Received 8 December 2025; accepted 30 December 2025. Date of publication 12 January 2026; date of current version 15 January 2026.

Digital Object Identifier 10.1109/OJCOMS.2026.3652319

Integrated Sensing and Communication for Blockage Detection in V2X Networks

SALEEMULLAH MEMON [1,2] (Member, IEEE), ALI KRAYANI [1] (Member, IEEE), PAMELA ZONTONE [1] (Member, IEEE), LUCIO MARCENARO [1] (Senior Member, IEEE), DAVID MARTÍN GÓMEZ [2] (Member, IEEE), AND CARLO REGAZZONI [1] (Senior Member, IEEE)

[1]Department of Electrical, Electronic, Telecommunications Engineering and Naval Architecture, University of Genoa, 16126 Genoa, Italy
[2]Department of Systems Engineering and Automation, University Carlos III of Madrid, 28911 Leganés, Spain

CORRESPONDING AUTHOR: S. MEMON (e-mail: saleemmemon1320@gmail.com)

This work was supported in part by the European Union's Horizon Europe Research and Innovation Programme under Grant 101121134 (B-prepared); in part by the European Union under the Italian National Recovery and Resilience Plan (PNRR) of NextGenerationEU Partnership on "Telecommunications of the Future" (PE00000001âĂŞProgram "RESTART") under Grant CUP E63C22002040007 (D.D. n.1549 of 11/10/2022); in part by the Spanish Government funded by MICIU/AEI/10.13039/501100011033 and ERDF/EU under Grant PID2024-157191OB-C21, and funded by MCIN/AEI/10.13039/501100011033 under Grant PID2022-140554OB-C32; and in part by the Comunidad de Madrid under Grant SEGVAUTO 5Gen-CM (TEC-2024/ECO-277) and Grant DIGIMATER-CM (TEC-2024/TEC-102).

ABSTRACT The integration of vehicle-to-everything (V2X) communication paradigms with sixth-generation (6G) wireless networks and artificial intelligence (AI) frameworks enables ultra-reliable, low-latency communication, which is essential for real-time decision-making in autonomous vehicles (AVs) and smart cities. Proprioceptive and exteroceptive sensors allow AVs to perceive both their internal states and external surroundings, ensuring rapid responses to critical events. Integrated sensing and communication (ISAC) enhances this capability by jointly leveraging perception and communication, enabling V2X systems to adapt intelligently to real-time emergencies. In this paper, we propose a probabilistic, data-driven, hierarchical, interactive, and explainable approach for an intelligent agent, i.e., a base station (BS), to learn the dynamic environmental perception from the 3D LiDAR point clouds and the strength of radio-frequency (RF) power signals between the connected BS and vehicles. An interactive coupled Markov jump particle filter (IC-MJPF) is proposed in the inference phase to leverage the probabilistic information provided by an interactive coupled generalized dynamic Bayesian network (IC-GDBN) to predict various types of LiDAR and RF power blockages, as well as to detect real-time abnormalities in an unsupervised manner arising from dynamic environmental changes. Experimental results demonstrate that the proposed approach consistently outperforms existing baseline studies, achieving superior performance in terms of blockage detection accuracy within 50 milliseconds across various blockage situations. These findings underscore the robustness and effectiveness of the proposed framework in addressing both physical and digital blockage challenges within the ISAC domain for connected V2X networks.

INDEX TERMS Integrated sensing and communication (ISAC), vehicle-to-everything (V2X), interactive coupled generalized dynamic Bayesian network (IC-GDBN), blockage detection, anomaly detection, 3D LiDAR point clouds.

I. INTRODUCTION

MODERN vehicular technologies have rapidly attracted significant attention from smart cities, industrial automation, and telecommunication industries, enabling the emergence of vehicle-to-everything (V2X) communications to enhance road safety, traffic management services, and driving comfort. V2X communication enables vehicles to exchange information with various external entities in real time, such as infrastructure (V2I), pedestrians (V2P), networks (V2N), and other vehicles (V2V). Encouraged by sixth-generation (6G) and beyond communication technologies, V2X is emerging as a cornerstone for the next generation of connected and autonomous vehicles (AVs) [1]. Although this technology offers numerous advantages for

intelligent transportation systems, it still faces several technical challenges, primarily related to security and performance.

Integrated sensing and communication (ISAC) has recently emerged as a paradigm feature of 6G, with the potential to support intelligent decision-making and adaptive learning in future V2X systems [2], [3]. This integration of sensing and communication allows vehicles to simultaneously perceive their environment and exchange information, thus paving the way toward a safer transportation system [4]. In the context of 6G and V2X applications, ISAC enables more efficient use of the spectrum and supports real-time decision-making by providing both environmental awareness and network connectivity within a single framework.

Modern vehicles are equipped with a range of proprioceptive and exteroceptive sensors such as, GPS, cameras, radar, LiDAR, communication modules, pedals, steering systems, and wheel speed sensors, that enable them to perceive both their internal states and external surroundings. By leveraging this rich, coexisting data, researchers have explored various tasks to enhance self-awareness and situational awareness in V2X systems. These tasks include jamming and spoofing detection [5], abnormal behavior detection [6], future blockage prediction [7], localization and severity estimation [8], object detection [9], collision avoidance [10], and emergency braking and activity recognition [11].

Sensor blockage, or occlusion, is a serious challenge for ISAC-enabled V2X communication systems. It occurs when a sensor is obstructed by physical objects such as buses, pedestrians, or other vehicles. Such situations may also arise due to sensor failures or operation under harsh and adverse weather conditions, such as wind, snow, rain, and fog. As a result, the system may fail to accurately perceive its surroundings, leading to errors or failures in navigation, continuous tracking, obstacle detection, and decision-making [12]. Moreover, high-frequency V2X communication systems rely on line-of-sight (LOS) links, which provide sufficient received signal power due to the shorter wavelength of the transmitted signals. These systems are highly susceptible to penetration loss, particularly in the millimeter-wave (mmWave) bands. Consequently, the blockage problem becomes more severe, posing significant challenges to the latency and reliability of such networks. For instance, if these links are blocked by moving objects, it can lead to abrupt performance degradation or even complete link disconnection [13].

In non-line-of-sight (NLOS) scenarios, signals experience absorption and multipath propagation caused by reflections, diffractions, and scattering from walls, buildings, trees, and other surrounding objects. Such interactions lead to significant attenuation, resulting in data loss, increased latency, and reduced reliability. Reconfigurable intelligent surface (RIS) technology has emerged as a promising approach to address these challenges by integrating the environment into the communication infrastructure [14]. Instead of increasing transmit power to overcome obstructions, RIS enables the physical surfaces to be configured so that signals are steered directly toward the intended receivers [15], [16], thereby mitigating the effects of blockages and the resulting NLOS propagation in wireless channels.

Although RIS is designed to overcome NLOS blockages, its effectiveness depends heavily on strategic deployment and dynamic reconfiguration. In complex environments, RIS panels must be positioned intelligently to mitigate blockages and maintain reliable connectivity [17]. Blockages also introduce multipath propagation, where signals reflect off multiple surfaces, creating interference and complicating signal processing. This necessitates sophisticated algorithms capable of handling dynamically varying signal paths. ISAC systems may further require supplemental infrastructure, such as relays, base stations (BSs), roadside units (RSUs), or autonomous aerial vehicles (UAVs), to sustain performance under lossy channel conditions, increasing both deployment complexity and cost [18], [19]. Moreover, RIS must adapt in real time to evolving blockage conditions, which calls for predictive optimization and low-latency control mechanisms. These challenges highlight the need for robust design strategies to ensure that RIS-enabled ISAC systems for 6G and beyond operate effectively in diverse, dynamic, and blockage-prone V2X environments.

Another possible way to address blockage challenges is to detect potential blockages at an early stage before they occur. This allows artificial agents to take proactive actions, thereby reducing the risk of sensor and communication disruptions. Considering this, some AI-driven approaches are studied in [20], which can be utilized to enhance security within integrated networks. These methods include machine learning (ML), deep learning (DL), and reinforcement learning (RL), which support blockage prediction by leveraging movement patterns, prior observations, and other relevant information to anticipate future blockages. By learning from sensorial data and adapting to dynamic conditions, V2X systems can make more informed real-time decisions. Consequently, the throughput and reliability of V2X communication networks improve, while latency decreases, contributing to safer and more efficient transportation systems.

In the recent literature, some sensorial information-based blockage predictions with AI techniques under the ISAC framework have been discussed. A 2D LiDAR point cloud and mmWave communication-based convolutional neural network (CNN) model for blockage prediction is proposed in [21]. Similarly, [22] and [23] propose a vision-based CNN and recurrent neural network (RNN) combined model, and a radar sensing-based CNN and long short-term memory (LSTM) combined model, respectively, for LOS blockage prediction between the BS and the user. Although effective in specific scenarios, these learning approaches typically involve high computational complexity, require large amounts of labeled data, face challenges in real-time blockage labeling, and struggle to adapt to new or unseen situations. Moreover, these approaches yield a closed box nature, lack the ability to handle interactions between multiple objects, and do not possess self-awareness

or explainable capabilities. Other approaches such as end-to-end deep probabilistic models [24] and crossmodal attention mechanisms [25] can offer moderate levels of explainability, primarily through uncertainty quantification and attention maps. However, both approaches are fundamentally grounded in supervised or self-supervised learning, relying on large labeled or paired datasets to achieve their performance which can be challengeable as environment varies.

On the other hand, human-inspired cognitive reasoning and learning techniques that work like a human brain, which learns from mistakes and new experiences and updates its internal model belief through continual learning as the environment changes, could be a suitable option to adopt in the ISAC domain for blockage challenges. This bio-inspired concept has been explored by neuroscientists such as Friston, Haykin, and Damasio [26], [27], [28]. The framework comprises six self-aware capabilities: initialization, memorization, prediction, anomaly detection, new model creation, and decision-making. This cognitive cycle approach could be achieved through probabilistic reasoning and data-driven approaches that enable adaptive learning, operate hierarchically from a low observation level to high abstraction levels with a message-passing architecture between random variables, ensure explainability by modeling transitions and uncertainties from hidden patterns, and accept multisensorial fusion, which enhances accuracy by integrating various sensor inputs [29], [30]. This integration allows the modeling of complex interactions, where the state from one model influences the probability distribution of states in another, providing a unified probabilistic representation that accounts for uncertainty and dependencies [31]. Motivated by these discussions, this paper aims to enhance V2X communication by leveraging real-time sensorial data. Instead of designing physical communication systems, a data-driven method, namely, a 3D LiDAR and radio-frequency (RF) received power-based interactive coupled generalized dynamic Bayesian network (IC-GDBN), is introduced to predict and detect both sensor and power blockages within the ISAC framework. The proposed probabilistic IC-GDBN learns hierarchically from temporal LiDAR point cloud sequences and dynamic RF power patterns. Rather than explicitly labeling blockage causes (e.g., extreme weather or sensor degradation) in a supervised manner, the model enables data-driven detection and interpretation of anomalies based on learned dynamic rules encoded within its dictionaries. To achieve this, an unsupervised growing neural gas (GNG) clustering approach is employed on vehicle positions and velocities, as well as on received RF power signals and their derivatives. A higher-level hierarchy within the IC-GDBN integrates LiDAR sensing with the communication domain, allowing semantic relationship between vehicle trajectory information and V2I communication signals. This mapping facilitates the joint understanding of RF power variations and vehicle motion behavior. During online inference, an interactive coupled Markov jump particle filter (IC-MJPF) is applied to the learned global IC-GDBN model. This representation and inference framework enables the BS to predict and detect blockages. When new environmental conditions emerge that deviate from the learned dynamics, the model identifies the resulting anomalies and provides explanations. This anomaly detection capability is a key enabler for online adaptive learning, as it allows the system to recognize abnormalities in real-time whether caused by sensor malfunctions, environmental disturbances, or other unforeseen factors and respond accordingly in a dynamic manner. This addresses an important gap in current ISAC literature, including the lack of coherent, unsupervised anomaly detection and the explainable techniques that can be used to adapt models in real time [29], [32]. Our approach introduces a form of self-awareness to the agent, enabling it to detect and interpret abnormalities without prior labeling, which is essential in real-time and unpredictable environments. We believe this contribution opens a pathway toward handling multimodal sensor failures and extreme conditions through intelligent, data-driven, and self-aware adaptation. The main contributions of this paper are summarized as follows:

1) A 3D LiDAR perception and RF received power-based blockage detection scheme is proposed for an ISAC-enabled V2X scenario by learning a probabilistic and explainable IC-GDBN model. The model captures the cross-correlation between RF power signals received from multiple vehicles and their corresponding 3D trajectories, which are extracted from 3D LiDAR point cloud sequences. The two data modalities are stochastically coupled at a higher level of abstraction to represent their semantic relationships.
2) To learn a 3D LiDAR dynamic sensing environment by an agent, multiple GDBN models are trained for both normal and blocked vehicle tracks; in other words, separate vocabularies are modeled for continuous and broken tracks. The interactions among multiple vehicles are captured at the sensing word-level hierarchy over time evolution. similarly, to represent the communication environment, multiple GDBN models are trained for normal and blocked received power patterns, with interactions among the power signals received from multiple vehicles encoded at the communication word-level hierarchy.
3) The BS evaluates whether the evolution of both vehicle trajectories and the received RF power signals aligns with the dynamic rules defined for normal and blockage scenarios within the IC-GDBN. In the event of anomalies, the BS identifies the cause, such as the detection of a new type of blockage in vehicle trajectories or LOS links.
4) The blockage prediction performance of the proposed IC-GDBN model is compared with existing CNN and LSTM methods. The IC-GDBN demonstrates superior accuracy, consistently achieving better predictions

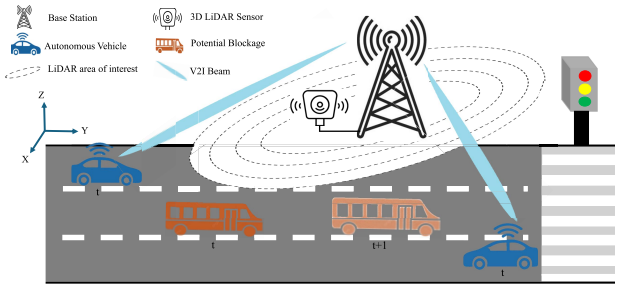


FIGURE 1. Illustration of a blockage scenario in the ISAC framework for V2X communication networks.

within 50 milliseconds across various blockage scenarios. This highlights its better learning ability and understanding of the complex dynamic environments.

The rest of this paper is structured as follows. The system model for the blockage environment is introduced in Section II. The description of the dataset used in this paper is given in Section III. The proposed method for blockage detection, including data preprocessing, representation of the RF environment, offline learning of IC-GDBN, and online inference through IC-MJPF, is described in detail in Section IV. The simulation results for the training and testing phases are discussed in Section V. Section VI concludes this work. Finally, the limitations and potential directions for further research are provided in Section VII.

II. ISAC ENABLED V2X SCENARIO

The dynamic system model, which represents the ISAC-enabled V2X scenario, is illustrated in Figure 1. It includes moving AVs, which serve as wireless transmitters, and a stationary BS that acts as a wireless roadside unit. The AVs communicate with the high-frequency BS through LOS V2I links. The BS utilizes a 3D LiDAR sensor, which accurately perceives the environmental awareness of static and other dynamic objects within its sensing range. The primary goal of LiDAR is to provide information about the mobility patterns, the location of AVs, dynamic scatters/blockages, pedestrians, and the range and angle information of other mobile objects at each time instant. A bus shown in the figure, moving from left to right at time t in lane 2, is expected to create a blockage shortly at time $t + 1$, for an AV moving in lane 3 from right to left side of the road. This sensor blockage issue leads to an inaccurate perception of the environment and significantly degrades LOS links, resulting in increased latency, packet loss, and reduced reliability in connected V2X networks. Furthermore, different types of dynamic blockages, such as short, long, and continuous blockages affecting the communication links, are illustrated in Figure 2. The duration of each blockage depends on the size, direction, speed of the mobile blockage and speed of AVs. With the help of environmental awareness enabled by the LiDAR sensor and dynamic RF received power patterns, the BS, acting as an intelligent agent, learns a probabilistic and hierarchical IC-GDBN model that encapsulates both sensing and communication environments. The agent leverages prior knowledge, encoded in the global dictionary, learned during the training phase, to predict and detect future blockages before their occurrence. In case of a predicted blockage, the intelligent agent can take proactive actions to avoid blockage situations. Consequently, this approach can increase reliability, reduce latency, and improve the performance of connected V2X networks.

III. DATA DESCRIPTION

This research is grounded in dataset scenarios collected in real-world environments, with particular emphasis on diverse V2X vehicle positions and RF power blockage situations. A detailed description of these datasets is provided in the following subsections.

A. DEEPSENSE 6G

DeepSense 6G is a large-scale, real-world multimodal dataset designed to advance machine learning research for 6G sensing and communication systems, and is publicly available at [33]. The dataset is tailored to perceive V2X environments, particularly within ISAC scenarios. To the best of current knowledge, it remains one of the most comprehensive resources available and is well suited for dynamic blockage analysis.

Scenario 32 is initially utilized for both training and testing. This street-level V2I scenario was collected in an outdoor wireless environment on a two-way city street along College Avenue in Tempe, Arizona, USA. The setup captures a diverse range of sensor blockages, with vehicles of varying sizes traveling in both directions at speeds up to 25 mph (40.6 km/h) across a street width of 13 meters. For our work, a time series of point cloud data from a fixed exteroceptive 3D LiDAR sensor is considered to capture the physical surroundings, enabling tracking of multiple vehicles for learning the temporal data-driven IC-GDBN model. The LiDAR sensor, mounted on a vehicle acting as a BS, has a sensing range of 100 meters and operates at a maximum spinning frequency of 20 Hz. In total, 3235 LiDAR point cloud samples were collected during this daytime scenario, providing sensor-level blockage information. In addition, the GPS position of the receiver is incorporated to obtain dynamic power measurements from detected vehicles under different channel fading conditions.

In addition, scenarios 33 and 34 from the DeepSense 6G dataset are considered for the model analysis, as these scenarios are well suited to the blockage detection task. These nighttime scenarios comprise approximately 8420 LiDAR point cloud sequences, enabling a detailed evaluation of LiDAR-based perception performance for detecting blockages and anomalies under varying illumination and environmental conditions. This allows a comprehensive assessment of the proposed blockage detection framework across diverse urban environments during both daytime and nighttime situations.

B. CAMPUS EXPERIMENTAL BLOCKAGE DATASET

For the validation purpose, an experimental setup is performed to collect the vehicular blockages dataset at Sabatini road, located near UC3M Campus, Leganes, Spain. The location represents a typical urban environment, featuring a two-way street configuration with two lanes. The road has a width of almost 7 meters. The street operates under a speed limit of 50 km/h and experiences regular traffic flow from vehicles of various sizes traveling at different speeds in both directions. The environment introduces diverse blockage situations that are characteristic of real-world LOS communication challenges for vehicles. In the experimental setup, two units were created: Unit 1 serves as the transmitting (Tx) unit, and Unit 2 serves as the receiving (Rx) unit. For simplicity, blockages are observed from fixed locations for both units as shown in Figure 3. Unit 2 is equipped with a Velodyne VLP-16 LiDAR sensor and an Rx XBee 3 Zigbee 3.0 module, while Unit 1 is equipped with a Tx XBee 3 Zigbee 3.0 module. Both units are placed on opposite sides of the road to observe the blockages. The LiDAR sensor continuously perceives the environment and generates detailed 3D point clouds representing all components present in the surroundings, including vehicles, pedestrians, and other potential obstacles. Simultaneously, the Tx and Rx XBee modules maintain continuous wireless communication, which facilitates real-time data transmission and monitoring. To ensure better RF LOS communication between the modules and improved LiDAR perception, the Tx unit is positioned at a height of 90 cm and the Rx unit at a height of 110 cm. Various types of blockages such as motorbikes, cars, and buses, are observed within the LiDAR area of interest during the data collection process, which affect both the LiDAR sensor perception and the LOS communication between the transmitter and receiver. Approximately 13, 700 data samples are recorded to observe blockages in the dynamic environment. To follow the proposed methodology outlined in Section IV, the LiDAR point clouds are used for vehicle tracking and detection, while the RF power values received by the BS are incorporated in accordance with the same methodology. This ensures a fair and consistent approach aligned with the other datasets.

IV. PROPOSED METHOD FOR BLOCKAGE DETECTION

This section primarily consists of three components: dataset preprocessing, the offline learning phase, and the online testing phase. The first part details the preprocessing and tracking of multiple vehicles using the joint probabilistic data association filter (JPDAF) approach applied to 3D LiDAR point clouds, along with the processing of received RF power signals under Rician and Rayleigh channel fading models. The offline learning phase focuses on training the joint world model (global model) for 3D LiDAR perception and RF power signals received by the BS. The online testing phase employs the IC-MJPF to achieve the blockage and abnormality detection objectives for both signals. The overall

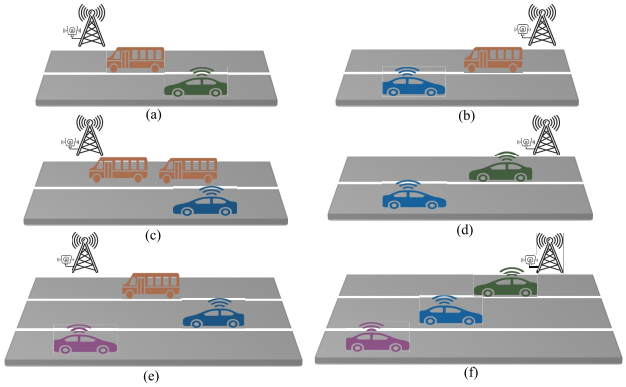


FIGURE 2. A schematic illustration of LiDAR and RF received power blockages in V2X networks: (a) large, continuous blockage; (b) large, short-duration blockage; (c) two large, continuous blockages; (d) small, short-duration blockage; (e) large blockage affecting two AVs; (f) short blockage for an AV in lane 2 and continuous blockage for an AV in lane 3.

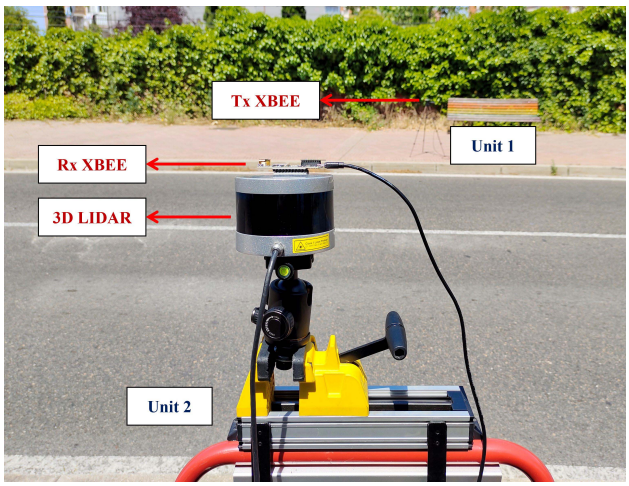


FIGURE 3. Campus experimental setup for blockage dataset collection.

pipeline is illustrated in Figure 4. The following subsections present the proposed methodology in detail.

A. PREPROCESSING 3D LIDAR DATA

To preprocess and extract the necessary features of dynamic objects from temporal sequences of LiDAR point cloud observations $\tilde{Z}^{l_n}_{t=1,2,...,T}$, where l_n denote the LiDAR observations of the n-th vehicle, t represents the time instant, and T indicates the total detection time of a vehicle, the raw data are first preprocessed using the LiDAR viewer toolbox in MATLAB R2025.a. The initial region of interest for scenario 32 of a Deepsense 6G dataset spans dimensions of -82 to 89 meters along the X-axis, -126 to 78 meters along the Y-axis, and -2 to 16 meters along the Z-axis in scenario 32 of the Deepsense 6G dataset. After applying filtering techniques such as ground removal, cropping, and neighbor denoising, the processed point clouds are reduced to dimensions of -28 to -4 meters along the X-axis, -13 to 35 meters along the Y-axis, and -1 to 1 meters along the Z-axis, with approximately 1000 points extracted from

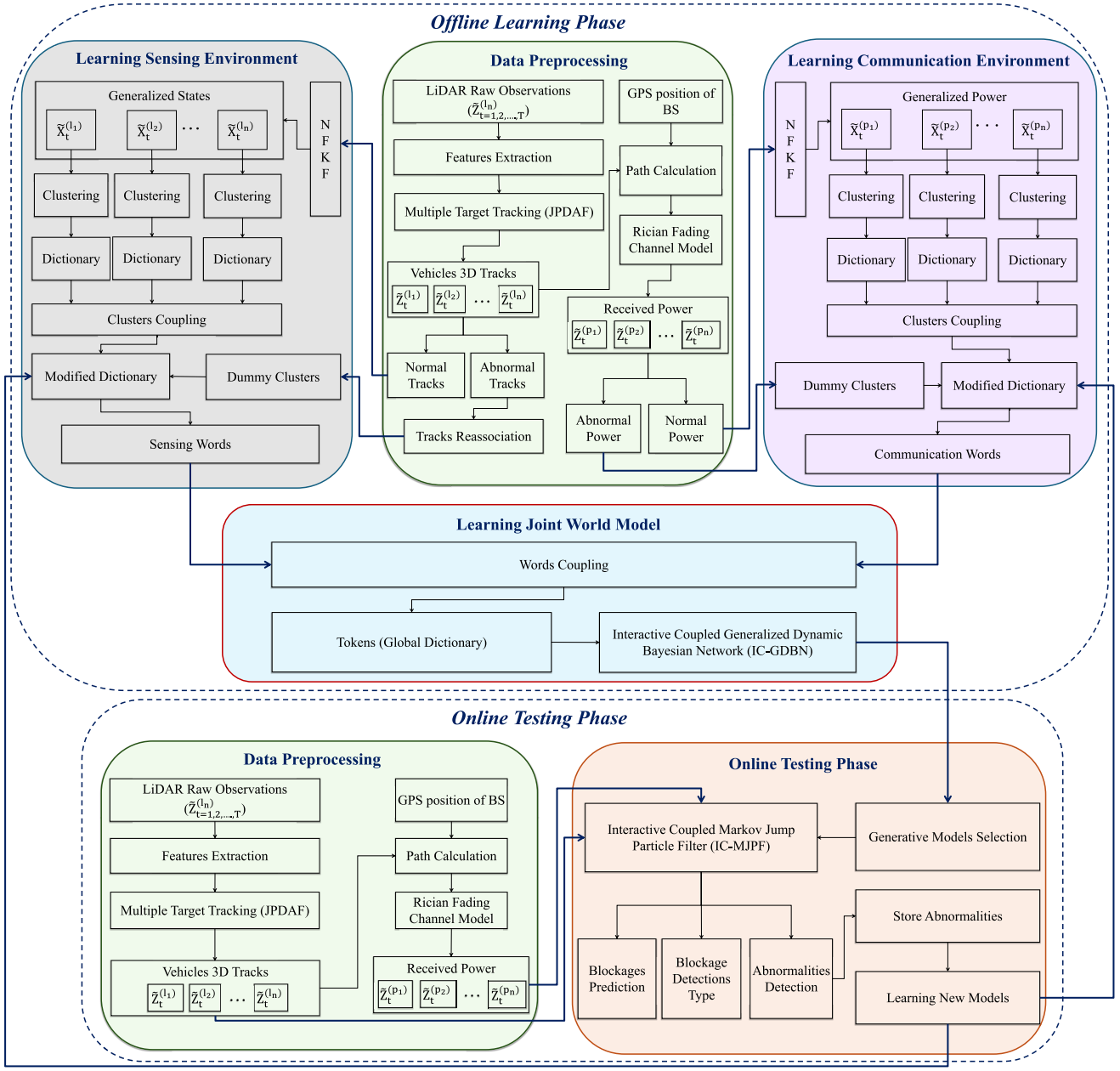


FIGURE 4. Proposed scheme for blockage detection in an ISAC environment by leveraging LiDAR and RF received power signals.

each point cloud in this scenario. These point clouds are still considered uncertain observations for the next processing step, as we are interested in tracking the vehicles and their corresponding positions that appear at different time instants within the environment. Furthermore, a similar strategy is applied to the other dataset scenarios to obtain meaningful point clouds.

To address the point cloud association problem, a JPDAF approach is utilized, ensuring accurate detection and tracking of multiple vehicles within the LiDAR field of view, since it better addresses the uncertainty issue, uses probabilistic weights based on prior estimates of observations, provides enhanced tracking accuracy, as well as can simultaneously detect and track various dynamic objects in the environment [34], [35].

In addition to continuous tracks, some discontinuous (broken) tracks caused by occlusions are also obtained by the JPDAF, which requires further processing. An approach is needed to estimate the missing positions of occluded vehicle tracks. For this purpose, an algorithm based on a federated learning framework for object re-identification has been proposed in [36]. In our work, we address this problem differently by introducing a re-association and re-identification approach, as shown in Algorithm 1, to handle vehicles with missing positions and associate them with the corresponding reappeared track. This algorithm records

Algorithm 1 Track Re-Association and Re-Identification

Input: Vehicle tracks (IDs, 3D positions, times) ← JPDAF output
Params: d_{th}, l_w, t_r, l_b
SameLane Function (x_1, x_2): return $|x_1 - x_2| \le l_w$ and $x_1, x_2 \in [l_{b,x}^{\min}, l_{b,x}^{\max}]$
for each detected track $n = 1, 2, \ldots, N$ **do**
  for each disappeared track $d = 1, 2, \ldots, D$ (not finished) **do**
    if d position $\notin l_b$ **then**
      Mark d as finished
    else
      $\Delta t \leftarrow t_n^{\text{start}} - t_d^{\text{last}}$; **continue** if $\Delta t \le 0$ or $\Delta t > t_r$
      Estimate velocity $\mathbf{v} \leftarrow (\mathbf{p}_d^{\text{last}} - \mathbf{p}_d^{\text{prev}})/(t_d^{\text{last}} - t_d^{\text{prev}})$
      Predict $\hat{\mathbf{p}}_d \leftarrow \mathbf{p}_d^{\text{last}} + \mathbf{v}\,\Delta t$
      Compute $dist \leftarrow \|\hat{\mathbf{p}}_d - \mathbf{p}_n^{\text{start}}\|$
      if $dist < d_{\text{th}}$ **and** SameLane$(p_{d,x}, p_{n,x})$ **then**
        Link track n with d; mark d as finished
        Update track ID for n; store predicted positions
      end if
    end if
  end for
  if n not linked **then**
    Add n to disappeared list (not finished)
  end if
end for
Output: Complete 3D tracks, missing positions, disappearance time

the disappearance and reappearance times of the blocked tracks, calculates velocity based on the last known positions and the time elapsed since disappearance, predicts missing positions at each time step, and computes the matching distance between the predicted and detected positions when the vehicle reappears in the scene. If the matching distance falls within a predefined threshold and the vehicle remains in the same lane, the track will be re-associated with the previously disappeared one. The algorithm yields complete 3D tracks, including missing positions, vehicle tracking IDs, and the time at which vehicles disappear. It is important to clarify that not all models are trained using data generated by this algorithm. Normal models are trained on complete 3D trajectories derived from real-time LiDAR sensor data. Algorithm 1 is specifically designed to identify and reconstruct missing positions, which are indicative of blockage scenarios. This information will later be used to train normal blockage models.

B. RF ENVIRONMENT REPRESENTATION

A detected AV transmits RF power to the BS through a LOS multipath fading model, i.e., Rician fading channel. Specifically, the model provides a realistic adjustment for modeling signal reception in environments dominated by a strong LOS component. Therefore, the received power in decibels (dB) can be expressed as:

$$P_{\text{rx}}(dB) = P_{\text{tx}} + G_{\text{tx}} + G_{\text{rx}} - \text{FSPL}(dB) + X_{\text{Rician}} \quad (1)$$

where, P_{tx} denotes the RF power in dBm transmitted by the detected vehicle, while G_{tx} and G_{rx} represent the gains of the transmitting and receiving antennas respectively, in dBi. FSPL and X_{Rician} refer to the free-space path loss and Rician fading components, respectively. The FSPL radio propagation model estimates the signal power loss that occurs during transmission due to distance and frequency over a LOS path in free space, defined in [37], and is calculated as:

$$\text{FSPL}(dB) = 20\log_{10}(d_{\text{km}}) + 20\log_{10}(f_{\text{MHz}}) + 32.44 \quad (2)$$

where d_{km} represents the distance between the BS and the transmitting AV in kilometers, f_{MHz} denotes the frequency in megahertz, and 32.44 is a constant that represents the path loss at a distance of 1km and a frequency of 1MHz. The fixed position of the BS is obtained from the available GPS dataset [33]. For a transmitting AV positions (x_i, y_i, z_i), the Euclidean distance (d_{km}) is calculated from the reference BS as:

$$d_{\text{km}} = \frac{1}{1000}\sqrt{(x_i - x_{\text{ref}})^2 + (y_i - y_{\text{ref}})^2 + (z_i - z_{\text{ref}})^2} \quad (3)$$

The Rician fading channel models fluctuations in the received RF signal strength due to scattering, diffraction, and reflections in environments with a strong LOS component. The probability density function followed by the received RF power signal (r) is obtained as:

$$p(r) = \frac{r}{\sigma^2}\exp\left(-\frac{r^2 + s^2}{2\sigma^2}\right) I_0\left(\frac{rs}{\sigma^2}\right), \quad (4)$$

$$r = |s + \sigma \cdot (N_R + jN_I)|, \quad (5)$$

In equations (4) and (5), r represents the amplitude of the received signal. The term $s = \frac{J}{J+1}$ represents the magnitude of the LOS path determined by the Rician J-factor. The standard deviation is given by $\sigma = \frac{1}{2}(J+1)$ and the variance of the Gaussian noise components is σ^2. The function $I_0(\cdot)$ denotes the modified Bessel function of the first kind and order zero. The variables N_R and N_I are independent, zero-mean Gaussian random variables with unit variance, representing the in-phase and quadrature components of the scattered signal, respectively.

C. LEARNING GLOBAL MODEL (OFFLINE PERCEPTION)

A basic generalized dynamic Bayesian network (GDBN) (see Figure 5(a)) offers a powerful framework for explaining and reasoning about the temporal evolution of complex systems. This three-level GDBN represents a probabilistic graphical model (PGM) comprising various nodes that represent random variables, with directed links indicating conditional dependencies between nodes that evolve over time. The arrows with orange shadow represent the inter-slice

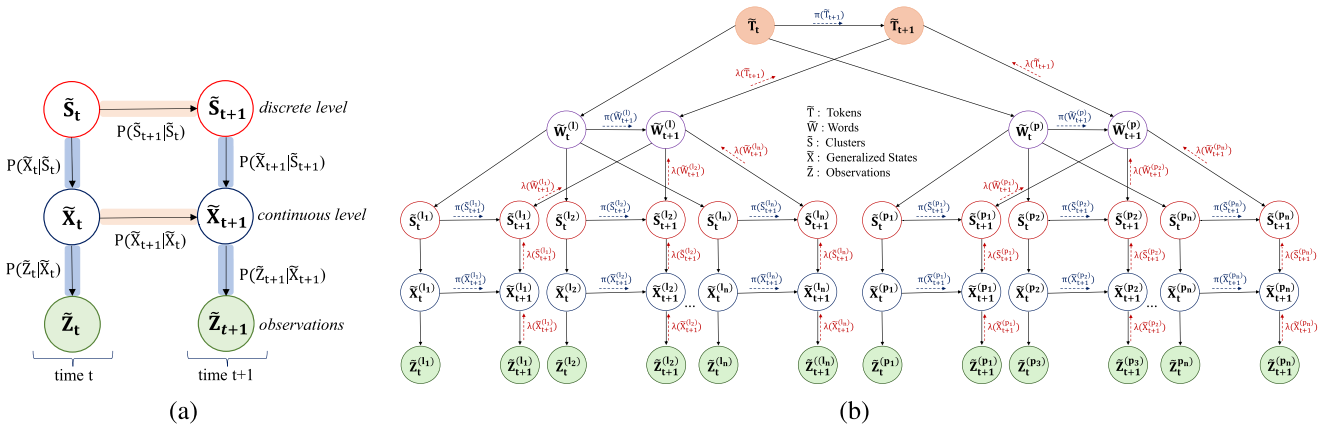


FIGURE 5. Modeling probabilistic dynamics in the ISAC framework: (a) The GDBN. (b) Illustration of the proposed IC-GDBN, composed of two coupled GDBNs that represent the sensing and communication signals received by the BS.

links, and the arrows with blue shadow represent the intra-slice links. The low-level node ($\tilde{Z}_t$) defines the observations obtained from the sensor, and the two above-level nodes ($\tilde{X}_t$) and ($\tilde{S}_t$) represent the continuous and discrete latent states, respectively. Due to their hierarchical structure, which encodes semantic relationships between low-level and high-level random variables, such causal PGMs have been widely used in the literature [6], [32].

1) INTERACTIVE COUPLED GDBN (IC-GDBN) MODEL

Our proposed IC-GDBN model for the blockage detection problem, encoding the temporal evolution and spatial relationships of both the LiDAR and RF received power signals, is illustrated in Figure 5(b). This integrated dynamic model incorporates two additional hierarchical levels of intelligence on top of the GDBN. This special case of hierarchical modelling represents semantic switching variables to encode the underlying reasons and better explain the predictions. The switching variables at the hierarchical structure encode the reasons behind blockage events. One branch represents the interaction between vehicles in terms of position information from the LiDAR sensor. This branch captures spatial interactions between vehicles, since blockage events occur when one vehicle creates a blockage for another. The other branch represents the interaction between power values, where semantic switching variables help to explain whether the RF power strength is reduced due to a blockage situation or remains normal. Therefore, the interaction-level variables distinguish whether blockage exists and, if so, what type of blockage has occurred. The additional hierarchical level at the top enables the BS to further explain whether the blockage event occurred due to power deficiency or LiDAR sensor detection, that is, whether the blockage results from physical object shadowing, absence of power signals, or the blockage has appeared in both the data modalities. This provides the semantic meaning needed for assessing blockage situations between physical objects and RF power. The model explicitly encodes conditional dependencies among random variables at multiple levels, enabling BS to understand cause–and–effect relationships. This, in turn, allows the BS to explain the predictions, identify errors, and adapt to dynamic environmental changes. Each discrete random variable (letter or cluster) represents a specific 3D position of the detected vehicle or the RF power strength signal received by the vehicle. A set of letters forms a word, which is also a discrete entity, representing an interaction among multiple vehicles, or an interaction among the RF power signals received from them. Therefore, the word-level hierarchy above the GDBN captures the interdependencies among multiple GDBN models. The dynamic interaction among multiple vehicles is learned through the LiDAR point cloud data in our previous work [38], [39]. A global dictionary, formed by a set of words (referred to as tokens are discrete symbols), represents the top abstraction level of IC-GDBN, aiming to achieve a joint objective by mapping tokens over time. Hence, the topmost hierarchy, built on interactive GDBNs, models the coupling between sensing and communication signal modalities through tokens that evolve temporally. We refer to this hierarchical global dynamic probabilistic model as the IC-GDBN.

In the graphical model, edges represent the dependency of discrete states on continuous generalized states (GSs). Through continuous neural message propagation within the graph, the BS can approximate the posterior distribution over latent variables. This propagation involves ascending signals originating from future states (brown arrows) and descending signals from past states (blue arrows). Thus, both future and past environmental states are continuously captured over time, enabling optimal online approximation of the posterior distribution. During the first-stage learning, a null-force Kalman filter (NFKF) is applied to the 3D positions, producing a 6D vector of GSs that contains the filtered positions and velocities of each tracked vehicle. The filter operates under the assumption that vehicles are unaffected by external forces [6]. Similarly, for RF received power signals, an NFKF yields a 2D vector of generalized powers (GPs) through the power derivative, enabling the agent to perceive

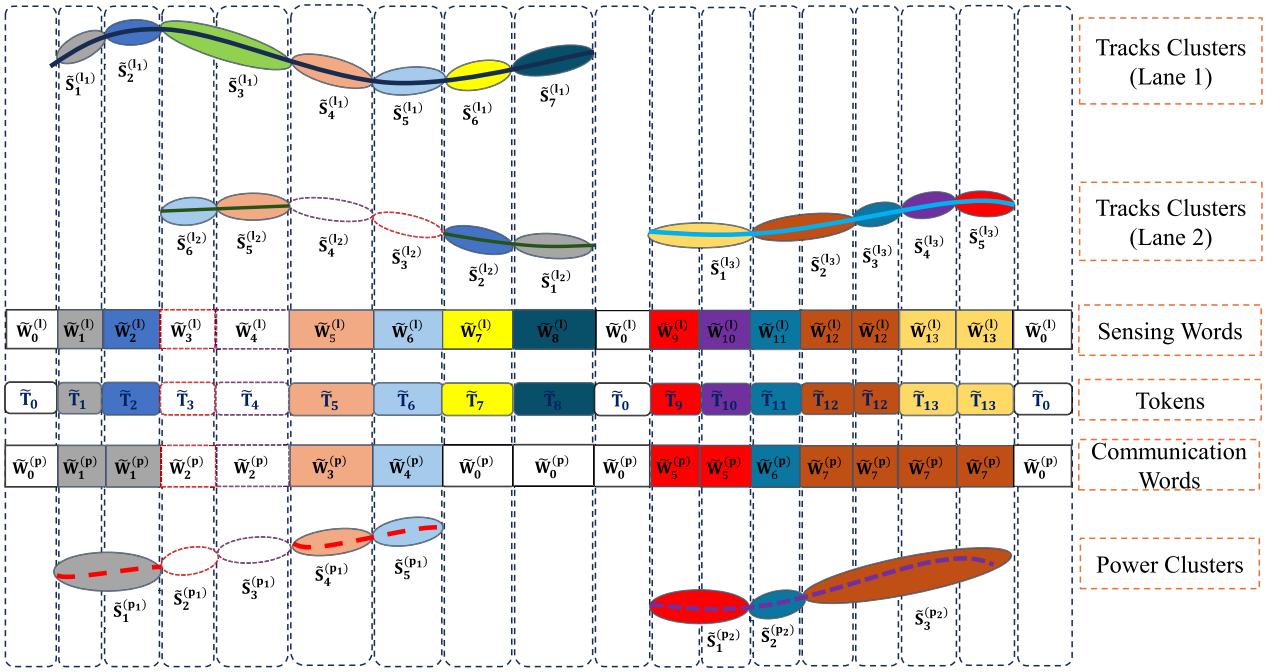


FIGURE 6. An example of semantic symbol creation at the word-level and token-level hierarchies in the proposed IC-GDBN, simultaneously encapsulating vehicles' 3D positions and their received power strengths.

how dynamic power varies over time. Here, we initiate the training phase using continuously detected vehicles (for normal tracks without blockages) and the normal RF power signals received at the BS.

The growing neural gas (GNG) algorithm has proven to be a powerful bio-inspired clustering method capable of learning complex data patterns and identifying similarities within intricate structures [40], [41]. As an unsupervised learning method, it employs a self-organizing neural network that incrementally and adaptively learns the structure of streaming data without requiring predefined knowledge of the cluster count [42]. Each node represents a neuron that maps the input data, and as learning progresses, new units are introduced to form a network that reflects the underlying topology of the data. In this work, we adopt GNG for its ability to support adaptive growth, operate effectively in dynamic 3D environments, and accommodate new observations without the need for model retraining. To learn a separate LiDAR sensing model for each detected vehicle and a separate communication model for RF received power, the GSs and GPs are separately given as an input to the GNG, which yields a different number of clusters as an output for each detected track, as well as for the power received by that vehicle.

The interactive hierarchy level above the GDBN models encodes interactions among multiple vehicles through a set of discrete nodes, referred to as sensing words, as illustrated in Figure 6. Similarly, interactions derived from the RF power received from multiple transmitting vehicles are represented as communication words. Finally, tokens that capture the relationship between sensing and communication words are encoded in the form of token symbols within the top hierarchy of the IC-GDBN.

2) GLOBAL DICTIONARY DEVELOPMENT

The global dictionary comprises the knowledge acquired at all hierarchical levels of the proposed IC-GDBN. It also includes the initially learned GDBN models, which serve as the basic vocabularies for both vehicle positions and their corresponding RF received power (see Figure 4). Here, we describe how the global dictionary is constructed, with the aim of detecting LiDAR and RF power blockages.

To explore the sensing environment, let $\mathbf{Z}^{(l_n)} = \{\tilde{Z}_t^{(l_1)}, \tilde{Z}_t^{(l_2)}, \ldots, \tilde{Z}_t^{(l_n)}\}$ denote the sequence of 3D positions of a detected vehicles trajectories at time t for l_n LiDAR tracks, where $n \in \{1, 2, \ldots, N\}$ and N is the total number of tracks obtained through the JPDAF filter, as discussed in the preprocessing step. A set of latent continuous GSs, consisting of the positions and velocities of each detected track (output of NFKF), is represented by $\mathbf{X}^{(l_n)} = \{\tilde{X}_t^{(l_1)}, \tilde{X}_t^{(l_2)}, \ldots, \tilde{X}_t^{(l_n)}\}$ and shown in Figure 5(b). Here, we incorporate the first-order derivative as part of the generalized variables to understand the dynamics of the environment as time evolves with limited computational effort, (i.e., $\tilde{X}_t^{(l_n)} = [\tilde{X}_t^{(l_n)}, \dot{\tilde{X}}_t^{(l_n)}]$, $\tilde{X}_t^{(l_n)} \in \mathbf{X}^{(l_n)}$). At the discrete level, let $\mathbf{S}^{(l_n)} = \{\tilde{S}_1^{(l_1)}, \tilde{S}_2^{(l_2)}, \ldots, \tilde{S}_M^{(l_n)}\}$ represent the set of discrete nodes (clusters, neurons, letters or superstates of a GDBN) corresponding to the dynamic 3D positions where M denotes the total number of discrete superstates in a given trajectory of the vehicle. These superstates are obtained using the unsupervised GNG algorithm when applied to the input GSs, $\mathbf{X}^{(l_n)}$, of each vehicle separately. where, $\tilde{S}_m^{(l_n)}$ signifies the

m-th cluster in $\mathbf{S}^{(l_n)}$, modeled as a Gaussian distribution: $\tilde{S}_1^{(l_1)} \sim \mathcal{N}(\mu_{\tilde{S}_m^{(l_n)}}, \Sigma_{\tilde{S}_m^{(l_n)}})$. To this end, the world model learns a basic GDBN model of vehicle trajectories detected from the LiDAR sensing environment.

Similarly, to explore the communication environment, let $\mathbf{Z}^{(p_n)} = \{\tilde{Z}_t^{(p_1)}, \tilde{Z}_t^{(p_2)}, \ldots, \tilde{Z}_t^{(p_n)}\}$ denote the series of RF power received by the BS at time t from transmitting vehicles, where p_n denotes the power received from n-th and $n \in \{1, 2, \ldots, N\}$. A set of hidden continuous GPs, consisting of the received RF power obtained from equation (1) and their rate of change, represented as $\mathbf{X}^{(p_n)} = \{\tilde{X}_t^{(p_1)}, \tilde{X}_t^{(p_2)}, \ldots, \tilde{X}_t^{(p_n)}\}$. Here, we again consider the first-order derivatives of power signals (i.e., $\tilde{X}_t^{(p_n)} = [\tilde{X}_t^{(p_n)}, \dot{\tilde{X}}_t^{(p_n)}]$, $\tilde{X}_t^{(p_n)} \in \mathbf{X}^{(p_n)}$). At the discrete level of a communication GDBN model, let $\mathbf{S}^{(p_n)} = \{\tilde{S}_1^{(p_1)}, \tilde{S}_2^{(p_2)}, \ldots, \tilde{S}_{M'}^{(p_n)}\}$ denote the set of power clusters corresponding to the received RF power values, where M' denotes the total number of power clusters generated by the GNG algorithm when the input GPs, $\mathbf{X}^{(p_n)}$, of each transmitting vehicle are applied to learn the dynamics of power from each transimitting vehicle separately. where, $\tilde{S}_{m'}^{(p_n)}$ denotes the m'-th cluster in $\mathbf{S}^{(p_n)}$, modeled as a Gaussian distribution: $\tilde{S}_1^{(p_1)} \sim \mathcal{N}(\mu_{\tilde{S}_{m'}^{(p_n)}}, \Sigma_{\tilde{S}_{m'}^{(p_n)}})$.

This clustering process is repeated for N continuously detected vehicle tracks from the JPDAF approach, as well as the RF power received from them. Consequently, a pool of GDBN models is learned, and two global sets, i.e., $\mathbf{S}^{(l)} = \{S^{(l_1)}, S^{(l_2)}, \ldots, S^{(l_N)}\}$ and $\mathbf{S}^{(p)} = \{S^{(p_1)}, S^{(p_2)}, \ldots, S^{(p_N)}\}$ are obtained for sensing and communication GDBN models, respectively. In addition to the number of clusters, the vocabulary of each GDBN model includes a set of temporal transition matrices (TTMs) for each time instant, an overall transition matrix (TM) capturing the probabilities of switching from one neuron to another, the maximum time duration spent in each neuron, as well as the neuron centroids and covariance matrices. Unlike our previous work, which considered the dictionary of only a single GDBN model [6], the proposed framework generalizes this formulation to multiple GDBN models.

The interactions among the N detected vehicles moving on the road lanes within the ISAC environment, as well as the RF power received from these vehicles, are represented through sensing and communication words, respectively. In the proposed model, these interactions are defined by structuring coupled hierarchies placed above the sensing and communication GDBNs, referred to as the sensing C-GDBN and communication C-GDBN, respectively, as illustrated in Figure 5(b).

A sensing word in the sensing environment refers to a group of clusters that encapsulates the relationships among discrete random variables of N tracked vehicles evolving over time t, and can be mathematically defined as $\mathbf{W}_t^{(l)} = [\tilde{S}_m^{(l_1)}, \tilde{S}_m^{(l_2)}, \ldots, \tilde{S}_m^{(l_N)}]$. Similarly, a communication word in the communication environment refers to a group of clusters that captures the relationships among the discrete superstates of RF power received from the N tracked vehicles evolving over time t, and can be mathematically defined as $\mathbf{W}_t^{(p)} = [\tilde{S}_m^{(p_1)}, \tilde{S}_m^{(p_2)}, \ldots, \tilde{S}_m^{(p_N)}]$.

Let $\mathbf{W}^{(l)} = \{\tilde{W}_1^{(l)}, \tilde{W}_2^{(l)}, \ldots, \tilde{W}_L^{(l)}\}$ denote the group of sensing words defining the 3D positions of detected vehicles from the LiDAR point clouds, where L is the total number of sensing words. In the same way, let $\mathbf{W}^{(p)} = \{\tilde{W}_1^{(p)}, \tilde{W}_2^{(p)}, \ldots, \tilde{W}_{L'}^{(p)}\}$ denote the group of communication words representing the RF power received by the agent from the detected vehicles, where L' symbolizes the total number of communication words.

With the help of these sensing and communication semantic representations, the BS can capture the interactions among moving vehicles across multiple road lanes, as well as the relationships among the RF power values received by these vehicles at each time instant, as illustrated in Figure 6. The agent obtains position and power information from vehicles traveling on a two-lane roadway, consistent with the real-world dataset used in this study. The sensing words are combined to describe a specific physical sensing configuration, which is subsequently associated with the received RF power signals to represent the corresponding communication configuration. This association establishes the ISAC paradigm for V2X systems.

In general, the dynamic rules for generating the words are as follows: a word with value zero indicates the absence of vehicles in the environment, while identical words are assigned when the same sequence of neurons is repeated. Otherwise, words are assigned in ascending order over the evolution of time. After obtaining the sensing words $\mathbf{W}^{(l)}$, we analyzed how vehicles positions are changing within the set of neurons to learn the interactive transition matrix (ITM) for LiDAR sensor by estimating the transition probabilities $\pi_{ij} = P(\tilde{W}_{t+1}^{(l)} = i \mid \tilde{W}_t^{(l)} = j)$, where $i, j \in \mathbf{W}^{(l)}$ over a passage of learning time. Hence, the word-level ITM, $\prod \mathbf{W}^{(l)}$, for the sensing environment can be defined as,

$$\Pi \mathbf{W}^{(l)} = \begin{bmatrix} \mathrm{P}\left(\tilde{\mathrm{W}}_1^{(l)} \mid \tilde{\mathrm{W}}_1^{(l)}\right) & \mathrm{P}\left(\tilde{\mathrm{W}}_2^{(l)} \mid \tilde{\mathrm{W}}_1^{(l)}\right) & \cdots & \mathrm{P}\left(\tilde{\mathrm{W}}_L^{(l)} \mid \tilde{\mathrm{W}}_1^{(l)}\right) \\ \mathrm{P}\left(\tilde{\mathrm{W}}_1^{(l)} \mid \tilde{\mathrm{W}}_2^{(l)}\right) & \mathrm{P}\left(\tilde{\mathrm{W}}_2^{(l)} \mid \tilde{\mathrm{W}}_2^{(l)}\right) & \cdots & \mathrm{P}\left(\tilde{\mathrm{W}}_L^{(l)} \mid \tilde{\mathrm{W}}_2^{(l)}\right) \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{P}\left(\tilde{\mathrm{W}}_1^{(l)} \mid \tilde{\mathrm{W}}_L^{(l)}\right) & \mathrm{P}\left(\tilde{\mathrm{W}}_2^{(l)} \mid \tilde{\mathrm{W}}_L^{(l)}\right) & \cdots & \mathrm{P}\left(\tilde{\mathrm{W}}_L^{(l)} \mid \tilde{\mathrm{W}}_L^{(l)}\right) \end{bmatrix} \quad (6)$$

where $0 \leq \mathrm{P}(\tilde{\mathrm{W}}_{l+1}^{(l)} \mid \tilde{\mathrm{W}}_l^{(l)}) \leq 1$, and $\sum_{l=1}^{L} \mathrm{P}(\tilde{\mathrm{W}}_{l+1}^{(l)} \mid \tilde{\mathrm{W}}_l^{(p)}) = 1$, $\forall l \in \{1, \ldots, L\}$.

Similarly, the communication words explaining how RF power patterns received from multiple vehicles are varying within the set of clusters, enabling the learning of an ITM for power signals by estimating the transition probabilities $\pi_{ij} = P(\tilde{W}_{t+1}^{(p)} = i \mid \tilde{W}_t^{(p)} = j)$, where $i, j \in \mathbf{W}^{(p)}$ as learning time progresses. Consequently, the word-level ITM, $\prod \mathbf{W}^{(p)}$,

for the communication environment can be defined as,

$$\Pi \mathbf{W}^{(p)} = \begin{bmatrix} \mathrm{P}\left(\tilde{\mathrm{W}}_1^{(p)} \mid \tilde{\mathrm{W}}_1^{(p)}\right) & \mathrm{P}\left(\tilde{\mathrm{W}}_2^{(p)} \mid \tilde{\mathrm{W}}_1^{(p)}\right) & \cdots & \mathrm{P}\left(\tilde{\mathrm{W}}_{L'}^{(p)} \mid \tilde{\mathrm{W}}_1^{(p)}\right) \\ \mathrm{P}\left(\tilde{\mathrm{W}}_1^{(p)} \mid \tilde{\mathrm{W}}_2^{(p)}\right) & \mathrm{P}\left(\tilde{\mathrm{W}}_2^{(p)} \mid \tilde{\mathrm{W}}_2^{(p)}\right) & \cdots & \mathrm{P}\left(\tilde{\mathrm{W}}_{L'}^{(p)} \mid \tilde{\mathrm{W}}_2^{(p)}\right) \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{P}\left(\tilde{\mathrm{W}}_1^{(p)} \mid \tilde{\mathrm{W}}_{L'}^{(p)}\right) & \mathrm{P}\left(\tilde{\mathrm{W}}_2^{(p)} \mid \tilde{\mathrm{W}}_{L'}^{(p)}\right) & \cdots & \mathrm{P}\left(\tilde{\mathrm{W}}_{L'}^{(p)} \mid \tilde{\mathrm{W}}_{L'}^{(p)}\right) \end{bmatrix} \tag{7}$$

where $0 \leq \mathrm{P}(\tilde{\mathrm{W}}_{l'+1}^{(p)} \mid \tilde{\mathrm{W}}_{l'}^{(p)}) \leq 1$, and $\sum_{l'=1}^{L'} \mathrm{P}(\tilde{\mathrm{W}}_{l'+1}^{(p)} \mid \tilde{\mathrm{W}}_{l'}^{(p)}) = 1, \quad \forall l' \in \{1, \ldots, L'\}$.

These time-varying ITMs capture how the transition probabilities of moving from one word to another change over time; some probabilities rise while others drop as time progresses, allowing the BS to track dynamic changes in the real environment.

To jointly represent semantic features related to vehicle positions and RF received power, we construct the top hierarchy of the proposed IC-GDBN, referring to it as the token-level hierarchy, as illustrated in Figure 5(b). Let $\mathbf{T} = \{\tilde{T}_1, \tilde{T}_2, \ldots, \tilde{T}_K\}$ denote a set of composite token symbols, where each token T_k encapsulates a pair of words, i.e., a pair of sensing word and a communication word, at a specific time step, where $k \in \{1, 2, \ldots, K\}$ and K represents the total number tokens. Formally, we define it as $\tilde{T}_k = [W_l^{(l)}, W_{l'}^{(p)}]$, where, $W_l^{(l)}$ refers to the l-th sensing word, while $W_p^{(l')}$ corresponds to l'-th communication word.

This formulation ensures that each token simultaneously captures both data modalities over time evolution. The dynamic rules for obtaining these tokens are the same as those used for the creation of words. Furthermore, to model how these tokens evolve within the word configurations as training time passes, we introduce an interactive coupled transition matrix (ICTM) by estimating the tokens transition probabilities $\pi_{ij} = P([W_{t+1}^{(l)}, W_{t+1}^{(p)}] = i \,|\, [W_t^{(l)}, W_t^{(p)}] = j)$, where $i, j \in \mathbf{T}$, which governs the progression and transformation of tokens within the system and hence the token-level ICTM, $\prod \mathbf{T}$, can be formulated as,

$$\Pi \mathbf{T} = \begin{bmatrix} \mathrm{P}\left(\tilde{\mathrm{T}}_1 \mid \tilde{\mathrm{T}}_1\right) & \mathrm{P}\left(\tilde{\mathrm{T}}_2 \mid \tilde{\mathrm{T}}_1\right) & \cdots & \mathrm{P}\left(\tilde{\mathrm{T}}_K \mid \tilde{\mathrm{T}}_1\right) \\ \mathrm{P}\left(\tilde{\mathrm{T}}_1 \mid \tilde{\mathrm{T}}_2\right) & \mathrm{P}\left(\tilde{\mathrm{T}}_2 \mid \tilde{\mathrm{T}}_2\right) & \cdots & \mathrm{P}\left(\tilde{\mathrm{T}}_K \mid \tilde{\mathrm{T}}_2\right) \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{P}\left(\tilde{\mathrm{T}}_1 \mid \tilde{\mathrm{T}}_K\right) & \mathrm{P}\left(\tilde{\mathrm{T}}_2 \mid \tilde{\mathrm{T}}_K\right) & \cdots & \mathrm{P}\left(\tilde{\mathrm{T}}_K \mid \tilde{\mathrm{T}}_K\right) \end{bmatrix} \tag{8}$$

where, $0 \leq \mathrm{P}(\tilde{\mathrm{T}}_{k+1} \mid \tilde{\mathrm{T}}_k) \leq 1$, and $\sum_{k=1}^{K} \mathrm{P}(\tilde{\mathrm{T}}_{k+1} \mid \tilde{\mathrm{T}}_k) = 1, \quad \forall k \in \{1, \ldots, K\}$.

Up to this point, the global dictionary of IC-GDBN is constructed from uninterrupted vehicle tracks and the corresponding RF power received by the BS. These continuous observations form the backbone of the learning process, capturing both spatial trajectories and RF received power signal characteristics under normal sensing and normal communication conditions. This joint vocabulary serves as the baseline for modeling dynamic behavior in environments free from blockages, occlusions, or disruptions, enabling the system to learn consistent patterns across both physical movement of the vehicles and RF power communication signals.

To address blockage scenarios, where direct LiDAR observations or RF power measurements are unavailable, we extend this framework by introducing dummy positions and dummy RF power values within dummy clusters. These surrogate entities allow the system to preserve structural coherence when direct observations are unavailable. When a track is broken due to occlusion, a track re-association and re-identification approach defined in Algorithm 1, is employed to estimate the missing positions for a disappeared track. This algorithm leverages temporal continuity, spatial proximity, and signal consistency to reconnect fragmented observations. The reconstructed data is then used to form the basic GDBN vocabularies for blockage situations, including mean, covariance, and transition probabilities in the TMs for both the sensing and communication models.

We refer to them as the normal blockage models, as the system learns to handle occlusion situations for both the vehicle positions and RF received power signals by the BS. To characterize different types of blockage, we introduce a set of configurable scenarios in which short-term or long-term occlusions may occur at various locations along the road lanes. Based on these configurations, we construct a pool of normal blockage models that span both physical (LiDAR-based) and communication (power feature-based) modalities. Each model captures a distinct blockage type, caused by a pedestrian, a car, or a bus and its impact on vehicle tracking and RF power reception. It should be noted that we trained separate normal models for each detected AV and for the RF power received from that AV. In addition, different configurations of normal blockage models were considered to represent possible blockage scenarios, covering both short- and long-term blockages. At the interaction hierarchy, the system assigns distinct sensing and communication words for each blockage model, following the same dynamic rule strategy used before for the normal case. These words reflect the altered semantics of sensor and signal behavior under occlusion, enabling the model to maintain interpretability despite missing or distorted inputs. This same strategy is extended to the token hierarchy, where blockage conditions across both data modalities are encoded into high-level semantic tokens. These tokens represent the integrated state of the environment, capturing both sensing and communication blockages in a unified symbolic form. By preserving the dynamic structure and vocabulary rules across hierarchies, the IC-GDBN remains robust and adaptive under diverse blockage scenarios.

The generalized tokens representing the vehicle positions and their associated RF received power, thereby bridging the sensing and communication domains. Each token captures a latent environmental state, reflecting either normal (uninterrupted) or blockage (occluded) conditions. The enriched global dictionary encompasses multiple layers of abstraction in the IC-GDBN framework: at the top-level, K token symbols and the ICTM; at the interaction-level, L sensing

and L' communication words including both normal and blockage specific vocabularies along with their ITMs; at the discrete-level, the multi-model basic GDBN dictionaries for vehicle positions and RF received power, capturing both uninterrupted and occluded scenarios; at the continuous-level, the continuous GSs, the dummy GSs, continuous GPs and dummy GPs; and finally, the raw observations of both modalities at the observation-level. This hierarchical structure can be viewed as a sequential chain, progressing from high-level tokens down to low-level observations. The proposed structure supports a hierarchical modeling framework that characterizes the dynamic evolution of sensing and communication environments, incorporating Markov chains across temporal and hierarchical layers. By employing the chain rule of probability and exploiting the conditional independencies encoded within the hierarchies, including those introduced by blockages, the joint probability distribution governing the dictionary can be formally formulated as:

$$\begin{aligned} &P\left(\tilde{\mathrm{T}}_t, \tilde{W}_t^{(l)}, \tilde{W}_t^{(p)}, \tilde{S}_t^{(ln)}, \tilde{S}_t^{(pn)}, \tilde{X}_t^{(ln)}, \tilde{X}_t^{(pn)}, \tilde{Z}_t^{(ln)}, \tilde{Z}_t^{(pn)}\right) \\ &= \prod_{t=1}^{T} P\left(\tilde{\mathrm{T}}_{t+1} \mid \tilde{\mathrm{T}}_t\right) P\left(\tilde{W}_{t+1}^{(l)} \mid \tilde{W}_t^{(l)}, \tilde{\mathrm{T}}_{t+1}\right) \\ &\times P\left(\tilde{W}_{t+1}^{(p)} \mid \tilde{W}_t^{(p)}, \tilde{\mathrm{T}}_{t+1}\right) P\left(\tilde{S}_{t+1}^{(ln)} \mid \tilde{S}_t^{(ln)}, \tilde{W}_{t+1}^{(l)}\right) \\ &\times P\left(\tilde{S}_{t+1}^{(pn)} \mid \tilde{S}_t^{(pn)}, \tilde{W}_{t+1}^{(p)}\right) P\left(\tilde{X}_{t+1}^{(ln)} \mid \tilde{X}_t^{(ln)}, \tilde{S}_{t+1}^{(ln)}\right) \\ &\times P\left(\tilde{X}_{t+1}^{(pn)} \mid \tilde{X}_t^{(pn)}, \tilde{S}_{t+1}^{(pn)}\right) P\left(\tilde{Z}_{t+1}^{(ln)} \mid \tilde{X}_{t+1}^{(ln)}\right) \\ &\times P\left(\tilde{Z}_{t+1}^{(pn)} \mid \tilde{X}_{t+1}^{(pn)}\right) \end{aligned} \tag{9}$$

Equation (9) formalizes the joint distribution by decomposing it into conditional probabilities that reflect both temporal transitions and hierarchical dependencies. This structure encodes conditional independencies while simultaneously capturing the dynamic interactions between both data modalities. Furthermore, the IC-GDBN model adopts a hierarchical message-passing architecture inspired by our previous work [5], [43], in which information is propagated across layers to allow efficient inference and integration of heterogeneous signals. In this way, the BS learns the proposed IC-GDBN probabilistically and hierarchically from the dynamic rules encoded in the global dictionary created above to make inferences, predict, and detect various types of blockages for vehicle positions and RF power signals during the online testing.

D. ONLINE TESTING PHASE

The online testing procedure is illustrated in Figure 4. The pre-processing strategy follows the same steps applied during the training phase, as described in Section IV. The testing data undergo LiDAR point-cloud processing, extraction of vehicle dynamic features, generation of multiple vehicle trajectories using JPDAF, and computation of received power signals using equation (1). For the testing phase, IC-MJPF is used to perform inference and detect different types of blockages in both data modalities. It is important to note that, during the offline training phase, the resource-constrained BS hardware requires sufficient memory to store dictionaries of trained models for multiple configurations. These include normal and blockage models derived from LiDAR-based vehicle positions as well as models trained from RF received power measurements. In the online phase, the IC-MJPF uses this information by selecting the best model for each situation from these pre-trained models to make inferences. The inference process is presented in Algorithm 2, where the particle update mechanism employs a sequential importance resampling (SIR) strategy to mitigate particle degeneracy and preserve diversity within the particle set. Particle weights are computed using the abnormality measures $KLDA^{(l)}$ and $KLDA^{(p)}$, assigning higher weights to particles exhibiting lower abnormality. This weighting scheme prioritizes particles that more closely reflect normal system behavior, thereby improving predictive accuracy, while penalizing those associated with higher abnormality levels. At each time instant t, the BS performs N_p predictions across both interaction levels, where N_p denotes the total number of particles propagated by the filter. The computational complexity for the prediction process is $O(N_p) + O(N_p)$. For discrete-level abnormalities detection, the complexity of computing $KLDA^{(l)}$ is $O(L \times N_p)$, while that of $KLDA^{(p)}$ is $O(L' \times N_p)$. This computational load remains manageable due to the resampling step, which ensures that only the most representative particles (i.e., those with the highest weights) are propagated to subsequent time steps.

1) PREDICTION AND PERCEPTION

The proposed approach aims to predict and detect various types of LiDAR and RF received power blockages. An interactive MJPF combines the Particle Filter and the Kalman Filter to perform temporal and hierarchical inference [44]. In our work, the BS predicts future blockages through the proposed IC-MJPF, based on the knowledge learned by the IC-GDBN during the offline learning phase, enabling the BS to make more complex interactive predictions. Before making predictions, it is essential to select the most suitable generative models for both data types, based on how well they represent the LiDAR and RF power measurements; this process enhances blockage prediction performance. For this purpose, we use the Euclidean distance to identify the closest superstates.

$$D_{t,cm}^{(\ell)} = \sqrt{\left(x_t - x_{c_m}^{(\ell)}\right)^2 + \left(y_t - y_{c_m}^{(\ell)}\right)^2 + \left(z_t - z_{c_m}^{(\ell)}\right)^2} \tag{10}$$

$$D_{t,c_{m'}}^{(p)} = \left| p_t - p_{c_{m'}}^{(p)} \right| \tag{11}$$

Equation (10) quantifies the spatial distance between the observed 3D vehicle positions (x_t, y_t, z_t) at time t and the centroid of the m-th superstate within each LiDAR GDBN dictionary ℓ, represented by coordinates $(x_{c_m}^{(\ell)}, y_{c_m}^{(\ell)}, z_{c_m}^{(\ell)})$, where $\ell \in \{1, 2, \ldots, \mathcal{L}\}$ and $\mathcal{L}$ denotes the total number of LiDAR data based trained GDBN models. Correspondingly,

equation (11) expresses the absolute difference between the observed power measurement p_t and the centroid power value $p_{c_{m'}}^{(p)}$ of the m'-th cluster in each power GDBN dictionary p, where $p \in \{1, 2, \ldots, \mathcal{P}\}$ and $\mathcal{P}$ denotes the total number of RF received power data based trained GDBN models.

At each time instant t and for each generative model, the closest superstates with minimum distances (errors) for both LiDAR and RF power are expressed as,

$$e_t^{(\ell)} = \min_{m \in \{1,\ldots,M\}} D_{t,c_m}^{(\ell)}, \quad e_t^{(p)} = \min_{m' \in \{1,\ldots,M'\}} D_{t,c_{m'}}^{(p)} \tag{12}$$

Finally, the optimal LiDAR ℓ^* and RF received power p^* generative models at time t with minimum errors are calculated as,

$$\ell^* = \arg \min_{\ell \in \{1,\ldots,\mathcal{L}\}} e_t^{(\ell)}, \quad p^* = \arg \min_{p \in \{1,\ldots,\mathcal{P}\}} e_t^{(p)} \tag{13}$$

Hence, to minimize computational complexity, the optimal models ℓ^* and p^* are selected for both LiDAR and RF received power observations at each time t to achieve better prediction accuracy.

2) ABNORMALITY SIGNALS

To evaluate whether the BS correctly predicts LiDAR and RF received power blockages, abnormality signals are used to quantify the divergence between the predicted and the estimated semantic symbols, based on models selection strategies. Since the global architecture encompasses both sensing and communication domains, our focus here is on the abnormality signals at the word-level branches. The abnormality signals in each branch assess how much the predicted sensing and communication words deviate from the estimated words based on the dynamic rules encoded in the global dictionary. The discrete abnormality signals at the superstate level have already been calculated using the Kullback-Leibler divergence abnormality (KLDA) in [45]. In this work, we emphasize the abnormality signals at the world-level hierarchies of both domains, using KLDA, which is computed as follows:

$$KLDA^{(l)} = \sum_{l=1}^{L} \pi\left(W_t^l\right) \log\left(\frac{\pi\left(W_t^l\right)}{\lambda\left(W_t^l\right)}\right) \tag{14}$$

$$KLDA^{(p)} = \sum_{l'=1}^{L'} \pi\left(W_t^p\right) \log\left(\frac{\pi\left(W_t^p\right)}{\lambda\left(W_t^p\right)}\right) \tag{15}$$

Equations (14) and (15) compute the KLDA abnormalities for both vehicle positions and RF received power respectively, where $\pi(W_t^l)$ and $\pi(W_t^p)$ denote the predictive support (i.e., expected messages) signify the transition evolution of words $((W_t^l)$ and $((W_t^p)$ picked from the transition matrices learned in previous experiences for both LiDAR and power data respectively. $\lambda(W_t^l)$ and $\lambda(W_t^p)$ denote the diagnostic support (evidence messages). If the observation confirms the prediction, the BS can identify the normal situation within the environment, otherwise something abnormal is occurred.

Algorithm 2 IC-MJPF

Input:
$\tilde{W}^{(l)}, \Pi\mathbf{W}^{(l)}, \tilde{W}^{(p)}, \Pi\mathbf{W}^{(p)}, \tilde{S}^{(ln)}, (\mu_{\tilde{S}_m^{(ln)}}, \Sigma_{\tilde{S}_m^{(ln)}}) \in$
$\tilde{S}^{(ln)}, \tilde{S}^{(pn)}, (\mu_{\tilde{S}_{m'}^{(pn)}}, \Sigma_{\tilde{S}_{m'}^{(pn)}}) \in \tilde{S}^{(pn)}, \tilde{X}^{(ln)}, \tilde{X}^{(pn)} \leftarrow$ Learned global dictionary
$\tilde{Z}^{(ln)}, \tilde{Z}^{(pn)} \leftarrow$ Testing measurements
$N_p \leftarrow$ Number of Particles

for $t = 1$: T **do**
  for $n_p = 1$: N_p **do**
    Interaction level prediction:
    Initialize particle weight: $w_t^{(n_p)} \leftarrow \frac{1}{N_p}$
    if $t = 1$ **then**
      Predict random words: $\tilde{W}_t^{(l,n_p)}, \tilde{W}_t^{(p,n_p)}$
      Obtain estimated superstates from $P(\tilde{S}_t^{(l)} \mid \tilde{W}_t^{(l,n_p)})$ and $P(\tilde{S}_t^{(p)} \mid \tilde{W}_t^{(p,n_p)})$
      Estimate current states from $P(\tilde{X}_t^{(l)} \mid \tilde{S}_t^{(l)})$ and $P(\tilde{X}_t^{(p)} \mid \tilde{S}_t^{(p)})$
      $\tilde{X}_t^{(l)} \leftarrow \tilde{X}_1^{(l)}$
    else
      Predict $\tilde{W}_t^{(l,n_p)}$ from $\Pi\mathbf{W}^{(l)}$
      Predict $\tilde{W}_t^{(p,n_p)}$ from $\Pi\mathbf{W}^{(p)}$
      Get superstates and states as before
      $\tilde{X}_t^{(l)} \leftarrow \tilde{X}_{t-1}^{(l)}$
    end if
    $\tilde{Z}_t^{(l)}, \tilde{Z}_t^{(p)} \leftarrow$ Receive current observations
    Find the best models l^* and p^* using (13)
    Compute estimated words from $P(\tilde{W}_t^{(l)} \mid \tilde{S}_t^{(l^*)})$ and $P(\tilde{W}_t^{(p)} \mid \tilde{S}_t^{(p^*)})$
    Compute $KLDA^{(l)}$ using (14)
    Compute $KLDA^{(p)}$ using (15)
    Update particle weight:
    $$w_t^{(n_p)} = \frac{w_t^{(n_p)}}{KLDA^{(l)} + KLDA^{(p)}}$$
  end for
  SIR resampling
end for
Output: $KLDA^{(l)}$, $KLDA^{(p)}$

Thus, comparing $\pi(W_t^l)$ with $\lambda(W_t^l)$ and $\pi(W_t^p)$ with $\lambda(W_t^p)$ allows BS to detect abnormal behaviors.

The IC-MJPF uses the optimal generative models to make blockage inferences for LiDAR and RF received power data, as well as detect their type of dynamic blockages in the environment. It propagates a number of particles, each one representing a hypothesis regarding the current state of the system. This process is defined by Algorithm 2, where initially, uniform weights for each particle are assigned. To make an inference at the interaction level, the first word is randomly sampled from a probability distribution. For each particle, the next word is predicted based on the transition probabilities of its estimated word. The estimated word at the current time step is then used to update the particle weights.

Particles are resampled based on their updated weights, and particles with higher weights are more likely to be selected.

3) BLOCKAGE DETECTION PROBABILITY

In our evaluation pipeline, we also assess blockage prediction performance at the latent word-level hierarchies for both sensing and communication configurations by comparing the most frequently occurring blockage word (representing a sequence of activated neurons) from the winning particles, against the ground truth using accuracy metrics. To formalize this, let $\tilde{W}_t^{(l_{pb})} = [\tilde{S}_m^{(l_{pb1})}, \tilde{S}_m^{(l_{pb2})}, \ldots, \tilde{S}_m^{(l_{pbN})}]$ and $\tilde{W}_t^{(p_{pb})} = [\tilde{S}_{m'}^{(p_{pb1})}, \tilde{S}_{m'}^{(p_{pb2})}, \ldots, \tilde{S}_{m'}^{(p_{pbN})}]$ denote the predicted blockage words, representing vehicle(s) 3D positional data and RF received power measurements at time t, respectively. The corresponding estimated words obtained from low-level observations are denoted by $\tilde{W}_t^{(l_e)} = [\tilde{S}_m^{(l_{e1})}, \tilde{S}_m^{(l_{e2})}, \ldots, \tilde{S}_m^{(l_{eN})}]$ and $\tilde{W}_t^{(p_e)} = [\tilde{S}_{m'}^{(p_{e1})}, \tilde{S}_{m'}^{(p_{e2})}, \ldots, \tilde{S}_{m'}^{(p_{eN})}]$. Let $\mathbf{W}^{(l)} = \{\tilde{W}_t^{(l_{pb})}\}_{t=1}^{L}$ and $\mathbf{W}^{(p)} = \{\tilde{W}_t^{(p_{pb})}\}_{t=1}^{L'}$ represent the total predicted words associated with the sensing and communication modalities, respectively. Then the corresponding estimated words are denoted by $\mathbf{W}^{(l_e)} = \{\tilde{W}_t^{(l_e)}\}_{t=1}^{L}$ and $\mathbf{W}^{(p_e)} = \{\tilde{W}_t^{(p_e)}\}_{t=1}^{L'}$, respectively. The Hamming distances for both data modalities i.e., vehicle(s) 3D positions and the received RF power by the BS to quantify the mismatch between predicted and estimated neuron sequences can be calculated as $H^{(l)} = \sum_{t=1}^{L} \mathbb{I}(\tilde{W}_t^{(l_{pb})} \neq \tilde{W}_t^{(l_e)})$ and $H^{(p)} = \sum_{t=1}^{L'} \mathbb{I}(\tilde{W}_t^{(p_{pb})} \neq \tilde{W}_t^{(p_e)})$, where $\mathbb{I}(\cdot)$ denotes the indicator function that returns 1 if the condition is true and 0 otherwise. These distances are then used to compute the corresponding accuracies for each modality, given by $A^{(l)} = 1 - \frac{H^{(l)}}{L}$ and $A^{(p)} = 1 - \frac{H^{(p)}}{L'}$, which represent the proportion of correctly matched neurons. Consequently, the blockage prediction probabilities for both the sensing and communication modalities are approximated by their respective accuracies, expressed as $\mathcal{P}^{(l)} \approx A^{(l)}$ and $\mathcal{P}^{(p)} \approx A^{(p)}$, respectively.

V. SIMULATION RESULTS AND DISCUSSION

This section evaluates the performance of our proposed IC-GDBN model within the ISAC framework. First, we analyze preprocessing and training results obtained at different hierarchical levels, followed by the online testing results achieved by the IC-MJPF. The performance of the GNG algorithm depends on the choice of hyperparameters, which govern adaptation stability and cluster formation. The parameter α regulates the update weight for 3D positions and RF power measurements, while β controls the update weight for velocities and power derivatives. These parameters determine the balance between stability and flexibility in node movement, larger values promote smoother and more stable cluster formation, whereas smaller values allow more dynamic adjustments, potentially increasing the number of nodes. The winner node learning rate ϵ_b and the neighbor learning rate ϵ_n influence how strongly nodes move toward incoming data samples. Higher learning rates accelerate adaptation but may lead to overly merged clusters, whereas lower values preserve finer structural distinctions. Since the detected 3D trajectories of the vehicles are more smoother than the more rapidly varying distributed power values affected by channel fading, assigning higher weights to α helps maintain stable spatial clusters while allowing a finer

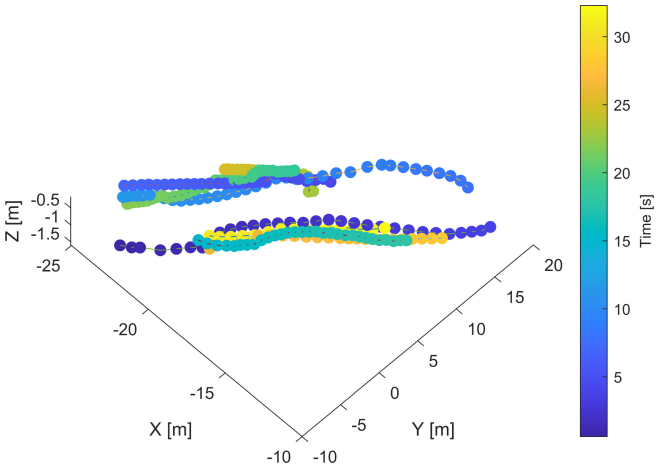


FIGURE 7. The output of the JPDAF approach, where multiple vehicles are detected from 3D LiDAR point clouds. The color bar represents the time of appearance of each vehicle.

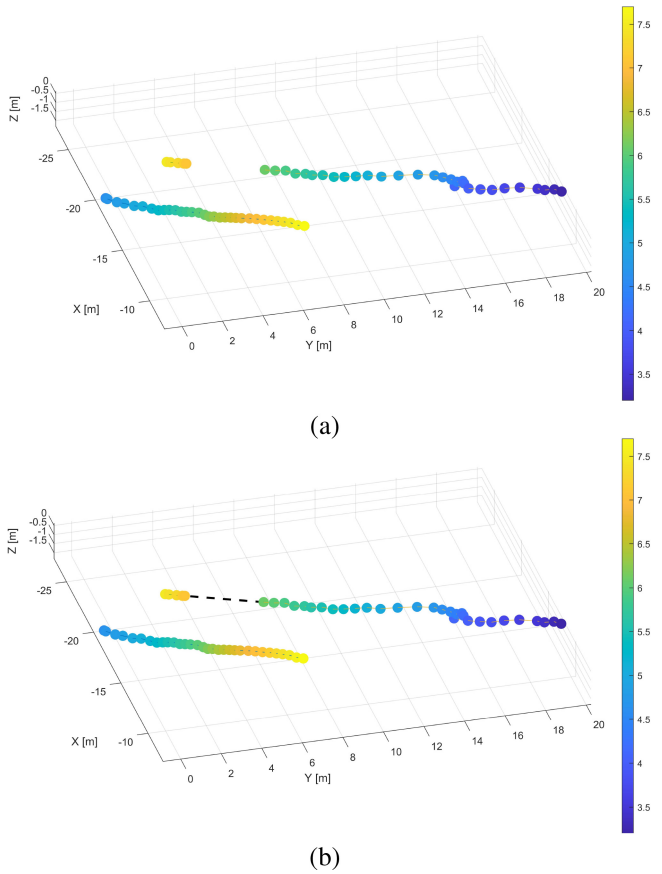


FIGURE 8. Blockage situations caused by dynamic occlusion in the adjacent lane: (a) an example of vehicle disappearance and reappearance; (b) re-association and re-identification of the same vehicle after disappearing from the sensing environment.

TABLE 1. Simulation parameters.

Symbol	Description	Value
P_{tx}	Transmit power	45 dBm
f_{MHz}	Frequency	2400 MHz
G_{tx}	Transmit antenna gain	10 dBi
G_{rx}	Receive antenna gain	10 dBi
α	Position/power weights	0.97
β	Velocity/power derivative weights	0.95
ϵ_b	Winner node learning rate	0.1
ϵ_n	Neighbor node learning rate	0.05
d_{th}	Distance threshold	5 meters
l_w	Lane width	3 meters
t_r	Re-association time	500 milliseconds
l_b	Boundary limits	$x: -30$ to -10 meters
		$y: -15$ to 35 meters
		$z: -1.5$ to 1.5 meters

granularity of clusters for power related data. This leads to coherent cluster structures and stable transition matrices, ultimately improving predictive performance. The number of clusters produced by GNG also depends on the underlying data distribution and dimensionality. Excessive cluster formation risks overfitting and increased computational cost, so the hyperparameter settings listed in the Table 1 are selected carefully to maintain a suitable balance.

In addition, the hyperparameters of the IC-MJPF during inference are primarily governed by the N_p, which defines the fundamental trade-off between prediction accuracy and computational cost. A larger particle set typically enhances prediction accuracy but increases computational load, whereas a smaller set reduces computational demand at the expense of accuracy. The appropriate choice of N_p depends strongly on the structure of the transition matrix, which governs the evolution of discrete states (sensing and communication words) and the likelihood of switching between models. A sparse transition matrix, such as that associated with dynamic RF power patterns, requires propagating a more N_p to maintain robustness. While, a smoother transition matrix, such as that associated with 3D vehicle position dynamics, does not necessitate an excessively large N_p. Consistent with the SIR strategy, only the highest weighted particles can be retained for prediction at the next time instant to reduce the computational complexity as discussed above. Later we will analyze the performance with varying N_p for both the data modalities.

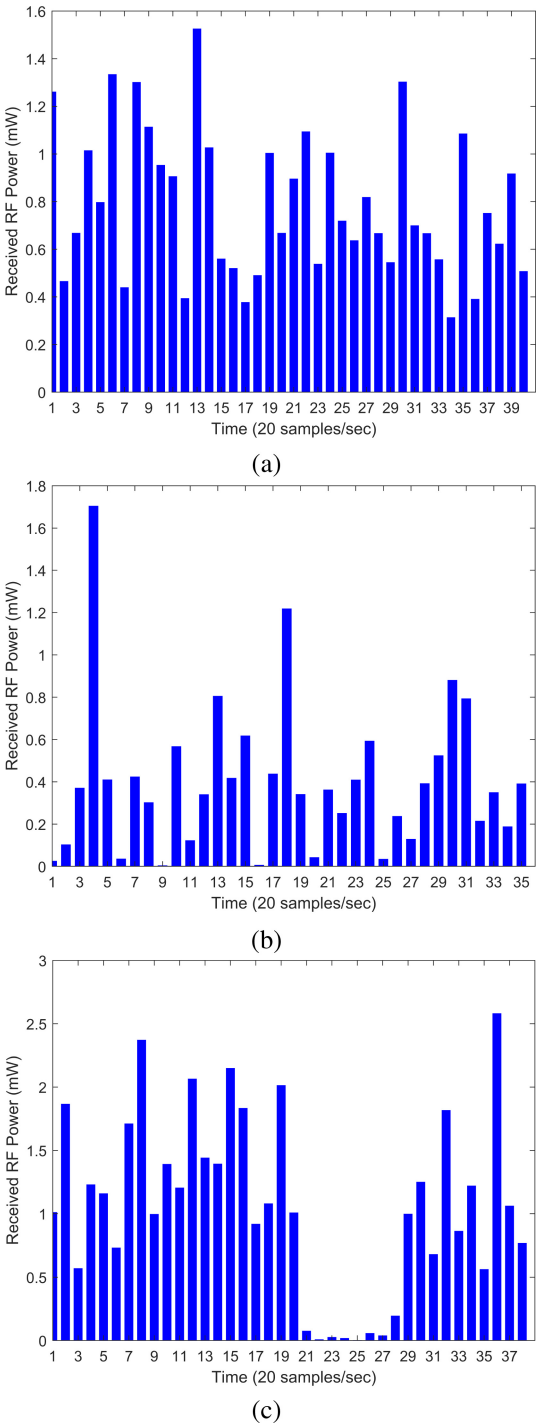


FIGURE 9. Examples of RF power received by the BS from transmitting AVs: (a) normal case under a Rician fading channel ($J = 10$); (b) normal case under a Rayleigh fading channel ($J = 0$); (c) LOS link blockage case.

A. DATA PREPROCESSING AND TRAINING RESULTS

Figure 7 illustrates the output of the JPDAF approach applied to a preprocessed sequence of 3D LiDAR frames, where multiple vehicles are tracked over time. The filter simultaneously maintains tracks of vehicles on a two-lane road within the environment. The JPDAF output comprises 3D trajectories of detected vehicles, including their x, y, and z positions, time of appearance, and tracking IDs. Only normal 3D tracks, those continuously detected without obstruction, are shown at each time instant. This information enables the BS to learn typical patterns of vehicle dynamics. Each 3D track corresponds to the longest confirmed trajectory that has persisted through several updates, while tentative tracks failing to meet the confirmation threshold are discarded. This

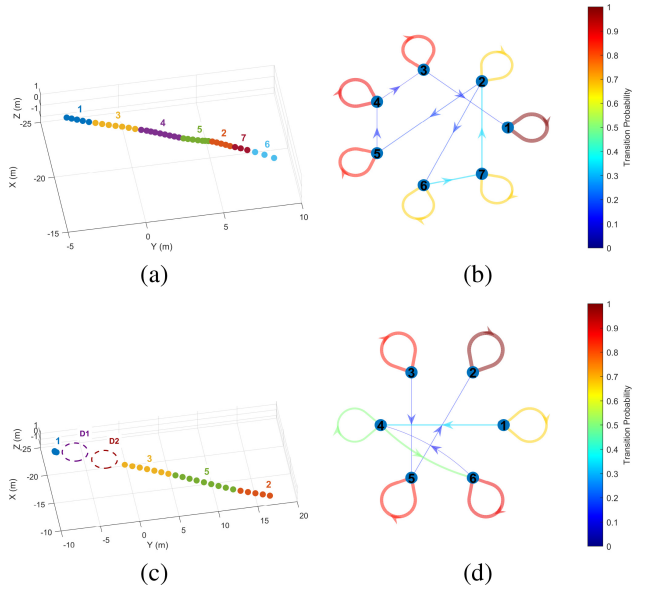


FIGURE 10. Unsupervised learning process: (a) clustering example of a detected vehicle using the JPDAF approach for a continuous normal track; (b) corresponding time-varying transition matrix of the vehicle in (a); (c) clustering example in the case of short-duration LiDAR blockage, where two dummy nodes (D1 and D2) are introduced; (d) corresponding time-varying transition matrix of the blocked vehicle in (c).

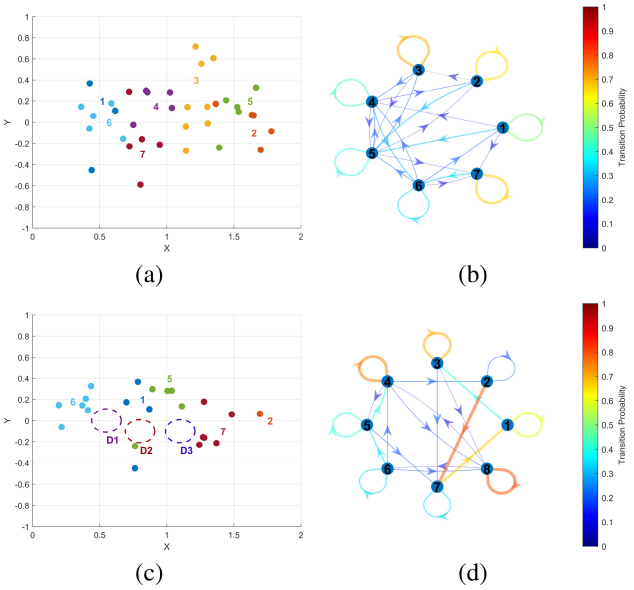


FIGURE 11. Unsupervised learning process: (a) clustering example of RF received power at the BS from a detected vehicle under normal conditions; (b) corresponding time-varying transition matrix of the RF received power in (a); (c) clustering example of RF received power in the case of long-duration power blockage, where three dummy nodes (D1, D2, and D3) are introduced; and (d) corresponding updated time-varying transition matrix of power blockage in (c).

ensures that only stable and temporally consistent tracks are retained, providing reliable input for time-series probabilistic modeling. The color bar represents the temporal information associated with the 3D position of each vehicle at the time of detection.

Figure 8(a) illustrates an example of a vehicle that temporarily disappears when the LiDAR sensor is occluded by blocking objects. The vehicle initially appears in the lane at time 3.2, disappears at time 6.1 due to occlusion caused by another object, and reappears at time 7.1 when the blockage has passed. All such disappeared tracks are processed using the re-association and re-identification Algorithm 1, which predicts the missing positions of the occluded tracks and re-associates the disappeared tracks with the newly observed ones based on the simulation parameters listed in Table 1. The output of the algorithm for the disappeared track shown in Figure 8(a) is presented in Figure 8(b). Through this process, the missing positions of all blockage-affected tracks are recovered and subsequently used to estimate the transition probabilities of the blockage models.

The RF power transmitted by the detected AVs and received at the BS is illustrated in Figure 9. These power levels are obtained under a realistic wireless environment influenced by Rician and Rayleigh fading, as modeled in (1). The corresponding simulation parameters are listed in Table 1. The temporal variation of the received RF power reflects the impact of the LOS component along with multipath NLOS effects caused by reflections, scattering, and diffraction from surrounding objects such as vehicles, buildings, trees, and pedestrians. Figure 9(a) presents the power received under a Rician fading channel with $J = 10$,

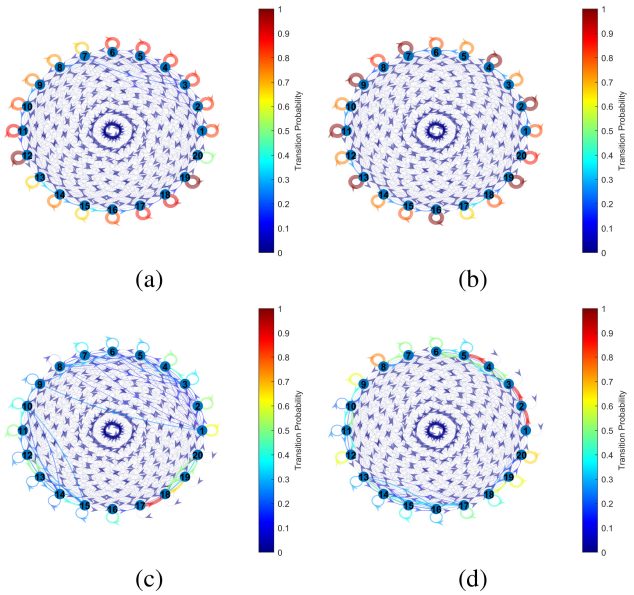


FIGURE 12. ITMs depicting probabilistic relationships between words denoting the vehicles' positions and RF received powers at the BS: (a) ITM for vehicles' position under normal situations; (b) ITM for vehicles' position under normal blockage situations; (c) ITM for RF received power under normal situations; (d) ITM for RF received power under normal blockage situations.

while Figure 9(b) shows the Rayleigh fading case with $J = 0$. Figure 9(c) depicts the power received when an obstruction blocks the LOS path. In this scenario, the power drops sharply to nearly zero, indicating that the AV is fully shadowed or blocked, and reliable detection or tracking is no longer feasible.

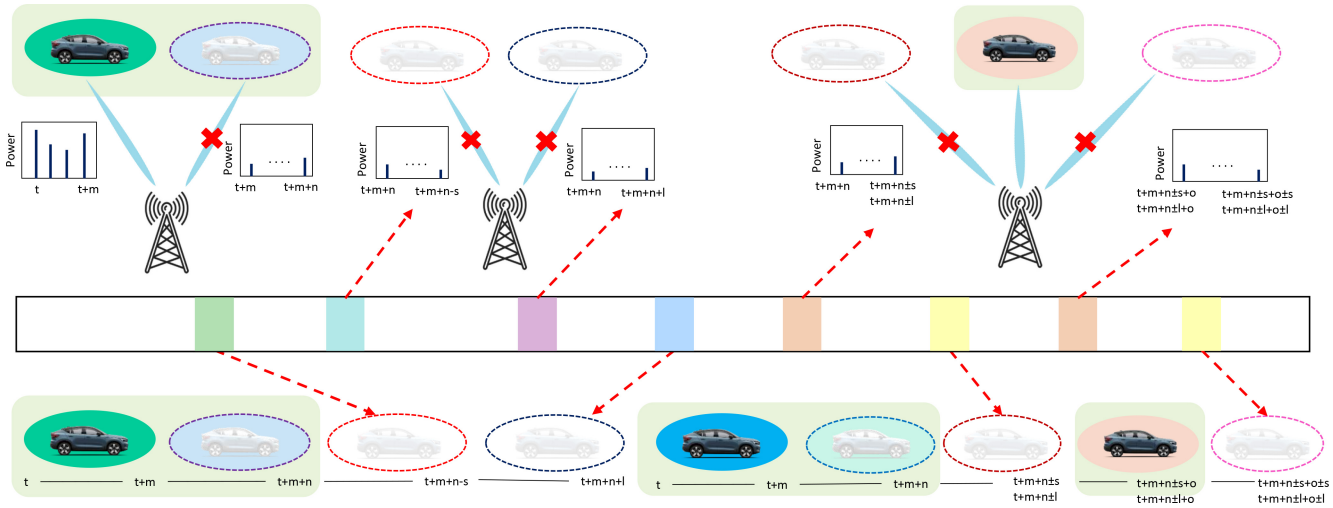


FIGURE 13. KLDA anomaly indicators at the interaction (word) levels, representing various blockage situations in both sensing and communication configurations.

The initial learning results of the GDBN models using the GNG approach for LiDAR sensing are shown in Figure 10. Figure 10(a) presents an example of GNG clusters for a normal vehicle track after providing input GSs. Each cluster represents a neuron or discrete node encapsulating the positions and velocities of the vehicles. Figure 10(b) illustrates the TM of the corresponding track shown in Figure 10(a). This quantifies the likelihood of the model switching from one neuron to another, revealing the temporal dynamics and behavioral patterns within the GSs, where the transition probabilities are indicated by a color bar. To learn a blockage model, we introduced dummy clusters representing the missing GSs obtained through the re-association and re-identification Algorithm 1, caused by LiDAR sensor blockages. An example of learning blockage is shown in Figure 10(c), where the dummy clusters $D1$ and $D2$ are introduced. The corresponding TM is presented in Figure 10(d), showing updated switching probabilities between normal and dummy clusters and vice versa. Additional such normal blockage models are generated with varying numbers of missing clusters, reflecting the fact that different clusters may disappear as the blockage size varies.

On the other hand, the GDBN models that learn from temporal sequences of received RF power patterns are illustrated in Figure 11. The GNG clusters obtained from the normal GDBN model using the input GPs are shown in Figure 11(a). The time-varying switching probabilities between the power clusters corresponding to these patterns are presented in Figure 11(b). The low likelihood values arise from multipath fading-induced fluctuations in the received power. The normal blockage model, in which three dummy nodes, $D1$, $D2$, and $D3$, represent the dummy GPs associated with LOS link blockages, is shown in Figure 11(c). The corresponding TM, which depicts the updated switching likelihoods between the normal and dummy power clusters (and vice versa), is provided in Figure 11(d).

Figure 12 shows the word-level ITMs for both the sensing and communication configurations under normal and normal blockage situations. Figures 12(a) and 12(c) illustrate the interaction transition probabilities among vehicles and among RF received power signals in normal operation, respectively. These matrices show the likelihood of moving from one discrete word state to another as time passes. Figures 12(b) and 12(d) depict the ITMs for both the sensing and communication domains under blockage conditions, reflecting transition dynamics behavior under dummy positions and power signals. These transition matrices serve as probabilistic models that help differentiate normal system behavior from blocked ones through the analysis of state transition patterns. It can be observed that the probabilities of remaining in the same word are high for vehicle positions but low for the received RF power. This is because most vehicles move along the same road lanes, maintaining consistent position states, whereas the RF received power fluctuates due to environmental factors.

B. TESTING CONFIGURATIONS

Testing new observations and making predictions should follow the same dynamic rules used to train the dynamic models from past experiences when blockages were present. The IC-MJPF provides two abnormality measures (see equations (14) and (15)) to identify sensor and power blockages, considering the different testing configurations defined below, and to determine whether the new testing data confirm the learned rules.

Reference Situation: In this case, the same set of detected vehicle tracks, both single and interactive, along with their corresponding RF received power (covering normal and blockage scenarios) are used as testing measurements to evaluate the blockage detection and abnormality detection capabilities of the learned IC-GDBN. These data are also employed to assess the proposed model's performance in terms of prediction accuracy and abnormality detection. This

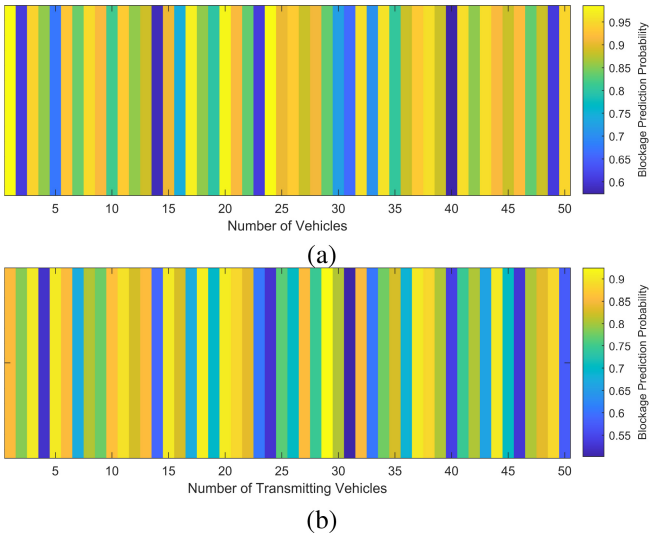


FIGURE 14. Blockage detection probabilities at latent interaction levels for both sensing and communication modalities: (a) LiDAR (vehicles' positions); (b) RF power received at the BS.

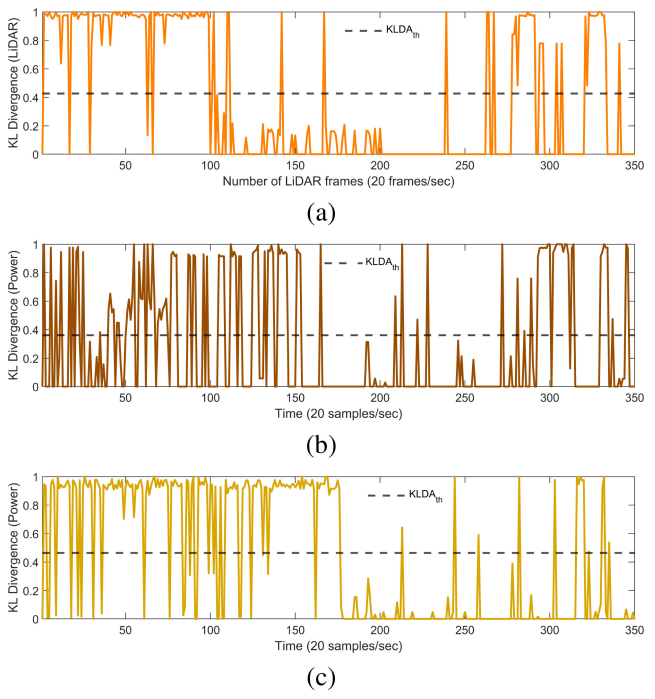


FIGURE 15. Abnormality signals: (a) KLDA detected at the LiDAR sensing branch; (b) KLDA detected at the RF power communication branch under a Rician fading channel ($J = 15$); (c) KLDA detected at the RF power branch under a Rayleigh fading channel ($J = 0$).

setup enables a comprehensive evaluation of the switching mechanism's behavior and joint blockage detection across both data modalities under the model(s) normal behavior.

1) SITUATION 1: DETECTING SHORT BLOCKAGES

This case defines an abnormal situation in which the blockage situation changes, specifically when its size decreases from $(t+m+n)$ to $(t+m+n-s)$, as illustrated in Figure 13. Here, $(t+m+n)$, represented by the green-shaded area,

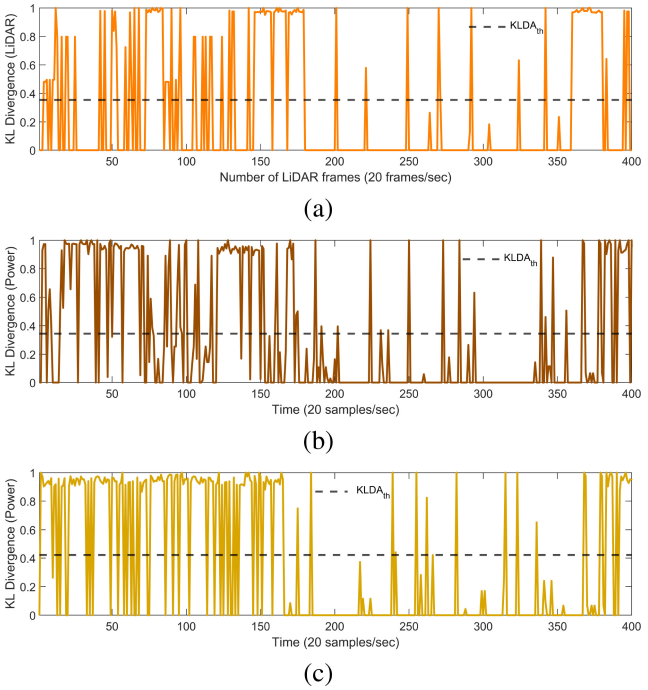


FIGURE 16. Abnormality signals: (a) KLDA detected at the LiDAR sensing branch; (b) KLDA detected at the RF power communication branch under a Rician fading channel ($J = 5$); (c) KLDA detected at the RF power branch under a Rayleigh fading channel ($J = 0$).

indicates that the IC-GDBN has prior knowledge of both normal and blockage scenarios during this interval, while s denotes the short time duration of the reduced blockage. This situation is visualized by an anomaly indicator, shown in light green for the vehicles' 3D positions. Similarly, the RF received power under this condition is depicted with the powder blue color.

2) SITUATION 2: DETECTING LONG BLOCKAGES

The dynamic situation may also arise when the blockage type changes, specifically, when the size of the blockage increases beyond the BS expectation, i.e., from $(t+m+n)$ to $(t+m+n+l)$, where l denotes the long time duration of the increased blockage. For example, two continuous vehicle blockages may occur, resulting in a compounded obstruction that persists longer than anticipated and introduces non-monotonic variations in the received signal strength. Such deviations may indicate unexpected environmental conditions, represented with the light purple color for sensing configurations and with the sky blue color as an anomaly indicator for communication configurations, as shown in Figure 13.

3) SITUATION 3: DETECTING UNKNOWN BLOCKAGES

This case describes an abnormal situation in which two or more blockages appear in the environment. These occlusions can vary in type and duration, either short or long, or a short-duration blockage may follow a long-duration one, and vice versa. For the LiDAR sensor, this situation is

represented by a light yellow anomaly indicator, while the corresponding RF received power is visualized using a light peach color, as shown in Figure 13. Considering such abnormal situations allows us to evaluate the performance and behavior of the proposed IC-GDBN under various sensing and communication blockage scenarios in a dynamic environment.

C. BLOCKAGE AND ABNORMALITY DETECTION PERFORMANCE

Figure 14 presents the blockage prediction probabilities for the detected vehicles along with their corresponding RF received power. The results indicate that the interactive models achieve better performance in identifying different types of blockages. In the case of position-based vehicle blockages, higher prediction accuracy is observed, as the model can reliably recognize high-probability sensing words within the global dictionary. Since the vehicles operate on two lanes, their spatial configurations are well represented in the dictionary, as illustrated in Figure 14(a). The blockage prediction probabilities for RF received power are shown in Figure 14(b). Compared to LiDAR-based physical blockages, RF power blockage detection is more sensitive due to the inherent randomness of power fluctuations and channel interference. This behavior indicates that power-related blockages are generally more challenging to detect than physical obstructions [13]. Furthermore, multiple examples are represented to capture diverse blockage conditions, as discussed earlier. In certain cases, lower prediction probabilities are observed when the blockage environment changes and deviates from the learned patterns encoded in the global dictionary of the IC-GDBN model.

The blockage prediction and KLDA abnormality detection results are shown in Figure 15, where initially 100 particles are propagated by the filter. In Figure 15(a), high abnormality levels appear at the beginning when the IC-GDBN model trained on the DeepSense 6G environment (Scenario 32) is tested on the UC3M campus experimental blockage dataset. Since this environment is new for the BS, its characteristics deviate from the rules encoded in the dictionaries. In this case, blockage prediction is not yet possible, and the system primarily identifies abnormalities from LiDAR frames 1 to approximately 100. Once new normal and blockage models are created and their corresponding rules are encoded in the IC-GDBN, the BS begins to predict blockages accurately from LiDAR frame 100, and onward until other unfamiliar blockage patterns appear, the model again marks them as abnormal due to their novelty. A similar pattern can also be observed for RF received power. Because the environment is new, the model initially outputs high abnormality values under the LOS Rician fading channel ($J = 15$) in Figure 15(b) and the NLOS Rayleigh fading channel ($J = 0$) in Figure 15(c). These abnormality peaks correspond to observations that diverge from the learned models. After generating new models for both fading channels, the abnormality peaks are decreased,

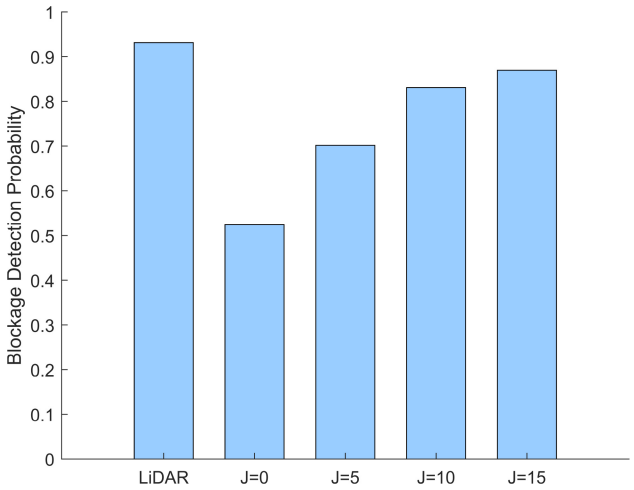


FIGURE 17. Blockage detection probabilities for LiDAR (vehicles' positions) and RF received power under Rayleigh fading and Rician fading channels with varying J-factors (dB).

indicating more reliable predictions until a new unfamiliar situation emerges. This demonstrates that the framework remains effective across environments once new rules are incorporated for previously unseen conditions.

Comparable behaviour appears in Figure 16 when the model is evaluated on DeepSense 6G Scenarios 33 and 34. In Figure 16(a), the initial abnormality levels are relatively low because the environment remains similar: vehicles follow the same lane structure as in Scenario 32, allowing some previously encoded rules to remain valid. However, differences in BS location and nighttime LiDAR operation introduce deviations that are detected as initial abnormalities. For the received power, the change in BS position produces unexpected signal dynamics and leads to high initial abnormality values, after which the model successfully detects RF received power blockages for both LOS and NLOS channels. For the LiDAR data, the system continues to operate effectively due to environmental similarity, although the altered BS position again introduces variations in received power. After learning the new rules, blockage prediction improves for both Rayleigh ($J = 0$) and Rician ($J = 5$) channels. Therefore, once the BS is equipped with rules adapted to the new environment, the model can be adjusted to provide reliable interpretation of LiDAR observations for vehicle positions as well as Rayleigh and Rician fading characteristics for RF received power. These results highlight the capability of the framework to operate across these new environments through novel rules adaptation.

The average blockage detection performance within 50 milliseconds, shown in the Figure 17, indicates that the BS more reliably detects vehicle positional blockages obtained from the LiDAR sensor compared to NLOS and LOS power blockages under different J-factor values. The model demonstrates a stronger capability in identifying positional blockages, achieving a probability greater than 0.93, whereas the power-based blockages exhibit more

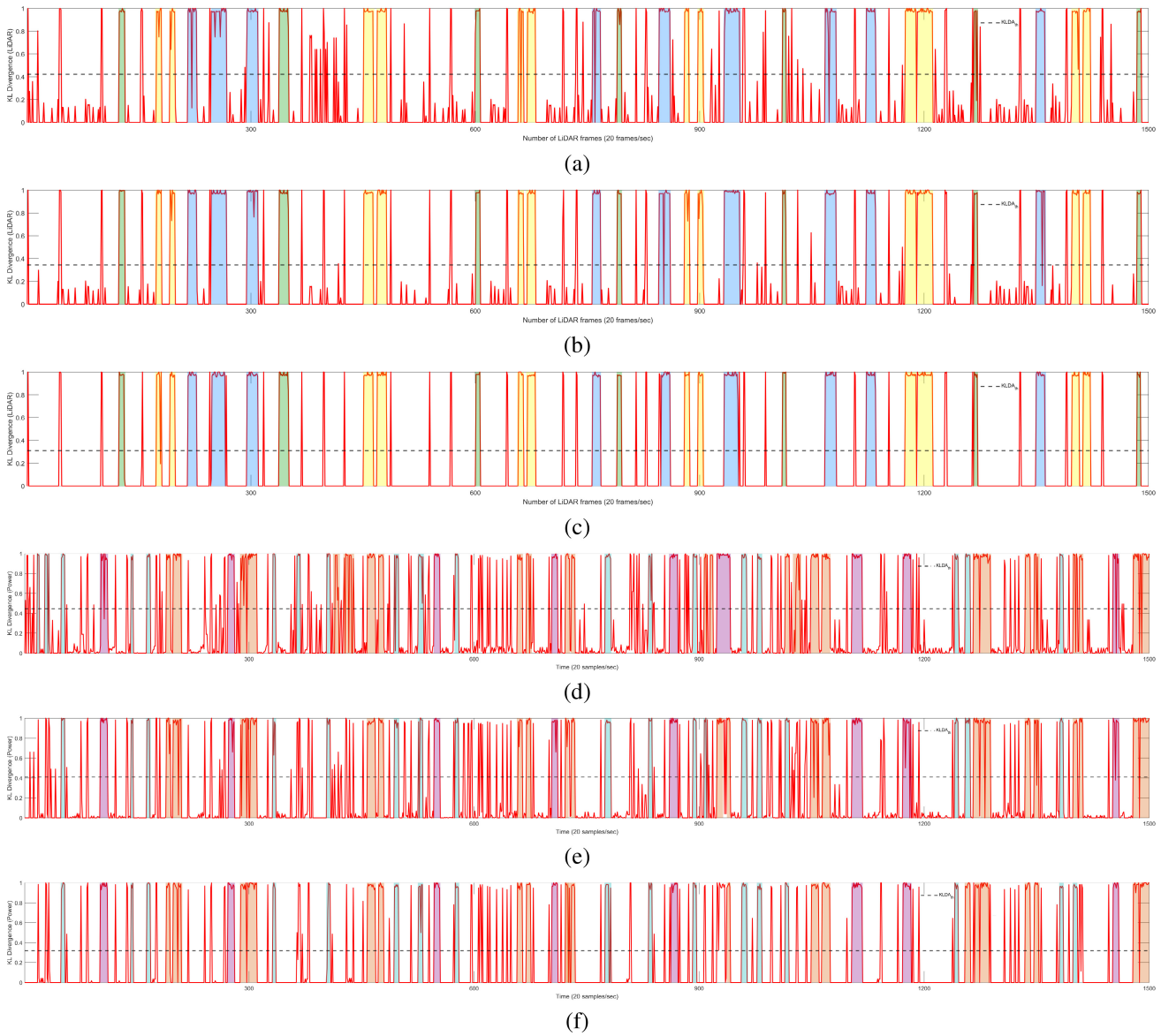


FIGURE 18. Abnormality signals: (a) KLDA detected at the LiDAR sensing branch at $N_p = 25$; (b) KLDA detected at the LiDAR sensing branch at $N_p = 50$; (c) KLDA detected at the LiDAR sensing branch at $N_p = 100$; (d) KLDA detected at the RF power branch under a Rician fading channel ($J = 10$) at $N_p = 25$; (e) KLDA detected at the RF power branch under a Rician fading channel ($J = 10$) at $N_p = 50$; (f) KLDA detected at the RF power branch under a Rician fading channel ($J = 10$) at $N_p = 100$.

dynamic behaviour. Under Rayleigh fading ($J = 0$), the blockage detection probability drops to 0.52, reflecting the absence of a dominant LOS component. For Rician fading, the probability increases to 0.70 at $J = 5$ dB, reaches an average of 0.83 at $J = 10$ dB, and further improves to 0.87 at $J = 15$ dB. These results confirm that stronger LOS components enable more reliable blockage prediction, whereas multipath-dominated channels reduce detection accuracy. This trend is consistent with the findings reported in [13], where LiDAR-based blockage prediction outperforms mmWave power-based blockage prediction in the DeepSense 6G dataset.

Figure 18 presents the blockage and abnormality detection performance on the Scenario 32 at varying N_p. When the likelihood that the predicted blockage lies within the predicted word is high, the filter outputs a low abnormality signal; otherwise, it produces a high abnormality signal. High abnormality levels are observed at time instants where the blockage size is short or long, or when two unseen blockages appear in the scene. These cases are highlighted using different indicators in the figure. It is further observed that abnormality levels are higher for smaller particle counts (e.g., $N_p = 25$), decrease for $N_p = 50$, and become nearly negligible for $N_p = 100$ for both LiDAR and RF received power observations. In addition, it can also be observed that more anomalies are detected for power signals compared to vehicle positions, consistent with trends observed in previous dataset scenarios. This is again due to the random nature of

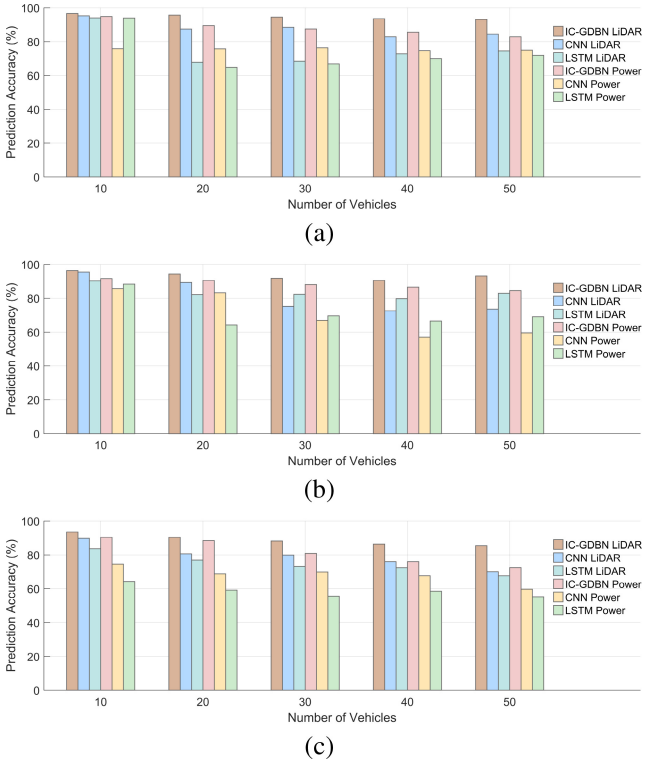


FIGURE 19. Performance comparison of IC-GDBN, CNN, and LSTM in terms of blockage prediction accuracy under (a) Situation 1, (b) Situation 2, and (c) Situation 3.

power signals, when the predicted words may not always align correctly with the estimated words.

D. COMPARISON PERFORMANCE

A further analysis has been conducted to compare the performance of the proposed IC-GDBN with baseline methods, including CNN and LSTM-based approaches, for both vehicle trajectories and the corresponding received power patterns. Both the CNN and LSTM models were trained using the same input vectors as those provided to the GNG, ensuring a fair basis for comparison. For both LiDAR sensing and communication configurations, 6D and 2D input vectors corresponding to positions and power information, respectively, were employed along with different blockage labels.

For CNN training, a convolutional layer was initially applied to detect data patterns, followed by a ReLU layer to introduce nonlinearities. For LSTM training, an LSTM layer was used to capture temporal dependencies and learn data patterns across each time step, followed by a batch normalization layer to stabilize and accelerate training. Both CNN and LSTM use a fully connected layer to map the learned features to the target classes and a Softmax layer to convert class scores into probability distributions.

The performance comparison in terms of blockage prediction accuracy for the previously defined scenarios is shown in Figure 19. It can be observed that IC-GDBN consistently outperforms the CNN and LSTM methods under various situations for both LiDAR and power signals. Specifically, the proposed method demonstrates the ability to predict different types of physical (positions) and digital (power) blockages within 50 milliseconds.

This indicates that the proposed model effectively learns the dynamic characteristics of signals across multiple hierarchical layers, thereby enabling accurate prediction of future behavior based on the learned patterns encapsulated within the IC-GDBN dictionary. While CNN and LSTM rely on supervised learning, where input vectors are paired with labeled data for each scenario during training, the IC-GDBN adopts an unsupervised learning strategy. This allows the model to discover underlying patterns without requiring explicit labels. Moreover, the IC-GDBN demonstrates the capability to explicitly learn and represent the relationships among hidden-layer random variables within the network, which facilitates abnormality detection in different situations defined in previous results. In contrast, LSTM and CNN models treat inter-layer dependencies as opaque closed boxes, making their internal reasoning difficult to interpret. This limitation constrains their ability to learn by understanding the model's internal mechanisms, an aspect essential for cognitive dynamic systems, which must continuously learn and adapt to environmental changes.

VI. CONCLUSION

In this paper, various types of dynamic blockages are detected using time-series LiDAR point clouds and power signal modalities, underscoring the role of integrated sensing and communication (ISAC) in vehicle-to-everything (V2X) applications. To jointly exploit these information sources, an interactive coupled generalized dynamic Bayesian network (IC-GDBN) is proposed, providing a probabilistic, hierarchical, and explainable framework that enables an intelligent base station (BS) to interpret dynamic environmental conditions from 3D LiDAR perception and RF power-signal variations between connected BS and surrounding vehicles. During the offline training phase, diverse normal and blockage models are constructed for both the signals in an unsupervised learning manner, where the dynamic rules are encoded into a global dictionary of the IC-GDBN. In the online inference phase, an interactive coupled Markov jump particle filter (IC-MJPF) leverages the probabilistic knowledge learned by the IC-GDBN to detect LiDAR and RF power blockages in both line-of-sight (LOS) and non-line-of-sight (NLOS) conditions. The framework identifies anomalous signals at higher hierarchical levels when environmental behavior deviates from encoded rules. Experimental results show that the model more reliably detects vehicle positional blockages compared to RF power-based dynamic blockages, which exhibit higher sensitivity. Additional evaluations demonstrate that the IC-GDBN outperforms existing convolutional neural networks (CNN) and long-term short memory (LSTM) approaches, achieving superior prediction accuracy within 50 milliseconds across a range of blockage scenarios.

VII. LIMITATIONS AND FUTURE WORK

Despite the contributions of our paper, this work has some limitations which provide opportunities for future research.

First, incremental learning is not incorporated at the current stage. The system instead follows a self-aware learning paradigm: during each experience, the IC-MJPF gives explanations based on the learned models. For incremental learning, anomalies and errors identified by the BS after each experience could be stored and later used to update the models offline, thereby reducing computational demands for resource-constrained hardware during the online adaptive learning.

Second, real-time adaptive learning under extreme weather conditions, such as dense fog, heavy rain, or other physical events, including sensor degradation, malicious attacks, or multimodal sensor failures, is not addressed in this study. Since the likelihood models are specifically tailored for blockage detection, such events cannot be explained within the present formulation. Nevertheless, the same reasoning strategy could be extended to these physical phenomena by developing corresponding likelihood models and integrating them into the learning framework. Addressing these physical events and training the models under such conditions represents a natural and important direction for future work, which we plan to investigate in subsequent studies.

Third, action selection, a crucial capability of self-aware agents is needed in the current approach. Instead, this study establishes a foundational step toward active mitigation strategies that can respond to detected abnormalities and blockages. Future extensions could incorporate decision-making or action selection adaptation, particularly for power blockages which are directly affecting the V2I communication. Potential actions include introducing alternative power-modulation schemes or deploying additional roadside units to reinforce signal strength for blocked vehicles. These capabilities can be enabled through active inference principles [43], wherein actions are selected probabilistically by minimizing surprise (prediction errors) and expected free energy. Such dynamic decision-making would balance exploration and exploitation, adapt to varying blockage conditions, and ultimately enhance V2X network performance by ensuring continuous vehicle–infrastructure connectivity.

REFERENCES

[1] G. Singh, A. Srivastava, V. A. Bohara, M. Noor-A-Rahim, Z. Liu, and D. Pesch, "Toward 6G-V2X: Aggregated RF-VLC for ultra-reliable and low-latency autonomous driving," *IEEE Commun. Stand. Mag.*, vol. 8, no. 4, pp. 80–87, Dec. 2024.

[2] M. Temiz and C. Masouros, "Unsupervised learning-based low-complexity integrated sensing and communication precoder design," *IEEE Open J. Commun. Soc.*, vol. 6, pp. 3543–3554, 2025.

[3] A. Magbool, V. Kumar, Q. Wu, M. Di Renzo, and M. F. Flanagan, "A survey on integrated sensing and communication with intelligent metasurfaces: Trends, challenges, and opportunities," *IEEE Open J. Commun. Soc.*, vol. 6, pp. 7270–7318, 2025.

[4] X. Cheng, D. Duan, S. Gao, and L. Yang, "Integrated sensing and communications (ISAC) for vehicular communication networks (VCN)," *IEEE Internet Things J.*, vol. 9, no. 23, pp. 23441–23451, Dec. 2022.

[5] A. Krayani, G. Barabino, L. Marcenaro, and C. Regazzoni, "Integrated sensing and communication for joint GPS spoofing and jamming detection in vehicular V2X networks," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, 2023, pp. 1–7.

[6] S. Memon, A. Krayani, P. Zontone, L. Marcenaro, D. M. Gomez, and C. Regazzoni, "Learning 3D LiDAR perception models for self-aware autonomous systems," in *Proc. 27th Int. Conf. Inf. Fusion (FUSION)*, 2024, pp. 1–8.

[7] M. Al-Quraan et al., "Enhancing reliability in federated mmWave networks: A practical and scalable solution using radar-aided dynamic blockage recognition," *IEEE Trans. Mobile Comput.*, vol. 23, no. 10, pp. 10146–10160, Oct. 2024.

[8] V. Ilçi, M. Yavuz, Y. Şişman, K. Par, and A. U. Peker, "Determination of optimum coordinate transformation parameters for GNSS and LiDAR-based localization in automated vehicles," *IEEE Trans. Intell. Transp. Syst.*, vol. 25, no. 10, pp. 13217–13230, Oct. 2024.

[9] J. Mia and M. H. Amini, "A secure object detection technique for intelligent transportation systems," *IEEE Open J. Intell. Transp. Syst.*, vol. 5, pp. 495–508, 2024.

[10] B. Wang, J. Zheng, N. Mitton, and C. Li, "A packet collision avoidance resource selection scheme for reliable intra-platoon message delivery in a C-V2X network," *IEEE Trans. Veh. Technol.*, vol. 74, no. 7, pp. 11366–11377, Jul. 2025.

[11] H.-Y. Lin and C.-H. Tseng, "Abnormal activity detection and classification of bus passengers with in-vehicle image sensing," *IEEE Access*, vol. 12, pp. 23057–23065, 2024.

[12] C. Zhang et al., "Detecting the anomalies in LiDAR pointcloud," 2023, *arXiv:2308.00187*.

[13] S. Wu, C. Chakrabarti, and A. Alkhateeb, "LiDAR-aided mobile blockage prediction in real-world millimeter wave systems," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, 2022, pp. 2631–2636.

[14] Y. Sun et al., "Multi-functional RIS-assisted semantic anti-jamming communication and computing in integrated aerial-ground networks," *IEEE J. Sel. Areas Commun.*, vol. 42, no. 12, pp. 3597–3617, Dec. 2024.

[15] L. Zhi et al., "Self-powered absorptive reconfigurable intelligent surfaces for securing satellite-terrestrial integrated networks," *China Commun.*, vol. 21, no. 9, pp. 276–291, Sep. 2024.

[16] J. Chen et al., "Secure wireless-powered zeRIS communications," *IEEE Trans. Wireless Commun.*, early access, Jul. 4, 2025, doi: 10.1109/TWC.2025.3582430.

[17] Y. Zhao and M. Jian, "Applications and challenges of reconfigurable intelligent surface for 6G networks," 2022, *arXiv:2108.13164*.

[18] G. Chopra and S. Ahmed, "RIS-assisted integrated sensing and communication: Applications, challenges and usecase scenario," *Discover Appl. Sci.*, vol. 7, no. 7, p. 650, 2025.

[19] M. Åström, P. Gentner, O. Haliloglu, B. Makki, and O. Tageman, "RIS in cellular networks–challenges and issues," 2024, *arXiv:2404.04753*.

[20] Z. Lin, Z. Feng, K. Guo, A. Nauman, D. Niyato, and J. Wang, "AI-driven seamless and massive access in space-air-ground integrated networks," *IEEE Wireless Commun.*, vol. 32, no. 3, pp. 72–79, Jun. 2025.

[21] S. Wu, C. Chakrabarti, and A. Alkhateeb, "Proactively predicting dynamic 6G link blockages using LiDAR and in-band signatures," *IEEE Open J. Commun. Soc.*, vol. 4, pp. 392–412, 2023.

[22] G. Charan, M. Alrabeiah, and A. Alkhateeb, "Vision-aided dynamic blockage prediction for 6G wireless communication networks," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, 2021, pp. 1–6.

[23] U. Demirhan and A. Alkhateeb, "Radar aided proactive blockage prediction in real-world millimeter wave systems," in *Proc. IEEE Int. Conf. Commun. (ICC)*, 2022, pp. 4547–4552.

[24] U. Utkarsh, D. C. Maddix, R. Ma, M. W. Mahoney, and Y. Wang, "End-to-end probabilistic framework for learning with hard constraints," 2025, *arXiv:2506.07003*.

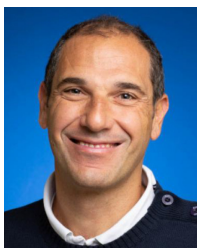
[25] Q. Chen, G. Huang, and Y. Wang, "The weighted cross-modal attention mechanism with sentiment prediction auxiliary task for multimodal sentiment analysis," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 30, no. 3, pp. 2689–2695, Jul. 2022.

[26] K. Friston, "The free-energy principle: A rough guide to the brain?" *Trends Cogn. Sci.*, vol. 13, no. 7, pp. 293–301, 2009.
[27] S. Haykin and J. M. Fuster, "On cognitive dynamic systems: Cognitive neuroscience and engineering learning from each other," *Proc. IEEE*, vol. 102, no. 4, pp. 608–628, Apr. 2014.
[28] A. R. Damasio, *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*. San Diego, CA, USA: Harcourt Brace, 1999.
[29] G. Slavic, P. Zontone, L. Marcenaro, D. M. Gómez, and C. Regazzoni, "Vehicle localization in an explainable dynamic Bayesian network framework for self-aware agents," *Inf. Fusion*, vol. 122, Oct. 2025, Art. no. 103136.
[30] A. S. Alemaw, G. Slavic, P. Zontone, L. Marcenaro, D. M. Gomez, and C. Regazzoni, "Modeling interactions between autonomous agents in a multi-agent self-awareness architecture," *IEEE Trans. Multimedia*, vol. 27, pp. 5035–5049, 2025.
[31] M. Scutari, "Learning Bayesian networks with the bnlearn R package," *J. Stat. Softw.*, vol. 35, pp. 1–22, Jul. 2010.
[32] A. Krayani, A. S. Alam, L. Marcenaro, A. Nallanathan, and C. Regazzoni, "Automatic jamming signal classification in cognitive UAV radios," *IEEE Trans. Veh. Technol.*, vol. 71, no. 12, pp. 12972–12988, Dec. 2022.
[33] A. Alkhateeb et al., "DeepSense 6G: A large-scale real-world multimodal sensing and communication dataset," *IEEE Commun. Mag.*, vol. 61, no. 9, pp. 122–128, Sep. 2023.
[34] T. Ashiwa, S. Yamada, R. Narita, M. Adachi, K. Sekiguchi, and K. Nonaka, "Pedestrian tracking accommodating point-cloud displacement between multiple LiDARs based on joint probabilistic data association filter," in *Proc. IEEE/SICE Int. Symp. Syst. Integr. (SII)*, 2024, pp. 881–887.
[35] J. Chen, J. He, G. Wang, and B. Peng, "A maritime multitarget tracking method with non-Gaussian measurement noises based on joint probabilistic data association," *IEEE Trans. Instrum. Meas.*, vol. 74, pp. 1–12, 2025.
[36] L. Kang, C. Zhao, and J. Huang, "Object re-identification based on federated incremental Subgradient proximal optimization," *IEEE Open J. Comput. Soc.*, vol. 6, pp. 60–71, 2025.
[37] H. A. H. Alobaidy et al., "Empowering extreme communication: Propagation characterization of a LoRa-based Internet of Things network using hybrid machine learning," *IEEE Open J. Commun. Soc.*, vol. 5, pp. 3997–4023, 2024.
[38] S. Memon, A. Krayani, P. Zontone, L. Marcenaro, D. M. Gomez, and C. Regazzoni, "Leveraging dynamic interaction models for anomalous behavior detection in 3D environments," in *Proc. 33rd Eur. Signal Process. Conf. (EUSIPCO)*, 2025, pp. 1677–1681.
[39] S. Memon, A. Krayani, P. Zontone, L. Marcenaro, D. M. Gomez, and C. Regazzoni, "LiDAR-based dynamic blockage prediction: A data-driven approach for learning interactive Bayesian models," in *Proc. IEEE Int. Workshop Technol. Def. Security (TechDefense)*, Nov. 2025.
[40] J. Benito-Picazo, E. J. Palomo, E. Domínguez, and A. D. Ramos, "Image clustering using a growing neural gas with forbidden regions," in *Proc. IEEE Int. Joint Conf. Neural Netw. (IJCNN)*, 2020, pp. 1–7.
[41] Q. Ding, P. Yin, J. Ai, and S. Han, "Incremental classification for myoelectric manifold representation with matrix-formed growing neural gas network," *IEEE Trans. Ind. Informat.*, vol. 20, no. 8, pp. 10065–10073, Aug. 2024.
[42] C. P. Ezenkwu and A. Starkey, "Unsupervised temporospatial neural architecture for sensorimotor map learning," *IEEE Trans. Cogn. Dev. Syst.*, vol. 13, no. 1, pp. 223–230, Mar. 2021.
[43] A. Krayani, F. Obite, A. S. Alam, L. Marcenaro, A. Nallanathan, and C. Regazzoni, "A semantic-aware resource allocation for emerging UAV-NOMA networks empowered by Bayesian active inference," *IEEE Trans. Cogn. Commun. Netw.*, early access, Nov. 27, 2025, doi: 10.1109/TCCN.2025.3637049.
[44] H. Iqbal, P. Marin, L. Marcenaro, D. M. Gómez, and C. Regazzoni, "Bayesian geometric-based interactions learning model for self-aware autonomous agents," *Signal Process.*, vol. 239, Feb. 2025, Art. no. 110237.
[45] A. Krayani, M. Baydoun, L. Marcenaro, A. S. Alam, and C. Regazzoni, "Self-learning Bayesian generative models for jammer detection in cognitive-UAV-radios," in *Proc. IEEE Glob. Commun. Conf. (GLOBECOM)*, 2020, pp. 1–7.

SALEEMULLAH MEMON (Member, IEEE) received the B.E. degree in electronic engineering from the Quaid-e-Awam University of Engineering, Science and Technology, Pakistan, in 2017, and the M.E. degree in electronics and communication engineering from the Beijing University of Posts and Telecommunications, China, in 2019. He is currently pursuing the Joint Ph.D. degree with the University of Genoa, Italy, and the University Carlos III of Madrid, Spain. He served as a Lecturer with the Department of Electronic Engineering, Mehran University of Engineering and Technology, Shaheed Zulfiqar Ali Bhutto Campus, Khairpur, Pakistan, from 2021 to 2023. His research interests include integrated sensing and communication, V2X, self-awareness for autonomous agents, sensor fusion, anomaly detection, and artificial intelligence.

ALI KRAYANI (Member, IEEE) received the bachelor's degree in telecommunication engineering from the Politecnico di Torino, Italy, in 2014, and the master's degree in telecommunication engineering from the University of Florence, Italy, in 2017. He is an Assistant Professor with the Department of Electrical, Electronic, Telecommunications Engineering, and Naval Architecture (DITEN), University of Genoa, Italy. In April 2022, he was awarded a joint Ph.D. degree with the University of Genoa and the Queen Mary University of London, London, U.K. He has worked as a software engineer in various companies and was a Postdoctoral Research Fellow with DITEN from 2021 to 2023. His current research interests include integrated sensing and communication, cognitive radios, AI-enabled radios, wireless communications (5G and 6G), UAV communications, physical layer security, IoT, semantic communications, UAV swarms, NOMA, federated learning, and artificial intelligence. In 2023, he received the Best Paper Award at the IEEE Wireless Communications and Networking Conference. He serves as a guest editor and a reviewer for several academic journals.

PAMELA ZONTONE (Member, IEEE) received the Laurea degree in electronic engineering and the Ph.D. degree in information and industrial engineering from the University of Udine, Italy, in 2004 and 2008, respectively. From 2009 to 2011, she was a Postdoctoral Fellow with the University of Trento, Italy. In 2017, she joined the DPIA, University of Udine. She is currently an Assistant Professor with the DITEN, University of Genoa, Italy. Her research interests include multimodal signal processing for autonomous systems, unsupervised techniques for anomaly detection, multidimensional signal processing, biophysical signal processing, and machine learning.



LUCIO MARCENARO (Senior Member, IEEE) received the degree in electronic engineering in 1999, and the Ph.D. degree in computer science and electronic engineering in 2003. He is currently an Associate Professor of Telecommunications with the Polytechnic School of Engineering, University of Genoa. He has over 20 years of experience in signal processing and image sequence analysis. He is the author of about 160 scientific papers related to signal processing for computer vision and cognitive radio. His main current research interests include video processing for event recognition, detection, and localization of objects in complex scenes, distributed heterogeneous sensors, and bio-inspired cognitive autonomous systems.

DAVID MARTÍN GÓMEZ (Member, IEEE) received the degree in industrial physics (automation) from the National University of Distance Education (UNED) in 2002, and the Ph.D. degree in computer science from the Spanish Council for Scientific Research (CSIC) and UNED, Spain, in 2008. He has been a member of the Intelligent Systems Laboratory since 2011. He is currently a Professor with the Carlos III University of Madrid. In 2014, he was awarded with the VII Barreiros Foundation Award to the best research in the automotive field. In 2015, the IEEE Society has awarded him as the Best Reviewer of the 18th IEEE International Conference on Intelligent Transportation Systems.

CARLO REGAZZONI (Senior Member, IEEE) is currently a Full Professor of Cognitive Telecommunications Systems with the DITEN, University of Genoa, Italy. He has been responsible of several national and EU funded research projects. He is also the coordinator of international Ph.D. courses on interactive and cognitive environments involving several European universities. He served as a general chair for several conferences and an associate/guest editor for several international technical journals. He has served in many roles in governance bodies for IEEE SPS. He is serving as the Vice President-Conferences for the IEEE Signal Processing Society from 2015 to 2017.