Article

An Efficient Distributed Approach for Cooperative Spectrum Sensing in Varying Interests Cognitive Radio Networks

Maria Trigka * and Elias Dritsas

Department of Computer Engineering and Informatics, University of Patras, 26504 Patras, Greece
* Correspondence: trigka@ceid.upatras.gr

Abstract: The rapid growth in wireless communications, coupled with insufficient utilization of the spectrum, led to the development of new wireless services and the promising technology of cognitive radio (CR) networks, which facilitate periodic access to the unoccupied spectrum bands and thus increases spectral efficiency. A fundamental task in CR networks is spectrum sensing, through which unauthorized secondary users (SUs) detect unoccupied bands in the spectrum. To achieve this, an accurate estimate of the power spectrum is necessary. From this perspective, and given that many other factors can affect individual detection, such as pathloss and receiver uncertainty, we aim to improve its estimate by exploiting the spatial diversity in the SUs’ observations. Spectrum sensing is treated as a parameters estimation problem, assuming that the parameters’ vector of each SU consists of some global and partially common parameters. To exploit this modeling, distributed and cooperative spectrum sensing is the subject of interest in this study. Diffusion techniques, and especially the Adapt-Then-Combine (ATC) method will be exploited, where each SU cooperates with a group of nodes in its neighborhood that share the same parameters of interest. We consider a network of three static PUs with overlapping power spectrums, and thus, frequency bands. The performance of the employed method will be evaluated under two scenarios: (i) when the PUs spectrum varies, since some frequency bands are not yet utilized, and (ii) when the frequency bands of the PUs are fixed, but there is a mobile SU in the network, changing regions and parameters of interest. Experimental results and performance analysis reveal the ATC algorithm robustness and efficiency.

Keywords: cognitive radio; spectrum sensing; diffusion strategies; adapt-then-combine; cooperation

1. Introduction

The radio frequency spectrum is an inherently limited resource. Many services provided to users every day are based on the broadcast of radio signals (radio, television, telephony, broadband internet, defense applications (e.g., radar), scientific uses such as meteorology, positioning (GPS), etc.). New technologies, such as the Internet of Things and Machine-to-Machine communication, appear that rely on the use of wireless radio systems (UMTS, Wi-Fi, WiMax, Digital TV, 5G, and B5G). Additionally, the transition to multimedia applications, and in general, the rapid growth in wireless communications have contributed to a great demand for the development of new wireless services for both the authorized and unauthorized frequency spectrums [1].

At the same time, the electromagnetic spectrum, because of greater and greater increases in demand, is in short supply. Most of it is occupied by already existing services and applications. Therefore, to assign a part of the electromagnetic spectrum to a new service, an existing one should be replaced, cease its operation, and thus free the requested spectrum band. However, the electromagnetic spectrum will still be fully occupied. On the other hand, a significant part of the already allocated spectrum has limited or sporadic usage, thus leading to its underuse.
All of the above create the requirement for continuously available spectrum frequencies without interference. Recent studies have shown that the current fixed spectrum assignment policy has imposed insufficient use of the spectrum. In recent years, dynamic spectrum management in 5G networks is more necessary, due to the increasing demand for free frequencies solving the problem of spectral inadequacy [2]. To address this problem, Cognitive Radio Networks have been brought to the foreground as a promising technology that facilitates periodic access to unoccupied frequency bands, contributing to the increase in spectrum efficiency [3].

According to the Federal Communications Commission (FCC), cognitive radio is a system that senses its operating electromagnetic environment and can dynamically and autonomously adjust its radio operating parameters to modify the operation of the system, such as to maximize the throughput, mitigate interference, facilitate interoperability and access the secondary networks. Hence, one main purpose of cognitive radio is the autonomous exploitation of the locally unused part of the spectrum (spectrum holes) with the aim of providing new access paths to it [4].

The main functions of the cognitive radio networks (spectrum sensing, management, mobility, sharing) are represented in a cycle called the cognitive radio cycle [5]. Cognitive Networks manage the problem of congestion frequencies by introducing the occasional use of frequency bands that are not occupied by authorized users and are known as spectrum holes. The CR network users can share the spectrum with primary users. Primary users have higher priority or authorization in the use of an available piece of the spectrum, and use the service to which the specific bandwidth is assigned, thus maximizing the degree of spectrum utilization [6]. Secondary users of the cognitive network have a lower priority and exploit the available spectrum in such a way that they do not create interference with primary users. The fundamental work of each of the SUs is to detect and identify the presence or absence of authorized users, known as primary users. It is usually achieved by sensing the RF environment through a process called spectrum sensing. The way of measuring and exploiting the spectral space is called spectrum coverage, which is stated as being a frequency band that is not used by its PU, at a specific time, in a specific geographic area [7]. CR networks, along with device-to-device communication, are two key technologies for efficient spectrum sharing, especially in next-generation networks [8,9].

In the current study, we will be concerned with cooperative sensing, and especially distributed learning and distributed estimation methods about intelligent wireless sensor networks similar to [10]. Specifically, our aim is to estimate the total spectrum of three PUs with overlapping frequency bands. Capitalizing on the relevant literature, the problem has been previously studied for either independent [11] or overlapping spectrums among two PUs [12]. Additionally, the main contribution of the presented analysis is the evaluation of the ATC diffusion-based method in more challenging and practical scenarios, where some frequency bands stop being utilized by some PUs, or where there is a mobile SU node in the network. The experimental results verify the efficiency of the method in such challenging conditions.

The rest of the paper is organized as follows. Section 2 describes the relevant spectrum sensing works with the subject under consideration. Furthermore, in Section 3, an analysis of the system model and the adopted methodology followed are made. In addition, in Section 4, we discuss the acquired research results. Finally, conclusions and future directions are outlined in Section 5.

The following notations are considered in this paper. Uppercase and lowercase bold letters denote matrices and vectors, respectively. \( \cdot^T \) and \( \cdot^H \) denote the transposition and the complex conjugate transposition of a matrix, \( I \) is the identity matrix, \( \lambda_{\text{max}}(\cdot) \) is the maximum eigenvalue of a matrix, \( \| \cdot \| \) denotes the Euclidean norm of a vector, \( E\{\cdot\} \) stands for the statistical expectation, \( A \otimes B \) captures the Kronecker product among the matrices, \( A \odot B \) captures the element-wise product among the matrices, \( A_{M \times N \times L} \) denotes a three-dimensional matrix of size \( M, N, L \), \( \text{size}(A) \) captures the dimensions (rows and
columns) of matrix $A$, $n = \text{find}(A(i,:)) \neq 0$ returns the columns of the $i$-th row of $A$ with non-zero elements, and $|n|$ stands for the length of vector $n$.

2. Spectrum Sensing Methods

In this section, we will present the most common spectrum detection techniques where each user individually decides if some frequency band in the spectrum is available for use. In particular, we will focus on the most popular in wireless communications, such as Energy Detector (ED), Cyclo-stationary Feature Detection (CFD), Compressive Detection (CD), and Matched Filter and Waveform-based technique [13,14]. Furthermore, we will present recent works on spectrum sensing exploiting the capabilities of machine/deep learning. Additionally, the section concludes with cooperative approaches.

2.1. Traditional Methods

An energy detector is a non-coherent method that identifies the primary signal based on the detected energy. It is the most common method due to its low computational and implementation complexity. Because of its simplicity and non-requirements for prior knowledge of the primary signals, the energy detector is the most popular technique in cooperatives spectrum sensing. The signal is detected by comparing the output of the detector to a threshold, which depends on the noise level [15]. However, the detection time to achieve a specific detection probability may be high. In addition to, the detection performance is subject to noise power uncertainty. An energy detector cannot be used to distinguish between the primary and CR signals. As a result, CR users need to be strongly synchronized and not transmit to the quiet period [6,16].

Cyclo-stationary Feature Detection (CFD) exploits the periodicity of the received PU signal to identify its presence. It is a detection method that takes advantage of the cyclo-stationary characteristics of the received signal, which is usually integrated into (i) sinusoidal carrier frequencies, (ii) pulse sequences, (iii) signal periodicity, or (iv) its statistical properties, such as mean value and autocorrelation [17]. Thus, CFD can distinguish between a telecommunication signal and the noise due to robustness at noise uncertainty, and performs better than the energy detector at low SNR. Although this requires prior knowledge of the signal characteristics, CFD can distinguish CR transmissions from various PU signals. It eliminates the requirement to synchronize the energy detector in cooperative detection. A weakness of this method is the high computational cost and long detection time. Because of these issues, it is less common than energy detectors in cooperative detection.

In the case of spectrum underutilization, compressive sensing techniques for cognitive radio networks can be used to approximate and recover the detected spectrum, which facilitates the detection of sparse primary signals in the wideband spectrum [18]. Hence, the compressive detection techniques provide promising solutions for the fast recovery of wideband signals and facilitate wideband detection with acceptable computational complexity. In wideband cooperative compressive detection, SU nodes individually perform compressive detection and cooperatively estimate the wideband spectrum by exchanging spectrum estimates, and iteratively reach a collective decision by exchanging local decisions [19].

Matched Filter [20] is the optimal method of detecting PUs signals when the transmitted signal is known. The main advantage is the short required time for achieving a certain probability compared to other methods that are discussed. The required number of samples increases with $O(1/\text{SNR})$ for a given false alarm probability $P_{fa}$ to low SNR. However, this requires the SU to modulate the received signals. Therefore, this requires perfect knowledge of the signal characteristics of PUs, such as bandwidth, operating frequency, modulation type, pulse shape, and format framework [21].

Waveform-based is usually used in wireless systems with known standards to help synchronization, or for other purposes. Such standards include preambles, midambles, spread-out sequences, etc. The preamble is a sequence transmitted before each burst.
The midamble is transmitted in the middle of a burst or a time slot. It uses the same model as the energy detector. In the presence of a known pattern, detection can be performed by correlating the received signal with a known copy of it. This method only applies to systems with known signal patterns and is called coherent detection. This method is more reliable, has a better time convergence, and the performance of the algorithm increases with the length of the standard signal [21].

2.2. Recent Methods and Motivation

Apart from traditional spectrum sensing methods, researchers have turned to machine/deep learning [22], and Reinforcement and Federated Learning for efficient spectrum sensing. A reliable machine learning-based spectrum sensing in CR networks is presented in [23]. Wang et al. surveyed dynamic spectrum allocation based on reinforcement learning algorithms in cognitive radio networks [24]. Additionally, in [25], it is suggested that a machine-learning-based opportunistic spectrum access approach exploits multi-armed bandit (a powerful reinforcement learning tool) and matching theory. In [26], machine learning techniques (support vector machine (SVM), random forest (RF), and K-nearest neighbors (KNN)) are used to detect a PU in Mobile CR networks.

Classical local techniques can also be part of a cooperative framework for spectrum sensing, which is usually employed to increase the detection reliability in either a centralized (e.g., Fusion Center), decentralized, or relay-assisted manner [27–29]. In [30,31], cooperative spectrum sensing methods are suggested based on an energy detector. Alternative methods perform spectrum sensing in CR networks based on adaptive LMS-based distributed cooperative methods [12,32,33]. Recently, in [34], spectrum sensing is solved via federated learning, which is exploited for data distribution and model training over many devices. The proposed method creates a common deep learning model for each user group in the iterative process.

In the era of 5G networks, a high trend toward distributed and cooperative schemes has emerged for important operations in the physical layer, with some remarkable examples being spectrum sensing [35], beamforming [36], and channel estimation in the mmWave spectrum [37,38]. Due to the persistent increase in smart devices, and the need for high data rates with low latency, 5G and beyond networks have become more important than ever. Essential technology for its successful implementation is direct Device-to-Device (D2D) communication (e.g., for extending capacity and enabling scalability) with Internet-of-Things (IoT) being a characteristic short-range application that will highly benefit from the capabilities of 5G networks. Focusing on spectrum sensing in a 5G IoT network, geographically distributed SUs can sense multiple channels owned by PUs and decide via cooperation on available spectrum holes using learning algorithms (i.e., centralized and distributed).

In this study, distributed cooperative spectrum sensing will concern us, since such schemes are adaptable and more robust against changes in topology, especially in mobile and scalable IoT systems [39]. Focusing on a mobile CR network with overlapping spectrum PUs, a Bayesian machine learning approach for collaborative spectrum sensing has been suggested in [40]. In a more recent study [41], the authors, considering clusters of unmanned aerial vehicles, elaborated a diffusion-based distributed cooperative spectrum sensing approach with adaptive weights, assuming a binary state model to describe channel occupancy.

In the same direction, this study focuses on the problem of multi-task [42] parameter estimation, assuming varying spectrum sensing and adopting a distributed diffusion-based approach from a well-established family of techniques, which are preferable in wireless sensor networks due to lower communication overhead and more improved performance, compared with the centralized and non-cooperative strategies. In particular, we exploit the distributed diffusion method ATC to cooperatively estimate the aggregated spectrum, assuming a more complex scenario of three PUs with overlapping spectrums. Hence, multiple clusters of users are derived, each of which relates to a common interest frequency.
band. According to [11], a multi-task CR network is formulated. Initially, we evaluate the efficiency of ATC, assuming that the spectrum of PUs is varying (i.e., we cease the use of some frequency bands) and that secondary users are static. Finally, an interesting aspect of this work constitutes the performance evaluation of the same method in the case where one SU node is mobile.

3. Materials and Methods

In this section, we will present a description of the system model that approximates the total power spectrum in a specific SU node, the main aspects of the adopted solution method, and the definition of the necessary structures for the selected case study.

3.1. System Model

We consider \( Q \) PUs and \( K \) SUs. The power spectrum emitted by each PU can be captured as a linear combination of some basis functions. Here, we choose Gaussian basis functions. Each SU, through spectrum sensing, essentially detects the total spectrum from all of the PUs of the area. The power spectrum from PU \( q \) is written as [43,44]

\[
S_q = \sum_{m=1}^{B} b_{qm} f_m(e^{j\omega})
\]

where \( f_m(e^{j\omega}) = e^{-\frac{(\omega - \omega_m)^2}{2\sigma^2_m}} \) and parameters \( \omega_m, \sigma_m \) are the central frequency and standard deviation (these parameters are part of the system design and thus a priori known), \( f_\omega = [f_1(e^{j\omega}), f_2(e^{j\omega}), f_3(e^{j\omega}), \ldots, f_B(e^{j\omega})] \) is a vector with the basis functions, scalars \( \{b_{qm}\} \) denote the coefficients of the basis expansion for user \( q \), and \( w_q = [b_{q1}, b_{q2}, \ldots, b_{qB}]^T \) is the vector with the factors involved in the linear combination of the basis functions. The representation in (1) can well approximate an essential part of the power spectrum if \( B \) is sufficiently large.

The power spectrum that a SU \( k \) detects by PU \( q \) is subjected to attenuation due to the propagation path loss denoted as \( p_{qk} \). Path loss coefficients are known and defined in advance in a training stage between PUs with each SU. Training is usually repeated at regular intervals because the coefficients may change (slowly) in time due to node movement. When the transmitted spectrum travels from the PU to the SU, then the power spectrum that is measured by the receiver of the SU \( k \) is \( p_{qk}S_q(e^{j\omega}) \).

Hence, the total power spectrum from all PUs at SU \( k \) is written as

\[
S_k^r = \sum_{q=1}^{Q} p_{qk} S_q(e^{j\omega}) + \sigma_k^2
\]

where \( \sigma_k^2 \) is the vector with the factors involved in the linear combination of the basis functions. The representation in (1) can well approximate an essential part of the power spectrum if \( B \) is sufficiently large.

\[
d_{k,r}(i) = u_{k,\omega_r} w_k^r + \sigma_k^2 + v_{k,r}, r = 1, 2, \ldots, L
\]
\[ d_{k,i} = \begin{bmatrix} d_{k,1}(i) - \sigma_k^2 \\ d_{k,2}(i) - \sigma_k^2 \\ \vdots \\ d_{k,L-1}(i) - \sigma_k^2 \\ d_{k,L}(i) - \sigma_k^2 \end{bmatrix}, \quad v_{k,i} = \begin{bmatrix} v_{k,1}(i) \\ v_{k,2}(i) \\ \vdots \\ v_{k,L-1}(i) \\ v_{k,L}(i) \end{bmatrix} \] (4)

\[ u_{k,\omega_1} \\ u_{k,\omega_2} \\ \vdots \\ u_{k,\omega_{L-1}} \\ u_{k,\omega_L} \] = \begin{bmatrix} p_{k,i} \otimes f_{\omega_1} \\ p_{k,i} \otimes f_{\omega_2} \\ \vdots \\ p_{k,i} \otimes f_{\omega_{L-1}} \\ p_{k,i} \otimes f_{\omega_L} \end{bmatrix} (L \times Q \cdot B) \] (5)

\[ d_{k,i} = U_{k,i}w_k^i + v_{k,i} \] (6)

where \( v_k \) is the measurement and/or model noise with zero mean and covariance matrix \( R_v \) of dimensions \( L \times L \). We have taken measurements at \( L \) different frequencies, and for this reason, the matrix has \( L \) rows. As a result, in (6), a linear model is obtained for estimating the parameters of interest in \( w_k^i \). All of the considered processing steps are also illustrated in Figure 1.

The path loss factor is approximated according to the formula

\[ p_{qk,i} = \left( \frac{d_{qk,i}}{d_0} \right)^n \] (7)

where \( d_{qk,i} \) is the Euclidean distance of SU \( k \) to PU \( q \) at \( i \), \( d_0 \) is the reference distance which is \( d_0 = 1 \), and \( n \) models the attenuation environment in the network. Hence, the path loss values between SU \( k \) and the \( Q \) PUs is captured in

\[ p_{k,i} = [p_{1k,i}, p_{2k,i}, p_{3k,i}, \ldots, p_{Qk,i}] \].

In the estimation of \( p_{k,i} \), we consider a respective measurement Gaussian noise of zero mean and standard deviation \( \sigma_p \); thus, \( \hat{p}_{k,i} = p_{k,i} + n_k \). When the SU \( k \) moves, the values of \( p_k \) change as its distance from the PUs changes, and consequently its values.

**Figure 1.** Processing steps involved in the system model.

To estimate the spectrum, it is sufficient to estimate the parameters vector that multiplies the basis functions. By relying on the network data \{\( d_{k,i}, U_{k,i} \}\), we treat the problem
as a parameter estimation of multiple interests and assume cooperation among the nodes to process the data in a distributed fashion based on a diffusion strategy, e.g., Adapt-Then-Combine (ATC). The aforementioned strategy well approximates the centralized solution when all nodes want to estimate the same vector of parameters \([12]\).

Each vector \(\{w_k^j\}_{k=1}^K\) might consist of parameters of global interest to the whole network, parameters of common interest to subsets of nodes including node \(k\), and parameters of local interest for node \(k\). In particular, subsets of parameters in \(w_k^j\) may account for:

- One global parameter vector related to the frequency band in the power spectrum of all PUs, which affects all SUs in the network.
- \(J\) common parameter vectors associated with frequency bands in the power spectrum of PUs that affect specific subgroups of nodes with partially or fully overlapped common interests.

Finally, in this study, we have not considered parameters of local interest.

In this way, considering a scenario where there are \(J\) different subsets of common parameters, the observation model provided in (6) can be reformulated as

\[
d_{k,i} = U_{k,i}w_k^j + \sum_{j \in l_k} U_{k,i,j}\varsigma_{k,j} + v_{k,i}\tag{9}
\]

where each node aims to solve through cooperation the following optimization problem

\[
\arg\min_{K} \sum_{k=1}^K E\{\|d_{k,i} - U_{k,i}w_k^j - \sum_{j \in l_k} U_{k,i,j}\varsigma_{k,j}\|^2\} \tag{10}
\]

with respect to \(w_k^j, \varsigma_{1,j}, \varsigma_{2,j}, \ldots, \varsigma_{J,j}\), where \(l_k\) is an ordered set of indices \(j\) associated with the vectors \(\varsigma_j\) that are of interest to node \(k\), \(U_{k,i}\) are matrices of dimensions \(L \times M_j\), \(L \times M_{kj}\) that consist of the columns of \(U_{k,i}\) associated with the \(w_k\) and the \(\varsigma_{k,j}\), respectively. Some statistical independence assumptions \([12]\) on the data are the following:

- \(v_{k,i}\) is temporally and spatially white noise, whose covariance matrix is \(R_{v_{k,i}} = E\{v_{k,i}v_{k,i}^H\}\), and which is independent of \(U_{k,i}\) for all \(k\) and \(i\), with \(k \in \{1,2,\ldots,K\}\) and \(i > 0\);
- \(U_{k,i}\) is independent of \(U_{k,j}\) with \(i, j > 0\) and \(i \neq j\) (temporal independence);
- \(U_{k,i}\) is independent of \(U_{m,j}\) with \(k, m \in \{1,2,\ldots,K\}\) and \(k \neq m\) (spatial independence refers to different SUs);
- \(U_{k,i}\) and \(U_{k,i,j}\) are independent for all \(k \in \{1,2,\ldots,K\}\) and \(j \in \{1,2,\ldots,J\}\) (independence among the global and common parameter vectors).

Under the assumption that all regressors \(U_{k,i}\) are spatially and temporally independent, the unbiased asymptotic convergence of the diffusion-based ATC LMS algorithm with the previous combiners is satisfied for any initial condition, and any choice of the combiners, if step size \(\mu\) takes values that are sufficiently small and that satisfy the inequality relation \(0 < \mu < 2/\lambda_{\text{max}}(R(U_k)), k = 1,2,\ldots,K\), where \(\lambda_{\text{max}}\) is the maximum eigenvalue of the autocorrelation matrix \(R(U_k)\) of the transformed measurements’ model in (9).

### 3.2. Distributed Solution: ATC Diffusion-Based LMS

Distributed and adaptive schemes are usually adopted to improve energy efficiency, robustness, and scalability. In this study, we will exploit the diffusion method ATC, which consists of an adaptation and a combination step \([12]\). In the following, we will describe the main steps of the method, which are also illustrated in Figure 2. Before presenting ATC, a list of useful designations is shown in Table 1. The main steps of the ATC algorithm are the following:

- Assume \(\phi_{k,\text{sys}}^{(0)} \{\phi_{k,j}^{(0)}\}_{j \in I_k}\) at each node \(k \in \{1,2,\ldots,K\}\).
For the estimation of $w^g_s$ and any $\xi_j^g$, choose $K \times K$ combination matrices $C^w$ and $C^g_j$ whose elements in each row $k$, i.e., $\{c^w_{k,ℓ}\}_{ℓ=1}^K$ and $\{c^g_{j,k}\}_{k=1}^K$, satisfy $c^w_{k,ℓ} = 0$ if $ℓ \not\in N_k$ and $\sum_{ℓ \in N_k} c^w_{k,ℓ} = 1$, $c^g_{j,k} = 0$ if $ℓ \not\in N_k \cap C_j$ and $\sum_{ℓ \in N_k \cap C_j} c^g_{j,k,ℓ} = 1$, respectively.

Table 1. List of designations.

| Notation | Description |
|----------|-------------|
| $K$      | Number of SU nodes |
| $Q$      | Number of PU nodes |
| $C_j$    | Set of SU nodes for common interest parameter vector $j = 1, 2, \ldots, J$ |
| $A$      | Adjacent matrix |
| $C^w$    | $K \times K$ keeps the global weights $c^w_{k,j}$ |
| $C^g_j$  | $M_j | C_j | C_j | C_j$ keeps the common weights $c^g_{j,k}$ |
| $M_g$    | Number of global interest parameters |
| $M_g$    | Number of common interest parameters |
| $M_k \leq M_g + \sum_{j=1}^J M_g$ | Maximum number of parameters of interest (global and common) |

Adaptation step at iteration $i$

$$
\begin{bmatrix}
\psi^{(i)}_k \\
\xi^{(i)}_k
\end{bmatrix}
= \begin{bmatrix}
\phi^{(i-1)}_k \\
\phi^{(i-1)}_{k,ℓ}
\end{bmatrix} + μ_k U_k^H d_{k,j} - U_{k,j} \begin{bmatrix}
\phi^{(i-1)}_{k,ℓ} \\
\phi^{(i-1)}_{k,ℓ}
\end{bmatrix}
$$

Combination step at iteration $i$

$$
\phi_{k,ℓ}^{(i)} = \sum_{ℓ \in N_k} c^w_{k,ℓ} \psi^{(i)}_k, \quad \phi_{k,ℓ}^{(i)} = \sum_{ℓ \in N_k \cap C_j} c^g_{j,k,ℓ} \xi^{(i)}_j
$$

for each $j \in I_k$, $\psi^{(i)}_k = \text{col}\{\{\xi^{(i)}_j\}_{j \in I_k}\}$ and $\phi^{(i)}_k = \text{col}\{\{\phi^{(i)}_{k,ℓ}\}_{j \in I_k}\}$. Once the algorithm terminates, $\phi_{k,ℓ}^{(i)}$ and $\phi_{k,ℓ}^{(i)}$ will be the estimates of the desired $w^g_s$ and $\xi_j^g$s. Assuming the clique topology, i.e., $|N_k \cap C_j| = |C_j|$ for all $k \in C_j$, the uniform combination rule creates the following combination weights:

$$
c^w_{k,ℓ} = \frac{1}{|N_k|},
$$

$$
c^g_{j,k} = \frac{1}{|N_k \cap C_j|}$$
Figure 2. Illustration of Adapt-Then-Combine method at SU node $k$.

The combination weights may be adaptive. Specifically, the following adaptive weighting mechanism is adopted,

$$
\gamma_{k,\ell}(i) = (1 - v)\gamma_{k,\ell}(i - 1) + v\|\phi^{(i)}_{k,\ell} - \phi^{(i-1)}_{k,\ell}\|^2,
$$

$$
\delta_{k,\ell}(i) = (1 - v)\delta_{k,\ell}(i - 1) + v\|\psi^{(i)}_{k,\ell} - \phi^{(i-1)}_{k,\ell}\|^2,
$$

where $v$ is a small positive factor in $[0, 1]$ and $\gamma_{k,\ell}, \delta_{k,\ell}$ are the variances in the estimation of the global and common interest parameters. Then, the weights corresponding to both global and common parameter estimation processes are computed according to

$$
\psi_{k,\ell}^{(i)} = \frac{\gamma_{k,\ell}^{-1}(i) - \sum_{m \in N_k(i)} \gamma_{k,m}^{-1}(i)}{\gamma_{k,\ell}^{-1}(i)},
$$

$$
\phi_{k,\ell}^{(i)} = \frac{\delta_{k,\ell}^{-1}(i) - \sum_{m \in N_k(i) \cap C_j} \delta_{k,m}^{-1}(i)}{\delta_{k,\ell}^{-1}(i)}.
$$

3.3. Case Study: Cognitive Radio Network with Three PUs of Overlapping Frequency Spectrum

We have considered a network of $Q = 3$ PUs with an overlapping power spectrum, as illustrated in Figure 3, and a set of $K = 11$ geographically distributed SUs that communicate via a connected network, as shown in Figure 4. The nodes are interconnected by a graph $G(V, E)$, where $V = \{1, 2, \ldots, K\}$ is the set of nodes and $E = \{(i, j) : \text{Node } i \in V \text{ is connected with node } j \in V\}$.

In the network, there is one global interest frequency band $d$ and $J = 6$ common interest frequency bands denoted as $a, b, c, e, z, h$, where $\text{PU}_1 = \{a, b, c, d\}$, $\text{PU}_2 = \{b, d, e, h\}$ and $\text{PU}_3 = \{c, d, e, z\}$.

Each of the following sets includes the nodes interested in the corresponding frequency band, and thus, the parameters of $w_k$ involved in the spectrum approximation. Based on this concept, we assume that there are the following sets:

- Frequency zone $a$: $C_1 = \{1, 2, 3, 4, 5, 6, 7\}$
- Frequency zone $b$: $C_2 = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 11\}$
- Frequency zone $c$: $C_3 = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$
- Frequency zone $d$: $C_4 = \{2, 3, 4, 5, 6, 7, 8, 9, 10\}$
- Frequency zone $e$: $C_5 = \{4, 5, 6, 7, 8, 9, 10\}$
- Frequency zone $f$: $C_6 = \{2, 3, 4, 5, 8, 9, 11\}$
According to these sets, SU1 is located at the coverage area of PU1 and is interested in frequency bands \(a, b, c, d\). SU4 is in the range of PU1, PU2, and PU3, and is interested in all frequency bands. SU2 is in the coverage of PU1 and PU2 and is interested in all zones except for \(z\). Based on the common interest sets \(C_j, j = 1, 2, \ldots, J\) with \(J = 6\), we construct a binary matrix whose columns refer to the number of global and common frequency bands or interest parameter vectors, and its rows capture the SUs in the network. Its size is \(K \times (J + 1)\) and is defined as \(C(k, j) = \begin{cases} 1 & \text{if } k \in C_j, \forall j \in 1, 2, \ldots, J + 1 \text{ and } k \in 1, 2, \ldots, K. \end{cases} \)

More specifically, this matrix is as follows:

\[
C = \begin{bmatrix}
1 & 1 & 1 & 1 & 0 & 0 & 0 \\
1 & 1 & 1 & 1 & 1 & 0 & 1 \\
1 & 1 & 1 & 1 & 1 & 0 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 0 & 0 \\
1 & 1 & 1 & 1 & 1 & 0 & 1 \\
1 & 0 & 1 & 1 & 1 & 1 & 1 \\
1 & 0 & 1 & 1 & 1 & 0 & 0 \\
1 & 0 & 1 & 0 & 1 & 0 & 1
\end{bmatrix}
\]

(17)

Moreover, we consider that for each of the frequency bands of common interest, a parameters’ vector of size \(M_{c_j} \times 1\) should be estimated, while for the global interest frequency band, one parameters vector of size \(M_{g} \times 1\) should be estimated. We create the matrix \(C_{k}^{\text{aug}}\) of size \(L \times (M_{g} + \sum_{j=1}^{J} M_{c_j})\) based on the interests matrix \(C\). At each row of the matrix \(C\), the first column element is duplicated \(M_{g}\) times, and each one of the rest elements is duplicated \(M_{c_j}\) times. Then, each row is duplicated \(L\) times. Hence, the augmented matrix \(C_{k}^{\text{aug}}\) is created and the “correct” input matrix at ATC is \(U_{k}^{\text{aug}} = C_{k}^{\text{aug}} U_{k,i}\) (point by point multiplication) at SU node \(k\) and time instant \(i\). Essentially, the \(C_{k}^{\text{aug}}\) fills with zeros the rows and columns of \(U_{k,i}\) that correspond to those common parameters, which are
out of the interest of node $k$. The acquired $U_{\text{aug}}$ will multiply the augmented parameters vector $w_k = \left[w_g \quad \mathbf{c}_1^T \quad \mathbf{c}_2^T \quad \cdots \quad \mathbf{c}_f^T \right]^T$ ($M_g \times 1$), which, based on (6), is used to collect the measurements $d_{ij}$. The performance evaluation of the ATC method is based on Mean Square Deviation (MSD), assuming different combination rules, i.e., a uniform clique and a relative variance with adaptive weights. The MSD is defined as

$$\text{MSD}_k = \lim_{i \to \infty} E\left\{\|w_k^o - \bar{w}_{k,i}\|^2\right\}$$

(18)

where $w_k^o, \bar{w}_{k,i}$ are respectively the optimal and estimated parameter vectors that contain in concatenate form the $M_g$ global and those $M_c$ common interest parameters that concern node $k$. An algorithmic overview of the above processing steps is outlined in Algorithms 1–3. More specifically, Algorithm 1 is utilized to generate the data that will be used as input to Algorithm 3, which estimates the parameters vector and the MSD for all users in the network.

Additionally, this algorithm exploits internally Algorithm 2, which implements the Equations (13)–(16) in order to create the matrices $C^g, C^c$ involved in the combination step of global and common (uniform and adaptive, respectively). If we set $t = 1$ (global) and $t = 2$ (common) in the Rule functions inside Algorithm 3, the results correspond to the uniform weights combination rule.

### Algorithm 1: Data generation

- Determine the number of PUs $Q$, SUs $K$
- Determine the $(x, y)$ position of SUs and PUs in the topology
- Determine the adjacent matrix $A$
- Determine the number of basis functions $B$, frequency channels $L$, $\omega_m$, $\sigma_m$, and $f_m(e^{i\omega_m})$; and the basis functions’ matrix $G \in \mathbb{R}^{B \times L}$
- Determine the parameters $M_g, \{M_c\}_{i=1}^j$ and set $m_i = \left[M_g \quad M_{c_1} \quad M_{c_2} \quad \cdots \quad M_{c_j}\right]$, with $M = \sum_{i=0}^{j+1} m_i(i)$
- Based on (7), determine path losses with $n = 2.5$, and from (8), construct matrix $P = \left[p_{1i} \quad p_{2i} \quad \cdots \quad p_{Kj}\right]$ of size $K \times Q$ and disturb it using the Gaussian noise of standard deviation $\sigma_p(k)$ and zero mean.
- Randomly generate the true parameters vector $w_{\text{aug}}^o \in \mathbb{R}^{M \times 1}$ and keep those that concern each $k$-th SU by $W(:, k) = C_{\text{arg}}(:, k) \odot \left(\frac{w_{\text{aug}}^o}{\|w_{\text{aug}}^o\|}\right), k = [1, 2, 3, \ldots, K]$, where $W \in \mathbb{R}^{M \times K}$
- Set noise standard deviation $\sigma_n$ to generate random measurement noise of zero mean Gaussian distribution per SU $k$ and iteration $i$
- Set step size $\mu$, the number of internal iterations $\text{iterat}$

**Measurements data generation**

```plaintext
for i = 1 : \text{iterat} do
    for k = 1 : K do
        $P_1 = P(k,:) + P_{\text{noise}}(:, k, i)$
        $U_k = C_{\text{arg}}(:, k) \odot (P_1 \odot G)$
        $U_{\text{aug}}(:, k, i) = U_k$
        $d(:, k, i) = U_k W(:, k, i) + d_{\text{noise}}(:, k, i)$
    end
end
```
Algorithm 2: Rule

Input: $A, c, v_s, K, t$
Output: $C_o$
Initialization: $C_o = O_{size(A)}$

if $t = 1$ then
    Uniform Rule: Global combination weights matrix
    for $i = 1 : K$ do
        $n = \text{find}(A(i,:) \neq 0)$
        for $j = 1 : |n|$ do
            $C_o(i, n(j)) = \frac{1}{|n|}$
        end
    end
else if $t = 2$ then
    Uniform Rule: Common combination weights matrix
    for $i = 1 : K$ do
        $A_c = A(c, c)$
        $n_c = \text{find}(A_c(i,:) \neq 0)$
        for $j = 1 : |n_c|$ do
            $C_o(i, n_c(j)) = \frac{1}{|n_c|}$
        end
    end
else if $t = 3$ then
    Relative Variance Rule: Global combination weights matrix
    $v_{\text{sigma}} = v_s^{-1}$
    for $i = 1 : K$ do
        $n = \text{find}(A(i,:) \neq 0)$
        $v_{n} = v_{\text{sigma}}(n)$
        for $j = 1 : |n|$ do
            $C_o(i, n(j)) = \frac{v_s(j)}{\sum_{j=1}^{\infty} v_s(j)}$
        end
    end
else if $t = 4$ then
    Relative Variance Rule: Common combination weights matrix
    $v_{\text{sigma}} = v_s^{-1}$
    $A_c = A(c, c)$
    for $i = 1 : K$ do
        $n_c = \text{find}(A_c(i,:) \neq 0)$
        $v_{n_c} = v_{\text{sigma}}(n_c)$
        for $j = 1 : |n_c|$ do
            $C_o(i, n_c(j)) = \frac{v_s(j)}{\sum_{j=1}^{\infty} v_s(j)}$
        end
    end
end
Algorithm 3: Diffusion-based Distributed Cooperative Method, Adapt-Then-Combine

\textbf{Output:} \( M S D : K \times (j + 1) \times \text{iterat} \), \( \hat{W}: M \times K \times \text{iterat} \)

\textbf{Input:} \( K, L, M, j, \text{iterat}, \mu, A, m, d, W, U_{\text{aug}}, C \)

\textbf{Stage I: Initialization}

\[ \hat{W} = O_{M \times K \times \text{iterat}}, \hat{W}(z;i,1) = \text{randn}(M,K,1), \epsilon = O_{L \times K \times \text{iterat}}, \text{MSD} = O_{K \times j+1 \times \text{iterat}} \]

\textbf{for} \( k = 1 : K \) \textbf{do}

\[ f = 0, r = 0 \]

\textbf{end}

\textbf{for} \( i = 1 : j + 1 \) \textbf{do}

\textbf{end}

\[ \text{MSD}(k,j,1) = \| \hat{W}(f : r,k,1) - W(f : r,k,1) \|^2 \]

\textbf{end}

\textbf{Stage II: Iterative Part}

\textbf{for} \( i = 2 : \text{iterat} \) \textbf{do}

\textbf{Adaptation Step per node}

\textbf{for} \( k = 1 : K \) \textbf{do}

\[ e(z;k,i) = d(z;k,i) - U_{\text{aug}}(z;k,i)\hat{W}(z,k,i-1) \]

\[ \hat{W}(z;k,i) = \hat{W}(z;k,i-1) + \mu U_{\text{aug}}(z;k,i)e(z;k,i) \]

\textbf{end}

\textbf{for} \( k = 1 : K \) \textbf{do}

\textbf{Global: Adaptive Weights Estimate}

\textbf{for} \( l = 1 : K \) \textbf{do}

\[ \gamma(k,l,i) = (1 - \nu)\gamma(k,l,i-1) + \nu(\| \hat{W}(1 : m_l(1),l,i) - \hat{W}(1 : m_l(1),k,i-1) \|^2) \]

\textbf{end}

\[ C^G = \text{Rule}(A, C, \gamma, 3) \]

\[ C^w = C^G \otimes I_{m(l)} \]

Select only global parameters vector from all users: \( \Phi_{w_g} = \hat{W}(1 : m_l(1),;i) \)

Concatenate the global subset of parameters from all users: \( \Phi_{w_g} = \Phi_{w_g}(\cdot) \)

\textbf{Combination step for Global:} \( \phi_{w_g} = C^w \Phi_{w_g}, \)

\[ \hat{W}(1 : m_l(1),;i) = \text{reshape}(\phi_{w_g}, m_l(1), K) \]

\textbf{Common: Adaptive Weights Estimate}

\[ z = 0, x = m_l(1) \]

\textbf{for} \( j = 1 : j \) \textbf{do}

\[ z = x + 1 \]

\[ x = x + m_l(j+1) \]

\textbf{for} \( l = 1 : K \) \textbf{do}

\[ \delta(k,l,i) = (1 - \nu)\delta(k,l,i-1) + \nu(\| \hat{W}(z : x,l,i) - \hat{W}(z : x,k,i-1) \|^2) \]

\textbf{end}

\[ C^C = \text{Rule}(A, C, \delta, 4) \]

\[ C^c = C^C \otimes I_{m(l+1)} \]

Select users interested for \( j \)-th subset of parameters: \( c = \text{find}(C; j + 1) \neq 0 \)

Select \( j \)-th subset of \( M_c \) common parameters from the users in \( c \) \( \Phi_{c_g} = \hat{W}(z : x,c,i) \)

Concatenate the \( j \)-th subset of common parameters from all users: \( \Phi_{c_g} = \Phi_{c_g}(\cdot) \)

\textbf{Combination step:} \( \phi_{c_g} = C^c \Phi_{c_g}, \)

\[ \hat{W}(z : x,c,i) = \text{reshape}(\phi_{c_g}, m_l(j+1), \text{size}(c,1)) \]

\textbf{end}

\textbf{end}

\textbf{MSD Error Estimate}

\textbf{for} \( k = 1 : K \) \textbf{do}

\[ f = 0, e = 0 \]

\textbf{end}

\textbf{for} \( j = 1 : j + 1 \) \textbf{do}

\[ f = e + 1, c = e + m_l(j) \]

\[ \text{MSD}(k,j,i) = \| \hat{W}(f : c,k,i) - W(f : c,k,i) \|^2 \]

\textbf{end}

\textbf{end}
4. Results and Discussion

In this section, we present the results of the ATC algorithm performance evaluation, which is used to estimate the $Q \cdot B = M_k = M_g + \sum_{i \in I_k} M_c$ parameters that approximate the total power spectrum emitted by the three PUs. Here, we assume that $M_1 = M_2 = \ldots = M_K = M$. The problem is translated into a multiple interests (global and common) parameter estimation problem.

The performance of the adopted method was evaluated using Monte Carlo simulations with 50 independent runs [45]. We simulated a network consisting of $Q = 3$ PUs and $K = 11$ SUs, as illustrated in Figure 3. Additionally, we have considered $B = 16$ Gaussian basis functions of amplitude normalized to one and standard deviation $\sigma_m = 0.05$. These functions are involved in (2), and especially, in the creation of $U_{k,i}$ in the linear model of (6). Furthermore, each SU scans $L = 50$ channels over the normalized frequency axis between 0 and 1. The noise $v_{k,i}$ in (6) or (9) is zero-mean Gaussian with a standard deviation varying between 0.04 and 0.16 for different $k$. Similarly, the additive noise $n_k$ in path loss estimation is generated with $\sigma_p \in [0.03, 1.25]$. In Figure 5, we illustrate the noise power of each SU node. The step-size $\mu_k$ of the LMS adaptation at each node is equal to 0.02 $\forall k$. Additionally, we selected $M_g = 10$, $M_{c_1} = 3$, $M_{c_2} = 5$, $M_{c_3} = 6$, $M_{c_4} = 7$, $M_{c_5} = 8$, $M_{c_6} = 9$.

Figure 4. An overview of the adjacent matrix.

Figure 5. Noise variance per node.
4.1. Static SU Nodes and Varying Spectrum

Here, the term varying spectrum is utilized to capture the case where the PUs ceases to use some of the frequent zones \( a, b, c, d, e, f, z, h \). In this case, we aim to examine the behavior of the ATC algorithm by monitoring if the parameters’ estimation error has the desired behavior. SUs are not aware of these changes, but through the measurements \( \{d_{i,j}, U_{k,i}\} \) and the estimation error of the involved parameters that approximate the total power spectrum, they learn of the changes that occur in the broadcasting spectrum of the PUs.

To evaluate the above conditions, we run the algorithm so that in different time intervals, the use of a particular frequency band is cancelled. The cancellation is achieved by setting zero values to the corresponding optimal parameters in the augmented vector \( w_k \) which is used internally in the algorithm for estimating the MSD. Specifically, we selected to cancel, in specific time intervals, the use of the following parameter vectors:

- 15–20: \( M_g \) parameters of \( w_g \),
- 60–70: \( M_{c_1} \) parameters of \( \varsigma_{1} \),
- 110–120: \( M_{c_2} \) parameters of \( \varsigma_{2} \).

In Figure 6, we present the average performance of MSD error for the aforementioned interest vectors. As it is shown, the ATC method identifies this change in the sensing environment, which is captured in the MSD error of these parameter vectors. For comparison, we also plot the performance of the non-cooperative LMS algorithm, setting the adjacent matrix as \( A = I_K \). The results verify the importance of cooperation in relation to individual estimation. Moreover, comparing the two methods for the estimation of the combination weights, the superiority of adaptive filtering is verified. The nodes working together to achieve spectrum estimation do not know in advance which neighboring nodes are affected by which model and which model affects their own data. A node with different interests may feed its neighbors with irrelevant data, that is, with incorrect estimates that alter their own. The nodes gradually adjust the weight they give to the estimates of that neighbor, which tends to zero, so the effect of its estimate becomes negligible. That is, the network is divided into clusters-groups according to the interests of the nodes. As we see, the collaboration between members of a group with the same interest leads to a better estimation of the parameters of interest. The weights’ values are a measure of trust that a node gives to its neighboring node.

Focusing on the convergence of the ATC, we observe that in the case of the global parameters vector, the algorithm starts to approach the steady state at iteration \( i = 1000 \). However, much more iterations are needed in the case of the two common interest parameter vectors for the MSD to attain a steady state. Additionally, an essential performance difference occurs from the average individual estimate of all nodes. Observing the MSD of the two common interest parameter vectors, it is lower than the MSD of the global parameters.

Figure 6. Cont.
4.2. Mobile CR Node

In this subsection, as a first approach, we evaluate the performance of ATC, assuming that there is one mobile SU node in the network. A moving SU changes position \((x,y)\). However, this change may not necessarily vary the frequency bands of interest and the neighbors. If changes occur, this implies that the corresponding structures of the ATC algorithm should be updated. These structures relate to the path loss coefficients between the mobile SU and the three PUs because it is location-dependent, the adjacency matrix \(A\) and the interests’ matrix \(C\). To check the area in which the moving node is located, we measure the node’s Euclidean distances from the three PUs and compare them with the corresponding transmission range of the PUs. Let’s assume that the distances are \(d_1, d_2,\) and \(d_3\), respectively, and the transmission ranges are \(R_1, R_2,\) and \(R_3\). For the specific experiment, we considered \(R_1 = 150, R_2 = 165,\) and \(R_3 = 170\). Based on Figure 3, we formulated the rules presented in Table 2, which helps us to update the interests’ matrix \(C\) and the structures that depend on it. Moreover, in Table 3, we assume the characteristic points of the SU1 trajectory and the corresponding frequency zones of interest derived from Table 2 under the selected values for the parameters \(R_1, R_2, R_3\).

### Table 2. Rules to determine the frequency zone of a mobile node.

| Rules                                                                 | Frequency Zone |
|----------------------------------------------------------------------|----------------|
| \(d_1 < R_1 \& d_2 > R_2 \& d_3 > R_3\)                             | \(a\)          |
| \(d_1 < R_1 \& d_2 < R_2 \& d_3 > R_3\)                             | \(b\)          |
| \(d_1 < R_1 \& d_2 > R_2 \& d_3 < R_3\)                             | \(c\)          |
| \(d_1 < R_1 \& d_2 < R_2 \& d_3 < R_3\)                             | \(d\)          |
| \(d_1 > R_1 \& d_2 < R_2 \& d_3 < R_3\)                             | \(e\)          |
| \(d_1 > R_1 \& d_2 > R_2 \& d_3 < R_3\)                             | \(z\)          |
| \(d_1 > R_1 \& d_2 < R_2 \& d_3 > R_3\)                             | \(h\)          |

Now, let us focus on Figure 7, which captures the MSD curves of those parameter vectors that correspond to the frequency zones (i.e., \(e, z, h\)) that started to be in the interests of the SU1 node and that relate to the common interest vectors \(\varsigma_4, \varsigma_5, \varsigma_6\). Focusing on \(\varsigma_4, \varsigma_5, \varsigma_6\) and the convergence of the ATC, we see that in the case of \(\varsigma_5, \varsigma_6\) the algorithm approaches the steady state earlier, at iteration \(i = 500\). However, more iterations are needed (at least \(i = 1000\)) in the case of \(\varsigma_4\) to attain the steady state. An essential performance distance is captured between the adaptive and the uniform weighting methods, where the adaptive weights win. Finally, we observe an abrupt fall in the MSD in the time instants 19 and 59, as indicated in Table 3, which shows that the algorithm identifies the new parameters of interest of SU1.
Table 3. Mobile CR node positions and frequency zones.

| Time Interval | Position     | Frequency Zone |
|---------------|--------------|----------------|
| 1–19          | (–40,170)    | \(a,b,c,d\)    |
| 20–39         | (120,150)    | \(a,b,c,d,e,h\) |
| 40–59         | (280,160)    | \(b,d,e,h\)    |
| 60–79         | (100,270)    | \(a,b,c,d,e,z,h\) |
| 80–99         | (110,300)    | \(a,b,c,d,e,z,h\) |
| 100–119       | (190,320)    | \(b,c,d,e,z,h\) |
| 120–139       | (220,390)    | \(c,d,e,z\)    |
| 140–159       | (130,390)    | \(c,d,e,z\)    |
| 160–500       | (10,310)     | \(a,b,c,d,e,z\) |

Figure 7. Network MSD (dB) for the common interest vectors \(\varsigma_4, \varsigma_5, \varsigma_6\).

Assuming the adaptive weights method in the combination part of the ATC, in Figure 8, we capture the mean weight error of the parameters of interest of the moving SU1, which as it moves, changes regions and frequency bands of interest. In each figure, we depict the evolution of the \(M_c\) parameters estimation error. Observing Table 3, we see that the time intervals where the mean error curve is flat correspond to regions where the SU1 is not interested in the respective frequency band.
Figure 8. Mean Weight Error from the true parameters for the common vectors $\{\varsigma_j\}_{j=1}^6$.

In concluding the evaluation part of the adopted ATC method, we see that the algorithm is able to identify the changes in the interests of a SU, which is reflected in the error curves. In either case, its performance remains intact, achieving low error values that attain zero.

5. Conclusions

In conclusion, the estimation of the total power spectrum of three PUs was investigated via the ATC diffusion algorithm, which was tackled as a multi-task problem, having promising results in each of the considered scenarios. We examined its efficiency for two different weighting rules for the coefficients in the combination step of the algorithm. In the case of the relative variance rule combined with the adaptive estimation filter, MSD achieved better performance than the uniform one. This is due to the fact that the noise profile of the nodes is derived based on the instant estimates of the parameters vector, both of them and their neighbors.

An open issue in the above scenarios could be the usage of more overlapping primary users, which would, however, eliminate the zone of global interest, and nodes would only estimate various common interest parameters. Additionally, a variation of the examined...
scenarios could concern the movement of more than one node, since, in practical scenarios, more SUs may be mobile.

Alternatively, we could determine the combination weights, taking into account only the variance of two communicating nodes and not of the whole neighborhood, and evaluate its benefits in the per node and network MSD in the adaptive combination weights. Our aim is to see the improvements to MSD by experimenting with the parameters involved in the ATC method and especially the step size $\mu$, which is an important parameter that impacts on the convergence of the algorithm. Finally, it would be interesting to re-evaluate the above parameter estimation problem, if we were using the Combine-Then-Adapt or LMS-type incremental methods.

**Author Contributions:** E.D. and M.T. conceived of the idea, designed and performed the experiments, analyzed the results, drafted the initial manuscript, and revised the final manuscript. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Conflicts of Interest:** The authors declare no conflict of interest.

**References**

1. Sivagurunathan, P.; Ramakrishnan, P.; Sathishkumar, N. Recent paradigms for efficient spectrum sensing in cognitive radio networks: Issues and challenges. *J. Phys. Conf. Ser.* 2021, 1717, 012057. [CrossRef]
2. Moussa, S.H.; Ismail, M.; Nordin, R.; Abdullah, N.F. Effective Wide Spectrum Sharing Techniques Relying on CR Technology toward 5G: A Survey. *J. Commun.* 2020, 15, 122–147. [CrossRef]
3. Hu, F.; Chen, B.; Zhu, K. Full spectrum sharing in cognitive radio networks toward 5G: A survey. *IEEE Access* 2018, 6, 15754–15776. [CrossRef]
4. Saint, M.; Brown, T.X. A dynamic policy license for flexible spectrum management. *Telecommun. Policy* 2019, 43, 23–37. [CrossRef]
5. Tarek, D.; Benslimane, A.; Darwish, M.; Kotb, A.M. Survey on spectrum sharing/ allocation for cognitive radio networks Internet of Things. *Egypt. Inform. J.* 2020, 21, 231–239. [CrossRef]
6. Kaur, R.; Buttar, A.S.; Anand, J. Spectrum sharing schemes in cognitive radio network: A survey. In Proceedings of the 2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA), Coimbatore, India, 29–31 March 2018; pp. 1279–1284.
7. Hassan, M.R.; Karmakar, G.C.; Kamruzzaman, J.; Srinivasan, B. Exclusive use spectrum access trading models in cognitive radio networks: A survey. *IEEE Commun. Surv. Tutor.* 2017, 19, 2192–2231. [CrossRef]
8. Zhang, L.; Xiao, M.; Wu, G.; Alam, M.; Liang, Y.C.; Li, S. A survey of advanced techniques for spectrum sharing in 5G networks. *IEEE Wirel. Commun.* 2017, 24, 44–51. [CrossRef]
9. Zhang, W.; Wang, C.X.; Ge, X.; Chen, Y. Enhanced 5G cognitive radio networks based on spectrum sharing and spectrum aggregation. *IEEE Trans. Commun.* 2018, 66, 6304–6316. [CrossRef]
10. Tan, F. The algorithms of distributed learning and distributed estimation about intelligent wireless sensor network. *Sensors* 2020, 20, 1302. [CrossRef]
11. Hajiabadi, M.; Khoshbin, H.; Abed Hodtani, G. Cooperative spectrum estimation over large-scale cognitive radio networks. *IET Signal Process.* 2017, 11, 1006–1014. [CrossRef]
12. Plata-Chaves, J.; Bogdanović, N.; Berberidis, K. Distributed Diffusion-Based LMS for Node-Specific Adaptive Parameter Estimation. *IEEE Trans. Signal Process.* 2015, 63, 3448–3460. [CrossRef]
13. Benazzouza, S.; Ridouani, M.; Salahdine, F.; Hayar, A. A survey on compressive spectrum sensing for cognitive radio networks. In Proceedings of the 2019 IEEE International Smart Cities Conference (ISC2), Casablanca, Morocco, 14–17 October 2019; pp. 535–541.
14. Arjoune, Y.; Kaabouch, N. A comprehensive survey on spectrum sensing in cognitive radio networks: Recent advances, new challenges, and future research directions. *Sensors* 2019, 19, 126. [CrossRef]
15. Pandit, S.; Singh, G. Spectrum sensing in cognitive radio networks: Potential challenges and future perspective. In *Spectrum Sharing in Cognitive Radio Networks*; Springer: Berlin/Heidelberg, Germany, 2017; pp. 35–75.
16. Kakalou, I.; Papadopoulou, D.; Xifilidis, T.; Psannis, K.E.; Siakavara, K.; Ishibashi, Y. A survey on spectrum sensing algorithms for cognitive radio networks. In Proceedings of the 2018 7th International Conference on Modern Circuits and Systems Technologies (MOCAST), Thessaloniki, Greece, 7–9 May 2018; pp. 1–4.
17. George, G.R.; Prema, S.C. Cyclostationary feature detection based blind approach for spectrum sensing and classification. *Radioengineering* 2019, 28, 298–303. [CrossRef]
18. Javed, J.N.; Khalil, M.; Shabbir, A. A survey on cognitive radio spectrum sensing: Classifications and performance comparison. In Proceedings of the 2019 International Conference on Innovative Computing (ICIC), Seoul, Korea, 26–29 August 2019; pp. 1–8.
9. Jin, F.; Qiu, T. Low Complexity Compressive Wideband Spectrum Sensing in Cognitive Radio. In Proceedings of the 2018 International Conference on Computer, Information and Telecommunication Systems (CITS), Colmar, France, 11–13 July 2018; pp. 1–5.

10. Odhavjibhai, B.A.; Rana, S. Analysis of Matched filter based spectrum sensing in cognitive radio. Int. Res. J. Eng. Technol. 2017, 4, 578–581.

11. Koekaya, K.; Develi, I. Spectrum sensing in cognitive radio networks: Threshold optimization and analysis. EURASIP J. Wirel. Commun. Netw. 2020, 2020, 1–19. [CrossRef]

12. Zheng, S.; Chen, S.; Qi, P.; Zhou, H.; Yang, X. Spectrum sensing based on deep learning classification for cognitive radios. China Commun. 2020, 17, 138–148. [CrossRef]

13. Shaht, H.A.; Koo, I. Reliable machine learning based spectrum sensing in cognitive radio networks. Wirel. Commun. Mob. Comput. 2018, 2018. [CrossRef]

14. Wang, Y.; Ye, Z.; Wan, P.; Zhao, J. A survey of dynamic spectrum allocation based on reinforcement learning algorithms in cognitive radio networks. Artif. Intell. Rev. 2019, 51, 493–506. [CrossRef]

15. Zhu, P.; Li, J.; Wang, D.; You, X. Machine-learning-based opportunistic spectrum access in cognitive radio networks. IEEE Wirel. Commun. 2020, 27, 38–44. [CrossRef]

16. Muñoz, E.C.; Pedraza, L.F.; Hernández, C.A. Machine Learning Techniques Based on Primary User Emulation Detection in Mobile Cognitive Radio Networks. Sensors 2022, 22, 4659. [CrossRef]

17. Gupta, M.; Verma, G.; Dubey, R.K. Cooperative spectrum sensing for cognitive radio based on adaptive threshold. In Proceedings of the 2016 Second International Conference on Computational Intelligence & Communication Technology (CICIT), Ghaziabad, India, 12–13 February 2016; pp. 444–448.

18. Sharma, V.; Joshi, S. A literature review on spectrum sensing in cognitive radio applications. In Proceedings of the 2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, 14–15 June 2018; pp. 883–893.

19. Arshid, K.; Jianbioa, Z.; Hussain, I.; Pathan, M.S.; Yaqub, M.; Jawad, A.; Munir, R.; Ahmed, F. Energy efficiency in cognitive radio network using cooperative spectrum sensing based on hybrid spectrum handoff. Egypt. Inform. J. 2022, in press. [CrossRef]

20. Ranjeeeth, M.; Nallagonda, S.; Anuradha, S. Optimization Analysis of Improved Energy Detection based Cooperative Spectrum Sensing Network in Nakagami-m and Weibull Fading Channels. J. Eng. Sci. Technol. Rev. 2017, 10, 114–121. [CrossRef]

21. Gao, R.; Qi, P.; Zhang, Z. Performance analysis of spectrum sensing schemes based on energy detector in generalized Gaussian noise. Signal Process. 2021, 181, 107893. [CrossRef]

22. Hajiboseini, A.; Ghorashi, S.A. Distributed spectrum sensing for cognitive radio sensor networks using diffusion adaptation. IEEE Sens. Lett. 2017, 1, 1–4. [CrossRef]

23. Al-Sayed, S.; Plata-Chaves, J.; Muma, M.; Moonen, M.; Zoubir, A.M. Node-specific diffusion LMS-based distributed detection over adaptive networks. IEEE Trans. Signal Process. 2017, 66, 682–697. [CrossRef]

24. Wyslekewa, M.; Bogucka, H.; Kliks, A. Federated Learning for 5G Radio Spectrum Sensing. Sensors 2022, 22, 198. [CrossRef]

25. He, H.; Jiang, H. Deep learning based energy efficiency optimization for distributed cooperative spectrum sensing. IEEE Wirel. Commun. 2019, 26, 32–39. [CrossRef]

26. Ben Smida, O.; Zaidi, S.; Affes, S.; Valaee, S. Robust distributed collaborative beamforming for wireless sensor networks with channel estimation impairments. Sensors 2019, 19, 1061. [CrossRef]

27. Trigka, M.; Mavrokefalidis, C.; Berberidis, K. A Distributed Sparse Channel Estimation Technique for mm Wave Massive MIMO Systems. In Proceedings of the 2021 29th European Signal Processing Conference (EUSIPCO), Dublin, Ireland, 23–27 August 2021; pp. 2154–2158.

28. Trigka, M.; Mavrokefalidis, C.; Berberidis, K. An efficient decentralized approach for mmWave Channel Estimation. In Proceedings of the 2021 International Balkan Conference on Communications and Networking (BalkanCom), Novi Sad, Serbia, 20–22 September 2021; pp. 31–35.

29. Gharib, A.; Ejaz, W.; Ibnkahla, M. Distributed spectrum sensing for iot networks: Architecture, challenges, and learning. IEEE Internet Things Mag. 2021, 4, 66–73. [CrossRef]

30. Xu, Y.; Cheng, P.; Chen, Z.; Li, Y.; Vucetic, B. Mobile collaborative spectrum sensing for heterogeneous networks: A Bayesian machine learning approach. IEEE Trans. Signal Process. 2018, 66, 5634–5647. [CrossRef]

31. Chi, W.; Wang, H.; Xie, W.; Zhang, P.; Ru, L. Research on Distributed Cooperative Intelligent Spectrum Sensing of UAV Cluster. Wirel. Commun. Mob. Comput. 2022. [CrossRef]

32. Nassif, R.; Vlaski, S.; Richard, C.; Chen, J.; Sayed, A.H. Multitask learning over graphs: An approach for distributed, streaming machine learning. IEEE Signal Process. Mag. 2020, 37, 14–25. [CrossRef]

33. Sayed, A.H. Diffusion adaptation over networks. In Academic Press Library in Signal Processing; Elsevier: Amsterdam, The Netherlands, 2014; Volume 3, pp. 323–453.

34. Modalavalasa, S.; Sahoo, U.K.; Sahoo, A.K.; Baraha, S. A review of robust distributed estimation strategies over wireless sensor networks. Signal Process. 2021, 188, 108150. [CrossRef]

35. Bogdanović, N.; Plata-Chaves, J.; Berberidis, K. Distributed diffusion-based LMS for node-specific parameter estimation over adaptive networks. In Proceedings of the 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Florence, Italy, 4–9 May 2014; pp. 7223–7227.