Abstract: Cognitive radio (CR) has been seen as a promising technology to make radio spectrum usage more effective by providing an opportunistic access for secondary users to the licensed spectrum areas. CR systems need to detect the presence of a primary user (PU) signal by continuously sensing the spectrum area of interest. Radiowave propagation effects like fading and shadowing often complicate the detection of PU because the PU signal can be weak in a particular area. By sensing the radio spectrum area of interest in a cooperative manner, the detection performance of PU signals can be increased and made more robust to channel fading. This paper studies distributed spectrum sensing in a cognitive radio context. We propose and analyse a distributed, combine and adapt type (CTA) of diffusion energy detection scheme where a central data processing unit is not needed and the required test statistics is estimated across the network. CR nodes fuse the power estimates of several neighbour nodes in the network. The PU signal is assumed to be in slow fading. The theoretical findings are verified through simulations. The proposed CTA algorithm is compared with the ring around distributed power estimation algorithm.

Key–Words: Cognitive radio, distributed estimation, distributed detection, adaptive networks, energy detection.

1 Introduction

In this paper, we study a distributed detection problem, where we have a number of nodes in the network sensing the spectrum area of interest.

The cognitive radio system is dynamic. The PU signal can be absent or present at any time. Often in practice the statistical information (for example conditional probability density of observations, prior probabilities of detection hypotheses, statistical behaviour of PU) may not available a priori for constructing a PU signal detection solution. The properties of the detection statistics may change in time.

We need to look for estimation and detection strategies which are able to react to possible changes in the properties of detection statistics and to learn the statistical information based on observations received by the nodes in the network. To reduce the computational complexity (memory requirements) and increase the learning speed, we look for methods, which support real time processing. One of the possible direction is to consider adaptive, on-line network learning methods. As new observations arrive, the estimated parameter is updated directly, without a need to re-run the network averaging process using all the observations from the past. This is a reasonable approach in cognitive radio, since we would like to avoid interference to the PU user and react to changes in a channel usage as soon as possible. On the other hand we would like to find the free spectrum opportunities as fast as possible.

Several proposed distributed spectrum sensing solutions make use of a central processing unit to collect together all the observations from all the nodes and make decisions about presence or absence of PU [1],[2],[3],[4]. We would like to remove such a node from the network. Instead we propose a power estimation solution, where the power estimates and measurements are fused in every cognitive radio network node, to allow the node to make decisions based on the data, which is available for the node. At every time instant new measurements and estimates become available, nodes in the network fuse the information and then make individual decisions about the present signal detection hypotheses.

Distributed adaptive estimation and detection schemes have been studied in several papers. The least mean square (LMS) and recursive least squares (RLS) based estimation schemes were analysed in [5],[6] and the consensus based schemes were handled in [7],[8],[9],[10]. Optimal distributed detection, based on diffusion type LMS and RLS estimation schemes, was studied in [11]. These two schemes rely on matched filter detection. However, in CR network
here we make an assumption, that we do not have any prior information about the waveform of the PU signal in the secondary nodes and hence we cannot design a matched filter. Therefore energy detection becomes a practical solution.

A ring network topology for recursive distributed energy detection without a fusion centre has been analysed in [12]. Compared to [12], in the current paper we propose a CTA diffusion based recursive calculation of the test statistics for the energy detectors. The test statistic in form of a converged power estimate is the soft information used for making the detection decision independently at every node. Such estimation strategy is able to track changes in the power of the received signal in time, as new samples are received. We propose a distributed parameter estimation model, which in Rayleigh fading channel becomes more robust to fading, compared to the case, where the test statistic is calculated over locally received signal samples only or over few nodes. The resulting energy detection performance is dependant on the performance of the recursive power estimation algorithm.

We organize the remainder of the paper as follows. In section II we review the system model and the bases of energy detection. In section III we derive an adaptive and distributed signal power estimation algorithm based on diffusion LMS strategy. In section IV we analyse the performance of the proposed distributed power estimation algorithm and the resulting energy detection. In section V we present our simulations results.

2 Distributed power estimation and detection

In classical detection theory, an energy detector can be used for detecting random signals in additive noise. For energy detection in a cognitive radio context, the type of PU signal can be completely unknown. The common assumption is that the noise power is known for being able to set the detection threshold. During a sensing time \( t \), an energy detector (ED) receives \( N \) samples of a signal \( x(n) \) from a specific frequency band [2]. The average energy of the received data samples is the test statistic \( T(x) \) of the ED, which compares \( T(x) \) to a predefined threshold \( \gamma \) and decides which of the hypotheses \( H_0 \) or \( H_1 \) is more likely.

We assume the following signal model at node \( k \):

\[
H_0 : E[|x_k(n)|^2] = E[|v_k(n)|^2] \\
H_1 : E[|x_k(n)|^2] = E[|\alpha_k|^2|s(n)|^2] + E[|v_k(n)|^2],
\]

where \( k = 1, 2, ..., K \) is the node number and \( n = 1, 2, ..., N \) is the sample index. \( v_k(n) \) is independent and identically distributed (i.i.d) circularly symmetric complex Gaussian (CSCG) noise with zero mean and variance \( E[|v_k(n)|^2] = \sigma_{v,k}^2 \), i.e. \( v(n) \sim \mathcal{CN}(0, \sigma_{v,k}^2) \).

The power of the emitted PU signal \( s(n) \) is denoted as \( E[|s(n)|^2] = S \), under \( H_1 \). The primary signal \( s(n) \) and the noise \( v_k(n) \) are assumed to be statistically independent. The PU signal passes through a slowly fading channel with gain \( \alpha_k(n) \). Note, that for implementing the energy detector, knowledge about the values of channel constants is not required. Noise power estimation is not considered in this paper. In this paper we make the following assumptions:

- (AS1) The \( x(n) \) is sensed by \( K \) nodes in the CR network.
- (AS2) The additive noise \( v_k(n) \) is uncorrelated in time and space and has the same power level over all the nodes in the CR network.
- (AS3) The number of performed iterations \( N \) is large enough.
- (AS4) The links between the CR nodes are ideal and not capacity restricted (no need to quantize the soft information).

We denote the received power estimate at node \( k \) and at iteration \( n \) as \( \hat{P}_k(n) \). For estimating the received signal power, we consider the CTA diffusion strategy [13]. In this strategy every CR node \( k \) shares the estimates (and also measurement is set so) with the neighbour nodes which are the connected to node \( k \). Every node \( k \) fuses the estimates (and measurements) from the neighbour nodes with the estimates (and measurements) from itself and updates its own estimate. In this work we assume the network topology to be fixed over the sensing time. Also we consider linear, fixed combinations of neighbour estimates and measurements at every node \( k \). We derive the CTA diffusion algorithm in three phases. First we consider local processing, when the nodes do not cooperate to estimate the received signal power jointly. Secondly we propose a global model for estimating the received signal power in cooperative manner, where all the observations are collected together to a FC for central processing. Finally we propose a fully distributed power estimation algorithm, where the nodes can observe the measurements and share the data only with their neighbour nodes.

2.1 Local and Global estimation

We start with the estimation of local received power when the nodes do not cooperate between each other.

![Image](image-url)
We are interested in estimating the parameter, denoted as $P_k^o$, in the form

$$P_k^o = E[|x_k(n)|^2].$$

(2)

By considering the standard cost function of Least Mean Square (LMS) type of filter [14], the mean square error of estimating the local received power adaptively, given the signal model (1), is

$$J_k(P) = E[|x_k(n)|^2 - P|^2].$$

(3)

Minimization of (3) with respect to parameter $P$ independently at every node $k$ results in the local solution, which is given in (2) and noted as $P_k^o$. The latter has the desired form of a test statistic of an energy detector. The local solution $P_k^o$ is optimal in the sense of minimum mean square error. $P_k^o$ will vary at every node $k$, since the expectation of $E[|x_k(n)|^2]$ varies due to channel gains. Using the standard steepest descent procedure we can find an iterative solution to the (3). The derivative of cost function (3) is

$$\nabla P J_k(P) = 2 (P - E[|x_k(n)|^2]).$$

(4)

We include the constant 2 into the step size $\mu$. Since usually the expected value in (4) is not known in practice, we replace the $E[|x_k(n)|^2]$ with its approximations $|x_k(n)|^2$ and we get the local LMS recursion

$$\hat{P}_k(n + 1) = \hat{P}_k(n) + \mu(|x_k(n)|^2 - \hat{P}_k(n)),$$

(5)

which is in the form of an exponential smoother.

According to model (1), the power of the PU signal is attenuated at every node $k$. The locally estimated power varies between nodes $k$. Therefore if the channel gain at node $k$ is low, the resulting energy detection performance is low. The result is opposite, when the node has a good channel gain. When nodes cooperate to estimate a common parameter $P^o$, the resulting detection performance can become more stable and robust to channel fading. To accomplish this purpose, we propose the following global parameter $P^o$ in the form

$$P^o = \frac{1}{K} \sum_{k=1}^{K} E[|x_k(n)|^2].$$

(6)

which is the average of the received power across the nodes $k \in K$ in the network. According to (1) and the assumption about the distribution of channel constants we observe that if the PU signal is present and when we have sufficient number of nodes in the CR network, the effect of varying channel gains is averaged out. The sum over channel gains converges close to its variance 1.

Similarly to the local cost (3), the corresponding global cost function can be given as:

$$J_{global}(P) = \sum_{k=1}^{K} E[|x_k(n)|^2 - P|^2].$$

(7)

where we have used the form of global cost as proposed in [13], [11], [5]. Minimization of the mean square error across the network (7) with respect to $P$ results in the optimal solution, which is denoted as $P^o$ and is given by (6). The observation process $|x_k(n)|^2$ is assumed to be stationary at node $k$ but the distributions of the observations vary across the nodes in the network.

An iterative solution for minimizing (7) can similarly be found using steepest descent method. The derivative of cost function (7) is

$$\nabla P J_{global}(P) = 2 \sum_{k=1}^{K} (P - E[|x_k(n)|^2]).$$

(8)

Similarly the constant 2 can be included in step size $\mu$. By replacing the moment $E[|x_k(n)|^2]$ with its approximation $|x_k(n)|^2$ we get the global LMS type recursion:

$$\hat{P}(n + 1) = \hat{P}(n) + \mu \left[ \sum_{k=1}^{K} |x_k(n)|^2 - \hat{P}(n) \right].$$

(9)

The algorithm (9) requires that all the observations are collected together to a fusion center for updating the recursion to compute a new estimate $\hat{P}(n + 1)$. Thus global information - data collected from all the nodes in the network is needed to be present for the algorithm to operate. Since the algorithm (9) is not distributed, we propose next the distributed strategy for the nodes to estimate $P^o$ based on the information what is available to the nodes.

### 2.2 Distributed Diffusion LMS estimation

We assume that $K$ nodes in the CR network are interested to estimate the scalar parameter $P^o$ in a distributed manner, where nodes can rely only on the information, what is available to them. Nodes do not have access to a global data. We need to find a way to approximate the global cost (7) in a distributed manner. The estimate $\hat{P}_k(n)$ of optimal (6) should be present at every node in the network for nodes being able independently to perform an energy detection.

The derivation of the CTA diffusion LMS type of algorithm follows the idea proposed in [6]. Let $N_k$ denote the neighbourhood group of node $k \in K$. We assume the connections between the nodes in the
neighbourhood are directed. Let us define a $K \times K$ matrix $C$, which is doubly stochastic (i.e. its rows and columns should sum up to 1). The non-negative element $c_{k,l}$ of matrix $C$ defines if a measurement from node $l$ (including node $k$) is available for node $k$. It holds that the element $c_{k,l} = 0$ if $l \notin N_k$.

By using the elements $c_{k,l}$ we can express the local cost and the corresponding local optimal solution in the neighbourhood of node $k$ as follows

$$J_k^{loc}(P) = \sum_{l \in N_k} c_{k,l} E[|x_k(n)|^2 - P|^2], \quad (10)$$

$$P_k^{loc} = \sum_{l \in N_k} c_{k,l} E[|x_k(n)|^2]. \quad (11)$$

Similarly as in [6] the global cost can be fractioned into the local cost of over the neighbourhood of node $k$ and local costs over the neighbourhood of other nodes. Using the completion of squares method to relate parameter $P$ and local optimal solution $P_k^{loc}$ the global cost function can be expressed as:

$$J_k^{glob}(P) = \sum_{l \in N_k} c_{k,l} E[|x_k(n)|^2 - P|^2] + \sum_{l \neq k} \|P - P_l^{loc}\|^2, \quad (12)$$

where we have used the fact that rows of $C$ sum up to 1. Let us define a $K \times K$ doubly stochastic matrix $B$. The non-negative element $b_{k,l}$ of matrix $B$ defines if data from node $l$ (including node $k$) is available for node $k$. It holds that $b_{k,l} = 0$ if $l \notin N_k$. With the help of elements $b_{k,l}$ the corresponding approximation of (12) in case of distributed estimation is given as

$$J_k^{dist}(P) = \sum_{l \in N_k} c_{k,l} E[|x_k(n)|^2 - P|^2] + \sum_{l \in N_k \setminus \{k\}} b_{k,l} \|P - \hat{P}_l\|^2. \quad (13)$$

In (13) the $P_l^{loc}$ has been replaced with the intermediate estimate $\hat{P}_l$ available at node $l$. The derivate of the cost function is (13) is

$$\nabla_P J_k^{dist}(P) = 2 \sum_{l \in N_k} c_{k,l} [P - E[|x_l(n)|^2]] + 2 \sum_{l \in N_k \setminus \{k\}} b_{k,l} [P - \hat{P}_l]. \quad (14)$$

The cost (13) can be used to obtain a recursion for the estimate of $P$ at node $k$, denoted as $\hat{P}_k(n)$. Using the steepest descent method, which is divided into two parts we get an iterative solution for (13) as follows:

$$\hat{\psi}_k(n) = \hat{P}_k(n) + \nu_k \sum_{l \in N_k \setminus \{k\}} b_{k,l} [\psi_l - \hat{P}_l(n)]$$

$$\hat{P}_k(n+1) = \hat{\psi}_k(n) + \mu_k \sum_{l \in N_k} c_{k,l} [E[|x_l(n)|^2 - \hat{P}_l(n)]]. \quad (15)$$

Here different step size at the nodes $k$ have been assigned and the constant 2 has been taken inside step sizes $\nu_k$ and $\mu_k$. In the first row of (15) we replace $\psi_l$ with $\hat{P}_l(n)$, which is available at time $n$. In the second row of (15) we replace $\hat{P}_k(n)$ by $\hat{\psi}_k(n)$. Thus the second row leads to

$$\hat{\psi}_k(n) = \left[1 - \nu_k \sum_{l \in N_k \setminus \{k\}} b_{k,l}\right] \hat{\psi}_k(n) + \nu_k \sum_{l \in N_k \setminus \{k\}} b_{k,l} \hat{\psi}_l(n). \quad (16)$$

Let us finally define a $K \times K$ doubly stochastic matrix $A$. The non-negative element $a_{k,l}$ defines if estimate from node $l$ (including node $k$) is available for node $k$. Thus for the elements of $A$ it holds that $a_{k,l} = 0$ if $l \notin N_k$. By taking $a_{k,k} = \left[1 - \nu_k \sum_{l \in N_k \setminus \{k\}} b_{k,l}\right]$ and $a_{k,l} = \nu_k b_{k,l}$ for $l \neq k$ we arrive to LMS type recursion what is called combine and adapt (CTA). We summarise it together with energy detection as Algorithm 1. In the CTA diffusion algorithm, the estimates

**Algorithm 1 Distributed CTA Diffusion Power Estimation**

1. **Start with** $\hat{P}_k(0) = P(0)$.
2. **Given non-negative real coefficients** $a_{k,l}, c_{k,l}$
3. **for every time instant** $n \geq 1$
4. **end for**
5. **for every node** $k = 1, ..., K$
6. **end for**

1. **Power estimation:**
2. **Detection decision:**
3. **end for**

(Refer to (43) for selecting the threshold)
\( k \) is used to calculate the new estimate \( \hat{P}_k(n + 1) \) at node \( k \), using the new observation available for node \( k \), at time instant \( n \). This is the incremental step.

3 Performance analysis

The performance analysis of the proposed algorithms is divided into three parts. First we analyse the mean and variance of CTA power estimates. Next we analyse the resulting energy detection performance. Let us note that for the theoretical analysis we need to know the values of the channel gains. For implementing the Algorithm 1 in practice, this knowledge is not required.

For more convenient notation we stack the estimates and observations from all the nodes into \( K \times 1 \) vectors as follows:

\[
\hat{P}(n) = \begin{bmatrix} \hat{P}_1(n) \\ \vdots \\ \hat{P}_K(n) \end{bmatrix}, \quad X(n) = \begin{bmatrix} |x_1(n)|^2 \\ \vdots \\ |x_K(n)|^2 \end{bmatrix}. \tag{17}
\]

Let us define additional matrix \( \mathcal{M} \) which holds the LMS algorithm step size parameters as follows

\[
\mathcal{M} = \text{diag}(\mu_1, \ldots, \mu_K). \tag{18}
\]

Then we can write the recursion in the following form

\[
\hat{P}(n + 1) = (I - \mathcal{M}) A \hat{P}(n) + \mathcal{M} C X(n). \tag{19}
\]

The initial estimate is \( \hat{P}(n) = (\hat{P}_k(0) \ldots \hat{P}_k(0))^T \).

For the CTA algorithm \( A = A_{\text{diff}} \), \( C = I \) or \( C = A_{\text{diff}} \), in case the measurements are exchanged between the nodes. Observe, that we can use the same recursion also for analysing the results in ring around topology [12], where \( A = A_{\text{ring}} \) and \( C = I \). For example when we have 3 nodes in the network, the corresponding ring around and diffusion topologies are given as follows

\[
A_{\text{ring}} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \quad A_{\text{diff}} = \begin{bmatrix} 0.5 & 0 & 0.5 \\ 0.5 & 0 & 0.5 \\ 0.5 & 0.5 & 0 \end{bmatrix}. \tag{20}
\]

Both the \( A_{\text{ring}} \) and \( A_{\text{diff}} \) are doubly stochastic. In this paper the diffusion topology is composed of local \( (A = I) \) and ring-around \( (A_{\text{diff}}) \) topologies, where we assign constant equal weights 0.5 for every allowed transition.

For evaluating the performance of the estimation algorithms and the resulting energy detection, we need to evaluate first the theoretical mean and variance of estimates \( \hat{P}(n) \).

3.1 Mean of estimation

According to signal model (1) let \( E(X(n)|H_i), i = 0, 1 \) denote the conditional mean under hypotheses \( H_0 \) (the mean when PU signal is absent) or under \( H_1 \) (PU signal is present) respectively. Considering the recursion (19), we have

\[
E(\hat{P}(n + 1)) = (I - \mathcal{M}) A E(\hat{P}(n)) + \mathcal{M} C E(X(n)), \tag{21}
\]

where the initial value is given as \( E(\hat{P}(0)) = \begin{bmatrix} \hat{P}_k(0) \ldots \hat{P}_k(0) \end{bmatrix}^T \).

We define a column vector \( w_k \), with dimension \( K \times 1 \) and which elements are zero, except the element \( k \) of vector \( w_k(k) = 1 \), i.e

\[
w_k = \text{col}(0, \ldots, 0, (w_k(k) = 1), \ldots 0). \tag{22}
\]

The conditional mean of \( \hat{P}_k(n) \) under hypothesis \( H_i \), for \( i = 0, 1 \) at node \( k \) can be found with the help of vector \( w_k \) as follows

\[
E(\hat{P}_k(n)|H_i) = w_k^T E(\hat{P}(n)|H_i) \quad \text{for } i = 0, 1. \tag{23}
\]

After iterating (21), the mean recursion can be given in the following equivalent form

\[
E(\hat{P}(n)) = [(I - \mathcal{M}) A]^n \hat{P}(0) + \sum_{i=0}^{n-1} [(I - \mathcal{M}) A]^i \mathcal{M} C E(X(n)). \tag{24}
\]

We are interested in the mean of the estimates in steady state, when the filter has converged, i.e when \( n \to \infty \). Thus according to (24) we need to analyse the asymptotic behaviour of the power of matrix \( [(I - \mathcal{M}) A]^n \) and the limit of geometric series \( \sum_{i=0}^{n-1} [(I - \mathcal{M}) A]^i \). Considering [15, Theorem 5.6.12], the power of matrix \( [(I - \mathcal{M}) A] \) converges asymptotically to zero if the matrix is stable. The matrix is stable if and only if the eigenvalues \( \lambda_i \) of matrix \( [(I - \mathcal{M}) A] \) are strictly inside the unit circle, i.e

\[
|\lambda_k (I - \mathcal{M}) A| < 1 \quad \text{for all } k = 1 \ldots K. \tag{25}
\]

Thus given the diffusion strategies with a doubly stochastic matrices \( A \) and \( C \) the convergence of power of the matrix \( [(I - \mathcal{M}) A] \) to zero is dependant on the selected step sizes in matrix \( \mathcal{M} \). The choice of step sizes should guarantee that the condition (25) holds.

In the recursion the matrix \( A \) is equal to doubly stochastic matrix \( A_{\text{ring}} \), \( A_{\text{diff}} \) or identity matrix \( I \). The matrix \( C \) is doubly stochastic or equals to \( I \) if no
measurements are fused. According to matrix spectral norm we can write for CTA algorithm that
\[ \| (I - \mathcal{M}) A \|_2 \leq \| (I - \mathcal{M}) \|_2 \| A \|_2. \]  
(26)
The spectral norm of doubly stochastic matrices \( A, C \) and identity matrix \( I \) is 1. Since the matrix \((I - \mathcal{M})\) is symmetric, we have that \( \| (I - \mathcal{M}) \|_2 = |\lambda_{max}((I - \mathcal{M}))| \). If the matrix on the LHS of (26) is also symmetric, then we can also replace the LHS with \( \| (I - \mathcal{M}) A \|_2 = |\lambda_{max}((I - \mathcal{M}) A)| \). We have
\[ |\lambda_{max}((I - \mathcal{M}) A)| \leq |\lambda_{max}(I - \mathcal{M})|. \]  
(27)
We conclude that in general if the step sizes \( \mu_k \) in matrix \( M \) are selected so that the spectrum of matrix \((I - \mathcal{M})\) is inside the unit circle, it holds also that the matrix \((I - \mathcal{M}) A\) is stable.

For the selection of \( \mu_k, k = 1 \ldots K \) so that the diagonal matrix \((I - \mathcal{M})\) is stable, we have the following condition
\[ |\lambda_k((I - \mathcal{M}) A)| = |1 - \mu_k| < 1 \quad \text{for all } k=1 \ldots K. \]  
(28)
Since in our model we have only one mode of convergence of the filter [14], \( \mu_k \) should be selected in the range:
\[ 0 < \mu_k < 2. \]  
(29)
The geometric series \( S_n = \sum_{i=0}^{n-1} [(I - \mathcal{M}) A]^i \) is generated by matrix \([(I - \mathcal{M}) A]^i\) and converges if and only if the condition (25) holds for all \( \lambda_i \). When it holds we can write the geometric series as follows
\[ S_n = [I - [(I - \mathcal{M}) A]]^{-1} [I - [(I - \mathcal{M}) A]^n]. \]  
(30)
Hence according to (25) as \( n \to \infty \) the power of matrix \([(I - \mathcal{M}) A]^n\) converges to zero. Thus by using the covered result of the geometric series and by noting the mean of \( \hat{P}(n) \) in steady state and under both hypotheses \( H_i, i = 0, 1 \) as \( E(\hat{P}(\infty)|H_i) \), we can write
\[ E(\hat{P}(\infty)|H_i) = [I - [(I - \mathcal{M}) A]]^{-1} \times \mathcal{M} E[X(n)|H_i], \]  
(31)
where the conditional expectations of observations are given by (1). Similarly to (23) we have that the mean of \( \hat{P}_k(n) \) in steady state is
\[ E(\hat{P}_k(\infty)|H_i) = w_k^T E(\hat{P}(\infty)|H_i) \quad \text{for } i=0,1. \]  
(32)

### 3.2 Variance of estimation
To find the recursion for the variance of \( \hat{P}_k(n) \), at node \( k \in K \), we start from the recursions (19), (21) and derive first the recursion for covariance matrix \( \text{Cov}(\hat{P}(n)) \). The covariance of \( \hat{P}(n) \) is defined as
\[ \text{Cov}(\hat{P}(n)) = E \left( \left( \hat{P}(n) - E[\hat{P}(n)] \right) \left( \hat{P}(n) - E[\hat{P}(n)] \right)^T \right). \]  
(33)
Let us note the conditional covariance of estimates under the hypothesis \( H_i, i = 0, 1 \) as \( \text{Cov}(\hat{P}(n+1)|H_i) \).
Similarly let \( \text{Cov}(X(n)|H_i) \) denote the conventional covariance of observations. After substituting (19) and (21) into (33), taking expectation and considering the fact that \( \hat{P}(n) \) is independent of observation vector \( X(n) \), it can be shown that the covariance recursion is
\[ \text{Cov}(\hat{P}(n+1)|H_i) = (I - \mathcal{M}) A \text{Cov}(\hat{P}(n)|H_i) \]  
\[ + \mathcal{M} C \text{Cov}(X(n)|H_i)C^T \mathcal{M}. \]  
(34)
where initial estimate of covariance matrix is noted by \( \text{Cov}(\hat{P}(0)|H_i), i = 0, 1 \). The covariance matrix of observations \( \text{Cov}(X(n)|H_i) \) is constant over time \( n \).

The covariance matrix \( \text{Cov}(X(n)|H_i) \) of \( K \times K \) has the following structure. When PU signal is present the main diagonal elements of matrix \( \text{Cov}(X(n)|H_i) \) - the variances of observations at node \( k \in K \) can be shown to:
\[ \text{Var}(|x_k(n)|^2|H_i) = (|\alpha_k|^2\sigma_x^2 + \sigma_v^2)^2. \]  
(35)
When PU signal is not present and according to Assumption 2 the variances of observations at node \( k \in K \) are given as
\[ \text{Var}(|x_k(n)|^2|H_0) = \sigma_v^4. \]  
(36)
When PU signal is present, the off diagonal elements of matrix \( \text{Cov}(X(n)|H_i) \) - the covariance of observations at nodes \( k \) and \( j \) if \( k, j \in K \) and \( i \neq j \) can be shown to:
\[ \text{Cov}(|x_k(n)|^2, |x_j(n)|^2|H_i) = |\alpha_k|^2|\alpha_j|^2\sigma_x^4. \]  
(37)
According to Assumption 2 the noise realizations \( v_k(n) \) and \( v_j(n) \) are uncorrelated in time and space for \( k, j \in K \) and \( i \neq j \). Thus when PU signal is absent the covariance of observations is
\[ \text{Cov}(|x_k(n)|^2, |x_j(n)|^2|H_0) = 0. \]  
(38)
for $k, j \in K$ and $i \neq j$.

The variance of $\hat{P}_k(n)$ at node $k$ under the hypothesis $H_i$, $i = 0, 1$, can be found by multiplying the recursion (34) with vector $w_k$ from the left and with vector $w_k$ from the right

$$\text{Var}(\hat{P}_k(n+1)|H_i) = w_k^T (I - \mathcal{M}) A \text{Cov}(\hat{P}(n)|H_i) \times A^T (I - \mathcal{M}) w_k + w_k^T \mathcal{M} \text{Cov}(X(n)|H_i) \times C^T \mathcal{M} w_k. \quad (39)$$

To derive the steady state value for $\text{Var}(\hat{P}_k(n)|H_i)$ we note that (34) is in a form of discrete time algebraic Riccati equation (DARE), [16, App. E]. Due to space constraints we skip the derivation details and note that the steady state variance $\text{Var}(\hat{P}_k(\infty)|H_i)$, $i = 0, 1$, at node $k \in K$ can be recovered by selecting the $\{k, k\}$ element of the steady state covariance matrix $\text{Cov}(\hat{P}(\infty)|H_i)$, which has been found as a solution to DARE. We have finally

$$\text{Var}(\hat{P}_k(\infty)|H_i) = w_k^T \left[ \text{Cov}(\hat{P}(\infty)|H_i) \right] w_k. \quad (40)$$

### 3.3 Detection Performance Analysis

As mentioned earlier the test statistic of the energy detector at node $k$ at time instant $n$ is estimated using CTA signal power estimation algorithms. Thus the resulting detection performance is dependant on the performance of the underlying estimation process. For deriving the formulas of probability of detection ($P_D$) and probability of false alarm ($P_{FA}$) we need to evaluate the probability density function (PDF) of the test statistic $\hat{P}_k(n)$ under both hypotheses $H_0$ and $H_1$.

Since the input signal is CSCG and in case $K = 1$, the test statistic of ED $\hat{P}_k(n)$ is local and under both hypothesis a Chi-Square distributed random variable with 2N degrees of freedom. The test statistic $\hat{P}_k(n)$ is obtained as a sum of a number of identically distributed variables and hence the CLT can be applied to approximate the Chi square distribution by a Gaussian distribution [17]. According to AS3 the number of samples is large enough, and the CLT is expected to apply.

The global test statistic $\hat{P}_k(n)$ in case of hypothesis $H_1$, is however estimated over independent, but not identically distributed variables. In such a case the Lyapunov CLT [18] still be applied over a large number of samples to result in a Gaussian approximation. We found in previous section that the variance $\hat{P}_k(n)$ in steady state is bounded. In [12] the formulas for the $P_{FA}$ and $P_{D,k}$ of the energy detector have been derived.

Using these results and by taking into account the (23) and (39), we provide approximate formulas for the resulting energy detection performance.

The probability of false alarm $P_{FA}$ of the energy detector under hypothesis $H_0$ (using the theoretical mean and variance of estimates under $H_0$) is given as follows

$$P_{FA} = Q \left( \frac{\gamma - E(\hat{P}(n)|H_0)}{\sqrt{\text{Var}(\hat{P}(n)|H_0)}} \right). \quad (41)$$

Based on the assumption AS2, we observe that the formula holds for every node $k \in K$.

The probability of detection of an energy detector under hypothesis $H_1$ (using the mean and variance of estimates under $H_1$) is correspondingly given as follows

$$P_{D,k} = Q \left( \frac{\gamma - E(\hat{P}(n)|H_1)}{\sqrt{\text{Var}(\hat{P}(n)|H_1)}} \right). \quad (42)$$

The sensing threshold is found from (41) by fixing the desired value of $P_{FA}$. Thus

$$\gamma = E(\hat{P}(n)|H_0) + Q^{-1}(P_{FA}) \sqrt{\text{Var}(\hat{P}(n)|H_0)}. \quad (43)$$

Due to the assumption AS2 the thresholds for every CR node $k$ are equal.

However calculation of the threshold requires knowledge about the moments of the estimation algorithm under hypothesis $H_0$ and these moments are dependant on the algorithm parameters (especially on the step size). In practice for the threshold calculation, the required moments can be calculated in advance using (23) and (39), known values of step size, noise power and then inserting these results into (43).

### 4 Simulation results

In the numerical simulation section we investigate the CTA type of power estimation algorithm (with constant step size). We compare the results with the previously proposed ring around [12] power estimation algorithm. Secondly we view the resulting energy detection performance of the CTA diffusion algorithm and compare to the ring around algorithm. In all these simulations the PU signal $s(n)$ is taken as QPSK with unit power $S$, under the active hypothesis $H_1$, the step size is: $\mu = 0.01$. 

---

**Note:** The above text is a natural representation of the document content. It has been formatted to ensure readability and coherence. The page number, ISBN, and any graphical elements are not included in the natural text representation.
4.1 Local and distributed power estimation

We start with the investigation of the estimation algorithms. The channel gain is assumed to be constant and obtained by: \( \alpha_k \sim \mathcal{C}N(0, 1) \). We use the same channel gain values for all the algorithms. All the nodes in the network receive \( N = 2000 \) samples. During samples \( n = 1 \ldots 1000 \) the PU signal with constant unit power \( S \) is present. The power \( S \) is attenuated by the channel gain \( |\alpha_k|^2 \). In sample range \( n = 1001 \ldots 2000 \) the PU signal is absent and only noise is present. Under both detection hypothesis the noise power is \( \sigma_n^2 = 1 \) and assumed to be the same in all the nodes. In the following simulations no measurements are exchanged thus, \( C = I \). When nodes do not cooperate, then the estimation results are highly dependent on the given channel gains and therefore vary across the nodes in the network. When nodes cooperate using to ring around topology, then the corresponding power estimates are given in the Fig. 1. All the estimated power values in the CR network of 10 nodes are plotted in one figure. In addition the optimal solution \( P^* \) has been calculated according to (6) using the given channel gains values and is added in the figure as a back dashed line. We see that the estimates of \( P^* \) fluctuate around the optimal solution, under both active detection hypotheses. Let us note, that it can be numerically verified, that the theoretical mean and variance of the ring around estimates, which are calculated using the formulas (23) and (39), match with the results, which are found using the formulas [12, Eq. 5] and [12, Eq. 10].

The CTA diffusion algorithm the estimates of the received power together with optimal solution \( P^* \) have been plotted in Fig. 2. Compared to the ring round, the variance of estimates of CTA algorithm is lower than the variance of estimates of ring around algorithm. We can conclude that the preciseness of power estimates increases when CTA algorithm is used. Thus also the resulting detection performance increases, what we show in next subsection.

4.2 Probability of detection

Next we investigate the probability of detection using the proposed distributed CTA power estimation algorithm. We compare the performance of 5 different network sizes: \( K = 1, 3, 10, 30, 50 \) nodes. The estimated and theoretical results of \( P_D \) of the last nodes in the set are compared, i.e at nodes \( k = K \). In the simulations the converged power estimate is used for detection i.e \( \hat{P}_k(\infty) \). The theoretical mean and variance of power estimates can hence be calculated directly using the steady state formulas (32), (40). The mean and covariance of the observation vector is taken under the detection hypothesis \( H_1 \) and the choices of values of matrices \( A \) and \( C \) to define CTA or ring around algorithms.

We set the desired \( P_{FA} = 10^{-4} \). The thresholds of the energy detectors at nodes \( k \in K \) are calculated using (43) and by using the corresponding theoretical steady state mean and variance of the power estimates under detection hypothesis \( H_0 \).

For estimating the \( P_D \) we use the Monte Carlo method [19]. The estimated \( P_D \) is compared with the theoretical \( P_D \). The latter is calculated using (42) and using the corresponding steady state mean and variance of the power estimates of the two algorithms under detection hypothesis \( H_1 \).

First we set \( C = I \). The detection performances of the ring around, and CTA algorithms are shown in Fig. 3, and in Fig. 4 respectively. We see that there is
a good match between estimated and theoretical $P_D$. Due to the smaller variance, the CTA algorithm outperforms the ring around algorithm. As the number of nodes in the network increases, about 4 dB is gained with respect to the noise power.

When also measurements from a neighbour nodes are available and we set $C = A_{diff}$ for CTA algorithm. The result is shown in Fig. 5. We see minor increase in the detection performance when additionally measurements are exchanged between the nodes. When the data transfer and processing capacity at the nodes is limited (energy constants etc), then the measurement exchange does not give significant improvement in resulting detection performance. However by fusing more estimates compared to the simplest ring around algorithm, we see notable improvement in resulting detection performance.

![Figure 3: Probability of detection, ring around, $C = I$](image3.png)

![Figure 4: Probability of detection, CTA, $C = I$](image4.png)

![Figure 5: Probability of detection, CTA topology, $C = A_{diff}$](image5.png)

## 5 Conclusion

In this paper we proposed a diffusion based distributed power estimation approach, that is applicable for CR networks for detecting the presence of PU signal. We derived CTA diffusion based power estimation algorithm for energy detection. The performance analysis of the derived algorithm was carried out and simulations were run. It was shown that the CTA diffusion power estimation algorithm outperforms the previously proposed ring around algorithm, while the effect of exchanging also measurements is rather small. The proposed algorithm is able to track changes in received signal power and is usable in cognitive radio systems.

**Acknowledgements:** This paper was partially supported by Estonian scholarship program Kristjan Jaak, which is funded and managed by Archimedes Foundation in collaboration with the Ministry of Education and Research of Estonia.

**References:**


