ON THE EFFECT OF SPATIAL CORRELATION ON DISTRIBUTED ENERGY DETECTION OF A STOCHASTIC PROCESS

Juan Augusto Maya[†‡*]

[†]CSC-CONICET, Buenos Aires, Argentina
jmaya@fi.uba.ar

Leonardo Rey Vega[†‡]

[‡]Universidad de Buenos Aires, Argentina
lrey@fi.uba.ar

ABSTRACT

We consider the problem of distributed detection over a network of sensors using energy measurements. Assuming that the problem is to detect the presence of stochastic signal with arbitrary temporal correlation, we exactly compute the characteristic function of the energy measurement at each sensing node. Is it is shown, that when time-bandwith product WT is large enough, the exact joint density function of the measurements at the sensor sites, can be closely approximated by a product of non-trivial density function that depends of some unknown parameters of the source signal and the sensor network. As those parameters can be easily estimated at each sensor node, we propose a modified GLRT. Through numerical experiments we show that this proposal is has performance close to the optimal one which use the exact joint density function and with a significant reduction in the consumption of network resources as bandwith and transmission power.

***Index Terms*—** sensor networks, distributed detection, characteristic function, energy detection.

1. INTRODUCTION

Wireless sensor networks (WSN) as a key technology in the emerging paradigm of Internet of Things (IoT) [1, 2] have received considerable attention in the past. In this context, we are interested in the problem of distributed source detection [3, 4]. We consider a set of nodes, with sensing, computing and communication capabilities, which jointly sense a phenomenon of interest and exchange information with their neighbors to obtain a final decision about the true state of nature. Moreover, we assume a fully distributed scenario in which there is no Fusion Center (FC) which centralizes the measurements and reach the final decision. Instead, the individual sensors reach a consensus decision about the state of nature. Clearly, this option is robust against node failures and attacks specially directed to the FC. In addition to this, communications between nodes are done locally, over typically short distances, saving energy and also bandwidth, e.g., employing spatial reuse of the frequency bands. When the source signal is on, the measurements taken at each sensor node are clearly correlated because these measurements are different noisy versions of the same signal. However, most of the works found in the literature assumes that the spatial measurements are independent [3, 5–7] neglecting the above mentioned dependence. In a fully distributing setting, the use of the optimal likelihood ratios becomes very costly in terms of energy and bandwidth when the measurements at the sensors are dependent. That is, the number of exchanges between the nodes needed to reach a consensus decision about the true state of the source signal could be large depending on the specific statistical model of the source. In this work, we consider a distributed detection problem in which each sensor can only sense the average energy during a given measurement temporal window. This is a common situation in some well-known applications such as spectrum sensing in cognitive radio [8]. By using the formalism of characteristic functions, we show that when the source signal is Gaussian (with arbitrary temporal correlation) and the product of the source signal bandwidth and the size of the measurement window at each sensor is large enough, then, the full joint distribution of the energy measurements taken at each sensor of the N nodes can be factorized as the product of N non-central chi-square distributions with different parameters (which are assumed to be unknown at each sensor node). As this approximation of the true measurements distribution is a product one, the implementation of a fully distributed detection procedure becomes simpler and more economical in energy and bandwidth resources. Moreover, the unknown parameters of each term in the approximated product pdf can be easily estimated at each sensor node, leading to the implementation of a simple distributed Generalized Likelihood Ratio Test (GLRT) in the network. Through numerical experiments, we also show how the proposed test has comparable performance with the optimal test with only a fraction of network resources.

2. PROBLEM SETTING

Consider N sensors distributed in geographical area with the source being located in the same area. It is assumed that this source emits a stochastic signal and its location is unknown for the sensing network. The goal of the sensing network is to

*Work supported by UBACyT20020170200283BA and PICT-2017-4533.

detect when this source is on or not. Each node of the network senses the environment during L measurement windows, each of duration M. During the l-th time windows (with $l \in [1:L]$), we can easily see that the signal at the position of each sensor $k \in [1:N]$ can be writen as:

$$\begin{aligned} \mathbf{y}_{k,l} &= \mathbf{v}_{k,l} \text{ under } \mathcal{H}_0 \\ \mathbf{y}_{k,l} &= h_{k,l}\mathbf{s}_l + \mathbf{v}_{k,l} \text{ under } \mathcal{H}_1 \end{aligned} \quad (1)$$

where $\mathbf{y}_{k,l} \in \mathbb{C}^M$ is the temporal signal at the position of sensor k and during time window l. Under $\mathcal{H}_0$ (the source is off), these signals are distributed as complex Gaussian circular noise, that is, for each k, l, $\mathbf{v}_{k,l} \sim \mathcal{CN}(\mathbf{0}, \sigma_v^2\mathbf{I}_M)$ $\sigma_v^2 > 0$. Across the sensors and measurement windows it is assumed that these vectors are uncorrelated and identically distributed. When the the unknown source is emitting (under $\mathcal{H}_1$), the signal at the position of each sensor is composed by noise $\mathbf{v}_{k,l}$ with the same characteristics as before, but also by signal $h_{k,l}\mathbf{s}_l$. Vector $\mathbf{s}_l \in \mathbb{C}^M$ is the signal emitted by the source during measurement window l. It is assumed that is a complex and circular Gaussian signal distributed as $\mathbf{s}_l \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Sigma}_{\mathbf{s},l})$, where $\mathbf{\Sigma}_{\mathbf{s},l}$ is the temporal correlation matrix of the source signal during measurement window l[1]. On the other hand, the coefficient $h_{k,l} \in \mathbb{C}$ quantifies the path-loss gain that exists between the position of the source and the k-th sensor position at measurement window l. This coefficient, is highly dependent of the scenario under study and it is typically unknown in practice. Notice that, as the signal part $\mathbf{s}_l$ is the same for each sensor (although affected by different path-loss coefficients h_k), the signals under $\mathcal{H}_1$ are temporally and spatially correlated during each measurement window. It is assumed that the nodes can only sense average energy during each measurement window. So, at each window, the sensors output the value $E_{k,l} \equiv \frac{1}{M}\|\mathbf{y}_{k,l}\|^2$. These energy values are then used by the sensing network to determine if the unknown source is actually on. In order to perform the hypothesis test we first need the joint probability distribution of the energy measurements $\mathbf{E}_l = [E_{1,l}, E_{2,l}, \ldots, E_{N,l}]^T$ for each measurement window.

3. ON THE DISTRIBUTION OF $\mathbf{E}_L$

Consider a fixed measurement window with $l \in [1:L]$. We can define the vector $\mathbf{y}_l = \left[\mathbf{y}_{1,l}^T, \mathbf{y}_{2,l}^T, \ldots, \mathbf{y}_{N,l}^T\right]^T$ as the NM-length vector that contains the signals during time measurement l at the N sensor positions. The exact joint density of vector $\mathbf{E}_l$ is a very difficult problem. It is well-known that it is related with the distribution of the trace of $\mathbf{Y}_l\mathbf{Y}_l^H$ where the rows of matrix $\mathbf{Y}_l$ are $\mathbf{y}_{1,l}^H, \ldots, \mathbf{y}_{N,l}^H$. Although this problem has attracted some interest in the wireless communications community (see [9] and [10]), most results consider the case in which the vectors $\mathbf{y}_{1,l}\mathbf{y}_{2,l}$ $\ldots \mathbf{y}_{N,l}$ are independent and identically distributed, which is not the case considered in this paper. The main reason behind this, is the temporal correlation of the source signal $\mathbf{s}_l$ and the presence of the path-loss coefficient between the source and the sensor locations. We will consider the exact computation of the characteristic function for vector $\mathbf{E}_l$, 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. Consider $\Psi_l(\boldsymbol{\omega}) \equiv \mathbb{E}_l[e^{j\boldsymbol{\omega}^T\mathbf{E}_l}]$ for $l \in [1:L]$ and $\boldsymbol{\omega} \in \mathbb{R}^N$. The following lemma, whose proof is relegated to the appendix, gives us the expression for $\Psi_l(\boldsymbol{\omega})$.

Lemma 1. The characteristic function $\Psi_l^i(\boldsymbol{\omega})$, where $i = 0, 1$ denotes the hypothesis $\mathcal{H}_i$ under consideration is given by:

$$\Psi_l^i(\boldsymbol{\omega}) = \prod_{k=1}^N \left(1 - j\tfrac{\omega_k\sigma_v^2}{M}\right)^{-M} G_i(\boldsymbol{\omega}),\ i = 0, 1, \quad (2)$$

where $G_0(\boldsymbol{\omega}) = 1$ and:

$$G_1(\boldsymbol{\omega}) = \prod_{i=1}^M \left(1 - j\frac{\lambda_{i,l}}{M}\sum_{k=1}^N \frac{|h_{k,l}|^2\omega_k}{1 - j\frac{\omega_k\sigma_v^2}{M}}\right)^{-1}, \quad (3)$$

where $\lambda_{1,l}, \lambda_{2,l}, \ldots, \lambda_{N,l}$ are the positive eigenvalues of source signal correlation matrix $\mathbf{\Sigma}_{\mathbf{s},l}$.

Under $\mathcal{H}_1$, $\Psi_l^1(\boldsymbol{\omega})$ does not correspond to N independent random variables. The values of $E_{k,l}$ are dependent random variables as expected even if source signal $\mathbf{s}_l$ is uncorrelated in time. This is a consequence of the spatial correlation induced by the noisy versions of the source signal at each sensor site. In principle, obtaining the exact joint density is very hard. However, assuming that $\lambda_{i,l} \neq \lambda_{j,l}$ with $i \neq j$, the marginal density for $E_{k,l}$ with $k \in [1:N]$ can be written as:

$$p_{E_{k,l}}^1(E_{k,l}) = \sum_{i=1}^M \prod_{j\neq i}^M \left(1 - \tfrac{\lambda_{j,l}|h_{k,l}|^2+\sigma_v^2}{\lambda_{i,l}|h_{k,l}|^2+\sigma_v^2}\right) \tfrac{M}{\lambda_{i,l}|h_{k,l}|^2+\sigma_v^2} \times \exp\left(-\tfrac{ME_{k,l}}{\lambda_{i,l}|h_{k,l}|^2+\sigma_v^2}\right),\ k \in [1:N]. \quad (4)$$

Although, the marginal densities are important, in order to implement a hypothesis test for detecting the source signal we need the full joint density of the energy measurements at the sensors. Although this problem is hard, some insights can be obtained when M is sufficiently large using the characteristic function $\Psi_l^1(\boldsymbol{\omega})$ in (2). Note that $M \approx WT$, where W is the bandwidth of the continuous time version of the source signal and T is the time duration of the sensing windows in which the energy measurement in each sensor is done. For several cases of practical interest WT will be large, and the following approximation will be useful:

$$1 - j\tfrac{\lambda_{i,l}}{M}\sum_{k=1}^N \tfrac{|h_{k,l}|^2\omega_k}{1-j\frac{\omega_k\sigma_v^2}{M}} \approx \exp\left(-j\tfrac{\lambda_{i,l}}{M}\sum_{k=1}^N \tfrac{|h_{k,l}|^2\omega_k}{1-j\frac{\omega_k\sigma_v^2}{M}}\right),$$

for $i \in [1:M]$ and $k \in [1:N]$. We can write:

$$\Psi_l^1(\boldsymbol{\omega}) \approx \prod_{k=1}^N \exp\left(\tfrac{j}{M}\tfrac{\mathrm{tr}(\mathbf{\Sigma}_{\mathbf{s},l})|h_{k,l}|^2\omega_k}{1-j\frac{\omega_k\sigma_v^2}{M}}\right)\left(1 - j\tfrac{\omega_k\sigma_v^2}{M}\right)^{-M}, \quad (5)$$

where we use the fact that $\sum_{i=1}^M \lambda_{i,l} = \mathrm{tr}\,(\mathbf{\Sigma}_{\mathbf{s},l})$. The most striking fact about this last expression is that, as M grows, the values of the entries of vector $\mathbf{E}_l$ becomes less statistically

[1] In order to simplify the developments we will assume that the temporal correlation across different measurement windows is negligible.

dependent between them, as (5) is the characteristic function of N independent random variables.

Remark 1. It is well known that each term in (5) is the characteristic function of a non-central chi-square random variable. On the other hand, under $\mathcal{H}_0$, it is already clear that the vector $\mathbf{E}_l$ has components independent and identically distributed as (2) shows. However, it is important to observe that under $\mathcal{H}_1$, the use of M large allows us to approximately consider that each $E_{k,l}$ is independent. Nevertheless, the approximation is not equivalent to approximate the full joint density of $\mathbf{E}_l$ by the product of its marginals given by (4). This shows, that the result of the approximation obtained when M is large is not trivial.

The fact that M is typically large in several application of interest shows that we can propose to approximate the true and exact likelihood ratio, using the joint pdf corresponding to (3) with a likelihood ratio test using the product pdf corresponding to (5). This has a twofold advantage. In first place, although not exact, the likelihood corresponding to (5) can be easily computed in closed form. Secondly, as the approximation naturally leads to a product pdf, it is well suited to a distributed detection scenario, where a likelihood ratio with a non-product pdf would be very costly to implement in terms of communications power and bandwidth.

4. APPROXIMATE LIKELIHOOD RATIO TEST

We will consider the special case in which $\mathbf{\Sigma}_{\mathbf{s},l} \equiv \mathbf{\Sigma}_{\mathbf{s}}$ for every $l = 1, \ldots, L$. That is, during the L measurement windows the statistical properties of the source signal are invariant. We also define the average variance $\sigma_{\mathbf{s}}^2 \equiv \frac{\mathrm{tr}(\mathbf{\Sigma}_{\mathbf{s}})}{M}$. Similarly, we will also assume that the channel characteristic between the source position and the sensors are also invariant during the L measurement windows, that is $h_{k,l} \equiv h_k$ for every $k = 1, \ldots, N$ and $l = 1, \ldots, L$. It is easy to show that $p^1_{\mathbf{E}_l}(\mathbf{E}_l) \approx \tilde{p}^1_{\mathbf{E}_l}(\mathbf{E}_l) \equiv \prod_{k=1}^N \tilde{p}^1_{E_{k,l}}(E_{k,l})$ where

$$\tilde{p}^1_{E_{k,l}}(E_{k,l}) = \tfrac{M}{\sigma_v^2} \exp\left(-\tfrac{M}{\sigma_v^2}(E_{k,l} + \sigma_{\mathbf{s}}^2|h_k|^2)\right) \times \left(\tfrac{E_{k,l}}{\sigma_{\mathbf{s}}^2|h_k|^2}\right)^{\frac{M-1}{2}} I_{M-1}\left(\tfrac{2M}{\sigma_v^2}\sqrt{\sigma_{\mathbf{s}}^2|h_k|^2 E_{k,l}}\right), \text{ and} \quad (6)$$

with $I_\beta(z)$ is the modified Bessel function of the first kind with integer parameter $\beta \geq 0$. The pdf $p^0_{\mathbf{E}_l}(\mathbf{E}_l)$ of $\mathbf{E}_l$ under $\mathcal{H}_0$ can be easily obtained from Lemma 1 as $p^0_{\mathbf{E}_l}(\mathbf{E}_l) = \prod_{k=1}^N p^0_{E_{k,l}}(E_{k,l})$, where each $p^0_{E_{k,l}}(E_{k,l})$ is the pdf of a central chi-square random variable with $2M$ degrees of freedom. We can write the log-likelihood ratio as:

$$T_{\mathrm{LRT}}(\mathbf{E}_1, \ldots, \mathbf{E}_L) = \sum_{l=1}^L \log\left(\frac{p^1_{\mathbf{E}_l}(\mathbf{E}_l)}{p^0_{\mathbf{E}_l}(\mathbf{E}_l)}\right) \quad (7)$$

In a practical situation is not realistic to assume that the channel gains h_k are known. A similar observation can also be made about the variance average $\sigma_{\mathbf{s}}^2$ of the source signal[2]. This call for handling the situation of these unknown parameters. Although several possibilities can be explored, we will consider the use of a GLRT [11]. We define $c_k \equiv \sigma_{\mathbf{s}}^2|h_k|^2$ for $k = 1, \ldots, N$ leading to the test:

$$T_{\mathrm{GLRT}}(\mathbf{E}_1, \ldots, \mathbf{E}_L) = \max_{\mathbf{c}\geq 0} \sum_{l=1}^L \log\left(\frac{p^1_{\mathbf{E}_l}(\mathbf{E}_l, \mathbf{c})}{p^0_{\mathbf{E}_l}(\mathbf{E}_l)}\right) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \gamma \quad (8)$$

with $\mathbf{c} = [c_1, c_2, \ldots, c_N]^T$ and where we used the fact that the unknown parameters are positive and only affect to energy measurements under $\mathcal{H}_1$. The threshold γ is determined by the desired probability of false alarm depending only on the distribution under $\mathcal{H}_0$. The GLRT proposed in this work is defined by replacing $p^1_{\mathbf{E}_l}(\mathbf{E}_l, \mathbf{c})$ by $\tilde{p}^1_{\mathbf{E}_l}(\mathbf{E}_l, \mathbf{c})$ in (8):

$$\tilde{T}_{\mathrm{GLRT}}(\mathbf{E}_1, \ldots, \mathbf{E}_L) = \sum_{k=1}^N \sum_{l=1}^L \log\left(\frac{\tilde{p}^1_{E_{k,l}}(E_{k,l}, \hat{c}_k)}{p^0_{E_{k,l}}(E_{k,l})}\right), \quad (9)$$

where $\hat{c}_k = \arg\max_{c_k \geq 0} \sum_{l=1}^L \log \tilde{p}^1_{E_{k,l}}(E_{k,l}, c_k)$, with $\tilde{p}^1_{E_{k,l}}(E_{k,l}, c_k)$ given by (6) with a slight abuse of notation. The statistic $\tilde{T}_{\mathrm{GLRT}}$ can be easily implemented in a distributed setting given that the inner sum in (9) is computed at each node without any communication with other nodes. Once this is done, it is required to compute the *spatial* sum $\sum_{k=1}^N(\cdot)$ over all the sensors in the network. Next $\bar{a} \equiv \sum_{k=1}^N a_k$ will represent that sum. Each sensor node produces a scalar value $a_k \in \mathbb{R}$, $k \in \mathcal{N} \equiv [1 : N]$ and it is necessary to compute $\tilde{a} = \frac{1}{N}\sum_{k=1}^N a_k$ (or the sum $\bar{a} = N\tilde{a}$) at each node with minimal resources to exchange information between the nodes. The spatial average can be computed via a consensus procedure such as in [12, 13]. Consider a network (modeled as a connected graph) $\mathcal{G} = (\mathcal{N}, \mathcal{E})$ consisting of a set of nodes $\mathcal{N}$ and a set of edges $\mathcal{E}$, where each edge $\{i, j\} \in \mathcal{E}$ is an unordered pair of distinct nodes. The set of neighbors of node i is denoted by $\mathcal{N}_i = \{j \in \mathcal{N} | \{i, j\} \in \mathcal{E}\}$. The average $\tilde{a}$ can be computed iteratively as, $t \in \mathbb{N}$:

$$a_k(t) = W_{kk}a_k(t-1) + \textstyle\sum_{j\in\mathcal{N}_k} W_{kj}a_j(t-1), k \in \mathcal{N}, \quad (10)$$

where $a_k(t)$ is the average after t iterations, $a_k(0) = a_k$ is the initial value and W_{kj} is the weight on $a_j(t-1)$ at the node k. Considering local transmissions only, i.e., each node broadcasts its local value at iteration t only to the nodes in its neighborhood, we have that for each $k \in \mathcal{N}$, $W_{kj} = 0$ for $j \notin \mathcal{N}_k$ and $j \neq k$. The weights are assumed to be symmetric with value $W_{kj} = W_{jk} = 1/\max(d_k, d_j)$, where d_k is the degree of node k, i.e., the number of neighbors of the node k. The convergence to the required average is guaranteed given that graph is not bipartite [12]. The selected stopping criterion for (10) is a fixed number of iterations N_{it}. Algorithm 1 summarizes the steps required to compute the statistic $\tilde{T}_{\mathrm{GLRT}}$.

5. NUMERICAL RESULTS

In this section we evaluate the performance of the exact and the approximated statistics in the context of a cognitive radio (CR) system [14]. In CR systems, unlicensed, or secondary

[2]The value of σ_v^2 is usually not known. However, it can be easily estimated under source silence periods to obtain a CFA receiver.

Algorithm 1 Distributed implementation of $\tilde{T}_{\text{GLRT}}$

for $k = 1, \ldots, N$ **do** *(simultaneously at each sensor)*
  Compute the local estimate $\hat{c}_k$.
  Compute $\tilde{T}_{\text{GLRT},k} \equiv \sum_{l=1}^{L} \log\left(\frac{\tilde{p}^1_{E_{k,l}}(E_{k,l},\hat{c}_k)}{p^0_{E_{k,l}}(E_{k,l})}\right)$.
  $\tilde{T}_{\text{GLRT},k} = \text{SPATIALSUM}(\{\tilde{T}_{\text{GLRT},j}\}_{j\in\mathcal{N}_k\cup\{k\}})$
  if $\tilde{T}_{\text{GLRT},k} < \gamma$ **then** Sensor k decides $\mathcal{H}_0$,
  else Sensor k decides $\mathcal{H}_1$.
  end if
end for
function $\text{SPATIALSUM}(\{a_j\}_{j\in\mathcal{N}_k\cup\{k\}})$ ▷ Compute $\bar{a}_k$.
  $t = 0$, $a_k(0) = a_k$ ▷ Initial condition for $t = 0$.
  while $t < N_{it}$ **do**
    $t = t + 1$
    Compute the spatial average $a_k(t)$ using (10).
  end while
  return $Na_k(t)$ ▷ Return the sum $\bar{a}_k$
end function

users (SU), sense the spectrum to detect the presence or absence of licensed, or a primary user (PU), in order to use the spectrum when it is available. The SUs, located at positions $\boldsymbol{r}_n^{\text{SU}} \in \mathbb{R}^2$, $n = 1, \ldots, N$, take energy measurements $\{E_{k,l}\}$, estimate the local parameter c_k, cooperate with the other SUs and finally make a decision about the presence or not of the PU, located at the unknown position $\boldsymbol{r}^{\text{PU}} \in \mathbb{R}^2$. The channel coefficient is assumed to be $h_n = \text{PL}(\|\boldsymbol{r}^{\text{PU}} - \boldsymbol{r}_n^{\text{SU}}\|)\eta_n$, where $\|\cdot\|$ computes the 2-norm, $\text{PL}(d) = d^{-\alpha/2}$ is the path-loss amplitude attenuation for a distance d and a path-loss exponent α, and η_n represents the shadowing-fading and the multi-path fading effects. We assume that the PU and the SUs are static during the observation time interval, thus, η_n is constant during that time. The SUs positions were uniformly randomly generated in an square area of 5km×5km. The distance from the PU to the center of the SU network is approximately 8 km. It was chosen such that the SU network is near to the edge the PU cell, that is, in low signal to interference and noise ratio (SINR) conditions. The parameters are: $N = 10$, $L = 12$, $N_0 = -174$ dBm/MHz, $\alpha = 5$, $W = 5$ MHz, and T takes the values $0.6, 1.33$ and 3μs to produce $M = 3, 8$, and 15, respectively. The transmit power of the PU is $\text{tr}(\boldsymbol{\Sigma}_{\mathbf{s}})/T = 10$mW when it is active and $\boldsymbol{\Sigma}_{\mathbf{s}}$ is a symmetric Toeplitz matrix with first row given by $[1, 0.5, 0.25, \ldots, 0.5^{N-1}]$. We assume $\eta_n = 1$ for all n. The graph that models the connections between SUs is shown in Fig. 1, with $|\mathcal{E}| = 20$ edges and $N_{it} = 20$ iterations are considered for Algorithm 1. The approximate implementation of the GLRT (9) needs to compute the MLE of the local parameter c_k at each node $k = 1, \ldots, N$. This is a one-dimensional nonlinear problem with bound constraint. To solve it, we implement a trust-region method [15] at each node. It usually takes few (4-7) iterations to converge to a local maximum.

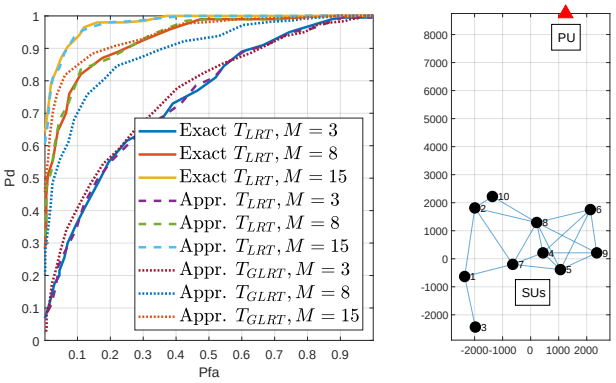

Fig. 1: ROC for exact and approximate computation of the GLRT (left). Network graph (right).

Thus, Algorithm 1 is run to cooperatively build the approximate GLRT at each node and, finally, make a decision. In Fig. 1 we show the receiver operating characteristic (ROC) of the presented statistics, where given a threshold γ and a statistic T, the detection and the false alarm probabilities are defined by $P_{\text{d}} = \mathbb{P}(T > \gamma)$ under $\mathcal{H}_1$ and $P_{\text{fa}} = \mathbb{P}(T > \gamma)$ under $\mathcal{H}_0$, respectively. These probabilities are computed using the method of Monte Carlo with 200 realizations. As the exact pdf of $\mathbf{E}_l$ under $\mathcal{H}_1$ is unknown, we evaluate the performance of the exact LRT by numerically computing the multidimensional (N dimensions in our case) inverse Laplace transform of (3). Although several algorithms exists [16], they are all numerically intensive and we are limited in the amount of nodes and realizations that we can simulate. In particular, we use the method of Gaver-Stehfest [16]. This computation complexity let us to compare the performance of the exact and the approximate statistics only when the parameter $\mathbf{c}$ is known[3]. In that case, we see that the exact LRT and the approximate LRT have similar performance for $M = 15, 8$ and even for $M = 3$. Finally, we also plot the performance of the GLRT for the same scenario. As expected, the fact of having to estimate the parameter $\mathbf{c}$ given that it is unknown introduces a penalty in the performance of this algorithm with respect to the unrealizable LRT.

6. CONCLUDING REMARKS

We considered the problem of detecting the presence of a stochastic signal in a distributed fashion over a network of sensors capable of acquiring energy measurements. From the mathematical analysis of the characteristic function of the measurements we obtained a simple approximation of their true probability density function. Through numerical experiments we showed that, despite the spatial correlation induced by the model, it is possible to implement a quasi-optimal hypothesis testing problem with large savings in network re-

[3]To compute the MLE at each node would require much more evaluations of the unknown pdf, which is computationally much more expensive.

sources.

A. PROOF OF LEMMA 1

It is clear that we can write:

$$\Psi_l(\boldsymbol{\omega}) = \mathbb{E}\left[e^{j\frac{1}{M}\sum_{k=1}^{N}\omega_k\|\mathbf{y}_{k,l}\|^2}\right]. \tag{11}$$

Assuming that $\boldsymbol{\Sigma}_{\mathbf{s},l} = \mathbf{Q}_l\boldsymbol{\Lambda}_l\mathbf{Q}_l^H$ where $\mathbf{Q}_l$ is unitary and $\boldsymbol{\Lambda}_l = \text{diag}(\lambda_{1,l}, \lambda_{2,l}, \ldots, \lambda_{N,l})$ with $\lambda_{k,l}$ the k-largest positive eigenvalue of $\boldsymbol{\Sigma}_{\mathbf{s},l}$, we can define $\tilde{\mathbf{y}}_l \equiv (\mathbf{I}_N \otimes \mathbf{Q}_l)\,\mathbf{y}_l$. It is inmediate to see that under $\mathcal{H}_0$, $\tilde{\mathbf{y}}_l \sim \mathcal{CN}(\mathbf{0}, \sigma_v^2\mathbf{I}_{NM})$ and under $\mathcal{H}_1$, $\tilde{\mathbf{y}}_l \sim \mathcal{CN}(\mathbf{0}, \mathbf{h}_l\mathbf{h}_l^H \otimes \boldsymbol{\Lambda}_l + \sigma_v^2\mathbf{I}_{NM})$. Moreover, we have that

$$\Psi_l(\boldsymbol{\omega}) = \mathbb{E}\left[e^{j\frac{1}{M}\sum_{k=1}^{N}\omega_k\|\tilde{\mathbf{y}}_{k,l}\|^2}\right] = \mathbb{E}\left[e^{j\frac{1}{M}\tilde{\mathbf{y}}_l^H\mathbf{A}_{\boldsymbol{\omega}}\tilde{\mathbf{y}}_l}\right], \tag{12}$$

where $\mathbf{A}_{\boldsymbol{\omega}} = \text{diag}(\omega_1, \omega_2, \ldots, \omega_N) \otimes \mathbf{I}_M$. As under both hypothesis, $\tilde{\mathbf{y}}_l$ is a complex and circular Gaussian random vector, it is easy to obtain that under $\mathcal{H}_0$:

$$\Psi_l^0(\boldsymbol{\omega}) = \prod_{k=1}^{N}\frac{1}{\left(1 - j\frac{\omega_k\sigma_v^2}{M}\right)^M}. \tag{13}$$

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

$$\Psi_l^1(\boldsymbol{\omega}) = \frac{1}{\left|\mathbf{I}_{NM} - j\frac{\left(\mathbf{h}_l\mathbf{h}_l^H \otimes \boldsymbol{\Lambda}_l + \sigma_v^2\mathbf{I}_{NM}\right)\mathbf{A}_{\boldsymbol{\omega}}}{M}\right|} \tag{14}$$

Note that we can write: $\left(\mathbf{h}_l\mathbf{h}_l^H \otimes \boldsymbol{\Lambda}_l\right)\mathbf{A}_{\boldsymbol{\omega}}$

$$\begin{aligned} &= \left(\mathbf{h}_l\mathbf{h}_l^H \otimes \boldsymbol{\Lambda}_l\right)\left(\text{diag}(\omega_1, \omega_2, \ldots, \omega_N) \otimes \mathbf{I}_M\right) \\ &= \left(\mathbf{h}_l\mathbf{h}_l^H \text{diag}(\omega_1, \omega_2, \ldots, \omega_N)\right) \otimes \boldsymbol{\Lambda}_l, \end{aligned} \tag{15}$$

where we have used that $(\mathbf{A}\otimes\mathbf{B})(\mathbf{C}\otimes\mathbf{D}) = (\mathbf{AC})\otimes(\mathbf{BD})$. Consider now the *conmutation matrix* $\mathbf{K}_{N,M}$ [17]. This matrix is a permutation matrix of size $NM \times NM$ with several interesting properies. 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 $\mathbf{A} \otimes \mathbf{B} = \mathbf{K}_{N,M}\left(\mathbf{B} \otimes \mathbf{A}\right)\mathbf{K}_{N,M}^T$. Using this matrix and the result of (15) we can express $\mathbf{I}_{NM} - j\frac{\left(\mathbf{h}_l\mathbf{h}_l^H \otimes \boldsymbol{\Lambda}_l + \sigma_v^2\mathbf{I}_{NM}\right)\mathbf{A}_{\boldsymbol{\omega}}}{M}$ as: $\mathbf{K}_{N,M}\big(\mathbf{I}_{NM} - j\frac{\boldsymbol{\Lambda}_l\otimes\left(\mathbf{h}_l\mathbf{h}_l^H\boldsymbol{\Omega}\right) + \sigma_v^2\mathbf{K}_{N,M}^T(\boldsymbol{\Omega}\otimes\mathbf{I}_M)\mathbf{K}_{N,M}}{M}\big)\mathbf{K}_{N,M}^T$, where we have defined $\boldsymbol{\Omega} = \text{diag}(\omega_1, \omega_2, \ldots, \omega_N)$. In first place note that, as $\mathbf{K}_{N,M}^T = \mathbf{K}_{M,N}$, $\mathbf{K}_{N,M}^T\left(\boldsymbol{\Omega} \otimes \mathbf{I}_M\right)\mathbf{K}_{N,M}$ transforms into $\mathbf{I}_M \otimes \boldsymbol{\Omega}$. In second place, the matrix $\boldsymbol{\Lambda}_l \otimes \left(\mathbf{h}_l\mathbf{h}_l^H\boldsymbol{\Omega}\right)$ is a block-diagonal matrix with diagonal blocks given by $\lambda_{i,l}\mathbf{h}_l\mathbf{h}_l^H\boldsymbol{\Omega}$ with $i = 1, \ldots, M$. As a consequence, and using the fact that $\left|\mathbf{K}_{N,M}\right|\left|\mathbf{K}_{N,M}^T\right| = 1$, (15) can be put as:

$$\Psi_l^1(\boldsymbol{\omega}) = \frac{1}{\prod_{i=1}^{M}\left|\mathbf{I}_N - j\frac{\sigma_v^2\boldsymbol{\Omega}}{M} - j\frac{\lambda_{i,l}\mathbf{h}_l\mathbf{h}_l^H\boldsymbol{\Omega}}{M}\right|}. \tag{16}$$

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

$$\Psi_l^1(\boldsymbol{\omega}) = \frac{1}{\prod_{i=1}^{M}\left|\mathbf{I}_N - j\frac{\sigma_v^2\boldsymbol{\Omega}}{M}\right|\left(1 - j\frac{1}{M}\lambda_{i,l}\mathbf{h}_l^H\left(\mathbf{I}_N - j\frac{\sigma_v^2\boldsymbol{\Omega}}{M}\right)^{-1}\mathbf{h}_l\right)}. \tag{17}$$

As all the matrices in the previous equation are diagonal we get the result of the Lemma.

B. REFERENCES

[1] J. Gubbi, R. Buyya, S. Marusic, and M. Palaniswami, "Internet of Things (IoT): A vision, architectural elements, and future directions," *Future generation computer systems*, vol. 29, no. 7, pp. 1645–1660, 2013.

[2] Ala Al-Fuqaha, Mohsen Guizani, Mehdi Mohammadi, Mohammed Aledhari, and Moussa Ayyash, "Internet of things: A survey on enabling technologies, protocols, and applications," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 4, pp. 2347–2376, 2015.

[3] D. Ciuonzo and P.S. Rossi, "Decision fusion with unknown sensor detection probability," *IEEE Signal Processing Letters*, vol. 21, no. 2, pp. 208–212, 2014.

[4] J.-F. Chamberland and V.V. Veeravalli, "Decentralized detection in sensor networks," *IEEE Transactions on Signal Processing*, vol. 51, no. 2, pp. 407–416, Feb 2003.

[5] S Hamed Hamed and Ali Peiravi, "Reliable distributed detection in multi-hop clustered wireless sensor networks," *IET Signal Processing*, vol. 6, no. 8, pp. 743–750, 2012.

[6] Soummya Kar, Ravi Tandon, H Vincent Poor, and Shuguang Cui, "Distributed detection in noisy sensor networks," in *Information Theory Proceedings (ISIT), 2011 IEEE International Symposium on*. IEEE, 2011, pp. 2856–2860.

[7] Shang Li and Xiaodong Wang, "Fully distributed sequential hypothesis testing: Algorithms and asymptotic analyses," *IEEE Transactions on Information Theory*, vol. 64, no. 4, pp. 2742–2758, 2018.

[8] Y. C. Liang, Y. Zeng, E. C. Y. Peh, and A. T. Hoang, "Sensing-throughput tradeoff for cognitive radio networks," *IEEE Transactions on Wireless Communications*, vol. 7, no. 4, pp. 1326–1337, Apr 2008.

[9] David Morales-Jimenez, Jose F. Paris, J. T. Entrambasaguas, and Kai-Kit Wong, "On the diagonal distribution of a complex wishart matrix and its application to the analysis of mimo systems," *IEEE Transactions on Communications*, vol. 59, no. 12, pp. 3475–3484, Dec 2011.

[10] Natalia Y. Ermolova and Olav Tirkkonen, "Distribution of diagonal elements of a general central complex wishart matrix," *IEEE Communications Letters*, vol. 16, no. 9, pp. 1373–1376, Sep 2012.

[11] Bernard C Levy, *Principles of signal detection and parameter estimation*, Springer, 2008.

[12] Lin Xiao and Stephen Boyd, "Fast linear iterations for distributed averaging," *Systems & Control Letters*, vol. 53, no. 1, pp. 65–78, 2004.

[13] Ali H Sayed, Sheng-Yuan Tu, Jianshu Chen, Xiaochuan Zhao, and Zaid J Towfic, "Diffusion strategies for adaptation and learning over networks: an examination of distributed strategies and network behavior," *IEEE Signal Processing Magazine*, vol. 30, no. 3, pp. 155–171, 2013.

[14] J. Lunden, V. Koivunen, and H.V. Poor, "Spectrum exploration and exploitation for cognitive radio: Recent advances," *Signal Processing Magazine, IEEE*, vol. 32, no. 3, pp. 123–140, May 2015.

[15] Mary Ann Branch, Thomas F Coleman, and Yuying Li, "A subspace, interior, and conjugate gradient method for large-scale bound-constrained minimization problems," *SIAM Journal on Scientific Computing*, vol. 21, no. 1, pp. 1–23, 1999.

[16] Joseph Abate and Ward Whitt, "A unified framework for numerically inverting laplace transforms," *INFORMS Journal on Computing*, vol. 18, no. 4, pp. 408–421, 2006.

[17] Jan R. Magnus and H. Neudecker, "The commutation matrix: Some properties and applications," *Annals of Statistics*, vol. 7, no. 2, pp. 381–394, Mar 1979, Zbl: 0414.62040.