Joint Spectrum Sensing and Jamming Detection with Correlated Channels in Cognitive Radio Networks

Jafar Mohammadi∗§, Sławomir Stańczak∗§, and Meng Zheng ‡

∗Communications and Information Theory Chair, Technische Universität Berlin, Einsteinufer 27, 10587 Berlin, Germany
§Fraunhofer Heinrich Hertz Institute, Einsteinufer 37, 10587 Berlin, Germany
‡Key Laboratory of Networked Control Systems, Chinese Academy of Sciences, Shenyang 110016, China

Email: {jafar.mohammadi, slawomir.stanczak}@mk.tu-berlin.de, and zhengmeng_6@sia.cn

Abstract—In a cognitive radio (CR) scenario, we study the joint problem of spectrum sensing and jamming detection. Modelling the scenario as a multiple hypothesis testing problem, we analyse the probability of detection of the optimal detector in the sense of Neyman-Pearson theorem. We derive one exact form in terms of a series and a closed-form version. Moreover, we evaluate the asymptotic probability of detection, as it results in a simpler form to handle. In all of the above analysis, we consider the spatially correlated observation data. We further consider two practical scenarios where first we have no knowledge of the jammer’s signal, and second where we have no knowledge of the noise power. We apply generalized likelihood ratio test in both of the cases. Simulation results confirm the accuracy of our asymptotic performance derivations.

Index Terms—Correlated signal detection, Jamming detection, Spectrum sensing, Likelihood ratio test.

I. INTRODUCTION

Cognitive radio (CR) systems are expected to tackle the resource dearth in wireless communication systems with particular focus on spectrum efficiency. Potential efficiency gain stem from the fact that the licensed spectrum, which is intended to be used by the primary users (PU), is not fully occupied by the PUs. In order to use the spectrum more efficiently, a set of users called the secondary users (SU) are supposed to exploit the empty time-frequency spots. According to the seminal works in this vast topic [1], [2], spectrum sensing is an etiquette accomplished by the SUs to detect the empty spectrum holes. Spectrum sensing legalizes the SUs to scan the spectrum, consequently, to listen to the PU. While this feature of CR systems enables the possibility of utilizing the spectrum in a more efficient manner, it may also introduce security vulnerabilities.

Researchers present different approaches for spectrum sensing. For instance, energy detection, matched-filter, and cyclostationarity-based detection are presented respectively in [2], [3], and [4]. Among those, energy detection is a long-established method since its performance is independent of the PU signal structure. In order to increase the accuracy and robustness of the energy detector, the authors in [5] propose a cooperative energy detection. They apply hard decision rule, an “OR” fusion rule [5], while in general the soft decision rules outperform the hard decision rules [6]. The common assumption when applying energy detector is that the observation signals are independent. In the case of spatially correlated signals, the authors in [7] show that the weighted energy detector is optimal for the case that one hypothesis has correlated observations while the other hypothesis has independent signals (null hypothesis). Moreover, [7] does not present the probability distribution of the weighted energy detector test statistic. Thus, no analytical results on the performance is established.

As far as security is concerned, one may note that certain peculiarities of CR networks make them quite vulnerable not only to the conventional threats targeted to wireless communications, but also to the specific type of attacks that are tailored to CR networks [8]. One of the most common attacks that could target CR networks is jamming. Jamming can be performed in different forms, for instance, constant jamming, random jamming, reactive jamming, and so on, which can target both PU and the SUs [8]. There are also more intelligently designed jammers that deceive the SUs by mimicking the signal characteristics of the PU to force the SUs to vacate the spectrum [9]. This is known as primary user emulation attacks (PUEA) in [9], [10].

In this work, we consider two of the main issues of the SUs, namely spectrum sensing and jamming detection, jointly. In order to model the two problems simultaneously, we adopt multiple hypothesis testing with four cases. Building upon the work of [7], we further derive the optimal detector when both of the hypotheses in each test are consists of correlated signals. Furthermore, we derive the exact and the asymptotic probability of detection of the optimal detector. We apply the generalized likelihood ratio test (GLRT) for practical scenarios, such as unknown power of the jammer and unknown
noise covariance matrix. Although there are similarities in the assumptions of this work and the research on PUEA, our work does not fall into that category due to its different approach and objective. The contributions of this paper can be summarized as follows:

1. We address the problem of joint detection of the jammer and the PU in the context of a multiple hypothesis testing problem and analyse the performance.
2. We derive the exact and the asymptotic probability distributions of the test statistic. For the former, we present the cumulative distribution function (CDF) in terms of Meijer’s G function.
3. We present a correlated GLRT solution for the practical scenarios of unknown jammer’s power and unknown noise’s power.

II. SYSTEM MODEL AND DEFINITIONS

We consider a pair of PU transceivers communicating in a licensed frequency band and K SU nodes performing cooperative spectrum sensing. Furthermore, there is a jammer which sends jamming signals in random intervals to interrupt the PU and the SU nodes as in Fig. 1. The main goal of the SU nodes is to determine the presence of the PU and discriminate it from the jammer. We assume the SU nodes send the data to a fusion center (FC) for the decision making process. The SU nodes are desired to distinguish the following spectrum states out of the observation signal: the channel is idle, only the PU is present, only the jammer is present, and both the PU and the jammer are present.

The sensed signal at the ith SU node, i ∈ {1, . . . , K}, under four hypotheses at the nth time instant, n ∈ {1, . . . , N}, is

\[ H_0 : x_i(n) = \nu_i(n) \]
\[ H_1 : x_i(n) = A_1 h_i(n) s(n) + \nu_i(n) \]
\[ H_2 : x_i(n) = A_2 g_i(n) u(n) + \nu_i(n) \]
\[ H_3 : x_i(n) = A_1 h_i(n) s(n) + A_2 g_i(n) u(n) + \nu_i(n) \]

where, \( s(n) \) and \( u(n) \) are complex signals with constant average power 1 from the PU and the jammer, respectively. \( h_i(n) \) and \( g_i(n) \) are channel coefficients between the PU and the ith SU node and between the jammer and the ith SU node, respectively. The assumptions on the channels and noise are provided as follows:

1) The channel coefficients \( h_i(n) \) and \( g_i(n) \) are assumed to be Rayleigh-distributed with \( E(|h_i(n)|^2) = E(|g_i(n)|^2) = \sigma_h^2 \), w. l. g. \(^1\)
2) We assume the channels are spatially correlated but independent temporally. The channel correlation coefficient matrix for \( g := [g_1, \ldots, g_K] \) and \( h := [h_1, \ldots, h_K] \) is defined as \( \Sigma := E[gg^H] = E[hh^H] \). Further, the \((i,j)\)th element of \( \Sigma \) is approximated by

\[
[\Sigma]_{i,j} = \begin{cases} 
  e^{-\rho d_{i,j}} & i \neq j \\
  1 & i = j
\end{cases}
\]  

(1)

where \( d_{i,j} \) is the Euclidean distance between nodes \( i \) and \( j \) and \( \rho \) is the correlation constant depending on the environment \([11]\).
3) Since it is reasonable to assume that the locations of the jammer and the PU are separated large enough\(^2\), it follows that the \( g_i \) and \( h_i \) are mutually independent.
4) The noise at the time \( n \) and the node \( i \) is depicted by \( \nu_i(n) \), which is assumed to be i.i.d. circularly symmetric Gaussian with \( CN(0, \sigma^2_n) \). Further, we assume that \( s(n) \), \( u(n) \), and \( \nu_i(n) \) are, w. l. g., independent.

The distribution of the observation signal under all hypotheses becomes,

\[ x \sim CN(0, P_i \Sigma + \sigma^2_n I), \quad x \in H_l \]  

(2)

where \( x := [x_1, \ldots, x_K]^T \) and for \( H_l, l \in \{0, 1, 2, 3\} \) we have \( P_1 := A_1^2 \sigma^2_h \) with \( A_1^2 := 0 \) and \( A_2^2 := A_1^2 + A_2^2 \). We assume \( A_1 \neq A_2 \) throughout this paper, since, distinguishing \( H_1 \) from \( H_2 \) using energy based detection is meaningless\(^3\).

III. OPTIMAL A POSTERIORI DETECTION

In the interest of optimally selecting the correct hypothesis, we are required to find a hypothesis that maximizes

\[ p(x|H_l)P_i(H_l), \quad l \in \{0, 1, 2, 3\}, \]  

(3)

where the \( P_i(H_l) \) is a priori probability of \( H_l \) \([13]\). In order to evaluate \( \arg \max_l p(x|H_l)P_i(H_l) \) we compare every two hypotheses using

\[ \frac{p(x|H_l)P_i(H_l)}{p(x|H_j)P_i(H_j)} \geq 1, \quad 0 \leq l < j \leq 3, \quad l \neq j \]  

(4)

Then we perform the binary search to find the maximum. We compare \( H_0 \) and \( H_1 \) and in parallel \( H_2 \) and \( H_3 \) using (4). Then we compare the “winners” of each of the previous tests to determine the overall selected hypothesis.

\(^1\)The general case of \( E(|h_i(n)|^2) \neq E(|g_i(n)|^2) \) with \( E(|h_i(n)|^2) = E(|g_i(n)|^2) \) and \( E(|g_i(n)|^2) = E(|h_i(n)|^2) \), \( \forall i,j \) can be considered with no extra elaboration.
\(^2\)More than half of the wavelength of the carrier frequency \([12]\).
\(^3\)One has to exploit other aspects of the jammer’s signal to be able to detect efficiently in such a condition \([9]\).
IV. PERFORMANCE ANALYSIS

In this section we study the performance of the detection tests in (4). Based on Section III, we need to make decisions between every $\mathcal{H}_l$ and $\mathcal{H}_j$, $l, j \in \{0, 1, 2, 3\}$. In the following, we consider three cases, namely, known average power of the jammer and the noise, unknown power of the jammer (with a given noise power), and unknown noise power (with a given the jammer power).

A. The Average Power of the Jammer and Noise are Known

In this case, based on the Neyman-Pearson theorem, one can apply log likelihood ratio (LLR) to achieve optimal detector. The following propositions present the optimal detector’s performance for two general hypotheses $\mathcal{H}_l$ and $\mathcal{H}_j$ where $l, j \in \{0, 1, 2, 3\}$.

Proposition 1: Assume that we are to decide between any two hypotheses $\mathcal{H}_l$ and $\mathcal{H}_j$ in (2). The Neyman-Pearson optimal test statistic is given by,

$$T_{l,j} = \sum_{n=1}^{N} \sum_{i=1}^{K} \lambda_i |z_i(n)|^2 \geq \vartheta_{l,j},$$

where $S_{l,j} = V \text{diag}[^\lambda_1, \ldots, \lambda_K] V^H$ with,

$$S_{l,j} := (P_j \Sigma + \sigma_n^2 I)^{-1} - (P_j \Sigma + \sigma_n^2 I)^{-1},$$

and $z := V^H x$ and $\text{diag}[^\lambda]$ returns a diagonal matrix with its argument as the diagonal elements. Moreover, $\vartheta_{l,j} := -\ln \frac{P_j|\mathcal{H}_j|}{P_j|\mathcal{H}_l|} - N \ln \det (P_j \Sigma + \sigma_n^2 I) + N \ln \det (P_j \Sigma + \sigma_n^2 I)$, is the decision threshold.

The proof is provided in the appendix VIII-A.

Note that for the case of correlated observations, conventional energy detector is not only suboptimal in the sense of Neyman-Pearson theorem, but also, the derivation of the probability distribution of the test statistic involves handling sum of correlated random variables [14].

To calculate a sensing performance measure, e.g. the probability of detection, we need to find the distribution of $T_{l,j}$. The following propositions yield the exact, a computationally simpler version of the exact, and the asymptotic distributions of $T_{l,j}$.

Proposition 2: The CDF of (5) conditioned on $\mathcal{H}_l$ and tested against $\mathcal{H}_j$, is given by,

$$P_l(T_{l,j} \leq \vartheta_{l,j}|\mathcal{H}_l) = C \sum_{q=0}^{\infty} \int_{0}^{\vartheta_{l,j}} \frac{\delta_q (NK+q-1)}{(\lambda_1 \theta_1)^{NK+q}} e^{-\frac{\lambda q}{\lambda_1 \theta_1}} (NK+q+1)^{\frac{q+1}{\lambda q}} dt, $$

where $C := \prod_{k=1}^{K} (\lambda k \theta_k)^{N}$, $\lambda_1 \theta_1 = \min_i \{\lambda_i \theta_i\}$ and we obtain $\delta_q$ using the following recursive formula,

$$\delta_0 = 1$$

$$\delta_{q+1} = \frac{N}{q+1} \sum_{k=1}^{K} \sum_{i=1}^{K} \left(1 - \frac{\lambda_i \theta_1}{\lambda_1 \theta_i}\right)^k \delta_{q+1-k},$$

and $\{\lambda_i \theta_i\}$ are defined as in the following,

The proof is provided in the appendix VIII-B.

In practice we need to truncate the series in (7) depending on the required precision [15]. In order to present a more compact and closed-form formula, we can apply the recent results in [16].

Corollary 1: The exact CDF of (5) conditioned on $\mathcal{H}_l$ and tested against $\mathcal{H}_j$, is given by

$$P_l(T_{l,j} \leq \vartheta_{l,j}|\mathcal{H}_l) = \left(\prod_{i=1}^{K} \frac{1}{2 \lambda_i \theta_i}\right)^{NK} G_1^{NK,0} \left[ e^{-\vartheta_{l,j}} \left[\sum_{K} \left(\frac{1}{\lambda K_1}, 1\right), 0 \right] \right]$$

in which $\Xi^{(1)}_K, \Xi^{(2)}_K \in \mathbb{R}_{1 \times NK}$ are defined as in the following,

$$\Xi^{(1)}_K := \{(1 + \frac{1}{2 \lambda_1 \theta_1}), \ldots, (1 + \frac{1}{2 \lambda_j \theta_j}), \ldots\}$$

and $G_{1+NK,0}$ is the Meijer’s G function [16].

The proof is provided in the appendix VIII-C.

Considering (8), we obtain the exact probability of detection of $\mathcal{H}_l$ while tested against $\mathcal{H}_j$ as

$$P_l(T_{l,j} > \vartheta_{l,j}|\mathcal{H}_l) = 1 - \left(\prod_{i=1}^{K} \frac{1}{2 \lambda_i \theta_i}\right)^{NK} G_1^{1+NK,1+NK} \left[ e^{-\vartheta_{l,j}} \left[\sum_{K} \left(\frac{1}{\lambda K_1}, 1\right), 0 \right] \right]$$

Calculating the exact forms might be computationally involved. Moreover, the asymptotic behaviour of the distribution of (5) delivers useful insights on studying optimum sensing strategies. Hence, in the following we present the distribution of (5) when $N \to \infty$.

Proposition 3: The asymptotic probability distribution of (5) when $N \to \infty$, conditioned on $\mathcal{H}_l$ tested against $\mathcal{H}_j$, is

$$T_{l,j} \sim \mathcal{N} \left[N \sum_{i=1}^{K} \theta_i \lambda_i, N \sum_{i=1}^{K} (\theta_i \lambda_i)^2\right],$$

where $\theta_i$s are eigenvalues of $P_j \Sigma + \sigma_n^2 I$.

The proof is provided in the appendix VIII-D.

Since we do not have the values for the a priori probabilities, to determine $\vartheta_{l,j}$ we proceed as in the following. First we define the probability of false alarm. Then for a fixed probability of false alarm we obtain the corresponding $\vartheta_{l,j}$, which can be used to compute probability of detection of different hypotheses. The probability of false alarm of $\mathcal{H}_l$ tested against $\mathcal{H}_j$, using the proposition 3, is given by

$$P_l(T_{l,j} > \vartheta_{l,j}|\mathcal{H}_j) = \left(\frac{\vartheta_{l,j} - N \sum_{i=1}^{K} \tilde{\theta}_i \lambda_i}{\sqrt{N \sum_{i=1}^{K} (\tilde{\theta}_i \lambda_i)^2}}\right),$$

where $\tilde{\theta}_i$s are the eigenvalues of $P_j \Sigma + \sigma_n^2 I$. 
B. The Average Power of the Jammer is Unknown

One of the most widely used methods to deal with an unknown parameter is GLRT. Basically in GLRT one obtains an estimate of the unknown parameter by applying maximum likelihood estimate (ML).

In typical scenarios, we have no information of the jammer, therefore, we need to estimate the covariance matrix of the observed signal under $H_2$ as well as $H_3$. In our channel model, however, we only need to know the average power of the jammer $P_j$, because the covariance matrix depends on the sensors’ locations.

In the following we estimate $P_j$, assuming that other parameters are known.

**Proposition 4:** Assuming the $H_2$ in (2), the ML estimate of $P_j$ is given by

$$\hat{P}_j = \frac{\sum_{i=1}^{K} \sum_{n=1}^{N} |z_i(n)|^2 - NK\sigma_j^2}{N \sum_{i=1}^{K} \tau_i},$$

where, $\tau_i$s are defined the eigenvalues of $\Sigma := V \text{diag}(\tau_1, \ldots, \tau_K) V^H$.

The proof is presented in appendix VIII-E.

Next we need to substitute (13) into (17). Therefore, the LRT for $H_2$ becomes as the following,

$$-N(\ln \det(P_j \Sigma + \sigma_j^2 I)) + N(\ln \det(\hat{P}_j \Sigma + \sigma_j^2 I))$$

$$- \sum_{n=1}^{N} x^H(n) ( (P_j \Sigma + \sigma_j^2 I)^{-1} - (\hat{P}_j \Sigma + \sigma_j^2 I)^{-1} ) x(n) \geq 0,$n

which is difficult to further analyse.

C. Unknown Noise Power

In the case of unknown noise, similar to the previous section, we use ML estimation. Based on our assumption, we know that noise covariance matrix is diagonal with equal diagonal elements. Hence, the matrix can be described with a single parameter.

**Proposition 5:** The ML estimate of $\sigma_j^2$ for $H_j$

$$\hat{\sigma}_j^2 = \text{arg max}_{\sigma_j^2} -NK \ln \pi - N \ln \det(P_j \Sigma + \sigma_j^2 I)$$

$$-N \text{Tr}( (P_j \Sigma + \sigma_j^2 I)^{-1} R )$$

in which $R := \frac{1}{N} \sum_{n=1}^{N} x(n)x^H(n)$ is the sample covariance matrix of the received signal and $\text{Tr}(\cdot)$ returns the trace of its argument. The ML estimated $\sigma_j^2$ is

$$\hat{\sigma}_j^2 = \frac{1}{NK} \sum_{i=1}^{K} \sum_{n=1}^{N} |z_i(n)|^2 - \frac{1}{K} \sum_{i=1}^{K} P_j \tau_i$$

The proof is similar to that of proposition 4 given in appendix VIII-E.

We need to substitute (15) in LRT to achieve the test statistic. This in turn is given in the following,

$$-N(\ln \det(P_j \Sigma + \hat{\sigma}_j^2 I)) + N(\ln \det(P_j \Sigma + \hat{\sigma}_j^2 I))$$

$$- \sum_{n=1}^{N} x^H(n) ( (P_j \Sigma + \hat{\sigma}_j^2 I)^{-1} - (P_j \Sigma + \hat{\sigma}_j^2 I)^{-1} ) x(n) \geq 0,$n

which is difficult to analyse.

V. Numerical Results

We provide numerical experiments to validate the results in this work. In our simulations we assume $K = 8$ SUs spread out in a square area of $10m \times 10m$ and $N = 10$ independent time samples unless otherwise is specified. The values for $\rho$ are 10 (almost independent) and 0.01 (highly correlated).

We first compare the exact CDF of $T_{i,j}$, given by (8), with the asymptotic CDF, presented by (11), to study the accuracy of the results. This is depicted in Fig. 2. One can observe that with the increase in $N$ and $K$ the asymptotic results tend to be more accurate.

For the case of known power of the jammer and noise, we depict the receiver operating characteristic (ROC) of $P_i(T_{1,0} \geq \vartheta_{1,0}|H_1)$ and $P_i(T_{2,1} \geq \vartheta_{2,1}|H_2)$ in Fig. 3. We notice that in the fully correlated case the weighted energy detector (WED) outperforms the conventional energy detector.

Unlike the case of $P_i(T_{2,1} \geq \vartheta_{2,1}|H_2)$, in the case of $P_i(T_{1,0} \geq \vartheta_{1,0}|H_1)$, the performance of WED boosts when correlation increases. One notes in the case of $P_i(T_{2,1} \geq \vartheta_{2,1}|H_2)$, both of $H_1$ and $H_2$ have correlated observations whereas in the case of $P_i(T_{1,0} \geq \vartheta_{1,0}|H_1)$, $H_1$ has independent observation. The performance of the case with unknown $P_j$

![Fig. 2. The CDF of $T$ versus its asymptotic CDF for the cases $N = 3, K = 3$ and $N = 10, K = 4$.](image)

is better than the case with unknown noise power, as it is illustrated in Fig. 4. The reason of performance degradation in GLRT in comparison to LRT is the error in ML estimation of the parameters. Since the ML estimation of noise appears in GLRT in comparison to LR T for $H_2$, the performance of the detector suffers more in comparison to the case that we incorporate the ML estimate of $P_j$, which appears only in $H_2$.

VI. Conclusion

We model the joint problem of spectrum sensing and jammer detection using multiple hypothesis testing. For the case of spatially correlated observations we derive the optimal test, which results in the weighted energy detector. Further,
we obtain the exact and an asymptotic probability of detection. Using GLRT, we study the practical cases of namely, unknown noise power and unknown jammer power.

VII. ACKNOWLEDGEMENT

The authors would like to thank Steffen Limmer, from Fraunhofer Heinrich Hertz Institute for helpful discussions.

VIII. APPENDIX

A. Proof of Proposition 1

First note that based on our assumptions in (2) we have,

\[ p(x|H_j) = \prod_{n=1}^{N} \frac{1}{\pi^R \det(P_j \Sigma + \sigma^2 I)} \times \exp(-x(n)^H (P_j \Sigma + \sigma^2 I)^{-1} x(n)). \]  

(16)

We substitute (16) into (4) and take the log to achieve,

\[ N \ln \det(P_j \Sigma + \sigma^2 I) - N \ln \det(P_\nu \Sigma + \sigma^2 I) \]

\[ + \sum_{n=1}^{N} x(n)^H S_{l,j} x(n) + \ln \frac{P_l(H_l)}{P_\nu(H_\nu)} \geq 0, \]

where, \( S_{l,j} \) is Hermitian and defined in (6). We proceed by keeping the data dependent parts in the LHS as,

\[ \sum_{n=1}^{N} x(n)^H S_{l,j} x(n) \geq \vartheta_{l,j}, \]

(18)

in which, \( \vartheta_{l,j} \) is the threshold that is defined in the proposition. We change the variable as \( z := V^H x \), where \( V \) is a unitary matrix given by the eigenvalue decomposition of \( S_{l,j} = V \text{diag}[\lambda_1, \ldots, \lambda_K] V^H \). Therefore, the test statistic becomes the weighted energy detector and is given as,

\[ \sum_{n=1}^{N} z(n)^H \text{diag}[\lambda_1, \ldots, \lambda_K] z(n) \geq \vartheta_{l,j}, \]

(19)

by which we can infer (5) directly.

B. Proof of Proposition 2

Let us break the problem into smaller pieces. We first find the distribution of \( t_i = \sum_{i=1}^{N} \lambda_i |z_i(n)|^2 \), then we use it to find \( T_{l,j} \).

Based on our assumptions across time the observation symbols are independent. Therefore, assuming \( x_i(n) \in H_l, \forall i \) then the distribution of \( z_i(n) \) is \( \mathcal{CN}(0, \theta_i) \), where \( \theta_i \) is the \( i \)th eigenvalue of \( (P_i \Sigma + \sigma_i^2 I) \). It follows then, the \( \lambda_i |z_i(n)|^2 \) is distributed according to \( \Gamma(1, \frac{1}{2\vartheta_i \lambda_i}) \). Thus, based on the rule of summation of i.i.d. gamma random variables, \( t_i \)'s distribution is given as \( \Gamma(N, \frac{1}{2\vartheta_i \lambda_i}) \).

In order to find the distribution of \( T_{l,j} = \sum_{i=1}^{K} t_i \) we need to have the distribution of the sum of independent but non-identical gamma random variables. This in turn, is given by the work in [15]. Therefore, we apply the results in [15] with our parameters and obtain (7).

C. Proof of Corollary 1

Proceeding along similar lines as in Appendix VIII-B, we find the probability of \( T = \sum_{i=1}^{K} t_i \), where \( t_i \sim \Gamma(N, \frac{1}{2\vartheta_i \lambda_i}) \). To this end, we directly apply Corollary 3 in [16].

D. Proof of Proposition 3

Since \( x_i \)'s are Gaussian the distribution of \( z \) is Gaussian as well, although with different correlation coefficient matrix. The correlation matrix of \( z \), conditioned on \( H_l \) is given by

\[ \Sigma_z = \text{diag}[\theta_1, \ldots, \theta_K], \]

where \( \theta_i \)'s are defined as \( P_j \Sigma + \sigma_j^2 I := V \text{diag}[\theta_1, \ldots, \theta_K] V^H \).

The test statistic is weighted sum of independent non-identical
gamma random variables. According to [17] the moment generating function of $T_n := \sum_{i=1}^{K} \lambda_i|z_i(n)|^2$ is given by
\[ M_z(s) = \prod_{i=1}^{K} (1 - s \lambda_i)^{-1}. \tag{20} \]
From (20) one can achieve the mean and the variance of $T_n$ as,
\[ E[T_n] = \sum_{i=1}^{K} \theta_i \lambda_i, \quad \text{Var}[T_n] = \sum_{i=1}^{K} (\theta_i \lambda_i)^2. \tag{21} \]
Based on the assumption that temporal observations, i.e. $T_n$, are i.i.d. and $N$ is sufficiently large, we apply the central limit theorem to approximate the distribution of $T$ as,
\[ T \sim \mathcal{N} \left( N \sum_{i=1}^{K} \theta_i \lambda_i, N \sum_{i=1}^{K} (\theta_i \lambda_i)^2 \right). \]

E. Proof of Proposition 4

We can take log of $\mathcal{H}_2$ in (2), then the ML problem becomes,
\[ \hat{P}_2 = \arg \max_{P_2 \geq 0} f(P_2) \tag{22} \]
\[ f(P_2) := -NK \ln \pi - N \sum_{i=1}^{K} \ln (P_2 \tau_i + \sigma_v^2) \tag{23} \]
\[ - \sum_{n=1}^{N} x^H(n) \mathbf{V}(P_2 \text{diag}(\tau_1, ..., \tau_K) + \sigma_v^2 \mathbf{I})^{-1} \mathbf{V}^H x(n). \]
Let us define $f'(P_2)$ and $f''(P_2)$ to be the first and second derivatives of $f(P_2)$, respectively. We denote the only solution of $f''(P_2) = 0$ as $P_2^*$. A simple analysis reveals:
\[ \begin{cases} f''(P_2) < 0 & P_2 \in [0, P_2^*) \tag{24} \\ f''(P_2) > 0 & P_2 \in (P_2^*, \infty) \end{cases} \]
Thus, $f(P_2)$ is concave for $P_2 \in [0, P_2^*)$ and then it becomes convex for $P_2 \in (P_2^*, \infty)$. Therefore, we have a unique local maximum on $P_2 \in [0, P_2^*)$ which we indicate it with $\hat{P}_2$. In the following we obtain this local maximum by solving,
\[ f'(P_2) = N \sum_{i=1}^{K} \tau_i (P_2 \tau_i + \sigma_v^2)^{-1} - \sum_{i=1}^{K} \sum_{n=1}^{N} \tau_i (P_2 \tau_i + \sigma_v^2)^{-2} |z_i(n)|^2 = 0 \tag{25} \]
\[ N \tau_i (P_2 \tau_i + \sigma_v^2)^{-1} - \tau_i \sum_{n=1}^{N} |z_i(n)|^2 = 0 \]
\[ = NK \sigma_v^2 + NP_2 \sum_{i=1}^{K} \tau_i - \sum_{i=1}^{K} \sum_{n=1}^{N} |z_i(n)|^2 = 0. \]
The equation in (25) yields $\hat{P}_2 = \infty$ and
\[ \hat{P}_2 = \frac{\sum_{i=1}^{K} \sum_{n=1}^{N} |z_i(n)|^2 - NK \sigma_v^2}{N \sum_{i=1}^{K} \tau_i}. \tag{26} \]
\(\hat{P}_2 = \infty\) results in a minimum for $f$, thus we only have one maximum which is (26). In the following we prove that it is the global maximum.

For $P_2 \in [\hat{P}_2, \infty]$ we note that $f'(P_2) \leq 0$, hence, $f(P_2)$ is non-increasing. This implies that $\hat{P}_2$ is the global maximum. In the trivial case of $\sum_{i=1}^{K} \sum_{n=1}^{N} |z_i(n)|^2 \leq NK \sigma_v^2$ due to the non-negativity constraint, we select $\hat{P}_2 = 0$. Therefore we conclude (13).

References