SPACE-TIME COHERENCE AND ITS EXACT NULL DISTRIBUTION

*Nick Klausner** *Mahmood Azimi-Sadjadi** *Louis Scharf** *Douglas Cochran*†

* Colorado State University
Fort Collins, CO 80523

† Arizona State University
Tempe, AZ 85281

ABSTRACT

This paper considers the problem of testing for the independence among multiple (≥ 2) random vectors with each random vector representing a time series captured at one sensor. Implementing the Generalized Likelihood Ratio Test involves testing the null hypothesis that the composite covariance matrix of the channels is block-diagonal through the use of a generalized Hadamard ratio. Using the theory of linear prediction and its connection with Gram determinants, it is shown that this generalized Hadamard ratio can be written as a product of scalars which are independently drawn from a beta distribution under the null hypothesis. This result is useful from a Monte Carlo analysis standpoint in that it is much more computationally efficient to form a product of scalar beta random variables than it is to compute the determinant of complex Wishart matrices.

***Index Terms*—** Generalized Likelihood Ratio Test, Multichannel Coherence, Multichannel Signal Detection, Test for Block Independence

1. 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, geological monitoring of seismic activity, as well as both active and passive radar and sonar applications. In [1], a non-parametric approach to multi-channel detection was proposed by defining the *Generalized Coherence* (GC) measure among multiple channels. The GC measure is shown to be a natural extension of the *Magnitude-Squared Coherence* (MSC) measure for more than two channels. Under the assumption that the observations from each channel contain white, complex normal noise, the authors derive 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 an additional channel. This work was recently extended in [2] by considering the detection of multiple temporally correlated time series. Forming a random vector containing the time series from all channels, the GLRT involves testing whether or not the space-time covariance matrix is block-diagonal through the use of a generalized Hadamard ratio involving a sample covariance matrix computed from multiple independent realizations.

In this paper, we consider the detection problem addressed in [2]. 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 truly block-diagonal. This extends the result in [1] to the case of multiple temporally correlated time series. Although the authors in [1] were able to derive exact distributions of the test statistic under the null hypothesis, doing so here under general circumstances is more difficult. However, one always has the option of determining the threshold needed to approximately achieve a desired false alarm probability through the use of Monte Carlo techniques. It is in this sense that the results of this paper become useful as we show that one simply needs to multiply scalar beta random variables to stochastically simulate the null distribution of the test statistic as opposed to computing the determinants of random matrices.

2. REVIEW OF THE LIKELIHOOD RATIO

The problem considered here is testing for the 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. 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 space-time 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 $R_{ij} = R_{ji}^H = E\left[\mathbf{x}_i\mathbf{x}_j^H\right] \in \mathbb{C}^{N \times N}$ a temporal cross-covariance matrix. This matrix not only characterizes the second-order temporal 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 boils down to testing whether or not the covariance matrix R is block-diagonal. Casting this problem into the standard inference framework, we consider the following hypothesis test

$$\begin{aligned} \mathcal{H}_0 &: \quad R \in \mathcal{R}_0 \\ \mathcal{H}_1 &: \quad R \in \mathcal{R} \end{aligned}$$

with $\mathcal{R}$ denoting the set of all PD Hermitian matrices and $\mathcal{R}_0$ denoting the set of all matrices in $\mathcal{R}$ which are additionally block-diagonal.

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 collection of random vectors

$$\mathbf{z}[m] = \left[\mathbf{x}_1^T[m] \ \cdots \mathbf{x}_L^T[m]\right]^T, \ \ m = 1, \ldots, M$$

has probability density function (PDF)

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

with $\hat{R}$ being an estimated composite covariance matrix

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

and $\hat{R}_{ij}$ being an M sample estimate of the matrix R_{ij}. The GLRT for this problem involves computing the likelihood ratio [2]

$$\begin{aligned} \Lambda &= \left(\frac{\max_{R \in \mathcal{R}_0} \prod_{m=1}^{M} f\left(\mathbf{z}[m]; R\right)}{\max_{R \in \mathcal{R}} \prod_{m=1}^{M} f\left(\mathbf{z}[m]; R\right)} \right)^{1/M} \\ &= \frac{\det\left(\hat{R}\right)}{\det\left(\hat{D}\right)} = \frac{\det\left(\hat{R}\right)}{\prod_{i=1}^{L} \det\left(\hat{R}_{ii}\right)} = \det\left(\hat{C}\right) \qquad (1) \end{aligned}$$

where $\hat{R}$ 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 a dimensionless matrix referred to as the coherence matrix [2].

2.1. 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 $T = \mathrm{blkdiag}\{T_1, \ldots, T_L\}$ with T_i any $N \times N$ invertible matrix [2]. This invariance property 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 $T = P \otimes I_N$ with P any $L \times L$ permutation matrix. This invariance property shows that the ordering in channel index has no influence on likelihood.

3. NULL DISTRIBUTION OF THE LIKELIHOOD RATIO

We'll begin by considering the random matrix Z^T

$$Z^T = \begin{bmatrix} \mathbf{x}_1[1] & \mathbf{x}_1[2] & \cdots & \mathbf{x}_1[M] \\ \mathbf{x}_2[1] & \mathbf{x}_2[2] & \cdots & \mathbf{x}_2[M] \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{x}_L[1] & \mathbf{x}_L[2] & \cdots & \mathbf{x}_L[M] \end{bmatrix} \in \mathbb{C}^{LN \times M}$$

so that each column in matrix Z corresponds to all M realizations of one particular random variable, $x_i[n]$. For any $i \geq 2$ and any $n = 0, \ldots, N-1$, it will become convenient to partition the data matrix Z as follows

$$Z = \left[\ Z_i \quad X_{in} \quad \mathbf{x}_{in} \quad \cdots\ \right]$$

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

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

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

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

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

$$Z^H Z = \left[\begin{array}{cccc} R_{ZZ} & R_{ZX} & \mathbf{r}_{Z\mathbf{x}} & \\ R_{ZX}^H & R_{XX} & \mathbf{r}_{X\mathbf{x}} & \cdots \\ \mathbf{r}_{Z\mathbf{x}}^H & \mathbf{r}_{X\mathbf{x}}^H & r_{\mathbf{x}\mathbf{x}} & \\ & \vdots & & \ddots \end{array} \right]$$

with entries that are defined as follows

$$\begin{aligned} R_{ZZ} = Z_i^H Z_i,\ R_{ZX} = Z_i^H X_{in},\ R_{XX} = X_{in}^H X_{in} \\ \mathbf{r}_{Z\mathbf{x}} = Z_i^H \mathbf{x}_{in},\ \mathbf{r}_{X\mathbf{x}} = X_{in}^H \mathbf{x}_{in} \\ r_{\mathbf{x}\mathbf{x}} = \mathbf{x}_{in}^H \mathbf{x}_{in} \end{aligned}$$

Gram determinants [3] 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 theory, it is straightforward to show that the determinant of the estimated composite covariance matrix can be written

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

where

$$\sigma_{in}^2(\hat{R}) = r_{\mathbf{x}\mathbf{x}} - \left[\begin{array}{cc} \mathbf{r}_{Z\mathbf{x}}^H & \mathbf{r}_{X\mathbf{x}}^H \end{array} \right] \left[\begin{array}{cc} R_{ZZ} & R_{ZX} \\ R_{ZX}^H & R_{XX} \end{array} \right]^{-1} \left[\begin{array}{c} \mathbf{r}_{Z\mathbf{x}} \\ \mathbf{r}_{X\mathbf{x}} \end{array} \right]$$

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

$$\begin{aligned} \sigma_{in}^2(\hat{R}) &= \mathbf{x}_{in}^H \left(I - [Z_i\ X_{in}] \left[\begin{array}{cc} R_{ZZ} & R_{ZX} \\ R_{ZX}^H & R_{XX} \end{array} \right]^{-1} \left[\begin{array}{c} Z_i^H \\ X_{in}^H \end{array} \right] \right) \mathbf{x}_{in} \\ &= \mathbf{x}_{in}^H \left(I - P_{ZX}\right) \mathbf{x}_{in} \\ &= \mathbf{x}_{in}^H P_{ZX}^{\perp} \mathbf{x}_{in} \end{aligned}$$

where P_{ZX} denotes the projection onto the $(i-1)N + n$ dimensional subspace $\langle ZX \rangle$ spanned by the vectors $[Z_i\ X_{in}]$. Moreover, using results for the inverse of a 2×2 block matrix, one can derive yet another equivalent expression for this term

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

where $P_X = X_{in} R_{XX}^{-1} X_{in}^H$ and $P_{P_X^{\perp} Z}$ denote the projection onto the n dimensional subspace spanned by the vectors X_{in} and the projection onto the $(i-1)N$ dimensional subspace spanned by the vectors $P_X^{\perp} Z_i$, 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

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

where

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

Recalling the relationship given in (2), note that this term can alternatively be expressed as follows

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

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

$$\begin{aligned} \Lambda = \frac{\det \hat{R}}{\det \hat{D}} &= \frac{\det\left(Z^H Z\right)}{\prod_{i=1}^{L} \det\left(X_{iN}^H X_{iN}\right)} \\ &= \frac{\det\left(X_{1N}^H X_{1N}\right)}{\det\left(X_{1N}^H X_{1N}\right)} \prod_{i=2}^{L} \frac{\prod_{n=0}^{N-1} \sigma_{in}^2(\hat{R})}{\det\left(X_{iN}^H X_{iN}\right)} \\ &= \prod_{i=2}^{L} \prod_{n=0}^{N-1} \frac{\sigma_{in}^2(\hat{R})}{\sigma_{in}^2(\hat{R}_{ii})} \\ &= \prod_{i=2}^{L} \prod_{n=0}^{N-1} \frac{\mathbf{x}_{in}^H P_{ZX}^{\perp} \mathbf{x}_{in}}{\mathbf{x}_{in}^H P_{ZX}^{\perp} \mathbf{x}_{in} + \mathbf{x}_{in}^H P_{P_X^{\perp} Z} \mathbf{x}_{in}} \end{aligned} \tag{3}$$

Each term within the product of this expression represents the ratio of the estimated variance of a 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. This process is depicted pictorially in Figure 1. Recalling the second invariance property in Section 2.1, one must keep in mind that the order in channel index one uses when constructing this sequence of estimation problems ultimately has no effect on the likelihood ratio.

To characterize the distribution of the likelihood ratio under the null hypothesis, 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 2.1, we note that under these circumstances we can always apply the linear transformation $T = D^{-1/2}$, 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 $\mathbf{x}_{in} \overset{iid}{\sim} \mathcal{CN}(\mathbf{0}, I_M)$.

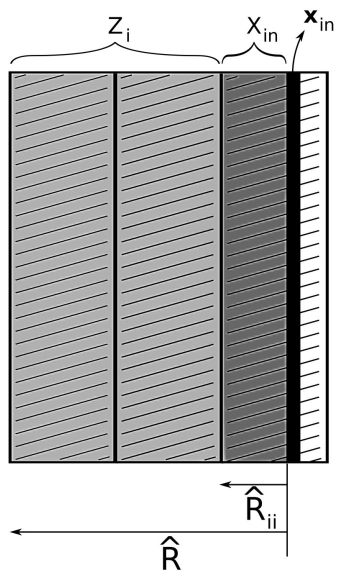

Fig. 1. Individual terms of the likelihood ratio involve linearly predicting the vector $\mathbf{x}_{in}$ using the data matrices Z_i and X_{in}.

Looking closely at (3), it is clear that the two projections $P^{\perp}_{ZX}\mathbf{x}_{in}$ and $P_{P^{\perp}_{X}Z}\mathbf{x}_{in}$ lie in two orthogonal subspaces of $\mathbb{C}^M$, i.e. $\langle P^{\perp}_{ZX}\mathbf{x}_{in}, P_{P^{\perp}_{X}Z}\mathbf{x}_{in}\rangle = 0$. A straightforward application of Cochran's Theorem [4] then shows that the two values $\mathbf{x}^H_{in}P^{\perp}_{ZX}\mathbf{x}_{in}$ and $\mathbf{x}^H_{in}P_{P^{\perp}_{X}Z}\mathbf{x}_{in}$ are statistically independent chi-squared random variables with degrees of freedom $2\ rank\left(P^{\perp}_{ZX}\right) = 2\alpha_{in}$ and $2\ rank\left(P_{P^{\perp}_{X}Z}\right) = 2\beta_i$, 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

$$\Lambda|\mathcal{H}_0 \overset{d}{=} \prod_{i=2}^{L}\prod_{n=0}^{N-1} Y_{in} \tag{4}$$

where $Y_{in} \sim$ Beta(α_{in}, β_i), all distributed independently of one another. 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. $Y_{in} \sim$ Beta$(\alpha_{in}/2, \beta_i/2)$. Equation (4) says, "under the null hypothesis, the likelihood ratio statistic is distributed as the product of beta random variables, Beta(α_{in}, β_i)".

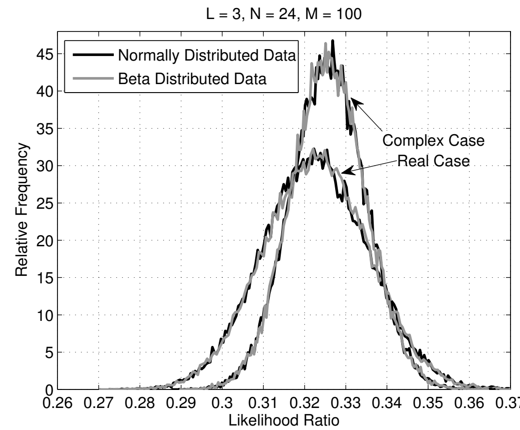

Fig. 2. Monte Carlo Results.

For both the real and complex-valued versions of the GLR and with $L = 3$, $N = 24$, and $M = 100$, Figure 2 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 ratio of determinants given in (1) while the lighter plots are generated by sampling from the appropriate beta distribution and forming the product given in (4). These figures show good agreement in the histograms demonstrating the fact that the null distribution of (1) can be stochastically generated by drawing independent beta random variables and forming their product.

4. CONCLUSIONS

Detecting the presence of common characteristics among two or more time-series channels is a problem that finds its uses in a wide range of applications. To make the detection method applicable under general circumstances, this paper considers a GLRT that simply 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, we show that this generalized Hadamard ratio can be written as a product of ratios involving the error variance of two 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. This result has several uses, one of them being the determination of thresholds needed to achieve given false alarm probabilities.

5. REFERENCES

[1] D. Cochran, H. Gish, and D. Sinno, "A geometric approach to multiple-channel signal detection," *IEEE Transactions on Signal Processing*, vol. 43, pp. 2049–2057, September 1995.

[2] D. Ramirez, J. Via, I. Santamaria, and L. Scharf, "Detection of spatially correlated Gaussian time series," *IEEE Transactions on Signal Processing*, vol. 58, no. 10, pp. 5006–5015, 2010.

[3] L. L. Scharf, *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis*, Addison-Wesley, 1991.

[4] W. Cochran, "The distribution of quadratic forms in a normal system, with applications to the analysis of covariance," *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 30, no. 2, pp. 178–191, 1934.