Detection of Spatially Correlated Time Series From a Network of Sensor Arrays

Nick Klausner, *Student Member, IEEE*, Mahmood R. Azimi-Sadjadi, *Senior Member, IEEE*, and Louis L. Scharf, *Life Fellow, IEEE*

***Abstract*—This paper addresses the problem of testing for the independence among multiple (≥ 2) random vectors. The generalized likelihood ratio test tests the null hypothesis that the composite covariance matrix of the channels is block-diagonal, using a generalized Hadamard ratio. Using the theory of Gram determinants, we show that this Hadamard ratio is stochastically equivalent to a product of scalars, which are independently drawn from a beta distribution under the null hypothesis. This result is then used to derive an asymptotic null distribution, which can be used to identify an appropriate threshold when the sample support is large. These results are then extended to the problem of detecting the presence of spatially correlated time series when each observer employs an array of sensors. Assuming wide-sense stationary processes in both time and space, the likelihood ratio is shown to involve a Hadamard ratio of an estimated cross-spectral matrix at every frequency/wavenumber pair. The proposed detector is compared to several alternative detectors, using simulated space-time fields.**

***Index Terms*—Broadband coherence, coherence, cross-spectral matrix, generalized coherence, generalized likelihood ratio test (GLRT), multichannel coherence, multichannel signal detection, spatio-temporal processing.**

I. Introduction

DETECTING the presence of a common but unknown signal among two or more data channels is a problem that finds its uses in many applications, including collaborative sensor networks [1], geological monitoring of seismic activity [2], radar [3], and sonar [4], [5]. Some detection systems in such situations use decision fusion to combine individual detection decisions into one global decision. However, this detection paradigm is sub-optimal as local decisions are based on the perspective of a single sensory system. Thus, methods that capture the coherent or mutual information among multiple data sets are needed.

Multichannel detection has been considered in [6], [7], and [8]. In [6], a geometric approach to multi-channel detection is proposed by defining the *generalized coherence* (GC) among multiple channels, which is shown to be a natural extension of the *magnitude-squared coherence* (MSC) for more than two channels. Under the assumption that the observations from each channel contain white, complex normal noise, the authors derived closed-form expressions for the null distributions of both the MSC and three-channel GC measures. This leads to a recursive formulation for finding the null distribution as one adds additional channels. A similar technique was considered in [7] by forming a *Generalized Likelihood Ratio Test* (GLRT) and using the assumption that observations are zero-mean, complex normal random vectors. Given multiple independent realizations of this random vector, the GLRT involves testing whether the sample covariance matrix has diagonal structure under the null hypothesis versus any arbitrary, positive-definite (PD) covariance structure under the alternative. In both [6] and [7], the detection statistic applies to *temporally white but spatially correlated* Gaussian sequences, and is given by the determinant of the sample covariance matrix over the product of its diagonal elements, i.e., a Hadamard ratio.

The work in [6] and [7] was recently extended in [8] by considering the detection of both *spatially and temporally correlated* time series. Given multiple independent realizations of a vector-valued time series, the GLRT of [8] tests whether or not the space-time covariance matrix is block-diagonal. The GLRT is a *generalized* Hadamard ratio involving the sample covariance matrix. Assuming temporally wide-sense stationary processes, and allowing the length of each time series to grow large, the test statistic is shown to be a function of frequency-dependent Hadamard ratios for narrowband cross spectral matrices. At each frequency this Hadamard ratio is a narrowband coherence statistic that measures linear dependence among the time series at that frequency. The log of each such narrowband coherence is integrated over the Nyquist band to produce the broadband coherence statistic. This GLRT is shown to exhibit many appealing properties including invariance to channel-by-channel filtering, a connection to mutual information for WSS Gaussian random processes, as well as providing a generalization of the MSC spectrum for more than two channels [8].

In this paper, we consider the detection problem addressed in [8]. Using the theory of linear prediction and its connection with Gram determinants, we show that the likelihood ratio can be written as a product of independent beta random variables under the null hypothesis that the covariance matrix is block-diagonal. This extends the result in [6] to the case of temporally correlated time series and makes it possible to set thresholds for false alarm control. Assuming that the number of indepen-

Manuscript received March 30, 2013; revised August 01, 2013 and August 01, 2013; accepted December 12, 2013. Date of publication January 09, 2014; date of current version February 24, 2014. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Antonio De Maio. This work was supported by the Office of Naval Research (ONR) under contracts N00014-12-C-0017 and N00014-12-1-0154 and the Air Force Office of Scientific Research (AFOSR) under contract FA 9550-10-1-0241.

N. Klausner and M. R. Azimi-Sadjadi are with the Department of Electrical and Computer Engineering, Colorado State University, Fort Collins, CO 80523 USA (e-mail: nklausne@engr.colostate.edu; azimi@engr.colostate.edu).

L. L. Scharf is with the Departments of Mathematics and Statistics, Colorado State University, Fort Collins, CO 80523 USA (e-mail: scharf@engr.colostate.edu).

Digital Object Identifier 10.1109/TSP.2014.2298833

dent realizations used to construct estimates of the covariance matrix grows large, we derive an asymptotic null distribution which may also be used to determine a threshold to approximately achieve a given false alarm probability. We then turn our attention to the detection of spatially correlated time series using a network of sensor arrays. Assuming WSS processes in both time and space, the likelihood ratio is shown to converge to a broadband integral of a log Hadamard ratio of a cross frequency-wavenumber spectrum, when the length of each time series and the number of sensors in each array grows large. This extends the frequency domain technique developed in [8] to the frequency/wavenumber domain when each observer employs an array of sensors. Although this result is asymptotic, it suggests frequency/wavenumber implementations of the GLRT detector even for low space and time support for measured signals.

The paper is organized as follows. Section II gives a brief overview of the GLRT and several of its invariance properties. Section III derives an equivalent expression for the likelihood ratio using Gram determinants and presents the asymptotic null distribution. In Section IV we derive the likelihood ratio for temporally/spatially WSS random processes and demonstrate its performance through simulation in Section V. Concluding remarks are made in Section VI.

II. Review of the Likelihood Ratio

The problem considered here is testing for independence among L random vectors $\{\mathbf{x}_i\}_{i=1}^{L}$ with each vector $\mathbf{x}_i = [x_i[0] \cdots x_i[N-1]]^T$ representing a length N time series captured at sensor i as depicted in Fig. 1. Assuming this collection of random vectors to be zero mean, the composite vector $\mathbf{z} = \left[\mathbf{x}_1^T \cdots \mathbf{x}_L^T\right]^T$ has covariance matrix

$$R = E\left[\mathbf{z}\mathbf{z}^H\right] = \begin{bmatrix} R_{11} & R_{12} & \cdots & R_{1L} \\ R_{12}^H & R_{22} & \cdots & R_{2L} \\ \vdots & \vdots & \ddots & \vdots \\ R_{1L}^H & R_{2L}^H & \cdots & R_{LL} \end{bmatrix} \in \mathbb{C}^{LN \times LN}$$

with $R_{ik} = R_{ki}^H = E\left[\mathbf{x}_i \mathbf{x}_k^H\right] \in \mathbb{C}^{N \times N}$. This matrix not only characterizes the second-order information for each channel individually but also captures the interdependence between every pair of channels. If the set of random vectors $\{\mathbf{x}_i\}_{i=1}^{L}$ is jointly proper complex normal, testing for independence among all L channels becomes a test of whether $R_{ik} = 0$ for all $i \neq k$. Consequently, the null hypothesis $\mathcal{H}_0$ is the hypothesis that R is block-diagonal versus the alternative $\mathcal{H}_1$ that it is not.

We now assume we are given an experiment producing M *iid* realizations $\{\mathbf{x}_i[m]\}_{m=1}^{M}$ of the random vector from each channel i, where

$$\mathbf{x}_i[m] = [x_i[0,m] \cdots x_i[N-1,m]]^T \in \mathbb{C}^N$$

The composite vectors $\mathbf{z}[m] = \left[\mathbf{x}_1^T[m] \cdots \mathbf{x}_L^T[m]\right]^T$ are organized into a data matrix $\mathcal{Z}$:

$$\mathcal{Z} = [\mathbf{z}[1] \cdots \mathbf{z}[M]] = \begin{bmatrix} \mathbf{x}_1[1] & \cdots & \mathbf{x}_1[M] \\ \vdots & \ddots & \vdots \\ \mathbf{x}_L[1] & \cdots & \mathbf{x}_L[M] \end{bmatrix} \in \mathbb{C}^{LN \times M}$$

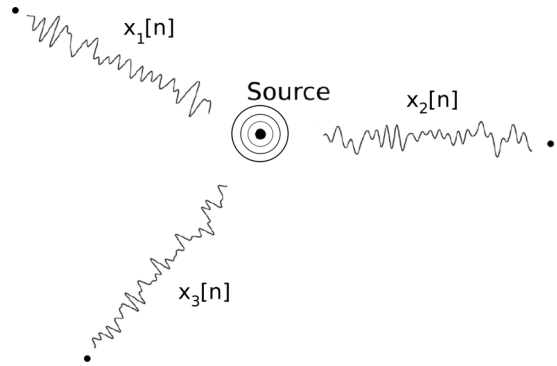


Fig. 1. In many cases, the random vector $\mathbf{x}_i$ represents the collection of a length N time series at one sensor location.

The probability density function (PDF) of $\mathcal{Z}$ is

$$\begin{aligned} f(\mathcal{Z}; R) &= \prod_{m=1}^{M} f\left(\mathbf{z}[m]; R\right) \\ &= \frac{1}{\pi^{LNM} \det\left(R\right)^M} \exp\left\{-M \mathrm{tr}\left(R^{-1}\hat{R}\right)\right\} \end{aligned}$$

In this expression $\hat{R}$ is the estimated composite covariance matrix

$$\begin{aligned} \hat{R} &= \frac{1}{M}\mathcal{Z}\mathcal{Z}^H = \frac{1}{M}\sum_{m=1}^{M} \mathbf{z}[m]\mathbf{z}^H[m] \\ &= \begin{bmatrix} \hat{R}_{11} & \hat{R}_{12} & \cdots & \hat{R}_{1L} \\ \hat{R}_{12}^H & \hat{R}_{22} & \cdots & \hat{R}_{2L} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{R}_{1L}^H & \hat{R}_{2L}^H & \cdots & \hat{R}_{LL} \end{bmatrix} \end{aligned}$$

and $\hat{R}_{ik}$ is an M sample estimate of the $N \times N$ cross-covariance matrix R_{ik}. The GLRT for this problem computes the likelihood ratio [8]

$$\begin{aligned} \Lambda &= \left(\frac{\sup_{R \in \mathcal{R}_0} f(\mathcal{Z}; R)}{\sup_{R \in \mathcal{R}} f(\mathcal{Z}; R)}\right)^{1/M} \\ &= \frac{\det\hat{R}}{\det\hat{D}} = \frac{\det\hat{R}}{\prod_{i=1}^{L} \det\hat{R}_{ii}} = \det\hat{C} \end{aligned} \tag{1}$$

where $\mathcal{R}$ denotes the set of all PD Hermitian matrices and $\mathcal{R}_0$ denotes the set of all matrices in $\mathcal{R}$ which are block-diagonal. Also, the matrices $\hat{R}$ and $\hat{D} = \mathrm{blkdiag}\left\{\hat{R}_{11}, \ldots, \hat{R}_{LL}\right\}$ are maximum likelihood (ML) estimates of R under the alternative and null hypotheses, respectively, and $\hat{C} = \hat{D}^{-1/2}\hat{R}\hat{D}^{-H/2}$ is referred to as the coherence matrix [8]. Note that the matrix $\hat{D}^{1/2}$ is a square-root matrix of $\hat{D}$ such that $\hat{D} = (\hat{D}^{1/2})(\hat{D}^{1/2})^H$. Then $\hat{D}^{-1/2} = (\hat{D}^{1/2})^{-1}$ and $\hat{D}^{-H/2} = (\hat{D}^{-1/2})^H$.

Finally, assuming that all time series are jointly WSS and using results on determinants of asymptotically large block-Toeplitz matrices, the authors in [8] extended this likelihood ratio to the frequency domain by noting that as $M, N \to \infty$,

$$\Lambda^{1/N} \to \exp\left\{\int_{-\pi}^{\pi} \ln \frac{\det \hat{S}(e^{j\theta})}{\prod_{i=1}^{L} \hat{S}_{ii}(e^{j\theta})} \frac{d\theta}{2\pi}\right\} \tag{2}$$

with $\hat{S}(e^{j\theta})$, $-\pi < \theta \leq \pi$, an estimated composite power spectral density matrix:

$$\hat{S}(e^{j\theta}) = \begin{bmatrix} \hat{S}_{11}(e^{j\theta}) & \hat{S}_{12}(e^{j\theta}) & \cdots & \hat{S}_{1L}(e^{j\theta}) \\ \hat{S}^*_{12}(e^{j\theta}) & \hat{S}_{22}(e^{j\theta}) & \cdots & \hat{S}_{2L}(e^{j\theta}) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{S}^*_{1L}(e^{j\theta}) & \hat{S}^*_{2L}(e^{j\theta}) & \cdots & \hat{S}_{LL}(e^{j\theta}) \end{bmatrix} \in \mathbb{C}^{L\times L}$$

Here, $\hat{S}_{ik}(e^{j\theta})$ is a quadratic estimate of the cross power spectrum between channels i and k at frequency θ. The likelihood ratio given in (2) is referred to as *broadband coherence* as the test statistic $\ln \Lambda^{1/N}$ is the log of the narrowband Hadamard ratio, $\det\hat{S}(e^{j\theta}) / \prod_{i=1}^{L} \hat{S}_{ii}(e^{j\theta})$, integrated over the Nyquist band.

A. Invariance Properties

Under suitable choices for the matrix T, the hypothesis testing problem and the likelihood ratio statistic given in (1) remain unchanged upon replacing the random vector $\mathbf{z}$ with $T\mathbf{z}$. Two examples of such linear transformations are given below.

- The set of all matrices T such that $T = \text{blkdiag}\{T_1, \ldots, T_L\}$ with T_i any $N \times N$ invertible matrix. This invariance property was noted in [8] and shows us that there exists no channel-by-channel invertible linear transformation, including scaling and filtering, that moves a covariance from $\mathcal{H}_0$ to $\mathcal{H}_1$ or vice versa.
- The set of all matrices T such that $T = P \otimes I_N$ with P any $L \times L$ permutation matrix. The invariance of the likelihood ratio to this class of linear transformations is easily seen by recalling properties of the determinant of a Kronecker product. This invariance property shows that the ordering in channel index has no influence on likelihood.

III. Stochastic Representation of the Likelihood Ratio Statistic

In this section we derive a formula for the likelihood ratio given in (1) consisting of a product of normalized prediction errors. The virtue of this formula is that it leads to an exact stochastic representation for $\det\hat{C}$ under the null hypothesis – a representation which may be used to set a threshold for false alarm control. This is then used to derive an asymptotic distribution valid for large M which may also be used to determine the threshold needed to approximately achieve a given false alarm probability.

A. The GLR Revisited

We'll begin by noting that, for any $i \geq 2$ and any $n = 0, \ldots, N-1$, the data matrix $\mathcal{Z}$ can be partitioned as follows

$$\mathcal{Z} = \begin{bmatrix} Z_i \\ X_{in} \\ \mathbf{x}_{in}^H \\ \vdots \end{bmatrix}$$

where the matrix $Z_i \in \mathbb{C}^{(i-1)N\times M}$ contains all M realizations of the time-series $\mathbf{x}_1, \ldots, \mathbf{x}_{i-1}$, up to sensor i:

$$Z_i = \begin{bmatrix} \mathbf{x}_1[1] & \mathbf{x}_1[2] & \cdots & \mathbf{x}_1[M] \\ \mathbf{x}_2[1] & \mathbf{x}_2[2] & \cdots & \mathbf{x}_2[M] \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{x}_{i-1}[1] & \mathbf{x}_{i-1}[2] & \cdots & \mathbf{x}_{i-1}[M] \end{bmatrix}.$$

The matrix $X_{in} \in \mathbb{C}^{n\times M}$ contains all M realizations of the time-series at sensor i up to temporal sample $n-1$,

$$X_{in} = \begin{bmatrix} x_i[0,1] & x_i[0,2] & \cdots & x_i[0,M] \\ x_i[1,1] & x_i[1,2] & \cdots & x_i[1,M] \\ \vdots & \vdots & \ddots & \vdots \\ x_i[n-1,1] & x_i[n-1,2] & \cdots & x_i[n-1,M] \end{bmatrix}$$

and the vector $\mathbf{x}_{in} = [x_i[n,1]\, x_i[n,2] \cdots x_i[n,M]]^H \in \mathbb{C}^M$ contains all M realizations of random variable $x_i[n]$. With this partition in the data matrix, the northwest corner of the Gram matrix $\mathcal{Z}\mathcal{Z}^H$ has the structure

$$\mathcal{Z}\mathcal{Z}^H = \begin{bmatrix} \hat{R}_{ZZ} & \hat{R}_{ZX} & \hat{\mathbf{r}}_{Z\mathbf{x}} & \\ \hat{R}^H_{ZX} & \hat{R}_{XX} & \hat{\mathbf{r}}_{X\mathbf{x}} & \cdots \\ \hat{\mathbf{r}}^H_{Z\mathbf{x}} & \hat{\mathbf{r}}^H_{X\mathbf{x}} & \hat{r}_{\mathbf{xx}} & \\ & \vdots & & \ddots \end{bmatrix},$$

with entries defined as follows

$$\begin{aligned} &\hat{R}_{ZZ} = Z_i Z_i^H,\ \hat{R}_{ZX} = Z_i X_{in}^H,\ \hat{R}_{XX} = X_{in} X_{in}^H \\ &\hat{\mathbf{r}}_{Z\mathbf{x}} = Z_i \mathbf{x}_{in},\ \hat{\mathbf{r}}_{X\mathbf{x}} = X_{in}\mathbf{x}_{in} \\ &\hat{r}_{\mathbf{xx}} = \mathbf{x}_{in}^H \mathbf{x}_{in} \end{aligned}$$

Gram determinants [9] are a technique commonly used to test whether or not a collection of vectors in an inner product space are linearly independent. Namely, a set of vectors are linearly independent if and only if the determinant of their Gram matrix is non-zero. Using this result, it is straightforward to show that the determinant of the estimated composite covariance matrix can be decomposed into a product of scalars as follows

$$M^{LN}\det\hat{R} = \det\left(\mathcal{Z}\mathcal{Z}^H\right) = \det\left(X_{1N}X_{1N}^H\right)\prod_{i=2}^{L}\prod_{n=0}^{N-1}\sigma^2_{in}(\hat{R})$$

where

$$\sigma^2_{in}(\hat{R}) = \hat{r}_{\mathbf{xx}} - \begin{bmatrix} \hat{\mathbf{r}}^H_{Z\mathbf{x}} & \hat{\mathbf{r}}^H_{X\mathbf{x}} \end{bmatrix} \begin{bmatrix} \hat{R}_{ZZ} & \hat{R}_{ZX} \\ \hat{R}^H_{ZX} & \hat{R}_{XX} \end{bmatrix}^{-1} \begin{bmatrix} \hat{\mathbf{r}}_{Z\mathbf{x}} \\ \hat{\mathbf{r}}_{X\mathbf{x}} \end{bmatrix}$$

Using the definition of these matrices given above, this term can be written

$$\begin{aligned} \sigma^2_{in}(\hat{R}) &= \mathbf{x}_{in}^H \left(I - \begin{bmatrix} Z_i^H X_{in}^H \end{bmatrix} \begin{bmatrix} \hat{R}_{ZZ} & \hat{R}_{ZX} \\ \hat{R}^H_{ZX} & \hat{R}_{XX} \end{bmatrix}^{-1} \begin{bmatrix} Z_i \\ X_{in} \end{bmatrix} \right) \mathbf{x}_{in} \\ &= \mathbf{x}_{in}^H \left(I - P_{ZX} \right) \mathbf{x}_{in} = \mathbf{x}_{in}^H P^{\perp}_{ZX} \mathbf{x}_{in} \end{aligned}$$

where P_{ZX} denotes the projection onto the $(i-1)N + n$ dimensional subspace $\langle ZX \rangle$ spanned by the columns of matrix $[Z_i^H\ X_{in}^H]$. Moreover, using results for the inverse of a 2×2

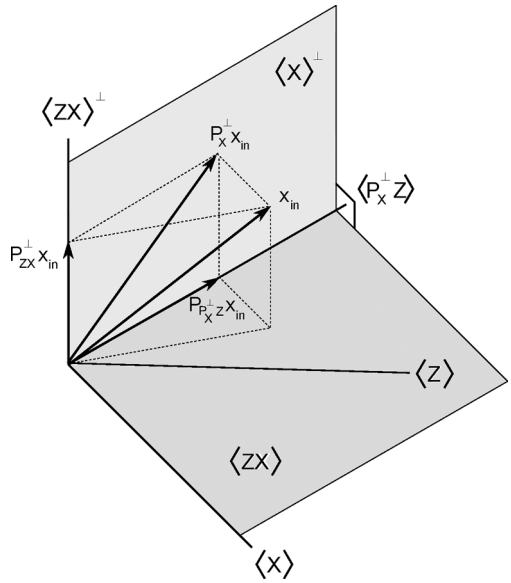


Fig. 2. Orthogonal decomposition of the projection $P_X^{\perp}\mathbf{x}_{in}$ into $P_{ZX}^{\perp}\mathbf{x}_{in}$ and $P_{P_X^{\perp}Z}\mathbf{x}_{in}$.

block matrix, one can derive yet another equivalent expression for this term

$$\begin{aligned}\sigma_{in}^2(\hat{R}) &= \mathbf{x}_{in}^H P_{ZX}^{\perp}\mathbf{x}_{in} \\ &= \mathbf{x}_{in}^H\left(P_X^{\perp} - P_X^{\perp}Z_i^H\left(Z_iP_X^{\perp}Z_i^H\right)^{-1}Z_iP_X^{\perp}\right)\mathbf{x}_{in} \\ &= \mathbf{x}_{in}^H P_X^{\perp}\mathbf{x}_{in} - \mathbf{x}_{in}^H P_{P_X^{\perp}Z}\mathbf{x}_{in} \end{aligned} \tag{3}$$

where $P_X = X_{in}^H\left(X_{in}X_{in}^H\right)^{-1}X_{in}$ and $P_{P_X^{\perp}Z}$ denote the projection onto the n dimensional subspace spanned by the columns of matrix X_{in}^H and the projection onto the $(i-1)N$ dimensional subspace spanned by the columns of matrix $P_X^{\perp}Z_i^H$, respectively.

To compute the determinant of the block-diagonal matrix $\hat{D}$ in the denominator of the likelihood ratio, one can take a very similar approach to show that

$$M^N\det\hat{R}_{ii} = \det\left(X_{iN}X_{iN}^H\right) = \prod_{n=0}^{N-1}\sigma_{in}^2(\hat{R}_{ii})$$

$$\begin{aligned}\text{where}\quad \sigma_{in}^2(\hat{R}_{ii}) &= \hat{r}_{\mathbf{xx}} - \hat{\mathbf{r}}_{X\mathbf{x}}^H\hat{R}_{XX}^{-1}\hat{\mathbf{r}}_{X\mathbf{x}} \\ &= \mathbf{x}_{in}^H\left(I - X_{in}^H\left(X_{in}X_{in}^H\right)^{-1}X_{in}\right)\mathbf{x}_{in} \\ &= \mathbf{x}_{in}^H P_X^{\perp}\mathbf{x}_{in}\end{aligned}$$

Recalling the relationship given in (3), note that this term can alternatively be expressed as follows:

$$\begin{aligned}\sigma_{in}^2(\hat{R}_{ii}) &= \mathbf{x}_{in}^H P_X^{\perp}\mathbf{x}_{in} \\ &= \mathbf{x}_{in}^H P_{ZX}^{\perp}\mathbf{x}_{in} + \mathbf{x}_{in}^H P_{P_X^{\perp}Z}\mathbf{x}_{in}\end{aligned}$$

The geometry of this decomposition is depicted in Fig. 2.

Using the decompositions of the determinants of these matrices, it is then straightforward to see that the likelihood ratio given in (1) can finally be written

$$\begin{aligned}\Lambda = \frac{\det\hat{R}}{\det\hat{D}} &= \frac{\det\left(\mathcal{Z}\mathcal{Z}^H\right)}{\prod_{i=1}^{L}\det\left(X_{iN}X_{iN}^H\right)} \\ &= \frac{\det\left(X_{1N}X_{1N}^H\right)}{\det\left(X_{1N}X_{1N}^H\right)}\prod_{i=2}^{L}\frac{\prod_{n=0}^{N-1}\sigma_{in}^2(\hat{R})}{\det\left(X_{iN}X_{iN}^H\right)} \\ &= \prod_{i=2}^{L}\prod_{n=0}^{N-1}\frac{\sigma_{in}^2(\hat{R})}{\sigma_{in}^2(\hat{R}_{ii})} \\ &= \prod_{i=2}^{L}\prod_{n=0}^{N-1}\frac{\mathbf{x}_{in}^H P_{ZX}^{\perp}\mathbf{x}_{in}}{\mathbf{x}_{in}^H P_{ZX}^{\perp}\mathbf{x}_{in} + \mathbf{x}_{in}^H P_{P_X^{\perp}Z}\mathbf{x}_{in}}\end{aligned} \tag{4}$$

Each term within the product of this expression is the ratio of the squared residual from two different linear least squares problems: the numerator, $\sigma_{in}^2(\hat{R})$, is found by regressing $x_i[n]$ onto all the random variables previous to it while the denominator, $\sigma_{in}^2(\hat{R}_{ii})$, is found by regressing $x_i[n]$ onto those random variables associated with channel i only. Recalling the second invariance property in Section II-A, one must keep in mind that the order in channel index used when constructing this sequence of estimation problems ultimately has no effect on the likelihood ratio. That is, switching orders of channels will lead to different estimation problems generally leading to different individual ratios within the product in (4). However, the product as a whole is invariant to order.

Fig. 2 depicts the geometry of the projections $P_{ZX}^{\perp}\mathbf{x}_{in}$ and $P_{P_X^{\perp}Z}\mathbf{x}_{in}$ used to construct each term of the likelihood ratio in (4). The squared length of the vector $P_{P_X^{\perp}Z}\mathbf{x}_{in}$ in essence represents the increase in mean-squared error incurred by excluding channels $\mathbf{x}_1,\ldots,\mathbf{x}_{i-1}$ from the estimation problem. The smaller the length of this vector relative to the length of $P_{ZX}^{\perp}\mathbf{x}_{in}$, the more evidence in support of the null hypothesis of independence as it indicates that one can just as accurately estimate $x_i[n]$ by ignoring the previous channels. Although the expression of the likelihood ratio given in (4) is no easier to compute than that in (1), it shows that (1) can be written as a product of scalar random variables whose null distribution is the subject of the upcoming section.

B. Stochastic Representation Under the Null Hypothesis

To characterize the distribution of the likelihood ratio under the null hypothesis $\mathcal{H}_0$, we begin by imposing the assumption that $\mathbf{z} \sim \mathcal{CN}(\mathbf{0}, D)$ for any $D = \text{blkdiag}\{R_{11},\ldots,R_{LL}\} \in \mathcal{R}_0$. Recalling the first invariance property in Section II-A, we note that under these circumstances we can always apply the linear transformation $T = D^{-1/2} = \text{blkdiag}\left\{R_{11}^{-1/2},\ldots,R_{LL}^{-1/2}\right\}$, a pre-whitener, to the random vector $\mathbf{z}$ without any consequence to the likelihood ratio. Thus, there is no loss in generality to assume that $D = I_{LN}$ (note that $I_{LN} \in \mathcal{R}_0$) or equivalently that $\mathbf{x}_{in} \overset{iid}{\sim} \mathcal{CN}(\mathbf{0}, I_M)$.

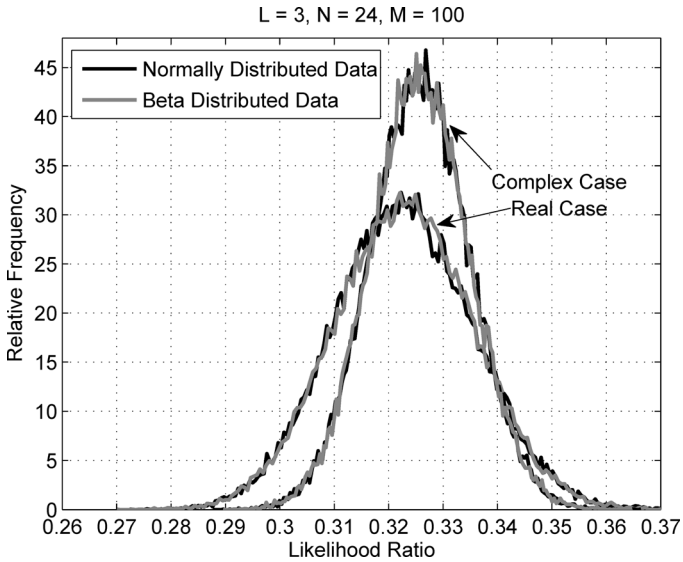


Fig. 3. Monte Carlo Results.

Looking again at Fig. 2, it is clear that the two projections $P_{ZX}^{\perp}\mathbf{x}_{in}$ and $P_{P_X^{\perp}Z}\mathbf{x}_{in}$ lie in two orthogonal subspaces of $\mathbb{C}^M$, i.e., $\langle P_{ZX}^{\perp}\mathbf{x}_{in}, P_{P_X^{\perp}Z}\mathbf{x}_{in}\rangle = 0$. A straightforward application of Cochran's Theorem [10] then shows that $2\mathbf{x}_{in}^H P_{ZX}^{\perp}\mathbf{x}_{in}$ and $2\mathbf{x}_{in}^H P_{P_X^{\perp}Z}\mathbf{x}_{in}$ are statistically independent chi-squared random variables with degrees of freedom $2\,rank\left(P_{ZX}^{\perp}\right) = 2\alpha_{in}$ and $2\,rank\left(P_{P_X^{\perp}Z}\right) = 2\beta_i$, respectively, where

$$\begin{aligned}\alpha_{in} &= M-(i-1)N-n\\ \beta_i &= (i-1)N\end{aligned}$$

Noting that if X and Y represent two independent chi-squared random variables with degrees of freedom ν_X and ν_Y, respectively, then the random variable $\frac{X}{X+Y}$ is distributed according to a beta distribution with parameters $\nu_X/2$ and $\nu_Y/2$, it then follows that

$$\Lambda|\mathcal{H}_0 \overset{\mathcal{D}}{=} \prod_{i=2}^{L}\prod_{n=0}^{N-1} Y_{in} \tag{5}$$

where $Y_{in} \sim \mathrm{Beta}(\alpha_{in},\beta_i)$, all distributed independently of one another. A discussion of the independence in these random variables is given in the appendix. Equation (5) says, "under the null hypothesis, the likelihood ratio statistic is distributed as the product of independent beta random variables, $\mathrm{Beta}(\alpha_{in},\beta_i)$." Note that if the assumption of a complex normal distribution for the data channels is replaced with a real-valued multivariate normal, we can modify the above statements accordingly by simply halving the parameters of these beta random variables, i.e., $Y_{in} \sim \mathrm{Beta}(\alpha_{in}/2,\beta_i/2)$.

For both the real and complex-valued versions of the GLR and with $L = 3$, $N = 24$, and $M = 100$, Fig. 3 displays histograms of Monte Carlo trials generated in two different fashions: the darker plots are generated by drawing data from a normal distribution, forming sample covariance matrices, and computing the generalized Hadamard ratio given in (1) while the lighter plots are generated by sampling from the appropriate beta distributions and forming the product given in (5). These figures show good agreement in the histograms, illustrating stochastically the mathematical fact that the likelihood ratio statistic in (1) is equal in distribution to a product of independent beta random variables. Fractions of threshold exceedances may then be counted to determine which thresholds return which false alarm rates.

C. Asymptotic Null Distribution

Given the definition of the beta function $B(x,y)$ and its connection with the gamma function $\Gamma(z)$

$$B(x,y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)} = \int_0^1 z^{x-1}(1-z)^{y-1}dz$$

every random variable within the product given in (5) has k^{th} raw moment

$$\begin{aligned}E\left[Y_{in}^k\right] &= \int_0^1 x^k f\left(x;\alpha_{in},\beta_i\right)dx\\ &= \frac{1}{B(\alpha_{in},\beta_i)}\int_0^1 x^{(\alpha_{in}+k)-1}(1-x)^{\beta_i-1}dx\\ &= \frac{B(\alpha_{in}+k,\beta_i)}{B(\alpha_{in},\beta_i)} = \frac{\Gamma(\alpha_{in}+k)\Gamma(\alpha_{in}+\beta_i)}{\Gamma(\alpha_{in})\Gamma(\alpha_{in}+\beta_i+k)}\end{aligned}$$

where $f(x;\alpha_{in},\beta_i)$ denotes the PDF of a beta random variable with parameters α_{in} and β_i. Using this fact along with the fact that these random variables are independent, one can see that the likelihood ratio has the following moments under the null hypothesis

$$\begin{aligned}E\left[\Lambda^k|\mathcal{H}_0\right] &= \prod_{i=2}^{L}\prod_{n=0}^{N-1} E\left[Y_{in}^k\right]\\ &= \prod_{i=2}^{L}\prod_{n=0}^{N-1}\frac{\Gamma(M-n)\Gamma(M-(i-1)N-n+k)}{\Gamma(M-(i-1)N-n)\Gamma(M-n+k)}\end{aligned}$$

If we define $\xi_n = (1-\rho)M-n$ with $0\le\rho\le 1$ an arbitrary real number, the characteristic function of the random variable $Z = -2\rho M\ln\Lambda$ can be written

$$\begin{aligned}\phi_Z(t) &= E\left[e^{jtZ}|\mathcal{H}_0\right] = E\left[\Lambda^{-2j\rho Mt}|\mathcal{H}_0\right]\\ &= \prod_{i=2}^{L}\prod_{n=0}^{N-1}\frac{\Gamma(\rho M+\xi_n)\Gamma(\rho M(1-2jt)+\xi_n-(i-1)N)}{\Gamma(\rho M+\xi_n-(i-1)N)\Gamma(\rho M(1-2jt)+\xi_n)}\end{aligned}$$

Its cumulant generating function is

$$\begin{aligned}\psi_Z(t) &= \ln\phi_Z(t)\\ &= \sum_{i=2}^{L}\sum_{n=0}^{N-1}\left[\ln\Gamma(\rho M+\xi_n)-\ln\Gamma\left(\rho M+\xi_n-(i-1)N\right)\right]\\ &\quad+\sum_{i=2}^{L}\sum_{n=0}^{N-1}\ln\Gamma\left(\rho M(1-2jt)+\xi_n-(i-1)N\right)\\ &\quad-\sum_{i=2}^{L}\sum_{n=0}^{N-1}\ln\Gamma\left(\rho M(1-2jt)+\xi_n\right)\end{aligned}$$

To investigate the properties of this cumulant generating function for large M, we employ the following asymptotic expansion [11], [12]

$$\ln\Gamma(z+a) = \frac{1}{2}\ln 2\pi + (z+a-\frac{1}{2})\ln z - z + \sum_{n=1}^{\infty}(-1)^{n+1}\frac{B_{n+1}(a)}{n(n+1)z^n}$$

where $B_n(x)$ denotes an n^{th} order Bernoulli polynomial. This series converges as $|z| \to \infty$ provided that $|\arg z| < \pi$. Assuming that the variable ρ does not go to zero as M becomes large (in fact we'll find that $(1-\rho)$ will be chosen to be $\mathcal{O}(M^{-1})$), then the expression given above can be used to expand the log-gamma functions in the cumulant generating function for large M. Using the 2*nd* order Bernoulli polynomial $B_2(x) = x^2 - x + \frac{1}{6}$ and after a bit of algebra, one obtains the asymptotic expression

$$\psi_Z(t) = -\frac{\nu}{2}\ln(1-2jt) + \omega_1(\rho)\left[(1-2jt)^{-1}-1\right] + \mathcal{O}(M^{-2})$$

where

$$\nu = 2\sum_{i=2}^{L}\sum_{n=0}^{N-1}(i-1)N = L^2N^2 - LN^2$$

$$\omega_1(\rho) = \frac{1}{2\rho M}\sum_{i=2}^{L}\sum_{n=0}^{N-1}\left[(i-1)^2N^2 + (1-2\xi_n)(i-1)N\right] = \frac{1}{2\rho}\left(-\nu(1-\rho) + \frac{L(L^2-1)N^3}{6M}\right)$$

Carefully counting the number of independent real parameters that must be estimated under the alternative and null hypotheses, respectively, one finds that $\dim(\mathcal{R}) = L^2N^2$ and $\dim(\mathcal{R}_0) = LN^2$. Thus, it is clear that the value ν defined above is related to the dimensions of these two spaces through the simple expression

$$\nu = \dim(\mathcal{R}) - \dim(\mathcal{R}_0)$$

The purpose behind the variable ρ in this story is to manipulate the higher order terms in the expansion so that one obtains a more accurate approximation [13]. Namely, it is clear that if we choose the following value for ρ

$$\rho^* = 1 - \frac{L(L^2-1)N^3}{6M\nu} = 1 - \frac{(L+1)N}{6M},$$

then the first order term in the asymptotic expansion of $\psi_Z(t)$ can be made to vanish as $\omega_1(\rho^*) = 0$. This effectively produces an approximation whose error is $\mathcal{O}(M^{-2})$ compared to an error that is $\mathcal{O}(M^{-1})$ if one were to set $\rho = 1$. Letting M tend to infinity and exponentiating the resulting cumulant generating function, we find that

$$\phi_Z(t) = e^{\psi_Z(t)} \overset{M\to\infty}{\longrightarrow} (1-2jt)^{-\nu/2}$$

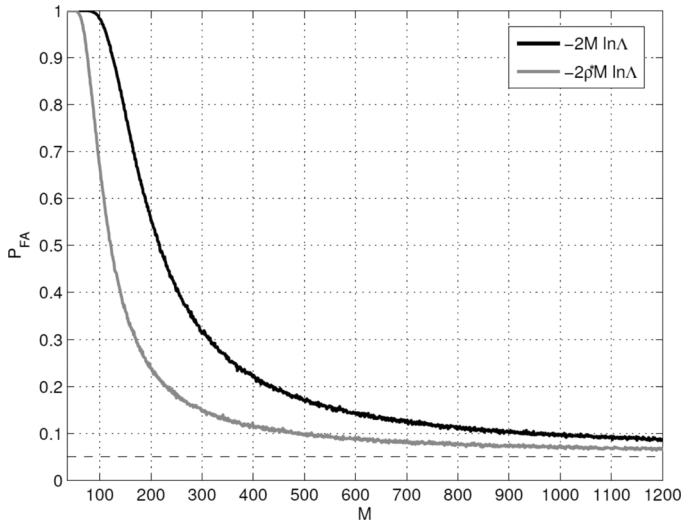


Fig. 4. Asymptotic Empirical False Alarm Probabilities, $P\left[-2\rho M\ln\Lambda > \eta\right]$ for $\rho = 1$ and $\rho = \rho^*$.

which one recognizes as the characteristic function of a chi-squared random variable with ν degrees of freedom. Thus, for large M it follows that

$$P\left[-2\rho^* M\ln\Lambda \le x\right] = P\left[\left(\frac{1}{3}(L+1)N - 2M\right)\ln\Lambda \le x\right] \overset{M\to\infty}{\longrightarrow} P\left[\chi^2_\nu \le x\right];\ \nu = \dim(\mathcal{R}) - \dim(\mathcal{R}_0)$$

With $L = 3$ and $N = 12$, Fig. 4 displays empirical false alarm probabilities versus M for the random variable $-2\rho M\ln\Lambda$ when $\rho = 1$ (shown by a darker line) and when $\rho = \rho^*$ (shown by a lighter line). Here the threshold is chosen from the asymptotic chi-squared distribution and set to achieve a desired false alarm rate of $P_{FA} = 0.05$ (shown by a dashed line). From the figure we can see that, by incorporating a scaling and allowing it to deviate from unity, one can achieve false alarm rates that are closer to the desired value for any finite M by selecting the appropriate value for ρ. Even for moderate values of L and N, however, one can also see that this approximation requires very large values for M to achieve convergence of the false alarm probability to the desired value of 0.05. This suggests that numerical integration of a stochastically generated product of betas as discussed in Section III-B or other approximation techniques such as saddlepoint approximations [14] are a practical alternative.

IV. Extensions to Vector-Valued Time Series

In certain examples of multi-channel detection applications, one may have the opportunity to observe multiple time series from each channel when deciding if a phenomenon common to all channels exists. One such example is a situation where several platforms each employ an array of sensors to take advantage of the spatial diversity such a sensing paradigm might have. As before, the problem considered here is to test for the independence among L random vectors, but we now assume that the random vector from each channel

$$\mathbf{x}_i = \left[\mathbf{x}_i^T[0]\ \cdots\ \mathbf{x}_i^T[N-1]\right]^T \in \mathbb{C}^{dN}$$

contains N samples from a d-dimensional vector-valued time series

$$\mathbf{x}_i[n] = [x_{i,0}[n] \ \cdots \ x_{i,d-1}[n]]^T \in \mathbb{C}^d$$

The likelihood ratio used to solve this problem is essentially the generalized Hadamard ratio given in (1). The only difference is that each block of the composite covariance matrix, R_{ik}, is now $dN \times dN$ rather than $N \times N$. In fact, all of the results of Sections II and III readily generalize to this situation upon replacing N with dN.

As described in [8], the extension of the GLRT to the frequency domain can be accomplished by first independently applying the linear transformation $T = F_N \otimes I_d$, with F_N denoting an $N \times N$ DFT matrix, to the data from each channel. Define the length-N DFT vector

$$\mathbf{f}_N(e^{j\theta_l}) = \frac{1}{\sqrt{N}} \left[1 \ e^{j\theta_l} \ \cdots \ e^{j(N-1)\theta_l}\right]^T$$

and the matrix $F_N(e^{j\theta_l}) = \mathbf{f}_N(e^{j\theta_l}) \otimes I_d \in \mathbb{C}^{dN \times d}$. Then the linear transformation

$$F_N^H(e^{j\theta_l})\mathbf{x}_i = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} e^{-j\theta_l n}\mathbf{x}_i[n]$$

simply corresponds to a unitary DFT analysis of the i^{th} channel at frequency $\theta_l = \frac{2\pi}{N}l$, $l = 0, \ldots, N-1$. Recalling the first invariance property in Section II-A, it follows that both sets of signals, $\{\mathbf{x}_i\}$ and $\{T\mathbf{x}_i\}$, share the same likelihood ratio so that (1) can be written

$$\Lambda = \det\left((I_L \otimes T)\hat{C}(I_L \otimes T)^H\right)$$

Introducing a permutation to the rows and columns of the matrix inside this determinant, the GLRT can be written as

$$\Lambda = \det\tilde{C}$$

$$\text{where} \quad \tilde{C} = \begin{bmatrix} \hat{C}(e^{j\theta_0}) & \cdots & \hat{C}(e^{j\theta_0}, e^{j\theta_{N-1}}) \\ \vdots & \ddots & \vdots \\ \hat{C}(e^{j\theta_{N-1}}, e^{j\theta_0}) & \cdots & \hat{C}(e^{j\theta_{N-1}}) \end{bmatrix}$$

is a matrix that not only captures the second-order information between channels but also between frequencies. The matrix $\hat{C}(e^{j\theta_l}, e^{j\theta_m}) \in \mathbb{C}^{dL \times dL}$ is an $L \times L$ block matrix consisting of $d \times d$ blocks of the form

$$\left\{\hat{C}(e^{j\theta_l}, e^{j\theta_m})\right\}_{i,k} = F_N^H(e^{j\theta_l})\hat{C}_{ik}F_N(e^{j\theta_m})$$

where $\hat{C}_{ik} = \hat{R}_{ii}^{-1/2}\hat{R}_{ik}\hat{R}_{kk}^{-H/2}$ and we use the convention $\hat{C}(e^{j\theta_l}) = \hat{C}(e^{j\theta_l}, e^{j\theta_l})$.

We now assume that all channels are temporally WSS in the sense that, for any pair of channels $\mathbf{x}_i$ and $\mathbf{x}_k$, there exists a matrix-valued sequence, $\{\Gamma_{ik}[l]\}$, such that

$$E\left[\mathbf{x}_i[n]\mathbf{x}_k^H[n+l]\right] = \Gamma_{ik}[l] \in \mathbb{C}^{d \times d}$$

Using results on large block-Toeplitz matrices [15], the matrix $\tilde{C}$ becomes asymptotically equivalent to the block-diagonal matrix

$$\text{blkdiag}\left\{\hat{C}(e^{j\theta_0}), \ldots, \hat{C}(e^{j\theta_{N-1}})\right\}$$

so that as N and M grow large but d and L remain fixed the GLRT becomes

$$\Lambda^{1/N} \overset{N\to\infty}{\longrightarrow} \exp\left\{\int_{-\pi}^{\pi} \ln\det\hat{C}(e^{j\theta})\frac{d\theta}{2\pi}\right\} = \exp\left\{\int_{-\pi}^{\pi} \ln\frac{\det\hat{S}(e^{j\theta})}{\prod_{i=1}^{L}\det\hat{S}_{ii}(e^{j\theta})}\frac{d\theta}{2\pi}\right\} \quad (6)$$

Again, the matrix $\hat{S}(e^{j\theta}) \in \mathbb{C}^{dL \times dL}$ is an $L \times L$ block-structured matrix consisting of $d \times d$ submatrices of the form

$$\left\{\hat{S}(e^{j\theta})\right\}_{i,k} = F_N^H(e^{j\theta})\hat{R}_{ik}F_N(e^{j\theta})$$

which is a quadratic estimate of the cross power spectral density matrix between channels i and k at frequency θ and we use the convention $\hat{S}_{ii}(e^{j\theta}) = \left\{\hat{S}(e^{j\theta})\right\}_{i,i}$. In other words, the matrix $\hat{S} = \left\{\hat{S}_{ik}\right\}_{i,k=1}^{L}$ is a cross spectral matrix of cross spectral matrices, with $\hat{S}_{ik}$ the cross spectral matrix $\left[\hat{S}_{ik}^{lm}\right]_{l,m=1}^{d}$.

The likelihood ratio given in (6) is a direct extension of the results in [8] to account for the situation being considered here and is not a particularly interesting result in that it simply corresponds to replacing every scalar-valued power spectral density estimate in (2) with a $d \times d$ matrix. Although this result is perfectly general in that nothing has been assumed about these vector-valued time series other than that they are temporally WSS, we proceed under the context of multiple-array detection in which case a notion of space can be ascribed to the time series of each channel.

To take advantage of the spatiotemporal properties of the problem, we now consider independently applying the linear transformation $T = F_N \otimes F_d$ to each channel (instead of the matrix $T = F_N \otimes I_d$ considered earlier) with F_d denoting a $d \times d$ DFT matrix. Note that pre-multiplying the vector $\mathbf{x}_i$ by the matrix T simply corresponds to the application of a 2-dimensional DFT, one applied temporally and the other spatially. For any frequency θ, we can then introduce a permutation of the rows and columns of the previously defined matrix $\hat{C}(e^{j\theta})$ so that

$$\det\hat{C}(e^{j\theta}) = \det\tilde{C}(e^{j\theta}) \quad (7)$$

where

$$\tilde{C}(e^{j\theta}) = \begin{bmatrix} \hat{C}(e^{j\theta}, e^{j\phi_0}) & \cdots & \hat{C}(e^{j\theta}, e^{j\phi_0}, e^{j\phi_{N-1}}) \\ \vdots & \ddots & \vdots \\ \hat{C}(e^{j\theta}, e^{j\phi_{N-1}}, e^{j\phi_0}) & \cdots & \hat{C}(e^{j\theta}, e^{j\phi_{N-1}}) \end{bmatrix} \quad (8)$$

and $\hat{C}(e^{j\theta}, e^{j\phi_l}) = \hat{C}(e^{j\theta}, e^{j\phi_l}, e^{j\phi_l})$. Define the length-$d$ DFT vector at frequency $\phi_l = \frac{2\pi}{d}l$ for $l = 0, \ldots, d-1$ as follows

$$\mathbf{f}_d(e^{j\phi_l}) = \frac{1}{\sqrt{d}} \left[1 \ e^{j\phi_l} \ \cdots \ e^{j(d-1)\phi_l}\right]^T.$$

Then the matrix $\hat{C}(e^{j\theta}, e^{j\phi_l}, e^{j\phi_m}) \in \mathbb{C}^{L\times L}$ has entries of the form

$$\left[\hat{C}(e^{j\theta}, e^{j\phi_l}, e^{j\phi_m})\right]_{i,k} = \mathbf{f}_d^H(e^{j\phi_l})F_N^H(e^{j\theta})\hat{C}_{ik}F_N(e^{j\theta})\mathbf{f}_d(e^{j\phi_m}) \quad (9)$$

When the entries of $\mathbf{x}_i[n]$ correspond to time series at different spatial locations, the frequency variable ϕ is often referred to as the *wavenumber* and, to avoid confusion with the variable θ, we will adopt this terminology. Equations (7) – (9) define a *frequency/wavenumber formula for null hypothesis testing in an L sensor network of d-sensor arrays.*

We now impose additional structure on the problem at hand by assuming that all channels are not only temporally WSS but spatially WSS as well so that the multivariate covariance function, $\{\Gamma_{ik}[l]\}$, considered earlier now corresponds to a sequence of Toeplitz matrices. That is, for any pair of channels $\mathbf{x}_i$ and $\mathbf{x}_k$, we now assume that there exists a two-dimensional sequence, $\{\gamma_{ik}[l,m]\}$, such that

$$E\left[x_{i,p}[n]x^*_{k,p+m}[n+l]\right] = \gamma_{ik}[l,m] \in \mathbb{C}$$

with l a temporal lag and m a spatial lag. An example of when this model would hold would be a set of L sensor suites, laid out in an arbitrary geometry, but with their respective d-element arrays laid out co-linearly.

Again invoking results on large block-Toeplitz matrices, it follows that the matrix $\tilde{C}(e^{j\theta})$ becomes asymptotically equivalent with the block-diagonal matrix

$$\text{blkdiag}\left\{\hat{C}(e^{j\theta}, e^{j\phi_0}), \ldots, \hat{C}(e^{j\theta}, e^{j\phi_{d-1}})\right\}$$

so that as M, N, and d grow large but L remains fixed the GLRT becomes

$$\begin{aligned}\Lambda^{\frac{1}{dN}} &= \det\left((I_L\otimes T)\hat{C}(I_L\otimes T)^H\right)^{\frac{1}{dN}} = \det\left(\tilde{C}\right)^{\frac{1}{dN}}\\ &\overset{N\to\infty}{\to} \exp\left\{\int_{-\pi}^{\pi}\ln\det\left(\hat{C}(e^{j\theta})\right)^{\frac{1}{d}}\frac{d\theta}{2\pi}\right\}\\ &= \exp\left\{\int_{-\pi}^{\pi}\ln\det\left(\tilde{C}(e^{j\theta})\right)^{\frac{1}{d}}\frac{d\theta}{2\pi}\right\}\\ &\overset{d\to\infty}{\to} \exp\left\{\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\ln\det\hat{C}(e^{j\theta},e^{j\phi})\frac{d\theta d\phi}{4\pi^2}\right\}\\ &= \exp\left\{\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\ln\frac{\det\hat{S}(e^{j\theta},e^{j\phi})}{\prod_{i=1}^{L}\hat{S}_{ii}(e^{j\theta},e^{j\phi})}\frac{d\theta d\phi}{4\pi^2}\right\} \quad (10)\end{aligned}$$

The matrix $\hat{S}(e^{j\theta}, e^{j\phi}) \in \mathbb{C}^{L\times L}$ has elements

$$\left[\hat{S}(e^{j\theta}, e^{j\phi})\right]_{i,k=1}^{L} = \mathbf{f}_d^H(e^{j\phi})F_N^H(e^{j\theta})\hat{R}_{ik}F_N(e^{j\theta})\mathbf{f}_d(e^{j\phi})$$

which is a quadratic estimate of the cross power spectral density between channels i and k in the frequency/wavenumber domain. Thus, we see that the GLRT involves the computation of a Hadamard ratio at each frequency/wavenumber pair (θ, ϕ), followed by broadband integration of its logarithm – a broadband-broadwavenumber coherence.

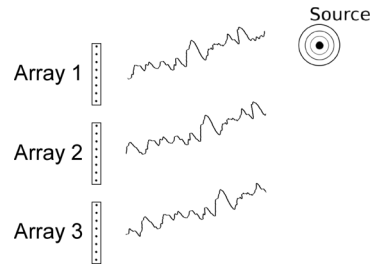


Fig. 5. Detection of a Source using Multiple Linear Arrays.

V. Simulation Results

In this section we provide simulation results to demonstrate a situation where the results of Section IV apply, and to demonstrate the improvement in detection performance that can be achieved under such a situation. Let's consider a network of $L = 3$ sensor arrays, each of which is a uniform linear array (ULA) of $d = 16$ sensor elements. Our aim is to generate a quite arbitrary field at sensor array 1, and then propagate this field to the other two sensor arrays as depicted in Fig. 5.

The propagating signal $s[n]$ for $n = 0, \ldots, MN-1$ is assumed to be a zero-mean WSS random process. As such, it follows that there exists an orthogonal increment process $\{\psi(\theta), -\pi < \theta \le \pi\}$ such that [16]

$$s[n] = \int_{-\pi}^{\pi} e^{jn\theta}d\psi(\theta)$$

The random measure $d\psi(\theta)$, which may be treated as a narrowband component of the signal $s[n]$ at the instantaneous frequency θ, is a complex normal random variable with covariance

$$E\left[d\psi(\theta)d\psi^*(\omega)\right] = \delta(\theta-\omega)\sigma_s^2(e^{j\theta})d\theta$$

and $\sigma_s^2(e^{j\theta})$ is the power spectral density

$$\sigma_s^2(e^{j\theta}) = \frac{1}{2\pi}\sum_{l=-\infty}^{\infty} e^{-jl\theta}E\left[s[n]s^*[n+l]\right]$$

This signal is then propagated to each sensor array so that the signal measured at the i^{th} sensor array, $s_i[n]$, may be written

$$s_i[n] = \int_{-\pi}^{\pi} e^{jn\theta}e^{-j\theta T_i}d\psi(\theta)$$

with T_i a bulk propagation delay and it may be assumed without any loss in generality that $T_1 = 0$. For this simulation, it is assumed that $s[n]$ arises from a first-order autoregressive process with coefficient ϕ and white noise variance σ^2 so that the spectral density $\sigma_s^2(e^{j\theta})$ may be written [16]

$$\sigma_s^2(e^{j\theta}) = \frac{\sigma^2}{2\pi}\frac{1}{|1-\phi e^{-j\theta}|^2}$$

The signal received at each array is then propagated as a planewave among its elements. At each d-element array a temporally colored nonpropagating noise component is added independently of all sensors so that the observation at the i^{th} array, $\mathbf{x}_i[n] \in \mathbb{C}^d$, may be written

$$\mathbf{x}_i[n] = \int_{-\pi}^{\pi} e^{jn\theta}e^{-j\theta T_i}\mathbf{a}(e^{j\theta})d\psi(\theta) + \mathbf{w}_i[n]$$

where the noise vector $\mathbf{w}_i[n]$ has the cross spectral matrix

$$\frac{1}{2\pi}\sum_{l=-\infty}^{\infty} e^{-jl\theta} E\left[\mathbf{w}_i[n]\mathbf{w}_i^H[n+l]\right] = \sigma_w^2(e^{j\theta})I_d$$

Also, the vector $\mathbf{a}(e^{j\theta})$ denotes the array response or steering vector

$$\mathbf{a}(e^{j\theta}) = \left[1\ e^{-j\theta\tau}\ \cdots\ e^{-j(d-1)\theta\tau}\right]^T$$

with τ a propagation delay dependent on the properties of the medium, the distance between sensor elements, and the Direction-of-Arrival (DOA) of the source. Consequently, each $d \times d$ block of the frequency-dependent spectral density matrix of the composite observation can then be written as follows

$$\left\{S(e^{j\theta})\right\}_{i,k} = \begin{cases} \sigma_s^2(e^{j\theta})\mathbf{a}(e^{j\theta})\mathbf{a}^H(e^{j\theta}) + \sigma_w^2(e^{j\theta})I_d & i = k \\ \sigma_s^2(e^{j\theta})e^{j\theta(T_k - T_i)}\mathbf{a}(e^{j\theta})\mathbf{a}^H(e^{j\theta}) & i \neq k \end{cases}$$

For this simulation, the sensor noise is generated by passing unit-variance white noise through a 5th-order FIR filter with complex weights $b_0, \ldots, b_5$ so that the noise spectral density may be written

$$\sigma_w^2(e^{j\theta}) = \frac{1}{2\pi}\left|\sum_{k=0}^{5} b_k e^{-jk\theta}\right|^2$$

Upon collecting all MN measurements at each sensor element, the data record is temporally partitioned into M non-overlapping copies of a time series of length $N = 24$. The likelihood ratio given in (10) (denoted "Frequency/Wavenumber Domain GLRT") is then used to discriminate situations where a source is present from those in which each sensor array observes its own correlated noise field only. The performance of this detector will be compared to the classical likelihood ratio given in (1) (denoted "Time Domain GLRT") as well as its frequency domain version given in (6) (denoted "Frequency Domain GLRT"). First, however, we demonstrate the stochastic nature of the squared residuals in the Time Domain GLRT.

As mentioned briefly at the beginning of Section IV, all of the results of Section III generalize to the situation considered here by simply replacing N with dN in which case the stochastic representation given in (5) becomes

$$\Lambda|\mathcal{H}_0 \overset{\mathcal{D}}{=} \prod_{i=2}^{L}\prod_{n=0}^{dN-1} Y_{in}$$
$$Y_{in} \sim \text{Beta}\left(M-(i-1)dN-n,(i-1)dN\right)$$

With $M = 1200$ realizations of the composite vector $\mathbf{z}[m]$ under the null hypothesis, Fig. 6 plots the ratio of squared residuals, $\sigma_{in}^2(\hat{R})$ to $\sigma_{in}^2(\hat{R}_{ii})$, given in (4)

$$\eta_{in} = \frac{\sigma_{in}^2(\hat{R})}{\sigma_{in}^2(\hat{R}_{ii})} = \frac{\mathbf{x}_{in}^H P_{ZX}^{\perp}\mathbf{x}_{in}}{\mathbf{x}_{in}^H P_{ZX}^{\perp}\mathbf{x}_{in} + \mathbf{x}_{in}^H P_{P_X^{\perp}Z}\mathbf{x}_{in}}$$

for $n = 0, \ldots, dN-1$ when (a) the channel index is $i = 2$ and (b) when it is $i = 3$. The dashed lines in both of these plots show a 95% confidence interval for each beta random variable Y_{in}, i.e., the interval $[a, b]$ such that $P[Y_{in} \leq a] = P[Y_{in} \geq b] = 0.025$. Likewise, with the presence of a source with signal-to-noise ratio $SNR = 10\log_{10}\sigma^2 = -3$ dB, Fig. 7 displays the same for one realization from the alternative hypothesis. Comparing these two figures, it is clear that this interval traps η_{in} with high probability under the null hypothesis, but does not under the alternative hypothesis where many values fall below the interval, signaling a deviation from independence.

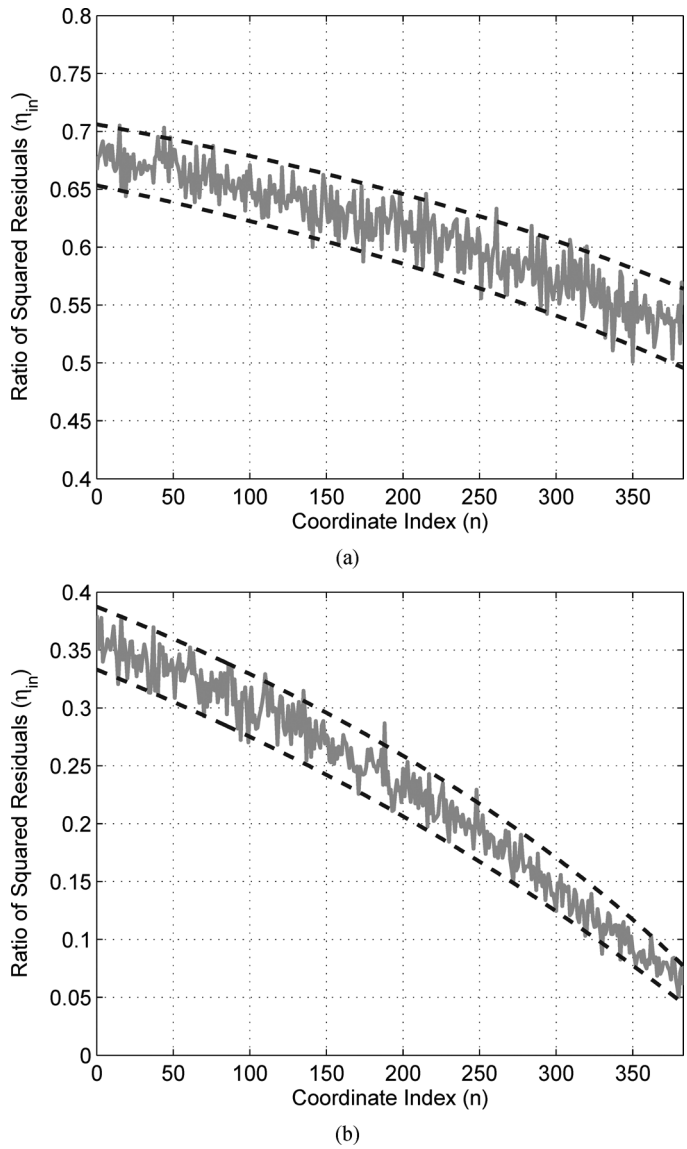
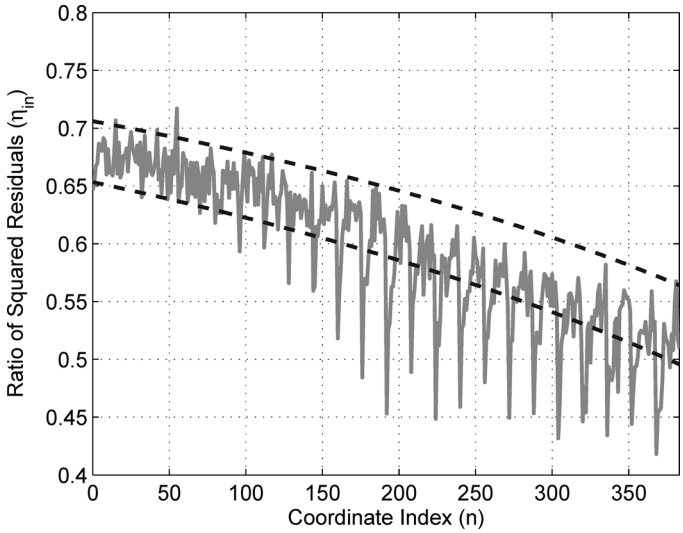
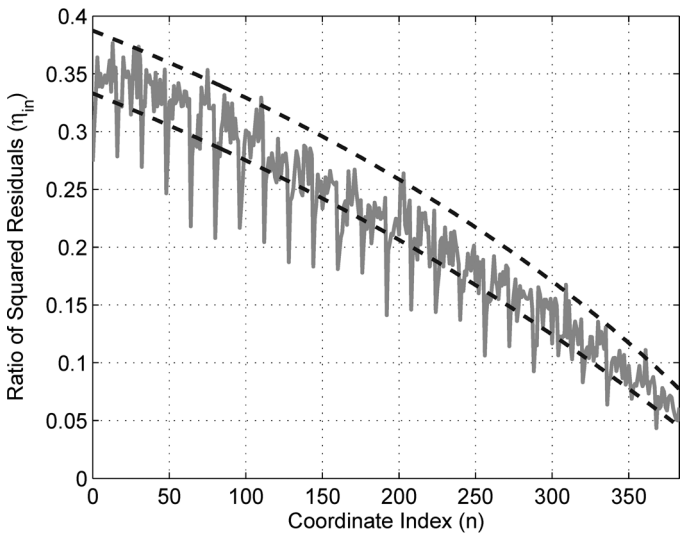


Fig. 6. Ratio of squared residuals under $\mathcal{H}_0$. (a) Channel Index $i = 2$. (b) Channel Index $i = 3$.

With $M = 1200$ and a -30 dB source, Fig. 8 displays the Receiver Operating Characteristic (ROC) curves for all three detection methods considered here. From Fig. 8 we can see that the Frequency/Wavenumber Domain GLRT exhibits a performance that exceeds that of the Frequency Domain GLRT when discriminating these two hypotheses while the performance of the Time Domain GLRT is particularly poor. This is most likely due to the fact that the Time Domain GLRT does not exploit wide-sense stationarity and its manifestation in a Toeplitz structure for the $N \times N$ or the $dN \times dN$ blocks of R. A true GLRT



(a)



(b)

Fig. 7. Ratio of squared residuals under $\mathcal{H}_1$. (a) Channel Index $i = 2$. (b) Channel Index $i = 3$.

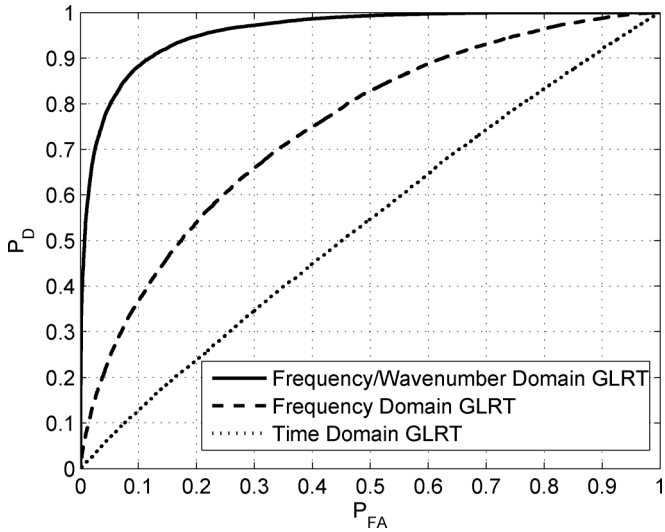


Fig. 8. Detection Performance with $M = 1200$ and $SNR = -30$ dB.

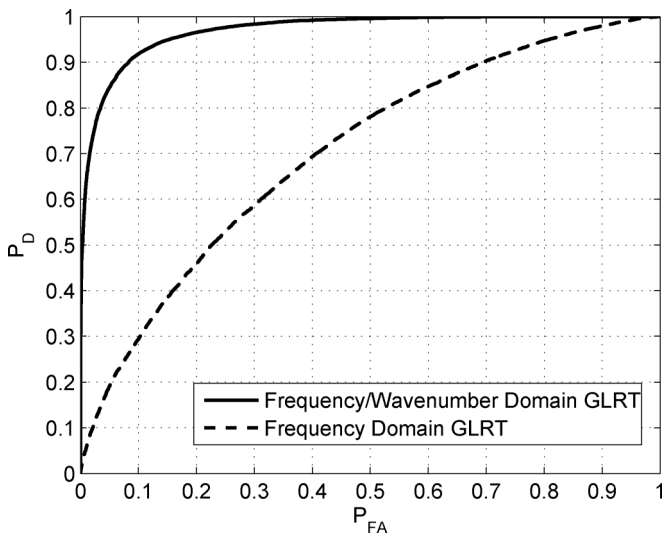


Fig. 9. Detection Performance with $M = 250$ and $SNR = -26$ dB.

for this case would use an ML estimate for Toeplitz matrices, an intractable problem with no analytical solution. So the time-domain GLRT, while generally applicable, is actually mis-matched to the WSS problem. On the other hand, the frequency-domain and frequency-wavenumber domain GLRTs, while not as generally applicable, are better matched to the WSS case. Moreover, these forms estimate cross-spectral matrices, which are approximately block-diagonal in the WSS case, and use only their block-diagonals. In other words, they exploit the assumed wide-sense stationarity by using only diagonal blocks of the cross spectral matrix. Asymptotically, this approaches a GLRT that is faithful to the assumptions of wide-sense stationarity.

Finally, Figs. 9 and 10 compare the performances of the Frequency/Wavenumber and Frequency Domain GLRTs for a successively smaller number of copies but with a higher power source. Note that the Time Domain GLRT has been excluded from these two studies because of insufficient sample support, i.e., M is too small to construct positive-definite covariance

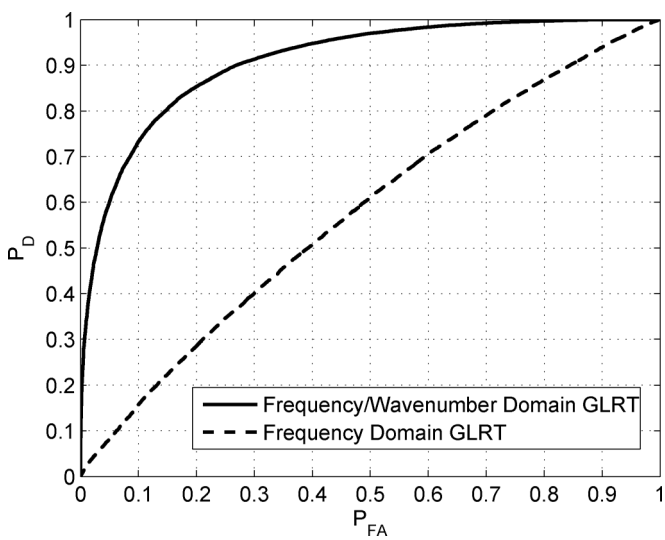


Fig. 10. Detection Performance with $M = 50$ and $SNR = -20$ dB.

estimates. Similar to the results of Fig. 8, we can again see that the Frequency/Wavenumber Domain GLRT outperforms in each case. Again, this is likely due to the fact that the likelihood ratio given in (10) is better matched to the (spatially) WSS case versus its alternative given in (6) which, while more generally applicable, does not fully exploit wide-sense stationarity. Thus, by taking advantage of the spatiotemporal properties of the problem at hand, we can see that the GLRT given in (10) presents an appealing likelihood ratio that exhibits improved detection performance when compared to the two alternatives considered in this paper.

VI. Conclusion

Detecting the presence of common characteristics among two or more data channels is a problem that finds its uses in a wide range of applications. This paper considers the GLRT of [8] which tests whether or not a composite covariance matrix is block-diagonal, through the use of a generalized Hadamard ratio. Using the theory of Gram determinants and their connection with linear prediction, this generalized Hadamard ratio is shown to be equivalent to a product of ratios involving error variances of linear predictors. Assuming that the underlying covariance matrix is truly block-diagonal, we then conclude that the likelihood ratio is statistically equivalent to a product of independent beta random variables under the null hypothesis. Building on this, we derive an asymptotic null distribution which can be used to select an appropriate threshold when the number of samples used to construct estimates of the covariance matrix grows large.

These results are then extended to the detection of spatially correlated time series from a network of sensor arrays. Applying Fourier analysis to each channel in both time and space and assuming both temporally and spatially WSS random processes, the GLRT is shown to be a Hadamard ratio of an estimated cross spectral matrix at every frequency/wavenumber pair when the length of each time series and the number of sensors in each array grows large. To demonstrate the proposed detection method, simulations consisting of 3 spatially separated ULAs sensing the presence of a wideband source were conducted. Through these simulations, it was shown that the proposed frequency/wavenumber domain technique presents an appealing version of the GLRT that can provide improved detection performance, especially in situations with low sample support.

Appendix

The purpose of this appendix is to establish the mutual independence of the random variables given in (5). Consider the random data matrix $X \in \mathbb{C}^{N\times M}$ with $M \geq N$, and without loss of generality $L = 1$,

$$X = \begin{bmatrix} x_0[1] & x_0[2] & \cdots & x_0[M] \\ x_1[1] & x_1[2] & \cdots & x_1[M] \\ \vdots & \vdots & \ddots & \vdots \\ x_{N-1}[1] & x_{N-1}[2] & \cdots & x_{N-1}[M] \end{bmatrix}$$

where $x_n[m] \overset{iid}{\sim} \mathcal{CN}(0,1)$ for $n = 0,\ldots,N-1$ and $m = 1,\ldots,M$. One can think of the n^{th} row of this matrix as an M-sample surrogate for the random variable x_n. As described in Section III, we may partition X

$$X = \begin{bmatrix} X_n \\ \mathbf{x}_n^H \\ \vdots \end{bmatrix}$$

with $X_n \in \mathbb{C}^{n\times M}$ containing all M realizations of random variables $x_0,\ldots,x_{n-1}$ and $\mathbf{x}_n \in \mathbb{C}^M$ containing all M realizations of random variable x_n. The determinant of the Gram matrix XX^H may then be written

$$\det\left(XX^H\right) = \prod_{n=0}^{N-1} \hat{\sigma}_n^2$$
$$\text{where} \quad \hat{\sigma}_n^2 = \mathbf{x}_n^H\left(I - X_n^H(X_nX_n^H)^{-1}X_n\right)\mathbf{x}_n = \mathbf{x}_n^H P_X^{\perp}\mathbf{x}_n$$

Now, conditioned on the data matrix X_n, the random scalar $y_n = 2\mathbf{x}_n^H P_X^{\perp}\mathbf{x}_n$ is a quadratic form involving standard complex normal random variables with a known, deterministic idempotent matrix so that $y_n \mid X_n \sim \chi^2_{2(M-n)}$. This conditional probability distribution is dependent on the number of rows (n) and columns (M) of X_n but in no way dependent on what value this data matrix actually takes making y_n statistically independent of X_n. As the sequence of preceding random variables $y_0,\ldots,y_{n-1}$ are all a function of X_n, this also implies pair-wise independence between y_n and $y_0,\ldots,y_{n-1}$. By induction on n, it then follows that the entire sequence of random variables $y_0,\ldots,y_{N-1}$ are mutually independent so that the random variable $2^N\det\left(XX^H\right)$ is distributed as the product of N independent chi-squared random variables with degrees of freedom $2M,\ 2(M-1),\ldots,\ 2(M-N+1)$. A less heuristic proof of this fact can be found in [17].

In a very similar manner we can now see that, although the random variables $\sigma^2_{in}(\hat{R})$ and $\sigma^2_{in}(\hat{R}_{ii})$ given in (4) are *functionally* dependent on the data matrix $[Z_i^H\ X_{in}^H]$ through the construction of the projection matrices $P_{ZX}^{\perp}$ and $P_X^{\perp}$, respectively, they are in fact *statistically* independent of this data matrix making the sequence of random variables given in (5) mutually independent.

Acknowledgment

The authors would like to thank Dr. D. Cochran and Dr. R. Butler for their invaluable comments and suggestions.

References

[1] A. Nasipui and K. Li, "Multisensor collaboration in wireless sensor networks for detection of spatially correlated signals," *Int. J. Mobile Netw. Design Innovat.*, vol. 1, pp. 215–223, 2006.

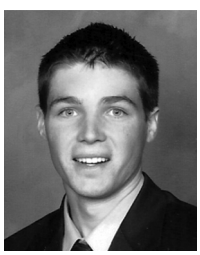
[2] G. Wagner and T. Owens, "Signal detection using multi-channel seismic data," *Bull. Seismolog. Soc. Amer.*, vol. 86, pp. 221–231, 1996.

[3] B. Liu, B. Chen, and J. Michels, "A GLRT for multichannel radar detection in the presence of both compound Gaussian clutter and additive white Gaussian noise," *Digit. Signal Process.*, vol. 15, no. 5, pp. 437–454, 2005.

[4] A. Clausen and D. Cochran, "Non-parametric multiple channel detection in deep ocean noise," in *Proc. 31st Asilomar Conf. Signals, Syst., Comput.*, Nov. 1997, pp. 850–854.

[5] N. Klausner and M. R. Azimi-Sadjadi, "Detection in multiple disparate systems using multi-channel coherence analysis," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 48, pp. 3554–3566, Oct. 2012.

[6] D. Cochran, H. Gish, and D. Sinno, "A geometric approach to multiple-channel signal detection," *IEEE Trans. Signal Process.*, vol. 43, pp. 2049–2057, Sep. 1995.
[7] A. Leshem and A. J. Van der Veen, "Multichannel detection of Gaussian signals with uncalibrated receivers," *IEEE Signal Process. Lett.*, vol. 8, no. 4, pp. 120–122, 2001.
[8] D. Ramirez, J. Via, I. Santamaria, and L. Scharf, "Detection of spatially correlated Gaussian time series," *IEEE Trans. Signal Process.*, vol. 58, no. 10, pp. 5006–5015, 2010.
[9] L. L. Scharf, *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis*. Reading, MA, USA: Addison-Wesley, 1991.
[10] W. Cochran, "The distribution of quadratic forms in a normal system, with applications to the analysis of covariance," *Math. Proc. Cambridge Philosoph. Soc.*, vol. 30, no. 2, pp. 178–191, 1934.
[11] R. Muirhead, *Aspects of Multivariate Statistical Theory*. Hoboken, NJ, USA: Wiley, 2005.
[12] A. Erdélyi, W. Magnus, F. Obergettinger, and F. Tricomi, *Higher Transcendental Functions*. New York, NY, USA: McGraw-Hill, 1953, vol. 1.
[13] G. Box, "A general distribution theory for a class of likelihood criteria," *Biometrika*, vol. 36, no. 4, pp. 317–346, 1949.
[14] R. Butler, *Saddlepoint Approximations with Applications*. New York, NY, USA: Cambridge Univ. Press, 2007.
[15] J. Gutierrez-Gutierrez and P. Crespo, "Asymptotically equivalent sequences of matrices and Hermitian block Toeplitz matrices with continuous symbols: Applications to MIMO systems," *IEEE Trans. Inf. Theory*, vol. 54, no. 12, pp. 5671–5680, 2008.
[16] P. Brockwell and R. Davis, *Time Series: Theory and Methods*, 2nd ed. New York, NY, USA: Springer, 2006.
[17] N. R. Goodman, "The distribution of the determinant of a complex Wishart distributed matrix," *Ann. Math. Statist.*, vol. 34, no. 1, pp. 178–180, 1963.



Nick Klausner (S'08) received the B.S. degrees in electrical engineering and economics from the Colorado School of Mines, Golden, CO, USA, in 2008 and the M.S. degree in electrical and computer engineering from Colorado State University, Fort Collins, CO, USA, in 2010. Since 2010, he has been working towards the Ph.D. degree in electrical and computer engineering at Colorado State University under the direction of Dr. M. R. Azimi-Sadjadi and Dr. L. Scharf.

His current research interests include statistical signal processing, digital signal processing, and digital image processing.

Mahmood R. Azimi-Sadjadi (M'81–SM'89) received the M.S. and Ph.D. degrees from the Imperial College of Science and Technology, University of London, London, U.K., in 1978 and 1982, respectively, both in electrical engineering with specialization in digital signal/image processing.

He is currently a Full Professor with the Electrical and Computer Engineering Department, Colorado State University (CSU), Fort Collins, CO, USA. He also serves as the Director of the Digital Signal/Image Laboratory, CSU. His main areas of interest include statistical signal and image processing, target detection, classification and tracking, sensor array processing, learning systems, and distributed sensor networks. He is the co-author of the book *Digital Filtering in One and Two Dimensions* (New York, NY, USA: Plenum, 1989).

Dr. Azimi-Sadjadi has served as an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING and the IEEE TRANSACTIONS ON NEURAL NETWORKS.

Louis L. Scharf (S'67–M'69–SM'77–F'86–LF'07) received the Ph.D. degree from the University of Washington, Seattle, WA, USA.

From 1971 to 1982, he served as Professor of Electrical Engineering and Statistics with Colorado State University (CSU), Ft. Collins, CO, USA. From 1982 to 1985, he was Professor and Chairman of Electrical and Computer Engineering, University of Rhode Island, Kingston, RI, USA. From 1985 to 2000, he was Professor of Electrical and Computer Engineering, University of Colorado, Boulder, CO, USA. In January 2001, he rejoined CSU as Professor of Electrical and Computer Engineering and Statistics. He is currently Research Professor of Mathematics at CSU. He has held several visiting positions here and abroad, including the Ecole Superieure dElectricite, Gif-sur-Yvette, France; Ecole Nationale Superieure des Telecommunications, Paris, France; EURECOM, Nice, France; the University of La Plata, La Plata, Argentina; Duke University, Durham, NC, USA; the University of Wisconsin, Madison, WI, USA; and the University of Tromso, Tromso, Norway. His interests are in statistical signal processing, as it applies to radar, sonar, power, cyber-security, and wireless communication.

Prof. Scharf was Technical Program Chair for the 1980 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Denver, CO; Tutorials Chair for ICASSP 2001, Salt Lake City, UT, USA; and Technical Program Chair for the 2002 Asilomar Conference on Signals, Systems, and Computers. He is past-Chair of the Fellow Committee for the IEEE Signal Processing Society. He has received numerous awards for his research contributions to statistical signal processing, including a College Research Award, an IEEE Distinguished Lectureship, an IEEE Third Millennium Medal, the Technical Achievement and Society Awards from the IEEE Signal Processing Society, and the Donald W. Tufts Award from the IEEE Underwater Acoustic Signal Processing Workshop.