

# A Semantic-Aware Resource Allocation for Emerging UAV-NOMA Networks Empowered by Bayesian Active Inference

Ali Krayani<sup>✉</sup>, Member, IEEE, Felix Obite<sup>✉</sup>, Member, IEEE, Atm S. Alam<sup>✉</sup>, Member, IEEE,  
Lucio Marcenaro<sup>✉</sup>, Senior Member, IEEE, Arumugam Nallanathan<sup>✉</sup>, Fellow, IEEE,  
and Carlo Regazzoni<sup>✉</sup>, Senior Member, IEEE

**Abstract**—Future wireless networks demand intelligent, data-intensive services with high reliability and low latency, motivating semantic-aware communication as a transformative paradigm. In this paper, we address the NP-hard problem of resource allocation in UAV-based NOMA networks by proposing a novel framework named Active Semantic Generalized Dynamic Bayesian Network (Active-SGDBN). Unlike conventional optimization or data-driven methods, our approach integrates expert domain knowledge, active inference, and semantic context-aware reasoning to enable real-time, interpretable decision-making. Specifically, we optimize power allocation based on UAV mobility to maximize system sum-rate while encoding semantic structures through unsupervised clustering and Bayesian inference. Simulation results demonstrate that Active-SGDBN achieves near-optimal performance and surpasses benchmark schemes in both sum-rate and bit error rate (BER), validating its effectiveness and generalization capability.

**Index Terms**—6G, Resource Allocation, UAV-NOMA, Active Inference, World Model.

## I. INTRODUCTION

Next-generation wireless networks, including 6G and beyond, are expected to support a variety of intelligent applications. Specifically, the explosive growth in traffic anticipated from Internet of Things (IoT) devices, smart cities,

This work was supported in part by the European Union under the Italian National Recovery and Resilience Plan (PNRR) of NextGenerationEU partnership on "Telecommunications of the Future" (PE00000001 - program "RESTART"), CUP E63C22002040007 - D.D. n.1549 of 11/10/2022, and in part by the Ministry of University and Research (MUR), National Recovery and Resilience Plan (NRRP), Mission 4, Component 2, Investment 1.5, project "RAISE - Robotics and AI for Socio-economic Empowerment" (ECS00000035).

Ali Krayani is with the Department of Electrical, Electronic, Telecommunications Engineering and Naval Architecture, University of Genoa, 16145 Genoa, Italy, and with the Italian National Inter-University Consortium for Telecommunications (CNIT), 43124 Parma, Italy. (e-mails: ali.krayani@ieee.org, ali.krayani@unige.it).

Felix Obite is with the Department of Electrical, Electronic, Telecommunications Engineering and Naval Architecture, University of Genoa, 16145 Genoa, Italy, and also with the School of Electronic Engineering and Computer Science, Queen Mary University of London, London E1 4NS, U.K. (e-mail: felix.obite@edu.unige.it).

Lucio Marcenaro and Carlo Regazzoni are with the Department of Electrical, Electronic, Telecommunications Engineering and Naval Architecture, University of Genoa, 16145 Genoa, Italy, and with the Italian National Inter-University Consortium for Telecommunications (CNIT), 43124 Parma, Italy. (e-mails: lucio.marcenaro@unige.it, carlo.regazzoni@unige.it).

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

Arumugam Nallanathan is with the School of Electronic Engineering and Computer Science, Queen Mary University of London, London E1 4NS, U.K., and also with the Department of Electronic Engineering, Kyung Hee University, Yongin-si, Gyeonggi-do 17104, Korea. (e-mail: a.nallanathan@qmul.ac.uk).

and vehicular technologies, including autonomous driving, as well as the metaverse, necessitates enhanced adaptive resource allocation. Similarly, toward the direction of fully autonomous missions requiring little human intervention, unmanned aerial vehicles (UAVs) are under investigation to operate independently in unknown and dynamic environments due to their flexibility and brilliant line-of-sight links [1]–[4]. In this regard, UAVs must exploit onboard sensors to perceive their surroundings and make intelligent real-time decisions. As a promising innovative multiple-access scheme, non-orthogonal multiple access (NOMA) has the potential to enhance spectrum efficiency, which is shown to outperform orthogonal multiple access in channel capacity [5]–[8]. NOMA utilizes the power domain by employing superposition coding and successive interference cancellation (SIC). It assigns different power levels to serve multiple users simultaneously for multiple access [9], [10]. The effectiveness of NOMA, however, is significantly influenced by the strategies used for resource allocation [11], [12]. The traditional SIC algorithm for NOMA is model-based and relies on knowledge of the underlying statistical model. In the uplink scenario, implementing SIC detection requires the UAV to have accurate knowledge of the channels between each user and the UAV. However, SIC's performance degrades in the presence of realistically imperfect channel state information (CSI) [13]. Alternatively, instead of depending on a model, the detection rule for SIC can be learned directly from data using a data-driven approach without requiring prior knowledge of the underlying channel model [14].

Traditionally, communication systems have focused mainly on accurately reproducing a transmitted message at the receiver end, without considering the meaning or significance of the message itself. Encoding techniques only account for randomness from the source, channel, and destination but do not incorporate any information related to the semantics of the message. Moreover, the transmitter and receiver typically do not utilize their memory or history of previous message transmissions or observed patterns in the data. By allowing the receiver to learn the underlying structure of message transmissions and leverage the context of previously received messages, the robustness of communication against channel and network irregularities can potentially be improved. To take advantage of historical data and learn contextual information, the concept of semantic communications can be utilized [15]. Semantic communication is an approach that aims to transform radio nodes into intelligent agents capable of extracting the underlying meaning or semantics from a dataset [16]–[18]. In

contrast to traditional communication systems, by extracting semantics, the nodes can leverage historical transmissions and learned patterns to optimize resource allocation and robustness. As pointed out by Shannon [19], addressing the communication of data semantics is imperative to maximize communication efficiency.

Finding optimal solutions that effectively capture the dynamic relationships of UAVs and the NOMA scheme, where multiple data symbols are transmitted and detected simultaneously, is extremely challenging due to their nonconvex nature. In such scenarios, the UAV's total flight time is divided into a series of discrete time intervals to simplify the computation. A set of resource allocation variables is optimized for each time interval to achieve optimal performance.

### A. State-of-the-Art

1) *Traditional Optimization-Based UAV-NOMA*: To further enhance system performance, several studies have explored numerical optimization for integrating NOMA and UAV technology [20]–[27]. Furthermore, most traditional optimization algorithms fall into the categories of iterative and heuristic methods [28]. A NOMA system with user quality of service requirements, focusing on enhancing weighted throughput, is analyzed in [29]. An iterative approach is used to derive a nearly optimal solution for maximizing channel selection and power allocation, adopting two users per subchannel. In [30], UAV positions are optimized using a successive convex approximation method to optimize the minimum sum rate of all ground users. Furthermore, [31] presents heuristic-based strategies for user association to maximize the spectral efficiency of a UAV-enabled NOMA system. However, when dealing with resource allocation and trajectory planning, the objective functions and constraints often become correlated and non-convex, posing significant challenges in deriving the optimal solution through conventional means. Converting the optimization problem into iterative or heuristic processes without a structured model can lead to system inefficiency. Consequently, conventional methods fail to provide the online adaptability needed for future dynamic systems.

2) *Machine Learning-Based UAV-NOMA*: To address optimization problems involving UAV trajectories that evolve over time, reinforcement learning (RL) has garnered significant research attention owing to its capability to optimize cumulative long-term benefits [32]. The authors in [33] developed a multi-agent reinforcement learning algorithm for allocating resources across networks with multiple UAVs, demonstrating rapid convergence compared to the conventional Q-learning approach. In the context of a UAV-based NOMA system, the study in [34] proposed an improved RL model optimizing the dynamic maneuver of a UAV with the goal of maximizing the sum rate of randomly moving NOMA users. The author in [35] employed two deep reinforcement learning (DRL) algorithms, one sample-efficient and one distributionally robust, to maximize the sum rate through joint optimization of the UAV location, RIS phase shifts, and uplink power allocation in a UAV-aided NOMA system. In [36], the author concentrates on employing a UAV as an adaptive RL learning agent within

an IoT area with the aim of maximizing the sum rate in a NOMA-UAV network.

Nonetheless, RL cannot deliver optimal performance for certain simple optimization problems where the global optimum can be derived via numerical optimization techniques [37]. Additionally, the exploration mechanisms employed by RL agents, such as epsilon-greedy, compel them to make suboptimal decisions with a small probability, even after convergence has been achieved. This characteristic compromises performance stability in the early stages of the training process [37].

3) *Semantic-Aware Resource Allocation*: With advances in deep learning (DL) and natural language processing (NLP), recent research efforts have been dedicated to semantic-aware resource allocation [38]–[42]. It is worth noting that the aforementioned literature does not integrate UAV-NOMA technology to exploit the full benefits of fully autonomous UAVs and overlooks the causal reasoning underlying data generation. Instead, the focus is limited to computing meaningful attributes that describe the observed data. Consequently, this approach fails to realize the advantages of semantic communication, as these methods lack semantic awareness of how the receiver utilizes the information.

### B. Motivation and Contributions

While numerical optimization algorithms have demonstrated strong performance over the years, their computational complexity poses significant challenges for real-time deployment, particularly in dynamic wireless environments. Likewise, deep learning (DL) models, although effective in data-driven scenarios, often suffer from poor generalization due to their reliance on extensive training data and lack of mechanisms for retaining accumulated knowledge over time [43]. As a result, these models struggle to adapt to novel situations, limiting their practical effectiveness in dynamic, real-world networks.

Semantic communication has emerged as a promising paradigm for 6G and beyond, aiming to enhance transmission efficiency by focusing on the meaning of the data rather than its raw representation. However, its goal-specific nature requires task-dependent customization, making it less suitable as a general-purpose framework for diverse communication tasks. This limitation can hinder its scalability across heterogeneous network applications.

Motivated by these gaps, we introduce a novel semantic-aware resource allocation framework for UAV-based NOMA networks, inspired by the theory of active inference from cognitive neuroscience [44]–[46]. The proposed framework, named Active Semantic Generalized Dynamic Bayesian Network (Active-SGDBN), integrates semantic representation, structured causal reasoning, and Bayesian active inference within a unified architecture. By embedding a virtual semantic layer atop the physical communication model, the system supports interpretable, context-aware decision-making and enables dynamic adaptation to new environments using minimal training data.

In particular, the main contributions of this article are outlined as follows:

- 1) **Semantic-Aware Resource Allocation via Active Inference:** We introduce Active-SGDBN, a data-driven framework that leverages semantic reasoning and Bayesian active inference to enable intelligent and adaptive resource allocation. By embedding semantic understanding into the communication loop, the framework goes beyond traditional optimization and learning methods, enabling UAVs to make context-aware decisions in real time.
- 2) **World Model Construction with Semantic Tokens:** We develop a generative world model by clustering expert-optimized trajectories and power allocations using a Growing Neural Gas (GNG) algorithm. This process captures the joint dynamics of UAV mobility and user association, and encodes them into a structured dictionary of symbolic semantic tokens. These tokens serve as interpretable representations that bridge the physical (distance) and digital (power) domains, facilitating more transparent and modular decision-making.
- 3) **Bayesian Active Inference for Online Policy Optimization:** We formulate a semantic-level active inference mechanism that enables the UAV to infer and refine power allocation strategies during online operation. By comparing predicted semantic symbols with perceived ones, the UAV continuously updates its actions to minimize discrepancies, allowing it to generalize well to new scenarios even with limited training data.
- 4) **Expert-Level Performance and Generalization:** We validate the effectiveness of Active-SGDBN through extensive simulations. Results show that our method matches the performance of exhaustive expert optimization in terms of sum-rate and bit error rate (BER), while significantly outperforming conventional baselines such as random allocation and Q-learning, especially in previously unseen environments.

The rest of this paper is organized as follows: In Section II, we introduce the system model and problem formulation. The proposed solution (Active-SGDBN) is presented in Section III. Section IV provides extensive numerical results to validate the effectiveness of the proposed solution. Finally, Section V concludes the paper.

## II. SYSTEM MODEL AND PROBLEM FORMULATION

### A. System Model

We examine a cellular uplink NOMA network consisting of a UAV-flying base station serving a group of  $M$  mobile ground users within a cell region, as illustrated in Fig. 1. In practical scenarios, a single UAV transmission proves economical for delay-tolerant services like emergency response and periodic environmental monitoring [47]. We model the transmission link with a UAV single receiving antenna and a single user transmitting antenna. This simplification is adopted to reduce system complexity. The UAV's constant flight duration is denoted as  $T$  in seconds (s), and its maximum flight speed is  $V_{max}$ , expressed in meters per second ( $m/s$ ). For ease of exposition, this duration is divided into  $N_t$  time slots, with each time slot having a duration  $\delta_t = T/N_t$ . This division ensures

that the time slots are sufficiently small, allowing the UAV to be constant within each time slot. Additionally, in accordance with the study in [48], it is assumed that the UAV maintains a constant altitude of  $H$  meters throughout its flight. For the sake of simplicity, we employ a three-dimensional Cartesian coordinate system, where the horizontal positions of the  $m$ -th user and the UAV at any given time slot  $n$  are represented by  $\mathbf{q}_m(n) = [x_m(n), y_m(n)]^T, m \in \mathcal{M} = \{1, 2, \dots, M\}$  and  $\mathbf{q}(n) = [x(n), y(n)]^T, n \in \mathcal{N} = \{n \mid 0 \leq n \leq N_t\}$ , respectively. Unlike conventional designs that rely on line-of-sight (LoS) or probabilistic LoS channels, this paper adopts a more practical segregated channel model, as described in [34]. In this model, if obstacles exist in the path between the UAV and user  $m$ , the channel follows a non-line-of-sight (NLoS) model. However, if there are no obstructions, the channel follows a LoS model. Therefore, the power gain  $g$  of the channel between user  $m$  and the UAV in each time slot  $n$  is represented by:

$$g_{m,b}[n] = \mu_{m,b}[n] \xi_{m,b}[n] \beta_{m,b}[n] d_m[n]^{-\alpha_{m,b}[n]}, \quad (1)$$

where  $\mu_{m,b}[n]$  compensates for the small-scale fading,  $\beta_{m,b}[n]$  is the reference power gain of the channel,  $\xi_{m,b}[n]$  is the shadowing component,  $\alpha_{m,b}[n]$  denotes the exponent for path loss,  $d$  is the distance between the UAV and the users, while  $b \in \{\text{LoS}, \text{NLoS}\}$  signifies the strong reliance of the propagation elements on LoS or NLoS situations. After establishing the uplink NOMA, the UAV receives the superimposed QPSK-modulated signal. QPSK modulation is selected for its simplicity and effectiveness in NOMA systems, providing a balanced trade-off between performance and interference.

Let  $B_w$  denote the subchannel bandwidth, at each time slot  $n \in \mathcal{N}$ . We employ the Shannon capacity equation to describe the communication link. Thus, for the uplink, the achievable data rate for user  $m \in \mathcal{M}$  in each UAV time slot  $n \in \mathcal{N}$  is expressed as:

$$R_m = B_w \log_2 \left( 1 + \frac{P_m |g_m|^2}{\sum_{j \neq m}^M P_j |g_j|^2 + N_0} \right), \quad (2)$$

where  $P_m$  and  $g_m$  are the transmit power and channel gain of user  $m$ , respectively, while  $N_0$  denotes the noise power. The UAV decodes the signals sequentially, starting with the user having the strongest channel gain (i.e., user  $\mathcal{M}$ , since it has the highest power level), and moving towards the user with the weakest channel gain. Since the users are ordered based on the strength of their received signals, the UAV begins by decoding the strongest signal first and subtracting it from the combined received signal. Next, the second strongest signal is decoded and subtracted from the combined signal. This process is repeated until all signals have been decoded [49].

The SIC procedure can be described as follows, the UAV decodes user  $\mathcal{M}$  first:

$$R_{\mathcal{M}} = B_w \log_2 \left( 1 + \frac{P_{\mathcal{M}} |g_{\mathcal{M}}|^2}{N_0} \right) \quad (3)$$

The UAV subtracts the signal of user  $\mathcal{M}$  to decode the user  $\mathcal{M}-1$  signal while treating other signals as noise:

$$R_{\mathcal{M}-1} = B_w \log_2 \left( 1 + \frac{P_{\mathcal{M}-1} |g_{\mathcal{M}-1}|^2}{P_{\mathcal{M}} |g_{\mathcal{M}}|^2 + N_0} \right) \quad (4)$$

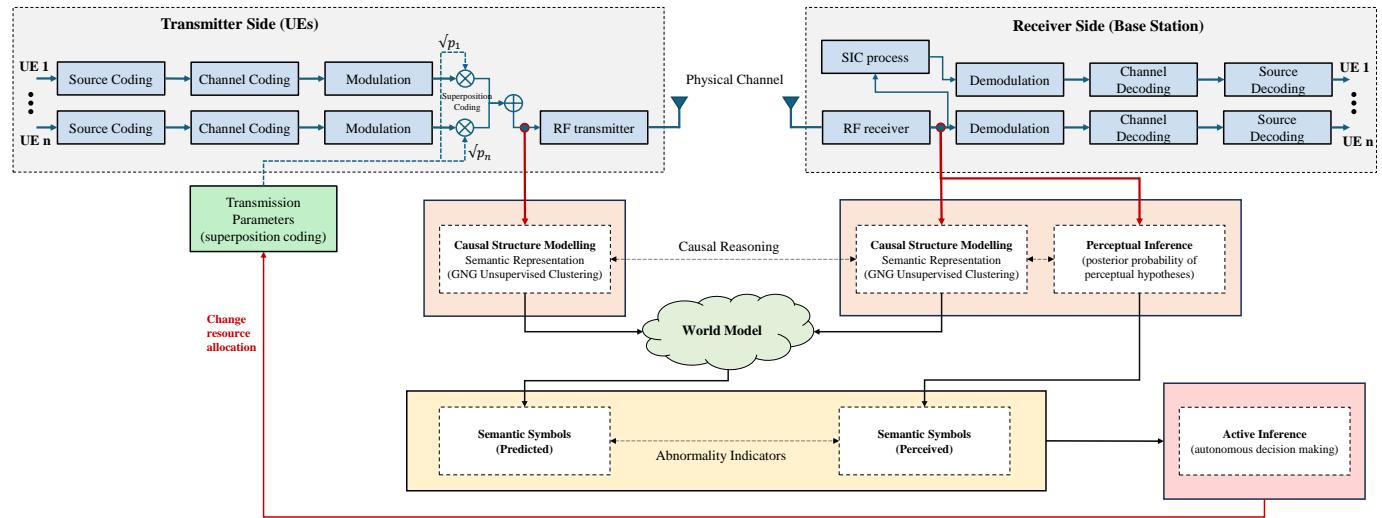


Fig. 1. The main framework of the proposed active inference-based semantic-aware resource allocation (Active-SGDBN).

The process continues iteratively and progressively to decode the user  $m$  signal, which is the weakest signal as defined by the general equation in (2). This iterative decoding procedure, where signals are decoded in a specific order, allows the UAV to maximize the total system sum rate and ensures that interference between users is effectively mitigated. Furthermore, to accurately represent the SIC process, it is necessary to take into account the specific order in which each user's signal is decoded [50], which is defined by a permutation function  $\pi_m : \{1, \dots, |\mathcal{U}_m|\} \rightarrow \mathcal{U}_m$ , with  $|\cdot|$  representing the cardinality of a finite set. This order is usually determined by the received signal power levels, where the strongest signal is decoded first. Nevertheless, the exhaustive search scheme proposed for this study evaluates all likely decoding orders to obtain the optimal configuration that maximizes the sum rate.

### B. Semantic Communication Model

The proposed framework illustrated in Fig. 1 incorporates a conventional communication system based on NOMA, augmented by a semantic virtual system. This virtual system facilitates semantic interaction between the transmitter and the receiver, thereby enhancing resource allocation. In contrast to other semantic communication systems that prioritize transmitted data, our framework transmits binary sequences akin to conventional systems, focusing on conferring semantic significance to the generated/received signals at both the transmitter and the receiver, irrespective of the data type being transmitted. Furthermore, the virtual semantic system analyzes the relationship between the generated/received signals and the distance between the transmitters and the receiver. This analysis provides insights into the expected signals to be transmitted and received based on user proximity to the receivers. The causal structure modeling provides a probabilistic representation of the patterns contained within the generated and received signals, which are identified through unsupervised clustering in semantic symbols. Within this framework, acquired knowledge pertains to the system's acquisition of the causal structure

and semantics associated with joint UAV mobility, user power allocation, and SIC ordering. This knowledge can be encoded in the world model, elucidating the potential evolution of both the physical and digital worlds based on the observed rules during training. Active inference plays a crucial role when the system is operational online, providing insight into how to approach such behaviour. The world model facilitates the generation of anticipated semantic symbols, enabling a comparison with the perceived semantic symbols from the received signal through perceptual inference. This assessment gauges the disparity between the resource allocation strategy and the generation of preferred semantic symbols and corresponding received physical signals.

### C. Problem Formulation

Our goal is to maximize the average sum-rate of all users over the UAV's duration, comprising of  $N$  time slots given by  $\frac{1}{N} \sum_{n=1}^N$ , by optimizing power allocation  $P$  based on the UAV's mobility  $Q$ , which can be expressed as follows:

$$\max_{\{Q, P\}} R_{\text{sum}} \triangleq \frac{1}{N} \sum_{n=1}^N \sum_{m=1}^M R_m(n) \quad (5)$$

$$\text{s.t.} \quad \text{C1: } \sum_{m=1}^M P_m(n) \leq p_{\max}^m, \quad \forall n, \quad (5a)$$

$$\text{C2: } P_m(n) \geq 0, \quad \forall m, \forall n, \quad (5b)$$

$$\text{C3: } Q(0) = Q(I), \quad (5c)$$

$$\text{s.t.} \quad \text{C4: } Q(N) = Q(0), \quad (5d)$$

$$\text{C5: } \|\dot{Q}(n)\| \leq V_{\max}, \quad \forall n \in [0, N], \quad (5e)$$

$$\text{C6: } \Pi(m), \quad \forall n, \quad (5f)$$

$$\text{C7: } |\mathcal{U}_m| \leq L, \quad m \in \mathcal{M}, \quad (5g)$$

where  $Q = \{q(n), \forall n\}$  represents the UAV mobility and  $P = \{P_m(n), \forall m, n\}$  is the power allocation. Constraint C1 ensures that the maximum power budget ( $P_{\max}$ ) is not

exceeded in each time slot. C2 ensures that the allocated power levels are non-negative. Constraint C3 defines the initial position of the UAV, while C4 guarantees that the UAV returns to the initial position after  $N$  time slots. Constraint C5 represents the derivative of the UAV's mobility  $Q(n)$  in each time slot  $n$ , and  $V_{max}$  represents the UAV's maximum velocity, which imposes that the UAV does not exceed its maximum speed. Constraint C6 guarantees the optimal uplink SIC decoding order. Because of the practical constraints of SIC and the complexity of decoding, the maximum number of users that can be multiplexed on each sub-channel is  $L$ , which is constraint C7.

#### D. Exhaustive Search Optimization

It is important to note that the optimization problem described in (5) is classified as NP-hard. This is because the UAV lacks prior knowledge of users' positions or CSI for future time slots due to the users' random mobility. Additionally, the scenario involving a mobile UAV further complicates the problem, making it difficult to solve using traditional numerical optimization algorithms. To address this challenge, we employ exhaustive search optimization to generate the training dataset, using it as expert domain knowledge. This strategy finds the global optimal solution by computing all possible solutions for UAV mobility and power allocation. While we acknowledge that the exhaustive search strategy is computationally intensive for generating the simulated dataset, it is a one-time cost that serves as a robust and accurate upper bound for validating our proposed solution. The Active-SGDBN agent leverages this data to learn an optimal strategy, producing a near-optimal solution without requiring exhaustive search computations during online deployment.

Choosing the variables for linear regression models is a major challenge for statistics [51]. Suppose we have  $N$  expanding variables. The easiest approach for choosing variables is by exhaustively searching all possible combinations, which involves estimating the combinations of all variables, given by  $2^N - 1 = {}_N C_1 + {}_N C_2 + \dots + {}_N C_K + \dots + {}_N C_N$ . This approach is called the exhaustive search [52]. The authors in [53] assert that any precise technique for selecting variables comes with a computational overhead of  $O(2^N)$ , which is applicable to the exhaustive search method [51].

For clarity and consistency, Table I summarizes the notation used throughout this paper.

### III. PROPOSED METHOD

In this section, we describe a data-driven approach to solving the optimization problem in equation (5), which we call Active-SGDBN. Interestingly, active inference is a generalization of Bayesian inference, where the goal is not just to infer the hidden or latent states that generate sensory data but to actively choose actions that will minimize future uncertainty or surprise. In information theory, this expected surprise is equivalent to entropy or uncertainty. Therefore, optimal behavior can be defined as acting in a way that resolves or reduces uncertainty [45].

TABLE I  
NOTATION SUMMARY

Symbol	Meaning
<i>Sets, indices, counts</i>	
$m, M$	User index; number of multiplexed users
$n, N$	Slot index; number of UAV slots (in Problem (5))
$t$	World-model (dictionary) time index
$k, K$	Token index; <b>number of tokens (dictionary size)</b>
$U$	Number of trajectory letters (clusters)
$F, F'$	Number of distance-letter / power-letter clusters
$L, L'$	Number of distance words / power words
$\mathcal{E} = \{E_n\}$	Training episodes (examples)
$D_{n,t}$	Distance set at $(n, t)$ : $\{d_{t,1}^{(n)}, \dots, d_{t,M}^{(n)}\}$
$Y_{n,t}$	Optimizer power solution at $(n, t)$ : $[p_{t,1}, \dots, p_{t,M}]$
<i>Geometry, timing, mobility</i>	
$T, \delta_t, N_t$	Flight duration, slot size $T/N_t$ , number of slots
$V_{max}, H$	UAV maximum speed; flight altitude
$q(n), q_m(n)$	UAV / user- $m$ 2D position at slot $n$
$d_m[n]$	UAV-user $m$ distance at slot $n$
$Q = \{q(n)\}$	UAV trajectory (sequence of positions)
<i>Channel and bandwidth</i>	
$g_{m,b}[n]$	Channel power gain ( $b \in \{\text{LoS, NLoS}\}$ )
$\mu_{m,b}[n], \xi_{m,b}[n]$	Small-scale fading; shadowing term
$\beta_{m,b}[n], \alpha_{m,b}[n]$	Reference gain; path-loss exponent
$B_w, N_0$	Sub-channel bandwidth; noise power
<i>NOMA uplink, rates, SIC</i>	
$P_m(n), P_{\max}$	User- $m$ transmit power at slot $n$ ; per-slot budget
$R_m(n), R_{\sum}$	User rate; average sum-rate objective
$\pi_m(\cdot)$	SIC decoding-order permutation on a resource
$U_m,  U_m , L$	Set, cardinality, and max users per resource (C7)
$\Pi(m)$	Optimal uplink SIC order (C6)
<i>Semantic world model (distance <math>d</math>, power <math>p</math>, trajectory <math>u</math>)</i>	
$l_{t,m}^d, l_{t,m}^p$	Distance/power letters (cluster indices)
$x_{t,m}^d, x_{t,m}^p$	Continuous latent states (distance / power)
$z_{t,m}^d, z_{t,m}^p$	Observations (distance / received-signal features)
$W_t^d, W_t^p$	Words: joint distance / power configs across users
$T_t, \mathcal{T}$	Token at time $t$ ; token set (size $K$ )
$\mathcal{L}^d, \mathcal{L}^p$	Sets of distance / power letters
$\mathcal{W}^d, \mathcal{W}^p$	Sets of distance / power words
$l_j^u$	Trajectory letter (cluster) index $j \in \{1, \dots, U\}$
<i>Probabilistic messages and decisions</i>	
$\pi(\cdot), \lambda(\cdot)$	Prior/belief (top-down) and likelihood (bottom-up)
$P(\cdot), \alpha$	Probability (mass/density); normalization constant
$D_{KL}[\cdot    \cdot]$	Kullback–Leibler divergence
$\Upsilon_t^{WP}$	Power-word semantic abnormality at time $t$
$a_t, a_t^*$	Action; abnormality-minimizing action
<i>Transitions, coupling, and optimization</i>	
$\Pi_T$	Token transition matrix $[P(T_k   T_{k'})]$
$\Pi_{QT}$	Trajectory-token coupling $[P(T_k   l_j^u)]$
$Q, P$	Decision variables: trajectory and power allocation (Prob. (5))
$C1-C7$	System/mobility/SIC/multiplexing constraints
$\dot{Q}(n)$	Slot-wise velocity used in the speed bound (C5)
<i>Evaluation shortcuts (Sec. IV-E/F)</i>	
$R_{\text{OPT}}, R_{\text{ACT}}$	Optimizer vs Active-SGDBN sum-rate
$\text{BER}_{\text{OPT}}, \text{BER}_{\text{ACT}}$	Optimizer vs Active-SGDBN BER

The proposed approach comprises two stages: 1) Learning the World Model during offline training, a perceptual learning process where the UAV learns generative models of word tokens representing both the radio and the physical environments, capturing the complex relationships and dependencies of the joint objective; and 2) an online Active-SGDBN stage that employs a Bayesian active inference for optimization of the power allocation based on UAV mobility. Fig. 2, presents the graphical representations of the proposed Active-SGDBN model.

### A. Offline Perception: Learning the World Model

Inspired by the discrete nature of words in NLP [54], which enables agents to communicate via discrete tokens, the framework involves an offline perception stage, as depicted in Fig. 2-(a). In this stage, we introduce a higher layer of intelligence into a hierarchical GDBN network in semantic spaces. The network captures the spatial relationships and temporal evolution of the UAV and NOMA users in both discrete and continuous states. In each UAV's time slot, a specific resource allocation must be optimized. Therefore, each letter represents a specific UAV position, power allocation, and SIC ordering. A group of letters forms a word, representing the resource allocation for the current state and time slot. A dictionary, which represents a group of words (referred to as word tokens), captures a high level of interaction and abstraction of the joint objective function by mapping words to users' QPSK signal constellations over time. The graphical model showcases how edges represent the dependence of discrete states on continuous generalized states. The graph is characterized by constant neuronal message passing, enabling the agent to approximate the posterior over hidden states. These messages consist of descending signals from prior states (blue arrows) and ascending signals from future states (red arrows). Essentially, both past and future environmental states are continuously captured over time, thus offering optimal online approximate posterior estimates.

The GNG algorithm is a powerful tool for learning the topological structure of data. It employs biologically-inspired methods for approximating signals. As a result, its incremental learning approach makes it well-suited for temporal clustering [55]. During the initial learning stage, we provide the UAV with an initial model based on the Unscented Kalman Filter (UKF) [56], which assumes that the external states evolve according to fixed rules for predicting the continuous states. Consequently, based on the received observation, the UAV's internal model generates initial prediction errors called generalized errors (GEs) [57]. These generalized errors are used for incrementally learning and updating new models that better capture the dynamics of the environment. The GNG-trained network produces clusters of nodes in both discrete and continuous generalized states, where a node represents a particular configuration of the UAV's state. These clusters create a generative model that encapsulates the complex dependencies among the UAV's trajectory, power levels, and the SIC decoding order.

### B. Dictionary Formation

The dictionary learning process (refer to Fig. 4) entails the formation of words that encode all possible combinations explored by the exhaustive search scheme for the UAV trajectory and power allocation. This process evaluates the sum rate for each configuration and generates a global dictionary that includes a word for all possible configurations explored in the search solution space. Each word in the dictionary represents a specific set of UAV positions and power allocations across all time steps. Subsequently, each word in the dictionary is mapped to a QPSK constellation, enabling

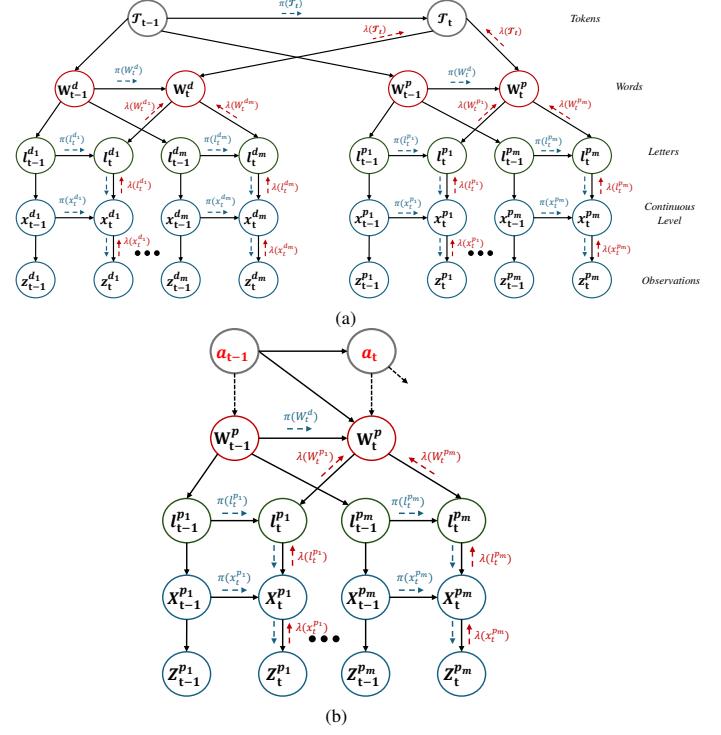


Fig. 2. Graphical representations of the proposed solution: (a) SGDBN, represents the offline stage (UAV's perception). (b) Active-SGDBN, the UAV's online active inference model.

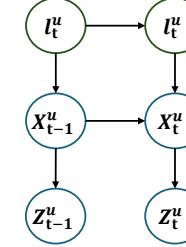


Fig. 3. The GDBN representing the UAV trajectory.

each configuration to be characterized as a cluster of QPSK symbols. The result is a comprehensive dictionary representing the complete set of possible configurations for a UAV-NOMA network within the defined search space. It is important to note that the GNG clustering is performed solely during the offline dictionary formation stage, and thus does not contribute to the online decision-making latency. The resulting online processing involves only token-based inference, which has negligible computational overhead.

Let  $\mathcal{E} = \{E_1, E_2, \dots, E_N\}$  represent a training set containing  $N$  examples of users' distribution in the cell, where  $E_n$  denotes the  $n$ -th example consisting of a set  $\mathcal{D}_{n,t} = \{d_{t,1}^n, d_{t,2}^n, \dots, d_{t,M}^n\}$  that represents the dynamic distances at each time instant  $t$  between the mobile UAV and the  $M$  users. These users are multiplexed on a specific sub-channel using NOMA. An exhaustive search optimizer is employed offline to solve the  $N$  examples in  $\mathcal{E}$ . This will produce a set,  $\mathcal{Y} = \{Y_{1,t}, Y_{2,t}, \dots, Y_{N,t}\}$ , representing the resulting demonstrations (or solutions). Here,  $Y_{n,t} = [p_{t,1}, p_{t,2}, \dots, p_{t,M}]$

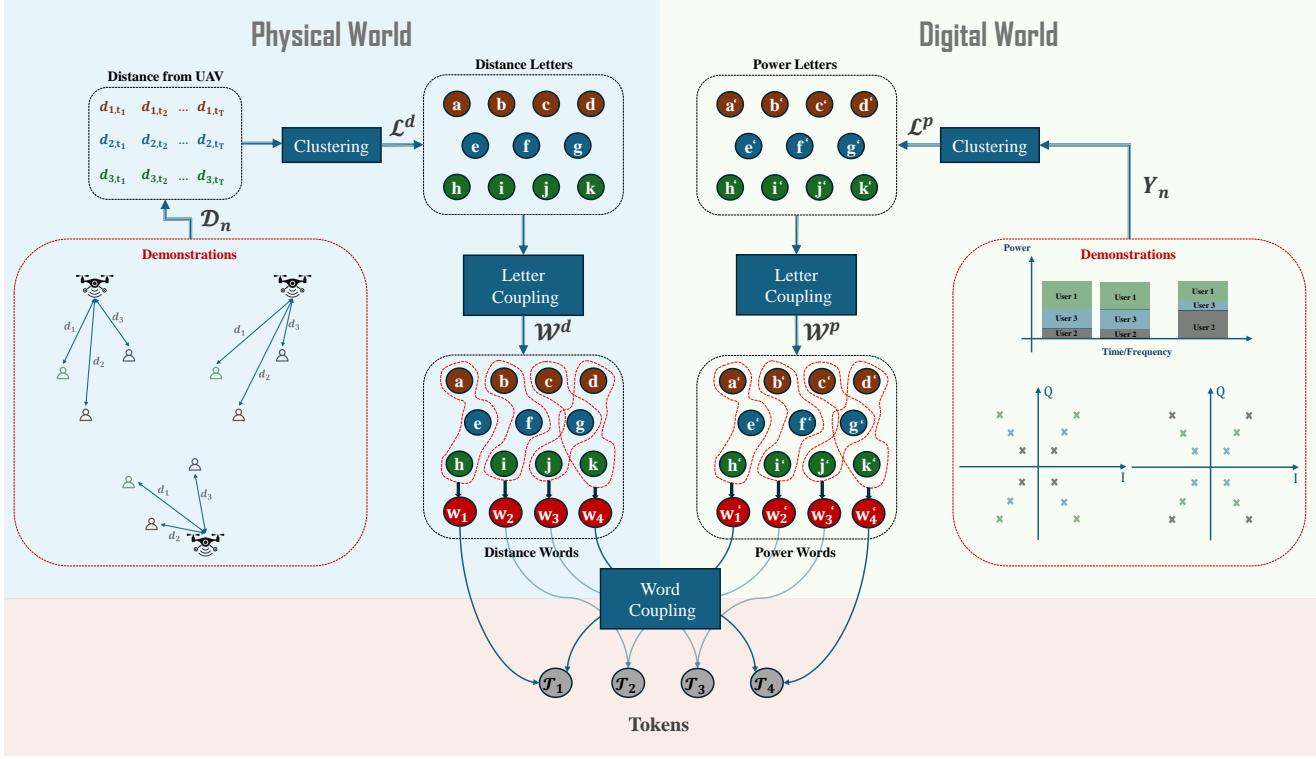


Fig. 4. A schematic representing the process of building the dictionary for the world model.

denotes the allocated power values for each example in  $\mathcal{E}$  at each time instant  $t$ . After the solutions are generated in  $\mathcal{Y}$  by the optimizer, a growing neural gas (GNG) is employed to explore the cluster patterns in the solutions.

Let  $\mathcal{L}^{d_m} = \{l_1^{d_m}, l_2^{d_m}, \dots, l_F^{d_m}\}$  be the set of clusters that correspond to the dynamic distances between the UAV and the  $m$ -th ground user. These clusters, referred to as letters, are generated by GNG when given input  $\mathcal{D}_n$ . Here  $l_f^{d_m}$  denotes the  $f$ -th letter in  $\mathcal{L}^{d_m}$  following a Gaussian distribution such that  $l_1^{d_m} \sim \mathcal{N}(\mu_{l_f^{d_m}}, \Sigma_{l_f^{d_m}})$ , and  $F$  is the total count of letters. Similarly, the set  $\mathcal{L}^{p_m} = \{l_1^{p_m}, l_2^{p_m}, \dots, l_F^{p_m}\}$  comprises letters produced by GNG using input  $\mathcal{Y}$  that represent the power values allocated to the  $m$ -th ground user based on its distance from the UAV. Each  $l_{f'}^{p_m} \in \mathcal{L}^{p_m}$  follows a Gaussian distribution such that  $l_{f'}^{p_m} \sim \mathcal{N}(\mu_{l_{f'}^{p_m}}, \Sigma_{l_{f'}^{p_m}})$ . This clustering process is repeated for each of the  $M$  users, resulting in two global sets  $\mathcal{L}^d = \{\mathcal{L}^{d_1}, \mathcal{L}^{d_2}, \dots, \mathcal{L}^{d_M}\}$  and  $\mathcal{L}^p = \{\mathcal{L}^{p_1}, \mathcal{L}^{p_2}, \dots, \mathcal{L}^{p_M}\}$ .

The letters representing the dynamic distances between the UAV and the  $M$  users are grouped together to form a word  $W_t^d = [l_1^{d_1}, l_2^{d_2}, \dots, l_M^{d_M}]$ . This word is synchronized with the word  $W_t^p = [l_{f_1}^{p_1}, l_{f_2}^{p_2}, \dots, l_{f_M}^{p_M}]$  that couples the letters of  $M$  users representing the allocated power values. In this way, the UAV, acting as a flying base station, treats the  $M$  users as independent entities that can be combined to form a specific physical configuration in the real world, which can then be linked with the power values representing the digital configuration in the digital world. Let  $\mathcal{W}^d = \{W_1^d, W_2^d, \dots, W_L^d\}$  represent the set of words denoting distances between the  $M$  users and the UAV, where  $L$  is the total number of distance words. Similarly, let  $\mathcal{W}^p = \{W_1^p, W_2^p, \dots, W_{L'}^p\}$  denote the

set of words representing power values allocated to the  $M$  users, where  $L'$  is the total number of power words. Therefore, distance words indicate how users are distributed on the ground and their proximity to the UAV, while power words represent the power values that the UAV must allocate based on the users' distribution.

To integrate both distance words and power words, we propose the introduction of the token set  $\mathcal{T} = \{\mathcal{T}_1, \mathcal{T}_2, \dots, \mathcal{T}_K\}$ , where  $\mathcal{T}_k = [W_k^d, W_k^p]$  and  $K$  is the total number of tokens. The evolution of the tokens can be tracked using the transition matrix, which is defined as:

$$\Pi_{\mathcal{T}} = \begin{bmatrix} P(\mathcal{T}_1|\mathcal{T}_1) & P(\mathcal{T}_1|\mathcal{T}_2) & \dots & P(\mathcal{T}_1|\mathcal{T}_K) \\ P(\mathcal{T}_2|\mathcal{T}_1) & P(\mathcal{T}_2|\mathcal{T}_2) & \dots & P(\mathcal{T}_2|\mathcal{T}_K) \\ \vdots & \vdots & \vdots & \vdots \\ P(\mathcal{T}_K|\mathcal{T}_1) & P(\mathcal{T}_K|\mathcal{T}_2) & \dots & P(\mathcal{T}_K|\mathcal{T}_K) \end{bmatrix}, \quad (6)$$

where  $0 \leq P(\mathcal{T}_k|\mathcal{T}_{k-1}) \leq 1$  and  $\sum_{k=1}^K P(\mathcal{T}_k|\mathcal{T}_{k-1}) = 1, \forall k$ .

The tokens represent a set of configurations of users' distribution and the allocated power values, linking the physical and digital worlds. The acquired dictionary includes tokens, words, letters, and measurements (i.e., real-time signals) and can be organized in a chain starting with tokens and ending with measurements. Therefore, the dictionary can be structured in a hierarchical representation that models the dynamic behaviour of the physical and digital worlds, incorporating Markov chains at both hierarchical and temporal levels.

From the chain rule of probability theory and using the conditional independencies encoded in the chain, the joint distribution function representing the dictionary can be expressed

as:

$$\begin{aligned} P(\mathcal{T}_t, W_t^d, W_t^p, l_t^{d_m}, l_t^{p_m}, x_t^{d_m}, x_t^{p_m}, z_t^{d_m}, z_t^{p_m}) = \\ \prod_{t=1}^T P(\mathcal{T}_t | \mathcal{T}_{t-1}) P(W_t^{d_m} | W_{t-1}^{d_m}, \mathcal{T}_t) P(W_t^{p_m} | W_{t-1}^{p_m}, \mathcal{T}_t) \\ P(l_t^{d_m} | l_{t-1}^{d_m}, W_t^{d_m}) P(l_t^{p_m} | l_{t-1}^{p_m}, W_t^{p_m}) P(x_t^{d_m} | x_{t-1}^{d_m}, l_t^{d_m}) \\ P(x_t^{p_m} | x_{t-1}^{p_m}, l_t^{p_m}) P(z_t^{d_m} | x_t^{d_m}) P(z_t^{p_m} | x_t^{p_m}). \quad (7) \end{aligned}$$

The proposed Generalized Dynamic Bayesian Network (GDBN) constitutes a comprehensive probabilistic model for the variables within the acquired dictionary, delineating joint distributions over these variables. This model encompasses all necessary information to address various probabilistic queries related to the specified variables. Such queries may include the interpretation of specific input data or, in the presence of utility information, the recommendation of the optimal course of action. Interpretation requires instantiating a set of variables that correspond to the input data, assessing their impact on the probabilities of designated hypotheses, and ultimately selecting the most probable combinations of these hypotheses. Furthermore, the GDBN can be regarded as a computational infrastructure for reasoning about the acquired knowledge base. In this capacity, the links within the network act as the primary mechanisms that direct and facilitate data flow throughout the processes of querying and updating beliefs. This distributed computational paradigm effectively leverages the independencies represented in the network, allowing for subtask decomposition and significantly reducing computational complexity. Consequently, any given variable within the network is able to compute its own distribution independently, without necessitating interaction with variables beyond its immediate neighborhood.

On the other hand, when considering UAV trajectories, let  $\mathcal{Q} = \{\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_N\}$  represent the set of  $N$  trajectories generated from the  $N$  examples mentioned earlier. This set of trajectories is used as input for the GNG algorithm to cluster them, resulting in a set of letters  $\mathcal{L}^u = \{l_1^u, l_2^u, \dots, l_U^u\}$ , where  $U$  is the total number of trajectory letters. The GDBN depicted in Fig. 3 organizes these letters to represent the discrete regions where the UAV may be located while also structuring the continuous variables that describe the UAV's specific positions within those regions. When the UAV begins to move, it predicts how the surrounding environment—both physical and digital—might change in response to its motion. This is achieved by learning an interactive transition matrix that encodes the probabilistic relationships between the UAV's actions and the world's tokens. This matrix helps explain how the world will evolve as the UAV moves. Such predictions facilitate the action selection process, allowing for sub-optimal power allocation and enabling adjustments to these strategies using active inference to achieve desired performance through a sub-ideal digital configuration. The interactive matrix is defined as follow:

$$\Pi_{\mathcal{Q}\mathcal{T}} = \begin{bmatrix} P(\mathcal{T}_1 | l_1^u) & P(\mathcal{T}_2 | l_1^u) & \dots & P(\mathcal{T}_K | l_1^u) \\ P(\mathcal{T}_1 | l_2^u) & P(\mathcal{T}_2 | l_2^u) & \dots & P(\mathcal{T}_K | l_2^u) \\ \vdots & \vdots & \vdots & \vdots \\ P(\mathcal{T}_1 | l_U^u) & P(\mathcal{T}_2 | l_U^u) & \dots & P(\mathcal{T}_K | l_U^u) \end{bmatrix}, \quad (8)$$

where  $0 \leq P(\mathcal{T}_k | l_j^u) \leq 1$  and  $\sum_{k=1}^K P(\mathcal{T}_k | l_j^u) = 1, \forall j, \forall k$ .

### C. Active decision-making enhanced with generative world model

The UAV incorporates a generative model, represented by a world model, where it seeks to gather the most evidence by making inferences from sensory data actively sampled from the environment. The UAV starts to move according to the actions recommended by the model (as illustrated in Fig. 3), which defines a specific trajectory behavior, such as circular motion or random walking. By choosing actions that influence its future movements, the UAV can anticipate the next letters (states) it may reach. This prediction of the UAV's letters is carried out by using the transition matrix, which encodes the probabilistic relationships between the UAV's letters according to  $P(l_t^u | l_{t-1}^u)$ . UAV's goal is to identify the most effective allocation strategy while functioning in real-time. It accomplishes this by predicting the tokens it may encounter based on its movement, which is represented by the UAV's letter ( $l_t^u$ ). By anticipating these tokens, the UAV can forecast the corresponding distance word related to its physical configuration and also predict the preferred power word connected to the digital configuration it is likely to achieve.

The UAV can anticipate the next expected token based on the predicted letter representing its subsequent state, denoted as  $l_t^u$ . To accomplish this, the UAV employs a Hierarchical Particle Filter (HPF) to propagate a set of equiprobable tokens, utilizing the token transition matrix defined in equation (8) as a proposal. The posterior probability associated with the predicted tokens is determined by:

$$\pi(\mathcal{T}_t) = P(\mathcal{T}_t | \mathcal{T}_{t-1}, l_{t-1}^u). \quad (9)$$

Thus, the prediction of the token encompasses both the anticipated distance word and the corresponding power word, which are associated with the relevant posterior probabilities:

$$\pi(W_t^{d_m}) = P(W_t^{d_m} | W_{t-1}^{d_m}, \mathcal{T}_t), \quad (10)$$

and,

$$\pi(W_t^{p_m}) = P(W_t^{p_m} | W_{t-1}^{p_m}, \mathcal{T}_t). \quad (11)$$

Consequently, the evolution of the physical distance between the UAV and the ground users can be predicted based on the anticipated words. The predicted letter, which denotes a discrete variable, is associated with the region where the distance between the UAV and user  $m$  may exist and is linked to the following posterior probability:

$$\pi(l_t^{d_m}) = P(l_t^{d_m} | l_{t-1}^{d_m}, W_t^{d_m}). \quad (12)$$

Furthermore, the predicted continuous variable representing the physical distance is also associated with the following posterior probability:

$$\pi(x_t^{d_m}) = P(x_t^{d_m}, l_t^{d_m} | z_t^{d_m}) = \int P(x_t^{d_m} | x_{t-1}^{d_m}, l_t^{d_m}) \lambda(x_{t-1}^{d_m}), \quad (13)$$

where,  $\lambda(x_{t-1}^{d_m}) = P(z_{t-1}^{d_m} | x_{t-1}^{d_m})$ .

Likewise, the evolution of the digital world, as indicated by the digital signals transmitted by ground users in accordance with their allocated power values, can be predicted using anticipated distance words. Consequently, the predicted discrete variables, which signify the regions where the potential power values of these digital signals may reside, are characterized by a specific probability distribution:

$$\pi(l_t^{p_m}) = P(l_t^{p_m} | l_{t-1}^{p_m}, W_t^{p_m}). \quad (14)$$

Additionally, the predicted continuous variables representing the physical signals' power are delineated by the following posterior distribution:

$$\pi(x_t^{p_m}) = P(x_t^{p_m}, l_t^{p_m} | z_t^{p_m}) = \int P(x_t^{p_m} | x_{t-1}^{p_m}, l_t^{p_m}) \lambda(x_{t-1}^{p_m}), \quad (15)$$

where,  $\lambda(x_{t-1}^{p_m}) = P(z_{t-1}^{p_m} | x_{t-1}^{p_m})$ .

Top-down predictive messages are compared with bottom-up diagnostic messages to calculate the differences between them. This comparison helps to adjust actions taken and update the internal representation of the world. Actual sensory input is utilized to generate the diagnostic messages that are then propagated upward in the hierarchy.

Consider a specific network in the left branch that represents the physical world in Fig. 2. This network consists of four nodes arranged in a chain:  $W_t^{d_m} \rightarrow l_t^{d_m} \rightarrow x_t^{d_m} \rightarrow z_t^{d_m}$ . If evidence  $z_t^{d_m}$  is observed, then according to Bayes' rule, the belief distribution of  $x_t^{d_m}$  can be expressed as:

$$\pi(x_t^{d_m}) = \alpha \pi(x_t^{d_m}) \lambda(x_t^{d_m}), \quad (16)$$

where,

$$\lambda(x_t^{d_m}) = P(z_t^{d_m} | x_t^{d_m}) \quad (17)$$

represents the diagnostic message that indicates the likelihood computed at  $z_t^{d_m}$ , which is transmitted as a message to  $x_t^{d_m}$ , enabling  $x_t^{d_m}$  to compute its belief distribution as defined in (20). In contrast, since  $x$  is not observed directly but is supported by indirect observation  $z$  of a descendant  $z$  of  $x$ , the likelihood  $\lambda(l_t^{d_m})$  can no longer be directly obtained from  $P(x_t^{d_m} | l_t^{d_m})$ , however, must reflect the  $P(z_t^{d_m} | x_t^{d_m})$  as well. Thus, conditioning on  $x_t^{d_m}$ , we can write:

$$\lambda(l_t^{d_m}) = P(z_t^{d_m} | x_t^{d_m}) P(x_t^{d_m} | l_t^{d_m}) = \lambda(x_t^{d_m}) P(x_t^{d_m} | l_t^{d_m}). \quad (18)$$

Using the fact that  $x_t^{d_m}$  separates  $l_t^{d_m}$  and  $z_t^{d_m}$ . Likewise, the diagnostic message computed at each letter  $l$  and transmitted to  $W$  can be computed as follows:

$$\lambda(W_t^d) = \prod_{m=1}^M \lambda(W_t^{d_m}) = P(z_t^{d_m} | x_t^{d_m}) P(x_t^{d_m} | l_t^{d_m}) \times P(l_t^{d_m} | W_t^{d_m}) = \lambda(l_t^{d_m}) P(l_t^{d_m} | W_t^{d_m}). \quad (19)$$

On the other hand, the right branch representing the digital world in Fig. 2, depicts the network consisting of four nodes arranged in a chain :  $W_t^{p_m} \rightarrow l_t^{p_m} \rightarrow x_t^{p_m} \rightarrow z_t^{p_m}$ . Thus, when evidence  $z_t^{p_m}$  is observed after performing the SIC process of the combined signal received by the UAV, that can split the combined signal to get the corresponding signals ( $z_t^{p_1}, \dots$ ,

$z_t^{p_m})$  of the users multiplexed on the same channel, the belief distribution of  $x_t^{p_m}$  can be calculated according to :

$$\pi(x_t^{p_m}) = \alpha \pi(x_t^{p_m}) \lambda(x_t^{p_m}), \quad (20)$$

where,

$$\lambda(x_t^{p_m}) = P(z_t^{p_m} | x_t^{p_m}), \quad (21)$$

represents the diagnostic message indicating the likelihood computed at  $z_t^{p_m}$  and transmitted as a message to the upper level, enabling  $x_t^{p_m}$  to compute its belief distribution. Consequently, the likelihood computed at  $x_t^{p_m}$  is computed and then transmitted to the upper level in the hierarchy as follows:

$$\lambda(l_t^{p_m}) = P(z_t^{p_m} | x_t^{p_m}) P(x_t^{p_m} | l_t^{p_m}) = \lambda(x_t^{p_m}) P(x_t^{p_m} | l_t^{p_m}). \quad (22)$$

After that, the diagnostic message computed at each letter and transmitted to the higher word level can be expressed as:

$$\lambda(W_t^p) = \prod_{m=1}^M \lambda(W_t^{p_m}) = P(z_t^{p_m} | x_t^{p_m}) P(x_t^{p_m} | l_t^{p_m}) \times P(l_t^{p_m} | W_t^{p_m}) = \lambda(l_t^{p_m}) P(l_t^{p_m} | W_t^{p_m}). \quad (23)$$

Finally, the diagnostic message transmitted to the Tokens level at the top of the hierarchy can be written as:

$$\lambda(\mathcal{T}_t) = \lambda(W_t^d) \lambda(W_t^p) P(W_t^d | \mathcal{T}_t) P(W_t^p | \mathcal{T}_t). \quad (24)$$

1) *Semantic Abnormality indicators:* To evaluate whether the actions taken have led to the desired semantic states, the UAV uses an abnormality indicator to calculate the similarities between the predicted desired semantic symbols and the perceived semantic symbols at different hierarchical levels. The world model comprises two branches: one representing the physical world and the other representing the digital world. Our focus will be on the abnormality indicator at the world level of both branches. The abnormality indicator in the physical branch assesses how much the current physical configuration deviates from the predicted one based on the UAV's motion. In contrast, the abnormality indicator in the digital branch evaluates how far the current digital configuration, after performing specific actions, is from the preferred (predicted) digital configuration. In this work, we will concentrate solely on actions taken in the digital world, leaving the integration of actions in both the digital and physical worlds through trajectory planning for future research. Therefore, we will emphasize the abnormality indicators at the word level of the digital branch to determine whether the allocated power values were appropriate or not.

The semantic abnormality indicator at the word level of the digital world is defined as follows:

$$\Upsilon_{W_t^p} = D_{KL}[\pi(W_t^p) || \lambda(W_t^p)] = \sum_W \pi(W_t^p) \log \left( \frac{\pi(W_t^p)}{\lambda(W_t^p)} \right), \quad (25)$$

where  $D_{KL}$  is the Kullback-Leibler (KL) divergence [58] that can be employed to assess the similarity between two probability density functions.

The objective while operating online is to minimize the Kullback-Leibler (KL) divergence between the predicted semantic symbol for the preferred word and the perceived

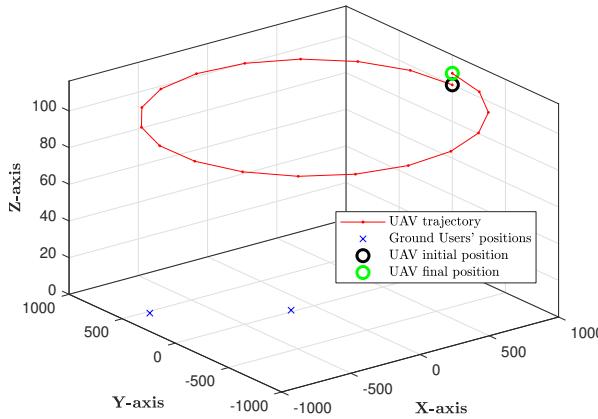


Fig. 5. An example of the system model containing two ground users.

semantic symbol for the generated word resulting from the actions performed, which is defined as follows:

$$a_t^* = \arg \min_{a_t} \Upsilon_{W_t^P}. \quad (26)$$

#### IV. SIMULATION RESULTS

In this section, we present the simulation results to evaluate the performance of the proposed framework. First, we outline the default system parameters used in the simulations. Next, we demonstrate the numerical results of the dictionary formation process discussed in Section III-B. Following that, we assess the effectiveness of the proposed method in adapting to new environmental conditions, with the goal of maximizing the sum-rate and minimizing bit error rates, supported by additional numerical results. Finally, we highlight the performance gains attained through active decision-making compared to various benchmarks.

##### A. Simulation Setup

TABLE II  
SIMULATION PARAMETERS

Parameter	Value
Cell radius	1000 m
UAV flight height ( $H$ )	100 m
UAV speed	5 m/s
Number of UAV time steps	10
Modulation scheme	QPSK
Channel model	LoS/NLoS [34]
SNR	[-30 dB, ..., +30 dB]
System Bandwidth, $B_w$	1.4 MHz
Number of sub-channels	6
Power budget of SUs $P_{max}$	20 W [30]
Maximum number of Ground Users per sub-channel $M$	$M = [2, 3, 4]$ , 0.01

Our study examines a cell area that covers a square region of  $1000 \times 1000$  m<sup>2</sup>, where ground users are uniformly distributed throughout the area. The UAV operates at a fixed altitude during the simulations, following a circular trajectory over the cell (an example with two ground users is depicted in Fig. 5). The ground users are categorized into near and far users in relation to the mobile UAV, which changes over time as the UAV moves across the area. This variability is due to the

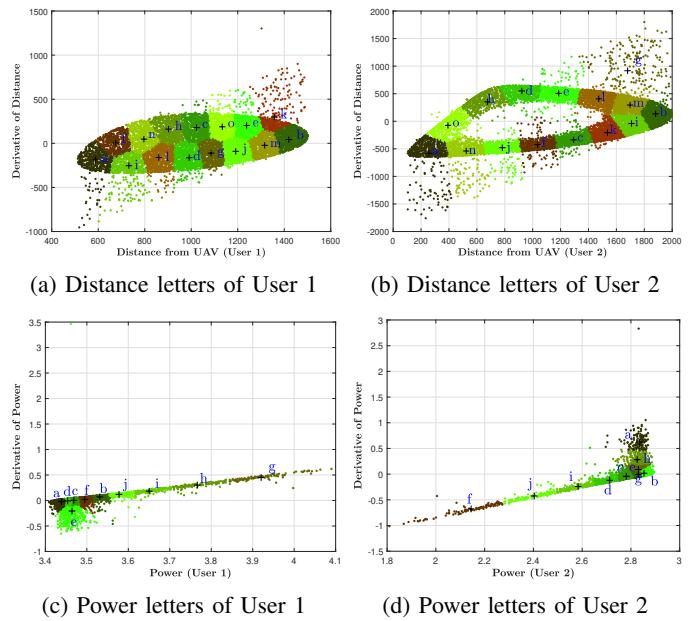


Fig. 6. The unsupervised clustering process involves analyzing 1000 solutions generated by the Optimizer. This results in a set of letters that represent both the physical and digital configurations, indicating the distance between UAV and ground users, as well as the power allocations determined by the optimizer.

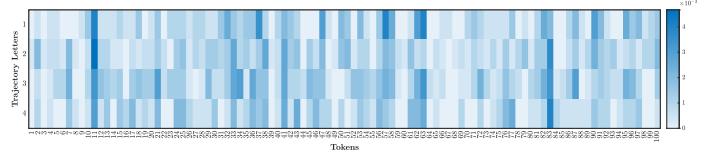


Fig. 7. An interactive transition matrix depicting the probabilistic relationships between letters denoting UAV trajectories (for simplicity, showing four distinct letters) and the acquired tokens (for simplicity, showing one hundred tokens).

UAV's mobility, meaning that a user classified as near at one moment may become a far user at a later time.

To prepare the training data, we generated numerous examples of the UAV flying over the cell populated with ground users. These examples were used as input for an optimizer (exhaustive search) in an offline process to determine the best power allocation strategies, which ensure a high sum-rate and minimum bit error rates. The resulting demonstrations (or solutions) from this optimization are then utilized to learn the dictionary, forming the world model as discussed in Section III. For testing purposes, we generated examples that differ from those used in training to evaluate the performance of the proposed method. A summary of the system parameters is provided in Table II.

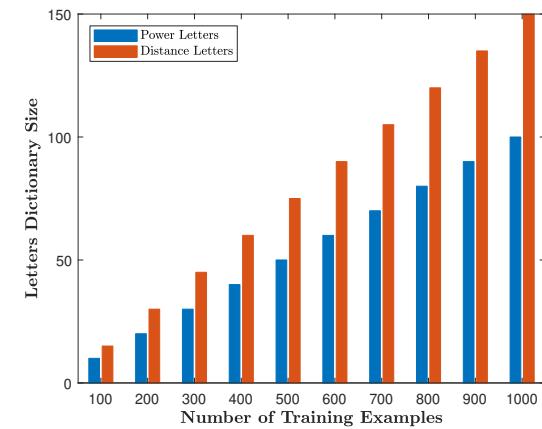
##### B. Dictionary Learning

The demonstrations of dynamically allocated power values, obtained through exhaustive search after analyzing training examples of the varying distances between mobile UAVs and ground users, are utilized to create a dictionary representing the world model. These dynamic distances and their cor-

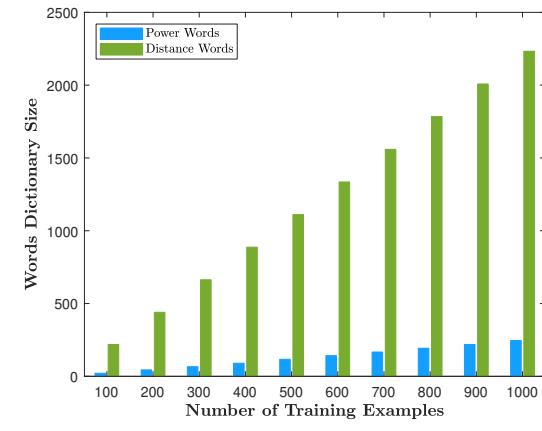
responding power values are clustered in an unsupervised manner using the Growing Neural Gas (GNG) algorithm. The resulting clusters for both distances and power values are depicted as "distance letters" and "power letters." Fig. 6 illustrates an example of the resulting letters when considering two ground users being multiplexed on the same subchannel. From this figure, it is evident that the distance letters are more spread out across the distance space compared to the power letters, as the distance space is broader than the power space. Additionally, both the dynamic distance and their corresponding derivatives and the power and their corresponding derivatives were analyzed to capture the dynamic changes, which evolve in response to mobility.

The distance and power letters are combined to form the distance and power words. This coupling process involves grouping the firing distance letter of a specific ground user at a given moment with the firing distance letter of another ground user at the same moment. The resulting distance words reflect the evolution of the distance letters over time. Similarly, to create power words, we group the firing power letter from a specific ground user at a particular moment with the firing power letter of another ground user at that same moment. Thus, the distance word can be understood as representing the physical configuration of the environment as the UAV moves. In contrast, the power word represents the digital configuration that reflects the digital landscape. By coupling the distance words and power words, we create a mapping between the physical and digital worlds, resulting in a set of tokens that symbolize the cross-correlation between the two realms. These tokens are then encoded in a transition matrix, which enables us to predict the dynamic evolution of both the physical and digital worlds, as well as the rules that govern that evolution. Additionally, the UAV trajectories stored in the training examples are clustered using GNG, resulting in a collection of clusters referred to as "trajectory letters." These trajectory letters are then coupled with the tokens, creating an interactive matrix that encodes the probabilistic relationships between the trajectory letters and the tokens, as illustrated in Fig. 7. The primary objective is to allow the UAV, once it determines its mode of motion, to predict how both the physical and digital worlds will evolve around it, thereby enhancing its power allocation strategies.

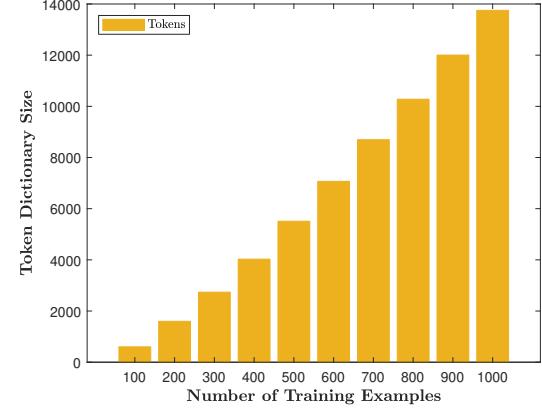
The dictionary size, encompassing the total number of letters, words, and tokens, is contingent upon the total number of training examples and demonstrations. It is essential to evaluate this size to ascertain whether performance may deteriorate or remain stable. We contend that the proposed approach does not require an extensive dictionary to achieve high generalization capabilities. Rather, it is effective with smaller dictionaries, owing to the adaptability facilitated by active inference, which anticipates future needs and refines actions based on predictive errors. Fig. 8-(a) illustrates the relationship between the number of distance letters and power letters as a function of the total number of training examples. It is clear that the total number of letters increases with the number of training examples. This is because more examples from both the physical and digital domains lead to the discovery of a greater variety of distances and power



(a) Letter Dictionary Size



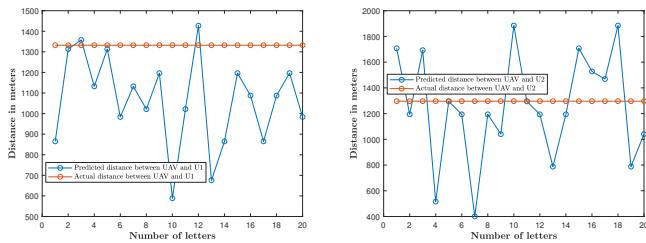
(b) Word Dictionary Size



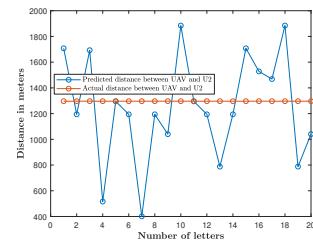
(c) Tokens Dictionary Size

Fig. 8. Dictionary size versus the total number of training examples, including: (a) the number of letters, (b) the number of words, and (c) the number of tokens.

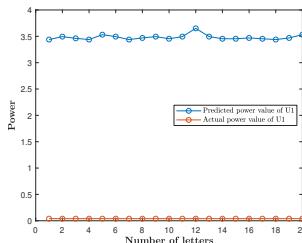
values. Additionally, we can see that the number of distance letters exceeds the number of power letters. This difference is attributed to the fact that the distance space is more expansive than the power space, which is constrained by the total power budget. Fig. 8-(b) shows how the dictionary size, including distance and power words, changes with the number of training examples. It is evident from the figure that the total number of words increases as the number of training examples rises.



(a) Distance between UAV and User 1

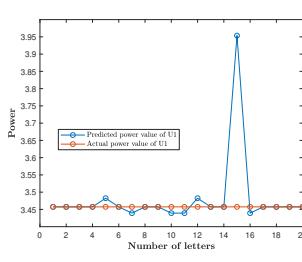


(b) Distance between UAV and User 2



(c) Power allocated to User 1

(d) Power allocated to User 2



(e) Corrected Power allocated to User 1

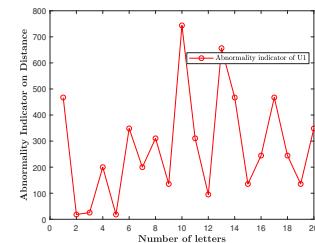
(f) Corrected Power allocated to User 2

Fig. 9. An example showing: the physical configuration representing the distances between the UAV and the ground users (User 1 and User 2) (refer to (a) and (b)), and the digital configuration representing the allocated power values to the ground users (refer to (c), (d)), as well as the corrected power allocated to the two users using the errors on power letters (refer to (e), (f)).

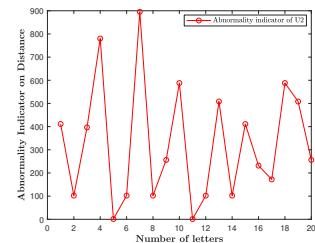
This increase can be attributed to the greater variety of letters available, as illustrated in Fig. 8-(a). Additionally, the number of distance words is significantly higher than that of power words. This discrepancy can be explained by the constrained power space, which leads to many repeated power words, resulting in a smaller overall count of unique power words. Fig. 8-(c) illustrates how the number of tokens varies with the total number of training examples. As expected, the number of tokens increases as the number of training examples grows. This is evident because the increase in the number of words shown in Fig. 8-(b) directly contributes to the rise in the number of tokens.

### C. Active decision-making

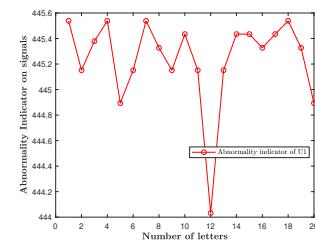
The UAV utilizes the GDBN model to determine possible actions that represent its movements. This allows it to predict future tokens that signify the anticipated physical and digital configurations, expressed through semantic words and letters. These represent the expected distances between the UAV and ground users, as well as the anticipated power values that



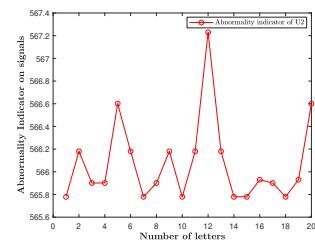
(a) Abnormality indicator on distance letters of User 1



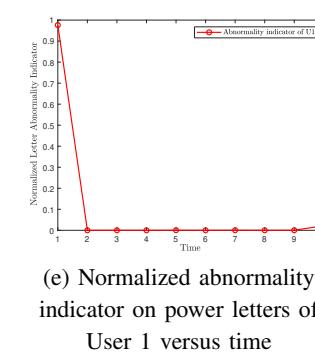
(b) Abnormality indicator on distance letters of User 2



(c) Abnormality indicator on power letters of User 1



(d) Abnormality indicator on power letters of User 2



(e) Normalized abnormality indicator on power letters of User 1 versus time

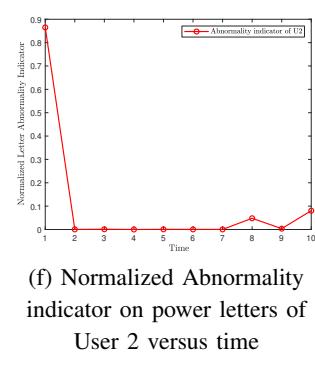


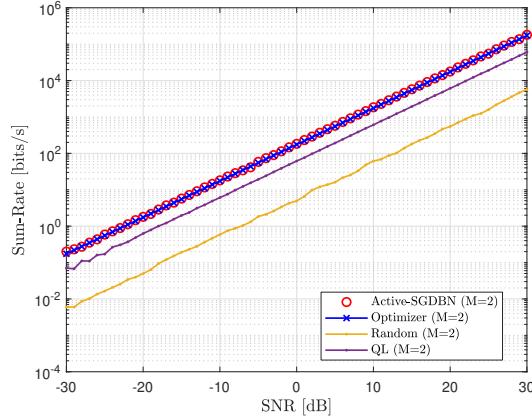
Fig. 10. An example showing: the abnormality measurements on distance and power letters (refer to (a), (b), (c), and (d)) corresponding to the physical and digital configurations, and the normalized abnormality indicators on power letters after correcting the allocated powers to the users (refer to (e), (f)).

may be allocated based on these predicted distances, which are influenced by the UAV's motion.

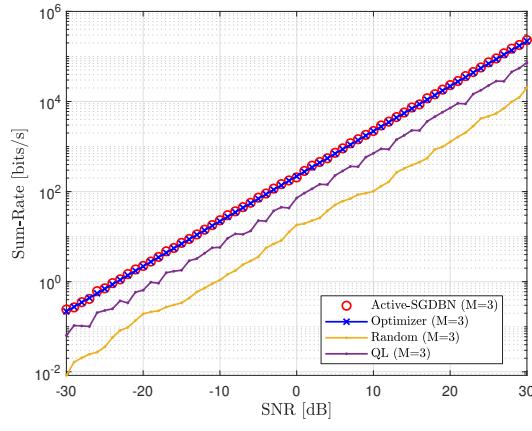
The predicted tokens, words, and letters serve as semantic symbols that the UAV will compare with perceived symbols after gathering new observations. This comparison enables the UAV to refine its actions and adapt to any unexpected situations.

Fig. 9-(a) and Fig. 9-(b) display examples of the predicted distances between the UAV and two ground users, which are multiplexed on the same sub-channel after the UAV performs an initial action representing its movement. Initially, the UAV selects very low power values, as demonstrated in Fig. 9-(c) and Fig. 9-(d). At this stage, the UAV is still trying to understand its surrounding environment, particularly in terms of the distances to the two users and how these measurements compare to its predictions, which represent the preferred environmental states.

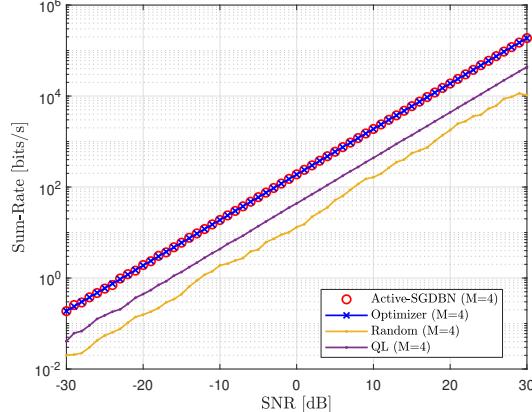
In Fig. 9-(c) and Fig. 9-(d), the initial power values (shown in orange) are still significantly different from the desired power values (shown in blue). The world model suggests the



(a) Sum-rate M=2



(b) Sum-rate M=3

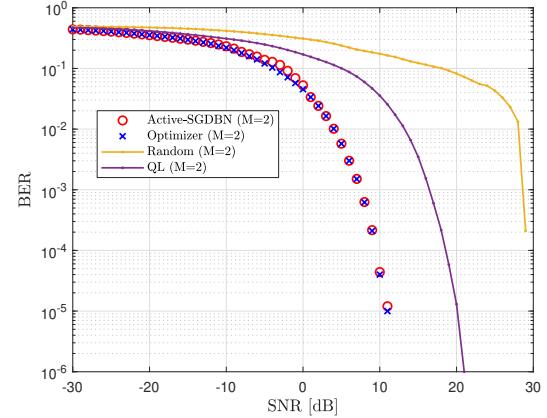


(c) Sum-rate M=4

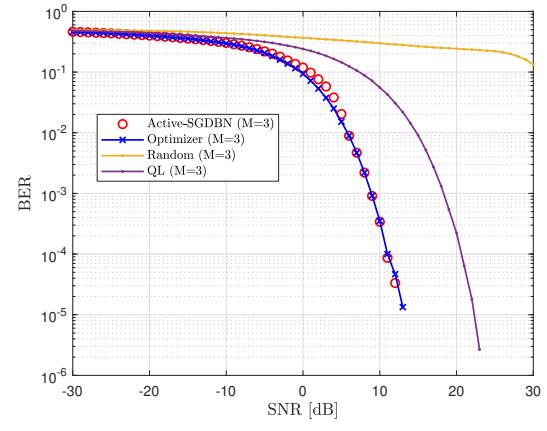
Fig. 11. Average Sum-rate versus SNR for different numbers of ground users multiplexed on the same sub-channel: (a) M = 2, (b) M = 3, (c) M = 4.

desired values based on the expected movements of the UAV and the anticipated changes in the physical environment.

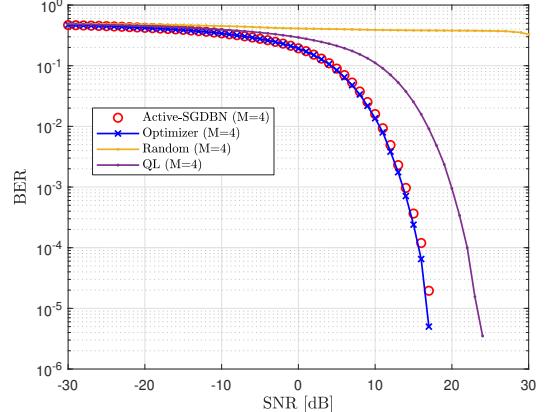
The UAV propagates a set of predicted tokens using the HPF, in the shown example with 20 particles. This set includes both words and letters and evaluates which of these can best explain the current physical configuration. From this evaluation, UAV deduces the best digital configuration through semantic abnormality indicators that measure how far the currently perceived semantic symbols differ from the preferred



(a) BER M=2



(b) BER M=3



(c) BER M=4

Fig. 12. Average BER versus SNR for different numbers of ground users multiplexed on the same sub-channel: (a) M = 2, (b) M = 3, (c) M = 4.

ones.

Fig. 10-(a) and Fig. 10-(b) illustrate the semantic abnormality indicators related to the distance values of User 1 and User 2, based on the 20 predicted letters. Meanwhile, Fig. 10-(c) and Fig. 10-(d) depict the semantic abnormality indicators for the power values of User 1 and User 2, again considering 20 predicted letters.

The UAV selects the predicted letter that results in the minimum abnormality and subsequently chooses the corre-

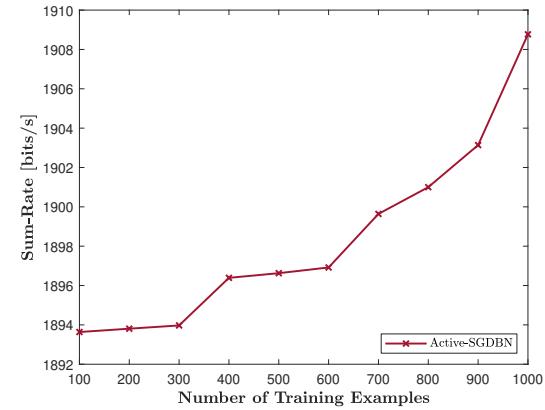
sponding word to use as a reference. This process allows the UAV to correct its initial actions using the prediction errors, resulting in new words that represent configurations (both physical and digital) that it has not encountered during training. This is illustrated in Fig. 9-(e) and Fig. 9-(f), which show how the corrected power actions aligned more closely with the preferred actions, represented by the predicted letters. This adjustment helps minimize abnormality, as illustrated in Fig. 10-(e) and Fig. 10-(f). Consequently, the UAV continues this adjustment process online after each movement and action, aiming to improve its performance in terms of sum-rate and bit error rate (BER) to match the optimizer's performance.

Fig. 11 illustrates the sum-rate values obtained from the proposed approach, SGDBN, against the signal-to-noise ratio (SNR). It also compares these results with those from the expert optimizer (i.e., exhaustive search), a random allocation method, and Q-learning (QL). The random approach involves randomly allocating power values to available users from the available power pool, which is the same pool used by the optimizer for a logical comparison. Q-learning was trained using the same dataset as SGDBN to create a large Q-table, ensuring a fair comparison. Additionally, Q-learning was equipped with an abnormality indicator to help it select the nearest state from the large Q-table during exploration, particularly when the agent encounters new states. The figure clearly shows that the sum rate increases as the SNR increases for all schemes. Notably, the proposed approach achieves the same sum rate as the optimizer and outperforms the other methods. This alignment with the optimizer's performance indicates that the UAV, equipped with a world model that suggests allocation strategies (actions) and utilizes active inference to correct previous actions, has effectively learned the strategic rules employed by the optimizer. Consequently, it can apply these rules to entirely new examples during testing, solving them quickly and efficiently without the need for external rewards or supervision.

Fig. 12 illustrates the bit error rates (BER) as a function of various signal-to-noise ratio (SNR) values for the proposed method. It compares these results to those produced by an expert optimizer, a random scheme, and Q-learning (QL). The BER is used to evaluate whether the allocated power values have allowed for the correct decoding of user signals. The figure clearly shows that the BER of the proposed approach matches that of the optimizer, indicating that the proposed method successfully learned the optimizer's rules for allocating power values to multiple users, even in entirely new environmental conditions. This demonstrates that the proposed approach enables the UAV to generalize suggested solutions from the world model into adaptive solutions for addressing new tasks and scenarios. Additionally, the figure reveals that the proposed method outperforms both the random scheme and QL in terms of BER.

#### D. Impact of training examples on performance

The proposed approach demonstrates the ability to generalize to new situations. However, this generalization capability typically relies on the number of training examples used to



(a) Average Sum-rate

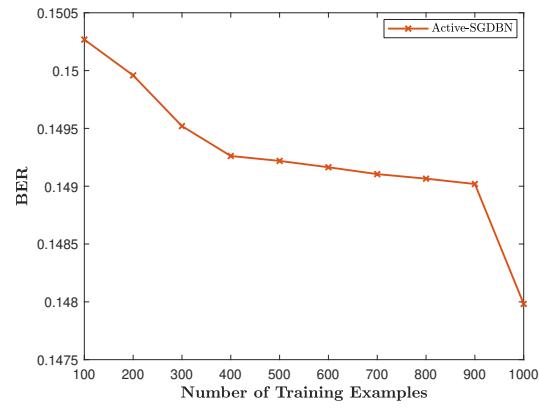


Fig. 13. Performance evaluation in terms of (a) average sum-rate and (b) average BER versus the number of training examples (dictionary size).

build predictive generative models, such as the world model we developed. It is known that as we increase the number of training examples, the model's effectiveness in regenerating data improves.

To evaluate the performance of the proposed approach and the associated world model, we conducted several tasks using varying numbers of training examples and different dictionary sizes. Fig. 13 illustrates the performance in terms of sum rate and Bit Error Rate (BER) against the number of training examples. Our observations indicate that the performance gap is minimal as the number of training examples increases, which validates the proposed method's ability to generalize effectively even with a limited number of examples. This aligns with the vision of transitioning towards smaller language models instead of larger models, especially for use in compact wireless devices.

#### E. Impact of UAV mobility patterns on performance

We evaluate *Active-SGDBN* alongside an exhaustive *Optimizer* over four UAV mobility patterns: circle, line (back-forth), random walk, and an intentionally irregular zigzag path. Each trajectory is resampled at equal arc length so that episodes share the same number of slots and per-slot displacement. The world model is trained with episodes from circle, line, and random walk. Testing always uses new episodes

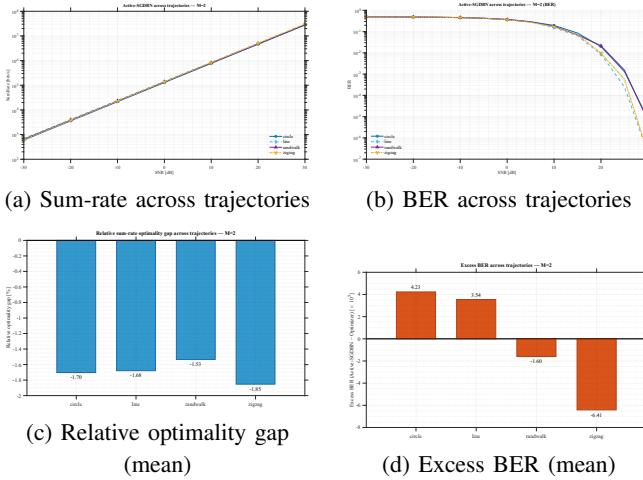


Fig. 14. UAV mobility impact,  $M = 2$ . (a)–(b) Active-SGDBN sum-rate/BER vs SNR across circle, line, random-walk, and zigzag—curves nearly coincide (trajectory-invariant). (c) Mean relative optimality gap ( $R_{\text{OPT}} - R_{\text{ACT}} / R_{\text{OPT}} \leq 2\%$ ). (d) Mean excess BER ( $\text{BER}_{\text{ACT}} - \text{BER}_{\text{OPT}} \approx 10^{-5}$ , near zero).

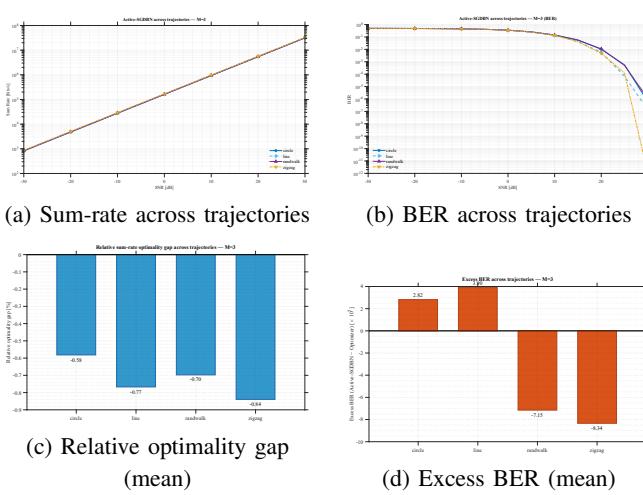


Fig. 15. UAV mobility impact,  $M = 3$ . (a)–(b) Active-SGDBN sum-rate/BER vs SNR across four trajectories—curves almost overlap. (c) Mean relative optimality gap < 1%. (d) Mean excess BER  $\approx 10^{-5}$ , near zero.

for all four paths. Thus, even for the families used during learning, evaluation is conducted on unseen realizations, while zigzag serves as an out-of-distribution (OOD) stress case. In the figures,  $R_{\text{ACT}}/\text{BER}_{\text{ACT}}$  denote the sum-rate/BER achieved by Active-SGDBN and  $R_{\text{OPT}}/\text{BER}_{\text{OPT}}$  those of the Optimizer under the same channel/user realizations.

The policy operates on a semantic state: the *sorted* user-UAV distances compactly encode the near/far structure that drives interference coupling and SIC effectiveness, while suppressing path-order effects. A tokenized world model provides an experience memory that binds recurring distance patterns to effective power-allocation/decoding-order patterns (tokens) distilled from the Optimizer, and—through trajectory labels—places a prior on how distance patterns tend to evolve

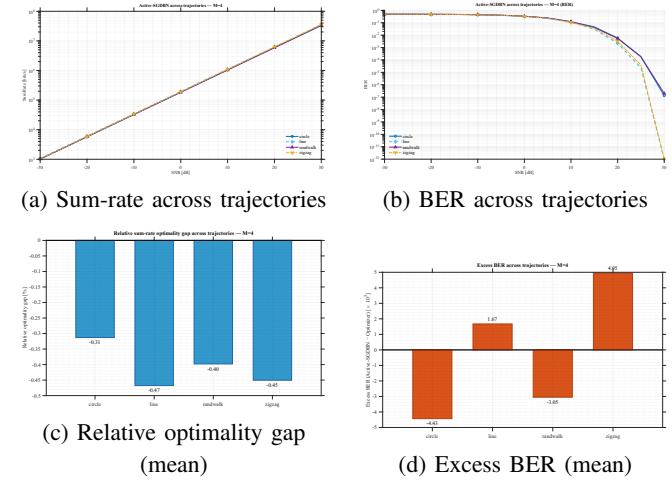


Fig. 16. UAV mobility impact,  $M = 4$ . (a)–(b) Active-SGDBN sum-rate/BER vs SNR across four trajectories—nearly identical trends. (c) Mean relative optimality gap < 0.6%. (d) Mean excess BER  $\approx 10^{-5}$ , near zero.

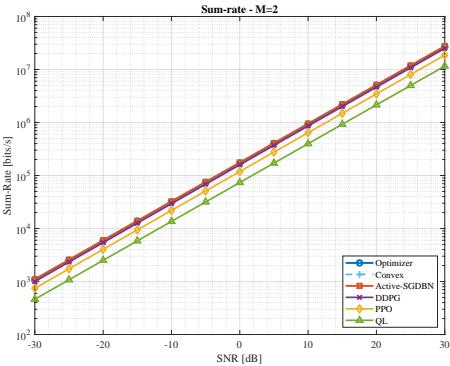
along a path. At run time, the agent retrieves a small set of relevant tokens for the current semantic state, anticipates short-term geometry changes under the current mobility family, and refines the selected token with a lightweight local adjustment. In effect, power is used to “shape” interference so the present slot behaves like well-solved cases in memory, without requiring explicit knowledge of the user distribution or a parametric channel law.

Figures 14, 15, and 16 summarize the outcome for  $M \in \{2, 3, 4\}$ . In panels (a)–(b), the Active-SGDBN sum-rate and BER curves across circle, line, random walk, and zigzag nearly coincide over the full SNR range, indicating that throughput and reliability are essentially insensitive to the mobility pattern. Panels (c)–(d) quantify this: the mean relative optimality gap between  $R_{\text{ACT}}$  and  $R_{\text{OPT}}$  remains small and trajectory-agnostic (a few percent for  $M=2$  and typically sub-percent for  $M \geq 3$ ), and the mean excess BER ( $\text{BER}_{\text{ACT}} - \text{BER}_{\text{OPT}}$ ) is near zero (order  $10^{-5}$ ).

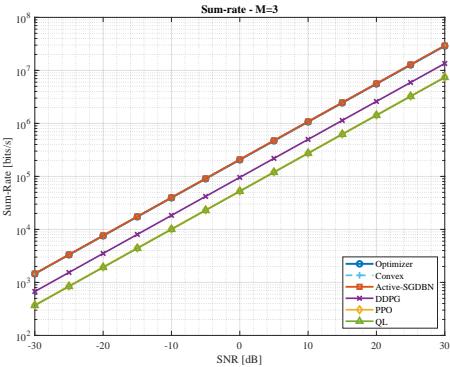
Overall, the semantic encoding of geometry, the trajectory-conditioned world model, and the on-line retrieval-plus-refinement step yield a policy that *generalizes* across mobility patterns. Changing the UAV path primarily reorders the sequence of geometries encountered; the learned prior anticipates these changes and the refinement reliably recovers high-quality power allocations and decoding orders on unseen episodes and even an OOD trajectory, without trajectory-specific tuning.

#### F. Comparison with advanced DRL under mixed-trajectory mobility

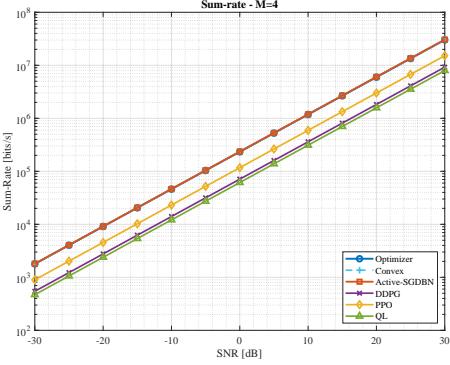
Figures 17(a)–(c) and 18(a)–(c) compare our *Active-SGDBN* with DDPG, PPO, and Q-learning (QL) under harder, unseen test cases and mixed-trajectory training (circle/line/random-walk/zigzag) to emulate realistic mobility and reduce overfitting. Across SNRs and user loads  $M = \{2, 3, 4\}$ , *Active-SGDBN* tracks the *Optimizer* in sum-rate while keeping BER



(a) Sum-rate M=2



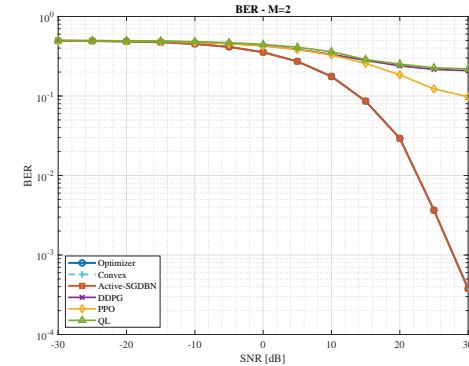
(b) Sum-rate M=3



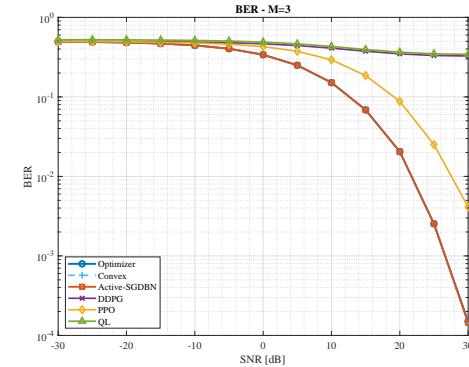
(c) Sum-rate M=4

Fig. 17. Average Sum-rate versus SNR for different numbers of ground users multiplexed on the same sub-channel: (a) M = 2, (b) M = 3, (c) M = 4.

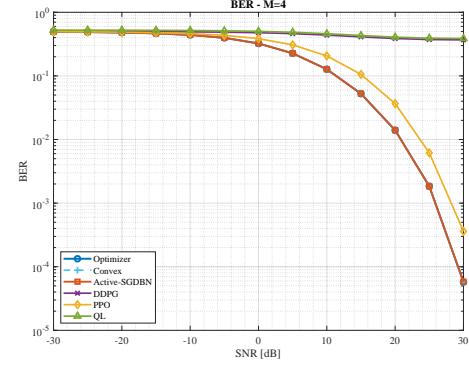
as low as the expert, consistently outperforming all DRL baselines; see the overlap between Active-SGDBN and Optimizer in (a)–(c) for both figures and the persistent gap to DDPG/PPO/QL. The advantage stems from *semantic world-modeling with active inference*: distances and powers are compressed into discrete “letters” by a compact GNG quantizer, composed into words/tokens that capture the joint UAV–user geometry and power policy; at run time a trajectory-conditioned prior (via token transitions) proposes actions, and abnormality indicators, i.e., KL divergences between *predicted* and *perceived* semantic symbols, drive online refinement of power decisions (closing the loop without costly exploration).



(a) BER M=2



(b) BER M=3



(c) BER M=4

Fig. 18. Average BER versus SNR for different numbers of ground users multiplexed on the same sub-channel: (a) M = 2, (b) M = 3, (c) M = 4.

This combination yields Optimizer-level performance with lightweight inference even when the mobility pattern shifts, demonstrating that our semantic, model-based agent generalizes more reliably than purely policy-gradient or value-based DRL under mixed-trajectory mobility.

## V. CONCLUSION

In this paper, we proposed a novel semantic context-aware reasoning framework, Active-SGDBN, for improved resource management in a UAV-based NOMA system. A global dictionary of semantic tokens was developed to encapsulate the maximum sum rate by jointly optimizing power

allocation based on UAV trajectory via exhaustive search optimization. The UAV utilizes the solutions generated by the expert (exhaustive search) to learn a set of training semantic samples that mimic the expert's optimization strategies using the GNG unsupervised clustering technique. These training samples constitute the UAV's beliefs and generative model. During online deployment, the UAV leverages this knowledge to make informed decisions by constantly updating its actions and beliefs based on incoming semantic data. We demonstrate that as the number of training samples increases, the number of semantic tokens also expands. Furthermore, we show that by analyzing abnormalities or errors related to the distances between the UAV and the power values of users, the UAV can predict tokens that minimize abnormalities to optimize its policies. Simulation results validate the effectiveness of the proposed Active-SGDBN, showing that it perfectly matches the optimizer (exhaustive search) while outperforming Q-learning and random optimization schemes in terms of BER and sum rate.

Future work will integrate trajectory planning with resource allocation in one semantic world model, expanding the action space to digital plus physical control and supporting multi-objective active-inference optimization.

## REFERENCES

- [1] T. Elmokadem and A. V. Savkin, "Towards fully autonomous uavs: A survey," *Sensors*, vol. 21, no. 18, p. 6223, 2021.
- [2] X. Liu, M. Chen, Y. Liu, Y. Chen, S. Cui, and L. Hanzo, "Artificial intelligence aided next-generation networks relying on uavs," *IEEE Wireless Communications*, vol. 28, no. 1, pp. 120–127, 2021.
- [3] Y. Su, L. Huang, and M. LiWang, "Joint power control and time allocation for uav-assisted iov networks over licensed and unlicensed spectrum," *IEEE Internet of Things Journal*, vol. 11, no. 1, pp. 1522–1533, Jan 2024.
- [4] S. Geng, Z. Wei, J. Zhao, F. Shen, J. Joung, and S. Sun, "Joint deployment and resource allocation for service provision in multi-uav-assisted wireless networks," *IEEE Internet of Things Journal*, vol. 11, no. 22, pp. 37269–37286, Nov 2024.
- [5] W. Li, H. Liang, C. Dong, X. Xu, P. Zhang, and K. Liu, "Non-orthogonal multiple access enhanced multi-user semantic communication," *IEEE Transactions on Cognitive Communications and Networking*, vol. 9, no. 6, pp. 1438–1453, 2023.
- [6] Z. Liu, J. Qi, Y. Shen, K. Ma, and X. Guan, "Maximizing energy efficiency in uav-assisted noma-mec networks," *IEEE Internet of Things Journal*, vol. 10, no. 24, pp. 22 208–22 222, Dec 2023.
- [7] R. Kumar, C. K. Singh, P. K. Upadhyay, A. M. Salhab, A. A. Nasir, and M. Masood, "Iot-inspired cooperative spectrum sharing with energy harvesting in uav-assisted noma networks: Deep learning assessment," *IEEE Internet of Things Journal*, vol. 10, no. 24, pp. 22 182–22 196, Dec 2023.
- [8] Y.-Y. Guo, X.-L. Tan, Y. Gao, J. Yang, and Z.-C. Rui, "A deep reinforcement approach for energy-efficient resource assignment in cooperative noma-enhanced cellular networks," *IEEE Internet of Things Journal*, vol. 10, no. 14, pp. 12 690–12 702, July 2023.
- [9] F. Obite, A. Kravani, A. S. Alam, L. Marcenaro, A. Nallanathan, and C. Regazzoni, "Intelligent resource allocation for uav-based cognitive noma networks: An active inference approach," in *2023 IEEE Future Networks World Forum (FNWF)*, 2023, pp. 1–7.
- [10] S. Pakravan, J.-Y. Chouinard, X. Li, M. Zeng, W. Hao, Q.-V. Pham, and O. A. Dobre, "Physical layer security for noma systems: Requirements, issues, and recommendations," *IEEE Internet of Things Journal*, vol. 10, no. 24, pp. 21 721–21 737, Dec 2023.
- [11] S. M. R. Islam, M. Zeng, O. A. Dobre, and K.-S. Kwak, "Resource allocation for downlink noma systems: Key techniques and open issues," *IEEE Wireless Communications*, vol. 25, no. 2, pp. 40–47, 2018.
- [12] Z. Ding, Z. Yang, P. Fan, and H. V. Poor, "On the performance of non-orthogonal multiple access in 5g systems with randomly deployed users," *IEEE Signal Processing Letters*, vol. 21, no. 12, pp. 1501–1505, 2014.
- [13] Z. Yang, Z. Ding, P. Fan, and G. K. Karagiannidis, "On the performance of non-orthogonal multiple access systems with partial channel information," *IEEE Transactions on Communications*, vol. 64, no. 2, pp. 654–667, 2016.
- [14] T. Van Luong, N. Shlezinger, C. Xu, T. M. Hoang, Y. C. Eldar, and L. Hanzo, "Deep learning based successive interference cancellation for the non-orthogonal downlink," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 11, pp. 11 876–11 888, 2022.
- [15] Q. Lan, D. Wen, Z. Zhang, Q. Zeng, X. Chen, P. Popovski, and K. Huang, "What is semantic communication? a view on conveying meaning in the era of machine intelligence," *Journal of Communications and Information Networks*, vol. 6, no. 4, pp. 336–371, 2021.
- [16] Y. Xiao, Y. Liao, Y. Li, G. Shi, H. V. Poor, W. Saad, M. Debbah, and M. Bennis, "Reasoning over the air: A reasoning-based implicit semantic-aware communication framework," *IEEE Transactions on Wireless Communications*, vol. 23, no. 4, pp. 3839–3855, 2024.
- [17] X. Xu, B. Xu, S. Han, C. Dong, H. Xiong, R. Meng, and P. Zhang, "Task-oriented and semantic-aware heterogeneous networks for artificial intelligence of things: Performance analysis and optimization," *IEEE Internet of Things Journal*, vol. 11, no. 1, pp. 228–242, Jan 2024.
- [18] R. Kaewpuang, M. Xu, W. Y. B. Lim, D. Niyyato, H. Yu, J. Kang, and X. Shen, "Cooperative resource management in quantum key distribution (qkd) networks for semantic communication," *IEEE Internet of Things Journal*, vol. 11, no. 3, pp. 4454–4469, Feb 2024.
- [19] C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, no. 3, pp. 379–423, 1948.
- [20] D.-T. Do, A.-T. Le, Y. Liu, and A. Jamalipour, "User grouping and energy harvesting in uav-noma system with af/df relaying," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 11, pp. 11 855–11 868, 2021.
- [21] X. Liu, J. Wang, N. Zhao, Y. Chen, S. Zhang, Z. Ding, and F. R. Yu, "Placement and power allocation for noma-uav networks," *IEEE Wireless Communications Letters*, vol. 8, no. 3, pp. 965–968, 2019.
- [22] D. Hu, Q. Zhang, Q. Li, and J. Qin, "Joint position, decoding order, and power allocation optimization in uav-based noma downlink communications," *IEEE Systems Journal*, vol. 14, no. 2, pp. 2949–2960, 2020.
- [23] A. A. Nasir, H. D. Tuan, T. Q. Duong, and H. V. Poor, "Uav-enabled communication using noma," *IEEE Transactions on Communications*, vol. 67, no. 7, pp. 5126–5138, 2019.
- [24] N. Zhao, X. Pang, Z. Li, Y. Chen, F. Li, Z. Ding, and M.-S. Alouini, "Joint trajectory and precoding optimization for uav-assisted noma networks," *IEEE Transactions on Communications*, vol. 67, no. 5, pp. 3723–3735, 2019.
- [25] B. Li, R. Zhang, and L. Yang, "Joint user grouping and power allocation for noma-based uav relaying networks," in *2021 IEEE International Conference on Communications Workshops (ICC Workshops)*, 2021, pp. 1–5.
- [26] W. Xu, J. Tian, L. Gu, and S. Tao, "Joint placement and power optimization of uav-relay in noma enabled maritime iot system," *Drones*, vol. 6, no. 10, p. 304, 2022.
- [27] M. Jia, Q. Gao, Q. Guo, and X. Gu, "Energy-efficiency power allocation design for uav-assisted spatial noma," *IEEE Internet of Things Journal*, vol. 8, no. 20, pp. 15 205–15 215, 2021.
- [28] H. Sun, X. Chen, Q. Shi, M. Hong, X. Fu, and N. D. Sidiropoulos, "Learning to optimize: Training deep neural networks for interference management," *IEEE Transactions on Signal Processing*, vol. 66, no. 20, pp. 5438–5453, 2018.
- [29] J. Zhu, J. Wang, Y. Huang, S. He, X. You, and L. Yang, "On optimal power allocation for downlink non-orthogonal multiple access systems," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 12, pp. 2744–2757, 2017.
- [30] Q. Wu, Y. Zeng, and R. Zhang, "Joint trajectory and communication design for multi-uav enabled wireless networks," *IEEE Transactions on Wireless Communications*, vol. 17, no. 3, pp. 2109–2121, 2018.
- [31] M. Liu, G. Gui, N. Zhao, J. Sun, H. Gacanin, and H. Sari, "Uav-aided air-to-ground cooperative nonorthogonal multiple access," *IEEE Internet of Things Journal*, vol. 7, no. 4, pp. 2704–2715, 2020.
- [32] P. Si, J. Zhao, K.-Y. Lam, and Q. Yang, "Uav-assisted semantic communication with hybrid action reinforcement learning," in *GLOBECOM 2023 - 2023 IEEE Global Communications Conference*, 2023, pp. 3801–3806.

- [33] J. Cui, Y. Liu, and A. Nallanathan, "Multi-agent reinforcement learning-based resource allocation for uav networks," *IEEE Transactions on Wireless Communications*, vol. 19, no. 2, pp. 729–743, 2020.
- [34] Y. Huang, X. Mo, J. Xu, L. Qiu, and Y. Zeng, "Online maneuver design for uav-enabled noma systems via reinforcement learning," in *2020 IEEE Wireless Communications and Networking Conference (WCNC)*, 2020, pp. 1–6.
- [35] J. Zhao, L. Yu, K. Cai, Y. Zhu, and Z. Han, "Ris-aided ground-aerial noma communications: A distributionally robust drl approach," *IEEE Journal on Selected Areas in Communications*, vol. 40, no. 4, pp. 1287–1301, 2022.
- [36] S. K. Mahmud, Y. Liu, Y. Chen, and K. K. Chai, "Adaptive reinforcement learning framework for noma-uav networks," *IEEE Communications Letters*, vol. 25, no. 9, pp. 2943–2947, 2021.
- [37] P. Si, L. Qian, J. Zhao, and K.-Y. Lam, "A hybrid framework of reinforcement learning and convex optimization for uav-based autonomous metaverse data collection," *IEEE Network*, vol. 37, no. 4, pp. 248–254, 2023.
- [38] Z. Yang, M. Chen, Z. Zhang, and C. Huang, "Energy efficient semantic communication over wireless networks with rate splitting," *IEEE Journal on Selected Areas in Communications*, vol. 41, no. 5, pp. 1484–1495, 2023.
- [39] L. Yan, Z. Qin, R. Zhang, Y. Li, and G. Y. Li, "Resource allocation for text semantic communications," *IEEE Wireless Communications Letters*, vol. 11, no. 7, pp. 1394–1398, 2022.
- [40] L. Yan, Z. Qin, R. Zhang, Y. Li, and G. Ye Li, "Qoe-aware resource allocation for semantic communication networks," in *GLOBECOM 2022 - 2022 IEEE Global Communications Conference*, 2022, pp. 3272–3277.
- [41] J. Su, Z. Liu, Y.-a. Xie, K. Ma, H. Du, J. Kang, and D. Niyato, "Semantic communication-based dynamic resource allocation in d2d vehicular networks," *IEEE Transactions on Vehicular Technology*, vol. 72, no. 8, pp. 10 784–10 796, 2023.
- [42] L. Xia, Y. Sun, X. Li, G. Feng, and M. A. Imran, "Wireless resource management in intelligent semantic communication networks," in *IEEE INFOCOM 2022 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, 2022, pp. 1–6.
- [43] C. Chacour, W. Saad, M. Debbah, Z. Han, and H. V. Poor, "Less data, more knowledge: Building next generation semantic communication networks," 2022.
- [44] K. Friston, "Embodied inference and spatial cognition," *Cognitive Processing*, vol. 13, pp. 171–177, 2012.
- [45] K. J. Friston, T. Parr, and B. de Vries, "The graphical brain: Belief propagation and active inference," *Network neuroscience*, vol. 1, no. 4, pp. 381–414, 2017.
- [46] T. Isomura, H. Shimazaki, and K. J. Friston, "Canonical neural networks perform active inference," *Communications Biology*, vol. 5, no. 1, p. 55, 2022.
- [47] Y. Zeng and R. Zhang, "Energy-efficient uav communication with trajectory optimization," *IEEE Transactions on Wireless Communications*, vol. 16, no. 6, pp. 3747–3760, 2017.
- [48] J. Lu, Y. Wang, T. Liu, Z. Zhuang, X. Zhou, F. Shu, and Z. Han, "Uav-enabled uplink non-orthogonal multiple access system: Joint deployment and power control," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 9, pp. 10 090–10 102, 2020.
- [49] F. A. Rabee, K. Davaslioglu, and R. Gitlin, "The optimum received power levels of uplink non-orthogonal multiple access (noma) signals," in *2017 IEEE 18th Wireless and Microwave Technology Conference (WAMICON)*, 2017, pp. 1–4.
- [50] L. Salaün, C. S. Chen, and M. Coupechoux, "Optimal joint subcarrier and power allocation in noma is strongly np-hard," in *2018 IEEE International Conference on Communications (ICC)*, 2018, pp. 1–7.
- [51] Y. Igarashi, H. Takenaka, Y. Nakanishi-Ohno, M. Uemura, S. Ikeda, and M. Okada, "Exhaustive search for sparse variable selection in linear regression," *Journal of the Physical Society of Japan*, vol. 87, no. 4, p. 044802, 2018.
- [52] Y. Igarashi, K. Nagata, T. Kuwatani, T. Omori, Y. Nakanishi-Ohno, and M. Okada, "Three levels of data-driven science," in *Journal of Physics: Conference Series*, vol. 699, no. 1. IOP Publishing, 2016, p. 012001.
- [53] T. M. Cover and J. M. Van Campenhout, "On the possible orderings in the measurement selection problem," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 7, no. 9, pp. 657–661, 1977.
- [54] M. Tucker, H. Li, S. Agrawal, D. Hughes, K. Sycara, M. Lewis, and J. A. Shah, "Emergent discrete communication in semantic spaces," *Advances in Neural Information Processing Systems*, vol. 34, pp. 10 574–10 586, 2021.
- [55] I. J. Sledge and J. M. Keller, "Growing neural gas for temporal clustering," in *2008 19th International Conference on Pattern Recognition*, 2008, pp. 1–4.
- [56] A. Krayani, A. S. Alam, L. Marcenaro, A. Nallanathan, and C. Regazzoni, "An emergent self-awareness module for physical layer security in cognitive uav radios," *IEEE Transactions on Cognitive Communications and Networking*, vol. 8, no. 2, pp. 888–906, 2022.
- [57] 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 *GLOBECOM 2020 - 2020 IEEE Global Communications Conference*, 2020, pp. 1–7.
- [58] R. Kulhavý, *Recursive nonlinear estimation: a geometric approach*. Springer, 1996.



**Ali Krayani** (Member, IEEE) is an Assistant Professor within the Department of Electrical, Electronic, Telecommunications Engineering, and Naval Architecture (DITEN) at the University of Genoa, Italy. He earned his Bachelor's degree in Telecommunication Engineering from the Politecnico di Torino, Italy, in 2014, followed by a Master's degree in Telecommunication Engineering from the University of Florence, Italy, in 2017. In April 2022, he was awarded a joint Ph.D. degree from the University of Genoa, Italy, and Queen Mary University of London, London, United Kingdom.

Dr. Krayani has worked as a Software Engineer in various companies and was a Postdoctoral Research Fellow at DITEN from 2021 to 2023. In 2023, he received the Best Paper Award at the IEEE Wireless Communications and Networking Conference (WCNC). He serves as a Guest Editor and reviewer for several academic journals. His current research interests include integrated sensing and communication, cognitive radios, AI-enabled radios, wireless communications (5G, 6G), UAV communications, physical layer security, IoT, semantic communications, UAV swarms, NOMA, federated learning, and artificial intelligence.



**Felix Obite** (Member, IEEE) received his Postgraduate Diploma (PGD) in Electronics and Telecommunication Engineering from Ahmadu Bello University, Zaria, Nigeria, in 2012. He earned his Master of Engineering degree in Electrical, Electronics, and Telecommunication from Universiti Teknologi Malaysia in 2017. Felix recently completed a joint PhD in Interactive and Cognitive Environments, with a focus on wireless communication, at Queen Mary University of London and the University of Genoa, Italy, in October 2024. His doctoral research lies at the intersection of signal processing, information theory, explainable Bayesian inference, and neuroscience-inspired active machine learning, with a focus on emergent wireless communication systems. He also has research interests in reinforcement learning, federated learning, and semantic communication.



**Atm S. Alam** (Member, IEEE) is a Lecturer (Assistant Professor) at the School of Electronic Engineering and Computer Science (EECS), Queen Mary University of London, UK. His research interests are in the areas of cognitive communications and computing, ML/AI in wireless communications, and emerging applications of cognitive communications and computing in verticals. With years of academic and research experience, Dr. Alam is adept at bridging research, innovations, and commercialisations via digital transformations for a better world. He is a member of Institute of Electrical and Electronic Engineering (IEEE) and a Fellow of Higher Education Academy (FHEA), UK.



**Lucio Marcenaro** (Senior Member, IEEE) is an Associate Professor of Telecommunications at the University of Genoa, Italy. He has 20 years of experience in image and video sequence analysis. He has authored about 130 technical papers on signal and video processing for computer vision. Lucio Marcenaro is associate editor of the IEEE Transactions on Image Processing and IEEE Transactions on Circuits and Systems for Video Technology, technical program co-chair for 13th International Conference on Distributed Smart Cameras (ICDSC) and for the first IEEE International Conference on Autonomous Systems (IEEE ICAS 2021), co-organizer for the 2019 Summer School on Signal Processing (S3P), General Chair of the Symposium on Signal Processing for Understanding Crowd Dynamics. He is Senior Member of IEEE and active within the IEEE Signal Processing Italy Chapter and Director of Student Services Committee (2018-2021).



**Carlo Regazzoni** (Senior Member, IEEE) is Full Professor of Cognitive Telecommunications Systems with the Department of Electrical, Electronic, Telecommunications Engineering and Naval Architecture (DITEN), University of Genoa, Italy. He has been responsible for 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 has been a co-editor of four edited books (Kluwer) on intelligent video surveillance. He

is an author/co-author of more than 100 articles in international scientific journals and more than 300 papers in peer-reviewed international conference proceedings.

Dr. Regazzoni served as the General Chair for several conferences and an Associate Editor and a Guest Editor for many international technical journals. He served in many roles within the governance bodies of the IEEE Signal Processing Society (SPS), being the IEEE SPS Vice President Conferences from 2015 to 2017. He has been an Associate Editor of several international journals, including the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, the IEEE TRANSACTIONS ON IMAGE PROCESSING, IEEE SIGNAL PROCESSING LETTERS, and others. He has been a Guest Editor of special issues in international journals, including the PROCEEDINGS OF THE IEEE, the IEEE Signal Processing Magazine, and others.



**Arumugam Nallanathan** (S'97-M'00-SM'05-F'17) is Professor of Wireless Communications and Head of the Communication Systems Research (CSR) group in the School of Electronic Engineering and Computer Science at Queen Mary University of London since September 2017. He was with the Department of Informatics at King's College London from December 2007 to August 2017, where he was Professor of Wireless Communications from April 2013 to August 2017 and a Visiting Professor from September 2017 till August 2020. He was an Assistant Professor in the Department of Electrical and Computer Engineering, National University of Singapore from August 2000 to December 2007. His research interests include Artificial Intelligence for Wireless Systems, Beyond 5G Wireless Networks and Internet of Things (IoT). He published more than 700 technical papers in scientific journals and international conferences. He is a co-recipient of the Best Paper Awards presented at the IEEE International Conference on Communications 2016 (ICC'2016), IEEE Global Communications Conference 2017 (GLOBECOM'2017) and IEEE Vehicular Technology Conference 2018 (VTC'2018). He is also a co-recipient of IEEE Communications Society Leonard G. Abraham Prize in 2022. He is an IEEE Distinguished Lecturer. He has been selected as a Web of Science Highly Cited Researcher in 2016, and 2022-2024.

He was a Senior Editor for IEEE Wireless Communications Letters, an Editor for IEEE Transactions on Wireless Communications, IEEE Transactions on Communications, IEEE Transactions on Vehicular Technology and IEEE Signal Processing Letters. He served as the Chair for the Signal Processing and Communication Electronics Technical Committee of IEEE Communications Society and Technical Program Chair and member of Technical Program Committees in numerous IEEE conferences. He received the IEEE Communications Society SPCE outstanding service award 2012 and IEEE Communications Society RCC outstanding service award 2014.