On Multisensor Detection of Improper Signals

Jitendra K. Tugnait, *Life Fellow, IEEE*

***Abstract*—We consider the problem of detecting the presence of a complex-valued, possibly improper, but unknown signal, common among two or more sensors (channels) in the presence of spatially independent, unknown, possibly improper and colored, noise. Past work on this problem is limited to signals observed in proper noise. A source of improper noise is in-phase/quadrature-phase (IQ) imbalance during down-conversion of bandpass noise to baseband. A binary hypothesis testing approach is formulated, and a generalized likelihood ratio test (GLRT) is derived using asymptotic distribution of a frequency-domain sufficient statistic, based on the discrete Fourier transform of an augmented measurement sequence. An asymptotic analytical solution for calculating the test threshold to yield a desired false alarm rate, is provided. The performance of the GLRT is analyzed by deriving an approximate asymptotic distribution of the GLRT statistic under a sequence of local alternative hypotheses. These results (GLRT and test threshold calculation) are then modified to address two special cases: detection of improper signals in proper noise, and detection of proper signals in proper noise. Simulation examples are presented in support of the proposed approaches.**

***Index Terms*—Improper complex random signals; generalized likelihood ratio test (GLRT); multichannel signal detection; spectral analysis.**

I. Introduction

We consider the problem of detecting the presence of a complex-valued, possibly improper, but unknown signal, common among two or more sensors. The unknown common signal is observed at multiple sensors in the presence of unknown, possibly improper and colored, noise that is independent across sensors.

A zero-mean complex-valued random sequence is called proper if the cross-correlation function of the sequence with its complex conjugate (called complementary correlation) is vanishing [1]. Typically, algorithms for complex signal processing in communications and statistical signal processing have been derived assuming that the complex signals are proper [1], [2]. However, this assumption of propriety is not always justified. For example, binary phase shift keying (BPSK), offset quadrature phase shift keying (QPSK), Gaussian minimum shift keying (GMSK) and amplitude shift keying (ASK) modulation based signals are improper [1]. Also, in-phase/quadrature-phase (IQ) imbalance during down-conversion of bandpass signals to baseband can result in impropriety in both signals and noise [3]. If the underlying signals are improper, much can be gained in performance if the information contained in the complementary correlation is also exploited [1], [4], [5].

A potential application of this problem is in spectrum sensing for cognitive radio (CR) to decide if the received signal, in addition to noise, contains signals from a single or multiple primary users (PUs); for other potential applications, see [6]. This is formulated as a binary hypothesis testing problem and is a well-investigated topic [7]. If the underlying communications channel is frequency-selective, the received signal will be colored. An irreducible source of noise at the receiver is thermal noise which is zero-mean white Gaussian and whose baseband-equivalent model is that of a zero-mean, white, proper complex Gaussian random process. Often, as a result of receive filters at the cognitive user's receiver, the filtered thermal noise will not be white. Before processing the (continuous-time) received noisy signal, at the receiver front-end, the noisy (baseband-equivalent) signal is passed through a receive filter $h_r(t)$. Let $\tilde{w}(t)$ denote the filtered noise and $w(t)$ denote the thermal noise. Then

$$\tilde{w}(t) = \int_{-\infty}^{\infty} w(t-\lambda) h_r(\lambda)\, d\lambda\,.$$

The autocorrelation function of $\tilde{w}(t)$ is

$$R_{\tilde{w}}(\tau) = E\{\tilde{w}(t)\tilde{w}^*(t-\tau)\} = \frac{N_0}{2}\int_{-\infty}^{\infty} h_r(t) h_r^*(t-\tau)\, dt$$

where $R_w(\tau) = \frac{N_0}{2}\delta(\tau)$, and superscript $*$ denotes the complex conjugate operation. If $h_r(t)$ is a square-root raised cosine filter with bandwidth $(1+\alpha)/(2T_s)$, $0 \le \alpha \le 1$, where T_s is the symbol interval, then $R_{\tilde{w}}(nT_s) = 0$ for $n = \pm 1, \pm 2, \cdots$. That is, $\{\tilde{w}(t)\big|_{t=nT_s}\}_n$ is white Gaussian sampled sequence when $\tilde{w}(t)$ is sampled at the symbol rate. But if one oversamples (more than one sample in T_s sec, called fractional sampling), then $\{\tilde{w}(t)\big|_{t=nT}\}_n$, $T < T_s$, is colored.

Thus, if the receive filters are not necessarily the ones that yield white noise at symbol rate sampling, or the sampling rate exceeds the symbol rate, the sampled filtered thermal noise will be a colored random sequence. For CR applications, active RC filters with tunable cut-off frequencies have been proposed in the literature [8], [9]. Sampling of the filtered thermal noise using such RC filters will invariably result in colored noise. The effect of noise correlation on spectrum sensing performance for a class of spectrum sensing algorithms has been analyzed in [10].

Several approaches exist for detection of random signals in random noise, based on differing signal and noise models [7]. A widely used model is that of temporally white but spatially correlated proper complex Gaussian PU signal in temporally and spatially uncorrelated proper complex Gaussian noise [11].

Manuscript received July 4, 2018; revised Oct. 21, 2018; accepted Dec. 4, 2018. The associate editor coordinating the review of this paper and approving it for publication was Prof. Hongbin Li.

J.K. Tugnait is with the Department of Electrical & Computer Engineering, 200 Broun Hall, Auburn University, Auburn, AL 36849, USA. Email: tugnajk@auburn.edu .

This work was supported by the National Science Foundation under Grant CCF-1617610.

A preliminary version of the material in this paper was presented in part at the *2017 IEEE Intern. Conf. Acoustics, Speech & Signal Proc. (ICASSP 2017)*, New Orleans, Louisiana, March 2017.

Temporally colored, proper signals in spatially uncorrelated but temporally correlated proper Gaussian noise have been considered in [12] assuming multiple independent realizations (snapshots) and Gaussian PU signals. The noise and signal model parameters and/or statistics are unknown in [12], hence, a generalized likelihood ratio test (GLRT) framework is used to derive a test statistic. The test threshold is selected via Monte Carlo simulations in [12] to yield a specified false alarm rate. The work of [12] has been extended in [6], where also one needs multiple independent realizations, with number of realizations becoming large for certain theoretical results, such as analytical threshold selection. The GLRT framework is also used in [11] to handle the unknown parameters. The approaches of [6], [11], [12] are based on time-domain consideration. A frequency-domain approach based on estimated power spectral density (PSD) of the measurements and the GLRT framework, has been proposed in [13] for temporally colored, proper signals in spatially uncorrelated but temporally correlated proper Gaussian noise, as in [12]. But unlike [12] where multiple independent realizations of noisy signal are required, only one data realization is needed in [13]. Moreover, an analytical formulation for test threshold selection is provided in [13], unlike [12].

Improper complex signals are considered in [14], which is limited to temporally white but spatially correlated improper complex Gaussian PU signals in temporally and spatially uncorrelated proper complex Gaussian noise. The model of [15] allows temporal correlation for both improper signal and proper noise. Both show improved performance compared to the case where improper signals are treated as proper. The model of [15] is limited to improper signals in spatially independent proper noise. In this paper we allow noise to be improper also.

A preliminary conference version of this paper is [16]. Our derivation of the GLRT in this paper is significantly different from that in [16]. In [16], the starting point is the estimated PSD of an augmented complex vector sequence consisting of the original measurement sequence augmented with its complex conjugate. In this paper, we first develop a frequency-domain sufficient statistic (Sec. II-C), and then under certain sufficient conditions ((A1)-(A3) in Sec. II-D), we derive the GLRT in Sec. III based on the frequency-domain sufficient statistic, in a principled way. The estimate of the PSD of the augmented measurements follows from the frequency-domain sufficient statistic as a maximum likelihood (ML) estimate (see (26) in Sec. III).

The main contributions of this paper are as follows. Detection of improper signals has been considered in [14] for temporally white but spatially correlated improper complex Gaussian signals in temporally and spatially uncorrelated proper complex Gaussian noise. In [15], both improper signal and proper noise are allowed to be temporally correlated. In this paper we allow noise to be improper also. We follow a frequency-domain approach by first developing a frequency-domain sufficient statistic based on the discrete Fourier transform (DFT) of the time-domain data. This approach is novel. For a large class of stationary random processes, the DFT of a given stretch of measurements is complex, asymptotically jointly proper Gaussian, and independent at distinct frequencies on an appropriate frequency grid [17]. The random process need not be Gaussian. This allows one to readily write down an expression for the likelihood function of observations in terms of the asymptotic distribution of its DFT. Detailed parametric modeling of the underlying signal or noise to capture any dependencies in the noisy signal samples is not essential, and neither is the (widely used) assumption that the observed signal is Gaussian. A binary hypothesis testing approach is formulated and GLRT is derived using the the frequency-domain sufficient statistic. An asymptotic analytical solution for calculating the test threshold is provided. The performance of the GLRT is also analyzed. The results are illustrated via computer simulations.

In Sec. II, we introduce notation and state the binary hypothesis testing problem of interest in this paper. We also develop a frequency-domain sufficient statistic (Sec. II-C), introduce certain sufficient conditions ((A1)-(A3) in Sec. II-D) needed later, and review asymptotic distribution of the sufficient statistic as number of measurement samples become large. We derive the GLRT in Sec. III based on the frequency-domain sufficient statistic and its asymptotic distribution. An asymptotic analytical solution for calculating the test threshold is provided in Sec. IV. In Sec. V, the performance of the GLRT is analyzed by deriving an approximate asymptotic distribution of the GLRT statistic under a sequence of local alternative hypotheses. The results of Sec. III are then modified to address two special cases in Sec. VI: detection of improper signals in proper noise (Sec. VI-A), and detection of proper signals in proper noise (Sec. VI-B). Simulation examples are presented in Sec. VII in support of the proposed approaches.

II. Preliminaries and System Model

Here we first introduce notation in Sec. II-A, and state the binary hypothesis testing problem of interest in this paper in Sec. II-B. Then we motivate a frequency-domain sufficient statistic in Sec. II-C. Certain reasonable sufficient conditions are presented in Sec. II-D, for use later in the derivation of the GLRT. In Sec. II-E we review asymptotic distribution of the sufficient statistic as number of measurement samples become large.

A. Notation

We use $\mathbf{S} \succeq 0$ and $\mathbf{S} \succ 0$ to denote that Hermitian $\mathbf{S}$ is positive semi-definite and positive definite, respectively. For a square matrix $\mathbf{A}$, $|\mathbf{A}|$ and $\mathrm{etr}(\mathbf{A})$ denote the determinant and the exponential of the trace of $\mathbf{A}$, respectively, i.e., $\mathrm{etr}(\mathbf{A}) = \exp(\mathrm{tr}(\mathbf{A}))$, $[\mathbf{B}_k]_{i:l,j:m}$ denotes the submatrix of the matrix $\mathbf{B}_k$ comprising its rows i through l and columns j through m, $[\mathbf{B}_k]_{ij}$ is its ijth element, $\mathbf{B}_{ij}$ is the ijth element of $\mathbf{B}$, and $\mathbf{I}$ is the identity matrix. The superscripts $*$, $\top$ and H denote the complex conjugate, transpose and Hermitian (conjugate transpose) operations, respectively, and E denotes the expectation operation. The set of integers 1 through n is denoted by $[1, n]$. The sets of real and complex numbers are denoted by $\mathbb{R}$ and $\mathbb{C}$, respectively. The notation $y = \mathcal{O}(g(x))$ means that there exists some finite real number

$b > 0$ such that $\lim_{x\to\infty} |y/g(x)| \le b$. Given square matrices $\mathbf{A}_i$, $i = 1, 2, \cdots n$, block-diag$\{\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_n\}$ denotes a square matrix with $\mathbf{A}_i$s along its main block-diagonal and zeros everywhere else.

The notation χ_n^2 represents central chi-square distribution with n degrees of freedom, whereas $\chi_n^2(\lambda)$ denotes the non-central chi-square distribution with n degrees of freedom and non-centrality parameter λ. The notation $\mathbf{x} \sim \mathcal{N}_c(\mathbf{m}, \Sigma)$ denotes a random vector $\mathbf{x}$ that is proper complex Gaussian with mean $\mathbf{m}$ and covariance Σ, and $\mathbf{x} \overset{a}{\sim} \mathcal{N}_c(\mathbf{m}, \Sigma)$ implies that $\mathbf{x}$ is asymptotically $\mathcal{N}_c(\mathbf{m}, \Sigma)$ as number of measurements tend to ∞. The notation $\overset{a}{\sim}$ applies to other distributions (χ^2, Wishart, etc) as well. Also, for real $\mathbf{x}$, $\mathbf{x} \sim \mathcal{N}_r(\mathbf{m}, \Sigma)$ denotes a real random vector $\mathbf{x}$ that is Gaussian with mean $\mathbf{m}$ and covariance Σ. The abbreviations w.p.1. and i.i.d. stand for with probability one, and independent and identically distributed, respectively.

B. Binary Hypotheses

Let $\mathbf{n}(t) \in \mathbb{C}^p$ denote a zero-mean spatially independent, stationary, possibly improper, random (noise) sequence and $\mathbf{s}(t) \in \mathbb{C}^p$ denote a zero-mean stationary, possibly improper random (signal) sequence which is independent of $\{\mathbf{n}(t)\}$. Both noise and signal may be non-Gaussian. Let $\mathcal{H}_0$ denote the null hypothesis that the user is receiving just noise, and $\mathcal{H}_1$ is the alternative that signal common to all sensors is also present. We consider the following binary hypothesis testing problem for the measurement sequence $\mathbf{x}(t)$, $t = 0, 1, \cdots, N-1$,

$$\begin{array}{ll} \mathcal{H}_0 : & \mathbf{x}(t) = \mathbf{n}(t), \ \ \text{noise only} \\ \mathcal{H}_1 : & \mathbf{x}(t) = \mathbf{s}(t) + \mathbf{n}(t), \ \ \text{signal and noise.} \end{array} \tag{1}$$

We assume that noise is independent across sensors.

C. Frequency-Domain Sufficient Statistic

Define the augmented complex process $\{\mathbf{y}(t)\}$ and the real-valued process $\{\mathbf{z}(t)\}$ as

$$\mathbf{y}(t) = \begin{bmatrix} \mathbf{x}(t) \\ \mathbf{x}^*(t) \end{bmatrix}, \quad \mathbf{z}(t) = \begin{bmatrix} \mathbf{x}_r(t) \\ \mathbf{x}_i(t) \end{bmatrix} \tag{2}$$

where $\mathbf{x}(t) = \mathbf{x}_r(t) + j\mathbf{x}_i(t)$, with $\mathbf{x}_r(t)$ and $\mathbf{x}_i(t)$ denoting its real and imaginary components, respectively. We have

$$\mathbf{y}(t) = \boldsymbol{\mathcal{T}}\mathbf{z}(t), \ \ \boldsymbol{\mathcal{T}} = \begin{bmatrix} \mathbf{I} & j\mathbf{I} \\ \mathbf{I} & -j\mathbf{I} \end{bmatrix} \in \mathbb{C}^{(2p)\times(2p)}. \tag{3}$$

Consider the (normalized) DFT $\mathbf{d}_z(f_n)$ of real-valued $\mathbf{z}(t)$, $t = 0, 1, \cdots, N-1$, given by

$$\mathbf{d}_z(f_n) := \frac{1}{\sqrt{N}} \sum_{t=0}^{N-1} \mathbf{z}(t) e^{-j2\pi f_n t} \tag{4}$$

where $f_n = n/N$, $n = 0, 1, \cdots, N-1$. Note that $\mathbf{d}_z(f_n)$ is periodic in n with period N, and is periodic in normalized frequency f_n with period 1. Since $\mathbf{z}(t)$ is real-valued, we have $\mathbf{d}_z^*(f_n) = \mathbf{d}_z(-f_n) = \mathbf{d}_z(1-f_n)$, so $\mathbf{d}_z(f_n)$ for $n = 0, 1, \cdots, (N/2)$ completely determines $\mathbf{d}_z(f_n)$ for all integers n.

As proved in [20, p. 280, Sec. 6.2], for any statistical inference problem, the complete sample is a sufficient statistic, and so is any one-to-one function of a sufficient statistic. Hence, given the complex-valued sample $\mathbf{x}(t)$, $t \in [0, N-1]$, the real-valued sample $\mathbf{z}(t)$, $t \in [0, N-1]$, is a sufficient statistic. Since the inverse DFT yields (one-to-one transformation)

$$\mathbf{z}(t) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} \mathbf{d}_z(f_n) e^{j2\pi f_n t}, \tag{5}$$

the set $\{\mathbf{d}_z(f_n)\}_{n=0}^{N-1}$ is a sufficient statistic, which can be further reduced to $\{\mathbf{d}_z(f_n)\}_{n=0}^{N/2}$, exploiting symmetries $\mathbf{d}_z^*(f_n) = \mathbf{d}_z(-f_n) = \mathbf{d}_z(1-f_n)$. As in (4), define

$$\mathbf{d}_y(f_n) := \frac{1}{\sqrt{N}} \sum_{t=0}^{N-1} \mathbf{y}(t) e^{-j2\pi f_n t}. \tag{6}$$

Then we have $\mathbf{d}_y(f_n) = \boldsymbol{\mathcal{T}}\mathbf{d}_z(f_n)$. Invoking [20, p. 280, Sec. 6.2], it follows that the set of complex-valued random vectors $\{\mathbf{d}_y(f_n)\}_{n=0}^{N/2}$ is a sufficient statistic for our binary hypothesis testing problem.

D. Model Assumptions

Here we assumptions on $\mathbf{z}(t)$ or $\mathbf{y}(t)$ used later in the paper.

(A1) **Assumption 2.6.1** [17]. Random sequence $\{\mathbf{z}(t)\} \in \mathbb{R}^{2p}$ is stationary with components $\mathbf{z}_a(t)$, $a = 1, 2, \cdots, 2p$, such that $E\{|\mathbf{z}_a(t)|^k\} < \infty$, satisfying

$$\sum_{\tau_1,\tau_2,\cdots,\tau_{k-1}=-\infty}^{\infty} |c_{\mathbf{z}:a_1,a_2,\cdots,a_k}(\tau_1, \tau_2, \cdots, \tau_{k-1})| < \infty \tag{7}$$

for $a_1, a_2, \cdots, a_k = 1, 2, \cdots, 2p$ and $k = 2, 3, \cdots$, where $c_{\mathbf{z}:a_1,a_2,\cdots,a_k}(\tau_1, \tau_2, \cdots, \tau_{k-1})$ is the joint cumulant function of order k of stationary $\{\mathbf{z}(t)\}$.

(A2) The PSD matrix of $\mathbf{y}(t)$, $\mathbf{S}_y(f) \succ 0$ for any $0 \le f \le 1$.

(A3) The PSD matrix $\mathbf{S}_y(f_n)$ is locally smooth such that it is constant over K ($\ge 2p$) consecutive frequency points, where $f_n = n/N$, $n \in [0, N-1]$.

See [18, Appendix A] (also [19, Appendix A]) for more details regarding satisfaction of Assumption 2.6.1 of [17] for a class of stationary improper sequences. If $\mathbf{x}(t)$ is Gaussian, then it is sufficient to verify (7) only for $k = 2$. Regarding assumption (A2), one can always add artificial proper white Gaussian noise to $\mathbf{x}(t)$ to achieve $\mathbf{S}_y(f) \succ 0$. Assumption (A3) is a standard assumption in PSD estimation literature [17].

E. Asymptotic Distribution of Sufficient Statistic

Under (A1), it follows from [17, Theorem 4.4.1] that asymptotically, $\mathbf{d}_z(f_n)$, $n = 1, 2, \cdots, (N/2)-1$, ($N$ even), are independent proper complex Gaussian $\mathcal{N}_c(\mathbf{0}, \mathbf{S}_z(f_n))$ random vectors, respectively, where $\mathbf{S}_z(f)$ is the PSD matrix of $\mathbf{z}(t)$. Also, asymptotically, $\mathbf{d}_z(f_0)$ and $\mathbf{d}_z(f_{N/2})$ (N even) are independent real Gaussian $\mathcal{N}_r(\mathbf{0}, \mathbf{S}_z(f_0))$ and $\mathcal{N}_r(\mathbf{0}, \mathbf{S}_z(f_{N/2}))$ random vectors, respectively, independent of the previous random vectors for $n = 1, 2, \cdots, (N/2)-1$. Using (3), it follows that $\mathbf{d}_y(f_n)$, $n = 1, 2, \cdots, (N/2) - 1$, ($N$ even), are asymptotically independent proper complex Gaussian $\mathcal{N}_c(\mathbf{0}, \mathbf{S}_y(f_n))$ random vectors, respectively, where [18]

$$\mathbf{S}_y(f) = \begin{bmatrix} \mathbf{S}_x(f) & \tilde{\mathbf{S}}_x(f) \\ \tilde{\mathbf{S}}_x^*(-f) & \mathbf{S}_x^*(-f) \end{bmatrix} \tag{8}$$

$$= \begin{bmatrix} \mathbf{S}_x(f) & \tilde{\mathbf{S}}_x(f) \\ \tilde{\mathbf{S}}_x^H(f) & \mathbf{S}_x^\top(-f) \end{bmatrix}, \tag{9}$$

$\mathbf{S}_x(f) = \sum_{\tau=-\infty}^{\infty} \mathbf{R}_x(\tau) e^{-j2\pi f\tau}$ denotes the PSD of $\{\mathbf{x}(t)\}$, $\mathbf{R}_x(\tau) = E\{\mathbf{x}(t+\tau)\mathbf{x}^H(t)\}$, $\tilde{\mathbf{S}}_x(f) = \sum_{\tau=-\infty}^{\infty} \tilde{\mathbf{R}}_x(\tau) e^{-j2\pi f\tau}$ is the complementary PSD (C-PSD) of $\{\mathbf{x}(t)\}$, and $\tilde{\mathbf{R}}_x(\tau) = E\{\mathbf{x}(t+\tau)\mathbf{x}^\top(t)\}$. While $\mathbf{d}_y(f_0)$ and $\mathbf{d}_y(f_{N/2})$, (N even), are independent Gaussian random vectors, also independent of $\mathbf{d}_y(f_n)$ for $n = 1, 2, \cdots, (N/2) - 1$, they are, in general, improper whereas $\mathbf{d}_y(f_n)$ for $n = 1, 2, \cdots, (N/2) - 1$ is proper. In our derivation of the GLRT in the next section, we will ignore these two frequency points.

III. Generalized Likelihood Ratio Test

In this section we derive the GLRT. We will use the frequency-domain sufficient statistic $\{\mathbf{d}_y(f_n)\}_{n=0}^{N/2}$ after discarding the end frequency points f_0 and $f_{N/2}$, as the asymptotic distribution of the latter is improper Gaussian, whereas the DFT at other frequencies is proper complex Gaussian.

Define

$$\mathbf{D} = [\mathbf{d}_y(f_1)\ \mathbf{d}_y(f_2)\ \cdots\ \mathbf{d}_y(f_{(N/2)-1})]^H \in \mathbb{C}^{((N/2)-1)\times(2p)}. \tag{10}$$

The asymptotic joint probability density function (pdf) of $\mathbf{D}$ is given by

$$f_{\mathbf{D}}(\mathbf{D}) = \prod_{n=1}^{(N/2)-1} \frac{\exp\left(-\mathbf{d}_y^H(f_n)\mathbf{S}_y^{-1}(f_n)\mathbf{d}_y(f_n)\right)}{\pi^{2p}\,|\mathbf{S}_y(f_n)|} \tag{11}$$

where we do not distinguish between a random vector/matrix and the values taken by them in our notation (for simplicity). The PSD matrix $\mathbf{S}_y(f_n)$ is unknown. Under $\mathcal{H}_0$, the ℓth component $\mathbf{x}_\ell(t)$ of $\mathbf{x}(t)$ is independent of $\mathbf{x}_m(t)$ for $\ell \neq m$. Let ($\ell = 1, 2, \cdots, p$)

$$\mathbf{S}_y^{(\ell)}(f) := \begin{bmatrix} [\mathbf{S}_y(f)]_{\ell\ell} & [\mathbf{S}_y(f)]_{\ell(\ell+p)} \\ [\mathbf{S}_y(f)]_{(\ell+p)\ell} & [\mathbf{S}_y(f)]_{(\ell+p)(\ell+p)} \end{bmatrix}. \tag{12}$$

Then in terms of $\mathbf{S}_x$ and $\tilde{\mathbf{S}}_x$,

$$\mathbf{S}_y^{(\ell)}(f) = \begin{bmatrix} [\mathbf{S}_x(f)]_{\ell\ell} & [\tilde{\mathbf{S}}_x(f)]_{\ell\ell} \\ [\tilde{\mathbf{S}}_x^*(-f)]_{\ell\ell} & [\mathbf{S}_x^*(-f)]_{\ell\ell} \end{bmatrix}. \tag{13}$$

Under $\mathcal{H}_0$, all entries in $\mathbf{S}_y(f)$ are zeros except for those in $\mathbf{S}_y^{(\ell)}(f)$, $\ell = 1, 2, \cdots, p$. Under $\mathcal{H}_1$, $\mathbf{x}(t)$ is improper with $\mathbf{S}_y(f) \succ 0$ (assumption (A2) of Sec. II-D) with no specific structure. Testing for the presence of an improper common signal in spatially independent improper noise is then reformulated as the problem

$$\begin{aligned} \mathcal{H}_0 :\ & \mathbf{S}_y(f_n) \text{ is such that } \mathbf{S}_x(\pm f_n) \text{ and } \tilde{\mathbf{S}}_x(\pm f_n) \\ & \quad \text{are diagonal } \ \forall n \in [1, (N/2)-1] \\ \mathcal{H}_1 :\ & \mathbf{S}_y(f_n) \succ \mathbf{0} \text{ with no specific structure} \\ & \quad \forall n \in [1, (N/2)-1]. \end{aligned} \tag{14}$$

By (8), (12) and (13), the above constraint under $\mathcal{H}_0$ is equivalent to $[\mathbf{S}_y(f_n)]_{\ell m} = 0\ \forall l, m$, except for the entries in $\mathbf{S}_y^{(q)}(f_n)$, $q = 1, 2, \cdots, p$, $\forall n \in [1, (N/2) - 1]$.

Now assume that $\mathbf{S}_y(f_n)$ is locally smooth (assumption (A3) of Sec. II-D), so that $\mathbf{S}_y(f_n)$ is (approximately) constant over $K = 2m_t + 1 \geq 2p$ consecutive frequency points. Pick

$$\tilde{f}_k = \frac{(k-1)K + m_t + 1}{N}, \qquad k = 1, 2, \cdots, M, \tag{15}$$

$$M = \left\lfloor \frac{\frac{N}{2} - m_t - 1}{K} \right\rfloor, \tag{16}$$

leading to M equally spaced frequencies $\tilde{f}_k$ in the interval $(0, 0.5)$, at intervals of K/N. It is assumed that for each $\tilde{f}_k$ (local smoothness),

$$\mathbf{S}_y(\tilde{f}_{k,\ell}) = \mathbf{S}_y(\tilde{f}_k) \ \text{ for } \ell = -m_t, -m_t + 1, \cdots, m_t\,, \tag{17}$$

where

$$\tilde{f}_{k,\ell} = \frac{(k-1)K + m_t + 1 + \ell}{N}. \tag{18}$$

Using (17) in (11), we have

$$\begin{aligned} f_{\mathbf{D}}(\mathbf{D}) &= \prod_{k=1}^{M} \Big[\prod_{\ell=-m_t}^{m_t} \frac{1}{\pi^p\,|\mathbf{S}_y(\tilde{f}_k)|} \\ &\qquad \times \exp\left(-\mathbf{d}_y^H(\tilde{f}_{k,\ell})\mathbf{S}_y^{-1}(\tilde{f}_k)\mathbf{d}_y(\tilde{f}_{k,\ell})\right) \Big] \qquad (19) \\ &= \prod_{k=1}^{M} \frac{\mathrm{etr}\left(-\mathbf{S}_y^{-1}(\tilde{f}_k)\tilde{\mathbf{D}}(\tilde{f}_k)\right)}{\pi^{2Kp}\,|\mathbf{S}_y(\tilde{f}_k)|^K} \\ &= \prod_{k=1}^{M} f_{\check{\mathbf{D}}(\tilde{f}_k)}(\check{\mathbf{D}}(\tilde{f}_k)) \qquad (20) \end{aligned}$$

where $K \times (2p)$ $\check{\mathbf{D}}(\tilde{f}_k)$ is

$$\check{\mathbf{D}}(\tilde{f}_k) = \left[\mathbf{d}_y(\tilde{f}_{k,-m_t})\ \mathbf{d}_y(\tilde{f}_{k,-m_t+1})\ \cdots\ \mathbf{d}_y(\tilde{f}_{k,m_t})\right]^H, \tag{21}$$

and $(2p) \times (2p)$ $\tilde{\mathbf{D}}(\tilde{f}_k)$ is

$$\tilde{\mathbf{D}}(\tilde{f}_k) = \sum_{\ell=-m_t}^{m_t} \mathbf{d}_y(\tilde{f}_{k,\ell})\mathbf{d}_y^H(\tilde{f}_{k,\ell}). \tag{22}$$

Explicitly indicating the dependence on the underlying hypothesis $\mathcal{H}_i$, $i = 0, 1$, the pdf of $\mathbf{D}$ under $\mathcal{H}_i$ is given by

$$f_{\mathbf{D}|\mathcal{H}_i}(\mathbf{D}|\mathcal{H}_i) = \prod_{k=1}^{M} f_{\check{\mathbf{D}}(\tilde{f}_k)|\mathcal{H}_i}(\check{\mathbf{D}}(\tilde{f}_k)|\mathcal{H}_i) \tag{23}$$

where

$$f_{\check{\mathbf{D}}(\tilde{f}_k)|\mathcal{H}_i}(\check{\mathbf{D}}(\tilde{f}_k)|\mathcal{H}_i) = \frac{\mathrm{etr}(-\mathbf{S}_y^{-1}(\tilde{f}_k)\tilde{\mathbf{D}}(\tilde{f}_k))}{\pi^{2Kp}\,|\mathbf{S}_y(\tilde{f}_k)|^K}. \tag{24}$$

Under $\mathcal{H}_1$ where there is no specific structure to the PSD matrix, the unknowns in (23)-(24) are Hermitian positive definite matrices $\mathbf{S}_y(\tilde{f}_k)$, $k = 1, 2, \cdots, M$. By [21, Theorem 1.10.4], for any positive definite $m \times m$ matrices $\mathbf{A}$ and $\mathbf{B}$, and for any scalars $a > 0$ and $b > 0$, one has

$$|\mathbf{A}|^{-b}\mathrm{etr}\left(-a\mathbf{A}^{-1}\mathbf{B}\right) \leq |a\mathbf{B}/b|^{-b}\exp(-mb), \tag{25}$$

with equality if and only if $\mathbf{A} = a\mathbf{B}/b$. Applying (25) to (24) with $a = 1$, $b = K$, $\mathbf{A} = \mathbf{S}_y(\tilde{f}_k)$, $\mathbf{B} = \tilde{\mathbf{D}}(\tilde{f}_k)$ and $m = 2p$, it follows that the ML estimate $\hat{\mathbf{S}}_y(\tilde{f}_k)$ of $\mathbf{S}_y(\tilde{f}_k)$ is given by

$$\hat{\mathbf{S}}_y(\tilde{f}_k) = \frac{1}{K}\tilde{\mathbf{D}}(\tilde{f}_k) = \frac{1}{K}\sum_{l=-m_t}^{m_t} \mathbf{d}_y(\tilde{f}_{k,\ell})\mathbf{d}_y^H(\tilde{f}_{k,\ell}). \quad (26)$$

Setting $\mathbf{S}_y(\tilde{f}_k) = \hat{\mathbf{S}}_y(\tilde{f}_k)$ in (23) and (24) yields

$$\sup_{\mathbf{S}_y(\tilde{f}_k) \text{ under } \mathcal{H}_1} f_{\mathbf{D}|\mathcal{H}_1}(\mathbf{D}|\mathcal{H}_1) = \frac{e^{-2MKp}}{\pi^{2MKp}} \prod_{k=1}^{M} |\hat{\mathbf{S}}_y(\tilde{f}_k)|^{-K}. \quad (27)$$

Exploiting the structure of $\mathbf{S}_y(\tilde{f}_k)$ under $\mathcal{H}_0$, the pdf of $\check{\mathbf{D}}(\tilde{f}_k)$ under $\mathcal{H}_0$ can be simplified as

$$f_{\check{\mathbf{D}}(\tilde{f}_k)|\mathcal{H}_0}(\check{\mathbf{D}}(\tilde{f}_k)|\mathcal{H}_0) = \prod_{q=1}^{p} \frac{\text{etr}(-\mathbf{S}_y^{-(q)}(\tilde{f}_k)\tilde{\mathbf{D}}^{(q)}(\tilde{f}_k))}{\pi^{2Kp}\,|\mathbf{S}_y^{(q)}(\tilde{f}_k)|^K} \quad (28)$$

where $\mathbf{S}_y^{(q)}(\tilde{f}_k)$ is defined as in (12), $\mathbf{S}_y^{-(q)}(\tilde{f}_k)$ is its inverse, and $\tilde{\mathbf{D}}^{(q)}(\tilde{f}_k)$ is defined similarly as ($q = 1, 2, \cdots, p$)

$$\tilde{\mathbf{D}}^{(q)}(\tilde{f}_k) := \begin{bmatrix} \tilde{\mathbf{D}}_{qq}(\tilde{f}_k) & \tilde{\mathbf{D}}_{q(q+p)}(\tilde{f}_k) \\ \tilde{\mathbf{D}}_{(q+p)q}(\tilde{f}_k) & \tilde{\mathbf{D}}_{(q+p)(q+p)}(\tilde{f}_k) \end{bmatrix}. \quad (29)$$

Under $\mathcal{H}_0$, the unknowns in (23) and (28) are Hermitian positive definite matrices $\mathbf{S}_y^{(q)}(\tilde{f}_k)$, $q = 1, 2, \cdots, p$, $k = 1, 2, \cdots, M$. Applying (25) to (28) with $a = 1$, $b = K$, $\mathbf{A} = \mathbf{S}_y^{(q)}(\tilde{f}_k)$, $\mathbf{B} = \tilde{\mathbf{D}}^{(q)}(\tilde{f}_k)$ and $m = 2$, it follows that the ML estimate $\hat{\mathbf{S}}_y^{(q)}(\tilde{f}_k)$ of $\mathbf{S}_y^{(q)}(\tilde{f}_k)$ is given by

$$\hat{\mathbf{S}}_y^{(q)}(\tilde{f}_k) = \frac{1}{K}\tilde{\mathbf{D}}^{(q)}(\tilde{f}_k) \in \mathbb{C}^{2\times 2} \quad (30)$$

where

$$\tilde{\mathbf{D}}_{i_1 i_2}^{(q)}(\tilde{f}_k) = \sum_{l=-m_t}^{m_t} [\mathbf{d}_y(\tilde{f}_{k,\ell})]_{i_1} [\mathbf{d}_y(\tilde{f}_{k,\ell})]_{i_2}^*. \quad (31)$$

Setting $\mathbf{S}_y^{(q)}(\tilde{f}_k) = \hat{\mathbf{S}}_y^{(q)}(\tilde{f}_k)$ in (23) and (28) leads to

$$\sup_{\mathbf{S}_y(\tilde{f}_k) \text{ under } \mathcal{H}_0} f_{\mathbf{D}|\mathcal{H}_0}(\mathbf{D}|\mathcal{H}_0) = \frac{e^{-2MKp}}{\pi^{2MKp}} \prod_{k=1}^{M}\prod_{q=1}^{p} |\hat{\mathbf{S}}_y^{(q)}(\tilde{f}_k)|^{-K} \quad (32)$$

Using (27) and (32) we obtain the GLRT

$$\mathcal{L} := \frac{\sup_{\mathbf{S}_y(\tilde{f}_k) \text{ under } \mathcal{H}_1} f_{\mathbf{D}|\mathcal{H}_1}(\mathbf{D}|\mathcal{H}_1)}{\sup_{\mathbf{S}_y(\tilde{f}_k) \text{ under } \mathcal{H}_0} f_{\mathbf{D}|\mathcal{H}_0}(\mathbf{D}|\mathcal{H}_0)} \quad (33)$$

$$= \prod_{k=1}^{M} \underbrace{\frac{\prod_{q=1}^{p} |\hat{\mathbf{S}}_y^{(q)}(\tilde{f}_k)|^K}{|\hat{\mathbf{S}}_y(\tilde{f}_k)|^K}}_{\mathcal{L}_k} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \tau_1 \quad (34)$$

where the threshold τ_1 is picked to achieve a pre-specified probability of false alarm $P_{fa} = P\{\mathcal{L} \geq \tau_1 \,|\, \mathcal{H}_0\}$. This requires pdf of $\mathcal{L}$ under $\mathcal{H}_0$ which is discussed in Sec. IV.

Note that $\hat{\mathbf{S}}_y(\tilde{f}_k)$ in (26) is an estimator of the PSD of $\mathbf{y}(t)$ at frequency $\tilde{f}_k$, based on unweighted smoothing in frequency-domain, as given in [17, Eqn. (7.3.2)]. Based on [17, Theorem 7.3.3], it is shown in [18, Sec. II-A.3] that as $N \to \infty$, $\hat{\mathbf{S}}_y(\tilde{f}_k)$ is distributed as $W_C\left(2p, K, K^{-1}\mathbf{S}_y(\tilde{f}_k)\right)$ (denoted as $\hat{\mathbf{S}}_y(\tilde{f}_k) \overset{a}{\sim} W_C(2p, K, K^{-1}\mathbf{S}_y(\tilde{f}_k))$) where $W_C\left(2p, K, K^{-1}\mathbf{S}_y(\tilde{f}_k)\right)$ denotes the complex Wishart distribution of dimension $2p$, degrees of freedom K, and mean value $\mathbf{S}_y(\tilde{f}_k)$. If a random matrix $\mathbf{X} \sim W_C(m, K, \mathbf{S}(f))$, then by [17, Sec. 4.2], $E\{\mathbf{X}\} = K\mathbf{S}(f)$, $\text{cov}\{\mathbf{X}_{jk}, \mathbf{X}_{rs}\} = K\mathbf{S}_{jr}(f)\mathbf{S}_{ks}^*(f)$, and for $K \geq m$, the probability density function (pdf) of $\mathbf{X}$ is given by

$$f_{\mathbf{X}}(\mathbf{X}) = \frac{1}{\Gamma_m(K)} \frac{1}{|\mathbf{S}(f)|^K} |\mathbf{X}|^{K-m} \,\text{etr}\{-\mathbf{S}^{-1}(f)\mathbf{X}\} \quad (35)$$

where the pdf (35) is defined for $\mathbf{S}(f) \succ 0$ and $\mathbf{X} \succeq 0$, and is otherwise zero, and

$$\Gamma_m(K) := \pi^{m(m-1)/2} \prod_{j=1}^{m} \Gamma(K - j + 1) \quad (36)$$

where $\Gamma(n)$ denotes the (complete) Gamma function $\Gamma(z) := \int_0^\infty t^{z-1}e^{-t}\,dt$.

Remark 1. Assumption (A3) is crucial in deriving (34). The true PSD $\mathbf{S}_y(f_n)$ is unknown and must be estimated from data, e.g., as in (26). One needs $K \geq 2p$ for $\hat{\mathbf{S}}_y(\tilde{f}_k)$ to be of full rank w.p.1. But since $\mathbf{d}_y(f_0)$ and $\mathbf{d}_y(f_{N/2})$ are statistically different from $\mathbf{d}_y(f_n)$ for $n = 1, 2, \cdots, (N/2)-1$, (the former are improper while the latter are proper), one can not average as in (26) by including $\mathbf{d}_y(f_0)$ and/or $\mathbf{d}_y(f_{N/2})$ therein. On the other hand, as N increases, we still have just two frequencies f_0 and $f_{N/2}$ where $\mathbf{d}_y(.)$ is improper, thereby precluding sufficient number of samples for estimators of $\mathbf{S}_y(f_0)$ or $\mathbf{S}_y(f_{N/2})$ to be of full rank. This leaves us with no choice but to exclude $\mathbf{d}_y(f_0)$ and $\mathbf{d}_y(f_{N/2})$ from the proposed GLRT. The impact of such an exclusion on the test performance is not clear. □

A. Invariance of GLRT

Note that $\mathcal{L}_k = \prod_{q=1}^{p} |\hat{\mathbf{S}}_y^{(q)}(\tilde{f}_k)|^K / |\hat{\mathbf{S}}_y(\tilde{f}_k)|^K$, defined in (34), is invariant to the transformation $\hat{\mathbf{S}}_y^{(q)}(\tilde{f}_k) \to \mathbf{A}_k^{(q)}\hat{\mathbf{S}}_y^{(q)}(\tilde{f}_k)\mathbf{A}_k^{(q)H}$ for any non-singular Hermitian $\mathbf{A}_k^{(q)} \in \mathbb{C}^{2\times 2}$, leaving the other entries of $\hat{\mathbf{S}}_y(\tilde{f}_k)$ unchanged. More explicitly, define ($k = 1, 2, \cdots, M$)

$$\mathbf{A}_k = \text{block}-\text{diag}\left\{\mathbf{A}_k^{(1)},\ \mathbf{A}_k^{(2)}, \cdots, \mathbf{A}_k^{(p)}\right\}, \quad (37)$$

and $\tilde{\mathbf{y}}(t) \in \mathbb{C}^{(2p)\times(2p)}$,

$$\tilde{\mathbf{y}}(t) = \left[\mathbf{x}_1(t)\ \mathbf{x}_1^*(t)\ \mathbf{x}_2(t)\ \mathbf{x}_2^*(t)\ \cdots\ \mathbf{x}_p(t)\ \mathbf{x}_p^*(t)\right]^\top. \quad (38)$$

Then there exists a permutation matrix $\mathbf{J}$ such that $\tilde{\mathbf{y}}(t) = \mathbf{J}\mathbf{y}(t)$. Define $\tilde{\mathbf{A}}_k = \mathbf{J}\mathbf{A}_k$. The statistic $\mathcal{L}_k$ is invariant to the transformation $\hat{\mathbf{S}}_y(\tilde{f}_k) \to \tilde{\mathbf{A}}_k\hat{\mathbf{S}}_y(\tilde{f}_k)\tilde{\mathbf{A}}_k^H$, as can be verified via direct substitution. To consider invariance of $\mathcal{L}$, define

$$\mathcal{S} = \text{block}-\text{diag}\left\{\hat{\mathbf{S}}_y(\tilde{f}_1),\ \hat{\mathbf{S}}_y(\tilde{f}_2),\ \cdots,\ \hat{\mathbf{S}}_y(\tilde{f}_M)\right\}, \quad (39)$$

and also define

$$\mathcal{A} = \text{block}-\text{diag}\left\{\mathbf{J}\mathbf{A}_1,\ \mathbf{J}\mathbf{A}_2,\ \cdots,\ \mathbf{J}\mathbf{A}_M\right\}. \quad (40)$$

Let $\mathcal{G}$ denote the group of transformations $\mathcal{G} = \{\mathcal{A} \text{ as in (40)}\}$, where $\mathcal{A}$ acts on $\mathcal{S}$ as $\mathcal{S} \to \mathcal{A}\mathcal{S}\mathcal{A}^H$. It is easily verified via direct substitution that under the action of group $\mathcal{G}$, $\mathcal{L}$ in (34) is invariant.

The invariance under $\mathcal{G}$ allows us to transform any $\hat{\mathbf{S}}_y(\tilde{f}_k)$ to $\tilde{\mathbf{S}}_y(\tilde{f}_k)$ such that $\tilde{\mathbf{S}}_y^{(q)}(\tilde{f}_k) \sim W_C(2, K, \mathbf{I})$ and $\tilde{\mathbf{S}}_y(\tilde{f}_k) \sim W_C(2p, K, \mathbf{I})$ under $\mathcal{H}_0$, by choosing $\mathbf{A}_k^{(q)} = \sqrt{K}(\mathbf{S}_y^{(q)}(\tilde{f}_k))^{-1/2}$ for $q \in [1,p]$. Then $\mathcal{L}$ is invariant, and the transformed $\tilde{\mathbf{S}}_y(\tilde{f}_k)$s now correspond to a proper i.i.d. (white) sequence $\{\mathbf{x}(t)\}$, which can be used to compute the test threshold via Monte Carlo simulations. This threshold is valid for any other PSD. However, in Sec. IV, we offer an analytical approach that is valid for large sample sizes.

IV. Threshold Selection

We now turn to determination of an asymptotic expansion of the distribution of $\mathcal{L}$ under $\mathcal{H}_0$ following [22]–[24]. Our main result is in Theorem 1, which allows us to calculate the test threshold analytically. Theorem 1 is stated below, and proved in Appendix A.

In the following, χ_n^2 denotes a random variable with central chi-square distribution with n degrees of freedom as well as the distribution itself. Also, $B_r(t)$ denotes the Bernoulli polynomial of degree r and order unity. The first five Bernoulli polynomials are ($B_0(t) = 1$):

$$\begin{aligned} &B_1(t) = t - (1/2), \quad B_2(t) = t^2 - t + (1/6), \\ &B_3(t) = t^3 - (3/2)t^2 + (1/2)t, \\ &B_4(t) = t^4 - 2t^3 + t^2 - (1/30), \\ &B_5(x) = t^5 - (5/2)t^4 + (5/3)t^3 - (1/6)t. \end{aligned}$$

Theorem 1. The GLRT for (14) is given by

$$2\rho \ln(\mathcal{L}) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \tau$$

where

$$\rho = 1 - \frac{2(p+1)}{3K}, \tag{41}$$

$$\ln(\mathcal{L}) = K \sum_{k=1}^{M} \Bigg(\big[\sum_{q=1}^{p} \ln|\hat{\mathbf{S}}_y^{(q)}(\tilde{f}_k)| \big] - \ln(|\hat{\mathbf{S}}_y(\tilde{f}_k)|) \Bigg). \tag{42}$$

The threshold τ is picked to achieve a pre-specified $P_{fa} = 1 - P\{2\rho \ln(\mathcal{L}) \le \tau \,|\, \mathcal{H}_0\}$ where

$$\begin{aligned} P\{2\rho \ln(\mathcal{L}) \le z \,|\, \mathcal{H}_0\} &= P\{\chi_\nu^2 \le z\} + \omega_2 \big[P\{\chi_{\nu+4}^2 \le z\} \\ &- P\{\chi_\nu^2 \le z\} \big] + \omega_3 \big[P\{\chi_{\nu+6}^2 \le z\} - P\{\chi_\nu^2 \le z\} \big] \\ &+ \big\{ \omega_4 \big[P\{\chi_{\nu+8}^2 \le z\} - P\{\chi_\nu^2 \le z\} \big] + \frac{1}{2}\omega_2^2 \big[P\{\chi_{\nu+8}^2 \le z\} \\ &- 2P\{\chi_{\nu+4}^2 \le z\} + P\{\chi_\nu^2 \le z\} \big] \big\} + \mathcal{O}(M/K^5), \end{aligned} \tag{43}$$

$$\nu = 4Mp(p-1) \tag{44}$$

and

$$\begin{aligned} \omega_r = &\frac{(-1)^{r+1} M}{r(r+1)(\rho K)^r} \Bigg\{ \Bigg(\sum_{l=1}^{2p} B_{r+1}((1-\rho)K + 1 - l) \Bigg) \\ &- p \Bigg(\sum_{l=1}^{2} B_{r+1}((1-\rho)K + 1 - l) \Bigg) \Bigg\}. \end{aligned} \tag{45}$$

The discussion in [18, Remark 1] applies here as well, and is restated below.

Remark 2. It follows from (16) that $MK = \mathcal{O}(N)$. As $N \to \infty$, we let both M and $K \to \infty$. Let $K = \mathcal{O}(N^\beta)$, $0 < \beta < 1$. Then $M = \mathcal{O}(N^{1-\beta})$. We also require $\lim_{N\to\infty}(M/K) = 0$, which implies that $0.5 < \beta < 1$. It is also worth noting that for a given p, we must have $K \ge 2p$, otherwise the Wishart density such as (35) does not exist. Therefore, by (41), $\lim_{N\to\infty} \rho = 1$, and by (45), $\omega_r = \mathcal{O}(M/K^r) \to 0$ for $r \ge 1$ as $N \to \infty$. Hence, by (43), as $N \to \infty$, under $\mathcal{H}_0$, we have

$$2\ln(\mathcal{L}) \sim \chi_{4Mp(p-1)}^2 \tag{46}$$

based on $P\{2\ln(\mathcal{L}) \le z \,|\, \mathcal{H}_0\} = P\{\chi_\nu^2 \le z\} + \mathcal{O}(M/K)$. As pointed out in [22]–[24], under $\mathcal{H}_0$, the distribution $\chi_{4Mp(p-1)}^2$ is often not accurate unless N is "quite large;" hence the extra terms and the Bartlett scale factor ρ in Theorem 1. Also, by Theorem 1, we have the following modifications to (43):

$$P\{2\rho \ln(\mathcal{L}) \le z \,|\, \mathcal{H}_0\} = P\{\chi_\nu^2 \le z\} + \mathcal{O}(M/K^2), \tag{47}$$

$$\begin{aligned} P\{2\rho \ln(\mathcal{L}) \le z \,|\, \mathcal{H}_0\} &= P\{\chi_\nu^2 \le z\} \\ &+ \omega_2 \big[P\{\chi_{\nu+4}^2 \le z\} - P\{\chi_\nu^2 \le z\} \big] + \mathcal{O}(M/K^3), \end{aligned} \tag{48}$$

and

$$\begin{aligned} P\{2\rho \ln(\mathcal{L}) \le z \,|\, \mathcal{H}_0\} &= P\{\chi_\nu^2 \le z\} \\ &+ \omega_2 \big[P\{\chi_{\nu+4}^2 \le z\} - P\{\chi_\nu^2 \le z\} \big] \\ &+ \omega_3 \big[P\{\chi_{\nu+6}^2 \le z\} - P\{\chi_\nu^2 \le z\} \big] + \mathcal{O}(M/K^4), \end{aligned} \tag{49}$$

where $\nu = 4Mp(p-1)$. How many terms to use is application dependent. □

Remark 3. For a given N, how should one choose M and K ($= 2m_t + 1$)? This is an unresolved issue. As noted in Remark 1, in (26) one averages over K samples, based on model assumption (A3). One must have $K \ge 2p$ for $\hat{\mathbf{S}}_y(\tilde{f}_k)$ to be of full rank w.p.1. Larger values of K will lead to smaller M and more accurate PSD estimators, provided that the true PSD is approximately constant over the chosen K samples. This is problem dependent, and exploratory PSD analysis should offer some guidance in this respect. In Sec. VII (see Fig. 4), we investigate the influence of various choices of M and K via simulations. This examples does show improving performance with increasing K. □

V. Performance: Probability of Detection

In this section, we derive asymptotic distribution of the test statistic under $\mathcal{H}_1$, so that one can calculate the power (probability of detection) of the GLRT. We first derive a result similar to (46) invoking [28] under local alternatives, and then heuristically extend it to an expression similar to (43). Arguments are offered for this heuristic extension, and later it is illustrated via simulations in Sec. VII and is shown to offer very good agreement with the empirical results.

Let $\mathbf{S}_{y0}(f_n)$ denote the restriction of $\mathbf{S}_y(f_n)$ to null hypothesis $\mathcal{H}_0$ (see (14)). Similar to [28], [29], consider a sequence of local alternatives $\mathcal{H}_{1K}$

$$\mathcal{H}_{1K} \,:\, \mathbf{S}_y(f_n) = \mathbf{S}_{y0}(f_n) + \frac{1}{\sqrt{K}}\mathbf{S}_{y*}(f_n), \;\; n \in [1,(N/2)-1] \tag{50}$$

where $\mathbf{S}_{y*}(f_n) \succeq \mathbf{0}$ is some fixed Hermitian matrix. As $K \to \infty$, the hypothesis sequence $\mathcal{H}_{1K}$ approaches $\mathcal{H}_0$, hence the term local alternative. Under a fixed (non-local) alternative, we would have a fixed $\mathbf{S}_y(f_n) = \mathbf{S}_{y0}(f_n) + \mathbf{S}_{y*}(f_n)$. As observed in [29] in a different context, the asymptotic distributions under a fixed non-local alternative hypothesis do not give good approximation when the alternative is close to the null hypothesis. Invoking [28, Theorem 1], one can prove Theorem 2.

Theorem 2. Under the sequence of local alternatives (50) for some fixed $\mathbf{S}_{y*}(f_n)$s, $n \in [1,(N/2)-1]$, where $\mathbf{S}_{y0}(f_n)$ and $\mathbf{S}_{y*}(f_n)$ satisfy model assumption (A3), the statistic $\ln(\mathcal{L})$ specified in (42) satisfies

$$P\{2\,\ln(\mathcal{L}) \le z \,|\, \mathcal{H}_1\} = P\{\chi^2_\nu(\lambda) \le z\} + \mathcal{O}(M/\sqrt{K}), \tag{51}$$

for some $\lambda > 0$, and $\nu = 4Mp(p-1)$ as in (44).

Proof. Consider $\mathcal{L}_k$ in (34) which is given by

$$\mathcal{L}_k = \frac{\sup_{\mathbf{S}_y(\tilde{f}_k) \text{ under } \mathcal{H}_1} f_{\check{\mathbf{D}}(\tilde{f}_k)|\mathcal{H}_1}(\check{\mathbf{D}}(\tilde{f}_k)|\mathcal{H}_1)}{\sup_{\mathbf{S}_y(\tilde{f}_k) \text{ under } \mathcal{H}_0} f_{\check{\mathbf{D}}(\tilde{f}_k)|\mathcal{H}_0}(\check{\mathbf{D}}(\tilde{f}_k)|\mathcal{H}_0)}. \tag{52}$$

The sample $\check{\mathbf{D}}(\tilde{f}_k)$ given by (21) consists of K independent and identically distributed, proper complex Gaussian vectors $\mathbf{d}_y(\tilde{f}_{k,\ell})$, $\ell = -m_t, -m_t+1, \cdots, m_t-1, m_t$, each $\sim \mathcal{N}_c\Big(\mathbf{0}, \mathbf{S}_y(\tilde{f}_k)\Big)$. The augmented real-valued vector $[\text{real}(\mathbf{d}_y^\top(\tilde{f}_{k,\ell}) \;\; \text{imag}(\mathbf{d}_y^\top(\tilde{f}_{k,\ell})]^\top] \in \mathbb{R}^{2p}$ is multivariate real Gaussian, and by Example 1 of [28], the derivatives up to and including those of order three, of $\ln(f_{\check{\mathbf{D}}(\tilde{f}_k)}(\check{\mathbf{D}}(\tilde{f}_k)))$ with respect to (w.r.t.) the elements of $\mathbf{S}_y(\tilde{f}_k)$, are bounded and integrable w.r.t. $\mathbf{d}_y(\tilde{f}_{k,\ell})$, $\ell = -m_t, -m_t+1, \cdots, m_t$. That is, all regularity conditions are satisfied for [28, Theorem 1] to hold true, which when applied to $\mathcal{L}_k$, leads to

$$P\{2\,\ln(\mathcal{L}_k) \le z \,|\, \mathcal{H}_1\} = P\{\chi^2_{\nu_k}(\lambda_k) \le z\} + \mathcal{O}(1/\sqrt{K}), \tag{53}$$

where $\lambda_k > 0$ and $\nu_k = 4p(p-1)$. Note that ν_k equals the number of unknowns under $\mathcal{H}_1$ minus the number of unknowns under $\mathcal{H}_0$. Since $\mathbf{S}_y(\tilde{f}_k) \in \mathbb{C}^{(2p)\times(2p)}$ is Hermitian, its $2p$ diagonal elements are real, with $2p$ unknowns, and off-diagonal elements are complex-valued with $2p(2p-1)/2$ real as well as imaginary unknowns in the upper triangle, leading to total $4p^2$ unknowns under $\mathcal{H}_1$. Under $\mathcal{H}_0$, there are p nonzero 2×2 Hermitian submatrices, with total $4p$ unknowns. In our later developments, we do not use the expression for λ_k as given in [28], except to note that $\lambda_k > 0$, and it equals zero only if $\mathbf{S}_{y*}(f_n) \equiv \mathbf{0}$. Since $\mathcal{L}_k$s are independent, $\ln(\mathcal{L}) = \sum_{k=1}^M \ln(\mathcal{L}_k)$, and for M independent chi-square random variables, $\sum_{k=1}^M \chi^2_{\nu_k}(\lambda_k) \sim \chi^2_\nu(\lambda)$ for $\lambda = \sum_{k=1}^M \lambda_k$ and $\nu = \sum_{k=1}^M \nu_k$, we have the desired result (51). $\square$

We now use an indirect approach to compute an approximation to λ (also used in [18], [19] in a different problem). We note that asymptotically, under $\mathcal{H}_1$, we must have $E\{2\,\ln(\mathcal{L})\} = 4Mp(p-1) + \lambda$, since $2\ln(\mathcal{L}) \overset{a}{\sim} \chi^2_{4Mp(p-1)}(\lambda)$ by Theorem 2, assuming $M/\sqrt{K} \to 0$ as $N \to \infty$. We now derive an expression for the non-centrality parameter λ via $E\{2\ln(\mathcal{L})\}$. In Appendix B, we establish Lemma 1.

Lemma 1. Under $\mathcal{H}_1$, we have

$$E\{2\ln(\mathcal{L}) \,|\, \mathcal{H}_1\} = 2K\Big[\Big(\sum_{k=1}^M \Big\{ \sum_{q=1}^p \ln\big|\mathbf{S}_y^{(q)}(\tilde{f}_k)\big|\Big\} - \ln\big|\mathbf{S}_y(\tilde{f}_k)\big|\Big) + \frac{2Mp(p-1)}{K} + \mathcal{O}\big(\frac{M}{K^2}\big)\Big]. \tag{54}$$

Using Lemma 1, we have an expression for the non-centrality parameter as

$$\begin{aligned}\lambda =& E\{2\ln(\mathcal{L} \,|\, \mathcal{H}_1\} - 4Mp(p-1)\\ =& 2K\sum_{k=1}^M \Big\{\Big(\sum_{q=1}^p \ln\big|\mathbf{S}_y^{(q)}(\tilde{f}_k)\big|\Big) - \ln\big|\mathbf{S}_y(\tilde{f}_k)\big|\Big\}\\ &+ \mathcal{O}(M/K).\end{aligned} \tag{55}$$

Remark 4. As discussed in Remark 1, under $\mathcal{H}_0$, the distribution $\chi^2_{4Mp(p-1)}$ is often not accurate, hence the extra terms and factor ρ in Theorem 1. Any expression for $P\{2\rho\,\ln(\mathcal{L}) \le \tau \,|\, \mathcal{H}_1\}$ must be such that, as the hypothesis $\mathcal{H}_1$ approaches $\mathcal{H}_0$, it approaches $P\{2\rho\,\ln(\mathcal{L}) \le \tau \,|\, \mathcal{H}_0\}$. Notice that as $\lambda \to 0$, i.e., $\mathcal{H}_1$ approaches $\mathcal{H}_0$, (51) tends to the distribution specified by (46). Just as Theorem 1 modifies the distribution $\chi^2_{2Mp^2}$ under $\mathcal{H}_0$, we follow (43) to modify the distribution $\chi^2_{2Mp^2}(\lambda)$ under $\mathcal{H}_1$, exploiting Theorem 1:

$$\begin{aligned}P\{2\rho\,\ln(\mathcal{L}) \le \tau \,|\, \mathcal{H}_1\} =& P\{\chi^2_\nu(\rho\lambda) \le \tau\}\\ &+ \omega_2\big[P\{\chi^2_{\nu+4}(\rho\lambda) \le \tau\} - P\{\chi^2_\nu(\rho\lambda) \le \tau\}\big]\\ &+ \omega_3\big[P\{\chi^2_{\nu+6}(\rho\lambda) \le \tau\} - P\{\chi^2_\nu(\rho\lambda) \le \tau\}\big]\\ &+ \Big\{\omega_4\big[P\{\chi^2_{\nu+8}(\rho\lambda) \le \tau\} - P\{\chi^2_\nu(\rho\lambda) \le \tau\}\big]\\ &+ \frac{1}{2}\omega_2^2\big[P\{\chi^2_{\nu+8}(\rho\lambda) \le \tau\} - 2P\{\chi^2_{\nu+4}(\rho\lambda) \le \tau\}\\ &+ P\{\chi^2_\nu(\rho\lambda) \le \tau\}\big]\Big\} + \mathcal{O}(M/K^5),\end{aligned} \tag{56}$$

where $\nu = 4Mp(p-1)$, and λ is given by (55). Under $\mathcal{H}_0$, $\lambda = 0$, and under $\mathcal{H}_1$, as the hypothesis $\mathcal{H}_1$ approaches $\mathcal{H}_0$, $\lambda \to 0$, since $\frac{1}{\sqrt{K}}\mathbf{S}_{y*}(\tilde{f}_k)$ becomes vanishingly small $\forall \tilde{f}_k$. Under both these cases, (43) and (56) become the same. Thus, use of (56) is well justified. We do need $M/\sqrt{K} \to 0$ as $N \to \infty$. $\square$

Remark 5. Theoretically more sound expressions and methodology (as opposed to heuristic arguments used in deriving (56)) for probability of detection may be found in [28]–[33] for a wide variety of detection problems (but not for the problem considered in this paper). In [30], [31] a moment-matching approach is used to approximate the non-null distribution via a beta random variable, whereas in [32], [33], a chi-square distribution approximation is utilized. All these approaches require a detailed examination of a series expansion of the characteristic (or moment generating) function of the test statistic under the alternative hypothesis. The application of such techniques to our problem is of considerable interest but appears to be rather involved, and is left for future research. Our (56) is illustrated via simulations in Sec. VII (see Fig. 5) where very good agreement with the empirical results can be seen. □

VI. GLRTs for Two Special Cases

Here the results of Secs. III and IV are modified to address two special cases: detection of improper signals in proper noise (Sec. VI-A), and detection of proper signals in proper noise (Sec. VI-B). The main results are in Theorem 3 for detection of improper signals in proper noise, and in Theorem 4 for detection of proper signals in proper noise. A version of Theorem 3 was previously presented in [15], but was derived directly from estimated PSD of augmented complex measurement vector sequence in an ad hoc fashion, whereas in Sec. VI-A, we derive it using the frequency-domain sufficient statistic introduced in Sec. II-C. A version of Theorem 4 was previously presented in [13], but was derived directly from estimated PSD of the complex measurement vector sequence (not augmented), whereas in Sec. VI-B, we derive it using the frequency-domain sufficient statistic introduced in Sec. II-C and based on augmented complex measurement vector sequence.

A. Improper Signals in Proper Noise

Here we consider the hypothesis testing problem (1) under the restriction that noise $\mathbf{n}(t)$ is known to be proper (e.g., thermal noise at the receiver with no IQ-imbalance), but signal could be improper. If $\{\mathbf{x}(t)\}$ is proper, its C-PSD vanishes, rendering $\mathbf{S}_y^{(\ell)}(f)$ defined in (13) diagonal

$$\mathbf{S}_y^{(\ell)}(f) = \begin{bmatrix} [\mathbf{S}_x(f)]_{\ell\ell} & 0 \\ 0 & [\mathbf{S}_x^*(-f)]_{\ell\ell} \end{bmatrix}. \tag{57}$$

We will denote this restriction of $\mathbf{S}_y^{(\ell)}(f)$ as $\mathbf{S}_y^{(\ell P)}(f)$, and as in (12), it is then given by

$$\mathbf{S}_y^{(\ell P)}(f) := \begin{bmatrix} [\mathbf{S}_y(f)]_{\ell\ell} & 0 \\ 0 & [\mathbf{S}_y(f)]_{(\ell+p)(\ell+p)} \end{bmatrix}. \tag{58}$$

Mimicking Sec. III, testing for the presence of an improper common signal in spatially independent proper noise is then reformulated as the problem

$$\begin{aligned} \mathcal{H}_0 : \quad & [\mathbf{S}_y(f_n)]_{\ell m} = 0 \ \forall l, m, \text{ except for the entries in} \\ & \quad \mathbf{S}_y^{(qP)}(f_n),\ q = 1, 2, \cdots, p,\ \forall n \in [1, (N/2) - 1] \\ \mathcal{H}_1 : \quad & \mathbf{S}_y(f_n) \succ 0 \text{ with no specific structure} \\ & \quad \forall n \in [1, (N/2) - 1]. \end{aligned} \tag{59}$$

Mimicking the developments in Sec. III leading to (34), for problem (59), we obtain the GLRT

$$\mathcal{L}^{(P)} := \frac{\sup_{\mathbf{S}_y(\tilde{f}_k) \text{ under } \mathcal{H}_1} f_{\mathbf{D}|\mathcal{H}_1}(\mathbf{D}|\mathcal{H}_1)}{\sup_{\mathbf{S}_y(\tilde{f}_k) \text{ under } \mathcal{H}_0} f_{\mathbf{D}|\mathcal{H}_0}(\mathbf{D}|\mathcal{H}_0)} \tag{60}$$

$$= \prod_{k=1}^{M} \frac{\prod_{q=1}^{p} \left([\hat{\mathbf{S}}_y(\tilde{f}_k)]_{qq} \times [\hat{\mathbf{S}}_y(\tilde{f}_k)]_{(q+p)(q+p)} \right)^K}{|\hat{\mathbf{S}}_y(\tilde{f}_k)|^K} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \tau_2, \tag{61}$$

where (cf. (30)-(31)),

$$[\hat{\mathbf{S}}_y(\tilde{f}_k)]_{ii} = \sum_{l=-m_t}^{m_t} |[\mathbf{d}_y(\tilde{f}_{k,\ell})]_i|^2. \tag{62}$$

Pick $\mathbf{A}_k^{(q)}$ of Sec. III-A to be diagonal. Then $\mathcal{L}^{(P)}$ is invariant under the group action $\mathcal{S} \to \mathcal{A}\mathcal{S}\mathcal{A}^H$, where $\mathcal{S}$ and $\mathcal{A}$ are as in Sec. III-A with $\mathbf{A}_k^{(q)}$ a non-singular, positive-definite diagonal matrix. In particular, if we pick $\mathbf{A}_k^{(q)} = \sqrt{K}\text{diag}\{([\mathbf{S}_y(\tilde{f}_k)]_{qq})^{-1/2},\ ([\mathbf{S}_y(\tilde{f}_k)]_{(q+p)(q+p)})^{-1/2}\}$, then $\mathcal{L}^{(P)}$ is invariant, and transformation of $\hat{\mathbf{S}}_y(\tilde{f}_k)$ to $\tilde{\mathbf{S}}_y(\tilde{f}_k)$ leads to $\tilde{\mathbf{S}}_y^{(q)}(\tilde{f}_k) \sim W_C(2, K, \mathbf{I})$ and $\tilde{\mathbf{S}}_y(\tilde{f}_k) \sim W_C(2p, K, \mathbf{I})$ under $\mathcal{H}_0$.

The counterpart to Lemma 3 turns out to be

$$E\left\{ \frac{1}{\left(\mathcal{L}^{(P)}\right)^h} \,|\, \mathcal{H}_0 \right\} = \frac{\prod_{k=1}^{2p} \Gamma^M(K(1+h) - k + 1)}{\Gamma^{2Mp}(K(1+h))} \times \frac{\Gamma^{2Mp}(K)}{\prod_{\ell=1}^{2p} \Gamma^M(K - \ell + 1)}. \tag{63}$$

Comparing (95) with (63), we find the correspondence

$$\begin{aligned} & a = 2Mp,\ b = 2Mp,\ x_k = K, \\ & \xi_k = -[(k-1)\,\text{mod}(2p)] \text{ for } k = 1, 2, \cdots, a, \\ & y_j = K,\ \eta_j = 0 \text{ for } j = 1, 2, \cdots, b, \\ & C = \frac{\Gamma^{2Mp}(K)}{\prod_{\ell=1}^{2p} \Gamma^M(K - \ell + 1)}. \end{aligned} \tag{64}$$

Comparing (63) and (95), we further have

$$\beta_k = (1-\rho)K \ \forall k, \quad \epsilon_j = (1-\rho)K \ \forall j. \tag{65}$$

Using the above values in Lemma 4, we have

$$\begin{aligned} & \sum_{k=1}^{a} \xi_k = -M \sum_{k=1}^{2p} (k-1) = -Mp(2p-1) \\ & \sum_{k=1}^{a} \xi_k^2 = M \sum_{k=1}^{2p} (k-1)^2 = Mp(2p-1)(4p-1)/3 \\ & \sum_{j=1}^{b} \eta_j = 0, \quad \sum_{j=1}^{b} \eta_j^2 = 0. \end{aligned}$$

Using these sums in (96) and (97) and simplifying, we obtain

$$\nu = 2Mp(2p-1), \quad \rho = 1 - \frac{2p+1}{3K}. \tag{66}$$

Similarly, it follows that

$$\sum_{k=1}^{a} \frac{B_{r+1}(\beta_k + \xi_k)}{(\rho x_k)^r} = M \sum_{l=1}^{2p} \frac{B_{r+1}((1-\rho)K + 1 - l)}{(\rho K)^r}, \tag{67}$$

$$\sum_{j=1}^{b} \frac{B_{r+1}(\epsilon_j + \eta_j)}{(\rho y_j)^r} = 2Mp \frac{B_{r+1}((1-\rho)K)}{(\rho K)^r}. \tag{68}$$

Therefore, we have

$$\omega_r = \frac{(-1)^{r+1} M}{r(r+1)(\rho K)^r} \left\{ \left(\sum_{l=1}^{2p} B_{r+1}((1-\rho)K + 1 - l) \right) - 2p B_{r+1}((1-\rho)K) \right\}. \tag{69}$$

Thus we have Theorem 3 applicable to problem (59).

Theorem 3. The GLRT for (59) is given by

$$2\rho \, \ln(\mathcal{L}^{(P)}) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \tau_3$$

where ρ and $\ln(\mathcal{L}^{(P)})$ are given by (66) and (70), respectively,

$$\ln(\mathcal{L}^{(P)}) = K \sum_{k=1}^{M} \Big(- \ln(|\hat{\mathbf{S}}_y(\tilde{f}_k)|) + \sum_{q=1}^{p} \big[\ln[\hat{\mathbf{S}}_y(\tilde{f}_k)]_{qq} + \ln[\hat{\mathbf{S}}_y(\tilde{f}_k)]_{(q+p)(q+p)} \big] \Big). \quad (70)$$

The threshold τ_3 is picked to achieve a pre-specified $P_{fa} = 1 - P\{2\rho \, \ln(\mathcal{L}^{(P)}) \le \tau_3 \,|\, \mathcal{H}_0\}$ where $P\{2\rho \, \ln(\mathcal{L}^{(P)}) \le \tau_3 \,|\, \mathcal{H}_0\}$ is given by (43) with $\mathcal{L}$ replaced by $\mathcal{L}^{(P)}$ and the various parameters needed therein specified in (66)-(69) •

B. Proper Signals in Proper Noise

Here we consider the case where both signal and noise are known to be proper, hence, $\{\mathbf{x}(t)\}$ is proper under both hypotheses. If $\{\mathbf{x}(t)\}$ is proper, its C-PSD vanishes, and therefore, $\mathbf{S}_y^{(\ell)}(f)$ satisfies (57) (and is denoted by $\mathbf{S}_y^{(\ell P)}(f)$, as in Sec. VI-A), and (8) is modified as

$$\mathbf{S}_y(f) = \begin{bmatrix} \mathbf{S}_x(f) & \mathbf{0} \\ \mathbf{0} & \mathbf{S}_x^*(-f) \end{bmatrix}. \quad (71)$$

Testing for the presence of a proper common signal in spatially independent proper noise is then reformulated as the problem

$$\begin{aligned} \mathcal{H}_0 : \quad & [\mathbf{S}_y(f_n)]_{\ell m} = 0 \; \forall l, m, \text{ except for the entries in} \\ & \mathbf{S}_y^{(qP)}(f_n), \; q = 1, 2, \cdots, p, \; \forall n \in [1, (N/2) - 1] \\ \mathcal{H}_1 : \quad & \mathbf{S}_y(f_n) \succ 0 \text{ is block-diagonal as in (71)} \\ & \forall n \in [1, (N/2) - 1]. \end{aligned} \quad (72)$$

Mimicking the developments in Sec. III leading to (34), for problem (72), we obtain the GLRT

$$\mathcal{L}^{(PP)} := \frac{\sup_{\mathbf{S}_y(\tilde{f}_k) \text{ under } \mathcal{H}_1} f_{\mathbf{D}|\mathcal{H}_1}(\mathbf{D}|\mathcal{H}_1)}{\sup_{\mathbf{S}_y(\tilde{f}_k) \text{ under } \mathcal{H}_0} f_{\mathbf{D}|\mathcal{H}_0}(\mathbf{D}|\mathcal{H}_0)} \quad (73)$$

$$= \prod_{k=1}^{M} \frac{\prod_{q=1}^{p} \Big([\hat{\mathbf{S}}_y(\tilde{f}_k)]_{qq} \times [\hat{\mathbf{S}}_y(\tilde{f}_k)]_{(q+p)(q+p)} \Big)^K}{|[\hat{\mathbf{S}}_y(\tilde{f}_k)]_{1:p,1:p}|^K \times |[\hat{\mathbf{S}}_y(\tilde{f}_k)]_{p+1:2p,p+1:2p}|^K} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \tau_4 \,. \quad (74)$$

As in Sec. VI-A, pick $\mathbf{A}_k^{(q)}$ of Sec. III-A to be diagonal. Then $\mathcal{L}^{(PP)}$ is invariant under the group action $\mathcal{S} \rightarrow \mathcal{A}\mathcal{S}\mathcal{A}^H$, where $\mathcal{S}$ and $\mathcal{A}$ are as in Sec. III-A with $\mathbf{A}_k^{(q)}$ a non-singular, positive-definite diagonal matrix. In particular, if we pick $\mathbf{A}_k^{(q)} = \sqrt{K}\text{diag}\{([\mathbf{S}_y(\tilde{f}_k)]_{qq})^{-1/2}, \; ([\mathbf{S}_y(\tilde{f}_k)]_{(q+p)(q+p)})^{-1/2}\}$, then $\mathcal{L}^{(PP)}$ is invariant, and transformation of $\hat{\mathbf{S}}_y(\tilde{f}_k)$ to $\tilde{\mathbf{S}}_y(\tilde{f}_k)$ leads to $\tilde{\mathbf{S}}_y^{(q)}(\tilde{f}_k) \sim W_C\left(2, K, \mathbf{I}\right)$ and $\tilde{\mathbf{S}}_y(\tilde{f}_k) \sim W_C\left(2p, K, \mathbf{I}\right)$ under $\mathcal{H}_0$.

The counterpart to Lemma 3 now is

$$E\left\{ \frac{1}{\left(\mathcal{L}^{(PP)}\right)^h} \,|\, \mathcal{H}_0 \right\} = \frac{\prod_{k=1}^{p} \Gamma^{2M}(K(1+h) - k + 1)}{\Gamma^{2Mp}(K(1+h))} \times \frac{\Gamma^{2Mp}(K)}{\prod_{\ell=1}^{p} \Gamma^{2M}(K - \ell + 1)} \,. \quad (75)$$

Comparing (95) with (75), we find the correspondence

$$\begin{aligned} & a = 2Mp, \; b = 2Mp, \; x_k = K, \\ & \xi_k = -[(k-1) \bmod (p)] \text{ for } k = 1, 2, \cdots, a, \\ & y_j = K, \; \eta_j = 0 \text{ for } j = 1, 2, \cdots, b, \\ & C = \frac{\Gamma^{2Mp}(K)}{\prod_{\ell=1}^{p} \Gamma^{2M}(K - \ell + 1)} \,. \end{aligned} \quad (76)$$

Comparing (75) and (95), we further have

$$\beta_k = (1-\rho)K \; \forall k, \quad \epsilon_j = (1-\rho)K \; \forall j. \quad (77)$$

Using the above values in Lemma 3, we have

$$\begin{aligned} & \sum_{k=1}^{a} \xi_k = -2M \sum_{k=1}^{p} (k-1) = -Mp(p-1) \\ & \sum_{k=1}^{a} \xi_k^2 = 2M \sum_{k=1}^{p} (k-1)^2 = Mp(p-1)(2p-1)/3 \\ & \sum_{j=1}^{b} \eta_j = 0, \quad \sum_{j=1}^{b} \eta_j^2 = 0. \end{aligned}$$

Using these sums in (96) and (97) and simplifying, we obtain

$$\nu = 2Mp(p-1), \quad \rho = 1 - \frac{p+1}{3K} \,. \quad (78)$$

Similarly, it follows that

$$\sum_{k=1}^{a} \frac{B_{r+1}(\beta_k + \xi_k)}{(\rho x_k)^r} = 2M \sum_{l=1}^{p} \frac{B_{r+1}((1-\rho)K + 1 - l)}{(\rho K)^r}, \quad (79)$$

$$\sum_{j=1}^{b} \frac{B_{r+1}(\epsilon_j + \eta_j)}{(\rho y_j)^r} = 2Mp \frac{B_{r+1}((1-\rho)K)}{(\rho K)^r}. \quad (80)$$

Therefore, we have

$$\omega_r = \frac{2(-1)^{r+1} M}{r(r+1)(\rho K)^r} \Big\{ \Big(\sum_{l=1}^{p} B_{r+1}((1-\rho)K + 1 - l) \Big) - pB_{r+1}((1-\rho)K) \Big\}. \quad (81)$$

Thus we have Theorem 4 applicable to problem (72).

Theorem 4. The GLRT for (72) is given by

$$2\rho \, \ln(\mathcal{L}^{(PP)}) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \tau_5$$

where ρ and $\ln(\mathcal{L}^{(PP)})$ are given by (78) and (82), respectively,

$$\ln(\mathcal{L}^{(PP)}) = K \sum_{k=1}^{M} \Big(- \ln(|[\hat{\mathbf{S}}_y(\tilde{f}_k)]_{1:p,1:p}|) - \ln(|[\hat{\mathbf{S}}_y(\tilde{f}_k)]_{p+1:2p,p+1:2p}|) + \sum_{q=1}^{p} \big[\ln[\hat{\mathbf{S}}_y(\tilde{f}_k)]_{qq} + \ln[\hat{\mathbf{S}}_y(\tilde{f}_k)]_{(q+p)(q+p)} \big] \Big). \tag{82}$$

The threshold τ_5 is picked to achieve a pre-specified $P_{fa} = 1 - P\{2\rho \ln(\mathcal{L}^{(PP)}) \leq \tau_5 \,|\, \mathcal{H}_0\}$ where $P\{2\rho \ln(\mathcal{L}^{(PP)}) \leq \tau_5 \,|\, \mathcal{H}_0\}$ is given by (43) with $\mathcal{L}$ replaced by $\mathcal{L}^{(PP)}$ and the various parameters needed therein specified in (78)-(81) •

VII. Simulation Examples

We now present some computer simulation examples to illustrate the proposed approach and the analytical results.

A. Signal and Noise Models

We generate stationary $\mathbf{x}(t) \in \mathbb{C}^p$ as a noisy improper signal given by

$$\mathbf{x}(t) = \mathbf{s}(t) + \mathbf{n}(t)\,, \tag{83}$$

where $\{\mathbf{n}(t)\}$ is spatially uncorrelated, colored, proper or improper complex Gaussian noise, and $\{\mathbf{s}(t)\}$ is the signal sequence. If proper, the noise sequence $\{\mathbf{n}(t)\} \in \mathbb{C}^p$ is generated as

$$\mathbf{n}(t) = \mathbf{n}_c(t) + \mathbf{n}_w(t)\,, \tag{84}$$

where $\mathbf{n}_w(t) \sim \mathcal{N}_c(\mathbf{0}, \sigma_w^2 \mathbf{I})$ is i.i.d., and $\mathbf{n}_c(t)$ is generated as follows. The various components of $\mathbf{n}_c(t)$ are i.i.d., and each component is generated by filtering an i.i.d. scalar sequence, distributed as $\mathcal{N}_c(0, \sigma_c^2)$, through a linear filter with impulse response $\{0.3, 1, 0.3\}$, normalized to unit norm. Therefore, $\mathbb{E}\{\|\mathbf{n}_c(t)\|^2\} = p\sigma_c^2$, leading to $\mathbb{E}\{\|\mathbf{n}(t)\|^2\} = p(\sigma_c^2 + \sigma_w^2) = p\sigma_n^2$. We pick $\sigma_w^2 = 0.2\sigma_n^2$ for a given value of σ_n^2. To generate improper noise sequence $\{\mathbf{n}(t)\}$, we still use (84) except that $\mathbf{n}_c(t)$ is generated differently. Let $\tilde{\mathbf{n}}_i(t)$, $i \in [1, p]$, denote p independent zero-mean white proper Gaussian sequences. We generate $[\mathbf{n}_c]_i(t) = a_I h_I(t) \circledast \tilde{\mathbf{n}}_i(t) + a_Q h_Q(t) \circledast \tilde{\mathbf{n}}_i^*(t)$ where $a_I = a_Q^* = (1 + j1)/\sqrt{2}$, $\circledast$ denotes convolution, $h_I(t) = [0.3\ 1\ 0.3]$, and $h_Q(t) = [0.4\ 1\ 0.5]$. Thus noise $\mathbf{n}_c(t)$, hence $\mathbf{n}(t)$, is spatially independent, improper Gaussian.

The signal $\{\mathbf{s}(t)\}$ is a filtered digital communications signal generated by passing an information sequence through a frequency-selective Rayleigh fading channel as follows:

$$\mathbf{s}(t) = \sum_{l=0}^{4} \mathbf{h}(l) d(t-l) \tag{85}$$

where $d(t)$ is a scalar i.i.d. BPSK information sequence, with equally likely binary values $d(t) \in \{-1, +1\}$, filtered through a random time-invariant, frequency-selective Rayleigh fading channel $\mathbf{h}(l) \in \mathbb{C}^p$ with 5 taps, equal power delay profile, mutually independent components, which are identically distributed zero-mean proper complex Gaussian random variables. For different ls, $\mathbf{h}\,(l)$s are mutually independent and identically distributed as $\mathbf{h}(l) \sim \mathcal{N}_c(\mathbf{0}, \sigma_h^2 \mathbf{I})$. Since $d(t)$ is a BPSK signal, it is improper. Since any linear filtering preserves propriety/impropriety property [1], we have an improper $\{\mathbf{s}(t)\}$ for BPSK $d(t)$.

We pick σ_h^2 to achieve the desired average signal-to-noise ratio (SNR) across p components (receive antennas in the communications context), defined as ratio of the sum of signal powers at the p antennas to the sum of noise powers:

$$\text{SNR} = E\{\|\mathbf{s}(t)\|^2\}/E\{\|\mathbf{n}(t)\|^2\}\,.$$

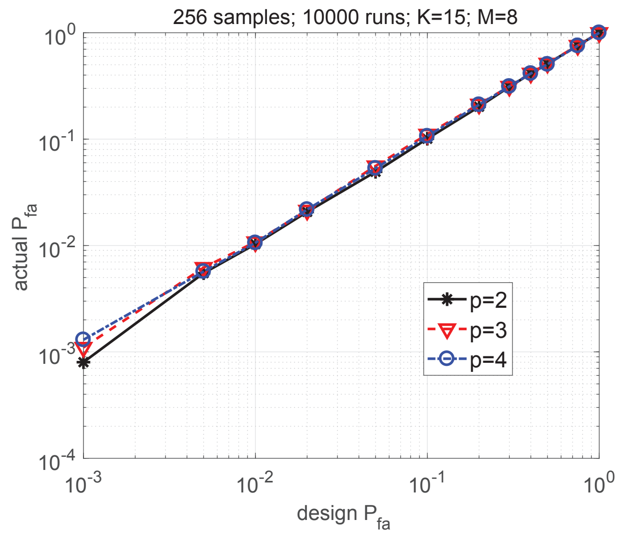

Fig. 1: Actual P_{fa} vs. design P_{fa} using the GLRT detector of Theorem 1, designed for improper signals in improper noise. $N = 256$, $K = 15$, $M = 8$.

B. Threshold Calculation

First we investigate the efficacy of Theorem 1 in computing the GLRT threshold for a given probability of false alarm P_{fa}. In this case, we set $\mathbf{x}(t) = \mathbf{n}(t)$, just noise, given by (84), where $\mathbf{n}(t)$ is improper. We consider up to four antennas ($p =$ 2, 3 or 4). We used unweighted smoothing in the frequency-domain to estimate PSD (see (26)). To estimate the PSD of augmented $\mathbf{y}(t)$, for $N = 256$, we choose $m_t = 7$ leading to $K = 15$ and $M = 8$. In Fig. 1, we compare the actual P_{fa} and design P_{fa} based on 10,000 runs for $N = 256$. The threshold values were calculated based on (43). It is seen that Theorem 1 is effective in accurately calculating the threshold value for $N = 256$, for all values of p used.

C. Detection Performance

Next, we show the receiver operating characteristic (ROC) curves (probability of detection P_d versus P_{fa}) to illustrate the detection performance, based on 10,000 runs. In this case, we set $\mathbf{x}(t) = \mathbf{s}(t) + \mathbf{n}(t)$, given by (83), where $\mathbf{n}(t)$ and $\mathbf{s}(t)$ are as in (84) and (85), respectively, with improper $\mathbf{n}_c(t)$, $d(t)$ is BPSK, and $p = 2$ or 3. Thus, both signal and noise are improper. The empirical probability of detection P_d versus empirical false-alarm rate P_{fa} results, based on 10,000 runs, are shown in Fig. 2 for three different SNR values (-10, -7.5

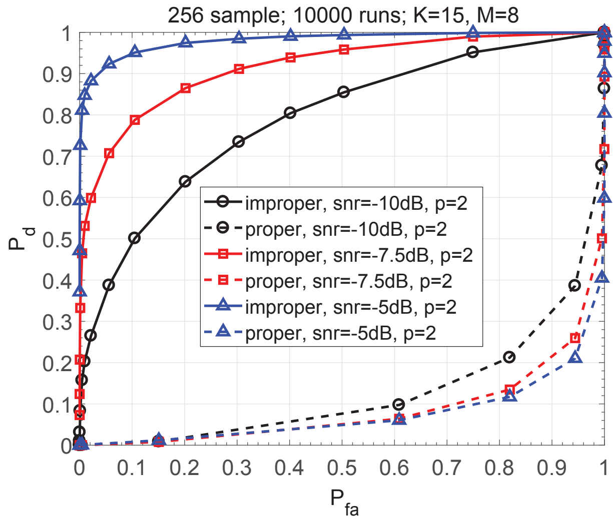

Fig. 2: ROC curves (probability of detection P_d versus P_{fa}) for improper signal in improper noise, $N = 256$, $K = 15$, $M = 8$, SNR= -10, -7.5 or -5dB. The curves labeled "improper" are based on the GLRT statistic $\ln(\mathcal{L})$ of (42) designed for improper signals in improper noise. The curves labeled "proper" are based on the GLRT statistic $\ln(\mathcal{L}^{(P)})$ of (70) (Theorem 3), designed for improper signals in proper noise.

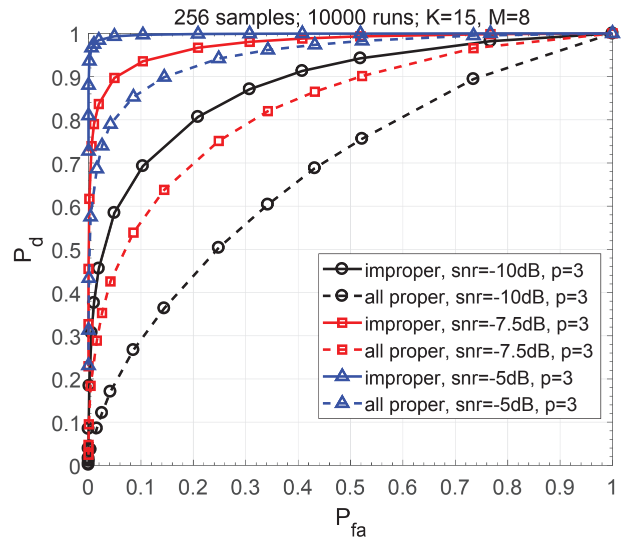

Fig. 3: ROC curves for improper signal in improper noise, $N = 256$, $K = 15$, $M = 8$, SNR= -10, -7.5 or -5dB. The curves labeled "improper" are based on the GLRT statistic $\ln(\mathcal{L})$ of (42) designed for improper signals in improper noise. The curves labeled "all proper" are based on the GLRT statistic $\ln(\mathcal{L}^{(PP)})$ of (82) (Theorem 4), designed for proper signals in proper noise.

or -5dB), $p = 2$, N=256, K=15 and M=8, using the GLRT statistic $\ln(\mathcal{L})$ of (42). These curves are labeled "improper" in Fig. 2. It is seen that the performance improves with increasing SNR, and our approach is able to detect improper random signals quite well at low SNRs.

We also show the ROC curves for the detector of Theorem 3 with GLRT statistic $\ln(\mathcal{L}^{(P)})$ of (70), which is designed for improper signals in proper noise. These curves are labeled "proper" in Fig. 2. There is a model mismatch as the noise is treated as proper by the detector whereas, in reality, it is improper. This detector applied to this problem treats improper noise as signal; under $\mathcal{H}_0$, $\ln(\mathcal{L}^{(P)})$ is not invariant to changes in impropriety of noise. As the ROC curves labeled "proper" show, the detector of Theorem 3 fails for this problem.

In Fig. 3, we show the ROC curves for the detector of Theorem 4 with GLRT statistic $\ln(\mathcal{L}^{(PP)})$ of (82), which is designed for proper signals in proper noise. These curves are labeled "all proper" in Fig. 3. The signal and noise, however, are as for Fig. 2, i.e., improper signal in improper noise, except that now $p = 3$. The statistic $\ln(\mathcal{L}^{(PP)})$ uses only the PSD of the measurements, ignoring the C-PSD. For comparison, we also show the ROC curves for the statistic $\ln(\mathcal{L})$ of Theorem 1, labeled "improper." It is seen that, unlike the GLRT statistic $\ln(\mathcal{L}^{(P)})$ shown in Fig. 2, the statistic $\ln(\mathcal{L}^{(PP)})$ works well, but not as well as $\ln(\mathcal{L})$, since the latter also exploits the C-PSD of the improper data. One expects the difference between the performances of $\ln(\mathcal{L})$ and $\ln(\mathcal{L}^{(PP)})$ to lessen with increasing N.

In Fig. 4, we investigate via simulations the effect of varying K $(= 2m_t + 1)$ on the performance of $\ln(\mathcal{L})$ for the example of Fig. 3. For a given N, as K increases, M must decrease. It is seen that as K increases, the performance improves (higher P_d for a given P_{fa}) for the presented example.

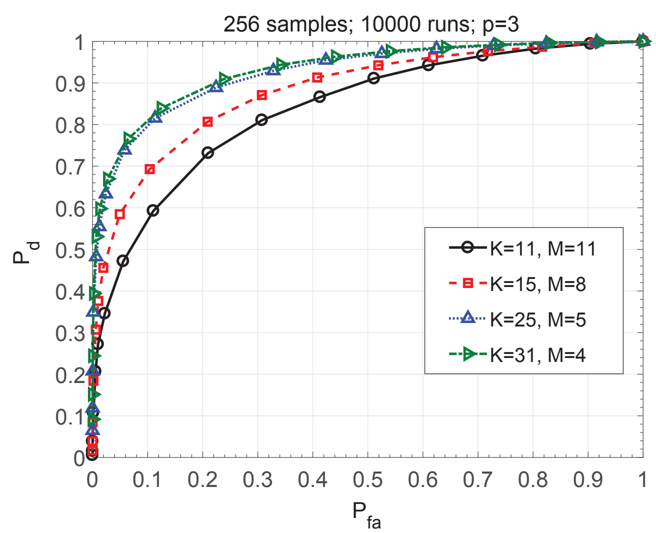

Fig. 4: ROC curves for improper signal in improper noise, with varying (K, M)=(11,11), (15,8), (25,5), (31,4): $p = 3$, $N = 256$, SNR= -10dB, 10,000 runs.

D. Corroboration of Performance Analysis

Here we compare the theoretical performance based on (56) with the simulations-based performance for the proposed detector $\ln(\mathcal{L})$. We use the improper signal in improper noise model of Fig. 3 with $p = 3$. Fig. 5 shows the probability of detection versus SNR results based on 10,000 runs under the set-up for Fig. 3, except that the channels after having been randomly generated, were then fixed for all runs – the results of Sec. V apply to processes with fixed PSDs. Fig. 5 shows the results based on 10,000 runs for N =128, 256 and 512, with corresponding values (K, M) = (11,5), (15,8) and (25,10), respectively, using (56) for theoretical results. It is seen from Fig. 5 that the agreement between the theoretical and simulation-based results is very good, although not "perfect," and it improves with increasing N.

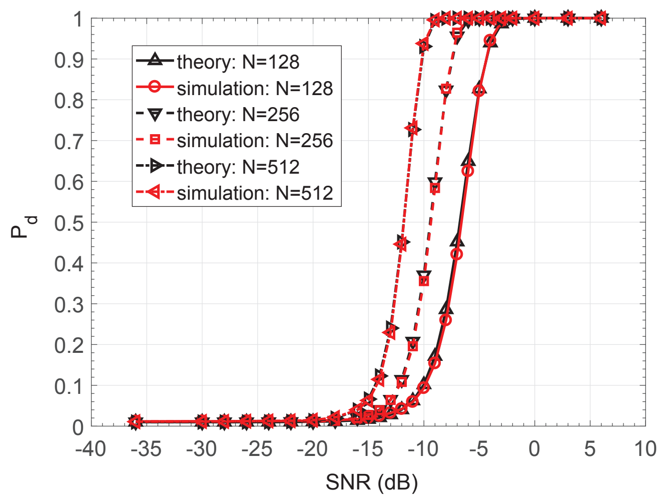

Fig. 5: Theoretical and empirical detection probability P_d vs. SNR, based on 10,000 runs: $p = 3$, $P_{fa} = 0.01$, $(N, K, M) = (128, 11, 5)$, $(256, 15, 8)$, $(512, 25, 10)$.

VIII. Conclusions

We investigated the problem of detection of a common improper signal among two or more sensors, in the presence of spatially independent improper noise. A source of improper noise is IQ imbalance during down-conversion of bandpass noise to baseband. Our proposed approach is based on the GLRT framework, and it exploits the asymptotic distribution of the DFT of an augmented measurement sequence. An analytical method for calculation of the test threshold was provided. We also analyzed the performance of the GLRT by deriving an approximate asymptotic distribution of the GLRT statistic under a sequence of local alternative hypotheses. Two special cases of detection of improper signals in proper noise, and of detection of proper signals in proper noise, were also investigated using the same theoretical approach. The signal or the noise need not be Gaussian. Simulation examples were presented which show the efficacy of analytical threshold calculation, and also show improved detection performance compared to the case where both signal and noise are treated as proper (even if they are improper).

Appendix A
Proof of Theorem 1

First we need to introduce some notation and recall/develop several auxiliary results.

As noted in [18, Appendix B], similar to the real case in [22, Def. 2.1.10], the complex multivariate gamma function $\Gamma_m(\alpha)$ is defined to be [25, Def. 2.1], [26, (83)]

$$\Gamma_m(\alpha) = \int_{\mathbf{X}=\mathbf{X}^H \succ 0} |\mathbf{X}|^{\alpha-m} \mathrm{etr}(-\mathbf{X})\, d\mathbf{X}, \;\; \mathrm{Re}(\alpha) > m-1 \tag{86}$$

where $\mathbf{X} \in \mathbb{C}^{m\times m}$. It turns out that [25, (1.2)], [26, (83)]

$$\Gamma_m(\alpha) = \pi^{m(m-1)/2} \prod_{i=1}^{m} \Gamma(\alpha - i + 1), \quad \mathrm{Re}(\alpha) > m-1\,. \tag{87}$$

Note that (87) is the same as the expression (36), except that here α can be a complex variable whereas in (36), $\alpha = K$, an integer.

The complex counterpart to [22, Theorem 2.1.11] is proved in [18, Lemma 3] and is stated below in Lemma 2.

Lemma 2. Let $\mathbf{X}$, $\mathbf{\Sigma} \in \mathbb{C}^{m\times m}$, $\mathbf{X} = \mathbf{X}^H \succ 0$ and $\mathbf{\Sigma} = \mathbf{\Sigma}^H \succ 0$. Then

$$\int_{\mathbf{X}=\mathbf{X}^H \succ 0} |\mathbf{X}|^{\alpha-m} \mathrm{etr}(-\mathbf{\Sigma}^{-1}\mathbf{X})\, d\mathbf{X} = \Gamma_m(\alpha)|\mathbf{\Sigma}|^{\alpha} \tag{88}$$

for $\mathrm{Re}(\alpha) > m-1$ •

We use Lemma 2 to prove the following result, needed to invoke [24].

Lemma 3. Under $\mathcal{H}_0$, for any h with $\mathrm{Re}(h) \geq 0$,

$$E\left\{\frac{1}{\mathcal{L}^h} \,|\, \mathcal{H}_0\right\} = \frac{\prod_{k=1}^{2p} \Gamma^M(K(1+h)-k+1)}{\prod_{j=1}^{2} \Gamma^{Mp}(K(1+h)-j+1)} \times \frac{\prod_{\ell=1}^{2} \Gamma^{Mp}(K-\ell+1)}{\prod_{\ell=1}^{2p} \Gamma^{M}(K-\ell+1)}\,. \tag{89}$$

Proof: First we use the transformation specified in Sec. III-A to obtain $\tilde{\mathbf{S}}_y(\tilde{f}_k) \sim W_C\left(2p, K, \mathbf{I}\right)$ under $\mathcal{H}_0$. This transformation leaves $\mathcal{L}$ unchanged, hence, we can compute $E\{\frac{1}{\mathcal{L}^h} \,|\, \mathcal{H}_0\}$ under this distribution. To simplify notation, we will use $\mathbf{Y}_k$ to denote $\tilde{\mathbf{S}}_y(\tilde{f}_k)$ in the rest of the proof. We have

$$\begin{aligned} E\left\{1/\mathcal{L}_k^h \,|\, \mathcal{H}_0\right\} &= \int \frac{|\mathbf{Y}_k|^{Kh+K-2p}}{\prod_{q=1}^{p} |\mathbf{Y}_k^{(q)}|^{Kh}} \frac{\mathrm{etr}\{-\mathbf{Y}_k\}}{\Gamma_{2p}(K)}\, d\mathbf{Y}_k \\ &= \frac{\Gamma_{2p}(K+Kh)}{\Gamma_{2p}(K)} E\left\{\prod_{q=1}^{p} \tilde{\mathbf{Y}}_k^{(q)}|^{-Kh}\right\} \\ &= \frac{\Gamma_{2p}(K+Kh)}{\Gamma_{2p}(K)} \prod_{q=1}^{p} E\left\{|\tilde{\mathbf{Y}}_k^{(q)}|^{-Kh}\right\}, \end{aligned} \tag{90}$$

where $\tilde{\mathbf{Y}}_k^{(q)}$ is defined similar to (12) and (29), $\tilde{\mathbf{Y}}_k \sim W_C\left(2p, K(1+h), \mathbf{I}\right)$, hence, $\tilde{\mathbf{Y}}_k^{(q)}$s are independent for $q \in [1,p]$ and $\tilde{\mathbf{Y}}_k^{(q)} \sim W_C\left(2, K(1+h), \mathbf{I}\right)$. We have

$$\begin{aligned} E\{|\tilde{\mathbf{Y}}_k^{(q)}|^{-Kh}\} &= \int |\mathbf{X}|^{-Kh} \frac{|\mathbf{X}|^{K(1+h)-2}\mathrm{etr}\left\{-X\right\}}{\Gamma_2(K(1+h))}\, d\mathbf{X} \\ &= \frac{1}{\Gamma_2(K(1+h))} \int |\mathbf{X}|^{K-2}\mathrm{etr}\left\{-X\right\} d\mathbf{X} \\ &= \frac{\Gamma_2(K)}{\Gamma_2(K(1+h))} \end{aligned} \tag{91}$$

where, in the last step above, we used Lemma 2 with $\alpha = K$, $m = 2$ and $|\mathbf{\Sigma}| = |\mathbf{I}| = 1$. Thus, we have

$$E\left\{1/\mathcal{L}_k^h \,|\, \mathcal{H}_0\right\} = \frac{\Gamma_{2p}(K+Kh)\Gamma_2^p(K)}{\Gamma_{2p}(K)\Gamma_2^p(K(1+h))}\,. \tag{92}$$

Now we use (87) with $m = 2p$ or 2, and i expressed as k or j in Γ_{2p}s and Γ_2s in (92), to obtain

$$E\left\{\frac{1}{\mathcal{L}_k^h} \mid \mathcal{H}_0\right\} = \frac{\prod_{\ell=1}^{2} \Gamma^p(K-\ell+1)}{\prod_{\ell=1}^{2p} \Gamma(K-\ell+1)} \times \frac{\prod_{k=1}^{2p} \Gamma(K(1+h)-k+1)}{\prod_{j=1}^{2} \Gamma^p(K(1+h)-j+1)} . \quad (93)$$

By (34) and independence of $\mathcal{L}_k$s for $1 \le k \le M$, we have

$$E\left\{1/\mathcal{L}^h \mid \mathcal{H}_0\right\} = \prod_{k=1}^{M} E\left\{1/\mathcal{L}_k^h \mid \mathcal{H}_0\right\} . \quad (94)$$

Using (93) and (94), we get the desired result. □

We also need Lemma 4 which follows from [22, Sec. 8.2.4], [23, Sec. 8.5.1], but is proved in [18, Lemma 9] in the stated form. In order to exploit Lemma 4 (stated next), we need to establish that $0 \le \mathcal{L}^{-1} \le 1$. Since $\hat{\mathbf{S}}_y(\tilde{f}_k) \succ 0$ w.p.1 for $K \ge 2p$ (hence $\hat{\mathbf{S}}_y^{(q)}(\tilde{f}_k) \succ 0$ w.p.1 $\forall q$), $\mathcal{L}^{-1} \ge 0$ follows immediately. By Fischer's inequality [27, p. 477], we have $|\hat{\mathbf{S}}_y(\tilde{f}_k)| \le \prod_{q=1}^{p} |\hat{\mathbf{S}}_y^{(q)}(\tilde{f}_k)|$ which implies $\mathcal{L}^{-1} \le 1$.

Lemma 4. Consider a random variable W ($0 \le W \le 1$) with hth moment

$$E\{W^h\} = C \left(\frac{\prod_{j=1}^{b} y_j^{y_j}}{\prod_{k=1}^{a} x_k^{x_k}} \right)^h \frac{\prod_{k=1}^{a} \Gamma(x_k(1+h)+\xi_k)}{\prod_{j=1}^{b} \Gamma(y_j(1+h)+\eta_j)} \quad (95)$$

where a and b are integers, C is a constant such that $E\{W^0\} = 1$ and $\sum_{k=1}^{a} x_k = \sum_{j=1}^{b} y_j$. Define

$$\nu = -2\Big[\sum_{k=1}^{a} \xi_k - \sum_{j=1}^{b} \eta_j - \frac{1}{2}(a-b)\Big] , \quad (96)$$

$$\rho = 1 - \frac{1}{\nu}\Big[\sum_{k=1}^{a} x_k^{-1}\big(\xi_k^2 - \xi_k + \frac{1}{6}\big) - \sum_{j=1}^{b} y_j^{-1}\big(\eta_j^2 - \eta_j + \frac{1}{6}\big)\Big] , \quad (97)$$

$$\beta_k = (1-\rho)x_k, \quad \epsilon_j = (1-\rho)y_j , \quad (98)$$

and

$$\omega_r = \frac{(-1)^{r+1}}{r(r+1)} \Big\{ \sum_{k=1}^{a} \frac{B_{r+1}(\beta_k+\xi_k)}{(\rho x_k)^r} - \sum_{j=1}^{b} \frac{B_{r+1}(\epsilon_j+\eta_j)}{(\rho y_j)^r} \Big\}. \quad (99)$$

Then

$$\begin{aligned} P\{-2\rho \ln(W) \le z\} &= P\{\chi_\nu^2 \le z\} + \omega_2 \left[P\{\chi_{\nu+4}^2 \le z\} - P\{\chi_\nu^2 \le z\}\right] \\ &+ \omega_3 \left[P\{\chi_{\nu+6}^2 \le z\} - P\{\chi_\nu^2 \le z\}\right] \\ &+ \Big\{\omega_4 \left[P\{\chi_{\nu+8}^2 \le z\} - P\{\chi_\nu^2 \le z\}\right] + \frac{1}{2}\omega_2^2 \\ &\times \left[P\{\chi_{\nu+8}^2 \le z\} - 2P\{\chi_{\nu+4}^2 \le z\} + P\{\chi_\nu^2 \le z\}\right] \Big\} \\ &+ \sum_{k=1}^{a} \mathcal{O}(x_k^{-5}) + \sum_{j=1}^{b} \mathcal{O}(y_j^{-5}) , \end{aligned} \quad (100)$$

where ρ is has been chosen to yield $\omega_1 = 0$. •

Comparing (95) with (89), we find the correspondence

$$\begin{aligned} & a = 2Mp,\ b = 2Mp,\ x_k = K, \\ & \xi_k = -[(k-1) \bmod (2p)] \text{ for } k = 1, 2, \cdots, a, \\ & y_j = K,\ \eta_j = -[(j-1) \bmod (2)] \text{ for } j = 1, 2, \cdots, b, \\ & C = \frac{\prod_{\ell=1}^{2} \Gamma^{Mp}(K-\ell+1)}{\prod_{\ell=1}^{2p} \Gamma^M(K-\ell+1)} . \end{aligned} \quad (101)$$

Comparing Lemmas 3 and 4, we further have

$$\beta_k = (1-\rho)K \ \forall k, \quad \epsilon_j = (1-\rho)K \ \forall j. \quad (102)$$

Furthermore, when $h = 0$, the right-side of (89) equals 1, i.e., C is such that $E\{W^0\} = E\{1/\mathcal{L}^0 \mid \mathcal{H}_0\} = 1$, as required by Lemma 4. Thus, Lemma 4 is applicable with $W = 1/\mathcal{L}$ and parameters specified in (101).

Using the above values in Lemma 4, we have

$$\begin{aligned} \sum_{k=1}^{a} \xi_k &= -M \sum_{k=1}^{2p} (k-1) = -Mp(2p-1) \\ \sum_{k=1}^{a} \xi_k^2 &= M \sum_{k=1}^{2p} (k-1)^2 = Mp(2p-1)(4p-1)/3 \\ \sum_{j=1}^{b} \eta_j &= -Mp \sum_{j=1}^{2} (j-1) = -Mp \\ \sum_{j=1}^{b} \eta_j^2 &= Mp \sum_{j=1}^{2} (j-1)^2 = Mp. \end{aligned}$$

Using these sums in (96) and (97) and simplifying, we obtain (41) and (44). Similarly, it follows that

$$\sum_{k=1}^{a} \frac{B_{r+1}(\beta_k+\xi_k)}{(\rho x_k)^r} = M \sum_{l=1}^{2p} \frac{B_{r+1}((1-\rho)K+1-l)}{(\rho K)^r}, \quad (103)$$

$$\sum_{j=1}^{b} \frac{B_{r+1}(\epsilon_j+\eta_j)}{(\rho y_j)^r} = Mp \sum_{l=1}^{2} \frac{B_{r+1}((1-\rho)K+1-l)}{(\rho K)^r}. \quad (104)$$

Substitution of (103) and (104) in (99) yields (45). We also have $\sum_{k=1}^{a} \mathcal{O}(x_k^{-5}) = \mathcal{O}(M/K^5)$ and $\sum_{j=1}^{b} \mathcal{O}(y_j^{-5}) = \mathcal{O}(M/K^5)$. It then follows from Lemma 4 that (43) holds true. This completes the proof of Theorem 1.

Appendix B
Proof Lemma 1

First we recall some useful results:

Lemma 5 [19, Lemma 4]. If $X \sim \chi_n^2(\lambda)$, then $E\{X\} = n+\lambda$. If $Z \sim \chi_n^2 = \chi_n^2(0)$, then $E\{\ln Z\} = \psi(\frac{n}{2}) + \ln 2$ where $\psi(x) = d \ln \Gamma(x)/dx$ is the digamma function and $\psi(x) = \ln(x) - \frac{1}{2x} - \frac{1}{12x^2} + \cdots$.

Lemma 6 [18, Lemma 5]. If $\mathbf{X} \sim W_C(p, K, \mathbf{\Sigma})$, $K \ge p$, then

$$E\{\ln(|\mathbf{X}|)\} = \ln(|\mathbf{\Sigma}|) + \sum_{\ell=0}^{p-1} \psi(K-\ell). \quad (105)$$

Now we turn to the proof of Lemma 1. As discussed in Sec. III, $\hat{\mathbf{S}}_y(\tilde{f}_k) \overset{a}{\sim} W_C(2p, K, K^{-1}\mathbf{S}_y(\tilde{f}_k))$. Therefore, by Lemma 6,

$$E\{\ln|\hat{\mathbf{S}}_y(\tilde{f}_k)|\} = \ln\big(\frac{|\mathbf{S}_y(\tilde{f}_k)|}{K^{2p}}\big) + \sum_{l=0}^{2p-1}\psi(K-l). \quad (106)$$

Since $\hat{\mathbf{S}}_y(\tilde{f}_k) \overset{a}{\sim} W_C(2p, K, K^{-1}\mathbf{S}_y(\tilde{f}_k))$, we have its 2×2 submatrix $\hat{\mathbf{S}}_y^{(q)}(\tilde{f}_k) \overset{a}{\sim} W_C(2, K, K^{-1}\mathbf{S}_y^{(q)}(\tilde{f}_k))$, $q = 1, 2, \cdots, p$. Therefore, by Lemma 6,

$$E\{\ln|\hat{\mathbf{S}}_y^{(q)}(\tilde{f}_k)|\} = \ln\big(\frac{|\mathbf{S}_y^{(q)}(\tilde{f}_k)|}{K^2}\big) + \psi(K) + \psi(K-1). \quad (107)$$

Using (42) and (106)-(107), we have

$$E\{2\ln(\mathcal{L})\,|\,\mathcal{H}_1\} = 2K\sum_{k=1}^{M}\Bigg(\big[\sum_{q=1}^{p}\ln|\mathbf{S}_y^{(q)}(\tilde{f}_k)|\big] - \ln(|\mathbf{S}_y(\tilde{f}_k)|) + d\Bigg) \quad (108)$$

where

$$d := p\,(\psi(K)+\psi(K-1)) - \sum_{l=0}^{2p-1}\psi(K-l). \quad (109)$$

Using Lemma 5 we have

$$\begin{aligned}\psi(K-l) &= \ln(K-l) - \frac{1}{2(K-l)} + \mathcal{O}(K^{-2}) \\ &= \ln(K) + \ln\left(1-\frac{l}{K}\right) - \frac{1}{2K(1-\frac{l}{K})} + \mathcal{O}(K^{-2}) \\ &= \ln K - \frac{l}{K} - \frac{1}{2K} + \mathcal{O}(K^{-2}) \quad (110)\end{aligned}$$

where we have used $\ln(1+x) = x + \mathcal{O}(x^2)$ and $(1-x)^{-1} = 1 + x + \mathcal{O}(x^2)$. By (109) and (110) we have

$$\begin{aligned}d &= -\frac{p}{K} + \frac{1}{K}\sum_{l=0}^{2p-1} l + \mathcal{O}(K^{-2}) \\ &= \frac{2p(p-1)}{K} + \mathcal{O}(K^{-2}). \quad (111)\end{aligned}$$

Substitution of (111) in (108) proves Lemma 1.

References

[1] P.J. Schreier and L.L. Scharf, *Statistical Signal Processing of Complex-Valued Data*, Cambridge, UK: Cambridge Univ. Press, 2010.

[2] F.D. Neeser and J.L. Massey, "Proper complex random processes with applications to information theory," *IEEE Trans. Inf. Theory*, vol. 39, pp. 1293-1302, July 1993.

[3] D. Tandur and M. Moonen, "Efficient compensation of transmitter and receiver IQ imbalance in OFDM systems," *EURASIP J. Adv. Signal Process.*, vol. 62, Article ID 106562, 14 pages, 2010.

[4] T. Adali, P.J. Schreier and L.L. Scharf, "Complex-valued signal processing: The proper way to deal with impropriety," *IEEE Trans. Signal Process.*, vol. 59, pp. 5101-5125, Nov. 2011.

[5] D.P. Mandic and V.S.L. Goh, *Complex Valued Nonlinear Adaptive Filters*. New York: Wiley, 2009.

[6] N. Klausner, M.R. Azimi-Sadjadi and L.L. Scharf, "Detection of spatially correlated time series from a network of sensor arrays," *IEEE Trans. Signal Process.*, vol. 62, no. 6, pp. 1396-1407, March 2014.

[7] E. Axell, G. Leus, E.G. Larsson and H.V. Poor, "Spectrum sensing for cognitive radio: State-of-the-art and recent advances," *IEEE Signal Process. Mag.*, vol. 29, pp. 101-116, May 2012.

[8] A. Vasilopoulos, G. Vitzilaios, G. Theodoratos and Y. Papananos, "A low-power wideband reconfigurable integrated active RC filter with 73 dB SFDR," *IEEE J. Solid-State Circuits* vol. 41, pp. 1997-2008, 2006.

[9] H. Shin and Y. Kim, "A CMOS active RC low-pass filter with simultaneously tunable high and low cutoff frequencies for IEEE 802.22 applications," *IEEE Trans. Circuits Syst. II Exp. Briefs*, vol. 57, pp. 85-89, 2010.

[10] S.K. Sharma, S. Chatzinotas and B. Ottersten, "The effect of noise correlation on fractional sampling based spectrum sensing," in *Proc. 2013 IEEE Intl. Conf. Commun. (ICC)*, Budapest, Hungary, June 2013, pp. 1183-1187.

[11] R. Zhang, T.J. Lim, Y.-C. Liang and Y. Zeng, "Multi-antenna based spectrum sensing for cognitive radios: A GLRT approach," *IEEE Trans. Commun.*, vol. 58, pp. 84-88, Jan. 2010.

[12] D. Ramirez, J. Via, I. Santamaria and L.L. Scharf, "Detection of spatially correlated Gaussian time series," *IEEE Trans. Signal Process.*, vol. 58, pp. 5006-5015, Oct. 2010.

[13] J.K. Tugnait, "Multichannel spectrum sensing via multivariate power spectrum analysis," in *Proc. 12th IEEE Intern. Workshop Signal Proc. Adv. Wireless Commun. (SPAWC 2011)*, San Francisco, CA, June 26-29, 2011, pp. 106-110.

[14] L. Huang, Y-H. Xiao and Q.T. Zhang, "Robust spectrum sensing for noncircular signal in multiantenna cognitive receivers," *IEEE Trans. Signal Process.*, vol. 63, pp. 498-511, Jan. 2015.

[15] J.K. Tugnait, "Multiantenna spectrum sensing for improper signals over frequency selective channels," in *Proc. 2016 IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP 2016)*, Shanghai, China, March 20-25, 2016, pp. 3781-3785.

[16] J.K. Tugnait, "Multisensor detection of improper signals in improper noise," in *Proc. 2017 IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP 2017)*, New Orleans, Louisiana, March 5-9, 2017, pp. 3939-3943.

[17] D.R. Brillinger, *Time Series: Data Analysis and Theory*, Expanded edition. New York: McGraw Hill, 1981.

[18] J.K. Tugnait and S.A. Bhaskar, "On testing for impropriety of multivariate complex-valued random sequences," *IEEE Trans. Signal Process.*, vol. 65, no. 11, pp. 2988-3003, Jun 1, 2017.

[19] J.K. Tugnait, "Comparing multivariate complex random signals: Algorithm, performance analysis and application," *IEEE Trans. Signal Process.*, vol. 64, no. 4, pp. 934-947, Feb. 15, 2016.

[20] G. Casella and R.L. Berger, *Statistical Inference*, 2nd edition. Pacific Grove, CA: Duxbury, 2002.

[21] M.S. Srivastava and C.G. Khatri, *An Introduction to Multivariate Statistics*. New York: Elsevier North Holland, 1979.

[22] R.J. Muirhead, *Aspects of Multivariate Statistical Theory*. New York: John Wiley, 1982.

[23] T.W. Anderson, *An Introduction to Multivariate Statistical Analysis*, Third edition. Hoboken, NJ: John Wiley, Wiley-Interscience, 2003.

[24] G.E.P. Box, "A general distribution theory for a class of likelihood criteria," *Biometrika*, vol. 36, pp. 317-246, Dec. 1949.

[25] D.K. Nagar and A.K. Gupta, "Expectations of functions of complex Wishart matrix," *Acta Applicandae Mathematicae*, vol. 113, pp. 265-288, 2011.

[26] A.T. James, "Distributions of matrix variates and latent roots derived from normal samples," *Ann. Math. Statist.*, vol. 35, pp. 475-501, 1964.

[27] R.A. Horn and C.R. Johnson, *Matrix Analysis*. Cambridge, UK: Cambridge Univ. Press, 1985.

[28] T. Hayakawa, "The likelihood ratio criterion for a composite hypothesis under a local alternative," *Biometrika*, vol. 62, no. 2, pp. 451-460, Aug. 1975.

[29] N. Sugiura, "Asymptotic non-null distributions of the likelihood ratio criteria for covariance matrix under local alternatives," *Annals of Statistics*, vol. 1, No. 4, pp. 718-728, 1973.

[30] L. Huang, C. Qian, Y. Xiao and Q.T. Zhang, "Performance analysis of volume-based spectrum sensing for cognitive radio," *IEEE Trans. Wireless Commun.*, vol. 14, pp. 317-330, Jan. 2015.

[31] L. Huang, Y. Xiao, H.C. So and J. Fang, "Accurate performance analysis of Hadamard ratio test for robust spectrum sensing," *IEEE Trans. Wireless Commun.*, vol. 14, pp. 750-758, Feb. 2015.

[32] Y.-H. Xiao, L. Huang, J. Xie and H.C. So, "Approximate asymptotic distribution of locally most powerful invariant test for independence: Complex case," *IEEE Trans. Inf. Theory*, vol. 63, pp. 1784-1799, March 2018.

[33] Y.-H. Xiao, L. Huang, J.-K. Zhang, J. Xie and H.C. So, "Performance analysis of locally most powerful invariant test for sphericity of Gaussian vectors in coherent MIMO radar," *IEEE Trans. Veh. Tech.*, vol. 67, pp. 5868-5882, July 2018.

Jitendra K. Tugnait (M'79–SM'93–F'94–LF'16) received the B.Sc.(Hons.) degree in electronics and electrical communication engineering from the Punjab Engineering College, Chandigarh, India in 1971, the M.S. and the E.E. degrees from Syracuse University, Syracuse, NY and the Ph.D. degree from the University of Illinois, Urbana-Champaign in 1973, 1974, and 1978, respectively, all in electrical engineering.

From 1978 to 1982 he was an Assistant Professor of Electrical and Computer Engineering at the University of Iowa, Iowa City, IA. He was with the Long Range Research Division of the Exxon Production Research Company, Houston, TX, from June 1982 to Sept. 1989. He joined the Department of Electrical & Computer Engineering, Auburn University, Auburn, AL, in September 1989 as a Professor. He currently holds the title of James B. Davis Professor. His current research interests are in statistical signal processing, wireless communications, and multiple target tracking.

Dr. Tugnait has served as Associate Editor of the IEEE Transactions on Automatic Control, the IEEE Transactions on Signal Processing, IEEE Signal Processing Letters, and the IEEE Transactions on Wireless Communications, and as Senior Area Editor of the IEEE Transactions on Signal Processing, and Senior Editor of IEEE Wireless Communications Letters.