produces

$$\gamma_0 = \frac{\beta}{\alpha(\alpha+\beta-3)+2} = \frac{\beta}{\alpha^2+\alpha\beta-3\alpha+2} \tag{105}$$

and

$$\lambda_0 = \frac{\alpha+\beta-1}{\alpha-1}\left[1 - \frac{\beta}{\alpha^2+\alpha\beta-3\alpha+2}\right]. \tag{106}$$

When signal is present, using (88) & (90) in (104) yields

$$\gamma_1 = \frac{\beta(\alpha+\beta-1)}{2\beta(\alpha+\beta-1)+(\alpha-2)[(\alpha-1)s+\alpha+\beta-1]^2} \tag{107}$$

for the shape parameter. The scale parameter can be described using γ_1 from (107) and the first moment of the GPD model,

$$\lambda_1 = \left(\frac{\alpha+\beta-1}{\alpha-1}+s\right)(1-\gamma_1). \tag{108}$$

In terms of N and K, the shape parameters are

$$\gamma_0 = \frac{N-1}{(K-2)(K-N+2)+2} \text{ and} \tag{109}$$

$$\gamma_1 = \frac{K(N-1)}{2K(N-1)+(K-N)[(K-N+1)s+K]^2}. \tag{110}$$

The scale parameters,

$$\lambda_0 = \frac{K}{K-N+1}(1-\gamma_0) \text{ and } \lambda_1 = \left[\frac{K}{K-N+1}+s\right](1-\gamma_1), \tag{111}$$

are again obtained from the first moment of the GPD model and the shape parameters.

The GPD model tends to an exponential distribution with mean $\lambda_i(1-\gamma_i)$ as $\gamma_i \to 0$. From (109) and (110), this is seen to occur as K or s increases, which corroborates the discussion in Sect. 4.2 on scintillation index.

4.5 J-divergence between two generalized Pareto distributions

The detection currency attained by an adaptive beamformer having a beta-distributed SINR loss can be approximated by the J-divergence between two GPD models using the parameters described in Sect. 4.4. The more general scaled-F distribution, which contains the GPD model as a special case, is required to evaluate the energy detector in Sect. 5.3. In support of this, the Kullback-Leibler (KL) divergences between two scaled-F distributions are presented in Sect. 6.4, with the J-divergence obtained from their sum, as shown in Sect. 6.5. The special case of two GPD models simplifies to

$$J = \gamma_1\left(1+\gamma_0^{-1}\right){}_2F_1\left(\gamma_1^{-1}, 1; 1+\gamma_1^{-1}, 1-\frac{\gamma_1\lambda_0}{\gamma_0\lambda_1}\right) - \gamma_1 \tag{112}$$

$$+\gamma_0\left(1+\gamma_1^{-1}\right){}_2F_1\left(\gamma_0^{-1}, 1; 1+\gamma_0^{-1}, 1-\frac{\gamma_0\lambda_1}{\gamma_1\lambda_0}\right) - \gamma_0 - 2 \quad \text{[unitless]} \tag{113}$$

for the J-divergence, with the detection currency obtained through JDC $= 5\log_{10} J$ [units: dB]. MATLAB® code for evaluating the Gauss hypergeometric function, ${}_2F_1(a,b;c;z)$, is found in App. C and MATLAB® code evaluating JDC for a single intensity sample formed from an SMI-MVDR adaptive beamformer in App. A.1.

5 Loss in detection currency for an SMI-MVDR adaptive beamformer

Most signal processing algorithms experience a loss in performance arising from estimation and mismatch errors. Estimation errors can arise when the processing exploits a model for which there are unknown parameters that must be estimated from the data. For example, many detection algorithms incorporate normalizers [8, Sects. 8.6 & 9.3] where the background power level is estimated and used to form a decision statistic that leads to a constant false alarm rate (CFAR). As might be expected, errors in the estimate result in a reduction in performance relative to a clairvoyant processor. In the detector, this can be reflected in a lower P_d (given a fixed P_f) than when the background power is known perfectly. When using the sonar equation to assess performance, it is common to describe the loss as the additional SINR required to maintain the desired (P_d, P_f) operating point.

Estimation of the ACM in adaptive beamforming is similar in principle to the case of a normalizer. However, because the forward modeling of JDC is straightforward, the ABF estimation-error loss (EEL) is defined as the decibel difference between the ideal performance and that achieved when the background is estimated,

$$\mathrm{EEL}_K = \mathrm{JDC}_\infty - \mathrm{JDC}_K \quad \text{[units: dB]}, \tag{114}$$

where the subscript indicates the number of snapshots used to estimate the ACM (assuming N is fixed). The clairvoyant-MVDR adaptive beamformer (i.e., when $K \to \infty$) attains an SINR of s and the beam response is exponentially distributed. This leads to

$$\mathrm{JDC}_\infty = 5 \log_{10}\left\{ \frac{s^2}{1+s} \right\} \quad \text{[units: dB]}, \tag{115}$$

which is the logarithmic form of the J-divergence in (30). The JDC for an SMI-MVDR adaptive beamformer is obtained using the generalized Pareto distribution (GPD) approximation as described in Sects. 4.4 & 4.5. The GPD approximation and some simpler ones for weak signals described in the following sections are compared to a numerical evaluation of JDC using the MATLAB® code found in App. A.3.

5.1 Estimation-error loss for the SMI-MVDR adaptive beamformer

The JDC achieved by the SMI-MVDR beamformer is compared to the performance of the clairvoyant-MVDR processor in Fig. 3 as a function of the clairvoyant-MVDR beam-output SINR for an array with $N = 32$ sensors and four values of K ranging from a minimum of $K = N + 1 = 33$ up to $K = 5N = 160$. The GPD approximations (solid lines) are seen to be quite accurate, except when the SINR and K are large, where it is optimistic relative to the numerical evaluation (dots), owing to a poor fit under H_0 at large values.[6] Although not shown, the GPD approximation also fails when $K = N$ owing to a poor fit at small values.

In addition to the GPD model, an approximation based on assuming the decision statistic is exponentially distributed is shown (dotted lines). The mean of the distribution, $\mu_1'(s) = r^{-1} + s$, is

[6] The JDC approximation for high SINR and large K can be improved by choosing the GPD model parameters under H_0 to more closely match the exact PDF in the vicinity of the average intensity when signal is present (at the expense of a poorer fit at lower values).

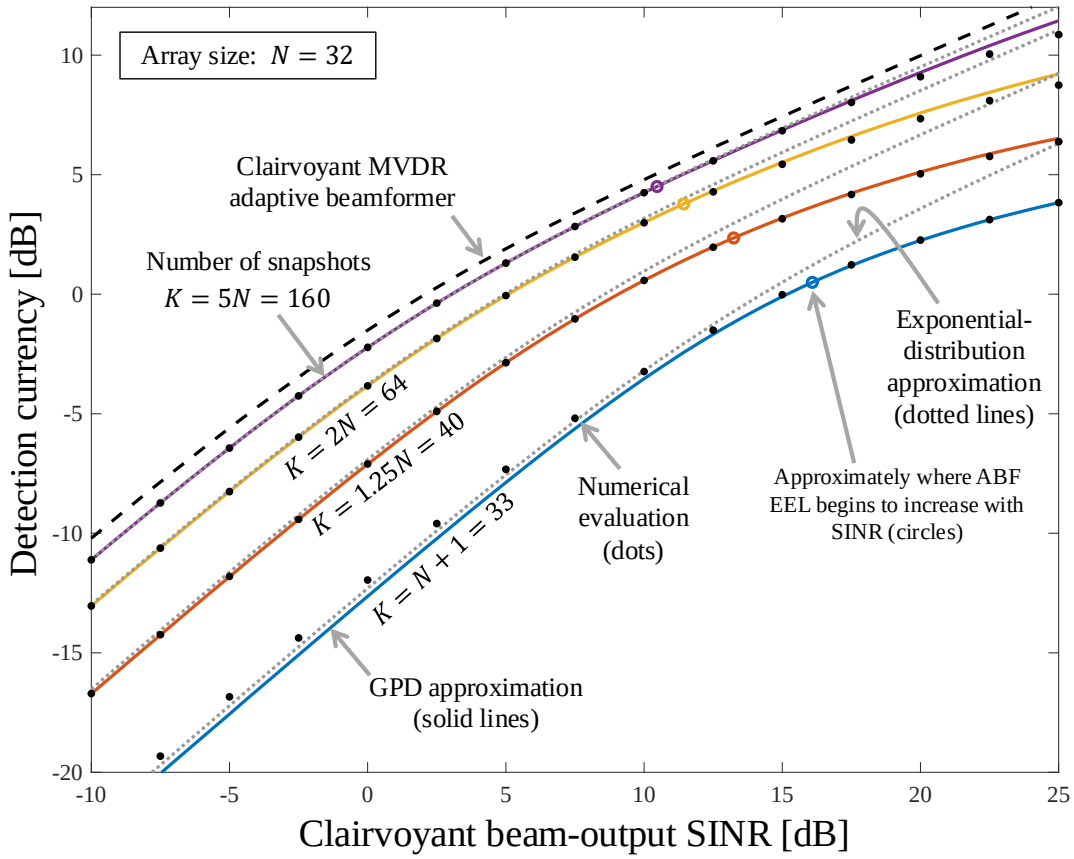

Figure 3: Detection currency as a function of the clairvoyant-MVDR beam-output SINR when estimating the array covariance matrix. Both approximations work well when the SINR is not too large, but eventually fail as SINR increases.

taken from (88), where

$$r = \frac{1}{E[1/\rho]} = \frac{\alpha - 1}{\alpha + \beta - 1} = \frac{K - N + 1}{K} \tag{116}$$

is one over the average beam-output noise power (i.e., $1/\mu_1'$ from (86)). This leads to using an SINR of rs in (115), which produces the approximation

$$\mathrm{JDC}_K \approx 5 \log_{10}\left\{\frac{(rs)^2}{1 + rs}\right\} \quad \text{[units: dB]} \tag{117}$$

and the ensuing estimation-error loss

$$\mathrm{EEL}_K \approx -10 \log_{10} r + 5 \log_{10}\left\{\frac{1 + rs}{1 + s}\right\} \quad \text{[units: dB]}. \tag{118}$$

As seen in Fig. 3, this simple approximation (dotted lines) is quite accurate for SINR up to about 10 dB. When only weak signals are of interest, as might be the case when many independent beam responses are combined across frequency in an energy detector, the (weak-signal) ABF estimation-

error loss in detection currency only depends on N and K,

$$\mathrm{EEL}_K \approx -10\log_{10} r \text{ for } s \ll 1 \tag{119}$$

$$= 10\log_{10}\left\{\frac{\alpha+\beta-1}{\alpha-1}\right\} = 10\log_{10}\left\{\frac{K}{K-N+1}\right\} \quad \text{[units: dB]}. \tag{120}$$

For many scenarios of interest, this simple approximation will be adequate.

Use of the exponential model is justified when $K \gg N$, because the ACM estimate is close to the true value and the beam response distribution tends toward an exponential. However, the accuracy of the JDC obtained from the exponential model for values of K close to N is surprising, because the noisy ACM estimate induces heavier tails in the beam response distribution. Using (92), the values of the scintillation index (SI) are ≈ 2.9, 1.2, 1.03, and 1.003 under H_0 for $N = 32$ and the four values of K considered, in increasing order. In particular, the $K = N + 1 = 33$ (blue line) case with SI ≈ 2.9 is significantly heavier tailed than the exponential distribution. At low SINR, the SI under H_1 will be similarly high. The results seen in Fig. 3 can only be explained by the distributions under the two hypotheses becoming heavier-tailed in such a way that the JDC stays approximately the same as the exponential model.

As SINR increases, the SI under H_1 tends down toward one and the exponential model over-predicts performance because it is not accounting for the heavier-tailed distribution under H_0. Although these regimes require use of the GPD approximation, it is straightforward to evaluate as demonstrated by the MATLAB® code in App. A.1. The importance of using the GPD approximation is seen in Fig. 3 by the increasing disparity between the ideal (dashed lines) and the achieved JDC as SINR increases above a certain point. EEL initially decreases with SINR because JDC for the clairvoyant-MVDR adaptive beamformer transitions into the regime where $J \approx s$ earlier than when the ACM is estimated. However, this changes when SINR reaches $\approx 10 - 5\log_{10} r$ [units: dB], which is noted on Fig. 3 by the circles. Above this point, EEL begins to increase with SINR. If maintaining a constant EEL with SINR is important, K needs to increase not only with N, but also with SINR when it is above this breakpoint.

The scintillation-index discussion in Sect. 4.2 (e.g., see (93)) indicated that SI decreases toward one with N when K is proportional to N (i.e., $c = N/K$ is held constant). This raises a concern that the GPD model may not be adequate at the small values of N that will be important in the analysis of the reduced-adaptivity processors mentioned in Sect. 4.1.1. To assess this, the fit of the GPD approximation is evaluated in Fig. 4 as a function of N with $K = 2N - 2$. This value of K was chosen so $E[1/\rho] = 1/r = 2$, which produces a 3 dB loss in the signal-to-*average*-interference-and-noise power ratio (i.e., using the average interference-and-noise power to form an SINR) and a constant value of JDC as a function of N in the exponential model (dotted lines). As seen in the figure, the exponential model of EEL from (118) is accurate when N is large or at low SINR. The over-prediction of performance at higher SINR (comparing the dotted lines to the black dots) seen in Fig. 3 is seen in Fig. 4 to be exacerbated by a reduction in N. Fortunately, the GPD model (solid lines) provides an accurate approximation at low SINR and only exhibits the previously seen small over-prediction at higher values of SINR (also see footnote 6 on pg. 23).

Returning to the choice of $K = 2N - 2$, this analysis suggests that the signal-to-average-interference-and-noise power ratio ($rs = s/E[1/\rho]$) is a better predictor of detection performance than the average SINR (i.e., $E[\rho s]$) when N is small. Setting $K = 2N - 3$ to elicit a constant $E[\rho] = 1/2$ with N and a 3 dB loss in the average SINR uses one less snapshot to estimate the

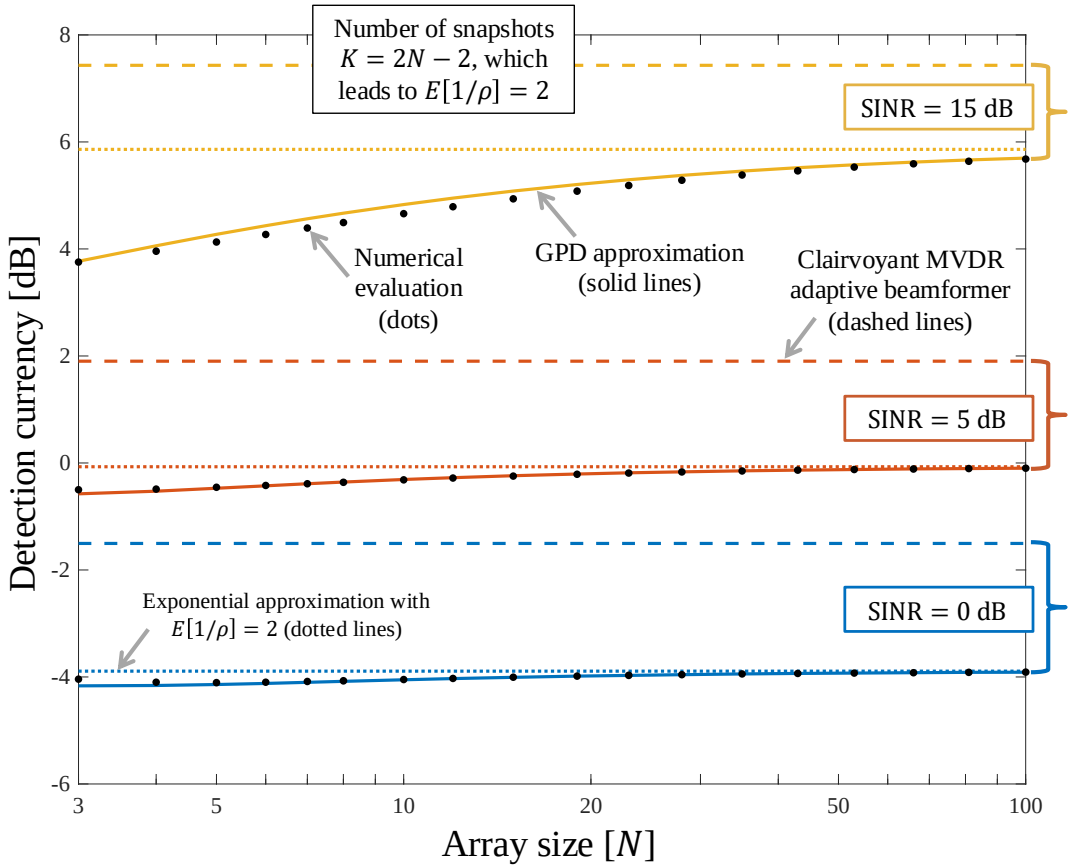

Figure 4: Detection currency as a function of N for various SINR when estimating the array covariance matrix with $K = 2N - 2$ snapshots.

ACM than keeping $E[1/\rho] = 2$ constant. As such the performance is below that seen in Fig. 4 for $K = 2N - 2$, implying a larger difference from constant performance with N. Although not shown in the figure, the disparity was observed to be worst when N and SINR were small and decrease as either variable increased.

5.2 Weak-signal ABF estimation-error loss in detection currency

When the beam-output SINR is below about 0 dB, the results seen in Fig. 3 suggest a single parameter (r) controls the ABF EEL through the model in (118). In the exponential-distribution approximation, setting r to one over the average noise power produced an accurate approximation to JDC when $K > N$ for these weak signals. However, it was not exact and produced larger errors when $K = N$. A more formal approach to obtaining r is pursued here to describe the weak-signal ABF EEL in detection currency in terms of the PDF on the SINR loss ρ and to address the case of $K = N$, where the other approximations failed.

For intensity-based decision statistics, the PDF of T when signal is present can be approximated by a first-order Maclaurin series,

$$f_1(t) \approx f_0(t) + sg(t), \tag{121}$$

where

$$g(t) = \frac{\partial}{\partial s} f_1(t)\bigg|_{s=0} \tag{122}$$

is the derivative of $f_1(t)$ with respect to s, evaluated at $s = 0$. The J-divergence in this scenario is then

$$J = \int [f_1(t) - f_0(t)] \log\left[\frac{f_1(t)}{f_0(t)}\right] \mathrm{d}t \approx s^2 \int \frac{g^2(t)}{f_0(t)} \mathrm{d}t, \tag{123}$$

which is approximately proportional to s^2 (using $\log(1+\epsilon) \approx \epsilon$ for $|\epsilon| \ll 1$). From this, it can be seen that the parameter controlling performance is

$$r = \sqrt{\int \frac{g^2(t)}{f_0(t)} \mathrm{d}t} \quad \text{[unitless]}. \tag{124}$$

Starting with (94), it can be shown that

$$g(t) = E\big[\rho^2(\rho t - 1)e^{-t\rho}\big] = tE\big[\rho^3 e^{-t\rho}\big] - E\big[\rho^2 e^{-t\rho}\big], \tag{125}$$

where the expectations are over ρ. Combined with the first definition of $f_0(t)$ in (96), this leads to

$$r = \sqrt{\int \frac{\{tE[\rho^3 e^{-t\rho}] - E[\rho^2 e^{-t\rho}]\}^2}{E[\rho e^{-\rho t}]} \mathrm{d}t} \quad \text{[unitless]}. \tag{126}$$

Although each of the expectations within (126) can be described in terms of a Kummer function, a numerical evaluation of the integrals formed by the expectations over ρ in the numerator and denominator may present a less complicated approach to obtaining r.

This formulation of the weak-signal ABF EEL parameter provides a means for assessing generic distributions on ρ and allows development of a more accurate approximation when $K = N$ in the SMI-MVDR processor, where (116) sets $r = 1/N$, but signficantly under-predicts performance. An empirical analysis of (126) for this scenario with $N \in [2, 100]$ led to the approximation

$$r \approx \frac{1.4\Big(1 - e^{-1.1N^{0.44}}\Big)}{N}, \tag{127}$$

which is greater than $1/N$ for $N \geq 2$, implying better performance than that predicted by the exponential-distribution approximation. The JDC for several values of SINR is shown in Fig. 5 as a function of N using (127) (solid lines), from a numerical evaluation of (126) (circles), and for the exponential-distribution approximation (dotted lines). The accuracy of (127) is evident, subject to SINR being small enough or N being large enough.

5.3 Approximating detection currency in an energy detector

When beam responses are combined incoherently over M snapshots in an energy detector, the same ACM estimate can be used in the MVDR beamforming filter vectors if the interference and

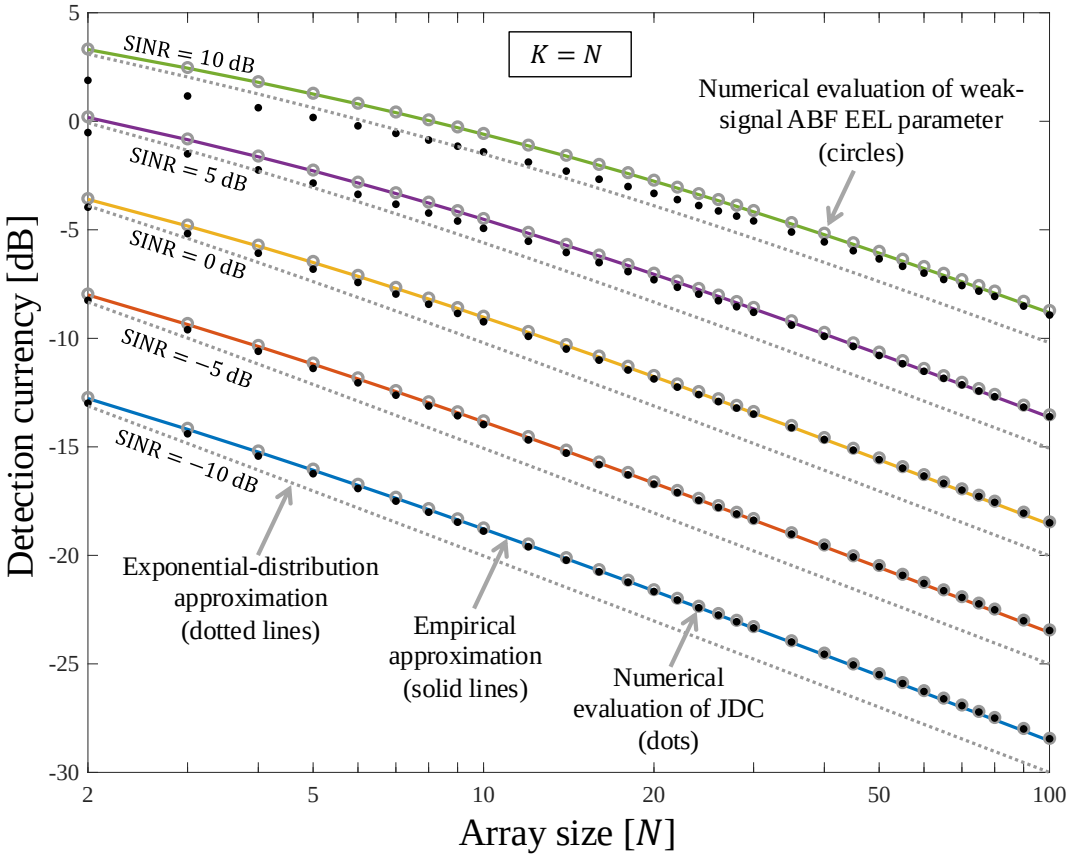

Figure 5: Detection currency as a function of array size when $K = N$ in the SMI-MVDR processor for various values of beam-output SINR.

noise background are the same for each snapshot. Conditioned on the ACM estimate, the normalized decision statistic,

$$T = \mathbf{d}_{\mathrm{s}}^H \mathbf{Q} \mathbf{d}_{\mathrm{s}} \sum_{m=1}^{M} |\mathbf{w}^H \mathbf{x}_m|^2 \quad \text{[unitless]} \tag{128}$$

$$\sim \text{Gamma}\big(M, \rho^{-1} + s\big)\,, \tag{129}$$

is gamma distributed with shape parameter M and scale parameter $\rho^{-1} + s$. The number of snapshots combined in the energy detector (M) is the time-bandwidth product and s is the SINR for a single snapshot, which is also assumed to be constant. The PDF of T under H_1 is a straightforward extension of (94) to an expectation over ρ of the gamma PDF,

$$f_1(t) = E_\rho \left[\frac{t^{M-1} e^{-t/(s+1/\rho)}}{\Gamma(M)(s + 1/\rho)^M} \right]. \tag{130}$$

The kth moment of T under H_1 is similarly obtained by generalizing (87) to

$$E_1[T^k] = E_\rho[E_1[T|\rho]] = E_\rho \left[\left(\frac{1}{\rho} + s \right)^k \right] \frac{\Gamma(M+k)}{\Gamma(M)}. \tag{131}$$

The PDF and kth moment under H_0 are then obtained by setting $s = 0$ in (130) and (131), respectively.

To account for integrating the intensities, the GPD model used to approximate JDC must be generalized to a scaled-F distribution, which is described in Sect. 6. The JDC between two scaled-F distributions is obtained by adding their Kullback-Leibler divergences, which are presented in Sect. 6.4 and combined to form J-divergence in Sect. 6.5. The parameters of the scaled-F approximations to the distribution of T under H_0 and H_1, obtained through moment matching, are presented in Sect. 6.2. MATLAB® code for evaluating JDC in this scenario can be found in App. A.2.

In Fig. 6, the scaled-F approximation to JDC (solid lines) is compared to a numerical evaluation of the integrals (dots) and a gamma-distribution approximation (dotted lines) for an array of $N = 32$ sensors and a beam-output SINR of $\approx 2.1\,\mathrm{dB}$, which yields JDC $= 0\,\mathrm{dB}$ when $M = 1$ in the clairvoyant-MVDR processor. As seen in the figure, the scaled-F approximation is significantly better than the gamma-distribution approximation, although the latter can be useful when $K \gg N$ and M is not too large. The gamma-distribution approximation simply adds $5\log_{10} M$ to the JDC obtained for a single intensity sample in (117).

Note that the example shown in Fig. 6 sets the number of snapshots used to estimate the ACM to

$$K = C\Big(N - 1 + M^{0.7-C/4}\Big)\,, \tag{132}$$

where $C \geq 1$. When $M = 1$, this term simplifies to $C = 1/c = K/N$. In the example presented here K increases with both N and M in such a way that the JDC loss is approximately constant with M. Although this is only expected to apply for $N = 32$ and this particular SINR, it illustrates that the scaled-F approximation is accurate when C is above about 1.25 and K is increased to keep a constant ABF EEL. Because the SINR was set to achieve 0 dB of JDC in the clairvoyant-MVDR processor when $M = 1$, the ABF EEL in JDC is the negative of the intercept point on the ordinate when $M = 1$. For example, using $C = 2$ causes a loss in JDC of slightly more than 2 dB for $M = 1$. The 2 dB ABF EEL is then approximately maintained as M increases when K is set according to (132).

5.4 Example: Combining beams containing correlated signals

The analyses presented in Sects. 4 and 5.1–5.3 were restricted to signals arriving from a single direction. The multipath scenario considered in Sect. 3 is evaluated here when the ACM is estimated. Unfortunately, it is not clear how to proceed with a theoretical analysis of this scenario where the same ACM estimate is used to form multiple beams from the same test snapshot and then combine them to form a decision statistic. The simplest approximation might be found in evaluating the beam-output SINR for each path, obtaining the J-divergence while accounting for ACM estimation as described in Sect. 5.1, and adding them under the assumption that they are independent. However, this is unlikely to be accurate when there is correlation between the paths or between the beams, as induced by use of a common ACM estimate.

As an initial alternative, a Monte-Carlo approach is used where the decision statistic is the weighted sum of the beam intensities pointing to the two signal paths. The weighting is chosen (clairvoyantly) as one over the beam-output interference-and-noise power. This accounts for the

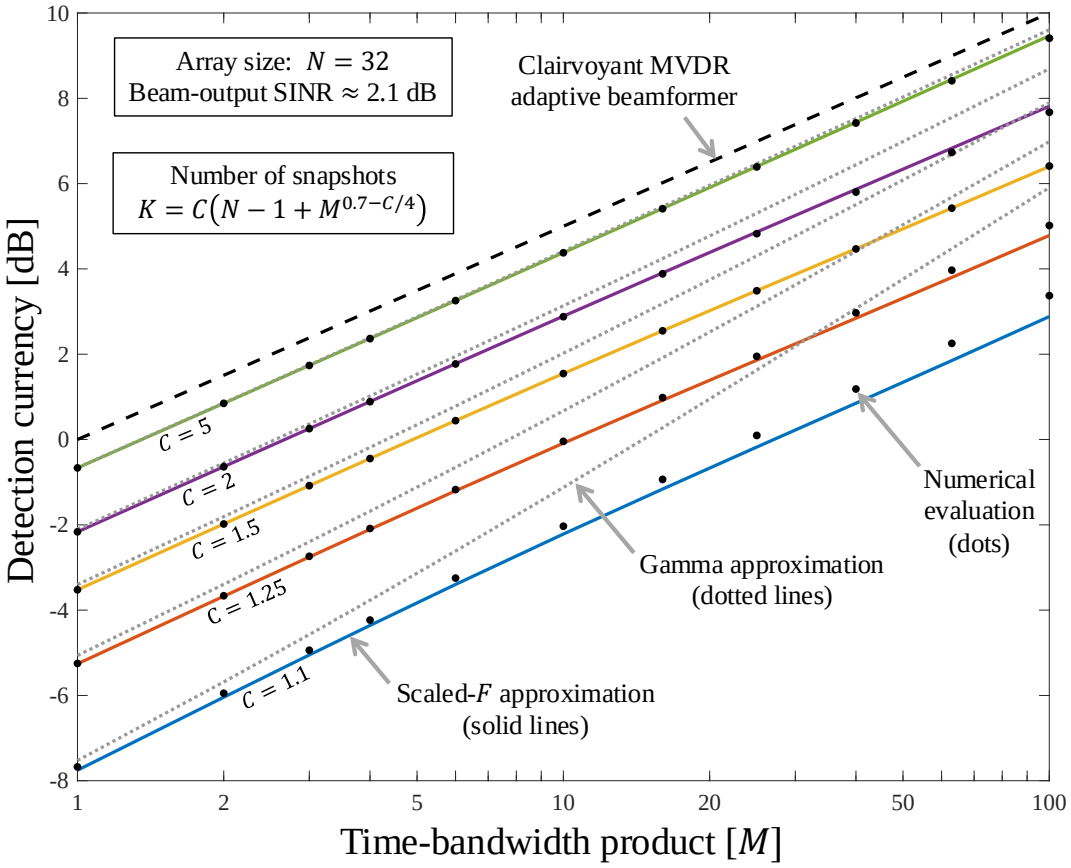

Figure 6: Detection currency as a function of the time-bandwidth product of an energy detector after using the same ACM estimate in an SMI-MVDR adaptive beamformer.

local background in which each signal path is observed and is similar to summing estimates of SINR. Although this may not be the optimal processor, it is a practical one that would typically be achieved by application of a (temporal) background-power-level normalizer. Assuming this interference-and-noise power is known limits the losses in this analysis to ACM estimation and correlation between the beams.

In the Monte-Carlo analysis, the decision statistic is obtained under the two detection hypotheses for 10^4 trials and fit to a scaled-F distribution using the first three moments, as described in Sect. 6.2. An example fit is shown in Fig. 7 for the test scenario of $N = 32$ and $K = 2N = 64$ with the secondary path at $Nu/2 = 4$ and a path-to-path correlation of $\rho = 0.5$. The accuracy of the approximation is evident visually and supported by Kolmogorov-Smirnov p-values of 0.98 and 0.81 under H_0 and H_1, respectively, when using the estimated parameters. Although this is not a comprehensive analysis, it provides some support for using the scaled-F distribution in the modeling.

The JDC for this decision statistic can then be approximated using the scaled-F model as described in Sect. 6.5 with the parameters estimated from the Monte-Carlo trials. The result is shown in Fig. 8 (reddish-brown lines with dots) for $N = 32$ with $K = 2N = 64$ and $K = 5N = 160$. The loss from having to estimate the ACM is not insignificant, as was seen in previous sections. In

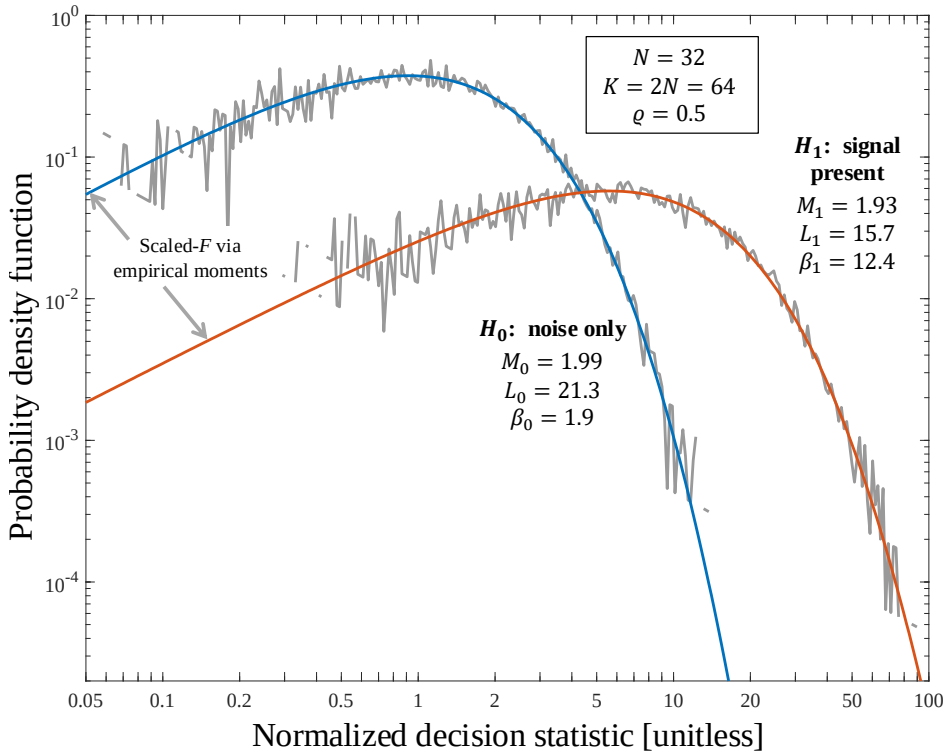

Figure 7: Histogram PDF estimates of the multi-beam decision statistic for detecting correlated signals under H_0 and H_1 along with scaled-F distributions using parameters estimated from the data.

comparison to using a Taylor-shaded CBF processor (violet line and 'x' marks), the SMI-MVDR processor is only seen to perform better when the secondary path is in the vicinity of the interference and for the $K = 5N$ case, which might be considered using a high-quality estimate of the ACM. The JDC for the CBF beam-intensity-sum was evaluated analytically using the results of Sect. 3.2.2 and also approximated ('x' marks) by fitting the decision statistics under H_0 and H_1 to gamma distributions.

Recalling from Sect. 1.1 that 5 dB of JDC represents an MDL operating point, it can be seen that using $K = 2N$ leads to performance below MDL for all cases except when the secondary path is coincident with the primary path (and this case reduces a nearly medium-quality operating point in the clairvoyant processor to a low-quality one when the ACM is estimated). Although using $K = 5N$ maintains an MDL operating point for regions where the interference does not impact the secondary path, the CBF with Taylor shading still performs better in these quiet regions. Near the interference, this assessment reverses with the CBF processor performing worse, which would also be the case if both signal paths were near a strong interferer.

Although the Monte-Carlo analysis provides an accurate assessment of the detection performance achievable when combining correlated multipath, it is computationally intensive enough to warrant exploring analytical approximations. First consider using the JDC achieved by the clairvoyant-MVDR adaptive beamformer and subtracting the basic ABF EEL from (120). As seen in Fig. 8 (triangles), this is very accurate for $K = 5N$, but slightly underestimates performance for $K = 2N$. Given the simplicity of the evaluation, however, it is a very useful initial approximation.

A more complicated alternative approach can be found by combining different parts of the clairvoyant-MVDR and the SMI-MVDR modeling to obtain the parameters of a scaled-F distribution, from which JDC for the normalized beam intensity sum can then be evaluated. The details are presented in Sect. 5.5 and the result seen in Fig. 8 (dotted lines). Although this approach is more accurate when $K = 2N$, it carries a caveat discussed at the end of Sect. 5.5 that it does not fully account for the correlation induced by using the same ACM estimate when pointed to different directions with the same test snapshot. It required an additional loss factor (0.35 dB in Fig. 8), which is small enough to suggest the basic approach will be useful in some scenarios. The accuracy might be improved by an analysis of the correlation between the SINR loss (i.e., the beta-distributed ρ) for beam responses formed from the same ACM estimate and the same test snapshot.

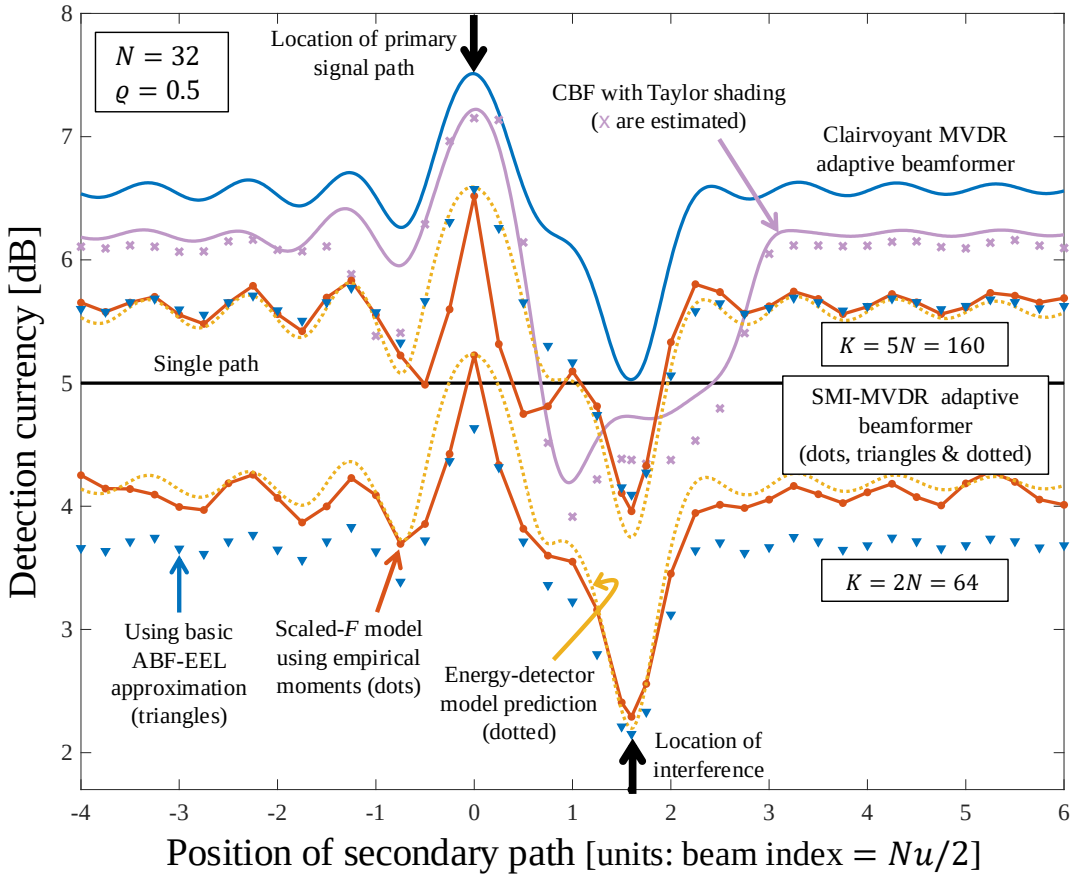

Figure 8: Detection currency when summing normalized beam intensities across two correlated multipath arrivals.

It is interesting to note that the Monte-Carlo and model-based approximations differ when the secondary path is near the primary path. Because both approaches involve approximations, it is not clear which should be believed. Examination of the Kolmogorov-Smirnov p-values in these areas revealed a weaker fit of the scaled-F model for at least one of the distributions, which suggests that the model-based approximations are more accurate in these regions.

5.5 Approximating performance for an incoherent multipath combiner

Suppose the detector decision statistic is a weighted sum of M beam intensities formed by applying the filter vectors $\mathbf{w}_1,\ldots,\mathbf{w}_M$ to the same test snapshot $\mathbf{x}$,

$$T = \sum_{m=1}^{M} b_m |\mathbf{w}_m^H \mathbf{x}|^2 = \mathbf{x}^H \mathbf{C} \mathbf{x} \quad \text{[unitless]}, \tag{133}$$

where the weights are assumed to be non-negative (i.e., $b_m \geq 0$) and crafted so T is unitless. This can be described as a matrix-vector quadratic form in $\mathbf{C}$ and $\mathbf{x}$ where

$$\mathbf{C} = \sum_{m=1}^{M} b_m \mathbf{w}_m \mathbf{w}_m^H \quad \text{[units: 1/power]}. \tag{134}$$

In this analysis, the weights are set to normalize each beam intensity to have unit mean under H_0. For the clairvoyant-MVDR adaptive beamformer, this results in

$$b_m = \mathbf{d}_m^H \mathbf{Q}^{-1} \mathbf{d}_m \quad \text{[units: 1/power]} \tag{135}$$

and

$$\mathbf{C} = \mathbf{Q}^{-1} \left(\sum_{m=1}^{M} \frac{\mathbf{d}_m \mathbf{d}_m^H}{\mathbf{d}_m^H \mathbf{Q}^{-1} \mathbf{d}_m} \right) \mathbf{Q}^{-1} \quad \text{[units: 1/power]}. \tag{136}$$

When $\mathbf{C}$ is deterministic, the quadratic form in (133) can usually be described as a mixture of exponential random variables. Here, however, it will be approximated by a gamma distribution, which only requires evaluation of its first two moments.

Suppose the test snapshot $\mathbf{x} \sim \mathcal{CN}(\mathbf{0}, \mathbf{R})$ as is the case under H_1. It is then straightforward to show that

$$E[T] = \mathrm{tr}\{\mathbf{RC}\} \quad \text{[unitless]} \tag{137}$$

and

$$E[T^2] = \mathrm{tr}^2\{\mathbf{RC}\} + \mathrm{tr}\{\mathbf{RCRC}\} \quad \text{[unitless]}, \tag{138}$$

where the latter is a simplification of the result shown in [16, pg. 564]. Under H_0, $\mathbf{R}$ in (137) and (138) is replaced by $\mathbf{Q}$. Using these to match moments with a gamma distribution produces

$$\alpha_0 = \frac{\mathrm{tr}^2\{\mathbf{QC}\}}{\mathrm{tr}\{\mathbf{QCQC}\}} \text{ and } \alpha_1 = \frac{\mathrm{tr}^2\{\mathbf{RC}\}}{\mathrm{tr}\{\mathbf{RCRC}\}} \tag{139}$$

for shape parameters and

$$\beta_0 = \frac{\mathrm{tr}\{\mathbf{QCQC}\}}{\mathrm{tr}\{\mathbf{QC}\}} \text{ and } \beta_1 = \frac{\mathrm{tr}\{\mathbf{RCRC}\}}{\mathrm{tr}\{\mathbf{RC}\}} \tag{140}$$

for scale parameters. At this point, it is assumed that $T \sim \mathrm{Gamma}(\alpha_i, \beta_i)$ under hypothesis H_i. Using the definition of $\mathbf{C}$ in (136), simplifications can be obtained for some of these parameters. For example, it can be shown that

$$\alpha_0 = \frac{M}{1 + \frac{1}{M} \sum \sum_{m \neq n} \frac{|\mathbf{d}_m^H \mathbf{Q}^{-1} \mathbf{d}_n|^2}{\mathbf{d}_m^H \mathbf{Q}^{-1} \mathbf{d}_m \cdot \mathbf{d}_n^H \mathbf{Q}^{-1} \mathbf{d}_n}}. \tag{141}$$

This demonstrates that α_0 can be as large as M when the beam responses are mutually independent (i.e., $\mathbf{d}_m^H \mathbf{Q}^{-1} \mathbf{d}_n = 0$ for $m \neq n$) and be as low as one when they are fully correlated (i.e., $\mathbf{d}_m = \mathbf{d}_n$ for all m and n). This is important because the gamma model accounts for the correlation between the beam responses induced by proximity of the signal paths.

The next step in this approximation is to assume the covariance matrix is estimated and describe T in the same way as the energy detector decision statistic was in Sect. 5.3. Dividing T by β_0 gives the result unit-scale under H_0. Replacing this unit scale with $1/\rho$ results in the distributions

$$\tilde{T} = \frac{T}{\beta_0} \sim \begin{cases} \text{Gamma}\big(\alpha_0, \rho^{-1}\big) & \text{under } H_0 \\ \text{Gamma}\big(\alpha_1, \rho^{-1} + s\big) & \text{under } H_1, \end{cases} \tag{142}$$

where the beam-output SINR (per independent intensity sample) is assumed to be

$$s \approx \frac{\beta_1}{\beta_0} - 1 = \frac{\text{tr}\{\mathbf{RCRC}\}}{\text{tr}\{\mathbf{RC}\}} \frac{\text{tr}\{\mathbf{QC}\}}{\text{tr}\{\mathbf{QCQC}\}} - 1. \tag{143}$$

This is now in the same form as the energy-detector results and (131) can be used to obtain the moments of $\tilde{T}$ under H_0 and H_1 with M set to α_i under hypothesis H_i, because it can differ across the hypotheses in this scenario. The first two moments of $\tilde{T}$ are then used to obtain the parameters of the scaled-F distribution under the two hypotheses following Sect. 6.2, setting $M_i = \alpha_i$ and obtaining L_i and λ_i for $i = 0$ and 1.

The final step is to evaluate JDC as described in Sect. 6.5 using the scaled-F parameters under H_0 and H_1. Unfortunately, this slightly overestimates the performance because the energy-detector model, although it handles a common ACM estimate, assumes there are M independent snapshots being combined in the sum. In this scenario, however, the beam intensities being summed are all from the same test snapshot; they come from different steering directions. The gamma approximations in (139) account for the correlation induced by the steering vectors, but they do not account for correlation induced by a shared ACM estimate. The effect of unaccounted correlation among the beam responses can be approximated by an additional loss in JDC. For example, the results presented in Fig. 8 were obtained by adding $5 \log_{10} 0.85 = -0.35\,\text{dB}$ to the scaled-F prediction (reducing M_0 and M_1 by a factor of 0.85 has a similar effect). This adjustment seemed to work well when $K = 2N$ for a few different values of N. However, it should depend on both K and N. A proper accounting of beam-to-beam correlation induced by estimation of the ACM is beyond the scope of this report.

6 The scaled-F distribution

The standard F distribution (see [8, Sect. 5.6.16]) has two degree-of-freedom (i.e., shape) parameters (ν_1 & ν_2) and does not contain a scale parameter. The parameterization employed here is obtained by multiplying a standard F-distributed random variable by λ and using shape parameters $M = \nu_1/2$ and $L = \nu_2/2$. The use of M and L over the traditional shape parameters comes from cell-averaging CFAR processors, which produce F-distributed random variables with $\nu_1 = 2M$ and $\nu_2 = 2L$ when M independent integrated intensities are normalized by a power estimate formed from L independent auxiliary data samples. This section contains the properties of the scaled-F distribution required to apply it as an approximation to an energy detector (i.e., integrated-intensity detector) formed from the beam responses of an SMI-MVDR processor using a common ACM estimate.

Readers who do not have immediate need for these details may wish to skip to the conclusions presented in Sect. 7.

6.1 Moments and distribution functions

The PDF of the scaled-F distribution with the (M, L, λ) parameterization is

$$f_{\text{sf}}(t; M, L, \lambda) = \frac{t^{M-1}\left[\frac{M}{L\lambda}\right]^M}{B(M,L)\left[1+\frac{Mt}{L\lambda}\right]^{M+L}} \quad \text{for} \quad t \geq 0, \tag{144}$$

where $B(M, L) = \Gamma(M)\Gamma(L)/\Gamma(M + L)$ is the beta function. The kth moment of the scaled-F distribution is simply λ^k times the kth moment of the standard F distribution (see [8, Sect. 5.6.16]),

$$\mu'_k = E\left[T^k\right] = \left[\frac{\lambda L}{M}\right]^k \frac{\Gamma(M+k)\Gamma(L-k)}{\Gamma(M)\Gamma(L)} \quad \text{for} \quad k < L. \tag{145}$$

As defined here, the scaled-F distribution simplifies to the standard F distribution when $\lambda = 1$, $M = \nu_1/2$ and $L = \nu_2/2$. The ratio-of-chi-squared-variates description of the standard F distribution can be used to characterize the scaled-F distribution through

$$T = \frac{\lambda G_M}{G_L}, \tag{146}$$

which is λ times the ratio (G_M/G_L) of two independent, unit-mean, gamma-distributed random variables with shape parameters M and L (i.e., $G_i \sim \text{Gamma}(i, 1/i)$). The GPD model arises when $M = 1$ and $L = 1/\gamma$, which suggests the scaled-F distribution might be useful in representing integrated intensities having a power-law decay in the tails of their PDF.

The cumulative distribution function (CDF) of a scaled-F distribution,

$$F_{\text{sf}}(t; M, L, \lambda) = \int_0^t f_{\text{sf}}(x; M, L, \lambda)\,\mathrm{d}x = \frac{1}{B(M,L)} \int_0^{\frac{Mt}{L\lambda}} \frac{u^{M-1}}{(1+u)^{M+L}}\,\mathrm{d}u \tag{147}$$

$$= \frac{B\left(M, L, \frac{Mt}{Mt+L\lambda}\right)}{B(M,L)} \quad \text{for} \quad t \geq 0, \tag{148}$$

can be described in terms of an incomplete beta function, $B(u, v, x)$, through use of the integral definition from [13, pg. 604, eq. 58:3:2].[7]

If $f_F(x; \nu_1, \nu_2)$ and $F_F(x; \nu_1, \nu_2)$ are, respectively, the PDF and CDF of the standard F distribution, then the scaled-F PDF is

$$f_{\text{sf}}(t; M, L, \lambda) = f_F(t/\lambda; 2M, 2L)/\lambda \tag{149}$$

and the CDF is

$$F_{\text{sf}}(t; M, L, \lambda) = F_F(t/\lambda; 2M, 2L). \tag{150}$$

[7] The function `betainc` in MATLAB® implements the version of the incomplete beta function normalized by $B(u, v)$, which directly yields the CDF.

6.2 Method-of-moments estimators

Method-of-moments estimators use the moments of a test distribution to obtain the parameters of a model distribution, by which the test distribution is to be approximated. This requires the solution of P moment equations (that are typically non-linear) when there are P unknown parameters in the model distribution. Although it is common to use the first P moments, fractional lower-order moments can be useful when the test-distribution moments are estimated from data and higher-order moments will produce a better fit in the tails of the distribution.

Let μ'_k be the kth moment of the test distribution; here using (131) in order to approximate the distribution of T. Define the moment ratio $\rho_k = \mu'_k/\mu'_{k-1}$ and initialize the recursion with $\rho_1 = \mu'_1$. Applying the method of moments approach to parameter estimation using the first three moments of the scaled-F distribution then yields

$$M_{\text{sf}} = \frac{2\rho_1(\rho_3 - \rho_2)}{\rho_1\rho_2 - 2\rho_1\rho_3 + \rho_2\rho_3}, \tag{151}$$

$$L_{\text{sf}} = \frac{M_{\text{sf}}(2\rho_2 - \rho_1) - \rho_1}{M_{\text{sf}}(\rho_2 - \rho_1) - \rho_1}, \text{ and} \tag{152}$$

$$\lambda_{\text{sf}} = \frac{\rho_1(L_{\text{sf}} - 1)}{L_{\text{sf}}}. \tag{153}$$

Note that the scaled-F parameters obtained here are denoted M_{sf}, L_{sf}, and λ_{sf} to distinguish them from the parameters of the test distribution (e.g., M_{sf} is not necessarily the same as the M found in the moments of T in (131)).

When the test distribution does not conform to the model distribution, the moment equations can be non-invertible (which is not the case here) or produce parameters that violate the conditions of the model distribution. For example, the scaled-F distribution requires M_{sf}, L_{sf}, and λ_{sf} to be positive and for $L_{\text{sf}} > k$ in order for the kth moment to exist. The method-of-moments approach fails when the solutions to (151)–(153) violate any of these conditions. In applying this approach to approximating the distribution of T for the integrated-intensity SMI-MVDR processor, the solutions were sometimes observed to violate the assumptions, particularly when M was large, K was not very large, or at very high SINR, which suggests the scaled-F distribution is not necessarily a good model in these scenarios.

An alternative approach that always provided valid solutions was found by forcing M_{sf} to equal the time-bandwidth product of the energy detector (i.e., setting $M_{\text{sf}} = M$). The remaining parameters of the scaled-F distribution are then obtained from (152) and (153), which only require use of the first two moments. Although this resulted in a slightly degraded fit when K was small, as can be seen in Fig. 6, it was otherwise very accurate.

For completeness, if L_{sf} is known and exceeds 2, the solutions for λ_{sf} and M_{sf} formed from the first two moments are

$$\lambda_{\text{sf}} = \frac{\mu'_1(L_{\text{sf}} - 1)}{L_{\text{sf}}} \text{ and } M_{\text{sf}} = \left[\frac{(L_{\text{sf}} - 2)\mu'_2}{(L_{\text{sf}} - 1)[\mu'_1]^2} - 1\right]^{-1}. \tag{154}$$

6.3 Maximum-likelihood estimator

The Monte-Carlo analysis presented in Sect. 5.4 produced data from which the parameters of the scaled-F distribution could be estimated. Although the method of moments could be applied as

described in Sect. 6.2, maximum-likelihood estimators were obtained through use of a generic multivariable function optimization routine. This allowed enforcing the requirements for the parameters to be positive by submitting their logarithms to the unconstrained optimization routine (i.e., finding the values of $\log M$, $\log L$, and $\log \lambda$ that maximize the log-likelihood function).

Given the proximity of the scaled-F distribution to a gamma distribution when L is large, the numerical optimization was initialized using the maximum likelihood estimates for the gamma model with an arbitrary value of L. MATLAB® code implementing the estimation can be found in App. B.2.

6.4 Kullback-Leibler divergences between two scaled-F distributions

The Kullback-Leibler divergence [17] from $f_1(t)$ to $f_0(t)$ is defined by the integral

$$I_{1:0} = \int f_1(t) \log\left[\frac{f_1(t)}{f_0(t)}\right] \mathrm{d}t = E_1[l_{1:0}(T)]\,, \tag{155}$$

where the two functions are probability density functions. It can also be described as the expected value of the log-likelihood ratio, $l_{1:0}(t) = \log\{f_1(t)/f_0(t)\}$, evaluated at T. The subscript on the expectation operator in (155) indicates the distribution under which it is performed and the subscript on the LLR function makes explicit which PDF is in the numerator and which is in the denominator. To simplify derivation of $I_{1:0}$, let the scaled-F PDF in (144) be defined under hypothesis H_i with parameters M_i, L_i, and λ_i for $i = 0$ and 1 according to

$$f_i(t) = \frac{t^{M_i-1} c_i^{M_i}}{B(M_i, L_i)[1 + c_i t]^{M_i+L_i}} \quad \text{for} \quad t \geq 0, \tag{156}$$

where $c_i = M_i/(L_i\lambda_i)$. The log-likelihood ratio between these scaled-F distributions is then

$$l_{1:0}(t) = \log\left[\frac{f_1(t)}{f_0(t)}\right] = (M_1 - M_0)\log t - (M_1 + L_1)\log(1 + c_1 t) + (M_0 + L_0)\log(1 + c_0 t) + \mathcal{C}_{1:0}, \tag{157}$$

where the constant terms have been grouped into

$$\mathcal{C}_{1:0} = \log\left[\frac{c_1^{M_1} B(M_0, L_0)}{c_0^{M_0} B(M_1, L_1)}\right]. \tag{158}$$

Examining (157), it can be seen that the expected value of $\log T$ and $\log(1 + c_j T)$ will be required under hypothesis H_i for i and j equal to 0 and 1. The first of these can be described in terms of the digamma function, $\psi(\cdot)$,

$$E_i[\log T] = \frac{c_i^{M_i}}{B(M_i, L_i)} \int_0^\infty \frac{t^{M_i-1} \log t}{(1 + c_i t)^{M_i+L_i}} \mathrm{d}t \tag{159}$$

$$= \psi(M_i) - \psi(L_i) - \log c_i. \tag{160}$$

The result in (160) can be obtained by decomposing T into the ratio of two independent gamma-distributed random variables, as was done in (146), and applying [15, pg. 577, eq. 4.352-1]. Alternatively, a direct solution of the integral can proceed by noting that the beta function can be

described by

$$B(a, b-a) = \int_0^\infty \frac{t^{a-1}}{(1+t)^b}\,\mathrm{d}t \tag{161}$$

from [15, pg. 917, eq. 8.380-3] and then that

$$\tfrac{\partial}{\partial a}B(a, b-a) = \int_0^\infty \frac{t^{a-1}\log t}{(1+t)^b}\,\mathrm{d}t. \tag{162}$$

The digamma functions in (160) arise from derivatives of the gamma functions comprising $B(a, b-a)$.

The solution to the "auto" term (i.e., when $i = j$),

$$E_i[\log(1 + c_i T)] = \psi(M_i + L_i) - \psi(L_i), \tag{163}$$

is obtained using [15, pg. 563, eq. 4.293-14]. The cross terms where the expectation occurs under H_i and $c_j \neq c_i$ in $\log(1 + c_j T)$ require use of an auxiliary function, $G(z; M, L)$, to be described in Sect. 6.6,

$$E_i[\log(1 + c_j T)] = \frac{\Gamma(M_i + L_i)c_i^{M_i}}{\Gamma(M_i)\Gamma(L_i)} \int_0^\infty \frac{t^{M_i-1}\log(1 + c_j t)}{(1 + c_i t)^{M_i+L_i}}\,\mathrm{d}t \tag{164}$$

$$= \frac{1}{M_i + L_i} G\left(\frac{c_i}{c_j}; M_i, L_i\right). \tag{165}$$

Setting $c_i = c_j$ in (165) can then be seen to produce (163) through use of (173).

Combining these results into (155) yields the Kullback-Leibler divergence between two scaled-F distributions,

$$I_{i:j} = (M_i - M_j)\psi(M_i) - (M_i + L_i)\psi(M_i + L_i) + (M_j + L_i)\psi(L_i) + \frac{(M_j + L_j)}{(M_i + L_i)} G\left(\frac{c_i}{c_j}; M_i, L_i\right) + M_j \log\left(\frac{c_i}{c_j}\right) + \log\left[\frac{B(M_j, L_j)}{B(M_i, L_i)}\right]. \tag{166}$$

6.5 J-divergence between two scaled-F distributions

By examining the definition of the J-divergence in (1), it can be seen that it is simply the sum of the two Kullback-Leibler divergences,

$$J = I_{1:0} + I_{0:1}. \tag{167}$$

Using (166) and carefully tracking the signs of each term, this can be seen to result in

$$J = \frac{(M_0 + L_0)}{(M_1 + L_1)} G\left(\frac{c_1}{c_0}; M_1, L_1\right) + \frac{(M_1 + L_1)}{(M_0 + L_0)} G\left(\frac{c_0}{c_1}; M_0, L_0\right) + (M_1 - M_0)\left[\psi(M_1) - \psi(M_0) + \psi(L_0) - \psi(L_1) - \log\left(\frac{c_1}{c_0}\right)\right] - (M_1 + L_1)[\psi(M_1 + L_1) - \psi(L_1)] - (M_0 + L_0)[\psi(M_0 + L_0) - \psi(L_0)]. \tag{168}$$

MATLAB® code evaluating J in (168) can be found in App. B.1.

6.6 Auxiliary function $G(z; M, L)$ for integer values of M

In Sect. 6.4, the function

$$G(z; M, L) = \frac{\Gamma(M+L+1)}{\Gamma(M)\Gamma(L)} \int_0^\infty \frac{t^{M-1} \log(1+t/z)}{(1+t)^{M+L}} \, \mathrm{d}t \tag{169}$$

was required to derive the Kullback-Leibler divergences between two scaled-F distributions. When M is an integer, the integral can be solved by applying integration by parts with $u = t^{M-1} \log(1+t/z)$ and $\mathrm{d}v = \mathrm{d}t/(1+t)^{M+L}$, so $\int u \, \mathrm{d}v = uv - \int v \, \mathrm{d}u$. Part of the $\int v \, \mathrm{d}u$ term is then recognized as being proportional to $G(z; M-1, L)$, which produces the following recursive solution,

$$G(z; M, L) = \frac{(M+L)}{(M+L-1)} \left[G(z; M-1, L) + \frac{1}{B(M,L)} \int_0^\infty \frac{t^{M-1}}{(1+t)^{M+L-1}(z+t)} \, \mathrm{d}t \right]. \tag{170}$$

Using [15, pg. 319, eq. 3.197-1], the remaining integral within the brackets is seen to be related to a Gauss hypergeometric function (${}_2F_1(a, b; c; x)$), which yields the solution

$$G(z; M, L) = \frac{(M+L)}{(M+L-1)} \left[G(z; M-1, L) + z^{M-1} {}_2F_1(M+L-1, M; M+L; 1-z) \right]. \tag{171}$$

The recursion stops at $M = 1$, for which

$$G(z; 1, L) = \frac{(L+1)}{L} {}_2F_1(L, 1; L+1; 1-z). \tag{172}$$

Numerical routines for evaluating the Gauss hypergeometric function are readily available (e.g., see the MATLAB® code found in App. C).

When $z = 1$, the iterative form in (171) simplifies to

$$G(1; M, L) = \frac{\Gamma(M+L+1)}{\Gamma(M)\Gamma(L)} \int_0^\infty \frac{t^{M-1} \log(1+t)}{(1+t)^{M+L}} \, \mathrm{d}t = (M+L)[\psi(M+L) - \psi(L)], \tag{173}$$

which can also be obtained from [15, pg. 563, eq. 4.293-14]. When the digamma function is not readily available, [13, pg. 451, eq. 44:5:4] provides a simpler form for the difference

$$\psi(M+L) - \psi(L) = \sum_{m=0}^{M-1} \frac{1}{L+m}, \tag{174}$$

which is found in (173) and in the J-divergence in (168).

When $G(z; M, L)$ must be evaluated for a non-integer value of M, a linear interpolation between the function values for $\lfloor M \rfloor$ and $\lceil M \rceil$ was found to be adequate for the approximations to JDC presented in Sect. 5.3.

7 Conclusions

The primary focus of this report has been the analysis of J-divergence detection currency (JDC) before and after array processing. Building on the results of [2], in which J-divergence was first applied to array processing, the analysis has been extended here to account for correlated multipath signal arrivals, shaded conventional beamforming, and the use of estimated covariance matrices in adaptive beamformers. With respect to multipath, the results are intuitively explained by the correlation among the paths and their proximity to each other and any interferences. The results for conventional beamforming similarly correspond to that expected based on the lower SINR they achieve relative to an optimal beamformer.

When the SINR loss of an adaptive beamformer using an estimated array covariance matrix (ACM) follows a beta distribution, as it does for the sample-matrix-inversion (SMI) MVDR processor, the reduction in performance was found to be accurately approximated for weak signals by using an average SINR in the basic formula for the JDC of a Gaussian-fluctuating signal in Gaussian noise (i.e., JDC $= 5\log_{10}\{s^2/(1+s)\}$). At higher values of SINR, where this approximation begins to degrade, accurate approximations were obtained through the use of the generalized-Pareto distribution (GPD) for single intensity samples and a scaled-F distribution for integrated intensities. Although it has not been explored here, the GPD and scaled-F approximations to the SMI-MVDR beam-response intensity or integrated intensity may be useful in evaluating the conventional P_f and P_d performance metrics.

Similar to the traditional use of beam-output SINR as a performance metric, the results presented here using JDC illustrate that use of an adaptive beamformer can be beneficial when signals of interest are near in angle to a strong interference. More importantly, they provide an assessment of the potential gain obtained by combining multipath and the loss incurred when the ACM must be estimated in a metric that is easily combined across independent measurements, is more closely related to detection performance than the average SINR, and can be evaluated throughout the signal and information processing chain.

References

[1] H. Jeffreys, *Theory of Probability*, 3rd ed. Oxford University Press, 1948.

[2] L. L. Scharf and P. H. Moose, "Information measures and performance bounds for array processors," *IEEE Transactions on Information Theory*, vol. IT-22, no. 1, pp. 11–21, January 1976.

[3] D. A. Abraham, "Modeling sonar performance using J-divergence," in *Proceedings of the Underwater Acoustics Conference & Exhibition*, Kalamata, Greece, 2023.

[4] ——, "Parameter estimation and performance modeling in generalized-pareto-distributed clutter," University of Washington, Applied Physics Laboratory, 1013 NE 40th Street, Seattle, WA 98105, Tech. Rpt. TR 2401, January 2024. [Online]. Available: http://apl.uw.edu/tr2401

[5] I. S. Reed, J. D. Mallett, and L. E. Brennan, "Rapid convergence rate in adaptive arrays," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-10, no. 6, pp. 853–863, November 1974.

[6] D. H. Johnson and D. E. Dudgeon, *Array Signal Processing: Concepts and Techniques.* Englewood Cliffs, New Jersey: Prentice Hall, 1993.

[7] H. L. Van Trees, *Optimum Array Processing.* New York: John Wiley & Sons, Inc., 2002.

[8] D. A. Abraham, *Underwater Acoustic Signal Processing: Modeling, Detection, and Estimation.* Springer, 2019.

[9] R. J. Urick, *Principles of Underwater Sound*, 3rd ed. New York: McGraw-Hill, Inc., 1983.

[10] I. Kirsteins and D. W. Tufts, "Rapidly adaptive nulling of interference," in *IEEE 1989 International Conference on Systems Engineering*, Aug 1989, pp. 269–272.

[11] Y. I. Abramovich, "Convergence analysis of linearly constrained SMI and LSMI adaptive algorithms," in *Proceedings of the IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium (Cat. No.00EX373)*, 2000, pp. 255–259.

[12] K. E. Wage and J. R. Buck, "SINR loss of the dominant mode rejection beamformer," in *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2015, pp. 2499–2503.

[13] K. Oldham, J. Myland, and J. Spanier, *An Atlas of Functions*, 2nd ed. New York: Springer Science, 2009.

[14] C. Richmond, "Derived PDF of maximum likelihood signal estimator which employs an estimated noise covariance," *IEEE Transactions on Signal Processing*, vol. 44, no. 2, pp. 305–315, 1996.

[15] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, 8th ed. Waltham, MA: Elsevier Academic Press, 2015, edited by D. Zwillinger.

[16] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory.* Prentice Hall PTR, 1993.

[17] S. Kullback and R. A. Leibler, "On information and sufficiency," *Annals of Mathematical Statistics*, vol. 22, no. 1, pp. 79–86, 1951.

A MATLAB® code for the detection currency when using an SMI-MVDR adaptive beamformer

A.1 Single intensity sample

```
function JDC=jdc_smi_mvdr(Sdb,N,K)
% JDC=jdc_smi_mvdr(Sdb,N,K)
%  J-divergence detection currency (JDC) in decibels for sample-matrix-inversion (SMI)
%  minimum-variance-distortionless-response (MVDR) adaptive beamformer
%  when testing a single intensity sample with a Gaussian-fluctuating signal
%  in Gaussian interference and noise
% Parameters: [scalar]
%  Sdb = beam-output SINR in decibels
%  N = size of the array
%  K = number of snapshots used to estimate the array covariance matrix (accurate for K>N)
%
s=10^(Sdb/10);
alp=K-N+2; bet=N-1; % Beta parameters for SINR loss
% GPD parameters through moment matching
gam0=bet/(alp^2+alp*bet-3*alp+2);
lam0=(alp+bet-1)*(1-gam0)/(alp-1);
gam1=bet*(alp+bet-1)/(2*bet*(alp+bet-1)+(alp-2)*((alp-1)*s+alp+bet-1)^2);
lam1=(s+(alp+bet-1)/(alp-1))*(1-gam1);
% JDC using J-div for scaled-F distributions
JDC=5*log10(jd_sfsf(1,1/gam0,lam0,1,1/gam1,lam1));
```

A.2 Energy detector with a common ACM estimate

```
function JDC=jdc_smi_edet(Sdb,N,K,M)
% JDC=jdc_smi_edet(Sdb,N,K,M)
%  J-divergence detection currency (JDC) in decibels for sample-matrix-inversion (SMI)
%  minimum-variance-distortionless-response (MVDR) adaptive beamformer
%  when a single estimated array covariance matrix (ACM) is applied to an energy detector
%  combining multiple independent intensities containing a Gaussian-fluctuating
%  signal in Gaussian interference and noise
% Parameters: [scalar]
%  Sdb = beam-output SINR in decibels for a single intensity sample
%  N = size of the array
%  K = number of snapshots used to estimate the array covariance matrix (accurate for K>N)
%  M = number of intensity samples integrated in the energy detector
%
s=10^(Sdb/10);
alp=K-N+2; bet=N-1; % Beta parameters for SINR loss
% Inverse moments of rho
up1=(alp+bet-1)/(alp-1); up2=up1*(alp+bet-2)/(alp-2);
% Scaled-F parameters under H0
M0=M; r1=M*up1; r2=(M+1)*up2/up1;
L0=(M0*(2*r2-r1)-r1)/(M0*(r2-r1)-r1); lam0=r1*(L0-1)/L0;
% Scaled-F parameters under H1
M1=M; r1=M*(up1+s);
r2=(M+1)*(up2+2*s*up1+s^2)/(up1+s);
L1=(M1*(2*r2-r1)-r1)/(M1*(r2-r1)-r1); lam1=r1*(L1-1)/L1;
% JDC using J-div for scaled-F distributions
JDC=5*log10(jd_sfsf(M0,L0,lam0,M1,L1,lam1));
```

A.3 Numerical evaluation for a beta-distributed SINR loss factor

```
function JDC=jdc_sinr_beta_num(Sdb,alp,bet,M)
% JDC=jdc_sinr_beta_num(Sdb,alp,bet,M)
%  J-divergence detection currency (JDC) in decibels for a gamma-distributed intensity
%  sum with a beta-distributed SINR loss factor as obtained through numerical integration
% Parameters: [scalar]
%  Sdb = clairvoyant-MVDR beam-output SINR in decibels for a single intensity sample
%  (alp,bet) = shape parameters of the beta-distributed SINR
%  M = number of intensity samples integrated in the energy detector
%
if nargin<4, M=1; end; s=10..^(Sdb/10);
% Function definitions
fillnaninf=@(x) fillmissing(x.*isfinite(x),'constant',0);
if M==1,
  fcdf=@(p,x,M,s) 1-exp(-x*p./(1+p*s));
  ecdf=@(p,x,M,s) exp(-x*p./(1+p*s));
  logfpdf=@(p,x,M,s) log(p./(1+p*s))-x*p./(1+p*s);
else
  fcdf=@(p,x,M,s) gamcdf(x,M,1../p+s);
  ecdf=@(p,x,M,s) gamcdf(x,M,1../p+s,'upper');
  logfpdf=@(p,x,M,s) M.*log(x.*p./(1+p.*s))-p.*x./(1+p.*s)-log(x)-gammaln(M);
end;
logbetapdf=@(p,alp,bet) (alp-1).*log(p)+(bet-1).*log(1-p)-betaln(alp,bet);
ffun=@(x,M,s,alp,bet) integral(@(p) exp(logfpdf(p,x,M,s)+logbetapdf(p,alp,bet)),0,1);
Efun=@(x,M,s,alp,bet) integral(@(p) exp(log(ecdf(p,x,M,s))+logbetapdf(p,alp,bet)),0,1);
Ffun=@(x,M,s,alp,bet) integral(@(p) exp(log(fcdf(p,x,M,s))+logbetapdf(p,alp,bet)),0,1);
% Find limits on integral
e=1e-4;
x0=gaminv(e,M,1); F0=Ffun(x0,M,0,alp,bet);
while F0>e,
  x0=x0/1.2; F0=1-Efun(x0,M,0,alp,bet);
end;
x1=gaminv(1-1e-4,M,1+s); E1=Efun(x1,M,s,alp,bet);
while E1>e,
  x1=x1*1.2; E1=Efun(x1,M,s,alp,bet);
end;
% Binary component of J-divergence
E0=Efun(x1,M,0,alp,bet); F1=Ffun(x0,M,s,alp,bet);
Jb=(F0-F1)*(fillnaninf(log(F0))-fillnaninf(log(F1)))+...
   (E1-E0)*(fillnaninf(log(E1))-fillnaninf(log(E0)));
% Get PDF values every 0.25 dB
Nx=max(50,ceil(4*10*log10(x1/x0)));
x=logspace(log10(x0),log10(x1),Nx);
for i=1:Nx,
  f0(i)=ffun(x(i),M,0,alp,bet);
  f1(i)=ffun(x(i),M,s,alp,bet);
end;
% Trapezoidal-rule integral & add binary component
JDC=5*log10(Jb+trapz(x,fillnaninf((f1-f0).*log(f1./f0))));
```

B MATLAB® code for the scaled-F distribution

B.1 J-divergence between two scaled-F distributions

```
function J=jd_sfsf(M0,L0,lam0,M1,L1,lam1)
% J=jd_sfsf(M0,L0,lam0,M1,L1,lam1)
%  J-divergence between two scaled-F distributions
% Parameters: [scalar]
%    (M0,L0,lam0) and (M1,L1,lam1) are shape and scale parameters under H0 & H1
%
c0=M0/(L0*lam0); c1=M1/(L1*lam1);
% Get auxiliary functions
[Gb,Ga]=Gfun_int(c1/c0,floor(M1)+1,L1); G1=interp1(floor(M1)+[0 1],[Ga Gb],M1);
[Gb,Ga]=Gfun_int(c0/c1,floor(M0)+1,L0); G0=interp1(floor(M0)+[0 1],[Ga Gb],M0);
% J-divergence
J=G1*(M0+L0)/(M1+L1)+G0*(M1+L1)/(M0+L0)...
  +(M1-M0)*(psi(M1)-psi(M0)+psi(L0)-psi(L1)-log(c1/c0))...
  -(M1+L1)*(psi(M1+L1)-psi(L1))-(M0+L0)*(psi(M0+L0)-psi(L0));
%--------------------------------------------------------------------------
function [Gb,Ga]=Gfun_int(z,M,L)
% [Gb,Ga]=Gfun_int(z,M,L)
% Returns the auxiliary functions Gb=G(z;M,L) and Ga=G(z;M-1,L) for integer values of M
if M==0,
  Ga=0; Gb=0;
else,
  Ga=Gfun_int(z,M-1,L);
  Gb=(Ga+hypgeo(M+L-1,M,M+L,1-z)*z^(M-1))*(M+L)/(M+L-1);
end
```

B.2 Maximum-likelihood estimation of the scaled-F distribution parameters

```
function [M,L,lam]=sf_mle(t)
% [M,L,lam]=sf_mle(t)
%    Maximum-likelihood estimates (MLEs) of the scaled-F distribution parameters
%    obtained using a numerical optimization algorithm
% Parameters:
%    t = vector of data samples
%    (M,L,lam) = parameter estimates of the scaled-F distribution
%

% Log of scaled-F PDF
logffun=@(m,l,b,t) (exp(m)-1)*mean(log(t))+exp(m).*(m-l-b)...
                  -(exp(m)+exp(l)).*mean(log(1+exp(m-l-b)*t))-betaln(exp(m),exp(l));
% Initialize with a gamma MLE for M and lam, fixed L
pg=gamfit(t(:)); p0=log([pg(1) 10 pg(2)]);
% Use an unconstrained optimization to find log(parameters)
pe=fminsearch(@(p) -logffun(p(1),p(2),p(3),t(:)),p0);
% Convert solution back to values that must be >0
M=exp(pe(1)); L=exp(pe(2)); lam=exp(pe(3));
```

C MATLAB® code for the Gauss hypergeometric function

```
function F=hypgeo(a,b,c,z,dig_acc)
% F=hypgeo(a,b,c,z,[dig_acc])
%   Gauss hypergeometric function 2F1(a,b;c;z) implemented for
%   scalar arguments using the algorithm from
%   "An Atlas of Functions", Spanier and Oldham, 1987
%   on pg 605, Sect. 60:8.
%
% Requires z<1 and c not equal to zero or a negative integer
% Default value of digits of accuracy is  dig_acc=8
%
if nargin<5, dig_acc=8; end;
F=1; d=abs(a)+abs(b)+abs(c);
if (z<0), % If negative, reflect onto (0,1)
  % Switch a & b if a>b (routine is more stable numerically)
  if a>b, aa=b; b=a; a=aa; end;
  % - remaining from Atlas
  F=(1-z)^(-a);
  b=c-b;
  z=z/(z-1);
end;
t=F; j=0; Bloop=true;
while (j<d)||Bloop,
  t=((a+j)*(b+j)/((c+j)*(1+j)))*t*z;
  j=j+1;
  F=F+t;
  Bloop=(abs(F)/(10^dig_acc))<(z*abs(t)/(1-z));
end;
F=F+z*t/(1-z);
```