

# An Exponentially-Tight Approximate Factorization of the Joint PDF of Statistical Dependent Measurements in Wireless Sensor Networks

JUAN AUGUSTO MAYA<sup>1</sup>, LEONARDO REY VEGA<sup>2,3</sup>,  
AND ANDREA M. TONELLO<sup>1</sup> (Senior Member, IEEE)

<sup>1</sup>Networked and Embedded Systems (NES), University of Klagenfurt, 9020 Klagenfurt, Austria

<sup>2</sup>Facultad de Ingeniería, Universidad de Buenos Aires, Buenos Aires C1053 CABA, Argentina

<sup>3</sup>Centro de Simulación Computacional (CSC), Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Buenos Aires C1425FQD, Argentina

CORRESPONDING AUTHOR: A. M. TONELLO (e-mail: andrea.tonello@aau.at)

The work of Juan Augusto Maya was supported by the Ubiquitous Sensing Lab, a Joint Laboratory between the University of Klagenfurt and Silicon Austria Labs. The work of Leonardo Rey Vega was supported by PICT under Grant 2020-Serie A-01336.

**ABSTRACT** We consider the distributed detection problem of a temporally correlated random radio source signal using a wireless sensor network capable of measuring the energy of the received signals. It is well-known that optimal tests in the Neyman-Pearson setting are based on likelihood ratio tests (LRT), which, in this set-up, evaluate the quotient between the probability density functions (PDF) of the measurements when the source signal is present and absent. When the source is present, the computation of the joint PDF of the energy measurements at the nodes is a challenging problem. This is due to the statistical dependence introduced to the received signals by the propagation through fading channels of the radio signal emitted by the source. We deal with this problem using the characteristic function of the (intractable) joint PDF, and proposing an approximation to it. We derive bounds for the approximation error in two wireless propagation scenarios, slow and fast fading, and show that the proposed approximation is exponentially tight with the number of nodes when the time-bandwidth product is sufficiently high. The approximation is used as a substitute of the exact joint PDF for building an approximate LRT, which performs better than other well-known detectors, as verified by Monte Carlo simulations.

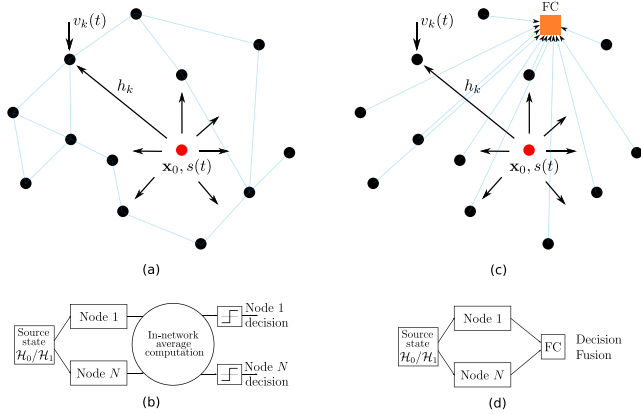
**INDEX TERMS** Distributed detection, wireless sensor networks, joint PDF factorization, statistically dependent observations.

## I. INTRODUCTION

WIRELESS sensor networks (WSNs) as a key technology in the emerging paradigm of Internet of Things (IoT) [1], [2], [3] have received considerable attention. Distributed signal processing is an important topic of research in this area, because efficient information processing in large networks of devices with limited communication, sensing, storage and computing capabilities has the potential of being cost-efficient and very robust [4]. Among the different signal processing tasks in WSNs, distributed detection is one of the most important [5], [6], [7].

In the distributed detection problem, a set of nodes sense the environment in search for the presence of a source signal, which is typically linked with some physical process

extended over the geographical area where the network is deployed. Through collaboration among the nodes, the network is expected to decide with high confidence if the above mentioned signal is present or not. A well-studied application of this general problem is the spectrum sensing task in cognitive radios [8]. An important issue is the presence and influence of the spatial and temporal correlation of the source signal on the implemented detection scheme. The source signal is sometimes modeled as a stochastic process which could present temporal correlation (e.g., cyclostationarities, [9]). On the other hand, when the source signal is present, the measurements taken at each sensor node are clearly correlated (and therefore, statistically dependent) because these measurements are different noisy versions



**FIGURE 1.** Two common network architectures. (a) Fully-distributed scenario. All nodes (black dots) compute a local statistic, exchange information through short-range communication links (light-blue lines), and lastly, build the final test statistic to make a decision. (b) A block diagram for the scenario in (a). (c) The network has a FC (orange square), where all the (quantized) measurements are sent from the nodes and the final decision is made. (d) A block diagram for the scenario in (c).

(affected by channel effects such as path-loss, shadowing, and fading) of the same random signal.

### A. MOTIVATION AND SOME PREVIOUS WORKS

It is well-known that the optimal test [10], which can take into account the correlation (or more generally, the statistical dependence) of the signals, is the likelihood ratio test (LRT), which is defined through the quotient between the probability density functions (PDF) of the measurements when the signal is present and not. However, in general, the exact implementation of the optimal LRTs is difficult in fully-distributed settings, where there is no fusion center (FC), and the cooperation between different sensor nodes is done through transmissions between neighbor nodes. See Fig. 1 (a) and (b) for a schematic representation of the problem. The reason is that those tests require network-wide interactions between the measurements taken at the sensor nodes. For example, when the signals captured at each sensor site are Gaussian, the optimal tests involve the computation of quadratic forms of the measurements. In fully-distributed consensus detection schemes [11], [12], [13], this is problematic when the covariance matrix of the measurements is not diagonal. The number of exchanges between the nodes in the network needed to construct the consensus decision could be large, which introduces severe penalties in power and bandwidth consumption, which are typically scarce resources in WSNs composed of typically inexpensive nodes. When a FC is present (see Fig. 1 (c) and (d)), the sensors could transmit their measurements to the FC, which is responsible of computing the final decision about the presence of the source signal [14], [15], [16]. This centralized architecture also presents some weaknesses as, for example, it lacks of robustness against the malfunctioning of a single device, given that a failure in the FC may severely degrade the performance of the system. Additionally, it requires that sensor nodes, typically battery powered devices,

communicate through orthogonal channels with the FC, in which case the energy and bandwidth resources increase with the amount of sensors and the area in which the network is deployed. This could be alleviated, in principle, using fusion decision algorithms [17], [18], [19], [20], [21], where the local measurements (or a statistic of them) are quantized to a few bits. These decisions are then communicated to a FC, where a typically suboptimal final test statistic is built and the final decision is made. However, the influence of spatial/temporal correlation in the construction of the local and final test statistic is not usually taken into account for their design, not modeled, or even discarded without an analytical justification.

### B. OUR CONTRIBUTIONS AND PAPER ORGANIZATION

Instead of restricting us to a specific architecture (with or without a FC), we begin in Section II formulating the problem of detecting an arbitrary temporal-correlated random source signal. We consider that the source signal and the additive noise at each sensor are Gaussian distributed signals, and that the sensors deliver energy measurements. Because the source signal is modeled as a random signal, the node's observations are statistically dependent, given that each sensor receives a noisy version of the same source signal propagated through the corresponding wireless channel. The exact PDF of the energy measurements at the sensor sites when the source signal is present, which corresponds to the diagonal entries of a Wishart-distributed matrix, is not known in closed form [22]. Although there exist some recent results [23], [24] in terms of multidimensional series that involve Laguerre polynomials, the expressions are not particularly easy to handle and compute (specifically in high-dimensional scenarios, i.e., many nodes, due to numerical precision issues), and they do not cover the more general case considered in this paper (the arbitrarily temporally correlated source signal case). From this starting point, our main contribution is the derivation of a tight approximation to the actual joint PDF of the energy measurements, which conveniently factorizes itself in the product of  $N$  (the number of sensors) PDFs, each of one is parameterized through parameters that each node can compute or obtain locally. This approximation is treated in two typical and important propagation scenarios: slow and fast fading. This finding presents two major advantages in distributed detection scenarios: i) It allows us to build simpler cooperative schemes for detecting a radio source by using any of the two network architectures shown in Fig. 1. ii) If desirable, it is possible to design simpler quantization schemes by relying on the local PDF in each node instead of the intractable joint PDF.

In Section III, we calculate in closed form the characteristic function (CF) of the energy measurements. From this exact CF, and under the hypothesis of a high time-bandwidth product  $WT$ , which is a common situation in several practical applications, the approximation to the true joint PDF of the energy measurements is derived. This product PDF is

non-trivial (i.e., it is not the product of the marginals of the true joint PDF) and can be used as a replacement of the true non-product joint PDF for implementing the LRT or quantization schemes, among others. In addition, this result implies that, when  $WT$  is large and for energy-based detectors, the temporal and spatial statistical dependence of the measurements are not critical in this set-up.

In Section IV, a theoretical study of the obtained approximation for the true PDF is also provided for both above-mentioned scenarios, showing the quality of the approximation in terms of the value of  $WT$  and the network size  $N$ . In this respect, the most important fact is that the approximation is exponentially tight with  $N$ . In Section V, in order to cope with the unknown remaining parameters in the obtained approximations, we obtain the corresponding generalized likelihood ratio tests (GLRT). Finally, in Section VI, some numerical simulations are also conducted to evaluate the performance of the obtained tests for detecting the presence of a radio source with unquantized energy measurements. The paper is finalized with some concluding remarks and future research directions. The proofs of some mathematical results are relegated to the appendices.

This work is an extension of our conference paper [25], where only preliminary results were presented. Specifically, we here provide new results for the fast fading scenario, the theoretical error approximation analysis and new simulations for both fast and slow fading scenarios. We also include the mathematical proofs of the results.

### C. NOTATION AND ACRONYMS

Vectors/matrices are denoted by lowercase/uppercase boldface. Also, depending on the context, with uppercase boldface letters we denote random vector and matrices. With  $\text{diag}(a_1, a_2, \dots, a_N)$  we denote the diagonal  $N \times N$  matrix with diagonal entries given by  $a_1, a_2, \dots, a_N$ . The trace and the determinant of square matrix  $\mathbf{A}$  are denoted by  $\text{Tr}(\mathbf{A})$  and  $|\mathbf{A}|$ , respectively. For a complex number  $z \in \mathbb{C}$ , we denote its magnitude with  $|z|$ . The use of  $|\cdot|$  for the determinant of a square matrix or for the magnitude of complex number will be clear from the context. The Kronecker product between matrices  $\mathbf{A}$  and  $\mathbf{B}$  is denoted as  $\mathbf{A} \otimes \mathbf{B}$ . A vector  $\mathbf{y}$  complex circular Gaussian distributed with mean  $\mathbf{a}$  and covariance matrix  $\mathbf{B}$  has PDF denoted by  $\mathbf{y} \sim \mathcal{CN}(\mathbf{a}, \mathbf{B})$ . The  $N \times N$  identity matrix is denoted as  $\mathbf{I}_N$ . Symbols  $P_{\text{fa}}$  and  $P_{\text{md}}$  denote the false-alarm probability and miss-detection probability of the statistical detection test considered. We use the big-O notation, that is,  $f(x) = \mathcal{O}(g(x))$  as  $x \rightarrow \infty$  and for  $g(x)$  a strictly positive function if  $|f(x)| \leq Kg(x)$  when  $x > x_0$  for  $K$  and  $x_0$  positive constants. The symbol  $\equiv$  is used to introduce the definition of a new quantity.

For easier reference, we list some common acronyms used in the text:

CF	Characteristic Function
CSI	Channel State Information

CROC	Complementary Receiver Operating Characteristic
FC	Fusion Center
GLRT	Generalized Likelihood Ratio Test
IoT	Internet of Things
LRT	Likelihood Ratio Test
MLE	Maximum Likelihood Estimator
PDF	Probability Density Function
SNR	Signal to Noise Ratio
UMPT	Uniformly Most Powerful Test
WSN	Wireless Sensor Network

## II. PROBLEM SETTING

Consider the problem of  $N$  sensors distributed in a given geographical area. All of them have sensing and communication capabilities. It is assumed that a radio source is located in the same area. The location of this radio source is unknown for the sensing network, and it is assumed that it emits an arbitrarily temporally-correlated Gaussian signal. Each node of the network senses the environment during  $L$  time-windows, each of duration  $M$  symbols. During the  $l$ -th time-window (with  $l \in [1 : L]$ ), the received signal at each sensor  $n \in [1 : N]$  can be written as:

$$\mathbf{y}_{n,l} = h_{n,l} \mathbf{s}_l + \mathbf{v}_{n,l}, \quad (1)$$

where for each  $n, l$ ,  $\mathbf{y}_{n,l} \in \mathbb{C}^M$  is the temporal baseband signal at sensor  $n$  and time-window  $l$ , which is composed of the noise signal  $\mathbf{v}_{n,l} \in \mathbb{C}^M$ , and the source signal (when present)  $\mathbf{s}_l \in \mathbb{C}^M$ . This source signal is affected by the channel gain  $h_{n,l} \in \mathbb{C}$  which includes, among other effects, the path-loss gain and fading that exists between the position of the source and the  $n$ -th sensor position at measurement window  $l$ . This coefficient, is highly dependent on the scenario under study (i.e., a high mobility situation or a stationary one) and it is typically unknown in practice or difficult and too complex to acquire in scenarios where the source does not cooperate to be detected. Notice that, as the signal  $\mathbf{s}_l$  is the same for each sensor (although affected by different channel coefficients  $h_{n,l}$ ), the signals  $\{\mathbf{y}_{n,l}\}_{n=1}^N$  received at each sensor node are spatially correlated for each measurement window  $l$ . Besides that, as the radio source signal  $\mathbf{s}_l$  can also be temporally correlated, the signals  $\mathbf{y}_{n,l}$  can also present temporal correlation.

It is assumed that the sensors have limited sensing capabilities and they can only sense the average energy during each measurement window. This is a practical and typical situation in several applications such as cognitive radio [26], and it was extensively studied [27], [28]. Energy detectors are typically available even in inexpensive hardware and they do not require coherent demodulation procedures, in which prior information related to the radio source's waveform or modulation is needed. Besides that, the optimal detection test for stochastic Gaussian signals (1) has the well-known estimator-correlator structure that leads to a *weighted* energy decision statistic [29], [30]. Therefore, dealing with the case where the nodes deliver only energy measurements

is reasonable, from a practical and a theoretical point of view. So, for each time-window  $l \in [1 : L]$ , each sensor  $n \in [1 : N]$ , outputs the value  $E_{n,l} \equiv \frac{1}{\beta(M)} \|\mathbf{y}_{n,l}\|^2$ , where  $\beta(M) > 0$  is an appropriate normalizing constant, that depends on the length of  $\mathbf{y}_{n,l}$ . This constant is increasing with  $M$  and some usual choices could be  $\beta(M) = M$  or  $\beta(M) = M^{1/2}$ . These energy values are then used, through the following hypothesis test, by the sensing network to determine if the radio source is actually transmitting or not:

$$\begin{cases} \mathcal{H}_0 : E_{n,l} = \frac{1}{\beta(M)} \|\mathbf{v}_{n,l}\|^2, & n \in [1 : N], l \in [1 : L] \\ \mathcal{H}_1 : E_{n,l} = \frac{1}{\beta(M)} \|h_{n,l}\mathbf{s}_l + \mathbf{v}_{n,l}\|^2. \end{cases} \quad (2)$$

It is well-known that the likelihood ratio between the PDFs under both hypotheses is a key element for optimal detection strategies. We denote with  $\mathbf{E}_l = [E_{1,l}, E_{2,l}, \dots, E_{N,l}]^T$  the energy measurements at each sensor site during time-window  $l$ , and with  $p_l^0(\mathbf{E}_l)$  and  $p_l^1(\mathbf{E}_l)$  the PDFs under  $\mathcal{H}_0$  and  $\mathcal{H}_1$ , respectively. We assume that the energy measurements  $\{\mathbf{E}_l\}_{l=1}^L$  are independent for different time windows  $l$ .<sup>1</sup> The optimal decision statistics in the Neyman-Pearson sense is [10]:

$$\log \frac{\prod_{l=1}^L p_l^1(\mathbf{E}_l)}{\prod_{l=1}^L p_l^0(\mathbf{E}_l)} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \tau, \quad (3)$$

where  $\tau$  is a threshold usually chosen to set the false alarm probability  $P_{\text{fa}} \equiv \mathbb{P}\{\text{decide for } \mathcal{H}_1 | \mathcal{H}_0\}$ . The miss-detection probability is defined by  $P_{\text{md}} \equiv \mathbb{P}\{\text{decide for } \mathcal{H}_0 | \mathcal{H}_1\}$ . For this reason and in order to analyze this hypothesis testing problem, we first need the joint PDF, over the sensing nodes, of the energy measurements  $\mathbf{E}_l$ , for each measurement window.

In order to closely investigate the PDFs under both hypotheses we will make some standard assumptions. Under  $\mathcal{H}_0$  (the radio source is not emitting), the signal  $\mathbf{v}_{n,l}$  is a complex Gaussian circular noise, that is, for each  $n, l$ ,  $\mathbf{v}_{n,l} \sim \mathcal{CN}(\mathbf{0}, \sigma_v^2 \mathbf{I}_M)$ ,  $\sigma_v^2 > 0$ . When the radio source is emitting (under  $\mathcal{H}_1$ ), the signal at each sensor comprises the noise signal  $\mathbf{v}_{n,l}$  with the same characteristics as before, and also the signal  $h_{n,l}\mathbf{s}_l$ . It is assumed that  $\mathbf{s}_l$  is a complex and circular Gaussian signal distributed as  $\mathbf{s}_l \sim \mathcal{CN}(\mathbf{0}, \Sigma_{s,l})$ , independent of the noise signal, and where  $\Sigma_{s,l} \in \mathbb{C}^{M \times M}$  is the temporal correlation matrix of the radio source signal during the measurement window  $l$ .

Although the Gaussian assumption may seem restrictive, it is widely used, even in situations where it is not strictly true. This is the case of, for example, cognitive radio applications, where the radio source's waveform is an orthogonal frequency-division multiplexing (OFDM) signal, with symbols taking values from a discrete set (e.g., QAM). In this example, the signal presents temporal correlation due to the cyclic prefix introduced by OFDM. We include numerical experiments of this scenario in Section VI. Moreover, as it is shown in the next section, the Gaussian

1. In practical scenarios, if  $M$  is set large enough, the temporal statistical dependence of the source signal between different time-windows vanishes.

assumption will permit us to obtain closed-form expressions of several important quantities and to deal with the spatial and temporal correlation of the measurements in the above detection problem.

The model assumed above and the results in Sections III and IV are, in principle, agnostic of the considered fusion setup (fully-distributed or centralized setup with a FC). Depending on the chosen fusion setup, the main differences will be in the demanded network communications resources to compute the test statistics derived from our PDF approximation.

*Remark 1:* It is straightforward to show that the performance of the LRT is invariant to the scaling coefficient  $\beta(M)$  in (2). That is, for a given value of  $P_{\text{fa}}$ ,  $P_{\text{md}}$  does not depend on  $\beta(M)$ . So, the assumed scaling coefficient has no practical consequence to the implementation of the LRT in (3). However, the scaling coefficient will be pivotal in Sections III and IV, where we will analyze the influence of the spatial dependence in  $\mathbf{E}_l$  under  $\mathcal{H}_1$ , when  $M$  is large.

### III. ON THE APPROXIMATION OF THE PDF OF $\mathbf{E}_L$

In this section, we first study the PDFs of the energy measurements under both hypotheses. We consider two scenarios:

1) *The slow fading scenario:* The channel gains between the source signal and sensor sites are fixed for all measurement windows, i.e.,  $h_{n,l} \equiv h_n \forall l \in [1 : L]$  and  $n \in [1 : N]$ . This is representative of a stationary environment.

2) *The fast fading scenario:* The channel gains between the source signal and sensor sites change for each measurement window according to a given distribution (e.g., Rayleigh distribution). This is representative of a high-mobility environment.

For these two scenarios, as obtaining the joint PDF of the measurements is a challenging problem, the study of the CF of  $\mathbf{E}_l$  conditioned on the values of channels gains between the source signal and the sensors proves to be critical. For that reason, we will proceed by obtaining a closed form expression of this CF and use it to obtain and mathematically justify our approximations. In the following, and except when this will be strictly needed to avoid confusion, we will consider a fixed measurement window with  $l \in [1 : L]$  of size  $M$  and we will omit the subscript  $l$  in all relevant quantities in order to use a less cumbersome notation. We will also define the vector of channel gains  $\mathbf{h} = [h_1, h_2, \dots, h_N]^T$ . We will begin analyzing the CF of  $\mathbf{E}$  for the slow fading scenario which is mathematically equivalent to the CF of  $\mathbf{E}$  conditioned on  $\mathbf{h}$ .

#### A. ON THE CHARACTERISTIC FUNCTION OF $\mathbf{E}$ CONDITIONED ON $\mathbf{H}$

We define the vector  $\mathbf{y} \equiv [\mathbf{y}_1^T, \mathbf{y}_2^T, \dots, \mathbf{y}_N^T]^T$  as the  $NM$ -length vector that contains the signals during the considered measurement window at the  $N$  sensor positions. Clearly,  $\mathbf{y} \sim \mathcal{CN}(\mathbf{0}, \sigma_v^2 \mathbf{I}_{NM})$  under  $\mathcal{H}_0$ . It is also not difficult to show that  $\mathbf{y} \sim \mathcal{CN}(\mathbf{0}, \mathbf{h}\mathbf{h}^H \otimes \Sigma_s + \sigma_v^2 \mathbf{I}_{NM})$  under  $\mathcal{H}_1$ .

The exact joint density of the vector  $\mathbf{E}$  is a very difficult problem [22], [31]. It is well-known that it is related to the distribution of the diagonal of  $\mathbf{Y}\mathbf{Y}^H$ , where  $\mathbf{Y}$  is the vertical concatenation of vectors  $\mathbf{y}_n^H$ ,  $n \in [1 : N]$ . Although this problem has attracted some interest in the wireless communications community (see [23] and [24]), most of the results consider the case in which the vectors  $\mathbf{y}_n$ ,  $n \in [1 : N]$  are independent and identically distributed (i.i.d.), which is not the case considered in this paper. The main reason behind this is the presence of the channel coefficient between source and sensor locations. This naturally generates a statistical dependence between vectors  $\mathbf{y}_n$ ,  $n \in [1 : N]$  which is difficult to characterize and quantify. If in addition, the radio source signal presents temporal correlation, this will also contribute to the mentioned dependence. Most solutions, in the i.i.d. case, consider inverse Fourier methods and series expansions for the joint density. This is motivated by the fact that the CF for the diagonal of  $\mathbf{Y}\mathbf{Y}^H$  can be easily computed in that case.

In this paper, we will perform the exact computation of the CF of  $\mathbf{E}$ , taking into account the specific signal model defined in the previous section which does not match the usual characteristics of the i.i.d. signal model assumed in most of the literature. The obtained CF will then be used, not for an exact series expansion for the PDF of  $\mathbf{E}$ , but to obtain a meaningful approximation to the PDF. As we will see, this approximation has very useful properties for the distributed detection problem in (2). Consider the CF of  $\mathbf{E}$  conditioned to  $\mathbf{h}$  and the hypothesis  $\mathcal{H}_i : \Psi^i(\omega|\mathbf{h}) \equiv \mathbb{E}[e^{j\omega^T \mathbf{E}}|\mathbf{h}, \mathcal{H}_i]$  where  $\omega \in \mathbb{R}^N$  and  $i = 0, 1$  denotes the true state of nature (i.e.,  $\mathcal{H}_0$  or  $\mathcal{H}_1$ ). The following lemma, proved in Appendix A, gives us the exact result for  $\Psi^i(\omega|\mathbf{h})$ .

*Lemma 1:* The CF of  $\mathbf{E}$  conditioned to  $\mathbf{h}$  and the hypotheses  $\mathcal{H}_0$  and  $\mathcal{H}_1$  are,<sup>2</sup> respectively,

$$\Psi^0(\omega|\mathbf{h}) = \prod_{n=1}^N \left(1 - j \frac{\omega_n \sigma_v^2}{\beta(M)}\right)^{-M}, \quad (4)$$

$$\Psi^1(\omega|\mathbf{h}) = \frac{\prod_{n=1}^N \left(1 - j \frac{\omega_n \sigma_v^2}{\beta(M)}\right)^{-M}}{\prod_{m=1}^M \left(1 - j \frac{\lambda_m}{\beta(M)} \sum_{n=1}^N \frac{|h_n|^2 \omega_n}{1 - j \frac{\omega_n \sigma_v^2}{\beta(M)}}\right)}, \quad (5)$$

where  $\lambda_1, \lambda_2, \dots, \lambda_N$  are the eigenvalues of the source signal covariance matrix  $\Sigma_s$ .

The previous lemma deserves some comments. In the first place, under  $\mathcal{H}_0$ , the energy vector density is distributed as  $N$  identical and independent central chi-square random variables with  $2M$  degrees of freedom and mean  $\frac{M\sigma_v^2}{\beta(M)}$ . Moreover, as the source is not transmitting, neither the CF or PDF depend on  $\mathbf{h}$ . Under  $\mathcal{H}_1$ ,  $\Psi^1(\omega|\mathbf{h})$  can not be factored as  $\prod_{n=1}^N \psi_n(\omega_n|\mathbf{h})$ , where each  $\psi_n$  with  $n \in [1 : N]$  is a characteristic function. As expected,  $\{E_n\}$

2. Strictly speaking the characteristic function under  $\mathcal{H}_0$  does not depend on  $\mathbf{h}$ . However, in order to use a uniform notation across the paper we will continue using  $\Psi^0(\omega|\mathbf{h})$ .

are dependent random variables independently of the time correlation characteristics of the source signal  $\mathbf{s}$ . From (5), this is true even when the source signal is uncorrelated in time, i.e.,  $\lambda_m = \sigma_s^2$  with  $m \in [1 : M]$ . This is a consequence of the spatial correlation induced by the common source signal present at each sensor site.

In principle, there is no known closed form expression for the joint density of  $\mathbf{E}$ . However it is easy to show, assuming  $\lambda_i \neq \lambda_j$ , with  $i \neq j$ , that the marginal density for  $E_n$  with  $n \in [1 : N]$  can be written as:

$$p^1(E_n|\mathbf{h}) = \sum_{m=1}^M \prod_{j \neq m}^M \left(1 - \frac{\lambda_j |h_n|^2 + \sigma_v^2}{\lambda_m |h_n|^2 + \sigma_v^2}\right)^{-1} \frac{\beta(M)}{\lambda_m |h_n|^2 + \sigma_v^2} \times \exp\left(-\frac{\beta(M)E_n}{\lambda_m |h_n|^2 + \sigma_v^2}\right), \quad n \in [1 : N]. \quad (6)$$

In the special case of an uncorrelated source signal  $\mathbf{s}$ , in which  $\lambda_m = \sigma_s^2$  for all  $m \in [1 : M]$ , we have:

$$p^1(E_n|\mathbf{h}) = \left(\frac{\beta(M)}{\sigma_s^2 |h_n|^2 + \sigma_v^2}\right)^M \frac{E_n^{M-1}}{(M-1)!} \times \exp\left(-\frac{\beta(M)E_n}{\sigma_s^2 |h_n|^2 + \sigma_v^2}\right), \quad n \in [1 : N], \quad (7)$$

that is, a central chi-square random variable with  $2M$  degrees of freedom and mean  $\frac{M(\sigma_s^2 |h_n|^2 + \sigma_v^2)}{\beta(M)}$ . Although the marginal densities are important, in order to implement a hypothesis test for detecting the source signal, we need the full joint PDF of the energy measurements at the sensors. Although, as we pointed out above, this problem is hard, some insights can be obtained when  $M$  is sufficiently large.

## B. APPROXIMATION OF THE JOINT DENSITY OF $\mathbf{E}$ FOR THE SLOW FADING SCENARIO

It is important to analyze, using the characteristic function  $\Psi^1(\omega|\mathbf{h})$  in (5), the case in which  $\beta(M)$  is large. This will happen when  $M$  is large as we have assumed that  $\beta(M)$  is an increasing function of  $M$ . Note that<sup>3</sup>  $M \approx WT$ , where  $W$  is the bandwidth of the continuous time version of the source signal and  $T$  is the analog time duration of the sensing window in which the energy measurement in each sensor is done [32], [33]. For practical applications (e.g., cognitive radio [8], [26]) the time-bandwidth product  $WT$  will not be small, and the following derivations will be useful. Under the assumption of  $M$  large, we will use the following first-order approximation of the exponential function:

$$1 - j \frac{\lambda_m}{\beta(M)} \sum_{n=1}^N \frac{|h_n|^2 \omega_n}{1 - j \frac{\omega_n \sigma_v^2}{\beta(M)}} \approx \exp\left(-j \frac{\lambda_m}{\beta(M)} \sum_{n=1}^N \frac{|h_n|^2 \omega_n}{1 - j \frac{\omega_n \sigma_v^2}{\beta(M)}}\right), \quad (8)$$

3. Note that  $M$  needs to be a positive integer. This is satisfied considering the largest integer smaller or equal to  $WT$ .

for  $m \in [1 : M]$  and  $n \in [1 : N]$ . Under this approximation,  $\Psi^1(\boldsymbol{\omega}|\mathbf{h}) \approx \hat{\Psi}^1(\boldsymbol{\omega}|\mathbf{h})$  where:

$$\hat{\Psi}^1(\boldsymbol{\omega}|\mathbf{h}) \equiv \frac{\prod_{m=1}^M \exp\left(j \frac{\lambda_m}{\beta(M)} \sum_{n=1}^N \frac{|h_n|^2 \omega_n}{1 - j \frac{\omega_n \sigma_v^2}{\beta(M)}}\right)}{\prod_{n=1}^N \left(1 - j \frac{\omega_n \sigma_v^2}{\beta(M)}\right)^M}. \quad (9)$$

Using that  $\sum_{m=1}^M \lambda_m = \text{Tr}(\boldsymbol{\Sigma}_s)$  we obtain:

$$\hat{\Psi}^1(\boldsymbol{\omega}|\mathbf{h}) = \prod_{n=1}^N \frac{\exp\left(\frac{j}{\beta(M)} \frac{\text{Tr}(\boldsymbol{\Sigma}_s) |h_n|^2 \omega_n}{1 - j \frac{\omega_n \sigma_v^2}{\beta(M)}}\right)}{\left(1 - j \frac{\omega_n \sigma_v^2}{\beta(M)}\right)^M}. \quad (10)$$

For a large but fixed value of  $M$ , notice that the approximation in (8) is a very good one when  $|\omega_n|$ ,  $n \in [1 : N]$  are small. Clearly, for larger values, the approximation is not so good. However, the denominator in (10) increases rapidly with  $|\omega_n|$ ,  $n \in [1 : N]$  when  $M$  is large and the scaling  $\beta(M)$  is chosen carefully. This behaviour seems to have the net effect that (10) approximates the true CF (5) reasonably well over large regions in the  $\boldsymbol{\omega}$  space. In Section IV, we will rigorously show that this is indeed the case by providing a bound for  $\sup_{\mathbf{E} \in \mathbb{R}_{>0}^N} |p^1(\mathbf{E}|\mathbf{h}) - \hat{p}^1(\mathbf{E}|\mathbf{h})|$  and obtaining its scaling behaviour with  $M$ .

The most striking fact about the last expression is that, as  $M$  grows, the entries of  $\mathbf{E}$  becomes less statistically dependent between them, as (10) is the CF of  $N$  independent random variables, each with characteristic function given by:

$$\hat{\psi}^1(\omega_n|h_n) \equiv \frac{\exp\left(\frac{j}{\beta(M)} \frac{\text{Tr}(\boldsymbol{\Sigma}_s) |h_n|^2 \omega_n}{1 - j \frac{\omega_n \sigma_v^2}{\beta(M)}}\right)}{\left(1 - j \frac{\omega_n \sigma_v^2}{\beta(M)}\right)^M}, \quad n \in [1 : N]. \quad (11)$$

It is well known that (11) is the characteristic function of a non-central chi-square random variable. In more precise terms, anti-transforming (10) we get  $p^1(\mathbf{E}|\mathbf{h}) \approx \hat{p}^1(\mathbf{E}|\mathbf{h})$ , where

$$\begin{aligned} \hat{p}^1(\mathbf{E}|\mathbf{h}) &\equiv \prod_{n=1}^N \frac{\beta(M)}{\sigma_v^2} \exp\left[-\frac{\beta(M)}{\sigma_v^2} \left(E_n + \frac{\text{Tr}(\boldsymbol{\Sigma}_s) |h_n|^2}{\beta(M)}\right)\right] \\ &\times \left(\frac{\beta(M) E_n}{\text{Tr}(\boldsymbol{\Sigma}_s) |h_n|^2}\right)^{\frac{M-1}{2}} I_{M-1}\left(\frac{2}{\sigma_v^2} \sqrt{\beta(M) \text{Tr}(\boldsymbol{\Sigma}_s) |h_n|^2 E_n}\right), \end{aligned} \quad (12)$$

where  $I_{M-1}$  is the modified Bessel function of the first kind and order  $M-1$ .

*Remark 2:* It is clear that under  $\mathcal{H}_0$ , the components of  $\mathbf{E}$  are independent and identically distributed. In addition, it is important to observe that, having  $M$  finite but large, allows us to approximately consider that  $\{E_n\}$  are independent also under  $\mathcal{H}_1$ . However, (12) is different to the product of its marginals, given by (6). This shows that the result of the approximation when  $M$  is finite but large is not trivial.

*Remark 3:* In applications for which  $M$  is large and (8) leads to a tight CF approximation (10), we can use (12) as

a substitute of the true (unknown and intractable) joint PDF of  $\mathbf{E}$  conditioned to  $\mathbf{h}$  and  $\mathcal{H}_1$ , for building an approximate LRT for the detection problem at hand. This has a twofold advantage. In the first place, although not exact, the approximate likelihood can be easily computed in closed form and we do not have to resort to any series approximation, which actually are not readily available for the considered general scenario. Secondly, as the approximation naturally leads to a factorized PDF, it is well suited for distributed detection scenarios. This is because each node is able to compute its local statistic without cooperation with other nodes, while the final test statistic is the network-wide average of those local statistics, which implies relatively low communication (energy and bandwidth) resources. Notice that likelihood ratio tests using non-product PDF will typically demand many communications resources and are not suitable for distributed scenarios. This will be the case even when, through the use of the multivariate central limit theorem, and selecting  $\beta(M) = \frac{1}{M}$ , the energy measurements are approximated by a multivariate Gaussian PDF, as it is done in several works [11], [26], [27], [34]. In those cases, the likelihood ratio will be a quadratic form depending on the precision matrix (the inverse of the covariance matrix), which will be very costly to compute in distributed scenarios [35].

### C. APPROXIMATION OF THE JOINT DENSITY OF $\mathbf{E}$ IN THE FAST FADING SCENARIO

In this case, we consider that the channel gains  $h_n$  are random variables. These random values remain fixed during a certain measurement window, but in the next window, they change again according to the same distribution. Contrary to the slow-fading case, here the channel gains are different for each window, and this has to be taken into account. Using the independence assumption of  $\mathbf{E}_l$  across the  $L$  measurements windows and that each  $\mathbf{E}_l$  conditioned on  $\mathbf{h}_l$  is independent of the vector channel gains at the other measurements windows, the approximate joint PDF of the  $L$  vector energy measurements  $\mathbf{E}_1, \mathbf{E}_2, \dots, \mathbf{E}_L$  conditioned on those channel gains can be expressed as:

$$\hat{p}^1(\mathbf{E}_1, \mathbf{E}_2, \dots, \mathbf{E}_L | \mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_L) = \prod_{l=1}^L \hat{p}^1(\mathbf{E}_l | \mathbf{h}_l), \quad (13)$$

where each  $\hat{p}^1(\mathbf{E}_l | \mathbf{h}_l)$  for  $l \in [1 : L]$  is given by (12). Clearly, assuming that the  $L$  channel gains are known is not a reasonable hypothesis. Assuming that they are unknown, and to devise a test like the GLRT is possible. However, its performance will not be good, as the number of parameters to estimate is  $NL$ , which is equal to the number of total energy measurements. Another possibility is to average (13) over  $\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_L$  and define<sup>4</sup>:

$$\hat{p}^1(\mathbf{E}_1, \mathbf{E}_2, \dots, \mathbf{E}_L) \equiv \mathbb{E} \left[ \prod_{l=1}^L \hat{p}^1(\mathbf{E}_l | \mathbf{h}_l) \right], \quad (14)$$

4. This averaging is only needed for the energy measurements PDF under  $\mathcal{H}_1$ , due to the resulting PDF under  $\mathcal{H}_0$  does not depend on the channel gains.

where expectation is with respect to  $p(\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_L)$ , the joint PDF of the channel vector gains. Assuming a high-mobility scenario, it is reasonable to consider a fast-fading model [36], where  $\mathbf{h}_l$ ,  $l \in [1 : L]$  are assumed to be independent and identically distributed random vectors. This allows us to write:

$$\hat{p}^1(\mathbf{E}_1, \mathbf{E}_2, \dots, \mathbf{E}_L) = \prod_{l=1}^L \mathbb{E}[\hat{p}^1(\mathbf{E}_l|\mathbf{h}_l)], \quad (15)$$

Then, we only need to analyze  $\mathbb{E}[\hat{p}^1(\mathbf{E}_l|\mathbf{h}_l)]$  for an arbitrary  $l \in [1 : L]$ . We will also assume that nodes are separated well enough. This means that the gains  $h_n$  (in the following, as we will be again analyzing an arbitrary measurement window, we will drop the sub-index  $l$ ) with  $n \in [1 : N]$  can be modeled as independent random variables. Under this assumption:

$$\mathbb{E}[\hat{p}^1(\mathbf{E}|\mathbf{h})] = \int \hat{p}^1(\mathbf{E}|\mathbf{h}) \prod_{n=1}^N p_n(h_n) dh_1 dh_2 \dots dh_N, \quad (16)$$

where  $p_n(h_n)$  is the PDF of the channel gain  $h_n$ . Several channel distributions of interest can be considered, such as Rayleigh and Nakagami [37]. In the following, we will consider that the channel gains are Rayleigh distributed (non-line-of-sight propagation scenario). Then, the squared gains  $|h_n|^2$  are exponentially distributed according to:

$$p_n(|h_n|^2) = \frac{1}{\sigma_n^2} \exp\left(-\frac{|h_n|^2}{\sigma_n^2}\right), \quad n \in [1 : N], \quad (17)$$

where  $\mathbb{E}[|h_n|^2] = \sigma_n^2 > 0$ . It is not easy to perform direct integration in (16). However, in the next lemma, using the CF in (10), we will be able to obtain  $\mathbb{E}[\hat{p}^1(\mathbf{E}|\mathbf{h})]$ . The proof is relegated to Appendix B.

*Lemma 2:* Assuming that  $p_n(|h_n|^2)$  for  $n \in [1 : N]$  is given by (17), and  $\hat{p}^1(\mathbf{E}|\mathbf{h})$  is given by (12), we have that (16) can be written as:

$$\begin{aligned} \mathbb{E}[\hat{p}^1(\mathbf{E}|\mathbf{h})] &= \prod_{n=1}^N \frac{\beta(M)}{\Gamma(M-1)} \frac{(\sigma_v^2 + \text{Tr}(\mathbf{\Sigma}_s)\sigma_n^2)^{M-2}}{(\text{Tr}(\mathbf{\Sigma}_s)\sigma_n^2)^{M-1}} \\ &\times \exp\left(-\frac{\beta(M)E_n}{\sigma_v^2 + \text{Tr}(\mathbf{\Sigma}_s)\sigma_n^2}\right) \gamma\left(M-1, \frac{\beta(M)\text{Tr}(\mathbf{\Sigma}_s)\sigma_n^2}{\sigma_v^2 + \text{Tr}(\mathbf{\Sigma}_s)\sigma_n^2} E_n\right), \end{aligned} \quad (18)$$

where  $\gamma(\alpha, x) \equiv \int_0^x t^{\alpha-1} e^{-t} dt$ ,  $\alpha, x \geq 0$  is the incomplete Gamma function.

The result of Lemma 2, jointly with (15), allows us to express an approximation to the true joint PDF for the full  $L$  measurements windows under  $\mathcal{H}_1$  in the fast fading scenario. In this way, we can use this approximation to implement a LRT as in (3) also for this scenario, in addition to the slow-fading one. The only issue that remains to be analyzed is the case of the unknown parameters in the derived PDF approximations in a practical application. In the slow fading scenario these parameters are mainly the fixed channel gains  $\mathbf{h}$  across the measurement windows. Meanwhile, in the fast fading scenario, the unknown parameters will be the values

of  $\sigma_n^2$ ,  $n \in [1 : N]$  which are also fixed for each measurement window. This will be studied in Section V.

#### IV. THEORETICAL ANALYSIS OF THE APPROXIMATIONS

The approximations (12) and (18) to the joint PDF of the energies sensed at the sensor sites during a measurement window seems to be satisfying from the practical point of view: they are closed-form expressions and are built on the product of PDFs, each of which can be computed locally at each sensor site. However, it is important to have some theoretical guarantees for these approximations. This is important on its own, but it is also motivated by the fact that no closed-form or even infinite series expressions are available for the exact joint PDF of  $\mathbf{E}$  under  $\mathcal{H}_1$ . Notice that even in the case where an infinite series expression is available for the mentioned joint PDF, as for a particular case of the considered problem [23], the fact that  $\mathbf{E}$  is an  $N$ -dimensional vector, where the amount of sensor nodes  $N$  is typically large, poses numerical issues for its computation. Moreover, understanding how good these approximations are for a given value of  $M$  permits to choose a proper energy scaling  $\beta(M)$ . Our main interest will be to obtain a meaningful bound for  $\sup_{\mathbf{E} \in \mathbb{R}_{\geq 0}^N} |p^1(\mathbf{E}|\mathbf{h}) - \hat{p}^1(\mathbf{E}|\mathbf{h})|$  as a function of  $M$ . This is loosely connected to the idea of a local limit result [38] for the true density  $p^1(\mathbf{E}|\mathbf{h})$  when  $M \rightarrow \infty$ . However, some differences are worth noting. In the first place, we are not interested in the limit of  $p^1(\mathbf{E}|\mathbf{h})$  or  $\hat{p}^1(\mathbf{E}|\mathbf{h})$  when  $M \rightarrow \infty$  (which could not even be well defined). We are only interested in computing how well these two expressions match when  $M$  is large. That is, we want some estimate on the rate at which these two expressions get closer. In the second place, under  $\mathcal{H}_1$ , and for each  $n \in [1 : N]$ ,  $E_n \equiv \frac{1}{\beta(M)} \|\mathbf{y}_n\|^2$  is not the sum of independent random variables as the components of each  $\mathbf{y}_n$  are dependent. Most results available in the literature about local limit theorems, or even limit theorems for distribution functions (like the celebrated Berry-Essen Theorem [39]) are restricted to the independent and identically distributed case. It should be clear that obtaining meaningful bounds on  $\sup_{\mathbf{E} \in \mathbb{R}_{\geq 0}^N} |p^1(\mathbf{E}|\mathbf{h}) - \hat{p}^1(\mathbf{E}|\mathbf{h})|$  could be more difficult or would require stronger technical conditions that getting similar results but for the distribution functions. However, optimal LRTs depend on the PDFs under both hypotheses, and not on the distribution functions. For this reason, we will study the term  $\sup_{\mathbf{E} \in \mathbb{R}_{\geq 0}^N} |p^1(\mathbf{E}|\mathbf{h}) - \hat{p}^1(\mathbf{E}|\mathbf{h})|$ . In our case, we will exploit the structure of the CFs given by (5) and (10).

From the fact that a PDF can be written as the  $N$ -dimensional inverse Fourier transform of its CF, we can easily get for all  $\mathbf{E} \in \mathbb{R}_{\geq 0}^N$ :

$$|p^1(\mathbf{E}|\mathbf{h}) - \hat{p}^1(\mathbf{E}|\mathbf{h})| \leq \frac{1}{(2\pi)^N} \int_{\mathbb{R}^N} |\Psi^1(\boldsymbol{\omega}|\mathbf{h}) - \hat{\Psi}^1(\boldsymbol{\omega}|\mathbf{h})| d\boldsymbol{\omega}, \quad (19)$$

from which we conclude that the  $L_1$  distance between the CFs  $\Psi^1(\boldsymbol{\omega}|\mathbf{h})$  and  $\hat{\Psi}^1(\boldsymbol{\omega}|\mathbf{h})$  is an upper bound for

$\sup_{\mathbf{E} \in \mathbb{R}_{\geq 0}^N} |p^1(\mathbf{E}|\mathbf{h}) - \hat{p}^1(\mathbf{E}|\mathbf{h})|$ . Let us define:

$$z_m(\boldsymbol{\omega}) \equiv j \frac{\lambda_m}{\beta(M)} \sum_{n=1}^N \frac{|h_n|^2 \omega_n}{1 - j \frac{\sigma_v^2 \omega_n}{\beta(M)}}, \quad m \in [1 : M]. \quad (20)$$

Using (5) and (10), we can write:

$$\begin{aligned} \left| \Psi^1(\boldsymbol{\omega}|\mathbf{h}) - \hat{\Psi}^1(\boldsymbol{\omega}|\mathbf{h}) \right| &\leq \prod_{n=1}^N \left| 1 - j \frac{\sigma_v^2 \omega_n}{\beta(M)} \right|^{-M} \\ &\quad \times \left| \prod_{m=1}^M \frac{1}{1 - z_m(\boldsymbol{\omega})} - \prod_{m=1}^M e^{z_m(\boldsymbol{\omega})} \right|. \end{aligned} \quad (21)$$

The following lemma is a consequence of the fact that  $\Re(z_m(\boldsymbol{\omega})) \leq 0$  for all  $\boldsymbol{\omega} \in \mathbb{R}^N$  and  $m \in [1 : M]$ , and it is proved in Appendix C:

*Lemma 3:* Let  $z_m(\boldsymbol{\omega})$ ,  $m \in [1 : M]$  defined in (20). Then for all  $\boldsymbol{\omega} \in \mathbb{R}^N$ :

$$\begin{aligned} \left| \prod_{m=1}^M \frac{1}{1 - z_m(\boldsymbol{\omega})} - \prod_{m=1}^M e^{z_m(\boldsymbol{\omega})} \right| \\ \leq \sum_{m=1}^M \min \left\{ 2, \left| 1 - z_m(\boldsymbol{\omega}) - e^{-z_m(\boldsymbol{\omega})} \right| \right\} \end{aligned} \quad (22)$$

The following lemma is important because it allows us to quantify the error on the first-order Taylor expansion of  $e^{-z}$  given by  $1 - z$  when  $\Re(z) \leq 0$ . We want to emphasize that as we are considering a Taylor expansion of an analytic complex valued function special care has to be taken to estimate the remainder of the expansion. The fact that the value  $z$  where we are computing the approximation is in a specific region of the complex plane allows us to obtain better estimates of the remainder. The proof can be found in Appendix D.

*Lemma 4:* Consider  $z \in \mathbb{C}$  such that  $\Re(z) \leq 0$ . Then:

$$|1 - z - e^{-z}| \leq |z|^2 e^{-\Re(z)} \quad (23)$$

At this point we can combine the results from Lemma 3 and 4 to get:

$$\begin{aligned} \left| \prod_{m=1}^M \frac{1}{1 - z_m(\boldsymbol{\omega})} - \prod_{m=1}^M e^{z_m(\boldsymbol{\omega})} \right| \\ \leq \sum_{m=1}^M \min \left\{ 2, |z_m(\boldsymbol{\omega})|^2 e^{-\Re(z_m(\boldsymbol{\omega}))} \right\}. \end{aligned} \quad (24)$$

It is easy to show that  $\max_{m \in [1 : M]} (-\inf_{\boldsymbol{\omega} \in \mathbb{R}^N} \Re(z_m(\boldsymbol{\omega}))) = \frac{\lambda_{\max}}{\sigma_v^2} \sum_{n=1}^N |h_n|^2$  where  $\lambda_{\max}$  is the maximum eigenvalue of  $\boldsymbol{\Sigma}_s$ . Then, from (24) we can get:

$$\begin{aligned} \left| \prod_{m=1}^M \frac{1}{1 - z_m(\boldsymbol{\omega})} - \prod_{m=1}^M e^{z_m(\boldsymbol{\omega})} \right| \\ \leq \min \left\{ 2M, e^{\frac{\lambda_{\max}}{\sigma_v^2} \sum_{n=1}^N |h_n|^2} \sum_{m=1}^M |z_m(\boldsymbol{\omega})|^2 \right\}. \end{aligned} \quad (25)$$

Using (20), the term  $\sum_{m=1}^M |z_m(\boldsymbol{\omega})|^2$  can be bounded as:

$$\begin{aligned} \sum_{m=1}^M |z_m(\boldsymbol{\omega})|^2 &= \sum_{m=1}^M \left| j \frac{\lambda_m}{\beta(M)} \sum_{n=1}^N \frac{|h_n|^2 \omega_n}{1 - j \frac{\sigma_v^2 \omega_n}{\beta(M)}} \right|^2 \\ &\leq \frac{\delta(\mathbf{h})M}{\beta^2(M)} \sum_{p=1}^N \sum_{r=1}^N \frac{|\omega_p| |\omega_r|}{\sqrt{1 + \frac{\sigma_v^4 \omega_p^2}{\beta^2(M)}} \sqrt{1 + \frac{\sigma_v^4 \omega_r^2}{\beta^2(M)}}}, \end{aligned}$$

where  $\delta(\mathbf{h}) \equiv \lambda_{\max}^2 \max_{n \in [1 : N]} |h_n|^4$ . In this way we can bound (25) as:

$$\begin{aligned} \left| \prod_{m=1}^M \frac{1}{1 - z_m(\boldsymbol{\omega})} - \prod_{m=1}^M e^{z_m(\boldsymbol{\omega})} \right| \\ \leq \min \left\{ 2M, \frac{\delta'(\mathbf{h})M}{\beta^2(M)} \sum_{p=1}^N \sum_{r=1}^N \frac{|\omega_p| |\omega_r|}{\sqrt{1 + \frac{\sigma_v^4 \omega_p^2}{\beta^2(M)}} \sqrt{1 + \frac{\sigma_v^4 \omega_r^2}{\beta^2(M)}}} \right\}, \end{aligned} \quad (26)$$

where  $\delta'(\mathbf{h}) \equiv \delta(\mathbf{h}) e^{\frac{\lambda_{\max}}{\sigma_v^2} \sum_{n=1}^N |h_n|^2}$ . At this point we can proceed integrating (21) using the bound in (26). The next theorem, whose proof is in Appendix E, follows:

*Theorem 1:* The maximal error between  $p^1(\mathbf{E}|\mathbf{h})$  and  $\hat{p}^1(\mathbf{E}|\mathbf{h})$  can be bounded as:

$$\begin{aligned} \sup_{\mathbf{E} \in \mathbb{R}_{\geq 0}^N} |p^1(\mathbf{E}|\mathbf{h}) - \hat{p}^1(\mathbf{E}|\mathbf{h})| \\ = \mathcal{O} \left( \min \left\{ \frac{\beta^N(M)}{M^{\frac{N}{2}-1}}, \delta'(\mathbf{h}) \frac{\beta^N(M)}{M^{\frac{N}{2}}} \right\} \right). \end{aligned} \quad (27)$$

From the above theorem we can obtain the following conclusions for the two scenarios we are considering:

#### 1) SLOW FADING SCENARIO

When  $M$  sufficiently large the second term in the RHS in (27) is the tighter term. We see that for  $\sup_{\mathbf{E} \in \mathbb{R}_{\geq 0}^N} |p^1(\mathbf{E}|\mathbf{h}) - \hat{p}^1(\mathbf{E}|\mathbf{h})| \rightarrow 0$  when  $M \rightarrow \infty$  for every value of  $\mathbf{h}$ , we need  $\frac{\beta(M)}{M^{\frac{1}{2}}} \rightarrow 0$ . This means that  $\beta(M) = M^{1/2-\epsilon}$  for  $0 < \epsilon < 1/2$  would be a valid choice. With this choice we get:

$$\sup_{\mathbf{E} \in \mathbb{R}_{\geq 0}^N} |p^1(\mathbf{E}|\mathbf{h}) - \hat{p}^1(\mathbf{E}|\mathbf{h})| = \delta'(\mathbf{h}) \mathcal{O} \left( \frac{1}{M^{N\epsilon}} \right) \quad (28)$$

which means that the approximation is exponentially fast in  $N$ , the size of the network.

#### 2) FAST FADING SCENARIO

In this case we should proceed with more care. We need to consider a bound on  $\sup_{\mathbf{E} \in \mathbb{R}_{\geq 0}^N} |\mathbb{E}[p^1(\mathbf{E}|\mathbf{h})] - \mathbb{E}[\hat{p}^1(\mathbf{E}|\mathbf{h})]|$ . We can write:

$$\begin{aligned} \sup_{\mathbf{E} \in \mathbb{R}_{\geq 0}^N} |\mathbb{E}[p^1(\mathbf{E}|\mathbf{h})] - \mathbb{E}[\hat{p}^1(\mathbf{E}|\mathbf{h})]| \\ \leq \mathbb{E} \left[ \sup_{\mathbf{E} \in \mathbb{R}_{\geq 0}^N} |p^1(\mathbf{E}|\mathbf{h}) - \hat{p}^1(\mathbf{E}|\mathbf{h})| \right]. \end{aligned} \quad (29)$$



At this point we can use (27). As the second term in the RHS of (27) includes the term  $e^{\frac{\lambda_{\max}}{\sigma_v^2} \sum_{n=1}^N |h_n|^2}$ , the simplest way to guarantee a non-trivial bound, independent of the channel gains distribution, is to use the first term in the RHS of (27). We obtain:

$$\sup_{\mathbf{E} \in \mathbb{R}_{\geq 0}^N} |\mathbb{E}[p^1(\mathbf{E}|\mathbf{h})] - \mathbb{E}[\hat{p}^1(\mathbf{E}|\mathbf{h})]| = \mathcal{O}\left(\frac{\beta^N(M)}{M^{\frac{N}{2}-1}}\right). \quad (30)$$

Clearly choosing  $\beta(M) = M^{1/2-\epsilon}$  with  $\frac{1}{N} < \epsilon < 1/2 - \frac{1}{N}$  we have:

$$\sup_{\mathbf{E} \in \mathbb{R}_{\geq 0}^N} |\mathbb{E}[p^1(\mathbf{E}|\mathbf{h})] - \mathbb{E}[\hat{p}^1(\mathbf{E}|\mathbf{h})]| = \mathcal{O}\left(\frac{1}{M^{N(\epsilon-\frac{1}{N})}}\right). \quad (31)$$

*Remark 4:* We see that in the fast fading scenario there is a small penalty in the possible rate of convergence of the maximal error of the approximation. This penalty decrease with the network size. It is worth to mention that this penalty appears because of the way we use (27) in (29). A more thoughtful use of (27) or even a direct treatment of  $\sup_{\mathbf{E} \in \mathbb{R}_{\geq 0}^N} |\mathbb{E}[p^1(\mathbf{E}|\mathbf{h})] - \mathbb{E}[\hat{p}^1(\mathbf{E}|\mathbf{h})]|$  using similar arguments to those in Appendix E will surely lead to a better scaling and even one without any penalty with respect to the slow fading case. However, the result obtained is already sufficient for us and an improvement of this result is out of scope for the present paper.

## V. PROPOSED TEST STATISTICS

In this section, we define two test statistics for the slow and fast fading scenarios, using the approximations to the joint PDF presented in Sections III-B and III-C, respectively. Those expressions depend on parameters that are typically unknown for the setup at hand, as for example, the power of the transmitting source, the channel gains in the slow fading case and the channel variance in the fast fading case.<sup>5</sup> This fact prevents the implementation of the LRT. In addition, the uniformly most powerful test (UMPT) [10] typically does not exist for problems with multidimensional unknown parameters as in the case of this work. Therefore, we resort to the GLRT [10], where the unknown parameters are estimated using the maximum likelihood estimator. We will consider the special case in which  $\Sigma_{s,l} \equiv \Sigma_s$  for every  $l = 1, \dots, L$ . That is, during the  $L$  measurement windows the statistical properties of the source signal are invariant.

### A. APPROXIMATE GLRT FOR THE SLOW FADING SCENARIO

In the slow fading case, we assume that the channel gains between the source position and the sensors are invariant during the  $L$  measurement windows. Therefore, we have that

5. The value of the noise variance  $\sigma_v^2$  is also usually not known. However, it can be easily estimated under source silence periods to obtain a constant false alarm rate test. For this reason, it will be assumed to be known.

$h_{n,l} = h_n$  for every  $n = 1, \dots, N$  and  $l = 1, \dots, L$ . We define  $c_n \equiv \text{Tr}(\Sigma_s)|h_n|^2$  for  $n = 1, \dots, N$  leading to the test:

$$\tilde{T}_{\text{GLRT-SF}}(\mathbf{E}_1, \dots, \mathbf{E}_L) = \sum_{n=1}^N \sum_{l=1}^L \log \left( \frac{\tilde{p}_{\text{SF}}^1(E_{n,l}; \hat{c}_n)}{p^0(E_{n,l})} \right) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\geq}} \tau, \quad (32)$$

where, from (12), we define:

$$\tilde{p}_{\text{SF}}^1(E_{n,l}; c_n) \equiv \frac{\beta(M)}{\sigma_v^2} \exp \left[ -\frac{\beta(M)}{\sigma_v^2} \left( E_n + \frac{c_n}{\beta(M)} \right) \right] \times \left( \frac{\beta(M)E_n}{c_n} \right)^{\frac{M-1}{2}} I_{M-1} \left( \frac{2}{\sigma_v^2} \sqrt{\beta(M)c_n E_n} \right) \quad (33)$$

and  $\hat{c}_n \equiv \arg \max_{c_n \geq 0} \sum_{l=1}^L \log \tilde{p}_{\text{SF}}^1(E_{n,l}; c_n)$ . We used the fact that the unknown parameters are positive and only affect the energy measurements under  $\mathcal{H}_1$ . The PDF  $p^0(E_{n,l})$  is given by:

$$p^0(E_{n,l}) \equiv \left( \frac{\beta(M)}{\sigma_v^2} \right)^M \frac{E_n^{M-1}}{(M-1)!} \exp \left( -\frac{\beta(M)E_n}{\sigma_v^2} \right). \quad (34)$$

The threshold  $\tau$  is given by the desired probability of false alarm  $P_{\text{fa}}$  depending only on the distribution of the test under  $\mathcal{H}_0$  (and independent of the unknown parameters under  $\mathcal{H}_1$ ) and its computed by Monte Carlo simulations using (34).

### B. APPROXIMATE GLRT FOR THE FAST FADING SCENARIO

In the fast fading case, we assume that the channel statistics between the source position and the sensors are invariant during the  $L$  measurement windows. Therefore, we have that  $\sigma_n^2$  is constant during the whole sensing time. We define  $d_n \equiv \text{Tr}(\Sigma_s)\sigma_n^2$  for  $n = 1, \dots, N$  leading to the test:

$$\tilde{T}_{\text{GLRT-FF}}(\mathbf{E}_1, \dots, \mathbf{E}_L) = \sum_{n=1}^N \sum_{l=1}^L \log \left( \frac{\tilde{p}_{\text{FF}}^1(E_{n,l}; \hat{d}_n)}{p^0(E_{n,l})} \right) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\geq}} \tau, \quad (35)$$

where, from (18), we define

$$\tilde{p}_{\text{FF}}^1(E_{n,l}; d_n) \equiv \frac{\beta(M)}{\Gamma(M-1)} \frac{(\sigma_v^2 + d_n)^{M-2}}{d_n^{M-1}} \times \exp \left( -\frac{\beta(M)E_n}{\sigma_v^2 + d_n} \right) \gamma \left( M-1, \frac{\beta(M)d_n}{\sigma_v^2(\sigma_v^2 + d_n)} E_n \right)$$

and  $\hat{d}_n \equiv \arg \max_{d_n \geq 0} \sum_{l=1}^L \log \tilde{p}_{\text{FF}}^1(E_{n,l}; d_n)$ . As in the previous case, we used the fact that the unknown parameters are positive and only affect the energy measurements under  $\mathcal{H}_1$ . The PDF  $p^0(E_{n,l})$  is defined in (34). Again, the threshold  $\tau$  is determined numerically by the desired  $P_{\text{fa}}$  using (34).

*Remark 5:* Both statistics  $\tilde{T}_{\text{GLRT-FF}}(\mathbf{E}_1, \dots, \mathbf{E}_L)$  and  $\tilde{T}_{\text{GLRT-SF}}(\mathbf{E}_1, \dots, \mathbf{E}_L)$  need to be implemented by the network in order to make a decision about the presence or not of the source signal. In any of the two network architectures (with or without FC) shown in Fig. 1, the sensor nodes need reporting communication channels for

**TABLE 1.** Selected parameters for the simulation setup.

$K$ (dB)	$\alpha$	$d_0$ (m)	$\sigma_\eta$	$W$ (MHz)	$T$ ( $\mu$ s)	$M$	$N_0$ (dBm/Hz)	$N$	$L$
-37	4	10	2	5	2	10	-174	100	20

building the statistics. In both cases, the sensor nodes compute the corresponding inner sum over index  $l$  in (32) and (35). Then, in the case of a network with a FC, the nodes communicate those quantities to the FC, where the outer sum over  $n$  in (32) and (35) is computed. In the case of a network without a FC, an average consensus algorithm (e.g., see [11], [40]) can be used to compute the outer sum cooperatively among the nodes via messages exchanges between the nodes. Once the statistics are computed, the decision is made. In this work, given that our main focus is on how to obtain a closed-form approximation of the joint density, and its use to compute the Neyman-Person test, we assume error-free communication channels<sup>6</sup> in each of the network architectures. Channel impairments, as considered for example in [18], [41], [42], [43], are left for future works. More importantly, the approximate factorization of the joint PDF proposed in (12) and (18) allows us to also factorize the LRT in (3) through the spatial index  $n$ , and efficiently compute the statistics in a distributed scenarios (with or without FC), as explained before, using communication resources only for computing the outer sum over  $n$ .

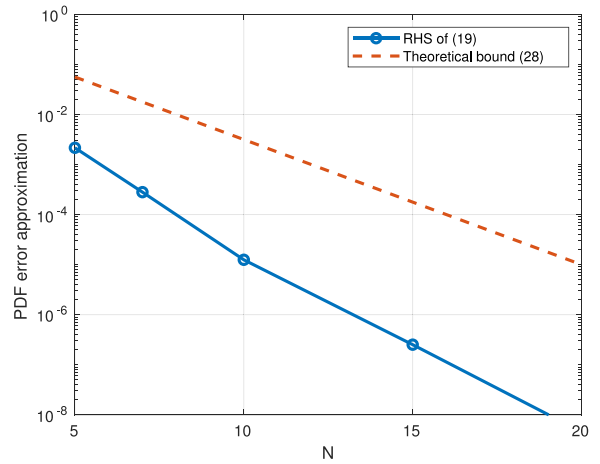
## VI. NUMERICAL RESULTS

In this section, we first assess the theoretical bounds presented in Section IV, and then compute the performance of the algorithms GLRT-SF and GLRT-FF proposed in the previous section.

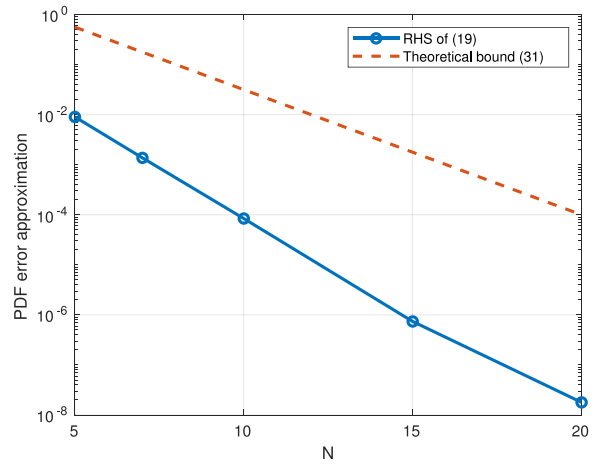
### A. THEORETICAL BOUNDS ASSESSMENT

Next, we numerically evaluate the theoretical bounds for the error approximation of the joint PDF of the energy measurements given in (28) and (31). Given that the true joint PDF of the energy measurements are unknown for the two considered scenarios (fast and slow fading), we numerically compute the bound in (19) based on the true characteristic functions of the corresponding joint PDFs. The right hand side (RHS) of (19), which involves a multidimensional integral, is computed via Monte Carlo simulations with  $10^7$  trials. On the other hand, as the theoretical bounds (28) and (31) are expressed in terms of the big-O notation, they do not provide the positive proportional constant of the bounds. Therefore, we are here interested in comparing the slopes of the curves. The bounds are plotted in Fig. 2 in semi-log graphs. The scenario parameters are selected as in the following section for the Gaussian source, with SNR = 10 dB and  $\epsilon = \frac{1}{4}$  (see also Table 1 for other relevant parameters). These curves verify that the error approximation of the joint PDF of the energy measurements decays exponentially with the number of nodes  $N$ . The

6. We also do not consider the quantization procedure needed in digital communication schemes.



(a) Slow fading.



(b) Fast fading.

**FIGURE 2.** Evaluation of the joint PDF approximation error bounds for slow and fast fading.

slightly difference in the slopes in each scenario can be explained in terms of the bounding techniques used in the derivation of the bounds, which are not tight. However, it is clear that the approximation proposed is very good for reasonable values of  $M$  and  $N$ .

### B. PERFORMANCE ANALYSIS OF GLRT-SF AND GLRT-FF

We consider the problem presented in Section II for both slow fading and fast fading scenarios. This problem can be motivated by a cooperative spectrum sensing application, where a set of secondary nodes senses a frequency band reserved to a primary user. The main idea is that, if the primary user is not using the band, then the secondary users could use it. As the primary user has priority in using this band, it is very important to have a robust detection procedure to avoid interference from the secondary user to the primary one. The slow and fast fading scenarios can model situations in which secondary nodes and/or

the primary user are stationary or moving terminals. The numerical performance of each algorithm is computed using  $10^4$  Monte Carlo runs. The data is generated following the model in (2). The channel variance  $\sigma_n^2$ , which is equal to the quotient between the received power (without considering the noise) and the transmitted power, is modeled using the path-loss/log-normal shadowing model [44],

$$\sigma_n^2(\text{dB}) = K - 10\alpha \log_{10}(d_n/d_0) - \eta_n, \quad (36)$$

where  $d_n$  is the distance between the source position and the  $n$ -th node position,  $K$  (in dB) is the path-loss attenuation at a certain distance  $d_0$ ,  $\alpha$  is the path-loss exponent, and  $\eta_n$  is a zero-mean Gaussian random variable with variance  $\sigma_\eta^2$  which models the shadowing effect. In both slow and fast fading scenarios, the large-scale shadowing effect is assumed to be the same for all the sensing intervals, i.e., the variance (36) remains constant in the whole sensed interval. Then, for each Monte Carlo run,  $\sigma_n^2$  is sampled following (36). Thus, we obtain the average performance of the algorithms with respect to network channel variance distribution (induced by the source-node distance distribution, to be defined next, and the path-loss/shadowing model).

The channel gains are determined as follows. In the slow fading case, it is assumed that the source and the sensors are static, so the channel gains  $\{h_n\}_{n=1}^N$  (which are the same for all time-windows  $l \in [1 : L]$ ) are i.i.d. sampled from the PDF  $\mathcal{CN}(0, \sigma_n^2)$  for each Monte Carlo run. On the other hand, in the fast fading scenario, it is assumed that the source and the sensors move, experiencing different small-scale fading realizations. This is modeled by sampling the channel gains  $\{h_{n,l}\}_{n=1, l=1}^{N,L}$  i.i.d. (in both sensor index  $n$  and time-window index  $l$ ) from the PDF  $\mathcal{CN}(0, \sigma_n^2)$ ,  $n \in [1 : N]$ , for each Monte Carlo run. In both slow and fast fading, the channel amplitudes are Rayleigh distributed, a model representative of non-line-of-sight propagation scenarios.

The chosen parameters for the simulation setup are shown in Table 1. The selected propagation model parameters ( $K$ ,  $\alpha$ ,  $d_0$  and  $\sigma_\eta$ ) are typical for outdoors scenarios [44, Ch. 2]. The signal-to-noise ratio is defined by  $\text{SNR} \equiv \frac{\sqrt{MP_s}\bar{\sigma}^2}{WN_0}$ , where  $\bar{\sigma}^2$  is the average variance of the channels (36) across the nodes. Notice that the source power  $P_s = \text{Tr}(\mathbf{\Sigma}_s)/T$  will be varied to be consistent with the corresponding SNR. The covariance matrix  $\mathbf{\Sigma}_s$  is set to be a Toeplitz matrix with first row  $\frac{P_s T}{M} [1, \rho, \dots, \rho^{M-1}]$ , where  $\rho = 0.5$ .

The source is assumed to be at (0, 0) m, and the location of each node in the plane ( $x, y$ ) is determined by  $x = d_n \cos(\zeta_n)$  and  $y = d_n \sin(\zeta_n)$ . The distance between the source and the  $n$ -th node  $d_n$  is assumed to be independently uniformly distributed in the log-scale,<sup>7</sup> i.e.,  $\log_{10}(d_n/d_0) \sim \text{U}[\log_{10}(800\text{m}/d_0), \log_{10}(8000\text{m}/d_0)]$ ,

7. The nodes distribution is selected such that nodes sensing relatively high power in a given frequency band wish to cooperate for detecting the source. This fits well in spectrum sensing schemes where cooperative nodes get some revenue for lending their sensing capabilities, as in [45].

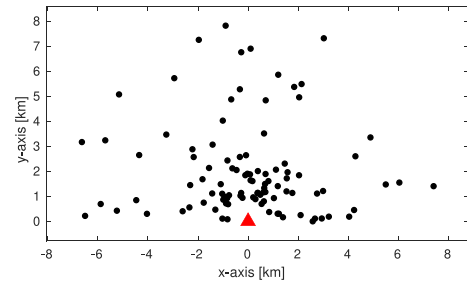


FIGURE 3. One realization of the sensor network for  $N = 100$  sensors represented with black dots. The source is depicted with a red triangle.

and  $\zeta_n$  is i.i.d. randomly selected from the uniform distribution  $\text{U}[0, \pi]$ . In Fig. 3, we show a realization of the sensor network. The computation of the MLE in (32) and (35), needed for implementing both GLRT-SF and GLRT-FF algorithms, cannot be done through a closed form formula. So, we need a numerical procedure for solving this one-dimensional nonlinear problem with a bound constraint (the positiveness of the parameters  $c_n$  and  $d_n$ ) at each node. We use here a trust-region algorithm [46] for which, depending on the scenario (fundamentally on the SNR) and the stopping criteria parameters (e.g., gradient tolerance and step tolerance), it typically takes 5-10 iterations to converge to a local maximum.

We compare the performance of the proposed algorithms against the ones shown in Table 2 (see the definitions of the acronyms in the third column). The CSI-SF and CSI-FF detectors are the version of GLRT-SF and GLRT-FF, respectively, for which the parameters of the channel (channel state information, CSI) and the source are perfectly known. These statistics, of course, are unrealizable in practice and they are included only as a reference for the comparison. Note that some performance loss of the GLRT-based algorithms is expected with respect to these *genie-aided* test statistics given the estimation errors of the unknown parameters.

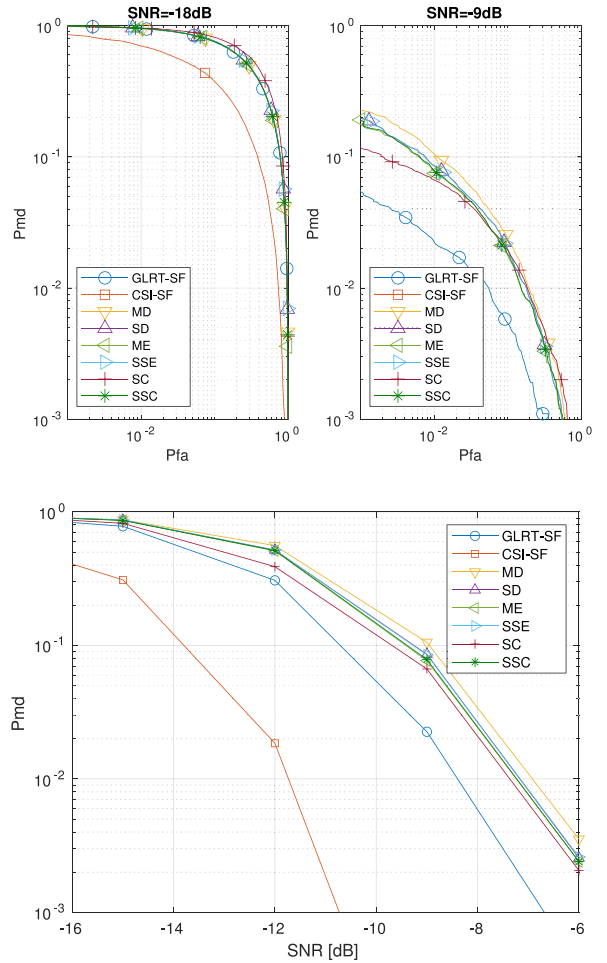
In Table 2, we also include the computational complexity of the test statistics considered in our simulations. It is straightforward to see that MD, SD, SC and SSC have complexity  $\mathcal{O}(NL)$ . Both ME and SSC require the computation of the eigenvalues of the sample covariance matrix of energy measurements at the nodes. This is computationally demanding, requiring  $\mathcal{O}(N^3)$  computations [47]. The cost of constructing the sample covariance matrix is  $\mathcal{O}(N^2L)$ . Therefore, the total computational complexity of these algorithms is  $\mathcal{O}(N^2(N + L))$ . The complexity for the genie-aided cases CSI-SF and CSI-FF in which the parameters  $c_n$  and  $d_n$  are known is  $\mathcal{O}(NL)$ . When those parameters are unknown, GLRT-SF and GLRT-FF estimate them. Then, we need to add the cost of estimating the unknown parameters  $c_n$  and  $d_n$  by a maximization procedure. It is important to note that the problem of obtaining the  $N$  optimal values of  $c_n$  or  $d_n$  can be decomposed into  $N$  scalar optimization problems, which are obviously significantly less complex

than a single  $N$ -dimensional optimization problem. This is a consequence of the factorized PDF obtained by our approximations. In our case, we used a trust-region algorithm which is based on the conjugate gradient method. This is an iterative method whose complexity depends on the accuracy required for the solution. It is known that if a tolerance of  $\xi$  is the solution required, the worst-case complexity is  $\mathcal{O}(\xi^{-\nu})$  where  $\nu \in [1, 2]$  depending on the properties of the function to be optimized [48]. In practice, the average complexity seems to be significantly lower. In our experience, for the problem considered in this paper and as explained above, the convergence to a local maximum was very fast, and the complexity of both GLRT-SF and GLRT-FF is dominated by  $\mathcal{O}(NL)$ .

The test statistic MD is the mean detector, i.e., the average of all network measurements, and is equivalent to the equal gain combining (EGC), typically used in low SNR regimes [49]. The test statistic SD is the square detector. We also include two eigenvalue-based detectors (ME and SSE) which are typically used in the present scenario. These detectors naturally consider the statistical dependence of the observations under  $\mathcal{H}_1$  at different sensor nodes, introduced by the random signal source, given that they are based on the eigenvalues  $\{\lambda_n\}_{n=1}^N$  of the sample covariance matrix of the observations  $\{\mathbf{E}_l\}_{l=1}^L$ . Nevertheless, judging by the analysis in the previous section, and the following numerical results, the statistical dependence appears to have a negligible impact in the detectors' performance when considering energy observations. Finally, we also include the SC detector [28], which selects the highest average measured energy among all the sensors, and the SSC detector [34].

In Fig. 4 and 5 (top figures), we plot the complementary receiver operating characteristic (CROC), i.e., the miss detection probability against the false alarm probability (both defined in Section II), for the algorithms in Table 2, in the slow fading and the fast fading scenario, respectively, for the indicated SNRs. Additionally, in the respective bottom figures, we plot the miss detection probability against the SNR, when  $P_{fa} = 10^{-2}$ . As expected, the genie-aided test statistics outperform the rest of the algorithms. Among the test statistics capable of being implemented in practice, GLRT-SF and GLRT-FF, achieve the best performance for each scenario. These gains seem to increase for large values of SNRs. This is important, because among the other methods there are some that explicitly exploit information about spatial correlation (as ME and SSE) and others that do not (as MD and SC). In addition, the proposed detectors demand similar communication resources to the simplest ones, and provide better performance than the ones that make use of the spatial correlation information among the nodes and also demand more network communication resources [35].

Finally, in Fig. 6, we also consider a case where the source signal is not Gaussian but a OFDM signal. We consider



**FIGURE 4.** Performance of the algorithms for slow fading channels. Top: Complementary receiver operating characteristics for SNR = -18dB (left) and SNR = -9dB (right). Notice that on the right plot, the CSI-SF curve is not seen because it falls outside below the range shown. Bottom: Miss detection probabilities vs SNR for false alarm probability  $P_{fa} = 0.01$ .

a NB-IoT and LTE-M scenarios, where typical small FFT sizes are used, such as 12 or 24, to allow for narrower subcarrier spacing and more efficient transmission of low-rate IoT data [52]. Notice that also a setting with such a small OFDM symbol length  $M$  is interesting as can be thought as a limiting case to our assumption that  $M \approx WT$  is large. In particular, in our case, the number of subcarriers is set to 12 (a resource block in Long Term Evolution cellular systems) and the cyclic prefix is defined to be 3 samples. Then, the OFDM symbol has a length of  $M = 15$ . We consider a 64-QAM as the subcarriers modulation. We can see that similar conclusions as for the Gaussian signal can be drawn, which validates the theoretical approach used in this work for this particular setup.

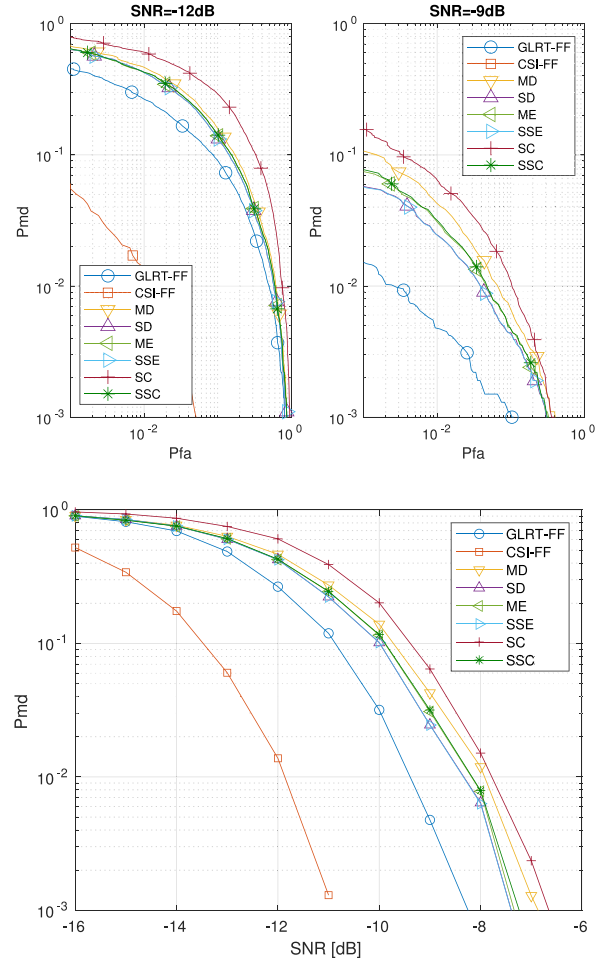
## VII. CONCLUDING REMARKS

We have considered the problem of energy-based distributed detection of a stochastic Gaussian radio source signal, for which the measurements at each sensor node present spatial

**TABLE 2.** Test statistics to be compared with GLRT-SF and GLRT-FF. † are genie-aided detectors.

Label	Test statistic/Ref.	Detector name / Observation	Complexity
GLRT-SF	Eq. (32)	GLRT-based detector for slow fading (SF)	$\mathcal{O}(NL)$
GLRT-FF	Eq. (35)	GLRT-based detector for fast fading (FF)	$\mathcal{O}(NL)$
CSI-SF†	Eq. (32) using the true value $c_n$ instead of $\hat{c}_n$ , $n \in [1, N]$ .	CSI detector for slow fading (SF).	$\mathcal{O}(NL)$
CSI-FF†	Eq. (35) using the true value $d_n$ instead of $\hat{d}_n$ , $n \in [1, N]$ .	CSI detector for fast fading (FF).	$\mathcal{O}(NL)$
MD	$1/(NL) \sum_{n=1}^N \sum_{l=1}^L E_{n,l}$	Mean detector.	$\mathcal{O}(NL)$
SD	$1/(NL) \sum_{n=1}^N \sum_{l=1}^L E_{n,l}^2$	Square detector.	$\mathcal{O}(NL)$
ME	$\lambda_{\max}$ , [50]	Maximum eigenvalue detector. $\lambda_{\max}$ is the maximum eigenvalue of the sample covariance matrix.	$\mathcal{O}(N^2(N+L))$
SSE	$\sum_{n=1}^N -\log \lambda_n^+ + \lambda_n^+$ , [51]	Subspace eigenvalue detector. $\lambda_n^+ \equiv \max(0, \lambda_n - 1)$	$\mathcal{O}(N^2(N+L))$
SC	$\max_{n \in [1, N]} m_{z,n}$ , [28]	Selection combining detector	$\mathcal{O}(NL)$
SSC	$\sum_{n=1}^N \max(m_{z,n}, 0)^2$ , [34]	Selection square combining detector	$\mathcal{O}(NL)$

and temporal correlation. Using the closed-form expression for the CF of the joint PDF of the energy measurement at each sensor site, we computed an approximation of the joint PDF (when the source signal is present), under the assumption of a large time-bandwidth product, and for both wireless scenarios of common interest: slow and fast fading. We also provided two deviation bounds for the obtained approximation respect to the true intractable joint PDF. These bounds prove to be exponential on the number of nodes in the network. The bounds indicate that the spatial and temporal correlation of the measurements at each node site are not extremely critical given that the exact joint PDF can be tightly approximated by a factorized PDF. These PDF approximations were used to implement likelihood ratio tests that show performance gains with respect to other usual schemes typically used in practice. Moreover, the product nature of the obtained PDF was shown to be valuable in implementing simpler cooperative schemes in distributed scenarios. It also can be an asset for designing simple quantization schemes relying on the local PDF at each node in order to cope, among other things, with the usual reporting channel impairments between the nodes and/or the FC. Another possible line of work could include the use of the approximation to analyze the outage probability of different receiving techniques of large MIMO systems with correlation on the transmitter and receiver antennas (see for example the results in [23] and [24]). The application of the obtained results for characterizing performance of distributed radar systems [53], can also be explored in the future.



**FIGURE 5.** Performance of the algorithms for fast fading channels. Top: Complementary receiver operating characteristics for SNR = -12dB (left) and SNR = -9dB (right). Notice that on the right plot, the CSI-FF curve is not seen because it falls outside below the range shown. Bottom: Miss detection probabilities vs SNR for false alarm probability  $P_{fa} = 0.01$ .

## APPENDIX A PROOF OF LEMMA 1

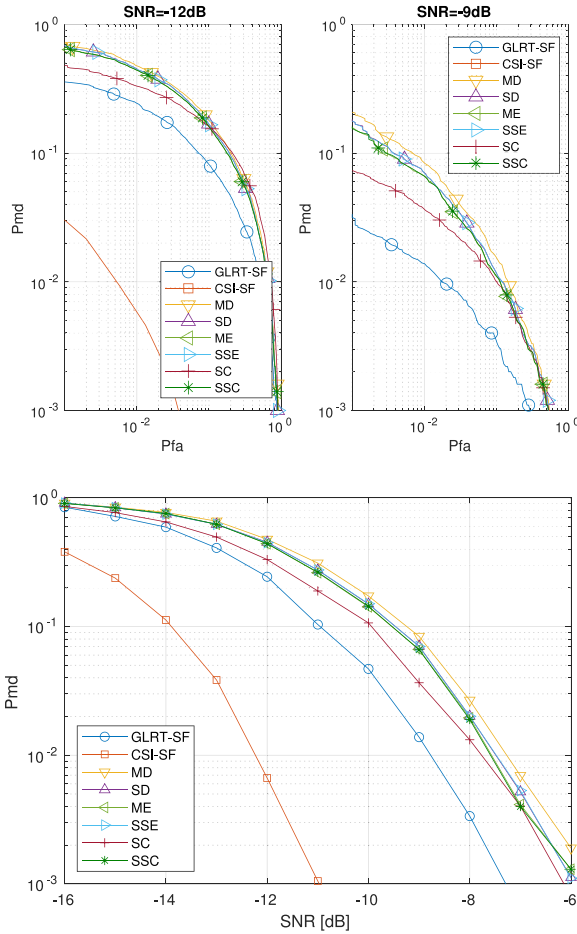
We can write (under both hypotheses  $\mathcal{H}_0$  and  $\mathcal{H}_1$ ):

$$\Psi(\omega|\mathbf{h}) = \mathbb{E} \left[ e^{j \frac{1}{\beta(M)} \sum_{n=1}^N \omega_n \|\mathbf{y}_n\|^2} \right]. \quad (37)$$

Assuming that  $\Sigma_s = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^H$  where  $\mathbf{Q}$  is unitary and  $\mathbf{\Lambda} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_M)$  with  $\lambda_m$  the  $m$ -largest positive eigenvalue of  $\Sigma_s$ , we can define  $\tilde{\mathbf{y}} \equiv (\mathbf{I}_N \otimes \mathbf{Q})\mathbf{y}$ . It is immediate to see that under  $\mathcal{H}_0$ ,  $\tilde{\mathbf{y}} \sim \mathcal{CN}(\mathbf{0}, \sigma_v^2 \mathbf{I}_{NM})$  and under  $\mathcal{H}_1$ ,  $\tilde{\mathbf{y}} \sim \mathcal{CN}(\mathbf{0}, \mathbf{h}\mathbf{h}^H \otimes \mathbf{\Lambda} + \sigma_v^2 \mathbf{I}_{NM})$ . Moreover, we have that

$$\Psi(\omega|\mathbf{h}) = \mathbb{E} \left[ e^{j \frac{1}{\beta(M)} \sum_{n=1}^N \omega_n \|\tilde{\mathbf{y}}_n\|^2} \right] = \mathbb{E} \left[ e^{j \frac{1}{\beta(M)} \tilde{\mathbf{y}}^H \mathbf{A}_\omega \tilde{\mathbf{y}}} \right], \quad (38)$$

where we have used that  $\mathbf{Q}$  preserves the norm and with  $\mathbf{A}_\omega = \text{diag}(\omega_1, \omega_2, \dots, \omega_N) \otimes \mathbf{I}_M$ . As under both hypotheses,



**FIGURE 6.** OFDM signal: Performance of the algorithms for slow fading channels when the source follows an OFDM modulation. Top: Complementary receiver operating characteristics for SNR = -12dB (left) and SNR = -9dB (right). Notice that on the right plot, the CSI-SF curve is not seen because it falls outside below the range shown. Bottom: Miss detection probabilities vs SNR for false alarm probability  $P_{fa} = 0.01$ .

$\tilde{\mathbf{y}}$  is a complex and circular Gaussian random vector, it is easy to obtain that under  $\mathcal{H}_0$ :

$$\Psi^0(\omega|\mathbf{h}) = \prod_{n=1}^N \frac{1}{\left(1 - j \frac{\omega_n \sigma_v^2}{\beta(M)}\right)^M}. \quad (39)$$

Under  $\mathcal{H}_1$  we obtain:

$$\Psi^1(\omega|\mathbf{h}) = \frac{1}{\left| \mathbf{I}_{NM} - j \frac{(\mathbf{h}\mathbf{h}^H \otimes \Lambda + \sigma_v^2 \mathbf{I}_{NM}) \mathbf{A} \omega}{\beta(M)} \right|} \quad (40)$$

Note that we can write:

$$\left( \mathbf{h}\mathbf{h}^H \otimes \Lambda \right) \mathbf{A} \omega = \left( \mathbf{h}\mathbf{h}^H \text{diag}(\omega_1, \omega_2, \dots, \omega_N) \right) \otimes \Lambda, \quad (41)$$

where we have used that  $(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = (\mathbf{AC}) \otimes (\mathbf{BD})$ . Consider now the *commutation matrix*  $\mathbf{K}_{N,M}$  [54]. This matrix is a permutation matrix of size  $NM \times NM$  with several interesting properties. The one we need is the following: for square matrices  $\mathbf{A} \in \mathbb{C}^{N \times N}$ ,  $\mathbf{B} \in \mathbb{C}^{M \times M}$ , we have  $\mathbf{A} \otimes \mathbf{B} =$

$\mathbf{K}_{N,M}(\mathbf{B} \otimes \mathbf{A})\mathbf{K}_{N,M}^T$ . Using this matrix and the result of (41) we can express  $\mathbf{I}_{NM} - j \frac{(\mathbf{h}\mathbf{h}^H \otimes \Lambda + \sigma_v^2 \mathbf{I}_{NM}) \mathbf{A} \omega}{\beta(M)}$  as:

$$\mathbf{K}_{N,M} \left( \mathbf{I}_{NM} - j \frac{\Lambda \otimes (\mathbf{h}\mathbf{h}^H \Omega) + \sigma_v^2 \mathbf{K}_{N,M}^T (\Omega \otimes \mathbf{I}_M) \mathbf{K}_{N,M}}{\beta(M)} \right) \times \mathbf{K}_{N,M}^T,$$

where we have defined  $\Omega = \text{diag}(\omega_1, \omega_2, \dots, \omega_N)$ . In first place note that, as  $\mathbf{K}_{N,M}^T = \mathbf{K}_{M,N}$ ,  $\mathbf{K}_{N,M}^T (\Omega \otimes \mathbf{I}_M) \mathbf{K}_{N,M}$  transforms into  $\mathbf{I}_M \otimes \Omega$ . In second place, the matrix  $\Lambda \otimes (\mathbf{h}\mathbf{h}^H \Omega)$  is a block-diagonal matrix with diagonal blocks given by  $\lambda_m \mathbf{h}\mathbf{h}^H \Omega$  with  $m = 1, \dots, M$ . As a consequence, and using the fact that  $|\mathbf{K}_{N,M}| |\mathbf{K}_{N,M}^T| = 1$ , (41) can be put as:

$$\Psi^1(\omega|\mathbf{h}) = \frac{1}{\prod_{m=1}^M \left| \mathbf{I}_N - j \frac{\sigma_v^2 \Omega}{\beta(M)} - j \frac{\lambda_m \mathbf{h}\mathbf{h}^H \Omega}{\beta(M)} \right|}. \quad (42)$$

Finally, using the fact that  $|\mathbf{A} + \mathbf{u}\mathbf{v}^H| = |\mathbf{A}|(1 + \mathbf{v}^H \mathbf{A}^{-1} \mathbf{u})$ , for every invertible matrix  $\mathbf{A}$  and vectors  $\mathbf{u}, \mathbf{v}$  of appropriate size, we get:

$$\Psi^1(\omega|\mathbf{h}) = \frac{\prod_{m=1}^M \left| \mathbf{I}_N - j \frac{\sigma_v^2 \Omega}{\beta(M)} \right|^{-1}}{\prod_{m=1}^M \left( 1 - j \frac{\lambda_m}{\beta(M)} \mathbf{h}^H \left( \mathbf{I}_N - j \frac{\sigma_v^2 \Omega}{\beta(M)} \right)^{-1} \mathbf{h} \right)}. \quad (43)$$

As  $\Omega$  is a diagonal matrix we can easily get the result of the Lemma.

## APPENDIX B PROOF OF LEMMA 2

As  $\hat{p}^1(\mathbf{E}|\mathbf{h})$  in (12) is product PDF and the channel gains are independent we can write:

$$\mathbb{E}[\hat{p}^1(\mathbf{E}|\mathbf{h})] = \prod_{n=1}^N \int_0^\infty \hat{p}^1(E_n | |h_n|^2) p_n(|h_n|^2) d|h_n|^2, \quad (44)$$

where

$$\hat{p}^1(E_n | |h_n|^2) = \frac{\beta(M)}{\sigma_v^2} \exp\left[-\frac{\beta(M)}{\sigma_v^2} \left(E_n + \frac{\text{Tr}(\Sigma_s) |h_n|^2}{\beta(M)}\right)\right] \times \left(\frac{\beta(M) E_n}{\text{Tr}(\Sigma_s) |h_n|^2}\right)^{\frac{M-1}{2}} I_{M-1}\left(\frac{2}{\sigma_v^2} \sqrt{\beta(M) \text{Tr}(\Sigma_s) |h_n|^2} E_n\right),$$

for  $n \in [1 : N]$ . Then, we need only to consider the computation of:

$$\mathbb{E}[\hat{p}^1(E_n | |h_n|^2)] = \int_0^\infty \hat{p}^1(E_n | |h_n|^2) p_n(|h_n|^2) d|h_n|^2 \quad (45)$$

Let us define the CF  $\tilde{\psi}_n^1(\omega_n)$  of  $\mathbb{E}[\hat{p}^1(E_n | |h_n|^2)]$  with  $n \in [1 : N]$ . Using Fubini's Theorem [55] we can write:

$$\begin{aligned} \tilde{\psi}_n^1(\omega_n) &= \int_0^\infty \mathbb{E}[\hat{p}^1(E_n | |h_n|^2)] e^{j\omega_n E_n} dE_n \\ &= \int_0^\infty \left[ \int_0^\infty \hat{p}^1(E_n | |h_n|^2) e^{j\omega_n E_n} dE_n \right] p_n(|h_n|^2) d|h_n|^2 \\ &= \mathbb{E}[\hat{\psi}^1(\omega_n | h_n)] \end{aligned} \quad (46)$$

where  $\hat{\psi}^1(\omega_n|h_n)$  was defined in (11). It is easy to show that  $\mathbb{E}[\hat{\psi}^1(\omega_n|h_n)]$  has a closed form when  $p_n(|h_n|^2)$  is as in (17):

$$\tilde{\psi}_n^1(\omega_n) = \frac{1}{\left(1 - j \frac{\sigma_v^2 \omega_n}{\beta(M)}\right)^{M-1}} \frac{1}{\left(1 - j \frac{\sigma_v^2 + \text{Tr}(\Sigma_s) \sigma_n^2}{\beta(M)} \omega_n\right)}. \quad (47)$$

Notice that  $\tilde{\psi}_n^1(\omega_n)$  is the product of two well-known CFs. In first place, the first term in the RHS of (46) is the CF of a Gamma distribution with shape  $M - 1$  and scale  $\frac{\sigma_v^2}{\beta(M)}$ . Similarly, the second term in the RHS of (46) is the CF of an exponential distribution with scale parameter  $\frac{\sigma_v^2 + \text{Tr}(\Sigma_s) \sigma_n^2}{\beta(M)}$ . Clearly the PDF  $\mathbb{E}[\hat{p}^1(E_n||h_n|^2)]$  which corresponds to the CF  $\tilde{\psi}_n^1(\omega_n)$  is the convolution of the above described pdfs. The calculation can be easily done to obtain for each  $n \in [1 : N]$ :

$$\begin{aligned} \mathbb{E}[\hat{p}^1(E_n||h_n|^2)] &= \frac{\beta(M)}{\Gamma(M-1)} \frac{(\sigma_v^2 + \text{Tr}(\Sigma_s) \sigma_n^2)^{M-2}}{(\text{Tr}(\Sigma_s) \sigma_n^2)^{M-1}} \\ &\times \exp\left(-\frac{\beta(M) E_n}{\sigma_v^2 + \text{Tr}(\Sigma_s) \sigma_n^2}\right) \gamma\left(M-1, \frac{\beta(M) \text{Tr}(\Sigma_s) \sigma_n^2}{\sigma_v^2 + \text{Tr}(\Sigma_s) \sigma_n^2} E_n\right), \end{aligned}$$

where  $\gamma(\alpha, x) \equiv \int_0^x t^{\alpha-1} e^{-t} dt$ ,  $\alpha, x \geq 0$  is the incomplete Gamma function. The final result can be easily obtained from here.

#### APPENDIX C PROOF OF LEMMA 3

As  $\Re(z_m(\omega)) \leq 0$  for all  $\omega \in \mathbb{R}^N$  and  $m \in [1 : M]$  we have:

$$\left| \frac{1}{1 - z_m(\omega)} \right| \leq 1, \quad \left| e^{z_m(\omega)} \right| \leq 1, \quad \forall \omega \in \mathbb{R}^N, \quad m \in [1 : M]. \quad (48)$$

We can write:

$$\begin{aligned} \prod_{m=1}^M \frac{1}{1 - z_m(\omega)} - \prod_{m=1}^M e^{z_m(\omega)} &= \left( \frac{1}{1 - z_M(\omega)} - e^{z_M(\omega)} \right) \\ &\times \prod_{m=1}^{M-1} e^{z_m(\omega)} + \frac{1}{1 - z_M(\omega)} \left( \prod_{m=1}^{M-1} \frac{1}{1 - z_m(\omega)} - \prod_{m=1}^{M-1} e^{z_m(\omega)} \right). \end{aligned} \quad (49)$$

Using (48) we obtain:

$$\begin{aligned} \left| \prod_{m=1}^M \frac{1}{1 - z_m(\omega)} - \prod_{m=1}^M e^{z_m(\omega)} \right| &\leq \left| \frac{1}{1 - z_M(\omega)} - e^{z_M(\omega)} \right| \\ &+ \left| \prod_{m=1}^{M-1} \frac{1}{1 - z_m(\omega)} - \prod_{m=1}^{M-1} e^{z_m(\omega)} \right|. \end{aligned} \quad (50)$$

Repeating this argument  $M$  times we get:

$$\left| \prod_{m=1}^M \frac{1}{1 - z_m(\omega)} - \prod_{m=1}^M e^{z_m(\omega)} \right| \leq \sum_{m=1}^M \left| \frac{1}{1 - z_m(\omega)} - e^{z_m(\omega)} \right| \quad (51)$$

$$\left| \frac{1}{1 - z_m(\omega)} - e^{z_m(\omega)} \right| \leq 2. \quad (52)$$

Also, using (48) we can write:

$$\begin{aligned} \left| \frac{1}{1 - z_m(\omega)} - e^{z_m(\omega)} \right| &= \left| \frac{1}{1 - z_m(\omega)} \right| \left| e^{z_m(\omega)} \right| \\ &\times \left| 1 - z_m(\omega) - e^{-z_m(\omega)} \right| \\ &\leq \left| 1 - z_m(\omega) - e^{-z_m(\omega)} \right|. \end{aligned} \quad (53)$$

Then, combining (51), (53) and using that for each  $m \in [1 : M]$  and  $\omega \in \mathbb{R}$  we have  $\left| \frac{1}{1 - z_m(\omega)} - e^{z_m(\omega)} \right| \leq 2$ , we get the desired result.

#### APPENDIX D PROOF OF LEMMA 4

Consider  $z \in \mathbb{C}$  such as  $\Re(z) \leq 0$  and  $e^{-tz}$  with  $t \in [0 : 1]$ . It is easy to show that:

$$z \int_0^1 e^{-tz} dt = 1 - e^{-z}. \quad (54)$$

At the same time we have:

$$\left| \int_0^1 e^{-tz} dt \right| \leq \int_0^1 e^{-t \Re(z)} dt \leq e^{-\Re(z)}, \quad (55)$$

where in the last inequality we used that  $\Re(z) \leq 0$ . Combining the last two equations we get  $|1 - e^{-z}| \leq |z| e^{-\Re(z)}$ . Similarly, we can write:

$$z \int_0^1 (e^{-tz} - 1) dt = 1 - z - e^{-z} \quad (56)$$

and from  $\Re(z) \leq 0$  and  $|1 - e^{-z}| \leq |z| e^{-\Re(z)}$ :

$$\int_0^1 |e^{-tz} - 1| dt \leq |z| e^{-\Re(z)}. \quad (57)$$

Finally, using (56) and (57) we get the desired result.

#### APPENDIX E PROOF OF THEOREM 1

The following bound can be trivially derived:

$$\begin{aligned} &\int_{\mathbb{R}^N} \left| \Psi^1(\omega|\mathbf{h}) - \hat{\Psi}^1(\omega|\mathbf{h}) \right| d\omega \\ &\leq \min \left\{ 2M \int_{\mathbb{R}^N} \prod_{n=1}^N \left| 1 - j \frac{\sigma_v^2 \omega_n}{\beta(M)} \right|^{-M} d\omega, \right. \\ &\quad \left. \frac{\delta'(\mathbf{h})M}{\beta^2(M)} \int_{\mathbb{R}^N} \sum_{p=1}^N \sum_{r=1}^N \frac{\prod_{n=1}^N \left| 1 - j \frac{\sigma_v^2 \omega_n}{\beta(M)} \right|^{-M} |\omega_p| |\omega_r|}{\sqrt{1 + \frac{\sigma_v^4 \omega_p^2}{\beta^2(M)}} \sqrt{1 + \frac{\sigma_v^4 \omega_r^2}{\beta^2(M)}}} d\omega \right\}. \end{aligned} \quad (58)$$

We will mainly analyze the second term in the min operator in the RHS of the above equation. The first term can be easily analyzed with the same arguments that will be given

in the following. First notice that the mentioned second term can be cast as the sum of  $N^2$  multidimensional integrals. Each one of those integrals can be written as:

$$\int_{\mathbb{R}^N} \left( \prod_{n=1}^N \frac{1}{\left(1 + \frac{\sigma_v^4 \omega_n^2}{\beta^2(M)}\right)^{M/2}} \right) \frac{|\omega_p| |\omega_r|}{\sqrt{1 + \frac{\sigma_v^4 \omega_p^2}{\beta^2(M)}} \sqrt{1 + \frac{\sigma_v^4 \omega_r^2}{\beta^2(M)}}} d\omega, \quad (59)$$

for each  $p, r \in [1 : N]$ . We need to analyze two cases:

#### 1) CASE 1: $P \neq R$

In this case (59) reduces to:

$$\left( \int_{-\infty}^{\infty} \frac{1}{\left(1 + \frac{\sigma_v^4 \omega^2}{\beta^2(M)}\right)^{M/2}} d\omega \right)^{N-2} \left( \int_{-\infty}^{\infty} \frac{|\omega|}{\left(1 + \frac{\sigma_v^4 \omega^2}{\beta^2(M)}\right)^{\frac{M+1}{2}}} d\omega \right)^2 \quad (60)$$

#### 2) CASE 1: $P = R$

In this case (59) reduces to:

$$\left( \int_{-\infty}^{\infty} \frac{1}{\left(1 + \frac{\sigma_v^4 \omega^2}{\beta^2(M)}\right)^{M/2}} d\omega \right)^{N-1} \int_{-\infty}^{\infty} \frac{\omega^2}{\left(1 + \frac{\sigma_v^4 \omega^2}{\beta^2(M)}\right)^{\frac{M+2}{2}}} d\omega. \quad (61)$$

From (60) and (61), we see that the calculation of (58) depends on the values of 3 different one-dimensional improper integrals. In first place we consider:

$$\int_{-\infty}^{\infty} \frac{|\omega|}{\left(1 + \frac{\sigma_v^4 \omega^2}{\beta^2(M)}\right)^{\frac{M+1}{2}}} d\omega = \frac{2\beta^2(M)}{\sigma_v^4(M-1)}, \quad (62)$$

where the final value is easily obtained from simple change of variables. In second place from [56] (Equation (3.251)) we can obtain<sup>8</sup>:

$$\int_{-\infty}^{\infty} \frac{1}{\left(1 + \frac{\sigma_v^4 \omega^2}{\beta^2(M)}\right)^{\frac{M}{2}}} d\omega = \frac{(M-3)!!\beta(M)\pi}{(M-2)!!\sigma_v^2}, \quad (63)$$

and

$$\int_{-\infty}^{\infty} \frac{\omega^2}{\left(1 + \frac{\sigma_v^4 \omega^2}{\beta^2(M)}\right)^{\frac{M+2}{2}}} d\omega = \frac{(M-3)!!\beta^3(M)\pi}{(M)!!\sigma_v^6}, \quad (64)$$

where  $n!!$  for  $n \in \mathbb{N}$  is the double factorial of  $n$ . It is well known that when  $n = 2l$  with  $l \in \mathbb{N}$ , then  $n!! = 2^l l!$  and when  $n = 2l - 1$ ,  $n!! = \frac{(2l)!}{2^l l!}$ . Using those facts, combining all the above results, and using Stirling approximation we

<sup>8</sup> We assume that  $M$  is even. Using complex contour integration we can obtain results when  $M$  is odd. However, as we are interesting in the results when  $M$  is large and in order to save space, we only consider this case.

can obtain the following bound for the second term in the RHS in (58):

$$\int_{\mathbb{R}^N} \sum_{p=1}^N \sum_{r=1}^N \frac{\prod_{n=1}^N \left|1 - j \frac{\sigma_v^2 \omega_n}{\beta(M)}\right|^{-M} |\omega_p| |\omega_r|}{\sqrt{1 + \frac{\sigma_v^4 \omega_p^2}{\beta^2(M)}} \sqrt{1 + \frac{\sigma_v^4 \omega_r^2}{\beta^2(M)}}} d\omega = \mathcal{O}\left(\frac{\beta^{N+2}(M)}{M^{\frac{N+2}{2}}}\right), \quad (65)$$

Finally we can get:

$$\frac{1}{(2\pi)^N} \int_{\mathbb{R}^N} \left| \Psi^1(\omega|\mathbf{h}) - \hat{\Psi}^1(\omega|\mathbf{h}) \right| d\omega = \mathcal{O}\left(\min\left\{\frac{\beta^N(M)}{M^{\frac{N}{2}-1}}, \delta'(\mathbf{h}) \frac{\beta^N(M)}{M^{\frac{N}{2}}}\right\}\right). \quad (66)$$

from which and (19) the desired result follows.

## REFERENCES

- [1] J. Gubbi, R. Buyya, S. Marusic, and M. Palaniswami, "Internet of Things (IoT): A vision, architectural elements, and future directions," *Future Gener. Comput. Syst.*, vol. 29, no. 7, pp. 1645–1660, Sep. 2013.
- [2] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, and M. Ayyash, "Internet of Things: A survey on enabling technologies, protocols, and applications," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 4, pp. 2347–2376, 4th Quart., 2015.
- [3] V. Gupta and S. De, "Collaborative multi-sensing in energy harvesting wireless sensor networks," *IEEE Trans. Signal Inf. Process. Over Netw.*, vol. 6, pp. 426–441, 2020, doi: [10.1109/TSIPN.2020.2995502](https://doi.org/10.1109/TSIPN.2020.2995502).
- [4] P. Di Lorenzo, S. Barbarossa, and S. Sardellitti, "Distributed signal processing and optimization based on in-network subspace projections," *IEEE Trans. Signal Process.*, vol. 68, pp. 2061–2076, 2020, doi: [10.1109/TSP.2020.2981204](https://doi.org/10.1109/TSP.2020.2981204).
- [5] S. P. Chepuri and G. Leus, "Sparse sensing for distributed detection," *IEEE Trans. Signal Process.*, vol. 64, no. 6, pp. 1446–1460, Mar. 2016.
- [6] D. Ciuonzo and P. S. Rossi, "Distributed detection of a non-cooperative target via generalized locally-optimum approaches," *Inf. Fus.*, vol. 36, pp. 261–274, Jul. 2017.
- [7] S. A. Aldalameh, S. O. Al-Jazzar, D. McLernon, S. A. R. Zaidi, and M. Ghogho, "Fusion rules for distributed detection in clustered wireless sensor networks with imperfect channels," *IEEE Trans. Signal Inf. Process. Netw.*, vol. 5, no. 3, pp. 585–597, Sep. 2019.
- [8] J. Lunden, V. Koivunen, and H. V. Poor, "Spectrum exploration and exploitation for cognitive radio: Recent advances," *IEEE Signal Process. Mag.*, vol. 32, no. 3, pp. 123–140, May 2015.
- [9] J. Lundén, V. Koivunen, A. Huttunen, and H. V. Poor, "Collaborative cyclostationary spectrum sensing for cognitive radio systems," *IEEE Trans. Signal Process.*, vol. 57, no. 11, pp. 4182–4195, Nov. 2009.
- [10] B. C. Levy, *Principles of Signal Detection and Parameter Estimation*. New York, NY, USA: Springer, 2008.
- [11] J. A. Maya and L. Rey Vega, "On fully-distributed composite tests with general parametric data distributions in sensor networks," *IEEE Trans. Signal Inf. Process. Over Netw.*, vol. 7, pp. 509–521, 2021, doi: [10.1109/TSIPN.2021.3101992](https://doi.org/10.1109/TSIPN.2021.3101992).
- [12] S. Al-Sayed, J. Plata-Chaves, M. Muma, M. Moonen, and A. Zoubir, "Node-specific diffusion LMS-based distributed detection over adaptive networks," *IEEE Trans. Signal Process.*, vol. 66, no. 3, pp. 682–697, Feb. 2018.
- [13] A. H. Sayed, S.-Y. Tu, J. Chen, X. Zhao, and Z. J. Towfic, "Diffusion strategies for adaptation and learning over networks: An examination of distributed strategies and network behavior," *IEEE Signal Process. Mag.*, vol. 30, no. 3, pp. 155–171, May 2013.
- [14] P. Willett, P. F. Swaszek, and R. S. Blum, "The good, bad and ugly: Distributed detection of a known signal in dependent Gaussian noise," *IEEE Trans. Signal Process.*, vol. 48, no. 12, pp. 3266–3279, Dec. 2000.



- [15] S. Li and X. Wang, "Fully distributed sequential hypothesis testing: Algorithms and asymptotic analyses," *IEEE Trans. Inf. Theory*, vol. 64, no. 4, pp. 2742–2758, Apr. 2018.
- [16] D. Ciuonzo, P. S. Rossi, and P. Willett, "Generalized Rao test for decentralized detection of an uncooperative target," *IEEE Signal Process. Lett.*, vol. 24, no. 5, pp. 678–682, May 2017.
- [17] D. Ciuonzo and P. S. Rossi, "Decision fusion with unknown sensor detection probability," *IEEE Signal Process. Lett.*, vol. 21, no. 2, pp. 208–212, Feb. 2014.
- [18] C. Li, G. Li, and P. K. Varshney, "Distributed detection of sparse stochastic signals with 1-bit data in tree-structured sensor networks," *IEEE Trans. Signal Process.*, vol. 68, pp. 2963–2976, 2020, doi: [10.1109/TSP.2020.2988598](https://doi.org/10.1109/TSP.2020.2988598).
- [19] D. Ciuonzo, P. S. Rossi, and P. K. Varshney, "Distributed detection in wireless sensor networks under multiplicative fading via generalized score tests," *IEEE Internet Things J.*, vol. 8, no. 11, pp. 9059–9071, Jun. 2021.
- [20] H. R. Ahmadi, N. Maleki, and A. Vosoughi, "On power allocation for distributed detection with correlated observations and linear fusion," *IEEE Trans. Veh. Technol.*, vol. 67, no. 9, pp. 8396–8410, Sep. 2018.
- [21] P. S. Rossi, D. Ciuonzo, K. Kansanen, and T. Ekman, "Performance analysis of energy detection for MIMO decision fusion in wireless sensor networks over arbitrary fading channels," *IEEE Trans. Wireless Commun.*, vol. 15, no. 11, pp. 7794–7806, Nov. 2016.
- [22] T. Royen, "Expansions for the multivariate chi-square distribution," *J. Multivar. Anal.*, vol. 38, no. 2, pp. 213–232, Aug. 1991.
- [23] D. Morales-Jimenez, J. F. Paris, J. T. Entrambasaguas, and K.-K. Wong, "On the diagonal distribution of a complex Wishart matrix and its application to the analysis of MIMO systems," *IEEE Trans. Commun.*, vol. 59, no. 12, pp. 3475–3484, Dec. 2011.
- [24] N. Y. Ermolova and O. Tirkkonen, "Distribution of diagonal elements of a general central complex Wishart matrix," *IEEE Commun. Lett.*, vol. 16, no. 9, pp. 1373–1376, Sep. 2012.
- [25] J. A. Maya and L. Rey Vega, "On the effect of spatial correlation on distributed energy detection of a stochastic process," in *Proc. IEEE ICASSP*, 2021, pp. 5045–5049.
- [26] Y.-C. Liang, Y. Zeng, E. C. Y. Peh, and A. T. Hoang, "Sensing-throughput tradeoff for cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 4, pp. 1326–1337, Apr. 2008.
- [27] J. Unnikrishnan and V. V. Veeravalli, "Cooperative sensing for primary detection in cognitive radio," *IEEE J. Sel. Topics Signal Process.*, vol. 2, no. 1, pp. 18–27, Feb. 2008.
- [28] F. F. Digham, M.-S. Alouini, and M. K. Simon, "On the energy detection of unknown signals over fading channels," in *Proc. IEEE ICC*, vol. 5, 2003, pp. 3575–3579.
- [29] S. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*, vol. 2, 1st ed. Upper Saddle River, NJ, USA: Prentice-Hall, 1998.
- [30] T. Kailath and H. V. Poor, "Detection of stochastic processes," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2230–2231, Oct. 1998.
- [31] D. R. Jensen, "The joint distribution of traces of Wishart matrices and some applications," *Ann. Math. Statist.*, vol. 41, no. 1, pp. 133–145, Feb. 1970.
- [32] H. J. Landau and H. O. Pollak, "Prolate spheroidal wave functions, Fourier analysis and uncertainty—Part III: The dimension of the space of essentially time and bandlimited signals," *Bell Syst. Techn. J.*, vol. 41, no. 4, pp. 1295–1336, Jul. 1962.
- [33] H. Urkowitz, "Energy detection of unknown deterministic signals," *Proc. IEEE*, vol. 55, no. 4, pp. 523–531, Apr. 1967.
- [34] J. A. Maya and A. M. Tonello, "A fully-distributed radio source detector for fast fading Rayleigh channels," in *Proc. IEEE 96th Veh. Technol. Conf. (VTC2022)*, 2022, pp. 1–7.
- [35] A. Wiesel and A. O. Hero, "Distributed covariance estimation in Gaussian graphical models," *IEEE Trans. Signal Process.*, vol. 60, no. 1, pp. 211–220, Jan. 2012.
- [36] D. Tse and P. Viswanath, *Fundamental of Wireless Communication*, 1st ed. Cambridge, U.K.: Cambridge Univ. Press, 2005.
- [37] T. Rappaport, *Wireless Communications: Principles and Practice*, 2nd ed. Upper Saddle River, NJ, USA: Prentice Hall, 2002.
- [38] V.V. Petrov, *Sums of Independent Random Variables*. Berlin, Germany: Springer, 1975.
- [39] W. Feller, *An Introduction to Probability Theory and Its Applications*, vol. 2, 2nd ed. New York, NJ, USA: Wiley, 1971.
- [40] L. Xiao and S. Boyd, "Fast linear iterations for distributed averaging," *Syst. Control Lett.*, vol. 53, no. 1, pp. 65–78, Sep. 2004.
- [41] R. Viswanathan and P. Varshney, "Distributed detection with multiple sensors part I. Fundamentals," *Proc. IEEE*, vol. 85, no. 1, pp. 54–63, Jan. 1997.
- [42] A. Patel, H. Ram, A. K. Jagannatham, and P. K. Varshney, "Robust cooperative spectrum sensing for MIMO cognitive radio networks under CSI uncertainty," *IEEE Trans. Signal Process.*, vol. 66, no. 1, pp. 18–33, Jan. 2018.
- [43] A. Chawla, A. Patel, A. K. Jagannatham, and P. K. Varshney, "Distributed detection in massive MIMO wireless sensor networks under perfect and imperfect CSI," *IEEE Trans. Signal Process.*, vol. 67, no. 15, pp. 4055–4068, Aug. 2019.
- [44] A. Goldsmith, *Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2005.
- [45] K. Kotobi and S. G. Bilen, "Secure blockchains for dynamic spectrum access: A decentralized database in moving cognitive radio networks enhances security and user access," *IEEE Veh. Technol. Mag.*, vol. 13, no. 1, pp. 32–39, Mar. 2018.
- [46] M. A. Branch, T. F. Coleman, and Y. Li, "A subspace, interior, and conjugate gradient method for large-scale bound-constrained minimization problems," *SIAM J. Sci. Comput.*, vol. 21, no. 1, pp. 1–23, 1999.
- [47] V. Y. Pan and Z. Q. Chen, "The complexity of the matrix eigenproblem," in *Proc. 31st Annu. ACM Symp. Theory Comput.*, New York, NY, USA, 1999, pp. 507–516.
- [48] S. Gratton, C. W. Royer, L. N. Vicente, and Z. Zhang, "Complexity and global rates of trust-region methods based on probabilistic models," *IMA J. Numer. Anal.*, vol. 38, no. 3, pp. 1579–1597, Jul. 2018.
- [49] R. Niu, B. Chen, and P. K. Varshney, "Fusion of decisions transmitted over Rayleigh fading channels in wireless sensor networks," *IEEE Trans. Signal Process.*, vol. 54, no. 3, pp. 1018–1027, Mar. 2006.
- [50] A. Taherpour, M. Nasiri-Kenari, and S. Gazor, "Multiple antenna spectrum sensing in cognitive radios," *IEEE Trans. Wireless Commun.*, vol. 9, no. 2, pp. 814–823, Feb. 2010.
- [51] 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, no. 1, pp. 84–88, Jan. 2010.
- [52] Y. P. E. Wang et al., "A primer on 3GPP narrowband Internet of Things," *IEEE Commun. Mag.*, vol. 55, no. 3, pp. 117–123, Mar. 2017.
- [53] S. Yang, Y. Lai, A. Jakobsson, and W. Yi, "Hybrid quantized signal detection with a bandwidth-constrained distributed radar system," *IEEE Trans. Aerosp. Electron. Syst.*, early access, Jul. 17, 2023, doi: [10.1109/TAES.2023.3296344](https://doi.org/10.1109/TAES.2023.3296344).
- [54] J. R. Magnus and H. Neudecker, "The commutation matrix: Some properties and applications," *Ann. Statist.*, vol. 7, no. 2, pp. 381–394, Mar. 1979.
- [55] P. Billingsley, *Probability and Measure*, 3rd ed. Hoboken, NJ, USA: Wiley, 1995.
- [56] I. S. Gradshteyn, I. M. Ryzhik, A. Jeffrey, and D. Zwillinger, *Table of Integrals, Series, and Products*, 6th ed. San Diego, CA, USA: Academic, 2000.