Extending Mininet-WiFi with Coordination-Aware MLO for Dense Multi-AP IEEE 802.11bn WLANs

Rania Sahraoui
SAMOVAR, Télécom SudParis
Institut Polytechnique de Paris
Palaiseau, France
rania.sahraoui@telecom-sudparis.eu

Fetia Bannour
SAMOVAR
ensIIE
Évry, France
fetia.bannour@telecom-sudparis.eu

Badii Jouaber
SAMOVAR, Télécom SudParis
Institut Polytechnique de Paris
Palaiseau, France
badii.jouaber@telecom-sudparis.eu

Djamal Zeghlache
SAMOVAR, Télécom SudParis
Institut Polytechnique de Paris
Palaiseau, France
djamal.zeghlache@telecom-sudparis.eu

***Abstract*—The growing adoption of IEEE 802.11 WLANs in industrial and high-density environments is intensifying contention, co-channel interference, and resource-management complexity. IEEE 802.11be (Wi-Fi 7) addresses these challenges through Multi-Link Operation (MLO), which enables a Multi-Link Device (MLD) to exploit multiple wireless links but introduces additional decisions on link activation, traffic distribution, and interference management. The emerging IEEE 802.11bn (Wi-Fi 8) further emphasizes Multi-AP Coordination (MAP-Co) to improve reliability and spectrum utilization across neighboring APs. As MLO support has not yet been incorporated into Mininet-WiFi, we first extend the platform with a bridge-relay-based MLO module that aggregates multiple wireless interfaces under a common U-MAC and supports EMLSR and EMLMR operation. Building on this extension, we propose an Agentic Coordination-Aware Multi-Link Operation scheme (A-CAMLO) that couples local traffic steering with network-wide interference mitigation. An MLO agent adapts link activation and traffic allocation at each AP, while a MAP-Co agent derives admissible per-link transmit-power bounds. Their decisions and observed outcomes are exchanged through a closed feedback loop to align local MLO adaptation with multi-AP coordination objectives. Experiments under increasing contention and overlapping-BSS interference show that, under overload, A-CAMLO achieves at least 13.5% higher median throughput than the other evaluated baselines. Under high interference, coordination increases A-CAMLO's energy efficiency to more than twice that of its uncoordinated operation, while reducing packet loss to nearly zero. These results suggest that A-CAMLO provides a favorable balance between throughput, energy efficiency, and reliability in dense WLANs.**

***Index Terms*—Wi-Fi 7; Wi-Fi 8; IEEE 802.11be; Multi-Link Operation; Multi-AP Coordination; Mininet-WiFi; DRL; SDN**

I. Introduction

Wireless local area networks are increasingly expected to support demanding applications such as immersive media, industrial automation, real-time monitoring, and latency-sensitive communications. These requirements continue to drive the evolution of IEEE 802.11 technologies, from IEEE 802.11be (Wi-Fi 7) toward the emerging IEEE 802.11bn amendment, commonly associated with Wi-Fi 8. While IEEE 802.11be introduces MLO as a key mechanism for improving throughput, latency, and reliability, the emerging IEEE 802.11bn amendment builds upon this multi-link foundation and places stronger emphasis on ultra-high reliability and coordinated operation among neighboring APs [1], [2].

A key Wi-Fi 7 feature is Multi-Link Operation (MLO), which enables a Multi-Link Device (MLD) to operate over multiple frequency links under a single logical association. By supporting parallel transmissions and flexible traffic distribution, MLO can improve throughput, latency, and resilience to congestion. Its performance depends on both radio configuration and traffic steering. Enhanced Multi-Link Multi-Radio (EMLMR) devices can use multiple links within the same scheduling interval, whereas Enhanced Multi-Link Single-Radio (EMLSR) devices use a single radio that switches among links. These configurations therefore introduce different trade-offs in throughput, latency, hardware complexity, energy consumption, and interference exposure [3], [4].

Efficient MLO operation requires adaptive decisions on link activation and traffic distribution, since static allocation may perform poorly under heterogeneous or time-varying congestion and channel conditions. Recent studies have investigated reinforcement learning, multi-armed bandits, and QoS-aware optimization for adaptive MLO traffic steering [5], [6], while continuous MLO has also been explored to improve spectrum utilization in congested environments [7].

The interaction between MLO and multi-AP deployments is also important. Inter-AP interference and coexistence with legacy Wi-Fi can significantly affect MLO performance, motivating coordinated resource management across neighboring APs [8]. MLO decisions are therefore coupled with network-wide interference conditions.

Despite these advances, most MLO studies rely on analytical or simulation-based frameworks, with limited support for integrated system-level implementations. In particular, Mininet-WiFi [9] does not provide native MLO support, including a unified MLD abstraction and coordination-aware multi-link traffic steering. This motivates extending Mininet-WiFi with controllable MLD capabilities.

To address this gap, we design, implement, and evaluate an MLO-capable Mininet-WiFi framework for dense and coordinated WLANs. The main contributions are as follows:

- We extend Mininet-WiFi with a software MLD abstraction that aggregates multiple wireless interfaces under a common U-MAC. The bridge-relay design emulates the principal data-plane behavior of EMLSR and EMLMR through configurable single and multi-radio link usage.

979-8-3195-0167-7/26/$31.00 ©2026 IEEE

- We propose a coordination-aware MLO framework that combines intelligent traffic steering with controller-assisted coordination. This design enables MLO decisions to adapt to network-wide conditions and strengthens their effectiveness in dense multi-AP WLANs.
- We conduct extensive simulations to investigate: (1) the behavior of different MLO configurations and policies under different contention levels; (2) the performance of the best-performing MLO policies under different interference levels to assess MLO in dense WLANs; and (3) the gains brought by coordination.

The remainder of this paper presents related work, the system model and emulation architecture, the coordination-aware MLO formulation and learning scheme, followed by the performance evaluation and conclusions.

II. Background and Related Work

Recent studies show that MLO performance strongly depends on the mechanisms used for link selection and traffic distribution. López-Raventós and Bellalta showed that dynamic, congestion-aware allocation can improve traffic delivery compared with static policies [10]. Experimental results similarly indicate that MLO can reduce latency when links experience balanced occupancy, but may degrade performance when traffic is assigned to a congested link [11]. Zou et al. further showed that traffic-to-link allocation directly affects delay in MLO/SLO coexistence scenarios [12]. Learning-based scheduling has also been explored, Abdelreheem et al. proposed an attention-enhanced DQN scheduler for synchronous N-STR MLDs to improve throughput, delay, fairness, and packet delivery [13].

A. Multi-AP Coordination and Centralized WLAN Control

MLO decisions become more complex in dense overlapping BSSs as the links selected by one MLD-AP affect the contention and interference experienced by neighboring APs. Activating additional links may improve local transmission opportunities, but can also increase airtime occupancy, inter-BSS contention, and interference exposure on channels shared with other BSSs. Consequently, uncoordinated local MLO decisions may degrade network-wide performance. Multi-AP coordination addresses this coupling by using network-wide information to coordinate channel access, spatial reuse, transmit power, and transmission scheduling [14]. Several studies investigate explicit coordination mechanisms. Zhang et al. proposed a centralized AP-controller architecture with learning-based channel access and proportional-fair resource sharing [15]. Nunez et al. combined TXOP sharing with coordinated spatial reuse, jointly selecting concurrent transmissions and transmit-power levels [16]. Haxhibeqiri et al. implemented a centralized coordinated-spatial-reuse controller that configures AP transmit power, modulation, and coding schemes for concurrent transmissions [17]. These works establish the benefits of direct multi-AP coordination, but they do not jointly manage MLO link activation and continuous U-MAC traffic allocation.

B. Joint MLO-Multi-AP Coordination and Emulation Gap

Recent studies have begun to jointly consider MLO and coordination across overlapping BSSs. Zhang et al. used a centralized AP controller to optimize AP-STA pairing and proportional-fair MLO radio-link allocation [18]. Ali and Bellalta proposed a federated reinforcement-learning framework in which neighboring BSSs collaboratively learn their MLO link-activation policies [19]. Kumar et al. introduced a graph-driven multi-AP scheduler that distributes MLO traffic across links while accounting for the loads generated in overlapping BSSs [20].

These studies show that MLO has been combined with coordination mainly through centralized or collaborative scheduling, where link activation or traffic allocation is optimized across multiple APs rather than decided independently at each AP. Although such approaches capture part of the coupling between MLO and neighboring BSSs, they generally optimize only a limited subset of the multi-AP coordination problem and do not jointly integrate controller-level coordination with local U-MAC traffic allocation in dense WLANs. In contrast, our framework combines multi-AP coordination with local U-MAC traffic allocation through a two-level control architecture. The controller determines network-wide coordination decisions, while each MLD-AP locally adapts transmit power, link activation, and traffic distribution. This control loop is realized within an executable Mininet-WiFi MLD emulation.

III. System Model and Coordination-Aware MLO Emulation

We consider an IEEE 802.11bn Wi-Fi 8 WLAN composed of a set of Multi-Link Access Points (MLD-APs), denoted by $\mathcal{N} = \{1, \ldots, N\}$, and a set of Multi-Link Stations (MLD-STAs), denoted by $\mathcal{M} = \{1, \ldots, M\}$. Each MLD-AP operates over a set of wireless links $\mathcal{F} = \{1, \ldots, F\}$, corresponding for instance to the 2.4 GHz, 5 GHz, and 6 GHz bands.

Each MLD follows the IEEE 802.11be MLO model, where one Upper MAC (U-MAC) coordinates transmissions over multiple Lower MACs (L-MACs). Each L-MAC is associated with one wireless interface and collects link-level measurements, while the U-MAC performs traffic steering across the available links. In this work, we emulate this architecture in Mininet-WiFi and extend it with an SDN-based coordination layer. The resulting model captures both local MLO traffic steering at each MLD-AP and the multi-AP coordination required to mitigate interference in dense deployments.

This section first describes the bridge-relay MLD emulation, then introduces the link and traffic-steering model, the considered MLO operating modes, and finally the coordination-aware multi-AP interference model used by the SDN controller.

A. MLO Emulation Architecture

In IEEE 802.11be, MLO relies on a logical MLD that aggregates multiple wireless interfaces under a single entity. This abstraction is not yet available in Mininet-WiFi, where APs and stations do not natively support coordinated simultaneous operation over multiple frequency links. To

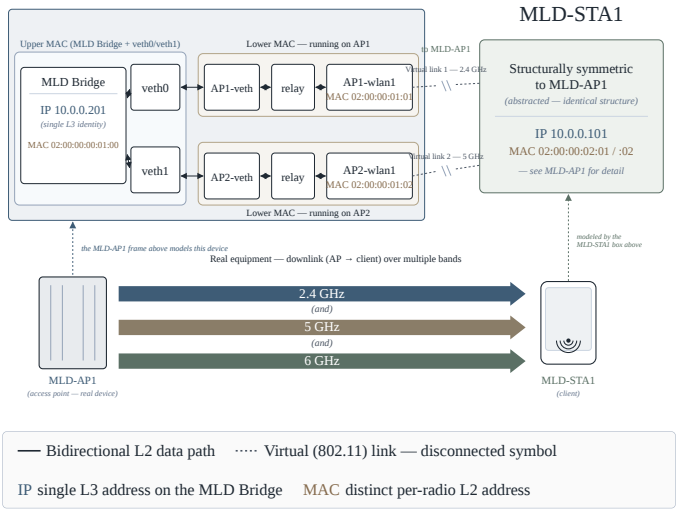


Fig. 1. Mininet-WiFi MLD emulation architecture.

reconstruct the MLD abstraction, we adopt the bridge-relay architecture illustrated in Fig. 1. Each MLD-AP or MLD-STA is implemented as a software Upper MAC (U-MAC) placed above multiple Lower MACs (L-MACs). The U-MAC is realized as a Linux host that aggregates the wireless interfaces through a central Linux bridge, which provides a unified logical identity for the MLD; each wireless interface represents one L-MAC and is attached to the bridge through a virtual Ethernet pair, namely $(\texttt{ap-veth}, \texttt{ap-wlan})$ for an MLD-AP and $(\texttt{sta-veth}, \texttt{sta-wlan})$ for an MLD-STA. However, Mininet-WiFi provides no default Layer-2 forwarding between the virtual Ethernet and wireless interfaces of such a pair, even though both are instantiated within the same network namespace. We therefore introduce a relay component that bridges this gap bidirectionally: outbound frames are forwarded from the virtual Ethernet side to the wireless side $(\texttt{veth}\rightarrow\texttt{wlan})$, and inbound frames in the reverse direction $(\texttt{wlan}\rightarrow\texttt{veth})$. During this process, the relay rewrites MAC addresses where needed so that packets traverse the correct wireless link while preserving the abstraction of a single aggregated MLD entity at the bridge/U-MAC level. This bridge-relay design preserves the semantics of independent wireless interfaces while enabling unified control at the U-MAC layer, which can thus steer traffic across the MLO links according to the selected operating mode. The source code is available at https://gitlab.com/sahraouirania389/mlo-mininet-wifi.

B. Link Model

Let $x_{n,m}[t] \in \{0,1\}$ denote the association between MLD-AP n and MLD-STA m at time slot t, where $x_{n,m}[t] = 1$ if STA m is associated with AP n, and 0 otherwise. Let B_f denote the bandwidth of link f. The achievable rate between AP n and STA m on link f is modeled as

$$C_{f,n,m}[t] = B_f \log_2 \left(1 + \mathrm{SINR}_{f,n,m}[t]\right), \tag{1}$$

where $\mathrm{SINR}_{f,n,m}[t]$ is the instantaneous signal-to-interference-plus-noise ratio observed on link (f, n, m).

We define a binary link-availability indicator $I_{f,n}[t] \in \{0,1\}$ as

$$I_{f,n}[t] = \begin{cases} 1, & \text{if link } f \text{ of AP } n \text{ is available at time } t, \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

The set of available MLO links for AP n is therefore

$$\mathcal{F}_n[t] = \{f \in \mathcal{F} \mid I_{f,n}[t] = 1\}. \tag{3}$$

C. U-MAC Traffic and Performance Model

When MLD-AP n has downlink traffic for an associated STA m, the U-MAC distributes the traffic over the available MLO links. Let $\rho_{m,n,f}[t] \in [0,1]$ denote the fraction of the traffic destined to STA m that is assigned by AP n to link f. Its feasibility with respect to AP-STA association, link availability, and traffic conservation is specified in Section IV.

Let $d_m[t]$ denote the traffic demand of STA m. The traffic rate assigned to link f is

$$\lambda_{m,n,f}[t] = \rho_{m,n,f}[t] d_m[t]. \tag{4}$$

The effectively served rate is limited by the assigned rate and the achievable link capacity:

$$\mu_{m,n,f}[t] = \min\left\{\lambda_{m,n,f}[t], C_{f,n,m}[t]\right\}. \tag{5}$$

The delay metric is defined as the flow-completion time of the complete traffic transfer from MLD-AP n to STA m:

$$D_{n,m}[t] = T^{\text{last-rx}}_{n,m}[t] - T^{\text{first-tx}}_{n,m}[t], \tag{6}$$

where $T^{\text{first-tx}}_{n,m}[t]$ denotes the transmission time of the first byte, and $T^{\text{last-rx}}_{n,m}[t]$ denotes the reception time of the last byte of the transfer across all active MLO links. The index t identifies the scheduling epoch in which the transfer is initiated, although the transfer may be completed in a subsequent epoch.

The total rate served by AP n over its associated STAs and MLO links is

$$\mu_n[t] = \sum_{m \in \mathcal{M}} \sum_{f \in \mathcal{F}} x_{n,m}[t] \mu_{m,n,f}[t]. \tag{7}$$

These quantities characterize the traffic delivered by the emulated U-MAC. Their use in the coordination-aware traffic-steering objective is presented in Section IV.

D. MLO Operating Modes

IEEE 802.11be supports different MLO configurations depending on the radio capabilities of the MLD and its simultaneous-transmission constraints [21], [22]. In this work, MLO operation is characterized by the radio configuration and the traffic-steering policy.

We consider Enhanced Multi-Link Multi-Radio (EMLMR) and Enhanced Multi-Link Single-Radio (EMLSR). EMLMR uses multiple radio interfaces and can exploit several links within the same scheduling interval, including STR-capable operation when simultaneous transmission and reception are supported [22]. In contrast, EMLSR relies on a single main radio that switches among available links [21], and is therefore used here as a single-radio N-STR baseline.

Static-MLO, Heuristic-MLO, and A-CAMLO are implemented as EMLMR traffic-steering policies, allowing traffic to be distributed across multiple active links, whereas the EMLSR baseline uses one link per transmission opportunity.

E. SDN-Based MAP-Co Coordination Context

In dense WLAN deployments, neighboring MLD-APs may operate on overlapping channels and create inter-AP interference [14]–[16]. This effect is amplified with MLO because an MLD-AP may occupy multiple frequency links, increasing the contention and interference experienced by neighboring APs. Local MLO traffic steering should therefore account for the surrounding multi-AP conditions.

In the proposed framework, the SDN controller collects per-link SINR, inter-AP RSSI, and AP-load measurements from the MLD-APs. Based on this network-wide information, the MAP-Co agent selects AP-STA associations and derives admissible per-link transmit-power bounds, within which the local Power-Control Agent selects the actual power, i.e., $P^{f,n}_{tx,\min,\mathrm{dBm}}[t] \le P^{f,n}_{tx,\mathrm{dBm}}[t] \le P^{f,n}_{tx,\max,\mathrm{dBm}}[t]$, $\forall f,n,t$. The achievable rate $C_{f,n,m}[t]$ in Eq. (1) captures inter-AP interference through the measured SINR, so stronger interference reduces the rate of the affected link. The controller output provided to MLD-AP n is the MAP-Co context

$$\mathbf{c}^{MAP}_n[t] = \left\{ x_{n,m}[t], P^{f,n}_{tx,\min,\mathrm{dBm}}[t], P^{f,n}_{tx,\max,\mathrm{dBm}}[t] \right\}_{m\in\mathcal{M},\, f\in\mathcal{F}}, \tag{8}$$

where $x_{n,m}[t]$ defines the AP-STA association and the power pair defines the admissible transmit-power interval for link f of AP n.

The resulting link conditions are provided to the U-MAC Traffic-Steering Agent through the link-level energy-efficiency indicator

$$\mathrm{EE}^{\mathrm{link}}_{f,n,m}[t] = \frac{C_{f,n,m}[t]}{P^{f,n}_c + p^{f,n}_{tx}[t]}, \tag{9}$$

where $P^{f,n}_c$ is the circuit power of the radio interface of AP n on link f, i.e., the achievable rate per unit of total radio-link power. Here, $\mathrm{EE}^{\mathrm{link}}_{f,n,m}[t]$ is used as a link-quality indicator based on the achievable rate $C_{f,n,m}[t]$, whereas the aggregate MLO energy efficiency in Eq. (10) is computed from the actually served rate $\mu_{m,n,f}[t]$. Consequently, the local traffic-steering decision is conditioned on the controller-provided association and power bounds, as well as on the SINR, rate, and energy-efficiency measurements resulting from the locally selected transmit powers; the corresponding coordination-aware problem is formulated in Section IV.

IV. Coordination-Aware MLO Traffic-Steering Formulation

Using the coordination context and link-level measurements defined above, each MLD-AP performs local U-MAC traffic steering at scheduling epoch t. The resulting optimization jointly determines the active MLO links and the traffic allocation while respecting the controller-defined association and transmit-power constraints.

A. Traffic-Allocation Problem

Under EMLMR operation, every configured link satisfying $I_{f,n}[t] = 1$ is eligible for simultaneous operation. The A-CAMLO policy selects a subset of the available links according to the current channel, traffic, and coordination conditions.

Let $a^{\mathrm{bin}}_{m,n,f}[t] \in \{0,1\}$ denote the learned link-activation decision for traffic transmitted from MLD-AP n to STA m over link f. In particular, $a^{\mathrm{bin}}_{m,n,f}[t] = 1$ indicates that link f is activated for the considered AP-STA transfer, whereas $a^{\mathrm{bin}}_{m,n,f}[t] = 0$ indicates that the link is not used.

Let $\rho_{m,n,f}[t] \in [0,1]$ denote the fraction of the traffic demand of STA m assigned to link f. The traffic-steering decision is therefore hybrid and contains both the binary activation variables and the continuous traffic-allocation ratios.

The aggregate MLO energy efficiency of the transfer between AP n and STA m is defined as

$$\mathrm{EE}^{\mathrm{MLO}}_{n,m}[t] = \frac{\sum_{f\in\mathcal{F}} \mu_{m,n,f}[t]}{\sum_{f\in\mathcal{F}} a^{\mathrm{bin}}_{m,n,f}[t]\left(P^{f,n}_c + p^{f,n}_{tx}[t]\right)}, \tag{10}$$

where $\mu_{m,n,f}[t]$ is the effectively served rate defined in Eq. (5), $P^{f,n}_c$ is the circuit power consumed by the radio interface associated with link f, and $p^{f,n}_{tx}[t]$ is the corresponding transmit power expressed in watts. Therefore, only links selected by the A-CAMLO policy contribute to the total radio-link power consumption. The instantaneous utility of an AP-STA transfer combines aggregate MLO energy efficiency with the flow-completion time defined in Eq. (6):

$$\mathcal{U}_{n,m}[t] = \theta_{\mathrm{EE}}\mathrm{EE}^{\mathrm{MLO}}_{n,m}[t] - \theta_D D_{n,m}[t], \tag{11}$$

where $\theta_{\mathrm{EE}} = 0.5$ and $\theta_D = 0.5$ control the relative scaling of energy efficiency and flow-completion time.

The coordination-aware MLO traffic-steering problem is formulated as

$$\max_{\{a^{\mathrm{bin}}_{m,n,f}[t],\rho_{m,n,f}[t]\}} \sum_{t=1}^{T}\sum_{n\in\mathcal{N}}\sum_{m\in\mathcal{M}} x_{n,m}[t]\mathcal{U}_{n,m}[t]. \tag{12}$$

The optimization is subject to

$$\begin{aligned}
&\text{(C1)} && \sum_{f\in\mathcal{F}} \rho_{m,n,f}[t] = x_{n,m}[t], && \forall m,n,t,\\
&\text{(C2)} && 0 \le \rho_{m,n,f}[t] \le a^{\mathrm{bin}}_{m,n,f}[t], && \forall m,n,f,t,\\
&\text{(C3)} && a^{\mathrm{bin}}_{m,n,f}[t] \le x_{n,m}[t] I_{f,n}[t], && \forall m,n,f,t,\\
&\text{(C4)} && x_{n,m}[t] \le \sum_{f\in\mathcal{F}} a^{\mathrm{bin}}_{m,n,f}[t] \le F x_{n,m}[t], && \forall m,n,t,\\
&\text{(C5)} && a^{\mathrm{bin}}_{m,n,f}[t] \in \{0,1\}, && \forall m,n,f,t.
\end{aligned} \tag{13}$$

Constraint (C1) ensures traffic conservation for every associated AP-STA pair and forces the traffic ratios to zero for non-associated pairs. Constraint (C2) allows traffic to be assigned only to links activated by the A-CAMLO policy. Constraint (C3) prevents the policy from activating a link that is unavailable or belongs to a non-associated AP-STA pair. Constraint (C4) requires the policy to activate at least one link for each associated STA, while allowing up to all F available EMLMR links to be selected. Constraint (C5) defines the binary activation variables. The assigned rate $\lambda_{m,n,f}[t]$ and the effectively served rate $\mu_{m,n,f}[t]$ are defined in Eqs. (4) and (5), respectively, and are thus omitted from the constraint set.

The formulation is coordination-aware because the association and admissible transmit-power bounds are supplied by the MAP-Co agent, whereas the activation and traffic-allocation decisions are learned locally by the A-CAMLO policy.

B. Traffic-Steering MDP

The hybrid traffic-steering problem is modeled as an MDP. In this section, the subscript TS denotes the Traffic-Steering Agent. For each STA m associated with AP n, the corresponding observation token is

$$\mathbf{o}_{m,n}[t] = \left[d_m[t], \left\{ \mathrm{SINR}_{f,n,m}[t], \mathrm{EE}^{\mathrm{link}}_{f,n,m}[t], I_{f,n}[t] \right\}_{f\in\mathcal{F}} \right]. \tag{14}$$

The state observed by the Traffic-Steering Agent at AP n contains the observation tokens of the STAs served in the current scheduling window:

$$\mathbf{s}_{TS,n}[t] = \{\mathbf{o}_{m,n}[t] \mid x_{n,m}[t] = 1\}. \tag{15}$$

The hybrid action contains the link-activation and traffic-allocation decisions for the associated STAs:

$$\mathbf{a}_{TS,n}[t] = \left[\left\{ a^{\mathrm{bin}}_{m,n,f}[t] \right\}_{m,f}, \left\{ \rho_{m,n,f}[t] \right\}_{m,f} \right]. \tag{16}$$

The binary component determines the subset of available EMLMR links used for each AP-STA transfer, while the continuous component distributes the complete traffic demand over the selected links.

The local reward of the Traffic-Steering Agent at AP n aggregates the utilities of the STAs served by that AP:

$$r_{TS,n}[t] = \sum_{m\in\mathcal{M}} x_{n,m}[t]\mathcal{U}_{n,m}[t]. \tag{17}$$

The transition dynamics depend on traffic arrivals, channel variations, link availability, neighboring-AP activity, MAP-Co decisions, actual transmit-power levels, and the hybrid activation and traffic-allocation actions applied by the A-CAMLO policy.

When MAP-Co coordination is disabled, the association and power configuration remain uncoordinated and the hybrid policy corresponds to uncoordinated A-CAMLO. When MAP-Co coordination is enabled, the same A-CAMLO policy operates under the controller-provided association and power bounds, yielding the proposed coordination-aware A-CAMLO scheme. The learning and action-generation mechanisms are presented in Section V.

V. Agentic PPO-Based Coordination-Aware AP-MLO Scheme

This section presents the A-CAMLO learning architecture for solving the coordination-aware MLO traffic-steering problem in Section IV. As shown in Fig. 2, the controller-level MAP-Co agent coordinates an AP-local U-MAC team composed of Power-Control and Traffic-Steering Agents. Using the state $\mathbf{s}_{TS}[t]$ in Eq. (15), the Traffic-Steering Agent selects the hybrid action $\mathbf{a}_{TS}[t]$ in Eq. (16), comprising link activation and traffic-allocation decisions. Its Transformer-based actor is trained with PPO under centralized training and decentralized execution (CTDE).

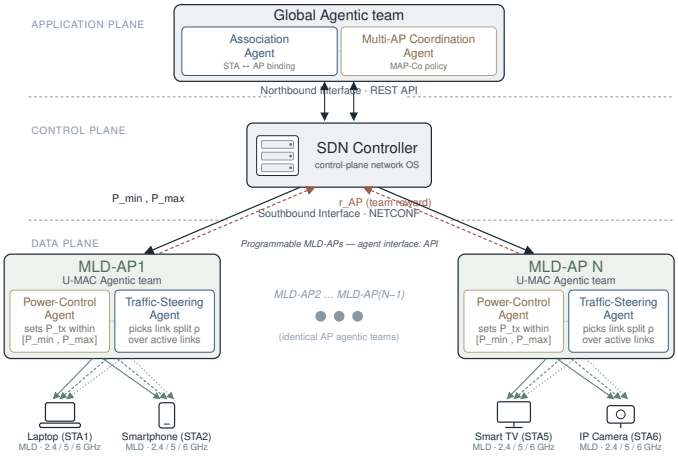


Fig. 2. A-CAMLO architecture coupling SDN-assisted MAP-Co coordination with local power control and hybrid MLO traffic steering.

A. AP-Team Observation Encoding

The Traffic-Steering Agent receives the local state $\mathbf{s}_{TS}[t]$ defined in Eq. (15). For a scheduling window containing N_w associated STAs, the per-STA observations are organized as a token matrix: $\mathbf{O}_{TS}[t] \in \mathbb{R}^{N_w \times d_T}$, where each row corresponds to one STA and d_T is the feature dimension. The link-quality features are measured after the local Power-Control Agent has applied its transmit-power decision. This representation allows the policy to capture dependencies among the STAs served by the same MLD-AP.

B. Transformer-Based Agentic Traffic Policy

Using the token representation $\mathbf{O}_{TS}[t]$ defined above, the Traffic-Steering Agent projects each STA token into a latent feature space. The resulting embeddings are processed by a Transformer encoder, whose self-attention mechanism captures interactions among the STAs served by the same MLD-AP. The encoded representation of each STA is then passed through a shared policy head to generate its traffic-steering output. This architecture enables joint consideration of the per-STA traffic demands and link conditions while preserving a separate decision for each STA.

$$\mathbf{Y}_{TS}[t] = \mathrm{MLP}\big(\mathrm{TrEnc}\big(\phi(\mathbf{O}_{TS}[t]\mathbf{W}_e + \mathbf{b}_e)\big)\big) \in \mathbb{R}^{N_w \times 2F}, \tag{18}$$

where $\mathrm{TrEnc}(\cdot)$ denotes the Transformer encoder, $\mathbf{W}_e$ and $\mathbf{b}_e$ are trainable parameters, and $\phi(\cdot)$ is a nonlinear activation. For each STA, the output is decomposed as $\mathbf{Y}_{TS}[t] = [\mathbf{z}^{act}[t], \mathbf{z}^{share}[t]]$, where $\mathbf{z}^{act}[t] \in \mathbb{R}^{N_w \times F}$ determines the candidate active links and $\mathbf{z}^{share}[t] \in \mathbb{R}^{N_w \times F}$ the traffic distribution over those links.

C. Hybrid Bernoulli-Dirichlet Action with Validation

The action is hybrid, comprising a binary link-activation decision $a^{\mathrm{bin}}_{m,n,f}[t] \in \{0,1\}$ and a continuous traffic-allocation vector. The activation decision is sampled as

$$a^{\mathrm{bin}}_{m,n,f}[t] \sim \mathrm{Bernoulli}\left(\sigma(z^{\mathrm{act}}_{m,f}[t])\right), \tag{19}$$

where $\sigma(\cdot)$ is the sigmoid function. The sampled decisions are masked by the link-availability indicators and validated to ensure that at least one available link remains active for each associated AP-STA pair.

Given the selected links, the Dirichlet concentration parameters are obtained from the traffic-share logits as $\alpha_{m,f}[t] = \mathrm{softplus}(z^{\mathrm{share}}_{m,f}[t]) + \epsilon$, with $\epsilon > 0$ for numerical stability. The traffic ratios are then sampled over the active-link subset:

$$\boldsymbol{\rho}_{m,n}[t] \sim \mathrm{Dirichlet}\left(\boldsymbol{\alpha}_m[t]\right). \tag{20}$$

Inactive links receive zero traffic, while the ratios over the selected links satisfy $\sum_{f\in\mathcal{F}} \rho_{m,n,f}[t] = 1$ and $\rho_{m,n,f}[t] \in [0,1]$. During evaluation, the deterministic Dirichlet mean is used over the active links. The complete hybrid action is $\mathbf{a}_{\mathrm{TS}}[t] = [\mathbf{a}^{\mathrm{bin}}[t], \boldsymbol{\rho}[t]]$.

D. Cooperative PPO with Team and Individual Critics

The AP-local controller comprises the Power-Control Agent P and the Traffic-Steering Agent TS. Each agent $i \in \mathcal{I} = \{P, \mathrm{TS}\}$ maintains an actor $\pi_{\theta_i}(a_i[t] \mid o_i[t])$, and the joint policy factorizes as

$$\pi_{\boldsymbol{\theta}}(\mathbf{a}[t] \mid \mathbf{s}[t]) = \prod_{i\in\mathcal{I}} \pi_{\theta_i}(a_i[t] \mid o_i[t]). \tag{21}$$

Training follows centralized training with decentralized execution. A team critic $Q_\psi(\mathbf{s}[t], \mathbf{a}[t])$ evaluates the joint decisions using the AP-team reward. Its TD(λ) target is

$$\begin{aligned} G^\lambda[t] = r_{\mathrm{AP}}[t] + \gamma(1-d_t)\big[&(1-\lambda)Q_{\psi^-}[t+1] \\ &+ \lambda G^\lambda[t+1]\big], \end{aligned} \tag{22}$$

and the corresponding critic loss and team surrogate are

$$\mathcal{L}_{\mathrm{team}} = \mathbb{E}_t\left[\mathrm{Huber}\big(Q_\psi[t] - G^\lambda[t]\big)\right], \tag{23}$$

$$\widehat{A}_{\mathrm{AP}}[t] = G^\lambda[t] - Q_\psi[t]. \tag{24}$$

Each agent also maintains an auxiliary individual critic $Q^i_{\omega_i}(o_i[t], a_i[t])$, trained using an individual TD(λ) return $G^\lambda_i[t]$ computed from its local reward:

$$\mathcal{L}^{\mathrm{ind}}_i = \mathbb{E}_t\left[\mathrm{Huber}\big(Q^i_{\omega_i}[t] - G^\lambda_i[t]\big)\right]. \tag{25}$$

For each actor, the PPO probability ratio is

$$\chi_i[t] = \frac{\pi_{\theta_i}(a_i[t] \mid o_i[t])}{\pi_{\theta^{\mathrm{old}}_i}(a_i[t] \mid o_i[t])}. \tag{26}$$

The joint actor loss is

$$\begin{aligned} \mathcal{L}_{\mathrm{actor}} = -\sum_{i\in\mathcal{I}} \mathbb{E}_t\Big[&\min\big(\chi_i[t]\widehat{A}_{\mathrm{AP}}[t], \\ &\bar{\chi}_i[t]\widehat{A}_{\mathrm{AP}}[t]\big) + \eta\mathcal{H}_i\Big], \end{aligned} \tag{27}$$

where $\bar{\chi}_i[t] = \mathrm{clip}(\chi_i[t], 1-\epsilon_{\mathrm{ppo}}, 1+\epsilon_{\mathrm{ppo}})$. The Power-Control actor uses a squashed Gaussian policy, while the Traffic-Steering actor uses a Bernoulli-Dirichlet policy. Training reuses each batch for K PPO epochs with shuffled minibatches and gradient clipping.

Algorithm 1 Cooperative PPO-Based AP-MLO Control

Require: Episodes E; horizon T; PPO epochs K; clip ϵ_{ppo}; discount γ; TD parameter λ
Ensure: Learned policies $\pi^*_{\theta_P}$ and $\pi^*_{\theta_{\mathrm{TS}}}$
1: Initialize actors π_{θ_P} and $\pi_{\theta_{\mathrm{TS}}}$; team critic Q^{AP}_ψ; target critic $Q^{\mathrm{AP}}_{\psi^-} \leftarrow Q^{\mathrm{AP}}_\psi$; individual critics $Q^P_{\omega_P}$ and $Q^{\mathrm{TS}}_{\omega_{\mathrm{TS}}}$
2: **for** $e = 1, \ldots, E$ **do**
3: Reset environment; clear buffer $\mathcal{B}_e$
4: **for** $t = 1, \ldots, T$ **do**
5: Compute association and MAP-Co decisions
6: Select power action within MAP-Co bounds
7: Generate and apply the validated MLO traffic action
8: Compute the individual and AP-team rewards
9: Store transition and behavior log-probabilities in $\mathcal{B}_e$
10: **end for**
11: Compute G^λ_t and $\widehat{A}_{\mathrm{AP}}[t] = G^\lambda_t - Q^{\mathrm{AP}}_\psi(\mathbf{s}_{\mathrm{AP}}[t], \mathbf{a}_{\mathrm{AP}}[t])$
12: Update Q^{AP}_ψ using the Huber loss; update target critic $Q^{\mathrm{AP}}_{\psi^-}$
13: Update $Q^P_{\omega_P}$ and $Q^{\mathrm{TS}}_{\omega_{\mathrm{TS}}}$ using their individual losses
14: Update π_{θ_P} and $\pi_{\theta_{\mathrm{TS}}}$ for K PPO epochs over $\mathcal{B}_e$
15: Report AP-team feedback to the controller agents
16: **end for**

E. Learning Procedure

Algorithm 1 summarizes the learning procedure. At each step, the controller team issues association and MAP-Co decisions, after which each AP orchestrator performs local adaptation: the Power-Control Agent sets the transmit power within the controller-defined bounds, and the Traffic-Steering Agent computes a validated MLO traffic split applied by the U-MAC. Lines 5–9 correspond to decision execution, reward computation, and transition collection, while Lines 11–14 perform advantage computation and the critic and actor updates. The resulting AP-team reward drives the update of the shared AP-team critic and both local actors, while aggregated feedback is reported to the controller team in Line 15 so that future MAP-Co decisions account for the local impact of traffic steering.

VI. PERFORMANCE EVALUATION

Mininet-WiFi is extended with an IEEE 802.11bn MLO architecture that supports additional frequency bands, including 6 GHz. This evaluation considers two links operating at 2.4 and 5 GHz. A Python application hosts the association and MAP-Co agents and obtains network measurements from the SDN controller through RESTCONF. The controller configures the AP-local Power-Control and Traffic-Steering Agents through a NETCONF southbound interface. Training uses two MLD-APs and 50 MLD-STAs. A-CAMLO is trained over 50-step episodes using Adam, with actor and critic learning rates of 10^{-3}, batch size 32, discount factor $\gamma = 0.95$, TD parameter $\lambda = \mathbf{0.95}$, and target-critic soft-update coefficient $\tau = 0.1$. Evaluation considers single-AP contention and multi-AP co-channel interference. Experiment 1 compares five MLO modes, whereas Experiment 2 focuses on H-MLO and A-CAMLO and reports results with MAP-Co disabled (*Off*) and enabled (*On*).

- **SLO**: Single-link operation on the 5 GHz link.

- **S-MLO**: EMLMR with a fixed 0.3/0.7 traffic split.
- **H-MLO**: SINR-based heuristic EMLMR traffic steering.
- **EMLSR**: Single-radio link switching per TXOP.
- **A-CAMLO**: A PPO-based hybrid policy for joint link activation and traffic allocation across multiple links.

A. Training Behavior

Fig. 3c presents the training evolution of the AP-team and Traffic-Steering critics for both MLD-APs. The losses decrease during the initial training stage and subsequently stabilize at low values, indicating stable critic fitting. The effectiveness of the resulting policy is assessed below using network-level performance metrics.

B. Experiment 1: Single-MLD-AP - Contention sensitivity

1) Setup: A single MLD-AP with two radio links (Link 1: 2.4 GHz, 20 MHz, $C_1 \approx 20$ Mbps; Link 2: 5 GHz, 40 MHz, $C_2 \approx 150$ Mbps) serves 25 MLD-STAs in a 20×30 m indoor area. At each scheduling window, $N_w = 8$ STAs are served concurrently. For a nominal load factor ρ, the per-STA demand is defined as $d_{\mathrm{STA}} = \rho(C_1 + C_2)/N_w$. The resulting demands are approximately 10.6 Mbps under underload ($\rho = 0.5$), 21.3 Mbps at saturation ($\rho = 1.0$), and 42.5 Mbps under overload ($\rho = 2.0$). Each configuration is repeated over multiple independent runs, and the total throughput delivered to each STA across its MLO links and its flow-completion time are reported as box plots.

2) Results and discussion: In the underloaded regime (50% offered load), Static MLO and A-CAMLO achieve the highest median throughput, reaching 20.01 and 20.00 Mbps, respectively, while all schemes maintain low flow-completion times between 1.17 and 1.25 ms. Under saturation, the multi-radio MLO schemes provide similar throughput (38.34–38.77 Mbps), but A-CAMLO reduces the median flow-completion time to 1.43 ms, compared with 1.96 ms for Heuristic MLO, 15.87 ms for SLO, and 36.04 ms for Static MLO. The benefit of adaptive steering becomes more pronounced under overload: A-CAMLO achieves the highest throughput of 62.20 Mbps, improving over Static MLO, Heuristic MLO, and SLO by approximately 13.5%, 67.6%, and 77.0%, respectively. It also limits the median flow-completion time to 1010.84 ms, below the corresponding values of 1222.40, 2562.67, and 2321.46 ms. Although EMLSR attains the lowest flow-completion time, its throughput remains substantially below that of A-CAMLO. These results show that A-CAMLO provides the most favorable throughput–flow-completion-time trade-off among the evaluated schemes as contention increases, particularly when the offered load exceeds the nominal AP capacity.

C. Experiment 2: Multi-MLD-AP - Interference sensitivity and MAP-Co Coordination

1) Setup: Experiment 2 focuses on H-MLO and A-CAMLO because both adapt their traffic-steering decisions to SINR variations, including those resulting from MAP-Co power control when coordination is enabled. This enables a direct comparison between rule-based and learning-based adaptive EMLMR policies under interference management. Two MLD-APs are deployed in the same 20×30 m area and operate on the same channel in each frequency band, creating OBSS conditions at saturation ($\rho = 1.0$). Interference intensity is varied through the inter-AP separation distance d across two levels: *medium* ($d = 10$ m, $\approx 50\%$ overlap) and *high* ($d = 6$ m, full overlap). Each configuration is evaluated with MAP-Co disabled (*Off*) and enabled (*On*). When enabled, the MAP-Co agent coordinates admissible per-link transmit-power bounds across the APs to mitigate inter-AP interference, within which the local Power-Control Agents select the actual transmit powers.

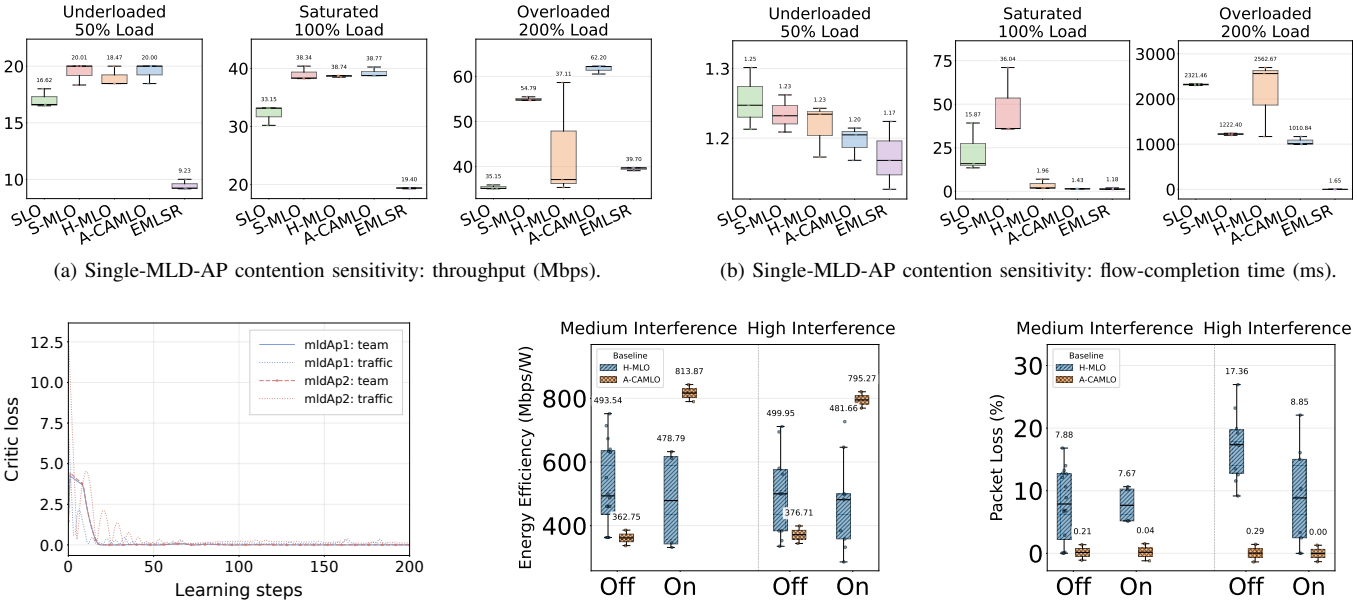
2) Results and discussion: Under *medium interference*, coordination affects the two traffic-steering policies differently. For Heuristic MLO, packet loss decreases from 7.88% to 7.67%, accompanied by a slight reduction in median energy efficiency from 493.54 to 478.79. In contrast, A-CAMLO increases its energy efficiency from 362.75 to 813.87 while reducing packet loss from 0.21% to 0.04%. Under *high interference*, coordination reduces the packet loss of Heuristic MLO from 17.36% to 8.85%, while its energy efficiency decreases slightly from 499.95 to 481.66. A-CAMLO, however, simultaneously increases its energy efficiency from 376.71 to 795.27 and reduces packet loss from 0.29% to approximately zero. These results indicate that Multi-AP coordination improves packet delivery for both adaptive policies, while A-CAMLO achieves larger coordination gains under the evaluated interference conditions by simultaneously improving energy efficiency and reliability in dense overlapping-BSS scenarios.

VII. Conclusion

In this study, we demonstrated the importance of jointly considering adaptive MLO traffic steering and Multi-AP coordination in dense IEEE 802.11bn WLANs. We addressed the lack of native MLO support in Mininet-WiFi by introducing a bridge-relay-based MLD abstraction that supports both EMLSR and EMLMR operation. Building on this extension, we proposed A-CAMLO, a two-level agentic framework that combines controller-level association and transmit-power coordination with AP-local power control, link activation, and traffic allocation. The resulting platform enables executable evaluation of coordination-aware MLO under increasing contention and overlapping-BSS interference. Our evaluations show that A-CAMLO provides the most favorable throughput–flow-completion-time trade-off under heavy traffic load, while MAP-Co coordination improves packet delivery for both adaptive policies and enhances the energy efficiency and reliability of A-CAMLO under severe interference. Future work will extend the framework toward a distributed SDN control architecture to improve scalability, resilience, and coordination across larger multi-domain deployments.

Acknowledgment

This work was funded by the Wi-FIP project under a grant awarded through the French I-DEMO call for projects (I-DEMO-52255).



(a) Single-MLD-AP contention sensitivity: throughput (Mbps).

(b) Single-MLD-AP contention sensitivity: flow-completion time (ms).

(c) A-CAMLO training behavior: traffic-agent and AP-team critic losses.

(d) Multi-MLD-AP interference sensitivity: energy efficiency (Mbps/W).

(e) Multi-MLD-AP interference sensitivity: packet loss.

Fig. 3. Evaluation results for contention sensitivity, critic training behavior, and multi-AP interference coordination.

REFERENCES

[1] 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, pp. 126–132, 2024.

[2] I. Val, D. López-Pérez, A. Kijanka, S. Schelstraete *et al.*, "Wi-fi 8 unveiled: Key features, multi-ap coordination, and the role of c-tdma," *IEEE Communications Magazine*, vol. 63, pp. 166–173, 2025.

[3] R. Sangeetha, H. C, A. KT, and A. Jaiswal, "Ieee 802.11be wlans: Performance assessment of multi-link operation with unsaturated queues," *Computer Networks*, vol. 281, p. 112187, 03 2026.

[4] S. R. Aishwarya Choorakuzhiyil, Kevin Ho, "Comparison of str and emlsr performance in wi-fi 7 multi-link operation," 2025.

[5] Y. Tai *et al.*, "Intelligent channel allocation for ieee 802.11be mlo: Multi-armed bandits meet llm," in *Proc. IEEE ICMLCN*, 2025.

[6] A. Kulshrestha, S. Shakkottai *et al.*, "Qos-aware multi-link edca optimization in ieee 802.11be networks," *IEEE Transactions on Network and Service Management*, 2025.

[7] G. Fontanesi, F. Wilhelmi, and L. G. Giordano, "Continuous multi-link operation: A contention-free mechanism for the unlicensed spectrum," in *2024 IEEE 25th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2024, pp. 766–770.

[8] S. Jung, S. Choi, H. Kim, Y. Yoon, and H.-K. Son, "Modeling the coexistence performance between wi-fi 7 and legacy wi-fi," in *NOMS 2024-2024 IEEE Network Operations and Management Symposium*, 2024, pp. 1–4.

[9] R. R. Fontes, S. Afzal, S. H. B. Brito, M. A. S. Santos, and C. E. Rothenberg, "Mininet-wifi: Emulating software-defined wireless networks," in *Proceedings of the 11th International Conference on Network and Service Management (CNSM)*, 2015, pp. 384–389.

[10] Á. López-Raventós and B. Bellalta, "Multi-link operation in IEEE 802.11be WLANs," *IEEE Wireless Communications*, vol. 29, no. 4, pp. 94–100, 2022.

[11] M. Carrascosa-Zamacois, G. Geraci, E. W. Knightly, and B. Bellalta, "Wi-Fi multi-link operation: An experimental study of latency and throughput," *IEEE/ACM Transactions on Networking*, vol. 32, no. 1, pp. 308–322, 2024.

[12] Y. Zou, M. Shen, H. Yin, and Y. Gao, "Delay analysis of multi-link devices coexisting with single-link devices in Wi-Fi 7," in *Proc. IEEE Conference on Computer Communications (INFOCOM)*, 2025, pp. 1–10.

[13] A. Abdelreheem, C. J. Sreenan, and A. H. Zahran, "Attention-enhanced DQN scheduling for multi-link devices in synchronous N-STR Wi-Fi 7 networks," in *Proc. International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWiM)*, 2025, pp. 659–666.

[14] 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.

[15] L. Zhang, H. Yin, S. Roy, and L. Cao, "Multiaccess point coordination for next-gen Wi-Fi networks aided by deep reinforcement learning," *IEEE Systems Journal*, vol. 17, no. 1, pp. 904–915, 2023.

[16] D. Nunez, F. Wilhelmi, S. Avallone, M. Smith, and B. Bellalta, "TXOP sharing with coordinated spatial reuse in multi-AP cooperative IEEE 802.11be WLANs," in *Proceedings of the IEEE 19th Annual Consumer Communications & Networking Conference (CCNC)*, 2022, pp. 864–870.

[17] J. Haxhibeqiri, X. Jiao, X. Shen, C. Pan, X. Jiang, J. Hoebeke, and I. Moerman, "Coordinated spatial reuse for WiFi networks: A centralized approach," in *Proceedings of the IEEE 20th International Conference on Factory Communication Systems (WFCS)*, 2024, pp. 1–8.

[18] L. Zhang, H. Yin, S. Roy, L. Cao, X. Gao, and V. Sathya, "IEEE 802.11be network throughput optimization with multilink operation and AP controller," *IEEE Internet of Things Journal*, vol. 11, no. 13, pp. 23 850–23 861, 2024.

[19] R. Ali and B. Bellalta, "A federated reinforcement learning framework for link activation in multi-link Wi-Fi networks," in *Proceedings of the IEEE International Black Sea Conference on Communications and Networking (BlackSeaCom)*, 2023, pp. 360–365.

[20] S. Kumar, E. García-Villegas, and D. Camps-Mur, "Multi-AP coordination for MLO scheduling in OBSS environments," in *Proceedings of the IEEE 102nd Vehicular Technology Conference (VTC2025-Fall)*, 2025, pp. 1–6.

[21] S. Avallone and P. Imputato, "Understanding the new enhanced multi-link single radio feature of ieee 802.11be wlans," *IEEE Journal on Selected Areas in Communications*, vol. 43, pp. 3683–3694, 2025.

[22] N. Korolev, I. Levitsky, I. Startsev, B. Bellalta, and E. Khorov, "Study of multi-link channel access without simultaneous transmit and receive in ieee 802.11be networks," *IEEE Access*, vol. 10, pp. 126 339–126 351, 2022.