Multi-Agent Reinforcement Learning Approach for Interference Optimization in Wi-Fi 8

Jamshid Bacha, Anaolij Zubow, and Falko Dressler
School of Electrical Engineering and Computer Science, TU Berlin, Germany
{bacha, zubow, dressler}@ccs-labs.org

***Abstract*—Interference management is becoming ever more complicated in multi-access networks such as Wi-Fi. Here, ultra-dense access point (AP) deployments add to the problem. With the introduction of beamforming, this can be in part reduced, however, only in combination with optimized nulling the operation of overlapping basic service set (OBSS) can be improved. In this paper, we present a machine learning (ML)-based solution for interference optimization for Wi-Fi 8. In particular, we present a multi-agent reinforcement learning (MARL) framework that uses a distributed approach to coordinate beamforming and nulling at nearby APs. We follow a two-step procedure: first, APs and nodes are clustered in order to reduce overhead due to exchange of channel state information; and secondly, the MARL framework configures beams for the next period. Simulation results confirm the advantages of our MARL approach.**

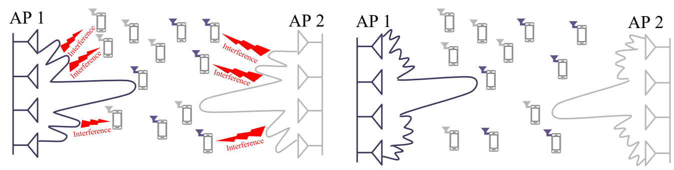
***Index Terms*—802.11bn, Overlapping basic service set, Co-ordinated beamforming, Interference optimization, Multi-agent reinforcement learning**

I. Introduction

The continuous growth of high-bandwidth and low-latency applications such as 4K video streaming, online gaming, and augmented reality and virtual reality (AR/VR) services is placing huge demands on Wi-Fi networks [1]. As user density and traffic volume increase, ensuring reliable and high-throughput Wi-Fi communication becomes a critical challenge, especially in indoor and urban scenarios. To address this, access point (AP) densification has emerged as a common deployment strategy, where multiple APs are installed within close proximity to improve signal coverage and throughput [2].

However, while the densification of APs can in general improve the signal for the users, the densification can also lead to a higher level of inter-AP interference, especially when multiple APs operate on the same channel [3]. In dense wireless local area network (WLAN) deployments, co-channel interference (CCI) from overlapping basic service set (OBSS) significantly reduces network efficiency and degrades user quality of service (QoS). This challenge is further increased by the limited availability of non-overlapping frequency channels in Wi-Fi bands, making it practically impossible to assign orthogonal channels to each AP in ultra-dense deployments [4]. As a result, multiple APs must operate on the same channel, leading to high CCI, which degrades network performance [5].

To improve spectrum efficiency and reduce interference, Wi-Fi 8 introduces the concept of coordinated beamforming (Co-BF) among multiple AP [6]. Specifically, Co-BF builds



(a) Beamforming and side lobes (b) Beamforming with null steering

Figure 1. Example of Co-BF: APs transmit simultaneously using (a) beamforming (to improve signal strength at receiver) and (b) null steering (to also reduce interference).

upon the spatial nulling mechanism used in multi-user multiple-input and multiple-output (MU-MIMO) in IEEE 802.11ac and extends it to multi-AP coordination. In this approach, each AP directs a beam toward its intended station while using the remaining spatial degrees of freedoms (DoFs) to suppress interference in the direction of other nearby non-associated stations to reduce interference on them.

However, Co-BF introduces its own set of challenges. An AP with N_t antennas can do one beam and N_{t-1} nulls per transmission slot. In practical scenarios, when the AP performs beamforming, there are also side lobes that interfere with nearby stations. However, the number of stations affected by these side-lobes interference is often higher than the number of available nulls at AP (cf. Figure 1a), which requires intelligent selection of the most disruptive directions to suppress those affected by these side lobes. For instance, an AP with four antennas can only select at most three stations to null (cf. Figure 1b).

Although nulling interference directions can improve signal quality for intended receivers, excessive nulling can also result in a significant drop in the overall data rate. This is because each null direction consumes spatial DoF, reducing the AP's ability to direct strong signals toward its intended stations. Frequent channel state information (CSI) updates are required, but this adds further overhead to network performance. Such as the artificial intelligence and machine learning (AIML) topic interest group (TIG) aims to enhance system goodput by reducing CSI feedback overhead and compression complexity through intelligent machine learning (ML) techniques without degrading performance [7].

Moreover, if there are large-scale scenarios where the number of APs and stations exceeds a manageable threshold, then handling the interference is even more complex if managed with a single reinforcement learning (RL) agent. Specifically, when the number of APs increases from a small-scale deployment to a realistic large-scale deployment (e.g., twenty to hundred APs with numerous associated stations), the computational overhead becomes prohibitive and it will be hard to handle the huge number of APs altogether [8]. Therefore, scalable and adaptive interference management mechanisms are urgently needed, particularly as network sizes continue to grow.

In this paper, we present a multi-agent reinforcement learning (MARL) framework with intelligent agents can provide a scalable and effective solution. By grouping AP based on the number of stations associated with each AP, the interference from non-associated APs, and the number of APs. These groups of APs will be managed by an independent single RL agent, and the system can better balance beamforming optimization and interference mitigation, ensuring higher network performance in ultra-dense environments. In our MARL framework, where each agent is responsible for a group of APs. This design addresses the scalability issue faced by single-agent approaches when managing large numbers of APs and stations.

Our key contributions can be summarized as follows:

- We present a centralized MARL framework for solving the cooperative beamforming problem plus nulling problem.
- We introduce a grouping algorithm based on the number of APs, station density, and real-time interference levels.
- We improve scalability in ultra-dense networks by limiting CSI exchange, channel sounding, and coordination signaling within local groups of APs.

II. Related Work

To address inter-AP interference and channel reuse challenges, different solutions are explored, such as MU-MIMO, introduced in IEEE 802.11ac and enhanced in 802.11ax [9]. MU-MIMO improves downlink throughput but its performance heavily depends on accurate CSI and the spatial separation of users, which is difficult to guarantee in ultra-dense environments [10]. Another solution is to use dynamic transmit power control (TPC), where APs adjust their transmission power levels to minimize interference while maintaining sufficient signal strength for their associated stations [11]. Although TPC reduces inter-AP interference to some extent, it can also reduce coverage and create dead zones if not carefully managed [12]. Without explicitly leveraging CSI and spatial interference patterns, these methods may fail to mitigate interference effectively in highly congested and spatially diverse Wi-Fi environments.

With the rise of artificial intelligence (AI), researchers are now using AI to improve Wi-Fi performance. For example, the performance challenges of channel access in OBSS environments under the upcoming IEEE 802.11bn standard are tackled by using MARL [13]. Several other studies in the field of 802.11 have investigated the application of MARL combined with knowledge transfer to enhance global awareness in both centralized and decentralized training setups. For instance, a distributed deep reinforcement learning (DRL) framework where device-to-device links adjust their channel selection and transmission power based on their own local data and outdated information from others is introduced in [14].

In addition, cooperative beamforming has been explored as part of multi-access point coordination (MAPC) in Wi-Fi 8. Some prior work in [15] applied multi-armed bandit (MAB) for cooperative spatial reuse scheduling, while [8] proposed a RL-based Co-BF framework combining beam steering and null steering. Both studies demonstrate the potential of ML to enhance interference management in dense Wi-Fi deployments, motivating our present MARL-based approach.

III. System Model

We consider an ultra-dense Wi-Fi network deployment where multiple APs equipped with a uniform linear array are placed in a fixed position and APs operate on the same channel using the full available spectrum, which can lead to increased interference if not properly managed. The stations are following random waypoint (RWP), a human walking speed of 1.4 m/s, commonly found in residential, office, or enterprise environments.

Each AP serves its associated stations using beamforming, with zero-forcing (ZF) as the primary technique. To enable interference-aware beamforming and null steering, the central controller acquires CSI by periodic channel sounding every 100 ms between each AP and relevant stations, including both associated and nearby non-associated stations. The received signal y at a station is modeled by the linear system

$$y = \mathbf{H} \cdot x + n, \tag{1}$$

where $\mathbf{H}$ is the channel matrix representing the wireless channel between the AP and the station, x is the transmitted signal vector, and n is the additive white Gaussian noise. Each AP computes its beamforming weight vector $\mathbf{w}$ using the ZF method as

$$\mathbf{w} = \left(\mathbf{H}\mathbf{H}^H\right)^{-1}\mathbf{H}^H\mathbf{e}, \tag{2}$$

where $\mathbf{H}$ is the combined channel matrix consisting of the channel vectors of the target station and the selected interference, $\mathbf{H}^H$ is its Hermitian transpose, and $\mathbf{e}$ is a selection vector that prioritizes the desired signal direction while steering nulls at others. $\mathbf{H}$ is defined as

$$\mathbf{H} = \left[\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_{N'_u}\right], \tag{3}$$

where each $\mathbf{h}_i$ is an $N_t \times 1$ channel vector representing the channel between the AP and a selected station. In conventional beamforming systems, achieving perfect nulling requires that the number of antennas N_t must be greater than or equal to the number of users N_u, i.e., $N_t \geq N_u$. However, in practical ultra-dense networks, the number of potential users often exceeds the available spatial DoF. To address this, our framework intelligently selects only the most disruptive directions N'_u for nulling based on real-time interference observations, maintains beamforming performance even when $N_u > N_t$. After calculating the $\mathbf{w}$, each AP serves one station per transmission slot and uses the remaining spatial DoF to nullify and reduce interference on other non-associated stations.

Table I
WIRELESS COMMUNICATION PARAMETERS

Parameter	**Value**
Frequency band	5 GHz
Channel bandwidth	160 MHz
Time slot	50 ms
AP to AP distance d_{AP}	20 m
AP to station distance r_{BSS}	10 m
Transmit power P_{tx}	20 dBm

This study also considers path loss as the power density decreases as it travels through space. We augment the standard free space path loss (FSPL) formulation with log-normal shadowing, such as

$$\mathrm{PL} = \begin{cases} \mathrm{PL_{fs}}(d) + X_\sigma, & \text{if } d \leq 10 \\ \mathrm{PL_{fs}}(10) + 10 \cdot 3.5 \cdot \log_{10}\left(\frac{d}{10}\right) + X_\sigma, & \text{if } d > 10 \end{cases}, \tag{4}$$

where X_σ represents the Gaussian noise term, which follows a log-normal distribution $X_\sigma \sim \mathcal{N}(0, \sigma^2)$. Including the noise term into the path loss equation helps simulate real-world signal fluctuations, making the predicted received signal strength (RSS) values more realistic and stable for the learning process.

The resulting signal-to-interference-plus-noise ratio (SINR) of station k is

$$\mathrm{SINR}_k = \frac{P_{\mathrm{tx},i} \cdot |\mathbf{h}_{k,i}^H \mathbf{w}_i|^2}{\sum_{j \neq i} P_{\mathrm{tx},j} \cdot |\mathbf{h}_{k,j}^H \mathbf{w}_j|^2 + P_{\mathrm{noise}}}, \tag{5}$$

where $P_{\mathrm{tx,i}}$ is the associated AP_i transmit power, $\mathbf{h}_{k,i} \in \mathbb{C}^{N_t \times 1}$ is the channel vector from associated AP_i to station k, and the Hermitian transpose $\mathbf{h}_{k,i}^H$ converts the channel vector into a row vector, allowing for the inner product with the beamforming vector $\mathbf{w}$. Similarly, the interference power of other APs is given by $P_{tx,j}$, and $\mathbf{h}_{k,j}$ is the channel vector from interfering AP_j to station k. Summing over all $j \neq i$ captures the total inter-AP interference. The total noise power with Thermal noise and noise figure is calculated as

$$P_{\mathrm{noise}} = -174 + 10 \cdot \log_{10}(B) + 7. \tag{6}$$

where B is the channel bandwidth. The network parameters used in this study are summarized in Table I.

IV. ACCESS POINT GROUPING

In dense Wi-Fi deployments, optimizing spatial reuse and minimizing interference is critical for maintaining network performance. Grouping APs plays a key role in enabling coordinated interference management. To accurately capture the interference relationship between APs, we group the APs based on the interference level, the most M APs in the group, and the number of stations in the group of APs.

First, we choose a master AP and start calculating the interference power from other nearest APs. Secondly, if stations receive interference power greater than a threshold (i.e., higher than –67 dBm [16]) from a neighboring AP are considered significant interferers and are considered in this group. Thirdly, if there are more APs whose interference power is higher than –67 dBm, we set the maximum group of APs to be 7.[1] Fourthly, we check if there are more than 100 stations in the group of APs, it will be considered as a separate group. This design allows the grouping process to consider not only load but also real interference coupling between APs.

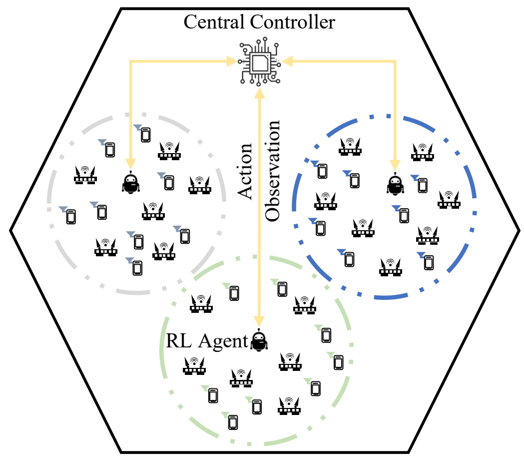


Figure 2. APs grouping and MARL central controller

To ensure robust grouping, the interference metric is averaged over time to avoid decisions based on short-term fluctuations, which results in more stable and reliable group formations. This is particularly beneficial in enterprise networks such as airports and urban areas, where AP deployments are dense, and interference sources change frequently. In our system, interference is handled independently within each group of APs by a central controller as shown in Figure 2.

Grouping before deploying the MARL significantly improves the efficiency of multi-agent learning. It reduces the size of the knowledge for each agent, allowing them to focus only on the relevant local environment. As a result, agents can learn more stable interference patterns and achieve better convergence. Furthermore, grouping prevents redundant use of RL agents. Random grouping could lead to unnecessary agents being deployed, which can increase complexity and overhead.

V. MULTI-AGENT REINFORCEMENT LEARNING

To optimize interference mitigation in dense Wi-Fi deployments, we adopt a MARL framework based on the multi-agent deep deterministic policy gradient (MADDPG) algorithm, enabling scalable and adaptive interference control through coordinated beamforming and null steering. This choice is motivated by the need for decentralized decision-making across multiple groups, while maintaining coordinated learning via a centralized critic. This helps each agent for the group of APs to learn better decisions while still working together with each

[1] In some first simulations, we identified that a max. number of AP per cluster should not exceed 7 in order to maintain reasonable performance.

other. The centralized critic sees all the information during training, so learning is more stable.

The MADDPG algorithm combines the advantages of deep Q-network (DQN) and deterministic policy gradient (DPG), making it effective for continuous control problems [17]. First, the number of agents $\mathbb{A}$ are decided based on the number of groups of APs, then the replay buffer $\mathcal{D} = 10^6$, the learning rate α, the discount factor $\gamma = 0.99$, and the target update rate $\tau = 10^{-3}$ are initialized for MARL.

State: In the MADDPG framework, each agent has access only to local observations of the group of APs. All APs collect CSI from their stations through periodic channel sounding using Equation (1). The local observation for a single group of APs for an agent is structured as

$$\mathcal{S}_{\mathbb{A}} = \big[(\phi, PL)_{s_k, AP_0}, (\Theta, PL)_{s_1 \dots s_n, AP_1}, \\ (\Theta, PL)_{s_1 \dots s_n, AP_2}, \dots, (\Theta, PL)_{s_1 \dots s_n, AP_{M-1}}\big], \quad (7)$$

where $(\phi, PL)_{s_k, AP_0}$ are the beam angle ϕ and path loss PL of the associated station k of AP_0 which is served at time slot t, $(\Theta, PL)_{s_1 \dots s_n, AP_1}$ are the null angles Θ and path losses PL of all non-assoiciated stations $s_1 \to s_n$ of AP_1, AP_2, till AP_{M-1} with respect to AP_0. M represents the number of APs in the group.

Action: After the observation collection from the network environment, each agent responsible for multiple APs in the group predicts an action $\mathcal{A}_{\mathbb{A}}$ corresponding to the number of available nulling DoFs for each AP in the group is defined as

$$\mathcal{A}_{\mathbb{A}} = [\Theta_1, \Theta_2, \dots, \Theta_{(N_t - 1) \times M}], \quad (8)$$

where $\Theta_1, \dots, \Theta_{(N_t-1) \times M}$ denote the specific null angles predicted by the agent for interference suppression for all the APs in a single group. N_{t-1} represents the number of spatial DoF available for nulling, which means the total number of null directions the agent can predict per time slot is $N_{t-1} \times M$. Each action $\mathcal{A}_1, \dots \mathcal{A}_{\mathbb{A}}$ are executed in each corresponding group. The direction of the beam is not predicted by the agent since each AP already knows its intended stations.

Reward: Designing the reward function is crucial for achieving satisfactory performance in DRL solutions. After passing the actions to each group, each AP performs precoding to beam towards its station and null the other directions. The reward function is designed to encourage the maximization of the network throughput while penalizing decisions that cause severe interference. Specifically, the reward is based on the sum-rate of all associated stations, and the total reward $\mathcal{R}_{\mathbb{A}}$ function from the group during the allocated time slot T_{slot} is

$$\mathcal{R}_{\mathbb{A}} = T_{\text{slot}} \sum_k B \cdot \log_2(1 + \text{SINR}). \quad (9)$$

This way the agent will learn to avoid to predict actions for an AP to null the most disruptive stations directions. After executing the actions and calculating the reward, all the information, i.e., local states $\mathcal{S}_{\mathbb{A}}$, actions $\mathcal{A}_{\mathbb{A}}$, and rewards $\mathcal{R}_{\mathbb{A}}$ of each agent are stored in a buffer.

Global State: To ensure cooperation and competition between agents in our MARL framework, we adopt centralized training with decentralized execution, as implemented in the MARL algorithm [18]. During training, the centralized critic has access to all the states $\mathcal{S}$ and actions $\mathcal{A}$, allowing it collaboration between agents. This enables the critic to accurately assess the joint impact of the actions on the environment. A single critic can help stabilize learning by providing a consistent evaluation signal, preventing individual agent biases from negatively impacting overall performance. During execution, each agent sees only its local observation $\mathcal{S}_{\mathbb{A}}$ and its own policy $\pi(\mathcal{A}_{\mathbb{A}}|\mathcal{S}_{\mathbb{A}})$, ensuring that the system can scale and operate under decentralized control without requiring full global knowledge.

Table II
MACHINE LEARNING PARAMETERS

Parameter	**Value**
State and Action space size	Dep. on clusters
Actor / critic hidden layers	[400, 300]
Actor, Critic learning rate	10^{-5}, 10^{-4}
Optimizer	Adam
Batch size	128

The parameters for the actor-critic networks in our MARL framework are summarized in Table II. rectified linear unit (ReLU) activation functions are used in the hidden layers to enable efficient backpropagation and avoid vanishing gradients. The loss function used to train the critic is mean square error (MSE), as commonly applied in value-based RL. We adopt the Ornstein-Uhlenbeck (OU) noise process, which helps maintain exploration diversity and stabilize convergence in continuous control tasks, as originally suggested in [17]. Grouping the APs before assigning agents reduces the dimensionality of each agent's observation and action space, allowing for faster learning and better scalability, and can even be deployed in a large Wi-Fi network.

VI. PERFORMANCE EVALUATION

We evaluate the performance of the proposed MARL framework for interference-aware beamforming in dense Wi-Fi deployments against single-agent reinforcement learning (SARL) and an oracle upper bound. The results demonstrate the scalability and effectiveness of our method, particularly in scenarios with high user density interference.

A. Single-Agent vs. Multi-Agent Reinforcement Learning

To evaluate the effectiveness of our proposed MARL framework, we compare it with the state-of-the-art SARL baseline for coordinated beamforming and interference suppression. The results are shown in Figure 3, which plots the aggregate OBSS throughput per time slot as a function of training episodes. This experiment was performed in a large-scale simulation environment consisting of two groups of APs. Each group contains 7 APs, and each AP is associated with 10 stations,

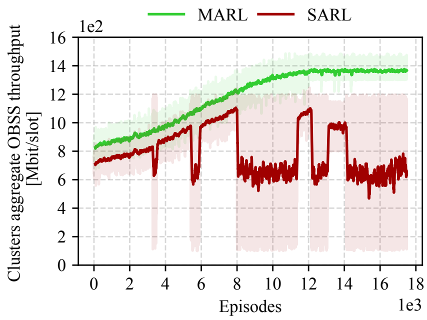


Figure 3. Single-agent vs. multi-agent reinforcement learning

resulting in a total of 14 APs and 140 stations. Every AP is equipped with 6 antennas, enabling one beamforming direction for its own station and up to five spatial nulls to suppress interference toward other stations.

The green curve represents the performance of the MARL approach, while the dark red curve corresponds to the SARL baseline. As shown, MARL not only achieves significantly higher aggregate throughput but also converges more smoothly over training. In contrast, SARL exhibits inconsistent learning behavior, with sharp drops around episodes $4,000$, $8,000$, and $14,000$. These instabilities arise due to the limitations of a single agent attempting to manage a large and complex environment involving many APs (i.e., 14 APs) and stations. This is because the single agent receives a very large observation and action space, making it difficult to extract relevant features and prioritize interference mitigation effectively. It often fails to make coordinated intelligent predictions of nulling decisions across all APs, especially when multiple AP interfere with each other. Our method first groups the APs and then uses MARL to reduce this complexity by assigning an agent per group. Thus, each agent only handles local coordination among at most 7 APs in its group, which simplifies the learning process and allows better performance.

MARL not only ensures better scalability but also allows the system to handle large deployments efficiently while maintaining throughput improvements. These advantages clearly demonstrate the strength of decentralized learning in dense Wi-Fi deployments with coordinated beamforming.

B. Interference handled by MARL

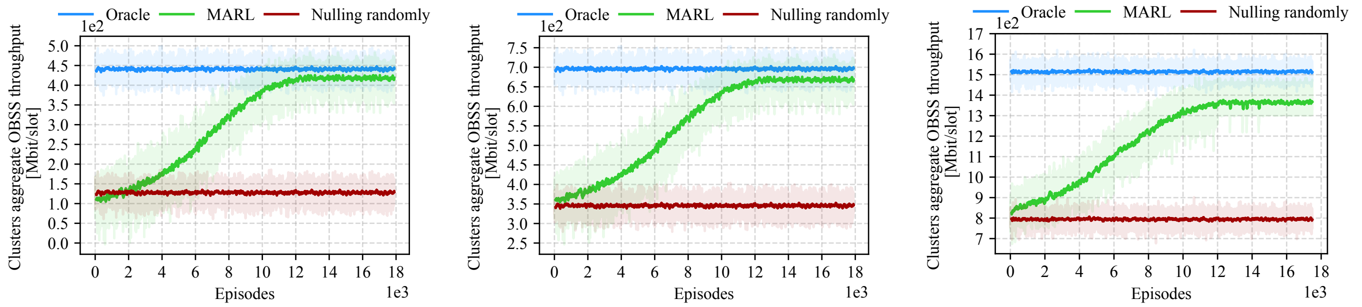
The results presented in Figure 4 compare the aggregate OBSS throughput across three distinct approaches: the proposed MARL method (green), an oracle with perfect CSI and nulling angle knowledge (blue), and a random nulling (red). In Figure 4a, the system comprises two clusters with a total of 13 APs, where the first cluster contains 6 APs and the second cluster 7 APs. Each AP is equipped with 6 antennas and serves between 5–15 stations, reflecting a dense and interference-prone environment. The oracle approach, as expected, achieves the highest throughput, reaching approximately 450 Mbit/slot, due to its ideal knowledge of CSI and optimal nulling angles. This serves as an upper bound for performance comparison. In contrast, the random nulling strategy performs poorly, with throughput limited to around 120 Mbit/slot, as it fails to mitigate interference effectively. The proposed MARL method demonstrates competitive performance, reaching up to 435 Mbit/slot, which is significantly higher than random nulling and approaches the oracle's performance. This improvement stems from the ability to dynamically learn and prioritize high-traffic interference while avoiding unnecessary nulling.

The scalability of the proposed MARL framework is further evaluated in a more complex five-cluster environment, where each cluster contains 3–7 APs, and each AP serves 15 stations (Figure 4b). In this scenario, the oracle maintains its superior performance, achieving throughput levels consistent with its ideal interference management capabilities. The random nulling approach, however, exhibits even poorer performance compared to the Oracle case, with throughput nearly 350 Mbit/slot. The MARL-based approach demonstrates remarkable robustness in this challenging environment. While its throughput does not reach the oracle's peak performance, as MARL makes decision based on the outdated CSI, but it consistently outperforms random nulling and achieving throughput levels around 650 Mbit/slot. The key to this success lies in the distributed nature of the MARL framework, where each agent independently learns to coordinate beamforming decisions based on local observations, thereby scaling effectively with the number of APs.

Finally, we examine the system performance in an ultra-dense seven cluster configuration, where each seven-group contains 4–7 APs with 5–15 stations per AP (Figure 4c). This scenario represents the most challenging interference environment studied, with significantly higher node density and spatial complexity compared to previous configurations. The oracle benchmark achieves stable throughput around 1500 Mbit/slot, demonstrating the theoretical upper bound for interference management in this ultra-dense scenario. Random nulling performance degrades severely in this environment, with throughput around 800 Mbit/slot, highlighting its complete inadequacy for dense deployments.

Several key observations emerge from this seven-cluster scenario. The MARL approach maintains relatively stable performance despite the increased network complexity, demonstrating its robustness to scaling effects. This stability stems from the distributed learning architecture, where each agent adapts to local interference patterns without requiring global network knowledge. The results reveal that the MARL throughput shows greater episode-to-episode stability in this seven-cluster scenario. This suggests the agents develop more generalized policies through exposure to diverse interference patterns during training.

As the number of clusters increases, the performance gap between the MARL approach and the Oracle becomes more pronounced. Although the MARL framework still outperforms single-agent baseline, these results highlight the importance of carefully designing clustering strategies to balance scalability, coordination complexity, and interference mitigation



(a) 2 clusters, each with 6-7 APs and 5-15 stations (b) 5 clusters, each with 3-7 APs and 15 stations (c) 7 clusters, each with 4-7 APs and 5-15 stations

Figure 4. MARL interference handling

effectiveness.

VII. Conclusion

We presented a MARL framework for interference management in dense Wi-Fi networks. By dynamically grouping APs based on real-time interference, AP, and station density, and assigning RL agents per group, our system effectively scales in ultra-dense deployments. The proposed approach improves network throughput by learning to suppress the most disruptive interference directions using coordinated nulling actions. Simulation results demonstrate that MARL outperforms SARL approaches in terms of throughput and learning stability, especially as the network scales. This work highlights the potential of intelligent coordination and grouping in next-generation Wi-Fi systems, and covers the way for adaptive, scalable interference management using ML. In the future, we will further investigate the gap between the MARL and the Oracle approach to reduce it further.

Acknowledgments

This work was supported by the Federal Ministry of Education and Research (BMBF, Germany) within the 6G Research and Innovation Cluster 6G-RIC under Grant 16KISK020K as well as by the German Research Foundation (DFG) within the project ML4WiFi under grant DR 639/28-1.

References

[1] Z. Li, R. Yang, X. Yang, J. Yang, X. Li, H. Lin, F. Qian, Y. Liu, Z. Liao, and D. Hu, “A Four-Year Retrospective of Mobile Access Bandwidth Evolution: The Inspiring, The Frustrating, and The Fluctuating,” *IEEE Transactions on Mobile Computing*, vol. 24, no. 8, pp. 7458–7474, Aug. 2025.

[2] M. Silva, J. Santos, and M. Curado, “The Path Towards Virtualized Wireless Communications: A Survey and Research Challenges,” *Journal of Network and Systems Management*, vol. 32, no. 1, Nov. 2023.

[3] S. Verma, T. K. Rodrigues, Y. Kawamoto, M. M. Fouda, and N. Kato, “A Survey on Multi-AP Coordination Approaches Over Emerging WLANs: Future Directions and Open Challenges,” *IEEE Communications Surveys & Tutorials*, vol. 26, no. 2, pp. 858–889, 2024.

[4] M. Z. Mahfouz, A. B. J. Kokkeler, A. Meijerink, and A. A. Glazunov, “Impact of Ultra-Narrowband Interference on Wi-Fi Links: An Experimental Study,” *IEEE Transactions on Wireless Communications*, vol. 20, no. 5, pp. 3016–3030, May 2021.

[5] A. Tusha and H. Arslan, “Interference Burden in Wireless Communications: A Comprehensive Survey from PHY Layer Perspective,” *IEEE Communications Surveys & Tutorials*, vol. 27, no. 4, pp. 2204–2246, Aug. 2025.

[6] L. Galati-Giordano, G. Geraci, M. Carrascosa, and B. Bellalta, “What will Wi-Fi 8 Be? A Primer on IEEE 802.11bn Ultra High Reliability,” *IEEE Communications Magazine*, vol. 62, no. 8, pp. 126–132, Aug. 2024.

[7] Z. He, “Toward AIML Enabled WiFi Beamforming CSI Feedback Compression: An Overview of IEEE 802.11 Standardization,” arXiv, eess.SP 2503.00412, Mar. 2025.

[8] J. Bacha, A. Zubow, S. Szott, K. Kosek-Szott, and F. Dressler, “Deep Reinforcement Learning based Interference Optimization for Coordinated Beamforming in Ultra-Dense Wi-Fi Networks,” *Elsevier Computer Communications*, vol. 242, p. 108 286, Oct. 2025.

[9] J. Chen, Y. Xiao, X. Lei, Y. Ding, H. Niu, K. Liu, Y. Zhong, S. Yang, V. Fusco, and W. Xiang, “A Survey on Directional Modulation: Opportunities, Challenges, Recent Advances, Implementations, and Future Trends,” *IEEE Internet of Things Journal*, vol. 12, no. 16, pp. 32 581–32 615, Aug. 2025.

[10] K. Bagadi, V. Annepu, N. Raju Challa, F. Benedetto, T. Shongwe, and K. Rabie, “Enhancing Performance of Massive MU-MIMO System With LR-RTS: A Low-Complexity Detection Algorithm,” *IEEE Access*, vol. 12, pp. 97 085–97 098, 2024.

[11] D. Nunez, P. Imputato, S. Avallone, M. Smith, and B. Bellalta, “Enabling Reliable Latency in Wi-Fi 8 Through Multi-AP Joint Scheduling,” *IEEE Open Journal of the Communications Society*, vol. 6, pp. 2090–2101, 2025.

[12] W. Cheng and Z. Yang, “Strengthening the Connection and Monitoring of Dead Zones: Implementing Drone-Based Monitoring Images Using 6G IoT and Deep Learning,” *Springer Wireless Personal Communications*, Jun. 2024.

[13] Z. Guo, Z. Chen, P. Liu, J. Luo, X. Yang, and X. Sun, “Multi-Agent Reinforcement Learning-Based Distributed Channel Access for Next Generation Wireless Networks,” *IEEE Journal on Selected Areas in Communications*, vol. 40, no. 5, pp. 1587–1599, May 2022.

[14] J. Tan, Y.-C. Liang, L. Zhang, and G. Feng, “Deep Reinforcement Learning for Joint Channel Selection and Power Control in D2D Networks,” *IEEE Transactions on Wireless Communications*, vol. 20, no. 2, pp. 1363–1378, Feb. 2021.

[15] M. Wojnar, W. Ciezobka, A. Tomaszewski, P. Chołda, K. Rusek, K. Kosek-Szott, J. Haxhibeqiri, J. Hoebeke, B. Bellalta, A. Zubow, F. Dressler, and S. Szott, “Coordinated Spatial Reuse Scheduling With Machine Learning in IEEE 802.11 MAPC Networks,” *IEEE Journal on Selected Areas in Communications*, Jul. 2025.

[16] M. Natkaniec and N. Bieryt, “An Analysis of the Mixed IEEE 802.11ax Wireless Networks in the 5 GHz Band,” *MDPI Sensors*, vol. 23, no. 10, p. 4964, May 2023.

[17] T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra, “Continuous control with deep reinforcement learning,” arXiv, cs.LG 1509.02971, Sep. 2015.

[18] T. Ikeda and T. Shibuya, “Centralized Training with Decentralized Execution Reinforcement Learning for Cooperative Multi-agent Systems with Communication Delay,” in *IEEE SICE 2022*, Aso, Japan: IEEE, Sep. 2022, pp. 135–140.