An Emergent Self-Awareness Module for Physical Layer Security in Cognitive UAV Radios

Ali Krayani, *Student Member, IEEE,* Atm S. Alam, *Member, IEEE,* Lucio Marcenaro, *Member, IEEE,* Arumugam Nallanathan, *Fellow, IEEE,* and Carlo Regazzoni, *Senior Member, IEEE*

***Abstract*—In this paper, we propose to introduce an emergent Self-Awareness (SA) module at the physical layer (PHY) in Cognitive Unmanned Aerial Vehicle (UAV) Radios to improve PHY security. SA is based on learning a hierarchical representation of the radio environment by means of a proposed Hierarchical Dynamic Bayesian Network (HDBN). It is shown how the acquired knowledge from previous experiences facilitate the radio spectrum perception and allow the radio to detect abnormal behaviours caused by jamming attacks. Detecting abnormalities realize a fundamental step towards growing up incrementally the radio's long-term memory. Deviations from predictions estimated during abnormal situations are used to characterize jammers at multiple levels and discover their dynamic behavioural rules. Besides, a proactive consequence can be drawn after estimating the jammer's signal to act efficiently by mitigating its effects on the received stimuli. Simulation results show that the introduction of the novel SA functionalities with the proposed HDBN framework provides the high accuracy of characterizing, detecting and predicting the jammer's activities.**

***Index Terms*—Cognitive Radio, Unmanned Aerial Vehicles, Jamming, Bayesian Filtering, Dynamic Bayesian Network**

I. Introduction

Recently, Unmanned Aerial Vehicles (UAVs) have attracted the attention of the telecommunication community and industry due to their remarkable features as deployment flexibility, mobility and dominant Line Of Sight (LOS) links [1]. The explosive number of wirelessly connected UAVs adopted in a wide range of applications including package delivery, traffic monitoring, and surveillance will overcrowd the radio spectrum resources and lead to spectrum scarcity [2]. Incorporating Cognitive Radio (CR) in UAV communications, which we refer to as Cognitive-UAV-Radios, has been firstly proposed to increase the spectrum efficiency [3]. Instead, this work study the integration of CR and UAV from the physical layer security perspective.

CR can provide a promising solution for UAVs to achieve the capability of reaching and maintaining connectivity with high degree of autonomy. Learning the radio environment and adjusting dynamically link parameters based on observation and previous experience are main characteristics of CR. The cognition cycle by which CR interacts with the external environment was first described by Joseph Mitola (the inventor of CR) in [4] and includes the following capabilities: 1) collect data by sensing the radio environment; 2) learn a representation of the collected data; 3) take a decision based on such representation to act on; 4) observe the environmental feedback in response to the action and update the acquired knowledge (autonomous incremental knowledge acquisition) to improve the future decision-action process, consequently. The cognition cycle is the fundamental building block of radio's cognition. However, a radio which does not subsume a certain level of **Self-Awareness (SA)** can not achieve the required cognition and thus can not be considered as CR, according to Mitola's declaration: "*A radio which should be termed as Cognitive must be Self-Aware*". SA as defined in [5] is the "ability of the radio to understand its own capabilities, i.e., to understand what it does and does not know, as well as the limits of its capabilities". SA is concerned with a radio's knowledge about itself and its environment. It is the ability of a radio to interpret the surrounding environment according to the knowledge encoded in its internal models and to adapt its behaviour according to the detected environmental changes to reach the dynamic equilibrium. *An important question that needs to be addressed is which SA representation should be used to realize computationally the cognition cycle of the original CR?* To answer this question we are proposing to introduce a novel SA module in CR that aims to organize the main functionalities of the cognition cycle by learning autonomously and incrementally emergent dynamic representations.

Another main challenge in designing CR is related to the physical layer security, due to the shared and dynamic radio environment that makes CR vulnerable to malicious attacks (e.g. jamming attacks) [6]. Such attacks might affect the system's performance drastically by making the CR learning behaviours wrong and taking mistaken actions. Besides, the dominant LOS links in UAV communications make the ground-to-air and air-to-ground channels more susceptible to terrestrial jammer threats [7]. All these motivations indicate the necessity that the security threats need to be investigated in Cognitive-UAV-Radios with more comprehensive solutions in detecting and mitigating the jamming attacks at the physical layer. Thus, the proposed SA module is specialized to enhance the physical layer security against jamming attacks. Nevertheless, functionalities of the SA module are general enough for

Ali Krayani is with the Department of Electrical, Electronic, Telecommunications Engineering and Naval Architecture, University of Genoa, 16145 Genoa, Italy, and also with the School of Electronic Engineering and Computer Science, Queen Mary University of London, London E1 4NS, U.K. (e-mails: ali.krayani@edu.unige.it, a.krayani@qmul.ac.uk).

Lucio Marcenaro and Carlo Regazzoni are with the Department of Electrical, Electronic, Telecommunications Engineering and Naval Architecture, University of Genoa, 16145 Genoa, Italy. (e-mails: lucio.marcenaro@unige.it, carlo.regazzoni@unige.it).

Atm S. Alam and Arumugam Nallanathan are with the School of Electronic Engineering and Computer Science, Queen Mary University of London, London E1 4NS, U.K. (e-mails: a.alam@qmul.ac.uk, a.nallanathan@qmul.ac.uk).

realizing other goals targeted by the radio and pave the road towards full Self-Aware Radio (i.e. original CR).

The proposed SA module firstly presented in [8] and inspired by [9], allows the radio to perform the following functionalities: ***i)*** **Learning Generative Models** autonomously by observing the occurring environmental changes; ***ii)*** **Radio Spectrum Perception** by predicting the future states of the spectrum and estimating the current states of the observed stimuli received from the environment; ***iii)*** **Abnormality Detection** which refers to the process of noticing any deviation from the normal situation (i.e similar to what it was learned from previous experience); ***iv)*** **Abnormality Characterization** by analyzing the new behaviour (the detected differences which are not seen before) and characterizing it to draw up the dynamic rules of how the new situation is evolving; ***v)*** **Incremental Learning** of new representations of the occurring environmental changes related to the detected signals which represent new behaviour; ***vi)*** **Action Selection** by removing the abnormal signal caused by a malicious user or by changing the policies of communication (changing the configuration: power, modulation, frequency, etc.); ***vii)*** **Learning Interactive Models** by observing multiple signals related to different sources (e.g. LTE and GPS signals) and learning the cross-correlation between them, or by learning the causality among different users (e.g. CR device and jammer) interacting inside the radio spectrum. These functionalities are dependent, while one incrementally leads to the other.

The SA module allows the radio to structure its own memories (brain) into an emergent hierarchical layered representation realized as a **Hierarchical Dynamic Bayesian Network (HDBN)**. This HDBN representation enables the radio to predict the spectrum's future states accurately using Bayesian filters (e.g. Particle Filter and Kalman Filter). We use a top-down inference prediction approach for the radio to anticipate the received (or sensed) signals. The bottom-up inference is used to match the predictions with the real measurement when a stimulus is received from the environment, leading to an updated perception (understanding) of the surrounding environment. This work will focus on the first six functionalities, while the last one will be studied more deeply in future investigation to keep the paper self-contained.

To summarize, our work contributes to the CR literature since: ***1)*** it is following a data-driven unsupervised approach by allowing the radio to build up knowledge about the radio spectrum from null memory. Hence, radio's autobiographical memories are grown up incrementally by observing real-time data and learning autonomously from the cognition cycle. ***2)*** In deep learning approaches where the hidden variables are considered as a black box which create results that are hard to understand or decisions that are not explainable. Differently, in the proposed module hidden variables are related by probabilistic relationships among them and with observations which allow to analyse and study their dynamic evolution. This improves the explainability of the learning and reasoning process respectively. ***3)*** it proposes a flexible SA module that can be implemented at different radio sides (Base Station, User Equipment) regardless of their role (receiver, transmitter or even sensor) and the PHY-layer level at which the module is installed (near or far the antenna). ***4)*** to the best of our knowledge this study is first of its kind which addresses the detection and characterization of the jamming signals in a probabilistic and incremental manner. ***5)*** it is relying on raw IQ data which are easy to extract. Also, using IQ data provides flexibility in implementing the proposed approach in different systems and environments.

The manuscript is organized as follows. The state of the art is reviewed in Section II. The system model is shown in Section III. The proposed SA module and experimental results are presented in Section IV and Section V respectively. Finally, conclusions are drawn in Section VI along with the future objectives.

II. Related Work

CR has attracted intensive attention from academia and industry since its introduction in 1999 [20]. While there is a rich literature on CR focusing on spectrum sensing and dynamic spectrum access, there have been few studies on the original vision of CR in being Self-Aware. As claimed in [5], [21] and [22], the original vision of CR goes beyond the spectrum sensing (which is, of course, one of the main components of the cognition cycle) and aims to improve Quality of Service (QoS) [23], Quality of Information (QoI) [24] and optimizing the wireless users' configuration [10]. Several recent studies (as in [10] and [11]) tried to highlight this fact by spotting the light on the original vision of CR and focusing on new functionalities rather than spectrum sensing, encouraged by the recent advances in Artificial Intelligence (AI) methods and their effectiveness in achieving detection, classification and prediction tasks that can empower the CR realization and support it to effectuate the desired functionalities. Nevertheless, the literature still lacks studies that help to realize the original Mitola's CR and to systematize the functionalities of the cognition cycle and learn incrementally through its consecutive iterations.

After 10 years of CR, a special issue [22] has been organized to provide an overview of the achievements, developments and challenges in this technology containing many references dealing with three functions of the cognition cycle: sensing, learning and decision making/action. An overview of the evolution of CR and the developments on intelligent radio during the last 20 years is provided in [12]. Authors in [10] provided a comprehensive overview of CR and Machine Learning (ML) to improve learning, perception and reconfiguration (components of the cognition cycle) and achieve intelligent wireless communications. A pathway to intelligent radio is presented in [11], where authors asked an important question of how to make CR more intelligent (which in our understanding is how to achieve the original CR) and proposed a structure based on learning and reasoning which give the radio the perception capability with reconfigurable functions to sense and act intelligently. However, these papers did not propose an explicit framework to reach the required functionalities incrementally and did not study the link between them to make the radio growing its memories from experience to achieve the real intelligence or cognition. The work in [13]

Functionalities	[10]	[11]	[12]	[13]	[14]	[15]	[16]	[17]	[18]	[19]	This Work
Learning from Radio Environment	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Radio Spectrum Perception	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Abnormality Detection	-	-	-	-	-	✓	✓	✓	✓	✓	✓
Abnormality Characterization	-	-	-	-	-	-	-	-	-	-	✓
Action Selection	✓	✓	✓	✓	✓	-	-	-	-	-	✓
Incremental Learning	-	-	-	-	-	-	-	-	-	-	✓

TABLE I: General comparison with the state-of-the-art. Publications that investigate the original vision of CR and recommend the ideas to achieve such vision without proposing an explicit framework are shaded.

discussed Game theory and reinforcement learning as learning mechanisms that provide the radio with the capability to learn from its past actions and those of others. In addition, an insight into reasoning frameworks in CR is provided and open research issues are viewed. The article concluded by addressing several challenges that may face the introduction of learning and reasoning in CR (e.g. complexity, reasoning time, etc.). Therefore, an efficient reasoning approach must select carefully the required knowledge within the cognition cycle and intertwines reasoning with learning. In addition to the mentioned learning and reasoning frameworks, we are proposing a combination of probabilistic reasoning and learning while the radio is observing the cognition cycle. Another interesting research is presented in [14] on the opportunities, challenges and future vision for the realization of a fully intelligent radio. An ML-based architecture with three hierarchical levels is proposed which enables the cognitive user to autonomously perceive, understand, and reason about the unknown environment. However, this work studied these levels independently without introducing a link among them to achieve an incremental perception, understanding and reasoning of the surrounding radio environment. In addition, using the feature as the Primary User's (PU's) transmit power might affect the performance of this method in practice (different users might have the same power level and will not allow the radio any more to identify between them).

Understanding the wireless environment is a crucial step in CR to achieve spectrum awareness and then SA [25]. This includes the ability of the radio to identify different signals (e.g. modulation classification) inside the spectrum and detect any unexpected behaviour while monitoring it. Such tasks support the radio to analyse and reason the radio environment and allow it to learn efficiently. For this purpose, a deep learning method based on Long Short Term Memory (LSTM) for automatic modulation classification (AMC) and spectrum anomaly detection based on an adversarial autoencoder (AAE) are proposed in [15] and [16] respectively. Authors in [17] proposed an unsupervised anomaly detection method for the CR using LSTM mixture density networks applied to time series data. Deep predictive coding neural networks for radio-frequency anomaly detection in wireless systems have been proposed in [18] and [19].

In this work, the proposed SA module organizes the main functionalities (Table I) of the cognition cycle and ensures an effective relationship among them by allowing the radio to learn in an incremental fashion and from experience (by observing the cognition cycle). In the considered scenario, the radio is focusing on the goal of enhancing the physical layer security. However, the proposed module can be adapted to reach any goal oriented by the radio ensuring its generalizability to different radio applications.

III. System Model

The system model depicted in Fig. 1 consists of a cellular-connected UAV with a 4G antenna and GPS receiver, acting as aerial user equipment and served by the ground Base Station (BS). A human operator commands and controls the UAV using the LTE cellular connectivity. Commands are sent to the UAV through BS via the Downlink (DL) channel. We assume that the ground-to-UAV link is always a Line-Of-Sight (LOS) under an Additive White Gaussian Noise (AWGN) channel condition. The 3GPP path loss model defined in [26] is adopted under the RMa-AV scenario. We consider the DL channel under the threat of a terrestrial jammer which aims to send false commands to alter the trajectory and take control of the UAV. The propagation model consisting of the LTE downlink transmitter, receiver and jammer is shown in Fig. 2. The BS continuously sends a Radio Frame (RF) of 10 ms duration to the active users (already synchronized with BS) in the cell. The BS allocates a specific number of sub-carriers to each user for a predetermined time which are referred to as Physical Resource Blocks (PRBs). We supposed that the GPS measures the 3D position every 50 ms and the UAV receives one PRB every 50 ms as well (assuming that the BS follows the third allocation scheme for UAV command & control (C2) data as mentioned in [27]) since the 3GPP specifies that efficient management of a UAV would require a maximum of 100 kb/s for C2 data, latency of 50 ms and inter-arrival time (defined also as Transmission Time Interval TTI) of 100 ms [28]. The commands (Pitch, Yaw, and Roll) are sent in the PRB over 9 consecutive sub-carriers in the frequency domain within 1 OFDM symbol in the time domain. We call these REs as a Resource Vector (RV) as shown in Figs. 3-4. The remaining sub-carriers and OFDM symbols of the PRB are related to other information sent to the UAV. For our analysis, only the RV is considered in which we are interested in studying the command signals. However, this can be simply extended to consider the whole PRB in future investigation. We assume that the jammer is smart and is aware of the transmission protocol and the resource allocation strategy performed by the BS. Hence, the jammer can locate and identify the PRBs allocated to the UAV inside the radio spectrum and attacks it consequently.

In our study, the data is extracted after the OFDM receiver block. More precisely, after the FFT, where the output of this block consists of all the Resource Elements (REs) which represent the time-frequency grid. At this level, the UAV can scan and sense the whole REs of the grid and capture the

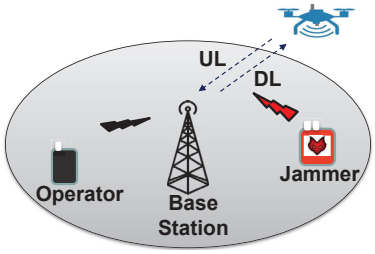


Fig. 1: Illustration of the system model.

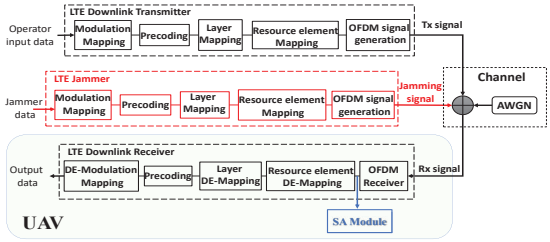


Fig. 2: Illustration of the Propagation Model including the Operator, Jammer and the UAV where the SA module is installed.

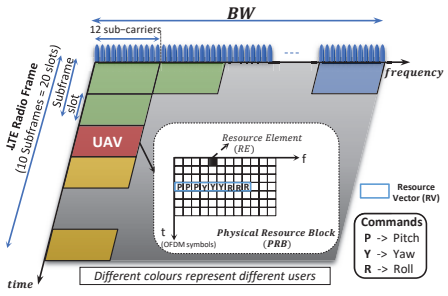


Fig. 3: LTE Physical resource allocation and RF structure.

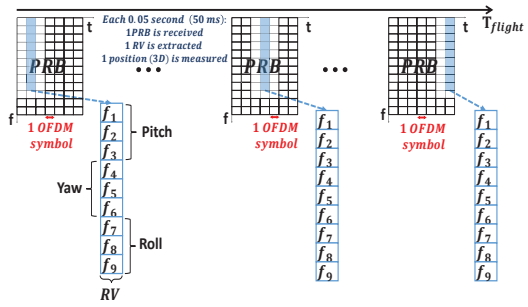


Fig. 4: Timing of the PRBs and RVs received by the UAV.

IQ data without any extra hardware (by exploiting the FFT). The SA module is installed at this level too enabling the UAV to perform all its functionalities presented in the following section.

IV. Proposed Self-Awareness (SA) Module

The proposed SA module illustrated in Fig. 5 involves several functionalities that allow building up knowledge about the radio environment and reach the goal oriented by the *radio* (in the considered system model the *radio* is the *UAV* augmented with the SA module). These functionalities are dependent, the one leads to the other in an incremental way and without any external supervision. The SA functionalities are discussed in details here below.

A. Learning Generative Dynamic Models

One of the most promising approaches towards developing algorithms that can analyze and understand data are Generative Models [29]. Generative Models are statistical models that aim to learn the joint distribution over a set of random variables from which data samples can be generated from that distribution. In this work, we propose to learn a generative model consisting of a Hierarchical Dynamic Bayesian Network (HDBN) [30], due to its ability to represent joint distributions of random variables at different abstraction and temporal levels. HDBN extends the standard Bayesian Networks by introducing the temporal slices where relationships among nodes at the same time instant provide causal links between different abstraction levels and links between consecutive slices represent causal temporal probabilities. The relationship among variables associated with nodes encoded through conditional probabilities is represented by directed edges. HDBN models provide also graphically decomposition of the joint probability directly related to causal relationships among different hidden and observable variables at consequent temporal instants. Therefore, the capability of learning such models from data is equivalent to be capable of generating data sequences at various abstraction levels (hierarchical inference) and temporal levels (temporal inference) coherent with the joint model. So, this makes it possible to individuate in learned models conditional distribution describing the dynamics (both linear and non-linear) over time as well as the semantic rules to associate hidden discrete symbols to variables describing continuous signals. The proposed HDBN whose graphical representation is depicted in Fig. 6-a can be associated with inference mechanisms that work simultaneously at different hierarchical levels, like switching models [31]. A main aspect of the proposed HDBNs is that the variables they include at all levels can be considered as Generalized random variables as allowing both advantages in representation and inference with dynamic systems. The concept of Generalized coordinates proposed by Friston in [32] and used here, implies to represent a pattern vector composed of the random variable per se and its temporal derivative allowing the radio to represent the dynamic rules of how the signal is evolving with time in terms of forces and to facilitate the prediction process. For example, the sensed signals associated to a certain radio situation are represented by nodes at the lower level of the hierarchy that can be defined as Generalized Observations ($\tilde{\mathbf{Z}}$), such that: $\tilde{Z}_t = \Big[\underbrace{I_{f_1},\ldots,I_{f_d},Q_{f_1},\ldots,Q_{f_d}}_{Z_t},\overbrace{\dot{I}_{f_1},\ldots,\dot{I}_{f_d},\dot{Q}_{f_1},\ldots,\dot{Q}_{f_d}}^{\dot{Z}_t}\Big]$, where $\tilde{Z}_t \in \tilde{\mathbf{Z}}$, d is the number of the sensed sub-carriers, and I, Q are in-phase and quadrature components of the sensed signal at different sub-carriers and $\dot{I}$, $\dot{Q}$ are the corresponding derivatives. Such variables are connected with

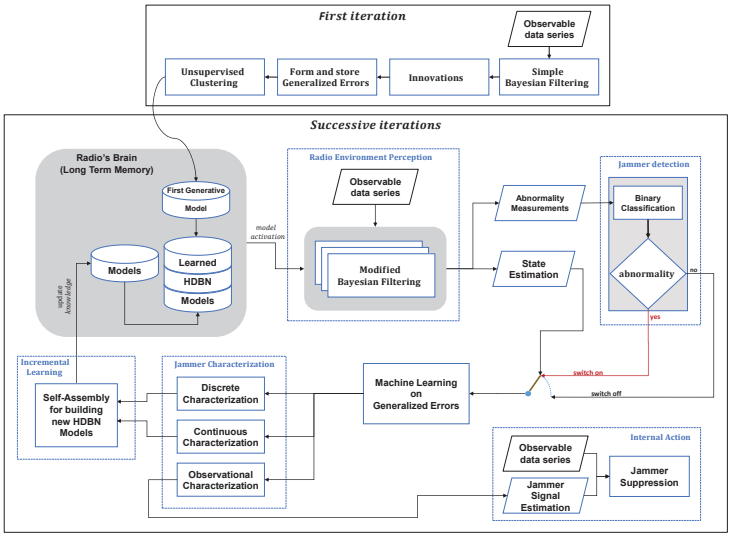


Fig. 5: The Proposed Self-Awareness (SA) Module.

hidden generalized states that are considered as the causes that generate observations and their changes.

At the very beginning (first iteration in Fig. 5) radio's brain is empty ($m = 0$). The radio is assumed to start as a quasi white blackboard where only knowledge describing a static condition assumption is available. Such initial generative model consists of a simplified HDBN of only two levels where the linear dynamic model assumes that generalized continuous variables of the environment (i.e. the state and its derivative) are not changing. The inference mechanism associated with such an HDBN uses a simpler Bayesian Filtering based on an Unmotivated Kalman Filter (UKF) (i.e. a KF that assumes that changes are only due to Gaussian noise). In UKF, predictions can be done using the following dynamic model:

$$X_t = X_{t-1} + v_t \tag{1}$$

where X_t are the predicted states and v_t is a Gaussian noise. The temporal innovations (deviations from the predicted derivatives) of a signal that violates the UKF assumption of static sequences (i.e. null derivative) can be calculated as: $\dot{X}_t = H^{-1}\left(\frac{Z_t - HX_{t-1}}{\Delta t}\right)$, where Δt is the sampling time and H is a matrix that maps hidden states to observations. The estimated elements X_t and $\dot{X}_t$ can be considered equivalent to filtered observations derived from I and Q radio signal features and they can be seen as a data series describing forces that caused violations of static hypothesis defined as Generalized errors ($\tilde{\mathbf{X}}$) expressed as:

$$\tilde{X}_t = \begin{bmatrix} X_t & \dot{X}_t \end{bmatrix} \tag{2}$$

where $\tilde{X}_t \in \tilde{\mathbf{X}}$. The HDBN model ($m = 1$) can be trained by using unsupervised clustering techniques (e.g. Growing Neural Gas (GNG), Self Organizing Maps (SOM)) to group those errors. Here, we employ the GNG algorithm that receives the Generalized errors $\tilde{\mathbf{X}}$ provided by the UKF model defined in (2) and produces a set of superstates (or clusters) $\mathbf{S}$ in which $\tilde{\mathbf{X}}$ are encoded, such that: $\boldsymbol{S^m} = \{S_1^m, S_2^m, \ldots, S_L^m\}$, where $S_k^m \in \boldsymbol{S^m}$ and L is the total number of superstates associated to the first model ($m = 1$). Additionally, each superstate is associated with mean value ($\mu_{S_k^m}$) and covariance matrix ($\Sigma_{S_k^m}$) of the set of hidden filtered $\tilde{X}_t$ samples clustered inside S_k^m, so providing the $P(\tilde{X}_t|S_t^m)$ link in the HDBN.

Within a slice learned, vertical links describe causal relationships between m_t, S_t^m, $\tilde{X}_t$ and $\tilde{Z}_t$ at a given time instant t. Besides, links between variables at consecutive time instants allow representing conditional temporal probabilities among generalized variables starting from obtained superstates related to model m, i.e. dynamic causality that drives changes in the signal. In particular, the LxL transition matrix Π defined as:

$$\Pi = \begin{bmatrix} \pi(S_t^m = S_1^m) \\ \vdots \\ \pi(S_t^m = S_L^m) \end{bmatrix} = \begin{bmatrix} \pi_{11} & \ldots & \pi_{1L} \\ \vdots & \ddots & \vdots \\ \pi_{L1} & \ldots & \pi_{LL} \end{bmatrix} \tag{3}$$

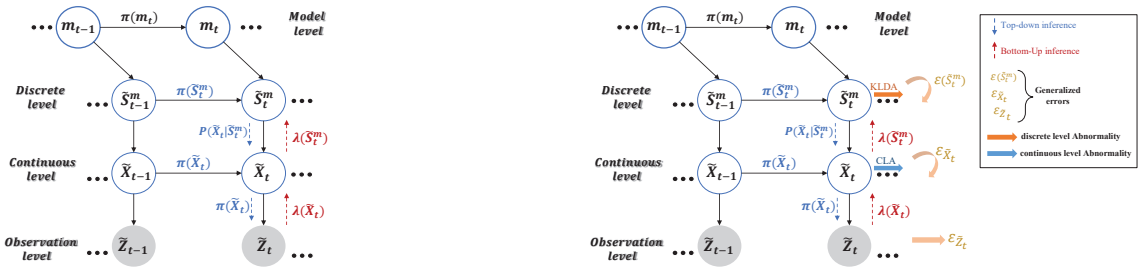
that embeds the dynamic causal models at the discrete level is learned by estimating the transition probabilities $\pi_{ij} = P(S_t^m = j|S_{t-1}^m = i)$, $i, j \in \boldsymbol{S^m}$ over a period of time. Such a probability allows the HDBN to embed knowledge describing the discrete dynamics of the signal, namely transitions from superstate i to superstate j as time evolves. To extend the concept of generalized variables to also include higher HDBN levels, generalized superstates $\tilde{\mathbf{S}}$ can be defined as follows: $\tilde{S_t^m} = [S_t^m \ \dot{S_t^m}] = [S_t^m \ E(S_t^m|S_{t-1}^m)]$, where, S_t^m stands for the current superstate and $E(S_t^m|S_{t-1}^m)$ represents the event of transiting to $S_t^m \in \mathbf{S^m}$ and conditioned to be in S_{t-1}^m in the previous time instant.

The dynamic causal models represented in the HDBN and formulated in generalized superstates $\tilde{\mathbf{S}}^m$ (hidden discrete variables) and generalized states $\tilde{\mathbf{X}}$ (continuous hidden variables) have the following forms:

$$\tilde{S}_t^m = f(\tilde{S}_{t-1}^m, m_{t-1}) + w_t = f(\pi_{ij}^m) + w_t \tag{4}$$

$$\tilde{X}_t = g(\tilde{X}_{t-1}, \tilde{S}_{t-1}^m, m_{t-1}) + w_t = A\tilde{X}_{t-1} + BU_{\tilde{S}_{t-1}^m} + w_t \tag{5}$$

The discrete non-linear function $f(.)$ determines the superstates temporal evolution based on the learned Π at the discrete level. On the other hand, the continuous linear function $g(.)$



(a) (b)

Fig. 6: (a) The Proposed Hierarchical Dynamic Bayesian Network (HDBN). (b) Abnormality measurements and generalized errors at hierarchical levels estimated during the real-time inference.

determines the state temporal evolution at the continuous level of the HDBN. Both $f(.)$ and $g(.)$ are subject to a process noise w_t which is assumed to be drawn from a zero multivariate normal distribution with covariance σ_t such that $w_t \sim \mathcal{N}(0, \sigma_t)$. In (5), $A = [A_1 \ A_2]^\intercal$ is a dynamic model matrix where $A_1 = [I_{2d} \ 0_{2d}]$ and $A_2 = [0_{2d} \ 0_{2d}]$. In addition, $B = [I_{2d} \ 0_{2d}]^\intercal$ is the control model and $U_{\tilde{S}_t^m}$ is the vector parameter describing the dynamics that the signal is expected to follow, when its states belong to a given cluster $\tilde{S}_t^m$ and it is defined as the mean of generalized states' derivative in the cluster, i.e. $\mu(\tilde{S}_t^m)$. So, Π encodes not only transitions between clusters but also jumps between different linear models (different $U_{\tilde{S}_t^m}$ vectors) at the continuous level.

The observation model stands for the bottom level of the HDBN and it is defined as:

$$\tilde{Z}_t = h(\tilde{X}_t) + v_t = H\tilde{X}_t + v_t \tag{6}$$

where $\tilde{Z}_t$ is the observable signal, $h(.)$ is a linear function that maps the hidden generalized state to the observed data and v_t is the measurement noise which is assumed to be zero mean Gaussian noise with covariance R_t such that $v_t \sim \mathcal{N}(0, R_t)$. In (6), $H = [H_1 \ H_2]$ is the observation matrix where $H_1 = [I_{2d} \ 0_{2d}]$, $H_2 = [0_{2d} \ 0_{2d}]$.

The model level in the HDBN showed in Fig. 6-a links the set of models ($m > 1$) stored in the radio's long term memory that are learned in previous experiences related to different radio situations ($m > 1$ depicts the successive iterations in Fig. 5). Such hierarchical representation provides the radio with the capability of predicting not only the dynamics of the continuous states (the direct cause of the observation/bottom level) but also to maintain coherent knowledge between discrete and continuous variables as well as to predict the dynamics at higher levels of abstractions enabling by that an anticipatory process in explaining the behaviour inside the radio spectrum deeply (at higher levels).

B. Radio Spectrum Perception

Classical probabilistic inference approaches (e.g. Belief propagation) make it possible to use Probabilistic Graphical Models (PGMs) like an HDBN for online predictions and estimations. However, here we employ a Modified Bayesian Filtering to integrate within such inference operation also multilevel abnormality measurements. It is known that prediction and estimation related inference in multilevel HDBNs like switching models [31] can be based on a combination of Particle Filter (PF) and Kalman Filter (KF). This allows inferring the radio environmental states at different hierarchical levels. Markov Jump Particle Filter (MJPF) firstly proposed in [33] is the type of inference methods that can be here applicable and from which a Modified Bayesian filtering block can be derived and used as shown in Fig. 5. The inference process concerning prediction and estimation within an MJPF considers the generative model of the HDBN by using its specific components, i.e. dynamic and likelihood inter-level models to parametrize the PF and the KF based on local generative properties required in their algorithmic steps (i.e. the knowledge acquired and stored in the radio's memory is used to drive the inference). In fact, the switching variables of the MJPF can be considered as a finite number of discrete superstates variables associated with the activated model and learned as described in the previous section. Moreover, the corresponding dynamics for each superstate is related to the linear model at the continuous level where the velocity (the average derivative component of the superstate in question) is encoded. PF is employed by the radio to perform superstate predictions due to its ability in dealing with non-linearity, while KF is employed to perform state predictions due to the linear relationship between the continuous variables of the dynamic model at the continuous level.

However, to allow Modified MJPF model to perform additional functionality, namely multilevel abnormality prediction, the local information flow of conditional generative predictions in the HDBN must be enriched by additional inference operations than those included in a classical MJPF. To understand the extension here proposed, it can be useful to recall that in hierarchical Bayesian models as MJPF, an inference can be described as a distributed message passing between nodes where top-down component of generative nature can be associated with both predictivity on temporal links (from slice to slice) and on hierarchical abstraction messages (within a slice or intra-slice). Generative nature in HDBNs also allows generating data series in a bottom-up way, i.e. messages from lower nodes to higher-level nodes. Thus, depending on the direction of message passing, a prognostic or predictive (top-down) component of inference can be distinguished from a diagnostic (bottom-up) one.

Temporal and semantic switching predictive top-down messages ($\pi(\tilde{X}_t)$, $\pi(\tilde{S}_t^m)$) depend on the knowledge learned in

dynamical models within an HDBN. The bottom-up inference is based on likelihood models and consists of passing messages ($\lambda(\tilde{X}_t)$, $\lambda(\tilde{S}_t^m)$) in a feed-backward manner to adjust the expectations (predictions provided by top-down messages) given a sequence of observations. Thus, the availability of synchronous top-down messages and bottom-up messages in classical Bayesian inference allows changing the belief in hidden variables in a distributed way by updating the belief using messages coming from connected variables. MJPF consists of two steps, prediction and update at each time instant t. In the prediction step, PF relies on the transition matrix as a proposal to predict the next superstate ($\tilde{S}_t^{m*}$) as mentioned in (4), by propagating a set of particles each being associated with a specific superstate and a weight $W_t^* = 1/N$, where N is the total number of particles at each t. Then for each particle ($.^*$), a KF is used to predict the continuous state as pointed out in (5). This prediction depends on the hypothesized superstate as (5) can be written as a conditional probability $P(\tilde{X}_t|\tilde{X}_{t-1}, \tilde{S}_t^{m*})$. Accordingly, the posterior probability associated with the predicted state is:

$$\pi(\tilde{X}_t) = P(\tilde{X}_t, \tilde{S}_t^{m*}|\tilde{Z}_{t-1}) = \int P(\tilde{X}_t|\tilde{X}_{t-1}, \tilde{S}_t^{m*}) \overbrace{\lambda(\tilde{X}_{t-1})}^{P(\tilde{Z}_{t-1}|\tilde{X}_{t-1})} d\tilde{X}_{t-1} \quad (7)$$

In the update step, the posterior probability $P(\tilde{X}_t, \tilde{S}_t^{m*}|\tilde{Z}_t)$ is corrected by using the message ($\lambda(\tilde{X}_t)$) backward propagated from the current observation $\tilde{Z}_t$, in the following way: $P(\tilde{X}_t, \tilde{S}_t^{m*}|\tilde{Z}_t) = \pi(\tilde{X}_t)\lambda(\tilde{X}_t)$. After that, the message $\lambda(\tilde{S}_t^m)$ backward propagated from the bottom level towards the discrete level can be used to update the belief in $\tilde{S}_t^m$, that can be calculated as:

$$\lambda(\tilde{S}_t^m) = \lambda(\tilde{X}_t)P(\tilde{X}_t|\tilde{S}_t^m) = P(\tilde{Z}_t|\tilde{X}_t)P(\tilde{X}_t|\tilde{S}_t^m) \quad (8)$$

where $P(\tilde{X}_t|\tilde{S}_t^m) \sim \mathcal{N}(\mu_{\tilde{S}_k^m}, \Sigma_{\tilde{S}_k^m})$ denotes a Gaussian distribution with mean $\mu_{\tilde{S}_k^m}$ and covariance $\Sigma_{\tilde{S}_k^m}$. While, $\lambda(\tilde{X}_t) \sim \mathcal{N}(\mu_{\tilde{Z}_t}, R)$ denotes a Gaussian distribution with mean $\mu_{\tilde{Z}_t}$ and covariance R. The multiplication between $\lambda(\tilde{X}_t)$ and $P(\tilde{X}_t|\tilde{S}_t^m)$ can be obtained by calculating the Battacharyya distance (D_B) as follows:

$$D_B\big(\lambda(\tilde{X}_t), P(\tilde{X}_t|\tilde{S}_t^m = \tilde{S}_k^m)\big) = -ln \int \sqrt{\lambda(\tilde{X}_t)P(\tilde{X}_t|\tilde{S}_t^m = \tilde{S}_k^m)}d\tilde{X}_t \quad (9)$$

where $\tilde{S}_k^m \in \tilde{\boldsymbol{S}}^{\boldsymbol{m}}$. The vector D_λ containing all the D_B values between $\lambda(\tilde{X}_t)$ and all the superstates in the set $\tilde{\boldsymbol{S}}^{\boldsymbol{m}}$ is here estimated as:

$$D_\lambda = \big[D_B(\lambda(\tilde{X}_t), P(\tilde{X}_t|\tilde{S}_t^m = \tilde{S}_1^m)), \ldots, D_B(\lambda(\tilde{X}_t), P(\tilde{X}_t|\tilde{S}_t^m = \tilde{S}_L^m))\big] \quad (10)$$

Therefore, the vector $\lambda(\tilde{S}_t^m)$ in terms of probability is:

$$\lambda(\tilde{S}_t^m) = \left[\frac{1/D_\lambda(1)}{1/\sum_{l=1}^L D_\lambda(l)}, \ldots, \frac{1/D_\lambda(L)}{1/\sum_{l=1}^L D_\lambda(l)}\right] \quad (11)$$

After calculating $\lambda(\tilde{S}_t^m)$ and unlike [33], herein the weight W_t^* of the particle $\tilde{S}_t^{m*}$ is updated using $W_t^* = W_t^*\lambda(\tilde{S}_t^{m*})$ and then normalized by considering the Sequential Importance Resampling (SIR) technique.

In classical MJPF, predictive and diagnostic messages are at the basis of Bayesian updating, i.e. are used to estimate an updated joint posterior at different levels. However, some information is lost in this process, namely an evaluation according to some probabilistic metric of the differences between two messages arriving at a given node. In fact, in such difference, i.e. the surprise or abnormality can be estimated based on the difference of expectation w.r.t evidence coming from data on a given variable. The following section defines how this is here obtained to provide Self-Aware radio with a general basis for abnormality detection.

C. Abnormality Detection

In the modified approach here proposed, classical Bayesian inference is enriched by another capability correlated to exploit and describe differences between top-down and bottom-up messages incoming into a generic node of the HDBN that provide hierarchical abnormality signals as shown in Fig. 6-b.

1) Kullback-Leibler-Divergence Abnormality ($KLDA$) at Discrete Level: in this case the abnormality is defined as a distance between the messages entering to node $\tilde{S}_t^m$, namely the predictive ($\pi(\tilde{S}_t^m)$) and the diagnostic ($\lambda(\tilde{S}_t^m)$) messages. Differences between the probability profiles of predictive support and evidence indicate that involved components of the generative model used to predict radio environment dynamics do not fit current observations, i.e. provide to the radio an awareness signal indicating whether and how much the current surrounding environment is behaving in a way different to the rules learned in the generalized model. Since these two messages represent discrete probability distributions, Kullback-Leibler-Divergence (KLD) is here proposed as probability distance measurement to calculate the difference between them as follows:

$$\boldsymbol{KLDA} = D_{KL}\big(\pi(\tilde{S}_t^m)||\lambda(\tilde{S}_t^m)\big) + D_{KL}\big(\lambda(\tilde{S}_t^m)||\pi(\tilde{S}_t^m)\big) \quad (12)$$

To computationally apply this distance in the run time of the MJPF, at each time instant t and after the updated process (KF) and resampling (PF) the histogram of the particles is extracted and the relative frequency of each particle is used to approximate the local new prior over $\tilde{S}_t^m$ as follows:

$$p(\tilde{S}_t^m = i) = \frac{y(\tilde{S}_t^m = i)}{N} \quad i \in \boldsymbol{S} \quad (13)$$

where $y(.)$ is the frequency (or number of occurrences of superstate i in the histogram) and N is the total number of particles. It is worth to note that $\lambda(\tilde{S}_t^m)$ is the same for all the particles propagated by PF at time instant t. As $\pi(\tilde{S}_t^m)$ is available as a row vector related to $\tilde{S}_t^m$ picked from the transition matrix, the predictive message can be approximated by selecting the row to be used in the $KLDA$ depending on the updated prior histogram previously introduced. Lets define $\mathbb{S}$ as the set of the winning particles (whose entries in the histogram is greater than zero), where:

$$\mathbb{S} = \{i|p(\tilde{S}_t^m = i) > 0\} \quad i \in \boldsymbol{S} \quad (14)$$

D_{KL} is calculated between $\lambda(\tilde{S}_t^m)$ and the rows of the transition matrix $\pi(\tilde{S}_t^m)$ related to the winning particles $\mathbb{S}$ (that are the most probable before update). Therefore, (12) becomes:

$$\boldsymbol{KLDA} = \sum_{i\in\mathbb{S}} \big[p(i) \sum_{j=1}^{L} \pi_{ij} \log(\frac{\pi_{ij}}{\lambda_j})\big] + \sum_{i\in\mathbb{S}} \big[p(i) \sum_{j=1}^{L} \lambda_j \log(\frac{\lambda_j}{\pi_{ij}})\big] \quad (15)$$

The space of the predicted discrete variables can be divided into two sub-sets. One corresponding to the normality region Ω_N, and the other to its complement (Ω_A), that can be interpreted as the subset of unexpected super-states, i.e. outside the model's trusted prediction area. Ω_N can be defined as the sub-set that contains the most probable particles (with high frequency of occurrence in the histogram) and satisfy the following condition:

$$\sum_{i\in\Omega_N} p(\tilde{s}_t = i) \geq \zeta \quad (16)$$

while, Ω_A is the sub-set that contains the less probable particles (with low frequency of occurrences) and satisfy the following condition:

$$\sum_{i\in\Omega_A} p(\tilde{s}_t = i) \leq 1 - \zeta \quad (17)$$

where, $\Omega_N \subset \mathbb{S}$, $\Omega_A \subset \mathbb{S}$ while $p(.)$ and $\mathbb{S}$ are defined in (13) and (14), respectively. ζ is the acceptance ratio that varies from 0 to 1. It is possible evaluating to what extent the probability mass function of the evidence message falls in the region Ω_N and its complement to 1 in Ω_A. A value α can be used to indicate the amount of support (i.e. the probability mass of $\lambda(\tilde{S}_t^m)$ in the normality region). The KLDA defined in (12) covers the global discrete state space, however by dividing the discrete space into two regions we can rewrite KLDA as:

$$\begin{aligned}\textit{KLDA} = \mathcal{KLD}_N + \mathcal{KLD}_A = \\ \sum_{i\in\Omega_N} p(\tilde{s}_t = i) \times \Big[D_{KL}\big(\pi(\tilde{S}_t^m = i)||\lambda(\tilde{S}_t^m)\big) + \\ D_{KL}\big(\lambda(\tilde{S}_t^m)||\pi(\tilde{S}_t^m = i)\big)\Big] \\ + \sum_{i\in\Omega_A} p(\tilde{s}_t = i) \times \Big[D_{KL}\big(\pi(\tilde{S}_t^m = i)||\lambda(\tilde{S}_t^m)\big) + \\ D_{KL}\big(\lambda(\tilde{S}_t^m)||\pi(\tilde{S}_t^m = i)\big)\Big] \quad (18)\end{aligned}$$

$\mathcal{KLD}_N$ measures the similarity between the set of particles with the highest prior probability and the probabilistic evidence, while $\mathcal{KLD}_A$ measures the similarity with low probability predictions. The abnormal situation can so be defined as:

$$\mathcal{KLD}_N > \mathcal{KLD}_A, \quad (19)$$

From (18),

$$\mathcal{KLD}_A = \textit{KLDA} - \mathcal{KLD}_N \quad (20)$$

After substituting (20) in (19) we have: $\mathcal{KLD}_N > \textit{KLDA} - \mathcal{KLD}_N$. Thus, $\textit{KLDA} < 2(\mathcal{KLD}_N)$. In addition, during a normal situation $\mathcal{KLD}_N$ is supposed to be small due to the fact that the observation often votes for the most probable particles, thus:

$$\mathcal{KLD}_N \leq (1 - \zeta)\alpha \quad (21)$$

where α is the support mass expected by $\lambda(\tilde{S}_t)$ to consider a prediction as not associated with abnormalities. Therefore the normal situation occurred if:

$$\textit{KLDA} < (1 - \zeta)\alpha \quad (22)$$

2) Abnormality at Continuous Level: At this level, the abnormality indicator can be defined as a distance between different probabilistic messages incoming to node $\tilde{X}_t$. It is based on the Bhattacharyya distance (D_B) between the predictive message $\pi(\tilde{X}_t)$ and the diagnostic message $\lambda(\tilde{X}_t)$ incoming from the observation level. Thus, the continuous level abnormality (CLA) is defined as:

$$\begin{aligned}CLA = D_B(\pi(\tilde{X}_t), \lambda(\tilde{X}_t)) = \\ - \ln \int \sqrt{P(\tilde{X}_t, \tilde{S}_t^{m*}|\tilde{Z}_{t-1}) P(\tilde{Z}_t|\tilde{X}_t)} d\tilde{X}_t \quad (23)\end{aligned}$$

$\pi(\tilde{X}_t)$ is distributed according to a multivariate Gaussian and it has the following probability density function (PDF):

$$\begin{aligned}P(\pi(\tilde{X}_t)) = \\ \frac{1}{\sqrt{(2\pi)^d |\Sigma_\pi|}} \exp\big[-\frac{1}{2}(\tilde{x_t} - \mu_\pi)^T \Sigma_\pi^{-1} (\tilde{x_t} - \mu_\pi)\big] \quad (24)\end{aligned}$$

The shape of $\pi(\tilde{X}_t)$ is characterized by the covariance matrix Σ_π and the Mahalanobis distance (D_M) of each data sample $\tilde{x}_t$ from the distribution's centroid. Thus, a good generative model can generate data samples $\tilde{x}_t$ that lie in the confidence region almost of the time. The confidence region is often represented as an ellipsoid around the data samples that satisfy the following condition:

$$(\tilde{x_t} - \mu_\pi)^T \Sigma_\pi^{-1} (\tilde{x_t} - \mu_\pi) \leq \chi_d^2 \quad (25)$$

where $(\tilde{x_t} - \mu_\pi)^T \Sigma_\pi^{-1} (\tilde{x_t} - \mu_\pi)$ is the Squared Mahalanobis distance and χ_d^2 is the quantile function of the chi-squared distribution with d degrees of freedom. Then,

$$\begin{cases} \tilde{x}_t \in \mathcal{R}_N, & \text{if } (\tilde{x_t} - \mu_\pi)^T \Sigma_\pi^{-1} (\tilde{x_t} - \mu_\pi) \leq \chi_d^2 \\ \tilde{x}_t \in \mathcal{R}_A, & \text{if } (\tilde{x_t} - \mu_\pi)^T \Sigma_\pi^{-1} (\tilde{x_t} - \mu_\pi) > \chi_d^2 \end{cases} \quad (26)$$

where $\mathcal{R}_N$ is the normal region and $\mathcal{R}_A$ is the abnormal region. From the Mahalanobis distances we can obtain the upper ($\tilde{x}_t^{UB}$) and lower ($\tilde{x}_t^{LB}$) bounds of $\pi(\tilde{X}_t)$. The area of each region can be written as:

$$\begin{aligned}\mathcal{R}_A = \\ \int_{-\infty}^{\tilde{x}_t^{LB}} \frac{1}{\sqrt{(2\pi)^d |\Sigma_\pi|}} \exp\big[-\frac{1}{2}(\tilde{x}_t - \mu_\pi)^T \Sigma_\pi^{-1} (\tilde{x}_t - \mu_\pi)\big] d\tilde{x}_t + \\ \int_{\tilde{x}_t^{UB}}^{+\infty} \frac{1}{\sqrt{(2\pi)^d |\Sigma_\pi|}} \exp\big[-\frac{1}{2}(\tilde{x}_t - \mu_\pi)^T \Sigma_\pi^{-1} (\tilde{x}_t - \mu_\pi)\big] d\tilde{x}_t, \quad (27)\end{aligned}$$

$$\mathcal{R}_N = \int_{\tilde{x}_t^{LB}}^{\tilde{x}_t^{UB}} \frac{1}{\sqrt{(2\pi)^d|\Sigma_\pi|}} \exp\big[-\frac{1}{2}(\tilde{x}_t-\mu_\pi)^T\Sigma_\pi^{-1}(\tilde{x}_t-\mu_\pi)\big]d\tilde{x}_t, \quad (28)$$

where the performance of the generative model depends upon these regions. A good generative model can generate more data samples that lie in $\mathcal{R}_N$, such that, $\mathcal{R}_N > \mathcal{R}_A$.

On the other side, the message ($\lambda(\tilde{X}_t)$) tells if the real signal matches the predicted signal. Thus, we aim to study how much the diagnostic message supports the predictive message generated by the generative model. This can be done by calculating the product between the two messages and then integrating it to obtain the corresponding area. After that we must see if the calculated area belongs to the normal or the abnormal region defined in (27) and (28) respectively, to make the correct decision. We propose to employ the D_B to calculate the similarity between the two messages. The analytical advantage of using D_B is due to its ability in measuring the overlap area between the two distributions through the Bhattacharyya coefficient ($\mathcal{BC}$) and then converting it to a distance metric. The $\mathcal{BC}$ is divided into 3 regions as follows:

$$\mathcal{BC} = \int_{-\infty}^{+\infty}\sqrt{\pi(\tilde{X}_t)\lambda(\tilde{X}_t)}d\tilde{X}_t = \overbrace{\int_{-\infty}^{\tilde{x}_t^{LB}}\sqrt{\pi(\tilde{X}_t)\lambda(\tilde{X}_t)}d\tilde{X}_t}^{\mathcal{BC}_A^1} + \overbrace{\int_{\tilde{x}_t^{LB}}^{\tilde{x}_t^{UB}}\sqrt{\pi(\tilde{X}_t)\lambda(\tilde{X}_t)}d\tilde{X}_t}^{\mathcal{BC}_N} + \underbrace{\int_{\tilde{x}_t^{UB}}^{+\infty}\sqrt{\pi(\tilde{X}_t)\lambda(\tilde{X}_t)}d\tilde{X}_t}_{\mathcal{BC}_A^2} \quad (29)$$

where $\mathcal{BC}_A$ is the overlap that falls inside the abnormal region $\mathcal{R}_A$ (i.e. $\mathcal{BC}_A \subset \mathcal{R}_A$) and $\mathcal{BC}_N$ is the overlap that falls in the normal region $\mathcal{R}_N$ ($\mathcal{BC}_N \subset \mathcal{R}_N$). The normal situation occurred if:

$$\mathcal{BC}_N > \mathcal{BC}_A \quad (30)$$

$\mathcal{BC}_A$ is the overlap between $\pi(\tilde{X}_t)$ and $\lambda(\tilde{X}_t)$ inside the region $\mathcal{R}_A$, which can be written as:

$$\mathcal{BC}_A = \int_{-\infty}^{\tilde{x}_t^{LB}}\sqrt{\pi(\tilde{X}_t)\lambda(\tilde{X}_t)}d\tilde{X}_t + \int_{\tilde{x}_t^{UB}}^{+\infty}\sqrt{\pi(\tilde{X}_t)\lambda(\tilde{X}_t)}d\tilde{X}_t \quad (31)$$

According to the product rule:

$$\mathcal{BC}_A = \int_{-\infty}^{\tilde{x}_t^{LB}}\sqrt{\pi(\tilde{X}_t)}d\tilde{X}_t \;.\; \int_{-\infty}^{\tilde{x}_t^{LB}}\sqrt{\lambda(\tilde{X}_t)}d\tilde{X}_t + \int_{\tilde{x}_t^{UB}}^{+\infty}\sqrt{\pi(\tilde{X}_t)}d\tilde{X}_t \;.\; \int_{\tilde{x}_t^{UB}}^{+\infty}\sqrt{\lambda(\tilde{X}_t)}d\tilde{X}_t \quad (32)$$

Then, according to the Cauchy–Schwarz inequality:

$$\begin{aligned}(\mathcal{BC}_A)^2 &\le \int_{-\infty}^{\tilde{x}_t^{LB}}\pi(\tilde{X}_t)d\tilde{X}_t \;.\; \int_{-\infty}^{\tilde{x}_t^{LB}}\lambda(\tilde{X}_t)d\tilde{X}_t + \\ &\int_{\tilde{x}_t^{UB}}^{+\infty}\pi(\tilde{X}_t)d\tilde{X}_t \;.\; \int_{\tilde{x}_t^{UB}}^{+\infty}\lambda(\tilde{X}_t)d\tilde{X}_t \\ &\le Q_{\tilde{x}_t^{LB}}(\tilde{X}_t) \;.\; [\beta_1(1-\alpha)] + Q_{\tilde{x}_t^{UB}}(\tilde{X}_t) \;.\; [\beta_2(1-\alpha)]\end{aligned} \quad (33)$$

Since $\pi(\tilde{X}_t)$ is a symmetric multivariate Gaussian we have: $\tilde{x}_t^{LB} = \tilde{x}_t^{UB} = \mathcal{T}$. Therefore,

$$\mathcal{BC}_A \le \sqrt{Q_\mathcal{T}(\tilde{X}_t) \;.\; [\beta_1(1-\alpha)] + Q_\mathcal{T}(\tilde{X}_t) \;.\; [\beta_2(1-\alpha)]} \quad (34)$$

where α is the area of $\lambda(\tilde{X}_t)$ falling inside the normal region $\mathcal{R}_N$ and $(\beta_1+\beta_2)=1$.
From (29):

$$\mathcal{BC}_N = \mathcal{BC} - \mathcal{BC}_A \quad (35)$$

After substituting in (30), we have: $\mathcal{BC} > 2(\mathcal{BC}_A)$. Hence, the normal situation occurred if:

$$\mathcal{BC} > \sqrt{Q_\mathcal{T}(\tilde{X}_t) \;.\; [\beta_1(1-\alpha)] + Q_\mathcal{T}(\tilde{X}_t) \;.\; [\beta_2(1-\alpha)]} \quad (36)$$

D. Abnormality Characterization

In the previous functionality, message-passing in the HDBN from slice to slice is exploited to calculate the abnormality measurements, while here the message-passing in the same slice (intra-slice) will be used to calculate the Generalized Errors (as shown in Fig. 6-b) from which the jammer can be characterized at multiple levels by means of machine learning.

1) Jammer Characterization at Discrete Level: At this level, the radio can characterize the abnormal situation by analysing the superstates' evolution based on the predictive messages ($\pi(\tilde{S}_t^m)$) and that based on the diagnostic messages ($\lambda(\tilde{S}_t^m)$). In this way two sets of superstates ($\mathbb{S}^\pi$ and $\mathbb{S}^\lambda$) are created, the first one contains the predicted superstates while the second contains the observed superstates. At each time instant t_j (the time when the jammer is detected) these superstates can be obtained as follows:

$$\begin{cases}\tilde{S}_{t_j}^\pi = \underset{\tilde{S}_{t_j}^m \in \mathbb{S}}{\operatorname{argmax}}\, y(\tilde{S}_{t_j}^m) \\ \tilde{S}_{t_j}^\lambda = \underset{\tilde{S}_{t_j}^m \in \boldsymbol{S}}{\operatorname{argmax}}\, \lambda(\tilde{S}_{t_j}^m)\end{cases} \quad (37)$$

where $y(\tilde{S}_t^m)$ is defined in (13), $\tilde{S}_{t_j}^\pi \in \mathbb{S}^\pi$ and $\tilde{S}_{t_j}^\lambda \in \mathbb{S}^\lambda$. Comparing between $\tilde{S}_{t_j}^\pi$ and $\tilde{S}_{t_j}^\lambda$ can help to understand how the jammer is affecting the superstates evolution. If $\tilde{S}_{t_j}^\lambda$ is not equal to $\tilde{S}_{t_j}^\pi$, this means that the jammer shifts the signal from $\tilde{S}_{t_j}^\pi$ to $\tilde{S}_{t_j}^\lambda$, otherwise the signal is manipulated by the jammer but kept in the expected superstate. In other words, the radio expects (predicts) that the signal's sample (i.e. OFDM symbol) will fall in a certain superstate based on the dynamic rules learned in previous experience (related to model m). However, during attacks, the jammer shifts the signal's sample

to another superstate or even manipulate the sample but keeps it inside the predicted superstate (the case when $\tilde{S}^{\pi}_{t_j}$=$\tilde{S}^{\lambda}_{t_j}$). This cross-correlation between the predictive support $\pi(\tilde{S}^{m}_{t_j})$ and the diagnostic support $\lambda(\tilde{S}^{m}_{t_j})$ allows the radio to understand the jammer's effect on the superstates (at discrete level) of the learned dynamic model and to discover the jammer's strategy or the dynamic rules it is following to attack the signal.

2) Jammer Characterization at Continuous Level: Characterizing the attack at the continuous level helps the radio to understand the jammer's force in terms of I and Q values and how much the jammer shifted the signal from one superstate to the other or from a specific superstates' centroid. This depends on the characterization done before at the discrete level. The characteristics at the discrete level can be forwarded towards the continuous level to calculate the generalized errors as follows:

$$\mathbb{D}_{t_j} = \begin{cases} \overbrace{\mu\big(\underset{\tilde{S}^{m}_{t_j}\in\mathbb{S}}{\operatorname{argmax}}\,\lambda(\tilde{S}^{m}_{t_j})\big) - \tilde{X}^{\lambda}_{t_j}}^{\textbf{Generalized Error 1 } (\varepsilon^{1}_{\tilde{X}_{t_j}})} \;\; if\;\; \tilde{S}^{\pi}_{t_j} = \tilde{S}^{\lambda}_{t_j} \\ \underbrace{\mu\big(\underset{\tilde{S}^{m}_{t_j}\in\boldsymbol{S}}{\operatorname{argmax}}\,\lambda(\tilde{S}^{m}_{t_j})\big) - \mu\big(\underset{\tilde{S}^{m}_{t_j}\in\boldsymbol{S}}{\operatorname{argmax}}\,\pi(\tilde{S}^{m}_{t_j})\big)}_{\textbf{Generalized Error 2 } (\varepsilon^{2}_{\tilde{X}_{t_j}})} \;\; if\;\; \tilde{S}^{\pi}_{t_j} \neq \tilde{S}^{\lambda}_{t_j} \end{cases} \tag{38}$$

where $\mathbb{D}_{t_j}$ is the generalized error containing the I and Q values and the corresponding derivatives at multiple subcarriers. In (38), if the jammer manipulates the signal but keep it in the same superstate (the expected one), the generalized error is equal to the mean value of ($\tilde{S}^{\lambda}_{t_j}$) subtracted from the generalized state associated with the most probable superstate in $\pi(\tilde{S}_{t_j})$. Otherwise, if the jammer shifts the signal from one superstate to another one, the generalized error is equal to the mean value of the current superstate ($\tilde{S}^{\lambda}_{t_j}$) subtracted from the mean value of the predicted superstate ($\tilde{S}^{\pi}_{t_j}$). The I-Q voting theory is employed to vote for the most probable IQ values encoded in $\mathbb{D}_{t_j}$ and obtained from (38). The radio will vote to similar $\mathbb{D}_{t_j}$ values, where the votes along with $\mathbb{D}_{t_j}$ and the corresponding superstates (the predicted ones) will be stored in a cell to be used later on. To understand how much the jammer shifted the signal with respect to the center of the expected superstate, the radio picks the $\mathbb{D}_{t_j}$ value which has the maximum number of votes from the cell stored during the real-time process and extract consequently the corresponding derivatives that realize the jammer's force (U_{jammer}).

3) Jammer Characterization at Observation Level (Observational Characterization): the characteristics obtained at the discrete level are forwarded towards the observation level to calculate the generalized error at this level ($\varepsilon_{\tilde{Z}_t}$) and explain such error as well. From the higher level, the radio can know which superstates of the model are affected by the jammer. Calculating the distance from the superstates' centroid allows to extract the source of the cause (jammer) that affected the shift noticed at higher levels. So, $\varepsilon_{\tilde{Z}_t}$ can be calculated in the following way:

$$\tilde{Z}^{J}_{t} = \overbrace{\tilde{Z}_t - H\mu(\underset{\tilde{S}_t\in\boldsymbol{S}}{\operatorname{argmax}}\,\lambda(\tilde{S}_t))}^{\textbf{Generalized Error 3 } (\varepsilon_{\tilde{Z}_t})} \tag{39}$$

which represent the jammer's Generalized State ($\tilde{Z}^{J}_{t}$) from which the radio can extract the jamming signal.

E. Incremental Learning of new models

The information obtained in the previous steps will be used here to incrementally learn the new model at two hierarchical levels, at the discrete level by updating the transition matrix of model ($m = 1$) and at the continuous level by updating the linear model associated to the reference model ($m = 1$).

1) Update Transition Matrix: The difference between variables whose belief is given by $\lambda(\tilde{S}^{m}_{t})$ and $\pi(\tilde{S}^{m}_{t})$ denotes the discrete Generalized Error ($\varepsilon(\tilde{S}^{m}_{t})$) which represents the $innovation$ provided at the discrete level by PF. As mentioned before $\lambda(\tilde{S}^{m}_{t})$ is a vector containing L elements and it is the same for all the particles propagated by the PF at time instant t. $\pi(\tilde{S}^{m}_{t})$ is a 1xL vector picked from the transition matrix and $\varepsilon(\tilde{S}^{m}_{t})$ is a KxL matrix where K is the total number of the elements in set $\mathbb{S}$ which is defined in (14). Therefore $\varepsilon(\tilde{S}^{m}_{t})$ can be defined as:

$$\varepsilon(\tilde{S}^{m}_{t}) = \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_K \end{bmatrix} = \begin{bmatrix} \lambda_1 - \pi_{\mathbb{S}_k 1}, & \lambda_2 - \pi_{\mathbb{S}_k 2}, & \dots & \lambda_L - \pi_{\mathbb{S}_k L} \\ \vdots & \vdots & \vdots & \vdots \\ \lambda_1 - \pi_{\mathbb{S}_K 1}, & \lambda_2 - \pi_{\mathbb{S}_K 2}, & \dots & \lambda_L - \pi_{\mathbb{S}_K L} \end{bmatrix} \tag{40}$$

where $\mathbb{S}_k$ denotes the k-th element (superstate) in the set $\mathbb{S}$. $\varepsilon(\tilde{S}^{m}_{t})$ is a zero-mean vector which contains positive and negative elements. At each time instant, t_j (i.e., when the jamming attack takes place, which is obtained from the abnormality signal classification), the i^{th} row vector related to the winning particle extracted from the transition matrix is updated following:

$$\Pi^{'}_{t_j} = \Pi + \varepsilon(\tilde{S}^{m}_{t_j}) = \begin{bmatrix} \pi_{11} + \varepsilon_1(1) & \pi_{12} + \varepsilon_1(2) & \dots & \pi_{1L} + \varepsilon_1(L) \\ \vdots & \vdots & \vdots & \vdots \\ \pi_{K1} + \varepsilon_K(1) & \pi_{K2} + \varepsilon_K(2) & \dots & \pi_{KL} + \varepsilon_K(L) \end{bmatrix} = \begin{bmatrix} \pi^{'}_{11} & \dots & \pi^{'}_{1L} \\ \vdots & \ddots & \vdots \\ \pi^{'}_{K1} & \dots & \pi^{'}_{KL} \end{bmatrix} \tag{41}$$

After considering the whole time where we detect jamming signal, the updated transition matrix will be:

$$\Pi^{''} = \mathbb{E}\left[\Pi^{'}_{t_j}, \dots, \Pi^{'}_{t_j+e}\right] \tag{42}$$

where $\Pi^{''}$ denotes the mean value of all the saved versions of the updated transition matrix Π' and $t_j + e$ is the time instant when the jammer ends the attack.

2) Update Dynamic Model: after characterizing the jammer at the continuous level to learn the rules it followed to attack the commands, we can update the dynamic rules of the original dynamic model by adding the jammer's force. Therefore, the updated dynamic model which represents the new situation (signal + jammer) can be written as follows:

$$\tilde{X}_t = A\tilde{X}_{t-1} + B\big(U_{\tilde{S}^m_{t-1}} + U_{jammer}\big) + w_t \qquad (43)$$

where A, B and w_t are the same as in (5) and U_{jammer} is obtained from jammer characterization at continuous level in Sec.IV-D2. In this way by using the updated dynamic model (m+1), the radio will be able to predict the jammer's effect on the commands in terms of I and Q values at multiple sub-carriers.

3) Relation b/w incremental learning and the detected abnormalities: Abnormalities are the surprising patterns in the observation that are previously not seen. Detecting abnormalities means that the radio is surprised by the new measurement. This surprise is due to the fact that the radio was expecting to receive signals that are generated according to the dynamic model encoded in its brain but in fact, improbable signals are received. The analytical meaning of incremental learning is related to the Free-energy principle [34]. The objective is to minimize the free-energy (i.e. the prediction error amount) by continuous correction of the dynamic model that represents the radio environment. This can be done by encoding a new probabilistic representation or by updating the current version of that representation. According to [34], the free-energy can be defined as a function of sensory data (observations) and dynamic model states (prediction):

$$\mathcal{F}(\tilde{Z}, \mathcal{M}) = < \mathcal{L}(t) >_q - \mathcal{H}(t) \qquad (44)$$

where the free-energy comprises the energy ($<\cdot>$) expected under a certain density q (that consists the statistical parameters of the generative model) and its entropy $\mathcal{H}$. $\mathcal{L}$ is the dynamic model encoded in the radio's brain that generates expected data samples and their causes. In the proposed framework (the generative model as HDBN), the free-energy can be reduced after calculating the generalized errors between top-down and bottom-up messages passing among hierarchical levels and consequently learning an appropriate model ($\mathcal{M}$) that optimize the predictions about hidden states generating observations ($\tilde{Z}$) and thus minimizing the free-energy. Hence, the solution is given by:

$$\mathcal{M} = \underset{\mathcal{M}}{\operatorname{argmin}}\, \mathcal{F}(\tilde{Z}, \mathcal{M}), \qquad (45)$$

which leads to the following optimization problem expressed in terms of abnormality measurements defined before:

$$\begin{cases} \pi(\tilde{S}^m_t) = \underset{\pi(\tilde{S}^m_t)}{\operatorname{argmin}} \Big(D_{KL}\big(\pi(\tilde{S}^m_t) || \lambda(\tilde{S}^m_t)\big) \Big) \\ \pi(\tilde{X}_t) = \underset{\pi(\tilde{X}_t)}{\operatorname{argmin}} -ln\Big[\mathcal{BC}\big(\pi(\tilde{X}_t(\tilde{S}^m_t)), \lambda(\tilde{X}_t(\tilde{S}^m_t))\big)\Big] \end{cases} \qquad (46)$$

The objective is to improve the radio's predictive ability at hierarchical levels. Prediction at the discrete level can be optimized if the measurement votes for the most probable particles (generated by the PF) all the time as discussed in Subsection IV-C1. During abnormal situation, the most probable particles are not voted by the observation providing by that a high D_{KL} as shown in (19). This means that the dynamic transition matrix ($\pi(\tilde{S}^m_t)$) must be changed in a way that it will be supported by the observation ($\lambda(\tilde{S}_t{}^m)$) which leads to minimize the $\mathcal{KLD}_{\mathcal{N}}$ defined in (18), and thus the overall $\mathcal{KLDA}$ will decrease. The rules of how the dynamic transition matrix must be updated is shown in Subsections IV-D1 and IV-E1. At the continuous level, the abnormality measurement can be optimized if the overlap area between the prediction $\pi(\tilde{X}_t)$ and the measurement $\lambda(\tilde{X}_t)$ in the normal region $\mathcal{R}_N$ is quasi-1 as discussed in Subsection IV-C2. During abnormal situation the overlap is very small (especially when the jammer attacks with high power). Thus, the prediction must be shifted towards the measurement, maximizing by that the overlap in the normal regions ($\mathcal{R}_N$) and minimizing it in the abnormal region ($\mathcal{R}_A$). Such shift realize the jammers power which can be learned during the jammer characterization process discussed in Subsection IV-D2 and consequently update the linear dynamic model as shown in Subsection IV-E2.

F. Action Selection

The jammer characterization done in the previous step at different hierarchical levels allows the radio to understand the jammer's nature of how (power), when (time) and where (frequency) it is attacking the commands. This offers several benefits to the radio including self-decision and self-action that support the radio to enhance the physical layer security. After extracting the jammer's signal the radio can suppress the jammer from the current observation and then overcome the issue of false commands or high error probabilities. This self-correction of the jammed signal realizes an auto-defence technique against the attacker threat without the help of other entity in the network. This reduces the time to act against the jammer, rather than sending feedback and then waiting for a response which increases the time of action during the real-time process. The Generalized Error defined in (39), produces the generalized state vector of the jammer ($\tilde{Z}^J_t$), which by the way can be fed to an unsupervised technique to be clustered allowing the radio not only to study statistically the effect of the jammer on the received commands (interaction b/w jammer and user) but also to study how the jammer's dynamics are evolving with time using probabilistic reasoning. This can be very useful when the jammer changes the strategy of the attacks, e.g. changes the output power while attacking the radio (to be investigated in future work). Regrading the proposed scenario we supposed that the jammer's output power will not change during the attacks. Thus, we supposed that there is only one cluster (superstate) of the jammer's dynamic model which encodes all the components of the $\tilde{Z}^J_t$. As done before after obtaining the superstates of the normal signal (commands) we can calculate the mean and covariance of the stand-alone superstate of the jammer from which the jammer's control vector (U_{jammer}) can be obtained. In addition, the radio can extract (or estimate) the jammer's signal from the $\tilde{Z}^J_t$ at multiple sub-carriers. $\tilde{Z}^J_t$ consists of the signal and

the noise and can be expressed as: $\tilde{Z}_t^J = \hat{J}_t + w_t$, thus, the jammer's signal can be estimated using: $\hat{J}_t = \tilde{Z}_t^J - \hat{w}_t$ where $\hat{J}_t$ is the estimated signal of the jammer extracted from $\tilde{Z}_t^J$, and $\hat{w}_t$ can be estimated using the reference model ($m = 1$). After extracting the jammer's signal, the radio can decide to suppress its effect (internal action) from observation before entering to the demodulator and decoder blocks. The observation that represents the jammed signal can be expressed as: $Z_t = Y_t + J_t + w_t$; where Z_t is the jammed signal, Y_t the normal signal, J_t the jammer's signal and w_t is the channel's noise. Therefore, the corrected signal $Z_t^\dagger$ will be decomposed as:

$$Z_t^\dagger = Z_t - \hat{J}_t = Z_t - \tilde{Z}_t^J + \hat{w}_t \tag{47}$$

V. Experimental Results

We conduct an extensive Monte Carlo simulation to evaluate the performance of the proposed framework using simulated data. Firstly, the trajectory of a quadcopter UAV is simulated based on [35]. A relationship between the commands and velocities of the UAV at different angles (Pitch, Yaw and Roll) is studied to generate the appropriate bits for simulating the LTE signal. Similarly, the altered trajectory is also extracted from the jammed signal. The LTE signal is generated according to the 3GPP specifications [36], and the important parameters are defined in Table II. The flight time of the UAV is $T_{flight} = 30s$ consisting of 600 samples due to the fact that the position is measured by the GPS every 50 ms. In addition, the UAV receives a PRB every 50ms and extracts the RV that contains a set of commands sent over 9 consecutive sub-carriers in 1 OFDM symbol. Thus, during the T_{flight} the UAV will receive 600 sets of commands, corresponding to 600 OFDM symbols in time domain (Fig. 4). Each received set of commands indicate how the UAV will move in the 3D space. The normal signal and the jammer are QPSK modulated. The output of the QPSK modulator for both is normalized based on the average power. The normal signal has average power $P_S = 1W$. While the average power of the jammer is P_J.

TABLE II: LTE simulation parameters

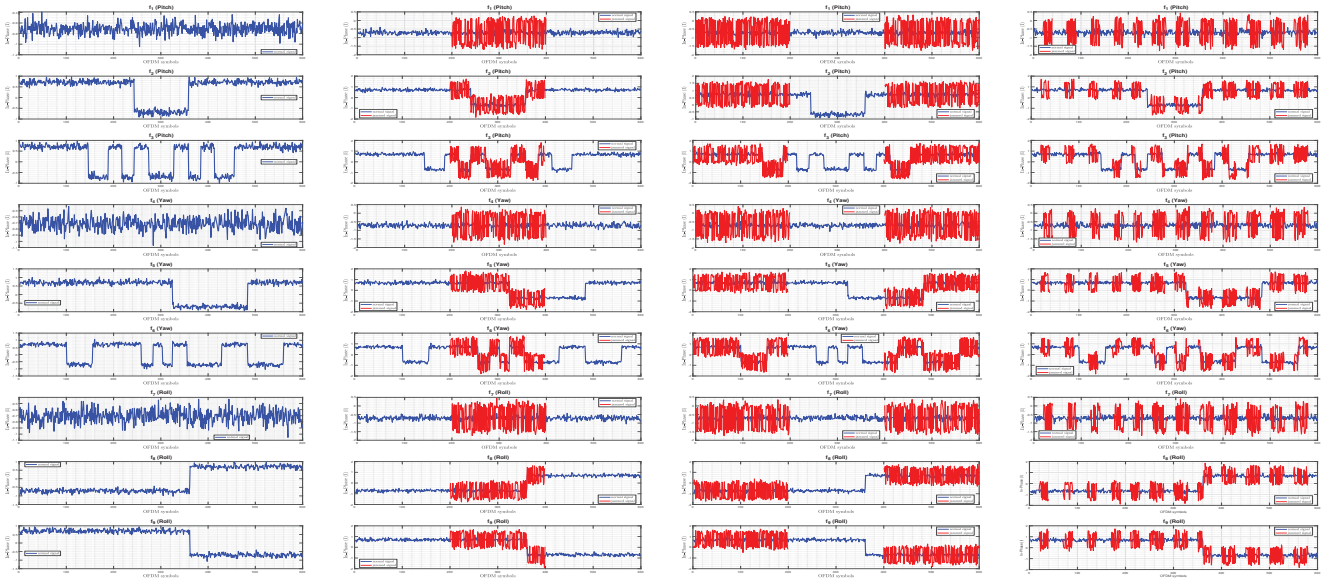
Parameter	BW	Duplex mode	Δf	Number of PRBs per BW	Sampling frequency	N_{FFT}	OFDM symbols per slot	CP length	SNR	Modulation	Channel	Total Radio Frames
Value	1.4 *MHz*	*FDD*	15 *kHz*	6	1.92 *MHz*	128	7	*Normal*	15 *dB*	*QPSK*	*AWGN*	600

Four different situations are considered, the first one is related to the **Reference Situation** representing a normal behaviour of the signal that carries original commands sent by the operator as shown in Fig. 7-a. The corresponding UAV trajectory is depicted in Fig. 8-a. The remaining situations are concerning the smart jammer behaviours in attacking the commands who has an average power $P_J = 1W$, and they are listed as follows: **Situation 1**, the jammer is attacking consecutively starting from time (in terms of OFDM symbols) $t = 200$ till $t = 400$ as shown in Fig. 7-b, where the altered UAV trajectory during the jamming attacks is shown in Fig. 8-b. In **Situation 2**, the jammer behaves in a dynamic fashion by attacking from $t = 1$ till $t = 200$, and from $t = 400$ till $t = 600$ as shown in Fig. 7-c, while, Fig. 7-d (**Situation 3**) illustrates a faster dynamic behaviours of the jammer.

Initially, the radio (UAV) does not have any knowledge (null memory) about the surrounding environment. Thus, at the beginning stage which is the *first iteration* exhibited in Fig. 5, the UAV predicts the future states of the spectrum supposing that the signals' states are static and do not change with time by employing UKF. Such an assumption leads to high abnormalities all the flight time since the UAV fails to predict the real states of the signal, as shown in Fig. 9-a. Accordingly, based on these predictions and by using the innovations (derivatives) produced by the UKF, the UAV will form and store the generalized errors. Then it will perform an unsupervised clustering method (the GNG algorithm) in an offline manner to learn and memorize the first generative model which represent the dynamics of the received commands during the normal situation. After that, the UAV is capable to predict the future states of the commands at multiple sub-carriers. This can be verified by calculating the abnormality signal during the normal situation. If the abnormality is quasi-zero, i.e. the learned model succeeded to capture the dynamic rules of the signal and allowed the UAV to perform correct predictions as shown in Fig. 9-b where the abnormality at the continuous level defined in (23) is showed. Testing new observations Z_t and predicting eventually, could follow the same rules with which the dynamic model has been learned from previous experience (reference situation) when the jammer was absent or could deviate due to the new rules caused by the jammer. Thus, in Situation 1, the UAV can detect any attack while predicting the future states of the commands and receiving the observations as shown in Fig. 10-a-b. As well as extracting the jammer's signal (Fig. 11-a) and act accordingly by mitigating its effect on the received commands (Fig. 11-b) which leads to auto-correction of the altered trajectory in Fig. 8-c.

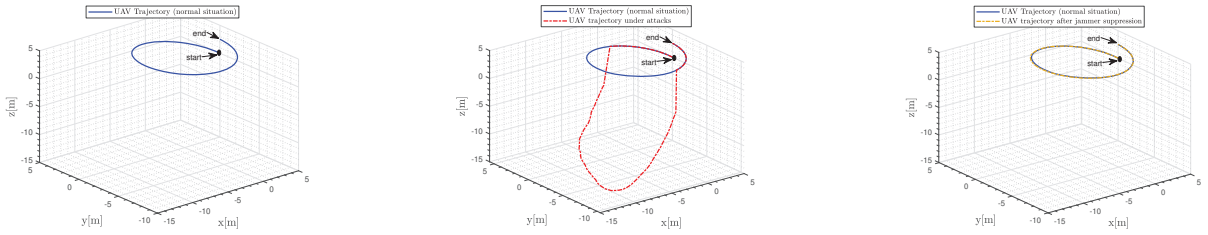
After this situation, the UAV can study the new behaviour (detected jammer) and learn incrementally a new dynamic model which represents the interaction of the jammer and the normal signal (commands). Facing a new situation (Situation 2), where the same jammer (detected before) is attacking the commands allow the radio to recognize it and predict its future activity inside the radio spectrum since it has already learned the rules of attacking. In this situation the UAV's memory contains two dynamic models: the UAV will *i)* switch between these models; *ii)* and select the best one that fits the observation. Fig. 10-c-d shows the abnormality measurements related to the reference model (learned during the reference situation) and that related to multiple models (reference model and the one learned incrementally that represents the situation where the jammer is ON). As expected the abnormality at both levels is decreased, which means that the UAV succeeded to learn the dynamic rules of the jammer in attacking the commands. Additionally, another new experience (Situation 3) has been tested by considering fast dynamics of attacking the commands by the same jammer. Fig. 10-e-f, confirms that the UAV has been succeeded in characterizing the jammer in question and learn its model incrementally, and then to predict its effect in future situations.

Further experiments are tested to validate the proposed approach, by varying the Jamming-to-Signal-Ratio (JSR) from



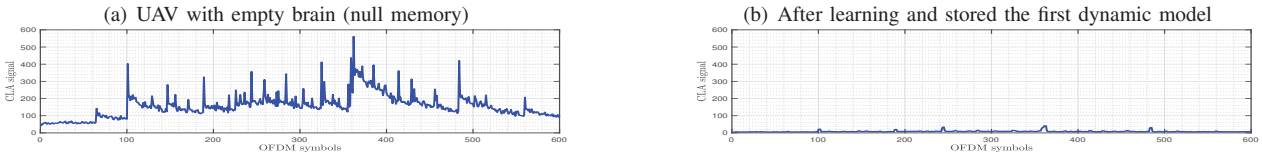
(a) Reference Situation (b) Situation 1 (c) Situation 2 (d) Situation 3

Fig. 7: Received Commands at multiple sub-carriers at different situations



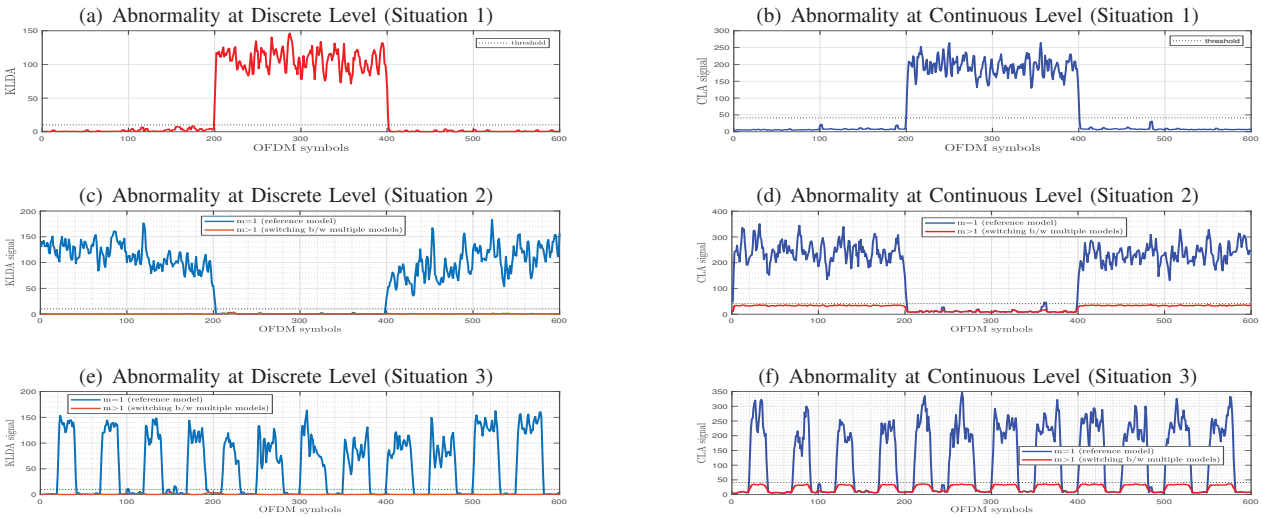
(a) Reference Situation (b) Situation 1 (c) Situation 1 (after Suppression)

Fig. 8: UAV Trajectories.



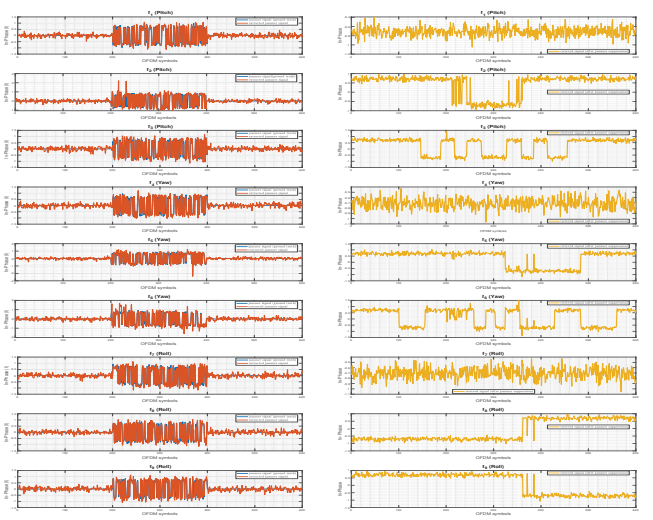
(a) UAV with empty brain (null memory) (b) After learning and stored the first dynamic model

Fig. 9: Abnormality Measurements at the continuous level



(a) Abnormality at Discrete Level (Situation 1) (b) Abnormality at Continuous Level (Situation 1)
(c) Abnormality at Discrete Level (Situation 2) (d) Abnormality at Continuous Level (Situation 2)
(e) Abnormality at Discrete Level (Situation 3) (f) Abnormality at Continuous Level (Situation 3)

Fig. 10: Jammer Detection at hierarchical levels during different situations



(a) Jammer's signal Extraction (b) Received signal after jammer suppression

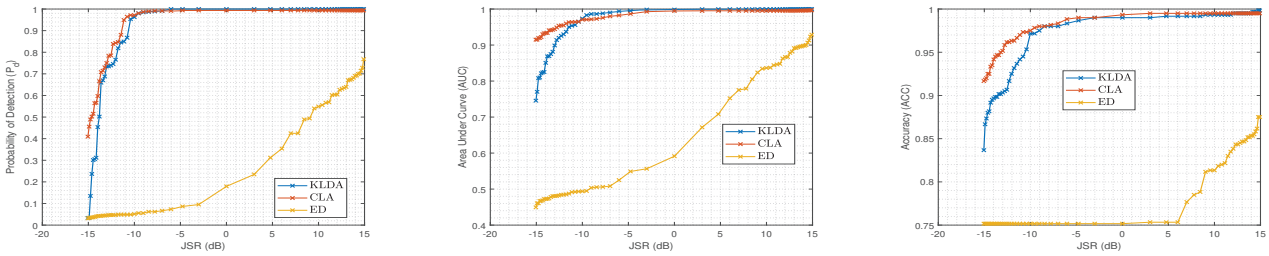
Fig. 11: Jammer extraction and jammer suppresion after characterizing the jamming attacks

$-15dB$ to $+15dB$. In all these experiments the jammer attacks dynamically all the 9 sub-carriers in frequency domain and certain OFDM symbols in time domain. Particularly, from $t = 1$ till $t = 150$, from $t = 200$ till $t = 350$, and from $t = 400$ till $t = 550$. In order to evaluate the performance of the proposed framework in detecting the jamming attacks referring to the reference model learned by the UAV, we used a range of confidence thresholds to build the corresponding ROC curves illustrated in Fig. 12-a along with the Area Under Curve (AUC) (Fig. 12-b) and Accuracy (ACC) (Fig. 12-c). The ROC curves show that the MJPF filtering on the reference model provides high detection probability (P_d) at both levels even when the jammer attacks with very low power. For example, with JSR = -5 dB, the probability of detection is almost 100% for both KLDA and CLA measurements. The high detection probabilities can be explained by the fact that the predictions performed at the higher level of the HDBN using PF were precise and accurate almost of the time considered in the analysis. The probability of detecting the jammer at the discrete level is calculated by using: $P_d = Pr\{KLDA > (1-\zeta)\alpha\}$. We estimated the acceptance ratio (ζ) as $\zeta = 0.8$ after observing the predicted superstates and the observed ones. This also implies that the learned model is accurate in predicting the discrete variables as can be observed from Fig. 10-a where the KLDA signal is quasi-zero (i.e. observation matches the prediction) in the time instants when the jammer is OFF. Directly estimating α is difficult here since the UAV does not has any prior knowledge about the jammer. However, it can be estimated by supposing that half of the probability mass function of the observation ($\lambda(\tilde{S}_t)$) falls in the normality region and the other half in the abnormality region (the concept of α is discussed in Section IV-C1). Thus, we set $\alpha = 0.5$. In this case, the threshold is adapted to different situations. On the other hand, the probability of detecting the jammer at the continuous level is calculated by using: $P_d = Pr\Big\{CLA > \sqrt{Q_{\mathcal{T}}(\tilde{X}_t)\ .\ [\beta_1(1-\alpha)] + Q_{\mathcal{T}}(\tilde{X}_t)\ .\ [\beta_2(1-\alpha)]}\Big\}$. In the numerical results, we chose $Q_{\mathcal{T}} = 0.2$ which is estimated after observing the *CLA* signal during normal situation shown in Fig. 9-b where the abnormal signal is quasi-0 implying that the learned generative model is accurate to predict continuous variables. Also here α is the area of $\lambda(\tilde{X}_t)$ falling inside the normal regions as discussed in Section IV-C2 and it is difficult to estimate it as claimed before so we set $\alpha = 0.5$ while $\beta_1 = 1$ and $\beta_2 = 0$. In addition, a comparison with the traditional Energy detector (ED) is provided as shown in Fig. 12 that shows how the proposed approach beats the Energy detector (ED) in detecting the jammer at different JSRs.

The characterization of the jammer is performed by the UAV during the first interval (from 1 to 150 OFDM symbols) and the successive periods are used to validate if the UAV succeeded to capture the dynamics of such attacks. This capability can be verified by calculating the Root Mean Square Error (RMSE) of the abnormality measurements (KLDA and CLA) and comparing these errors in two cases. The first case ($m = 1$) is when the UAV relay on the reference model (the first dynamic model) while the second case ($m > 1$) is based on switching between the reference model and the one learned incrementally which encodes the dynamic rules of the jammer in question. RMSE is the difference between prediction and evidence realizing the prediction accuracy, high RMSE means that the evidence does not match the prediction while low RMSE stands for the fact that prediction matches the evidence. Also, RMSE depends on the abnormality level which increases as the jammer's power increase as shown in Fig. 13 (blue curves). In the second case, by switching between two models the UAV can predict the jammer's effect in future OFDM symbols and cause a decrement of the RMSE respectively. This can be seen in Fig. 13 (red plots) where the RMSE of KLDA (Fig. 13-a) and RMSE of CLA (Fig. 13-b) converge and are somehow stable and decreased with respect to the RMSE related to the reference model. This implies that the abnormality level is stable too since the UAV is predicting correctly the jammer's activity and not detecting unexpected behaviours any more (not surprised any more).

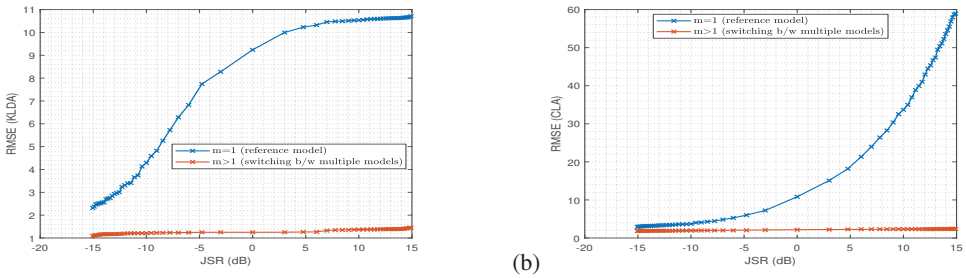
VI. Conclusion and Future Work

In this paper, we proposed a general framework for Cognitive UAV Radios by introducing an emergent data-driven Self-Awareness (SA) module to enhance the physical layer security. This module allows the radio to build up its own memories incrementally by observing the stimulus received from the radio environment and learning with reasoning a hierarchical representation of such observation. The radio augmented with SA is capable to predict the radio environment and identify any abnormality within a received signal. The SA module is also capable of characterizing the abnormal situation caused by a jammer while studying the rules of how (Power), when (Time) and where (Frequency) the jammer is attacking. Additionally, an incremental learning process is proposed for the radio to learn a new model that represents the new situation, decide and act efficiently by suppressing the jamming signal. All



(a) (b) (c)

Fig. 12: Performance comparison between the SA module (KLDA, CLA) and Energy detector (ED): ROC curves (a) and the corresponding AUC (b) and ACC (c)



(a) (b)

Fig. 13: RMSE comparison versus JSR by using only the reference model (blue curves) and switching between two models (red curves). (a) RMSEs of KLDA. (b) RMSEs of CLA

these tasks are performed by the radio itself in an incremental approach without any external supervision, and the proposed framework is generalized enough to be employed in different radio applications. We show that the proposed HDBN framework accurately characterizes the jammer's behaviours in different situations while the probability of detection is significantly high even at low Jamming-to-Signal-Ratio. The results also show that after learning the jammer's behaviors, the UAV with the proposed framework can correctly predict the future activities of the jammer, which can eventually help in mitigating any future attacks. In the future, we will analyse the jammer classification in a more practical multiple jammers scenario as well as introducing interactive learning models between the jammer and the user to design an anti-jamming technique.

REFERENCES

[1] J. Du, W. Xu, Y. Deng, A. Nallanathan, and L. Vandendorpe. Energy-saving UAV-Assisted Multiuser Communications with Massive MIMO Hybrid Beamforming. *IEEE Communications Letters*, pages 1–1, 2020.

[2] D. Bin, M. Ruijiu, N. Kong, and D. Sun. Distributed Data Fusion for On-Scene Signal Sensing with a Multi-UAV System. *IEEE Transactions on Control of Network Systems*, pages 1–1, 2020.

[3] X. Liang, W. Xu, H. Gao, M. Pan, J. Lin, Q. Deng, and P. Zhang. Throughput Optimization for Cognitive UAV Networks: A Three-Dimensional-Location-Aware Approach. *IEEE Wireless Communications Letters*, pages 1–1, 2020.

[4] J. Mitola and G. Q. Maguire. Cognitive radio: making software radios more personal. *IEEE Personal Communications*, 6(4):13–18, Aug 1999.

[5] Shujun Li and Mieczyslaw Kokar. *Flexible Adaptation in Cognitive Radios*. 01 2013.

[6] X. Wang, J. Wang, Y. Xu, J. Chen, L. Jia, X. Liu, and Y. Yang. Dynamic Spectrum Anti-Jamming Communications: Challenges and Opportunities. *IEEE Communications Magazine*, 58(2):79–85, February 2020.

[7] G. Zhang, Q. Wu, M. Cui, and R. Zhang. Securing UAV Communications via Joint Trajectory and Power Control. *IEEE Transactions on Wireless Communications*, 18(2):1376–1389, 2019.

[8] A. Toma, A. Krayani, M. Farrukh, H. Qi, L. Marcenaro, Y. Gao, and C. S. Regazzoni. AI-Based Abnormality Detection at the PHY-Layer of Cognitive Radio by Learning Generative Models. *IEEE Transactions on Cognitive Communications and Networking*, 6(1):21–34, 2020.

[9] C. Regazzoni, L. Marcenaro, D. Campo, and B. Rinner. Multisensorial Generative and Descriptive Self-Awareness Models for Autonomous Systems. *Proceedings of the IEEE*, pages 1–24, 2020.

[10] X. Zhou, M. Sun, G. Y. Li, and B. Fred Juang. Intelligent wireless communications enabled by cognitive radio and machine learning. *China Communications*, 15(12):16–48, Dec 2018.

[11] Z. Qin and G. Y. Li. Pathway to Intelligent Radio. *IEEE Wireless Communications*, 27(1):9–15, February 2020.

[12] Z. Qin, X. Zhou, L. Zhang, Y. Gao, Y. Liang, and G. Y. Li. 20 Years of Evolution From Cognitive to Intelligent Communications. *IEEE Transactions on Cognitive Communications and Networking*, 6(1):6–20, March 2020.

[13] L. Gavrilovska, V. Atanasovski, I. Macaluso, and L. A. DaSilva. Learning and Reasoning in Cognitive Radio Networks. *IEEE Communications Surveys Tutorials*, 15(4):1761–1777, Fourth 2013.

[14] P. Cheng, Z. Chen, M. Ding, Y. Li, B. Vucetic, and D. Niyato. Spectrum Intelligent Radio: Technology, Development, and Future Trends. *IEEE Communications Magazine*, 58(1):12–18, January 2020.

[15] S. Rajendran, W. Meert, D. Giustiniano, V. Lenders, and S. Pollin. Deep Learning Models for Wireless Signal Classification With Distributed Low-Cost Spectrum Sensors. *IEEE Transactions on Cognitive Communications and Networking*, 4(3):433–445, Sep. 2018.

[16] S. Rajendran, W. Meert, V. Lenders, and S. Pollin. Unsupervised Wireless Spectrum Anomaly Detection With Interpretable Features. *IEEE Transactions on Cognitive Communications and Networking*, 5(3):637–647, Sep. 2019.

[17] M. Walton, M. Ayache, L. Straatemeier, D. Gebhardt, and B. Migliori. Unsupervised Anomaly Detection for Digital Radio Frequency Transmissions. In *2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA)*, pages 826–832, Dec 2017.

[18] N. Tandiya, A. Jauhar, V. Marojevic, and J. H. Reed. Deep Predictive Coding Neural Network for RF Anomaly Detection in Wireless Networks. In *2018 IEEE International Conference on Communications Workshops (ICC Workshops)*, pages 1–6, May 2018.

[19] M. A. Conn and D. Josyula. Radio Frequency Classification and Anomaly Detection using Convolutional Neural Networks. In *2019 IEEE Radar Conference (RadarConf)*, pages 1–6, April 2019.

[20] J. Mitola. Cognitive Radio an integrated agent architecture for software

defined radio. *Ph.D. dissertation, Royal Institute of Technology (KTH), Kista, Sweden*, 2000.

[21] Lise Safatly, Mario Bkassiny, Mohammed Husseini, and Ali El-Hajj. Cognitive Radio Transceivers: RF, Spectrum Sensing, and Learning Algorithms Review. *International Journal of Antennas and Propagation*, 2014, 03 2014.

[22] Jacques. Palicot, Joseph. Mitola, Zander. (Zhongding) Lei, and Friedrich. K. Jondral. Special issue on 10 years of cognitive radio: state-of-the-art and perspectives. *EURASIP Journal on Wireless Communications and Networking*, 2012.

[23] V. Bassoo and N. Khedun. Improving the quality of service for users in cognitive radio network using priority queueing analysis. *IET Communications*, 10(9):1063–1070, 2016.

[24] J. Mitola. Cognitive Radio Architecture Evolution. *Proceedings of the IEEE*, 97(4):626–641, 2009.

[25] S. K. Jayaweera. *The Cognitive Radio*, pages 27–41. Wiley, 2015.

[26] 3GPP TR 36.777. Technical specification group radio access network: study on enhanced LTE support for aerial vehicles. V15.0.0, Dec., 2017.

[27] H. C. Nguyen, R. Amorim, J. Wigard, I. Z. KováCs, T. B. Sørensen, and P. E. Mogensen. How to Ensure Reliable Connectivity for Aerial Vehicles Over Cellular Networks. *IEEE Access*, 6:12304–12317, 2018.

[28] J. Stanczak, D. Kozioł, I. Z. Kovács, J. Wigard, M. Wimmer, and R. Amorim. Enhanced Unmanned Aerial Vehicle Communication Support in LTE-Advanced. In *2018 IEEE Conference on Standards for Communications and Networking (CSCN)*, pages 1–6, Oct 2018.

[29] X. Gao, Z.-Y. Zhang, and L.-M. Duan. A quantum machine learning algorithm based on generative models. *Science Advances*, 4(12), 2018.

[30] Kevin Murphy. *Dynamic Bayesian Networks: Representation, Inference and Learning*. PhD thesis, 01 2002.

[31] E. Fox, E. B. Sudderth, M. I. Jordan, and A. S. Willsky. Bayesian nonparametric inference of switching dynamic linear models. *IEEE Transactions on Signal Processing*, 59(4):1569–1585, 2011.

[32] Bhashyam Balaji and Karl Friston. Bayesian State Estimation Using Generalized Coordinates. *Proc SPIE*, 8050, 05 2011.

[33] M. Baydoun, D. Campo, V. Sanguineti, L. Marcenaro, A. Cavallaro, and C. Regazzoni. Learning Switching Models for Abnormality Detection for Autonomous Driving. In *2018 21st International Conference on Information Fusion (FUSION)*, pages 2606–2613, July 2018.

[34] Harriet Feldman and Karl Friston. Attention, Uncertainty, and Free-Energy. *Frontiers in Human Neuroscience*, 4:215, 2010.

[35] N. Michael, D. Mellinger, Q. Lindsey, and V. Kumar. The GRASP Multiple Micro-UAV Testbed. *IEEE Robotics Automation Magazine*, 17(3):56–65, Sep. 2010.

[36] 3GPP. Evolved Universal Terrestrial Radio Access (E-UTRA); LTE physical layer; General description (Release 14). Technical Specification (TS) 36.201, 3rd Generation Partnership Project (3GPP), 2017.