

RL-Based Mobility Load Balancing for Energy Efficiency in Multi-Band Cellular Systems

Ahmed El Soukkary and Karim G. Seddik

Electronics and Communications Engineering Department, American University in Cairo, Egypt

Email: {ahmedelsoukkary, kseddik}@aucegypt.edu

Abstract—This paper presents a mobility-aware framework for energy-efficient load balancing in homogeneous multi-band cellular networks. The proposed approach combines a newly proposed user-association algorithm, termed *SIRRA-HO* (*SINR-Ratio-Based Resource Allocation under Handover Constraints*), with a Proximal Policy Optimization (PPO)-based power-control scheme for high-frequency bands. By prioritizing users with larger discrepancies between their best and second-best feasible serving bands, *SIRRA-HO* mitigates the loss caused by suboptimal association and improves mobility stability.

The framework explicitly captures time-varying user locations, realistic handover feasibility, and dynamic propagation conditions. Simulation results show that the joint scheme achieves more than 24% energy-efficiency gain relative to the baseline, while maintaining low dropped-user levels. They further show that RL-based power adaptation provides the dominant energy-efficiency improvement, whereas the proposed *SIRRA-HO* algorithm contributes mainly through reduced handover activity and improved mobility robustness. Overall, the results highlight the effectiveness of combining the proposed handover-aware user association with RL-based power control for next-generation cellular networks.

Index Terms—Reinforcement Learning, Cellular Networks, Machine Learning, Mobility Load Balancing, Handover

I. INTRODUCTION

The exponential growth of mobile data traffic has made energy efficiency a central concern for cellular network operators, both environmentally and economically. Mobile networks consumed on the order of 290 TWh of electricity in 2023, with base stations accounting for 70–80% of this figure, and global traffic is projected to more than double by 2030 [1]–[3]. Reducing network energy consumption while meeting quality-of-service (QoS) and fairness requirements is therefore a key design objective. Traditional resource-management strategies based on static configurations and heuristic rules struggle to exploit the energy-saving potential of modern multi-band deployments [4], [5]. This has motivated a growing body of work on optimization and learning-based approaches, including convex programming, dual decomposition, and reinforcement learning, for joint user association, power control, and cell-activity management [6]. Architecturally, the Open RAN (O-RAN) framework further enables such approaches by providing open interfaces and logically centralised RAN Intelligent Controllers (RICs), allowing closed-loop optimization logic to be deployed as modular *xApp/rApp* applications independent of vendor-specific base-station stacks [7].

Despite this progress, much of the existing literature either focuses on heterogeneous or ultra-dense deployments, or adopts a static snapshot of the network that abstracts away user mobility and handover dynamics [8]. Homogeneous multi-band networks, where each base station operates multiple

frequency layer, offer underexplored degrees of freedom for energy-aware traffic steering, particularly when user mobility is taken into account. This paper addresses that gap by extending our prior static-association framework [9] to the mobility case, explicitly modelling time-varying user locations, inter-cell and inter-band handovers, and their associated costs, within a unified energy-efficient load-balancing formulation.

The main contributions of this paper are as follows:

- We propose a new mobility-aware user-association algorithm, termed *SIRRA-HO* (*SINR-Ratio-Based Resource Allocation under Handover Constraints*), for homogeneous multi-band cellular networks. The algorithm explicitly incorporates handover-feasible candidate sets and prioritizes users with larger discrepancies between their best and second-best feasible bands, thereby reducing the degradation caused by suboptimal association.
- We combine the proposed user-association scheme with a PPO-based power-allocation framework for the high-frequency bands, enabling dynamic power adaptation to improve system energy efficiency while accounting for QoS and load-balancing objectives.
- We evaluate the proposed framework under different propagation environments and propagation scenarios, and show that the joint scheme achieves more than 24% energy-efficiency gain while the proposed UA provides a clear benefit in reducing handover activity.

The remainder of this paper is organized as follows. Section II reviews the related literature. Section III presents the system model and problem formulation. Section IV describes the proposed mobility-aware user-association and power-allocation framework. Section V reports the simulation results. Finally, Section VI concludes the paper and outlines directions for future work.

II. LITERATURE REVIEW

Energy-efficient load balancing in cellular networks has been studied through both model-based optimization and learning-based control. Early work largely focused on centralized formulations for joint user association, power allocation, and cell activity management, especially in heterogeneous and ultra-dense networks. Representative examples include mixed-integer and dual-decomposition based approaches for millimeter-wave and heterogeneous systems [10], [11], as well as alternating and iterative optimization methods for jointly improving load distribution, utility, and energy efficiency [12], [13]. These studies established the benefits of coordinated association and power control, while also highlighting the re-

liance of such centralized methods on accurate system models and their growing computational burden in dense deployments.

To improve scalability, later studies explored distributed and more practical association schemes. For example, distributed user-association and message-passing methods were proposed to reduce signaling overhead while still improving network-wide energy efficiency and load balancing [14], [15]. At the same time, mobility-aware load balancing began to attract attention. In dense LTE networks, load-balanced user association has been studied with explicit attention to handover reduction and load distribution, showing that mobility and association decisions should be treated jointly rather than through purely static metrics [16]. This direction is reinforced by recent literature mapping studies, which show that machine learning for handover optimization has become a major research theme in LTE and 5G networks, with reinforcement learning emerging as one of the dominant approaches [17].

More recently, reinforcement learning and deep reinforcement learning have been used to address the complexity and adaptivity limitations of classical optimization. Q-learning and multi-agent learning have been applied to joint user association, power allocation, and cell activation in dense heterogeneous networks, with reported gains in energy efficiency, fairness, and QoS [18], [19]. Deep and multi-agent RL methods have further extended this line of work to decentralized decision-making and more complex radio settings [20]–[23]. In the mobility context, reinforcement learning has also been used for user association and handover in mmWave-enabled networks, where explicit handover costs and rapidly varying channels make sequential decision-making particularly important [24].

Another relevant recent direction concerns intelligent control architectures for future RANs. In particular, hierarchical optimization frameworks for Open RAN have been proposed for intelligent user association and scheduling, showing how optimization and learning can be embedded into modular control loops compatible with RIC-based deployment [25]. In parallel, work on homogeneous multi-band networks has started to examine how traffic can be steered across multiple frequency layers to improve load balancing. A notable example is [8], which uses multi-objective and meta-reinforcement learning to adapt handover and reselection parameters in mobility settings. However, that work primarily targets throughput balancing rather than system-level energy efficiency.

Overall, the literature shows strong progress in optimization, distributed control, and learning-based mobility management. However, a gap remains at the intersection of three aspects: explicit energy-efficiency optimization, homogeneous multi-band cellular deployments, and time-varying user mobility. This paper addresses that gap by focusing on energy-efficient load balancing in the mobility case, where user association decisions evolve over time and must account for both traffic distribution and handover-related effects.

III. SYSTEM MODEL AND PROBLEM FORMULATION

We consider the downlink of a homogeneous multi-band cellular network with base stations $\mathcal{N} = \{1, \dots, N\}$ and users $\mathcal{U} = \{1, \dots, U\}$. Each base station operates a low-frequency band and a high-frequency band; let $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$ denote the set

of all bands, where $\mathcal{B}_1$ and $\mathcal{B}_2$ are the low- and high-frequency bands, respectively. Time is slotted as $t = 1, \dots, T$, and users are mobile over the considered horizon. Let $H_{b,u}(t)$ denote the effective channel gain experienced by user u on band b at slot t , capturing the time-varying propagation conditions induced by mobility.

Let $x_{b,u}(t) \in \{0, 1\}$ indicate whether user u is associated with band b at slot t , and let $P_b(t)$ be the transmit power of band b . For $b \in \mathcal{B}_1$, the received SINR is $\gamma_{b,u}(t) = \frac{P_b(t)H_{b,u}(t)}{\sum_{b' \in \mathcal{B}_1 \setminus \{b\}} P_{b'}(t)H_{b',u}(t) + \sigma^2}$. The load of band b is $y_b(t) = \sum_{u \in \mathcal{U}} x_{b,u}(t)$, and assuming equal bandwidth sharing, the achievable rate is $R_{b,u}(t) = \frac{W_b}{y_b(t)} \log_2(1 + \gamma_{b,u}(t))$.

To incorporate mobility, association updates are restricted by handover feasibility. Let $b_u(t-1)$ be the serving band of user u at slot $t-1$, and let $M_{b,u}(t)$ denote the measurement quantity used for mobility decisions. A candidate band b is feasible at slot t only if it satisfies the Event A3 condition over the time-to-trigger (TTT) window [26], i.e., $M_{b,u}(\tau) \geq M_{b_u(t-1),u}(\tau) + \text{HOM}$ for all $\tau \in \{t - \text{TTT} + 1, \dots, t\}$, where HOM is the handover margin in dB. Accordingly, the feasible candidate set is $\mathcal{B}_u^{\text{HO}}(t) = \{b_u(t-1)\} \cup \{b \in \mathcal{B} : e_{b,u}(t) = 1\}$ for $t \geq 2$, while $\mathcal{B}_u^{\text{HO}}(1) = \mathcal{B}$.

At each slot, the decision variables are the association variables $\{x_{b,u}(t)\}$ and the transmit powers of the high-frequency bands $\{P_b(t) : b \in \mathcal{B}_2\}$, while low-frequency powers are fixed and each high-frequency power is selected from a feasible set $\mathcal{P}_b$. The objective is to maximize the time-aggregated energy efficiency,

$$\text{EE} = \frac{\sum_{t=1}^T \sum_{b \in \mathcal{B}} \sum_{u \in \mathcal{U}} x_{b,u}(t) R_{b,u}(t)}{\sum_{t=1}^T \sum_{b \in \mathcal{B}} P_b(t)}, \quad (1)$$

subject to single-band association, power feasibility, per-slot QoS, and handover constraints:

$$\max_{\{x_{b,u}(t)\}, \{P_b(t)\}} \text{EE} \quad (2a)$$

$$\text{s.t.} \quad \sum_{b \in \mathcal{B}} x_{b,u}(t) \leq 1, \quad \forall u, t, \quad (2b)$$

$$P_b(t) \in \mathcal{P}_b, \quad \forall b \in \mathcal{B}_2, \forall t, \quad (2c)$$

$$\sum_{b \in \mathcal{B}} x_{b,u}(t) R_{b,u}(t) \geq R_u^{\min}, \quad \forall u, t, \quad (2d)$$

$$x_{b,u}(t) = 0, \quad \forall b \notin \mathcal{B}_u^{\text{HO}}(t), \forall u, t \geq 2, \quad (2e)$$

$$x_{b,u}(t) \in \{0, 1\}, \quad \forall b, u, t. \quad (2f)$$

The resulting problem is a mixed-integer nonlinear program, due to the binary association variables, nonlinear SINR-rate coupling, and the time dependence introduced by user motion and handover feasibility. We therefore decompose it into two coupled subproblems: a user-association subproblem that selects feasible serving bands under QoS and mobility constraints, and a power-allocation subproblem that adjusts high-band transmit powers over time to improve energy efficiency under the resulting load distribution.

IV. PROPOSED SOLUTION

To solve the mobility-aware problem efficiently, we decompose it into a user-association (UA) stage executed at every

slot and a power-allocation (PA) stage updated at a slower timescale.

A. Proposed UA: SIRRA-HO

The proposed UA algorithm, termed *SIRRA-HO* (*SINR-Ratio-Based Resource Allocation under Handover Constraints*), prioritizes users according to the discrepancy between their best and second-best feasible bands, so as to minimize the degradation caused by suboptimal association decisions. At slot t , for each unassigned user u , we define the effective candidate set as $\mathcal{B}_u(t) = \mathcal{B}_u^{\text{HO}}(t) \cap \mathcal{B}_{\text{open}}$, where $\mathcal{B}_{\text{open}}$ is the set of bands still open to new assignments. If $\mathcal{B}_u(t) = \emptyset$, the user is dropped at that slot. Otherwise, let $b_1(u)$ and $b_2(u)$ denote the best and second-best feasible bands in terms of instantaneous SINR, with corresponding values $\gamma_1(u)$ and $\gamma_2(u)$. For users with at least two feasible bands, we compute the priority metric $\mu_u(t) = \gamma_1(u)/\gamma_2(u)$ and process users in descending order of $\mu_u(t)$, so that users with larger best-versus-second-best discrepancy are assigned earlier. This reduces the loss incurred by forcing highly sensitive users to suboptimal bands.

Users with only one feasible band are handled separately and prioritized according to their only feasible SINR. The algorithm therefore maintains two queues: one for singleton-candidate users ordered by $\gamma_1(u)$, and one for users with at least two candidates ordered by $\mu_u(t)$. At each step, the algorithm selects the head user with the larger best-band SINR, assigns it to $b_1(u)$ if its QoS requirement is satisfied, and updates band availability accordingly. As in the static case, if admitting more users to a band would violate the rate constraints of already associated users, that band is closed to further assignments. In this way, the proposed mobility-aware UA preserves the low-complexity greedy structure of the static algorithm while explicitly accounting for handover-feasible candidate sets. The complete step-by-step pseudocode is given in Algorithm 1.

B. Benchmark UA: Max-SINR-HO

For benchmarking, we consider a mobility-aware Max-SINR user-association baseline, termed *Max-SINR-HO*. At each slot t , user u is assigned to the handover-feasible band with the largest instantaneous SINR, namely $b_u^*(t) = \arg \max_{b \in \mathcal{B}_u^{\text{HO}}(t)} \gamma_{b,u}(t)$. If the resulting assignment violates QoS constraints, the affected user is reassigned iteratively to its next-best feasible band within $\mathcal{B}_u^{\text{HO}}(t)$ until feasibility is restored or no candidate band remains.

Thus, Max-SINR-HO follows the conventional Max-SINR association rule while enforcing the same handover-feasibility constraints as the proposed SIRRA-HO scheme, ensuring a fair benchmark.

C. Reinforcement Learning for Power Allocation

The PA subproblem is modeled as a reinforcement-learning task and solved using PPO [27]. To avoid overly frequent power changes under mobility, the agent updates the high-band power vector once every T_d slots.

Algorithm 1: Proposed UA (SIRRA-HO) at slot t

Input: SINR values $\gamma_{b,u}(t)$; HO-feasible sets $\mathcal{B}_u^{\text{HO}}(t)$ for all $u \in \mathcal{U}$ and $b \in \mathcal{B}$

Output: User association $\{x_{b,u}(t)\}$

```

1 Initialize  $x_{b,u}(t) \leftarrow 0$  for all  $b \in \mathcal{B}$ ,  $u \in \mathcal{U}$ 
2 Set  $\mathcal{U}_{\text{unassigned}} \leftarrow \mathcal{U}$ ,  $\mathcal{B}_{\text{open}} \leftarrow \mathcal{B}$ 
3 while  $\mathcal{U}_{\text{unassigned}} \neq \emptyset$  and  $\mathcal{B}_{\text{open}} \neq \emptyset$  do
4   Initialize two empty priority queues  $\mathcal{Q}_1$  and  $\mathcal{Q}_2$ 
5   foreach  $u \in \mathcal{U}_{\text{unassigned}}$  do
6      $\mathcal{B}_u \leftarrow \mathcal{B}_u^{\text{HO}}(t) \cap \mathcal{B}_{\text{open}}$ 
7     if  $\mathcal{B}_u = \emptyset$  then
8       Drop user  $u$  at slot  $t$  and remove it from
         $\mathcal{U}_{\text{unassigned}}$ 
9     end
10    Let  $b_1(u) \leftarrow \arg \max_{b \in \mathcal{B}_u} \gamma_{b,u}(t)$  and
       $\gamma_1(u) \leftarrow \gamma_{b_1(u),u}(t)$ 
11    if  $|\mathcal{B}_u| = 1$  then
12      Insert user  $u$  into  $\mathcal{Q}_1$  with priority key
         $\gamma_1(u)$ 
13    else
14      Let  $b_2(u)$  be the second-best band in  $\mathcal{B}_u$ 
        and  $\gamma_2(u) \leftarrow \gamma_{b_2(u),u}(t)$ 
15      Compute metric  $\mu_u(t) = \gamma_1(u)/\gamma_2(u)$ 
16      Insert user  $u$  into  $\mathcal{Q}_2$  with priority key
         $\mu_u(t)$ 
17    end
18  end
19  while  $\mathcal{Q}_1 \neq \emptyset$  or  $\mathcal{Q}_2 \neq \emptyset$  do
20    if  $\mathcal{Q}_1 = \emptyset$  then
21       $u^* \leftarrow \text{PopHead}(\mathcal{Q}_2)$ 
22    else if  $\mathcal{Q}_2 = \emptyset$  then
23       $u^* \leftarrow \text{PopHead}(\mathcal{Q}_1)$ 
24    else
25       $\hat{u}_1 \leftarrow \text{Head}(\mathcal{Q}_1)$ ,  $\hat{u}_2 \leftarrow \text{Head}(\mathcal{Q}_2)$ 
26      if  $\gamma_1(\hat{u}_1) \geq \gamma_1(\hat{u}_2)$  then
27         $u^* \leftarrow \text{PopHead}(\mathcal{Q}_1)$ 
28      else
29         $u^* \leftarrow \text{PopHead}(\mathcal{Q}_2)$ 
30      end
31    end
32    if band  $b_1(u^*)$  can satisfy  $u^*$ 's demands then
33       $x_{b_1(u^*),u^*}(t) \leftarrow 1$ 
34      Remove  $u^*$  from  $\mathcal{U}_{\text{unassigned}}$ 
35    end
36    else if No open band in  $\mathcal{B}_{u^*}$  can satisfy  $u^*$ 's
      demand then
37      Remove  $u^*$  from  $\mathcal{U}_{\text{unassigned}}$ 
38    end
39    if assigning new users to  $b_1(u^*)$  violates the
      rate constraints of its already-connected users
      then
40      Remove  $b_1(u^*)$  from  $\mathcal{B}_{\text{open}}$ .
41    end
42  end
43 end
44 return  $\{x_{b,u}(t)\}$ 

```

a) *States*: The RL state is defined at the beginning of each decision step k , where the decision instants are $t_k = (k - 1)T_d + 1$, $k = 1, \dots, K$, and $K = \lceil T/T_d \rceil$. As in the static case, the state consists of the normalized user density and per-band load-quality indicators,

$$\mathbf{s}_k = \mathbf{s}(t_k) = [U/U_0, g_{b_1}(t_k), \dots, g_{b_{2N}}(t_k)], \quad (3)$$

where U_0 is the reference user count used for normalization and $g_b(t_k)$ is the proportion of users on band b whose RSRP exceeds a predefined threshold.

b) *Actions*: At each decision step k , the agent selects a feasible high-band power vector

$$\mathbf{a}_k = [P_b(k) \mid b \in \mathcal{B}_2, P_b(k) \in \mathcal{P}_b], \quad (4)$$

which is held fixed over the block $\mathcal{T}_k = \{t_k, \dots, \min(t_k + T_d - 1, T)\}$. Within each block, the UA algorithm is executed at every slot using the selected power vector.

c) *Reward*: To avoid penalizing short transient disconnections caused by mobility or handover execution, a user is counted as dropped only if it remains unconnected for T_{con} consecutive slots. Letting $D(t)$ denote the resulting dropped-user fraction, and $J(t) = \frac{(\sum_{b \in \mathcal{B}} y_b(t))^2}{|\mathcal{B}| \sum_{b \in \mathcal{B}} y_b^2(t)}$ denote Jain's fairness index [28] over the band loads, the reward at decision step k is defined as the block-average

$$r_k = \frac{1}{|\mathcal{T}_k|} \sum_{t \in \mathcal{T}_k} (\widehat{\text{EE}}(t) - \beta D(t) + \alpha J(t)), \quad (5)$$

where $\widehat{\text{EE}}(t)$ is the normalized instantaneous energy efficiency, and α and β weight fairness and connectivity penalties, respectively. This reward encourages energy-efficient operation while maintaining balanced load distribution and limiting persistent user outages.

V. SIMULATION RESULTS AND ANALYSIS

This section evaluates the proposed mobility-aware framework under time-varying channels and handover constraints. Four KPIs are considered: energy efficiency (EE), dropped-user percentage, Jain's fairness index, and handover rate. We first compare the different UA/PA configurations under a representative propagation setting, then examine the effect of the load-balancing weight α , and finally study the impact of increasingly NLOS-dominated propagation conditions.

A. Simulation setup

The network consists of three base stations deployed over a 1 km $\times$ 1 km area with inter-site distance 500 m. Each base station operates two bands: a low-frequency band at 800 MHz with bandwidth 5 MHz and a high-frequency band at 2.6 GHz with bandwidth 10 MHz. The low-band transmit powers are fixed at 20 W, while each high-band power is either fixed to 40 W in the full-power schemes or selected by RL from the discrete set $\mathcal{P}_b = \{0, 20, 22.5, \dots, 40\}$ W. For all KPI-versus-users results, the number of users is swept from 10 to 610.

Each user moves along a linear track of length 200 m with a random orientation and a constant speed 1 m/s, which corresponds to a time horizon of $T = 695$ slots. Each track is divided into segments, and each segment is labeled NLOS with probability p_{NLOS} and LOS otherwise; the corresponding channel parameters are generated using the appropriate

Parameter	Value	Parameter	Value
Steps per episode	30	Clip factor	0.2
Experience horizon	15	Mini-batch size	5
Number of epochs	3	Entropy loss weight	0.1
GAE factor	0.95	Discount factor	0.995

TABLE I: PPO hyperparameters for mobility-aware power allocation.

QuaDRiGa UMa LOS/NLOS scenarios [29]. The minimum user-rate requirement is independently sampled from 50 to 150 kbps. Event A3 handover feasibility is enforced using $\text{TTT} = 3$ slots and $\text{HOM} = 3$ dB, and the connectivity window used in the dropped-user definition is set to $T_{\text{con}} = 2$ slots. For RL-based power allocation, the PPO agent updates the high-band power vector every $T_d = 24$ slots and keeps it fixed between two decision epochs. All agents are trained for 1000 episodes using the hyperparameters in Table I. The RSRP threshold used in the good-user fractions is adapted as $\text{RSRP}_{\text{thresh}}(p_{\text{NLOS}}) = 10^{(-8 - 2.9p_{\text{NLOS}})} \text{ W}$. This expression was obtained empirically from preliminary simulations by tracking how the RSRP distribution changes with p_{NLOS} and selecting a simple monotonic fit that keeps the good-user fractions g_b within a useful dynamic range, rather than saturating near zero in strongly NLOS scenarios.

B. Results for $p_{\text{NLOS}} = 0.9$.

We first compare the four UA/PA configurations at $p_{\text{NLOS}} = 0.9$ with $(\alpha, \beta) = (1, 1)$, where the baseline UA is Max-SINR-HO and the proposed UA is SIRRA-HO. Figures 1–4 show EE, dropped-user percentage, load balancing fairness, and handover rate, respectively, as functions of the number of users. The figures show that the dominant EE improvement comes from RL-based PA: RL increases EE by 23.13% with Max-SINR and by 23.26% with SIRRA-HO, whereas replacing Max-SINR by SIRRA-HO yields a smaller additional gain of 0.90% under full power and 1.01% under RL power; applying both changes together gives an overall EE gain of 24.39%.

The dropped-user percentage remains low across all configurations. In particular, RL leaves the dropped-user percentage unchanged with Max-SINR and changes it by only 0.21 percentage points (pp) with SIRRA-HO, while replacing Max-SINR by SIRRA-HO changes the dropped-user percentage by 0.65 pp under full power and 0.87 pp under RL power; the overall joint change is therefore 0.87 pp. For load balancing fairness, RL reduces Jain's index by 4.00 pp with Max-SINR and by 4.37 pp with SIRRA-HO, whereas SIRRA-HO recovers 0.57 pp under full power and 0.20 pp under RL power, leading to a net fairness change of -3.80 pp when both mechanisms are combined.

The clearest advantage of SIRRA-HO appears in the handover metric. RL alone has only a minor effect on handover rate, reducing it by 0.11 pp with Max-SINR and by 0.03 pp with SIRRA-HO, while replacing Max-SINR by SIRRA-HO reduces the handover rate by 1.34 pp under full power and 1.27 pp under RL power; jointly, the two changes reduce the handover rate by 1.37 pp.

C. Effect of the load-balancing weight α

To examine the influence of the fairness term in the reward, we compare the joint proposed UA + RL PA scheme at

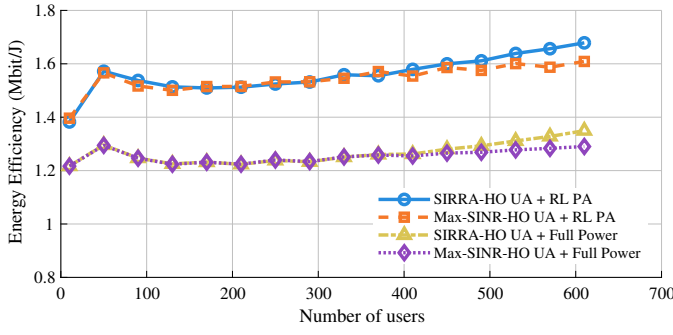


Fig. 1: Energy efficiency comparison for $p_{\text{NLOS}} = 0.9$, and $(\alpha, \beta) = (1, 1)$.

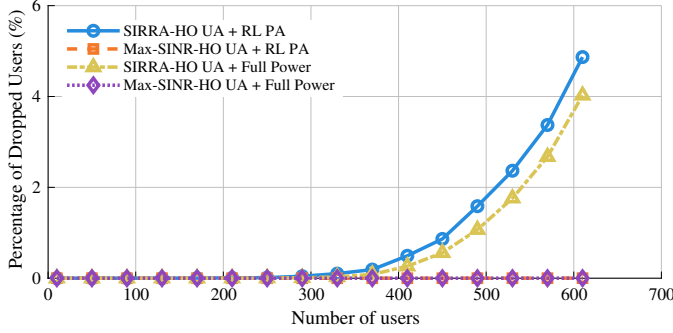


Fig. 2: Dropped users comparison for $p_{\text{NLOS}} = 0.9$, and $(\alpha, \beta) = (1, 1)$.

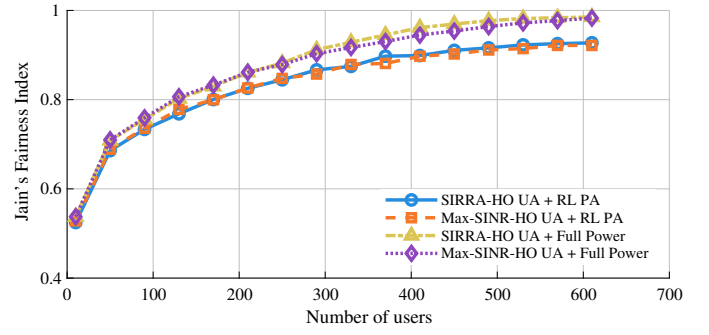


Fig. 3: Fairness index comparison for $p_{\text{NLOS}} = 0.9$, and $(\alpha, \beta) = (1, 1)$.

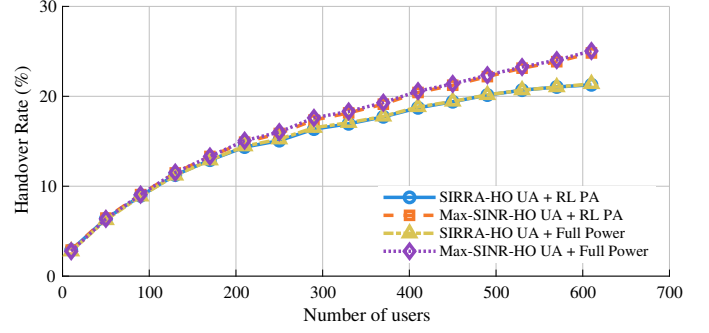


Fig. 4: Handover rate comparison for $p_{\text{NLOS}} = 0.9$, and $(\alpha, \beta) = (1, 1)$.

$p_{\text{NLOS}} = 0.5$ for two values of the fairness weight, namely $\alpha = 1$ and $\alpha = 4.5$. Table II shows that increasing α slightly reduces the EE gain from 27.23% to 25.98%, while leaving the impact on dropped users unchanged at +0.34 percentage points (pp). At the same time, the fairness degradation is marginally reduced from -5.17 pp to -5.01 pp, indicating that a stronger fairness emphasis can partially recover load balance without eliminating the energy-efficiency benefit of RL. The handover-rate impact remains unchanged at -0.08 pp, which suggests that α mainly affects load redistribution rather than mobility-triggered association updates.

Metric	$\alpha = 1$	$\alpha = 4.5$
EE Gain	27.23%	25.98%
Impact on Dropped Users	+0.34 pp	+0.34 pp
Impact on Fairness	-5.17 pp	-5.01 pp
Impact on Handover Rate	-0.08 pp	-0.08 pp

TABLE II: Effect of the fairness weight α on the joint UA-PA gain ($p_{\text{NLOS}} = 0.5$, $\beta = 1$).

D. Sensitivity to p_{NLOS}

To assess robustness, we sweep the segment-level NLOS probability p_{NLOS} while keeping the remaining parameters fixed. For each value of p_{NLOS} , the four KPIs are evaluated over the full simulation horizon for all configurations. Figure 5 summarizes how the joint UA-PA effect varies with the propagation environment. As expected, the benefit of the proposed framework is influenced by the severity of NLOS propagation. While the EE gain shows a decreasing trend as p_{NLOS} increases, it still exceeds 24% in the considered range, which demonstrates the robustness of the proposed joint UA-PA design even under NLOS-dominated conditions. Similarly, the dropped-user performance tends to worsen at larger p_{NLOS} ,

reflecting the tighter feasibility imposed by minimum-rate constraints in NLOS-dominated tracks. Load balancing fairness is also negatively affected in difficult propagation conditions, suggesting that energy-oriented decisions increasingly concentrate traffic on a smaller subset of bands. In contrast, the relative handover-rate improvement becomes more pronounced as p_{NLOS} grows, indicating that the proposed mobility-aware UA remains effective in mitigating unnecessary handovers even in more challenging environments.

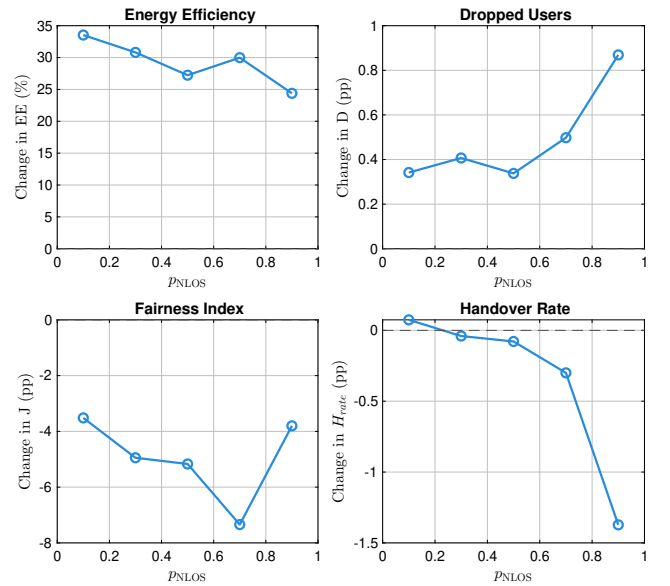


Fig. 5: Joint effect of proposed UA and RL PA on the metrics versus p_{NLOS}

VI. CONCLUSIONS AND FUTURE WORK

This paper presented a mobility-aware framework for energy-efficient load balancing in homogeneous multi-band cellular networks, combining a SINR-ratio-based user association algorithm with PPO-based power allocation on high-frequency bands. The system model explicitly accounts for LOS/NLOS propagation, equal-bandwidth sharing, LTE Event A3 handover-feasibility constraints, and a comprehensive set of KPIs including energy efficiency, dropped-user fraction, load-balancing fairness, and handover rate. The decomposed design, with slot-level association and slower-timescale power control, yields a practical solution for the otherwise intractable joint UA-PA optimization.

Simulation results with QuaDRiGa-based channels demonstrate that the proposed SIRRA-HO+RL scheme can substantially improve energy efficiency compared to fixed full-power operation while keeping dropped-user percentages and fairness close to baseline levels. Importantly, the RL-based PA also reduces the average handover rate relative to Max-SINR full-power configurations, indicating that mobility robustness and energy efficiency can be improved jointly rather than being conflicting objectives.

Future work will consider decentralized realizations of the proposed framework, for example through multi-agent RL or distributed association with limited inter-BS signaling, to enhance scalability and robustness in dense deployments. Another extension is to incorporate MIMO spatial degrees of freedom, such as beamforming and spatial multiplexing, into the state, action, and reward design, enabling joint optimization over association, power, and spatial resources for further gains in energy efficiency and mobility performance.

REFERENCES

- [1] "Mobile net zero report 2025," 2025. [Online]. Available: <https://view.gsma.com/mobile-net-zero-2025/>
- [2] A. Zappone *et al.*, "Conceiving sustainable 6g networks," CentraleSupélec, Tech. Rep., 2024, white paper on energy consumption and sustainability of mobile networks. [Online]. Available: <https://12s.centralesupelec.fr/>
- [3] "Ericsson mobility report, june 2025," 2025, mobile data traffic and 5G outlook to 2030. [Online]. Available: <https://www.ericsson.com/en/reports-and-papers/mobility-report>
- [4] S. Buzzi, C.-L. I, T. E. Klein, H. V. Poor, C. Yang, and A. Zappone, "A survey of energy-efficient techniques for 5g networks and challenges ahead," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 4, pp. 697–709, 2016.
- [5] F. Hussain, S. A. Hassan, R. Hussain, and E. Hossain, "Machine learning for resource management in cellular and iot networks: Potentials, current solutions, and open challenges," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 2, pp. 1251–1275, 2020.
- [6] A. Alwarafy, M. Abdallah, B. S. Çiftler, A. Al-Fuqaha, and M. Hamdi, "The frontiers of deep reinforcement learning for resource management in future wireless hetnets: Techniques, challenges, and research directions," *IEEE Open Journal of the Communications Society*, vol. 3, pp. 322–365, 2022.
- [7] O-RAN Alliance, "O-RAN: Towards an Open and Smart RAN," 2018. [Online]. Available: <https://mediastorage.o-ran.org/white-papers/O-RAN.White-Paper-2018-10.pdf>
- [8] A. Feriani, D. Wu, Y. T. Xu, J. Li, S. Jang, E. Hossain, X. Liu, and G. Dudek, "Multiobjective load balancing for multiband downlink cellular networks: A meta- reinforcement learning approach," *IEEE Journal on Selected Areas in Communications*, vol. 40, no. 9, pp. 2614–2629, 2022.
- [9] A. El Soukkary and K. G. Seddik, "Energy efficient load balancing in multiband cellular networks via reinforcement learning," in *2025 IEEE 36th International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, 2025, pp. 1–6.
- [10] H. Zhang, S. Huang, C. Jiang, K. Long, V. C. M. Leung, and H. V. Poor, "Energy efficient user association and power allocation in millimeter-wave-based ultra dense networks with energy harvesting base stations," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 9, pp. 1936–1947, 2017.
- [11] T. Zhou, Z. Liu, J. Zhao, C. Li, and L. Yang, "Joint user association and power control for load balancing in downlink heterogeneous cellular networks," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 3, pp. 2582–2593, 2018.
- [12] X. Huang, W. Xu, H. Shen, H. Zhang, and X. You, "Utility-energy efficiency oriented user association with power control in heterogeneous networks," *IEEE Wireless Communications Letters*, vol. 7, no. 4, pp. 526–529, 2018.
- [13] T. Zhou, Y. Fu, D. Qin, X. Li, and C. Li, "Joint user association and bs operation for green communications in ultra-dense heterogeneous networks," *IEEE Transactions on Vehicular Technology*, vol. 73, no. 2, pp. 2305–2319, 2024.
- [14] S. H. Lee, M. Kim, H. Shin, and I. Lee, "Belief propagation for energy efficiency maximization in wireless heterogeneous networks," *IEEE Transactions on Wireless Communications*, vol. 20, no. 1, pp. 56–68, 2021.
- [15] R. Chatzigeorgiou and A. Bletsas, "Distributed, inference-based, energy efficient user association with convergence guarantees," in *ICC 2023 - IEEE International Conference on Communications*, 2023, pp. 3228–3233.
- [16] S. Biswas, A. Gupta, and S. Chakraborty, "Load-balanced user associations in dense LTE networks," *Computer Networks*, vol. 189, p. 107928, 2021.
- [17] V. Párraga-Villamar, D. S. Guzmán, F. Grijalva, and P. A. Lupera Morillo, "A systematic literature mapping of machine learning for handover optimization in cellular networks," *IEEE Access*, vol. 13, pp. 144 699–144 732, 2025.
- [18] Y. L. Lee, W. L. Tan, S. B. Y. Lau, T. C. Chuah, A. A. El-Saleh, and D. Qin, "Joint cell activation and user association for backhaul load balancing in green hetnets," *IEEE Wireless Communications Letters*, vol. 9, no. 9, pp. 1486–1490, 2020.
- [19] D. Li, H. Zhang, K. Long, W. Huangfu, J. Dong, and A. Nallanathan, "User association and power allocation based on q-learning in ultra dense heterogeneous networks," in *2019 IEEE Global Communications Conference (GLOBECOM)*, 2019, pp. 1–5.
- [20] D. Kim, H. Jung, and I.-H. Lee, "Deep learning-based spectral and energy efficiency optimization for comp in hetnets," in *2023 IEEE Globecom Workshops (GC Wkshps)*, 2023, pp. 1970–1975.
- [21] J. Moon, S. Kim, H. Ju, and B. Shim, "Energy-efficient user association in mmwave/thz ultra-dense network via multi-agent deep reinforcement learning," *IEEE Transactions on Green Communications and Networking*, vol. 7, no. 2, pp. 692–706, 2023.
- [22] X. Zhao, Y. Cao, H. Chen, Z. Huang, and D. Wang, "Multi-objective resource allocation based on deep reinforcement learning in hetnets," in *2022 IEEE 8th International Conference on Computer and Communications (ICCC)*, 2022, pp. 574–578.
- [23] S. R. M. Abouamasha, "Load balancing in mobile networks using deep reinforcement learning and traffic prediction," Master's thesis, American University in Cairo, 2025. [Online]. Available: <https://fount.aucegypt.edu/etds/2495>
- [24] A. Alizadeh and M. Vu, "Reinforcement learning for user association and handover in mmWave-enabled networks," *IEEE Transactions on Wireless Communications*, vol. 21, no. 11, pp. 9712–9728, 2022.
- [25] F. Kavehmadavani, T. X. Vu, V.-D. Nguyen, and S. Chatzinotas, "Intelligent user association and scheduling in open RAN: A hierarchical optimization framework," *IEEE Transactions on Communications*, vol. 73, no. 11, pp. 11 574–11 589, 2025.
- [26] ETSI, "ETSI TS 136 331 V17.7.0 (2024-02): LTE; E-UTRA; Radio Resource Control (RRC); Protocol specification," European Telecommunications Standards Institute (ETSI), Tech. Rep. TS 136 331 V17.7.0, 2024. [Online]. Available: https://www.etsi.org/deliver/etsi_ts/136300_136399/136331/17.07.00_60/ts_136331v170700p.pdf
- [27] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal policy optimization algorithms," *arXiv preprint arXiv:1707.06347*, 2017. [Online]. Available: <https://arxiv.org/abs/1707.06347>
- [28] R. Jain, D.-M. Chiu, and W. Hawe, "A quantitative measure of fairness and discrimination for resource allocation in shared computer systems," Digital Equipment Corporation, Tech. Rep. DEC-TR-301, 1984.
- [29] S. Jaeckel, L. Raschkowski, K. BÄrner, L. Thiele, F. Burkhardt, and E. Eberlein, "Quadriga – quasi deterministic radio channel generator, user manual and documentation," Fraunhofer Heinrich Hertz Institute, Tech. Rep. v2.8.1, 2023.