eng

MDPI

Article

Dynamic Uplink Power Control for Cell-Free Massive MIMO

Hussein A. Jasim [1,2,*], Mohd Fadlee A. Rasid [1,3], Fazirulhisyam Hashim [1,3] and Syamsiah Mashohor [1]

[1] Department of Computer and Communication Systems Engineering, Faculty of Engineering, Universiti Putra Malaysia (UPM), Serdang 43400, Selangor, Malaysia; fadlee@upm.edu.my (M.F.A.R.); fazirul@upm.edu.my (F.H.); syamsiah@upm.edu.my (S.M.)
[2] Computer Sciences Department, College of CSIT, University of Basrah, Basrah 61003, Iraq
[3] Wireless and Photonics Networks Research Centre (WiPNET), Faculty of Engineering, Universiti Putra Malaysia (UPM), Serdang 43400, Selangor, Malaysia
* Correspondence: gs61712@student.upm.edu.my or hussein.jasim@uobasrah.edu.iq

Abstract

Dynamic uplink power allocation is a critical challenge in cell-free massive MIMO (CF-mMIMO) networks, where distributed access points (APs) jointly serve multiple user equipment (UEs) under mobility, time-varying propagation conditions, and strong inter-user interference. Conventional optimization-based methods can improve fairness or spectral efficiency, but they often require repeated numerical solving and are usually designed for a specific objective. Learning-based approaches can reduce online decision time after training; however, their effectiveness depends strongly on the reward design and the selected operating objective. In response to these challenges, we propose a Deep Hybrid Intelligent (DHI) architecture designed to evaluate dynamic uplink power management within cell-free massive MIMO environments. The framework uses Soft Actor-Critic (SAC) learning to generate continuous uplink transmit-power decisions and evaluates objective-specific configurations for fairness, signal-to-interference-plus-noise ratio (SINR) improvement, and spectral-efficiency enhancement. In addition, three optimization-based strategies, namely max-min fairness, max-product SINR optimization, and max-sum-rate maximization, are incorporated to analyze the trade-off among fairness, signal quality, throughput, and computational cost. Limited-memory Broyden-Fletcher-Goldfarb-Shanno with bound constraints (L-BFGS-B) optimization is employed for the max-product and max-sum-rate objectives, while the max-min strategy is evaluated through a fairness-oriented feasibility procedure. Simulation results show that the fairness-oriented configuration achieves the highest Jain's fairness index, reaching 0.989 at 120 access points, whereas the sum-rate-oriented configuration provides stronger SINR and user-rate performance. The results also indicate execution-time reductions of 51.6%, 83.7%, and 85.0% for the evaluated max-min, max-product, and max-sum-rate strategies, respectively, compared with conventional optimization-based implementations. These execution-time gains are accompanied by a clear performance trade-off: the max-min strategy provides the strongest fairness behavior, the max-sum-rate strategy improves total spectral efficiency and user-rate performance, and the max-product strategy offers a balanced operating point between collective SINR improvement and user-service balance. Therefore, the proposed framework does not optimize only computational speed, but also clarifies the trade-off among execution time, SINR, spectral efficiency, and fairness under dynamic uplink CF-mMIMO conditions. These results indicate that this architecture serves as an adaptable platform to evaluate dynamic uplink power distribution across CF-mMIMO networks.

Check for updates

Academic Editors: Yo-Sheng Lin and Mário Silva

Received: 10 June 2026
Revised: 1 July 2026
Accepted: 20 July 2026
Published: 22 July 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

Keywords: cell-free massive MIMO; uplink power control; Soft Actor-Critic; spectral efficiency; Jain's fairness index; L-BFGS-B; dynamic resource allocation

1. Introduction

Next-generation wireless networks require efficient radio resource management techniques that can support high spectral efficiency, reliable connectivity, and balanced user service under dynamic propagation conditions [1]. Cell-free massive multiple-input multiple-output (CF-mMIMO) has emerged as a promising network architecture because a large number of distributed access points (APs) jointly serve multiple user equipment (UEs) without relying on conventional cell boundaries [2], as shown in Figure 1. This distributed cooperation improves coverage uniformity, reduces the impact of shadow fading, and mitigates inter-user interference, particularly in dense wireless deployments [3].

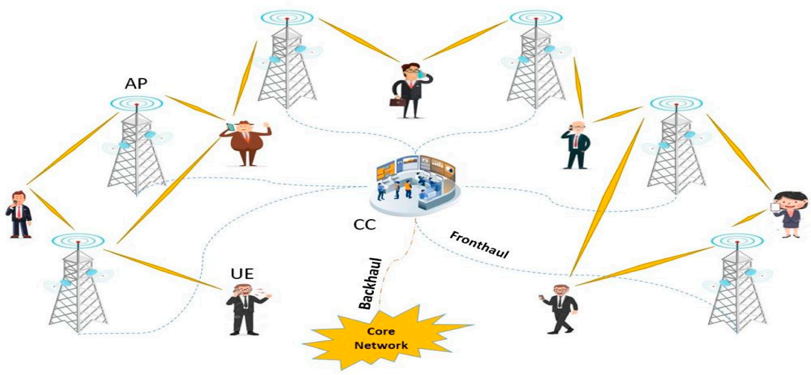

Figure 1. General architecture of cell-free massive MIMO system.

However, the performance advantages of CF-mMIMO strongly depend on how the uplink transmit power is allocated among users [4]. In the uplink, inappropriate power allocation may increase interference, degrade the signal-to-interference-plus-noise ratio (SINR), reduce spectral efficiency, and create unfair service among users, particularly when user mobility and time-varying channel conditions are considered [5].

Uplink power control in CF-mMIMO systems is inherently complex due to the tight coupling between users; increasing one UE's transmit power improves its own link but simultaneously raises the interference level for others [6]. This inter-user dependence grows highly intricate as more APs and active UEs join the network. Several optimization-based approaches have been proposed to address this problem, including max-min fairness, proportional fairness, sum-rate maximization, and weighted minimum mean-square error (WMMSE)-based techniques [7]. Fairness-oriented methods, such as max-min and proportional fairness schemes, are effective in protecting users with weak channel conditions and improving service uniformity. Nevertheless, such strategies can degrade the aggregate network throughput, given that they prioritize channeling resources toward users with poorer channel conditions [8]. In contrast, sum-rate and WMMSE-based methods can improve spectral efficiency and overall system performance, but they usually require repeated iterative computation and may increase the online computational burden in large-scale or dynamic CF-mMIMO environments [7,9]. Therefore, although classical optimization methods remain important for CF-mMIMO power control, many are objective-specific and computationally demanding when real-time adaptation is required [10].

Recent interference-aware power-allocation studies further show the importance of jointly considering interference mitigation and computational efficiency in wireless resource allocation. For example, hybrid optimization-based approaches have been investigated for interference-aware power allocation in GFDM- and OFDM-based cognitive radio networks, where particle swarm optimization and simulated annealing are combined to improve power-allocation performance under interference constraints. Although such methods demonstrate the effectiveness of hybrid optimization for interference-aware radio-resource management, they are still generally designed for specific network architectures and require iterative search procedures. In contrast, the present study focuses on uplink CF-mMIMO and combines SAC-based continuous power-control learning with objective-specific refinement strategies to analyze fairness-oriented, SINR-oriented, and throughput-oriented behavior under a common dynamic simulation environment [11].

Recently, learning-based power-control methods have been introduced to improve adaptability and reduce online computational complexity in CF-mMIMO systems [12]. Deep learning and deep reinforcement learning (DRL) approaches can learn the relationship between network states and power-control decisions, allowing faster inference after training [13]. For example, deep learning-based methods have been used to predict power allocation with lower complexity, while deep reinforcement learning (DRL)-based methods, such as deep Q-network (DQN) and deep deterministic policy gradient (DDPG) [14], have been applied to dynamic power-control problems in CF-mMIMO networks [15]. These studies are closely related to the present work because they address adaptive power allocation under changing network conditions. Nevertheless, many existing learning-based methods are designed around a fixed objective, such as sum spectral-efficiency maximization, or are limited to a specific transmission scenario. As a result, they do not fully address the need for a unified uplink power-control framework that can flexibly operate under different system priorities, including fairness-oriented, SINR-oriented, and throughput-oriented operation.

Based on the existing literature, three main research gaps can be identified. First, many classical and learning-based power-control methods are objective-specific and cannot easily switch among fairness maximization, SINR improvement, and sum-rate maximization. Second, conventional optimization-based methods can provide strong performance, but they often require repeated numerical solving, which limits their suitability for real-time dynamic CF-mMIMO operation. Third, several learning-based methods improve adaptability, but they do not sufficiently combine reinforcement learning with selective constrained optimization to improve objective performance while avoiding unnecessary online computational cost. These gaps motivate the development of a unified and computationally efficient framework for dynamic uplink power allocation under changing CF-mMIMO channel conditions. Table 1 summarizes the main related studies and highlights their objectives, methods, advantages, and limitations.

Table 1. Overview of existing literature on power control techniques in cell-free massive MIMO networks.

Ref.	Year	Method	Main Objective	Relevance to Uplink Power Control	Limitation/Research Gap	How This Study Differs
[6]	2025	Max-min power control	Fairness improvement	Addresses fairness-oriented uplink power allocation	May reduce total spectral efficiency when weak users are prioritized	Evaluates fairness-oriented behavior together with SINR- and throughput-oriented strategies
[7]	2024	WMMSE-based optimization	Spectral-efficiency improvement	Provides an optimization-based baseline for power-control design	Requires repeated iterative computation under changing channel conditions	Compares learning-based decision-making with optimization-based strategies under the same CF-mMIMO environment

https://doi.org/10.3390/eng7070357

Table 1. *Cont.*

Ref.	Year	Method	Main Objective	Relevance to Uplink Power Control	Limitation/Research Gap	How This Study Differs
[9]	2020	SEPFM/ proportional fairness	Spectral-efficiency–fairness balance	Studies the trade-off between throughput and user fairness	Limited adaptability under dynamic mobility and changing channel conditions	Evaluates fairness, SINR, spectral efficiency, and execution time under mobile CF-mMIMO conditions
[10]	2024	DQN-based power control	Dynamic power-control learning	Applies DRL to CF-mMIMO power-control problems	Discrete action representation may limit fine-grained power allocation	Uses continuous SAC action space for UE transmit-power control
[11]	2026	Hybrid PSO-SA optimization	Interference-aware power allocation	Shows the role of hybrid optimization in interference-aware wireless resource allocation	Designed for GFDM/OFDM cognitive radio networks and requires iterative search	This study focuses on dynamic uplink CF-mMIMO and combines SAC-based continuous control with objective-specific refinement
[12]	2021	DDPG-based DRL	Sum spectral-efficiency maximization	Studies dynamic CF-mMIMO power control using DRL	Mainly focuses on throughput-oriented learning	Considers fairness-oriented, SINR-oriented, and throughput-oriented evaluation metrics
[13]	2021	Deep learning-based power prediction	Fast power-control inference	Shows that learning can reduce online decision time	Performance depends strongly on training data and selected objective	Investigates SAC reward configurations and compares them with objective-specific optimization strategies
[15]	2021	DDPG-based power control	Low-complexity learning-based decision-making	Demonstrates the use of DRL for adaptive power-control decisions	Usually depends on a fixed objective or scenario-specific learning setting	Uses SAC-based continuous control and evaluates multiple reward configurations

To address these limitations, this study develops a Deep Hybrid Intelligent (DHI) framework for dynamic uplink power-control analysis in CF-mMIMO systems. The proposed framework uses Soft Actor-Critic (SAC) learning to generate continuous uplink transmit-power decisions under time-varying channel and mobility conditions. Instead of treating uplink power allocation as a single fixed-objective problem, the study evaluates objective-specific operating configurations that reflect different system priorities, including fairness-oriented, SINR-oriented, and spectral-efficiency-oriented behavior. Furthermore, we evaluate three distinct optimization algorithms for power control: max-min fairness, max-product SINR, and aggregate sum-rate maximization. These are incorporated to provide objective-guided comparison and to clarify the trade-off among user fairness, signal quality, total throughput, and computational complexity. The max-product and max-sum-rate objectives are solved using L-BFGS-B optimization, whereas the max-min strategy is evaluated through a feasibility-based fairness-oriented procedure. This design allows learning-based and optimization-based power-control behavior to be examined under a common dynamic CF-mMIMO simulation environment.

The core contributions of this work are outlined below:

- Design a dynamic uplink power allocation architecture for CF-mMIMO networks using the Soft Actor-Critic (SAC) algorithm, where the learning agent generates continuous UE transmit-power decisions under time-varying channel and mobility conditions.
- Objective-specific reward configurations are investigated to support different uplink power-control goals, including fairness-oriented, SINR-oriented, and spectral-efficiency-oriented operation within a common CF-mMIMO simulation environment.
- Three optimization-based power-control strategies are incorporated for objective-guided comparison, namely max-min fairness, max-product SINR optimization, and

max-sum-rate maximization, enabling a clearer assessment of the trade-off among fairness, signal quality, and throughput.
- L-BFGS-B optimization is employed for the max-product and max-sum-rate objectives, while the max-min strategy is evaluated through a feasibility-based fairness-oriented optimization procedure. This allows the computational behavior of learning-based and optimization-based power-control approaches to be compared under the same network configuration.
- The proposed evaluation framework is assessed using multiple performance indicators, including uplink power distribution, SINR, spectral efficiency, user-rate behavior, Jain's fairness index, and execution time, providing a broader view of the performance–complexity trade-off in dynamic uplink CF-mMIMO systems.

This paper is structured in the following manner. Section 2 outlines the foundational system model and mathematical formulation before introducing the proposed DHI framework, learning algorithm, and optimization methodology. It also covers the solution approach, computational complexity, and simulation configuration. Next, Section 3 analyzes the simulation outcomes and contrasts the performance of our strategies with existing benchmarks. Section 4 wraps up the paper by summarizing conclusions, limitations, and future research paths.

2. DHI Model

The CF-mMIMO system considered in this study consists of distributed APs that jointly serve multiple single-antenna UEs over a large coverage area. The APs are connected to a central processing unit (CPU) through fronthaul links, allowing the received uplink signals to be processed jointly. This study focuses exclusively on the uplink transmission scenario, where each UE transmits its data symbols to the distributed APs, and the CPU performs joint signal processing and performance evaluation. The uplink channel is modeled using path loss, lognormal shadow fading, and small-scale Rayleigh fading, while the SINR and spectral efficiency are computed based on the resulting signal and interference conditions.

The architecture of the proposed DHI scheme for managing CF-mMIMO uplink power is mapped out in Figure 2. Within this framework, an actor-critic mechanism drives the SAC agent's environmental interactions: the actor network determines continuous uplink power levels, while the critic network assesses the performance of these decisions. Once an action is constrained within the valid UE transmission power limits, the environment updates its channel states to calculate the resulting SINR, spectral efficiency, fairness metrics, and reward values. These interaction experiences are saved in a replay buffer, which provides samples during training updates to iteratively optimize both networks.

Once the SAC actor outputs an initial valid power vector, a dedicated refinement module evaluates or tunes this vector using three distinct criteria: max-min, max-product, and max-sum-rate. We implement the fairness-focused max-min approach via a feasibility-checked bisection method. Conversely, we solve the max-product and max-sum-rate formulations using the L-BFGS-B algorithm under individual UE power thresholds. This dual-stage processing highlights the hybrid nature of our framework, combining the rapid initial decision-making of reinforcement learning with targeted algorithmic refinement from the optimization module. Instead, it provides a common evaluation structure for comparing learning-based and optimization-based uplink power-control behaviors in terms of SINR, spectral efficiency, Jain's fairness index, and execution time. This structure allows the study to analyze the trade-off among fairness, signal-quality improvement, throughput maximization, and computational efficiency in a dynamic uplink CF-mMIMO environment.

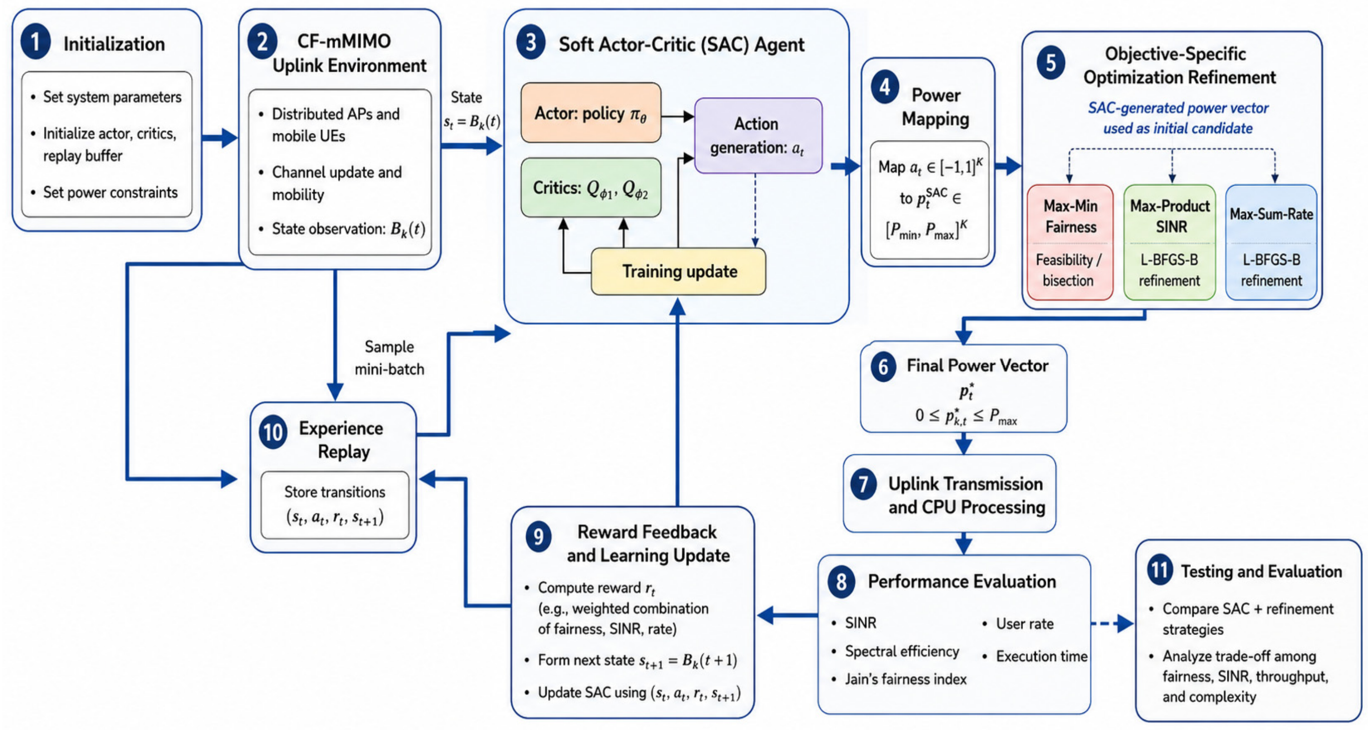

Figure 2. Structural diagram of the developed DHI framework combining SAC-driven uplink power control and target-specific optimization comparisons in CF-mMIMO networks.

3. Mathematical Model

In this uplink CF-mMIMO configuration, M distributed APs serve K single-antenna UEs. These APs route data to a central processing unit (CPU) for centralized signal processing [3]. We represent the uplink transmission power of the $k-th$ UE during time interval t as $P_k(t)$ [16], where

$$P_{min} \le P_k(t) \le P_{max}, \qquad k = 1, 2, \ldots, K. \tag{1}$$

The uplink power-allocation vector is written as [17]:

$$P(t) = [p_1(t), p_2(t), \ldots, p_K(t)]^T \tag{2}$$

The channel coefficient between AP m and UE k is represented as [18]:

$$h_{m,k}(t) = \sqrt{\beta_{m,k}(t)} g_{m,k}(t) \tag{3}$$

In this formulation, the small-scale Rayleigh fading is captured by $g_{m,k}(t)$, whereas $\beta_{m,k}(t)$ represents large-scale fading, combining path loss and shadowing. Because user movement continuously alters the distance between APs and UEs, the simulation updates these large-scale fading coefficients dynamically at each step.

The received uplink signal at AP m can be expressed as [10]:

$$\mathcal{Y}_m(t) = \sum_{i=1}^{k} \sqrt{p_i(t)} h_{m,i}(t) s_i(t) + n_m(t) \tag{4}$$

where $s_i(t)$ is the transmitted symbol of UE i, and $n_m(t)$ is the additive noise at AP m. At the CPU, the received signals from the distributed APs are combined to detect the transmitted signal of each UE.

To express the effective uplink SINR for the $k-th$ UE as [18]:

$$\gamma_k(t) = \frac{p_k(t)\left|\sum_{m=1}^{M} v_{m,k}^*(t)h_{m,k}(t)\right|^2}{\sum_{i=1,i\neq k}^{K} p_i(t)\left|\sum_{m=1}^{M} v_{m,i}^*(t)h_{m,i}(t)\right|^2 + \sigma^2\sum_{m=1}^{M}\left|v_{m,k}(t)\right|^2} \quad (5)$$

where $v_{m,k}(t)$ is the receive combining coefficient associated with UE k at AP m, and σ^2 denotes the noise power. This expression shows that the SINR of each UE depends not only on its own transmit power, but also on the interference produced by the other UEs.

To find the uplink spectral efficiency achieved by UE k, we apply the formulation [16]:

$$SE_k(t) = \frac{\tau_u}{\tau_c} log_2(1+\gamma_k(t)), \quad (6)$$

where τ_u denotes the number of uplink data symbols within one coherence block, and τ_c denotes the coherence block length in symbols. In this study,τ_c is used as a common spectral-efficiency normalization parameter for the simulated uplink transmission interval; it should not be interpreted as a fixed 10 ms frame duration. User mobility is captured through time-varying AP-UE distances and updated large-scale fading values, while the same coherence-block structure is assumed for all UEs during each simulation time step.

$$SE_{sum}(t) = \sum_{k=1}^{K} SE_k(t) \quad (7)$$

To evaluate user-service balance, Jain's fairness index is used [19]:

$$J(t) = \frac{\left(\sum_{k=1}^{K} SE_k(t)\right)^2}{K\sum_{k=1}^{K} SE_k^2(t)} \quad (8)$$

The value of $J(t)$ lies between $1/K$ and 1, where a value closer to 1 indicates a more uniform distribution of spectral efficiency among UEs.

Based on these performance measures, three objective-specific uplink power-control strategies are considered. The first is the max-min fairness objective, which aims to improve the weakest UE rate:

$$\underset{P}{max}\ \underset{k}{min}\, SE_k(t) \quad (9)$$

subject to

$$P_{min} \leq P_k(t) \leq P_{max}, \qquad \forall k. \quad (10)$$

The second is the max-product SINR objective, which improves the collective SINR behavior of the system [10]:

$$\underset{p}{max} \prod_{k=1}^{K}(1+\gamma_k(t)) \quad (11)$$

subject to

$$P_{min} \leq P_k(t) \leq P_{max}, \qquad \forall k. \quad (12)$$

Equivalently, this objective can be solved in logarithmic form to improve numerical stability:

$$\underset{p}{max} \sum_{k=1}^{K} log_2(1+\gamma_k(t)) \quad (13)$$

The third is the max-sum-rate objective, which maximizes the total uplink spectral efficiency [10]:

$$\max_{p} \sum_{k=1}^{K} SE_k(t), \tag{14}$$

subject to

$$P_{min} \leq P_k(t) \leq P_{max}, \qquad \forall k. \tag{15}$$

When a total uplink power budget is imposed, the following additional constraint can be considered:

$$\sum_{k=1}^{K} p_k(t) \leq P_{tot} \tag{16}$$

These objective functions provide the mathematical basis for evaluating the trade-off among fairness, SINR improvement, and spectral-efficiency maximization in the considered uplink CF-mMIMO system.

4. Learning Model Description

This study integrates the Soft Actor-Critic (SAC) algorithm to handle the learning tasks required for dynamic uplink power management in cell-free massive MIMO systems. SAC is uniquely suited for continuous control environments where the decision variable is a continuous transmit-power vector rather than a set of discrete actions. Step by step, the agent monitors the instantaneous channel states of the system and assigns specific uplink power levels to all active users.

The observation vector used by the SAC agent is represented by the aggregated large-scale fading information of the UEs. Let the state at time step t be denoted as [20]:

$$s_t = B_k(t) \tag{17}$$

where $B_k(t) = [B_1(t), B_2(t), \ldots, B_k(t)]$ represents the UE-level large-scale fading vector. Although UE locations and mobility parameters are not used as explicit SAC input features, they influence the environment indirectly through distance-dependent channel variations. This is because UE movement changes the AP–UE distances over time, which in turn modifies the large-scale fading coefficients, interference behavior, SINR, and spectral efficiency. Therefore, the SAC agent learns power-control decisions under dynamic channel conditions produced by user mobility.

In the present implementation, the SAC observation vector is based on the UE-level aggregated large-scale fading state $B_k(t)$. This choice provides a compact and slowly varying representation of the CF-mMIMO channel condition and reduces the dimensionality of the learning problem. However, it should be noted that $B_k(t)$ does not explicitly represent all instantaneous physical-layer effects. In particular, small-scale fading, instantaneous interference coupling, previous SINR, previous power allocation, and UE trajectory information are not directly included as separate input features to the SAC actor. Instead, these effects are reflected indirectly through the environment update, where UE mobility changes the AP-UE distances, updates the channel-dependent quantities, and affects the resulting SINR, spectral efficiency, and reward.

To clarify this modeling choice, the current SAC formulation can be interpreted as a compact partially observed state representation rather than a complete channel-state representation. A more informative observation vector for future implementation can be expressed as:

$$s_t^{aug} = [B_t\ ,\ \gamma_{t-1}\ ,\ I_{t-1}\ , P_{t-1}\ ,\ \Delta u_t] \tag{18}$$

where B_t denotes the large-scale fading vector, γ_{t-1} denotes the previous SINR vector, I_{t-1} denotes the previous interference indicator, P_{t-1} denotes the previous uplink power-allocation vector, and Δu_t represents mobility-related displacement or trajectory information. This augmented state representation would allow the SAC agent to observe both slow channel variations and short-term interference/mobility effects more explicitly. Therefore, extending the SAC observation space in this direction is identified as an important improvement for future work.

The SAC actor network generates a continuous action vector [21],

$$a_t = [a_1(t), a_2(t), \ldots, a_k(t)] \tag{19}$$

where each action component satisfies $a_k(t) \in [-1, 1]$. Since the physical uplink transmit power must satisfy the feasible power range, the action is rescaled into the actual UE transmit-power vector as [22]:

$$p_k(t) = \frac{a_k(t)+1}{2}(P_{max} - P_{min}) + P_{min}, \tag{20}$$

where $p_k(t)$ is the uplink transmit power of UE k, P_{min} is the minimum transmit power, and P_{max} is the maximum allowable transmit power. The resulting power vector is expressed as [15]:

$$P_t = [p_1(t), p_2(t), \ldots, p_k(t)]^T \tag{21}$$

Once the system applies the power vector, it updates user positions (if mobility is enabled), recalculates the channel states, and determines the resulting SINR and spectral efficiency. The reward is then calculated according to the selected reward configuration. In this implementation, the reward design supports direct performance-based objectives, such as minimum spectral efficiency, mean spectral efficiency, sum spectral efficiency, geometric-mean spectral efficiency, and channel-capacity-based reward. It also supports temporal reward formulations based on either SINR or spectral efficiency, where the spectral efficiency is computed from the uplink SINR of each user equipment (UE) according to the adopted system model. These reward configurations allow the learning process to emphasize different performance behaviors without changing the basic SAC control structure.

During training, the SAC agent stores the transition tuple

$$(s_t, a_t, r_t, s_{t+1}) \tag{22}$$

After saving the transition in the replay buffer, the system applies off-policy learning to update both the actor and critic networks. The critic networks estimate the expected return of the selected power-control action, while the actor network is updated to maximize the expected reward with entropy regularization. The entropy term encourages exploration and reduces the risk of premature convergence to poor power-allocation patterns. After training, the learned SAC policy can generate uplink power-control decisions directly from the observed fading state, which reduces the need for repeated numerical optimization during online inference.

5. Optimization Problem

The dynamic uplink power allocation problem is formulated as a continuous-control reinforcement learning problem. At each time step t, the SAC agent observes the channel-dependent state s_t and generates an action vector a_t,, which is mapped into the feasible uplink power-allocation vector P_t. The objective of the learning agent is to find a policy π_θ that maximizes the expected discounted return over time while satisfying the transmit-power constraints of all UEs.

https://doi.org/10.3390/eng7070357

The general reinforcement learning objective is expressed as:

$$\pi_\theta^* = \underset{\pi_\theta}{argmax} \mathbb{E}_{\pi_\theta}\left[\sum_{t=0}^{\infty} \gamma^t r_t\right] \tag{23}$$

where π_θ is the SAC policy parameterized by θ, r_t is the immediate reward at time step t, and $\gamma \in [0,1]$ is the discount factor. The generated power allocation must satisfy the individual UE power constraints:

$$P_{min} \le P_k(t) \le P_{max}, \qquad \forall k \in \{1,2,\ldots,K\} \tag{24}$$

When a total uplink power budget is considered, the following additional constraint can be included [22]:

$$\sum_{k=1}^{K} p_k(t) \le P_{tot} \tag{25}$$

In this study, the reward function is not restricted to a single fixed mathematical form. Instead, several reward configurations are used to examine how different learning objectives affect uplink power-control behavior. The direct performance-based reward configurations are defined as follows.

For the channel-capacity-based reward, the reward is computed from the sum of logarithmic SINR terms [23]:

$$r_t^{cap} = \sum_{k=1}^{K} log_2(1 + \gamma_k(t)) \tag{26}$$

where $\gamma_k(t)$ denotes the uplink SINR of UE k at time step t.

For the minimum spectral-efficiency reward, the objective emphasizes the weakest UE [24]:

$$r_t^{minSE} = \underset{k}{min} SE_k(t) \tag{27}$$

For the mean spectral-efficiency reward, the reward is expressed as [22]:

$$r_t^{meanSE} = \frac{1}{K}\sum_{k=1}^{K} SE_k(t) \tag{28}$$

For the sum spectral-efficiency reward, the objective emphasizes total throughput:

$$r_t^{sumSE} = \sum_{k=1}^{K} SE_k(t) \tag{29}$$

For the geometric-mean spectral-efficiency reward, the reward is written as:

$$r_t^{geoSE} = \left(\prod_{k=1}^{K} SE_k(t)\right)^{1/K} \tag{30}$$

These direct reward forms allow the SAC policy to emphasize different operating behaviors. The minimum-SE reward encourages fairness-oriented power allocation, the sum-SE reward encourages throughput-oriented behavior, while the mean-SE and geometric-mean rewards provide intermediate trade-offs between user balance and total spectral efficiency.

In addition to direct performance-based rewards, temporal reward formulations are used to evaluate improvement relative to historical performance. Let $x_k(t)$ represent the selected performance indicator, which may be either SINR or spectral efficiency. Let $\overline{x}_k(t)$

denote the moving average of the same indicator over a temporal window of size W. The delta-based temporal reward can be expressed as:

$$r_t^{\triangle} = \mathcal{O}(x_k(t) - \overline{x}_k(t)) \tag{31}$$

where $\mathcal{O}(\cdot)$ denotes an aggregation operation such as mean, sum, minimum, maximum, or geometric mean. The relative temporal reward is given by [21]:

$$r_t^{rel} = \mathcal{O}\left(\frac{x_k(t)}{\overline{x}_k(t) + \epsilon}\right) \tag{32}$$

where ϵ epsilonϵ is a small positive value used to avoid division by zero. Exponential and logarithmic variants are also considered to control the sensitivity of the reward to changes in SINR or spectral efficiency. These temporal reward configurations allow the learning process to capture whether the current power-allocation decision improves or degrades the system performance compared with recent historical behavior.

Accordingly, the SAC optimization problem can be summarized as [25]:

$$\underset{\pi_\theta}{max}\mathbb{E}_{\pi_\theta}\left[\sum_{t=0}^{\infty} \gamma^t r_t^{(q)}\right] \tag{33}$$

subject to

$$P_{min} \leq P_k(t) \leq P_{max,} \qquad \forall k, \tag{34}$$

where $r_t^{(q)}$ denotes the selected reward configuration. This formulation allows the same SAC learning structure to be evaluated under different objective-specific reward designs, including fairness-oriented, SINR-oriented, and spectral-efficiency-oriented behavior.

6. Solution Method

The proposed solution method is implemented through a common CF-mMIMO simulation environment in which learning-based and optimization-based power-control strategies are evaluated under the same network configuration. The solution procedure consists of two main evaluation paths: the SAC-based dynamic power-control path and the objective-specific optimization path.

In the SAC-based path, the learning agent observes the current channel-dependent state st, represented by the UE-level large-scale fading vector. The actor network generates a continuous action vector, which is rescaled into the feasible uplink transmit-power vector p(t). After applying this power vector, the environment updates the UE positions when mobility is enabled, recalculates the channel-dependent variables, and computes the resulting SINR and spectral efficiency. The reward is then obtained according to the selected reward configuration, and the transition (st, at, rt, st + 1) is stored in the replay buffer for actor–critic updates. In this way, the SAC agent learns a dynamic mapping between the observed fading state and the uplink power-allocation decision.

In parallel with the SAC-based learning path, three optimization-based power-control strategies are implemented for objective-guided comparison. The first strategy is the max-min fairness method, which focuses on user fairness by maximizing the lowest achievable transmission rate across the network. This method is evaluated using a feasibility-based bisection procedure, where candidate SINR or rate values are tested under the transmit-power constraints until a feasible power vector is obtained. The max-min solution is therefore used to represent fairness-oriented optimization behavior.

The second strategy is the max-product SINR method. This strategy aims to improve the collective SINR behavior of the system by maximizing the product-related SINR objective across users. In the implementation, the max-product problem is solved using the L-BFGS-B algorithm under individual UE power bounds. The optimization is initialized using a uniform power vector and then iteratively updated to obtain a bounded solution that improves the selected SINR-based objective.

Lastly, the max-sum-rate strategy focuses on maximizing the collective spectral efficiency of the network. Similar to the max-product strategy, the max-sum-rate problem is solved using L-BFGS-B with individual transmit-power constraints. The objective function is formulated as the negative sum spectral efficiency so that the minimization procedure of L-BFGS-B can be used to obtain a power vector that increases the total achievable rate.

Based on this structure, the proposed evaluation framework does not assume that a single SAC policy directly solves all three optimization objectives. Instead, the SAC-based policy and the three optimization-based strategies are evaluated within the same CF-mMIMO environment to compare their behavior under dynamic uplink conditions. This design allows the study to analyze the trade-off among fairness, SINR improvement, throughput maximization, and computational execution time. It also provides a clearer distinction between learning-based power-control decisions and classical optimization-based solutions, as shown in Figure 3.

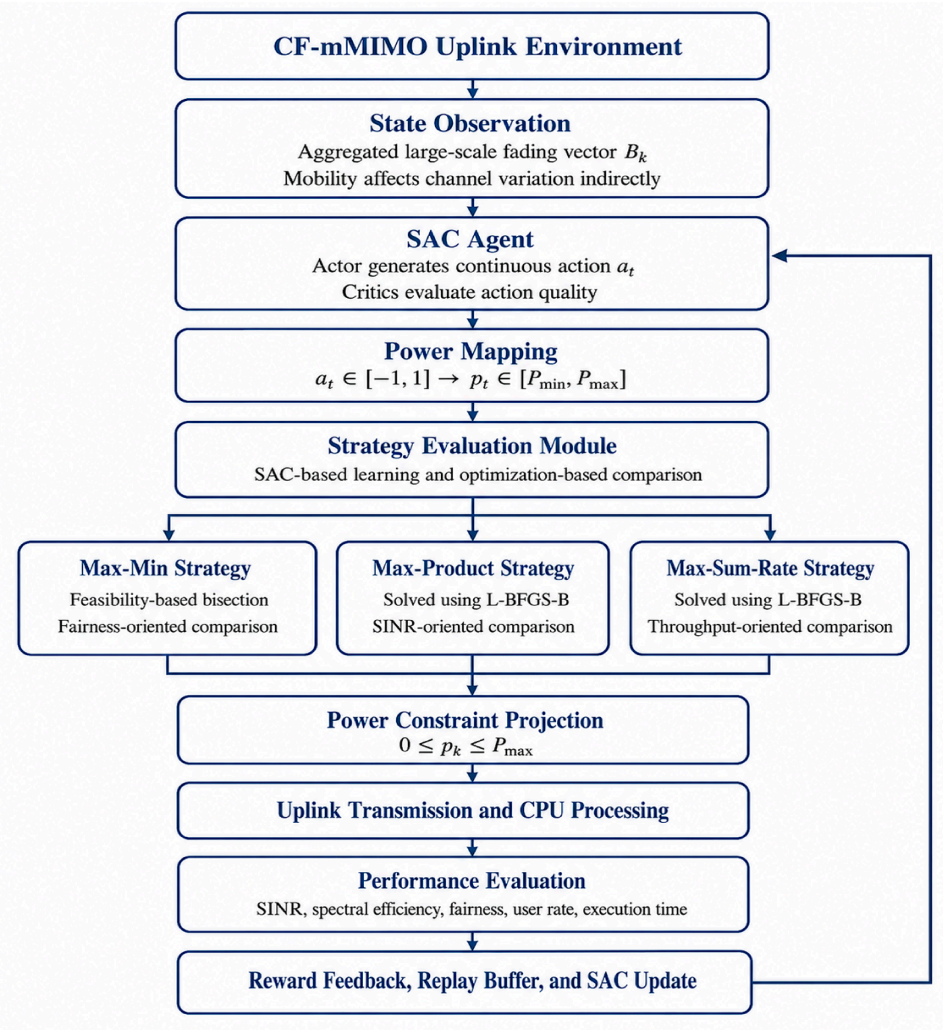

Figure 3. Structural layout of power control in DHI model.

7. Complexity Analysis

Two main elements determine the computational complexity of the evaluation framework: the SAC learning module and the optimization-based power control. For a CF-mMIMO network with M APs and K UEs, four key factors dictate this processing cost: the total AP-UE links, the number of users undergoing power tuning, the size of the SAC agent's neural networks, and the iteration counts required by the benchmark algorithms.

For the SAC-based component, the main offline training cost comes from repeated interaction with the CF-mMIMO environment and from updating the actor and critic networks. At every interval, the system updates the channel-dependent conditions, reviews the implemented power vector, and determines the resulting SINR, spectral efficiency, and reward values. The cost of evaluating the channel and interference terms generally increases with the AP–UE dimensions and can be expressed as a function of M and K. In addition, the SAC update requires forward and backward propagation through the actor and critic networks. If C_{NN} denotes the computational cost of one neural-network update and T_{tr} denotes the number of training time steps, the overall SAC training complexity can be approximately expressed as [15]:

$$O(T_{tr}(C_{env}(M,K) + C_{NN})), \tag{35}$$

where $C_{env}(M,K)$ represents the environment evaluation cost, including channel update, SINR calculation, spectral-efficiency calculation, and reward computation.

After training, the SAC inference stage is significantly less expensive because it only requires a forward pass through the actor network to generate the uplink power vector. Therefore, the online inference complexity of the SAC policy can be written as [20]:

$$O(C_{actor}), \tag{36}$$

where C_{actor} is the cost of the actor-network forward pass. Since the actor output dimension is K, the final action-to-power mapping has a linear cost of $O(K)$. Thus, the trained SAC policy can generate power-allocation decisions without repeatedly solving a numerical optimization problem at every time step.

For the max-min optimization strategy, the power-allocation problem is solved using a feasibility-based bisection procedure. If N_{bis} denotes the number of bisection iterations and $C_{feas}(K)$ denotes the cost of solving the feasibility problem for K UEs, the complexity can be approximated as [16]:

$$O(N_{bis}\ C_{feas}(K)\Big) \tag{37}$$

This strategy can provide strong fairness-oriented behavior, but its repeated feasibility solving increases the computational burden, especially when the number of UEs increases or when the algorithm is executed repeatedly under mobility.

To resolve the bounded optimization challenges in the max-product and max-sum-rate formulations, we employ the L-BFGS-B method. Let N_{LBFGS} denote the number of L-BFGS-B iterations and $C_{obj}(M,K)$ denote the cost of evaluating the selected objective function and its associated SINR or spectral-efficiency terms. The corresponding complexity can be approximated as [26]:

$$O\Big(N_{LBFGS}\, C_{obj}(M,K)\Big) \tag{38}$$

The memory cost of L-BFGS-B is lower than full BFGS because it stores only a limited number of correction vectors rather than the full Hessian matrix. Therefore, it is more suitable for bounded power-control problems with multiple UEs. However, it still requires iterative objective evaluations during online optimization.

Based on this analysis, the SAC-based approach shifts most of the computational burden to the offline training phase, while the trained policy provides fast online inference. In contrast, the optimization-based strategies can provide objective-specific solutions but require repeated numerical solving when the channel conditions change. This difference explains why the proposed evaluation considers four main performance and complexity indicators: SINR, spectral efficiency, Jain's fairness index, and execution time. This comparison highlights the trade-off between objective-specific performance and online execution-time efficiency in dynamic uplink CF-mMIMO systems.

To complement the theoretical complexity analysis, an implementation-level summary is provided to clarify the practical computational cost of the proposed DHI framework. The SAC component shifts most of the computational burden to the offline training stage, whereas the optimization-based methods incur repeated online iterative computation. In the present implementation, the SAC policy uses a compact multilayer perceptron with two hidden layers, and the replay buffer stores 100,000 transitions. Based on the used state and action dimensions, the replay-buffer storage requirement is approximately tens of megabytes, while the neural-network parameter count remains relatively small compared with the iterative optimization cost. The measured execution time of each strategy is reported in the Section 9 to quantify the online computational cost required to generate uplink power-control decisions.

Table 2 presents the computational overhead for the DHI framework. The replay buffer drives the main memory footprint, while the policy and critic networks remain lightweight. This confirms that the SAC-based component is computationally lightweight during online inference, whereas the optimization-based strategies require repeated iterative solving. Therefore, the proposed framework is more suitable for scenarios where a one-time offline training cost is acceptable in exchange for lower online decision complexity.

Table 2. Implementation-level computational summary of the proposed DHI framework.

Component	Quantitative Summary	Practical Implication
SAC actor network	Two hidden layers [64, 64]; approximately 10,432 parameters	Compact policy network with low inference cost
SAC twin critics	Two critic networks with approximately 16,770 trainable parameters in total	Moderate offline training cost
Replay buffer	100,000 transitions	Supports off-policy learning
Replay-buffer memory	Approximately 39 MB (float32-based estimate)	Moderate memory requirement
SAC inference stage	Single forward pass plus action-to-power mapping	Low online decision cost
Max-Min strategy	Feasibility-based bisection	Higher online iterative cost
Max-Product strategy	L-BFGS-B iterative refinement	Higher online iterative cost than SAC inference
Max-Sum-Rate strategy	L-BFGS-B iterative refinement	Higher online iterative cost than SAC inference
Execution time	Reported in the corresponding results table	Supports practical comparison among strategies

Pseudo Algorithm 1 outlines the evaluation procedure for the DHI framework. The SAC agent learns to map observed channel states directly to a continuous UE transmit-power vector. For a direct comparison, we test the max-min, max-product, and max-sum-rate strategies under identical CF-mMIMO conditions. We solve the max-min problem using a feasibility-based bisection method, and handle the max-product and max-sum-rate problems using L-BFGS-B. This setup ensures a fair comparison between learning-based and optimization-based power control regarding SINR, spectral efficiency, fairness, and execution speed.

Algorithm 1. Dynamic uplink power allocation via the proposed DHI framework

```
Initialize SAC actor, critic networks, target critics, and replay buffer
Initialize the CF-mMIMO environment and power constraints
Initialize AP and UE locations, large-scale fading, signal, and interference terms
Select the SAC reward configuration for uplink power-control learning
for each episode do
    Reset the environment and observe the initial state st
    for each time step t do
        Sample action at from the SAC policy
        Map at to the uplink power vector pt
        Apply pt in the CF-mMIMO environment
        Update UE positions mobility
        Compute SINR, spectral efficiency, fairness, and reward rt
        Observe the next state st + 1
        Store (st, at, rt, st + 1) in the replay buffer
        Update the critic networks, update the actor network, and target critics
   Evaluate the max-min optimization strategy using feasibility-based bisection
      Evaluate the max-product optimization strategy using L-BFGS-B
      Evaluate the max-sum-rate optimization strategy using L-BFGS-B
   Record uplink power, SINR, spectral efficiency, fairness, and execution time
        Set st ← st + 1
    end for
end for
```

8. Simulation Setup

An uplink CF-mMIMO simulation environment featuring distributed APs and mobile UEs served as the testbed to evaluate the proposed framework. The network consisted of M = 64 single-antenna APs and K = 32 single-antenna UEs randomly distributed over a 1000 × 1000 m simulation area. The APs jointly served the UEs through a central processing unit, and uplink power-control decisions were evaluated under time-varying channel conditions caused by user mobility.

Large-scale fading was modeled using a three-slope path-loss model with lognormal shadowing. Simulation parameters include an 8 dB shadow fading standard deviation and a 100 m decorrelation distance. Independent Rayleigh fading represents small-scale fading, while the system uses a 20 MHz bandwidth and a 9 dB noise figure, resulting in an approximate noise power of −92 dBm. In the present simulation, the uplink SINR is calculated according to the adopted CF-mMIMO signal and interference model, without imposing an explicit hardware-related SINR ceiling. Therefore, no error-vector-magnitude (EVM)-based SINR clipping or maximum SINR saturation threshold is applied. The reported SINR values should therefore be interpreted as idealized simulation results under the assumed channel, interference, and noise model. Practical RF impairments and EVM-limited SINR saturation are considered as possible extensions for future work. The number of orthogonal pilot symbols, denoted by τp, was set to 20. Each UE transmit power was constrained within the range 0–100 mW, with the initial UE power set to 100 mW.

User mobility was modeled using the random-waypoint mobility model. The simulation assigns each UE a random speed chosen uniformly between 0.5 and 5 m/s, using a 1 s mobility time step. Mobility was enabled during evaluation to capture changes in AP–UE distances, large-scale fading, SINR, and spectral-efficiency behavior over time. Temporal reward smoothing was applied using a window size of 10. The parameters

$r\alpha = 1$ and $r\beta$ denote reward-weighting coefficients used to balance the contribution of the current performance term and the temporal-improvement term in the reward calculation, respectively. In this study, both coefficients were set to 1, i.e., $r\alpha = 1$ and $r\beta = 1$, giving equal weighting to the two reward components.

The simulation parameters used in this study are summarized in Table 3.

Table 3. Simulation parameters of the uplink CF-mMIMO system.

Parameter	Value
Number of APs, M	64
Number of UEs, K	32
Simulation area	1000×1000 m
Number of orthogonal pilots, τp	20
Path-loss model	Three-slope model
Shadow fading standard deviation, σsf	8 dB
Shadow fading decorrelation distance	100 m
Small-scale fading	Rayleigh fading
System bandwidth	20 MHz
Noise figure	9 dB
Noise power	≈ -92 dBm
UE transmit-power range	0–100 mW
Initial UE transmit power	100 mW
Mobility model	Random waypoint
UE speed range	0.5–5 m/s
Mobility time step	1 s
Temporal reward window size	10
Temporal reward-weighting coefficients, $r\alpha$, $r\beta$	1, 1

The DHI hyperparameters were tuned using Optuna over predefined search ranges. Hyperparameter tuning was performed to identify a stable training configuration for the dynamic uplink power-control environment. The search space included the discount factor γ, learning rate, batch size, replay buffer size, learning-start threshold, training frequency, Polyak averaging coefficient τ, gradient-update frequency, entropy coefficient, target entropy, network architecture, and optimizer. In addition, SGD was used as the optimizer in the implemented SAC configuration. Although adaptive optimizers are commonly used in many SAC implementations, SGD was retained in this study after the hyperparameter-tuning stage because it provided stable training behavior for the considered compact actor-critic network and large replay-buffer setting. The main hyperparameter search space is summarized in Table 4.

Table 4. DHI hyperparameter search space used for Optuna tuning.

Hyperparameter	Setting
Discount factor, γ	{0.9, 0.95, 0.98, 0.99, 0.995, 0.999, 0.9999}
Learning rate	10^{-5} to 1, logarithmic sampling
Batch size	{16, 32, 64, 128, 256, 512, 1024, 2048}
Replay buffer size	$\{10^4, 10^5, 10^6\}$
Learning starts	{0, 10, 100, 1000}
Training frequency	{1, 4, 8, 16, 32, 64, 128, 256, 512}
Gradient steps	Equal to training frequency
Polyak coefficient, τ	{0.001, 0.005, 0.01, 0.02, 0.05, 0.08}
Entropy coefficient	Auto
Target entropy	Auto
Network architecture	Small, medium, big, mixed, or large
Optimizer	SGD
Temporal reward window	10
Temporal reward-weighting coefficients, $r\alpha$, $r\beta$	1, 1

9. Results and Discussion

Five key indicators measure the performance of the proposed DHI-based uplink power-control framework: uplink power distribution, spectral efficiency, SINR, Jain's fairness index, and computational execution time. These metrics were selected because uplink power control in CF-mMIMO should not be assessed only by throughput improvement. A practical power-control method must also control inter-user interference, maintain reliable signal quality, support balanced service among UEs, and reduce the computational burden required for online decision-making.

The evaluated strategies represent different operating behaviors. The fairness-oriented configuration emphasizes the protection of weaker UEs and aims to improve service balance. By focusing on the system's collective SINR, the max-product configuration strikes a practical balance between user fairness and signal quality. The max-sum-rate configuration prioritizes total spectral-efficiency improvement and user-rate enhancement, but it may produce a wider performance variation among users. Therefore, the following results are analyzed not only to identify the best-performing strategy for a single metric, but also to clarify the trade-off among fairness, SINR, throughput, and computational complexity.

To address the effect of reward design on the SAC-based power-control behavior, a reward-configuration ablation analysis was added. The purpose of this analysis is to clarify how each reward formulation contributes to the final operating behavior of the proposed DHI framework. As summarized in Table 5, the minimum-SE reward emphasizes weak-user protection and is therefore more suitable for fairness-oriented operation. The sum-SE reward emphasizes total throughput and supports rate-oriented power allocation. The mean-SE and geometric-mean-SE rewards provide intermediate behavior by balancing average service quality and user-service uniformity. The channel-capacity reward encourages SINR-based rate improvement, while the temporal reward formulations help evaluate whether the current power-control decision improves the recent historical performance under mobility and time-varying channel conditions. This ablation clarifies that the final behavior of the SAC agent is strongly influenced by the selected reward configuration, and that no single reward is optimal for all objectives.

Table 5. Qualitative ablation of SAC reward configurations.

Reward Configuration	Main Optimization Emphasis	Expected Contribution to Final Behavior	Possible Trade-Off
Minimum-SE reward	Weak-user protection	Improves lower-tail user performance and supports fairness-oriented power allocation	May reduce total spectral efficiency
Mean-SE reward	Average service quality	Improves average user spectral efficiency	May not fully protect the weakest users
Sum-SE reward	Throughput maximization	Increases total spectral efficiency and user-rate performance	May reduce fairness for weak users
Geometric-mean-SE reward	Balance between fairness and throughput	Provides a compromise between weak-user protection and total spectral efficiency	May not maximize either fairness or throughput
Channel-capacity reward	SINR/rate improvement	Encourages higher SINR-based rate behavior	May favor users with stronger channels
Temporal reward	Improvement over recent history	Stabilizes learning under mobility and time-varying channel conditions	Sensitive to window size and scaling

To clarify the evaluation logic, the three DHI-based strategies are first summarized according to their main objective, expected performance strength, and possible limitation, as shown in Table 6. This comparison is important because the proposed framework is not intended to show that one strategy dominates all metrics. Instead, each strategy reflects a different operating priority in uplink CF-mMIMO power control.

Table 6. Performance role of the evaluated DHI-based strategies.

Strategy	Main Objective	Main Strength	Possible Limitation
DHI-Max-Min	Fairness-oriented uplink power control	Protects weaker UEs and improves service balance	May reduce total throughput compared with throughput-oriented strategies
DHI-Max-Product	SINR-oriented balanced optimization	Improves collective SINR behavior and provides a compromise between fairness and throughput	May not achieve the highest sum spectral efficiency
DHI-Max-Sum-Rate	Total throughput maximization	Achieves stronger user-rate and spectral-efficiency performance	May increase performance variation among UEs

Table 7 summarizes the expected best strategy for each performance metric. This table provides a compact interpretation of the trade-off observed in the simulation results and helps connect the numerical results with the practical objective of each power-control strategy.

Table 7. Best strategy according to the main evaluation metrics.

Performance Metric	Best-Performing Strategy	Explanation
Jain's fairness index	DHI-Max-Min	The strategy prioritizes weaker UEs and therefore improves service uniformity
Weak-user protection	DHI-Max-Min	The max-min objective focuses on enhancing the performance of the worst-served user
SINR behavior	DHI-Max-Product/ DHI-Max-Sum-Rate	These strategies allocate power to improve signal-quality-related performance
Total spectral efficiency	DHI-Max-Sum-Rate	The objective directly maximizes the total achievable rate
User-rate performance	DHI-Max-Sum-Rate	Throughput-oriented allocation favors stronger rate improvement
Balanced performance	DHI-Max-Product	Provides an intermediate operating point between fairness and throughput
Computational efficiency	SAC-based learned configuration	The trained policy reduces the need for repeated numerical solving during inference

Figure 4 presents the cumulative distribution of uplink power under the DHI-Max-Min, DHI-Max-Product, and DHI-Max-Sum-Rate strategies. The figure shows how each strategy distributes transmit power among the UEs according to its optimization objective. In Figure 4a, the mean UE uplink power under DHI-Max-Min is more tightly distributed, indicating that this strategy tends to allocate power more uniformly across users. This behavior reflects the fairness-focused nature of max-min power control, which prevents large performance imbalances between UEs.

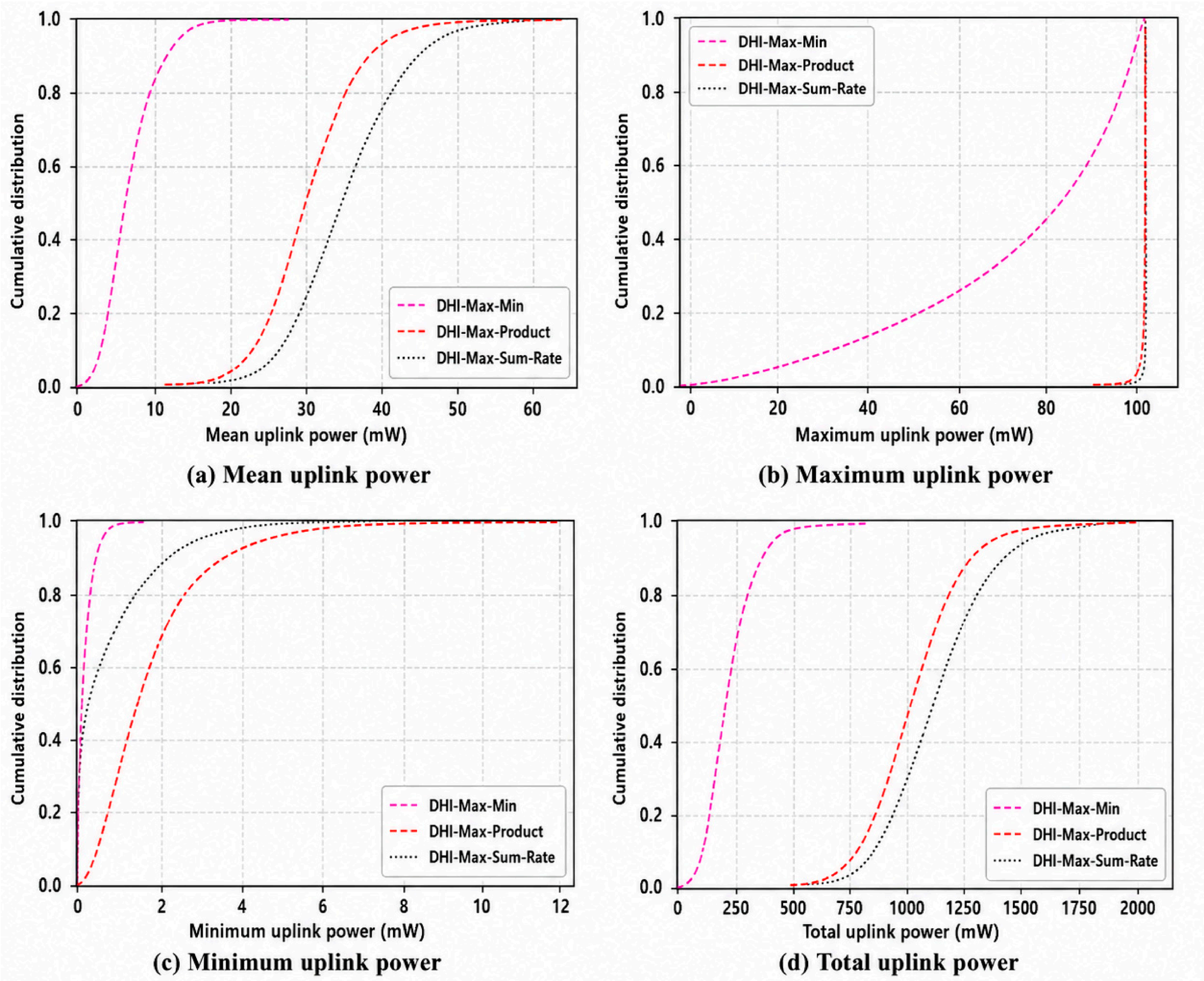

Figure 4. Cumulative distribution of uplink power in DHI model.

Conversely, DHI-Max-Product and DHI-Max-Sum-Rate distribute power across a much wider range. This indicates that these strategies allocate transmit power more selectively according to SINR improvement or throughput maximization. In Figure 4b, the maximum UE power is generally higher for the max-product and max-sum-rate strategies, which suggests that these methods allow stronger power allocation to users or channel conditions that contribute more to signal-quality or rate improvement. Figure 4c further shows that DHI-Max-Min better preserves the minimum UE power level, confirming its ability to protect weaker users.

Figure 4d compares the total uplink power distribution. The results indicate that DHI-Max-Min uses a more controlled aggregate power profile, while DHI-Max-Product and DHI-Max-Sum-Rate may require higher total transmit power to improve SINR or spectral-efficiency performance. Therefore, Figure 4 demonstrates the first major trade-off among the evaluated strategies: DHI-Max-Min provides more balanced power allocation, whereas DHI-Max-Product and DHI-Max-Sum-Rate apply more aggressive power allocation to improve signal quality and throughput.

Figure 5 Uplink spectral efficiency CDF for the DHI-Max-Min, DHI-Max-Product, and DHI-Max-Sum-Rate strategies. Each curve maps the probability that the achieved spectral efficiency remains below a given value. To provide a numerical interpretation, the 10th, 50th, and 90th percentile values are reported for the mean, maximum, minimum, and total spectral-efficiency distributions. The results show that DHI-Max-Sum-Rate provides the strongest mean and total spectral-efficiency performance, while DHI-Max-Product achieves nearly comparable spectral-efficiency behavior with better lower-tail protection in the minimum-SE distribution. DHI-Max-Min shows lower spectral-efficiency values because it prioritizes fairness-oriented operation rather than throughput maximization. Pairwise Kolmogorov–Smirnov tests were also applied to evaluate the statistical separation among the CDF distributions.

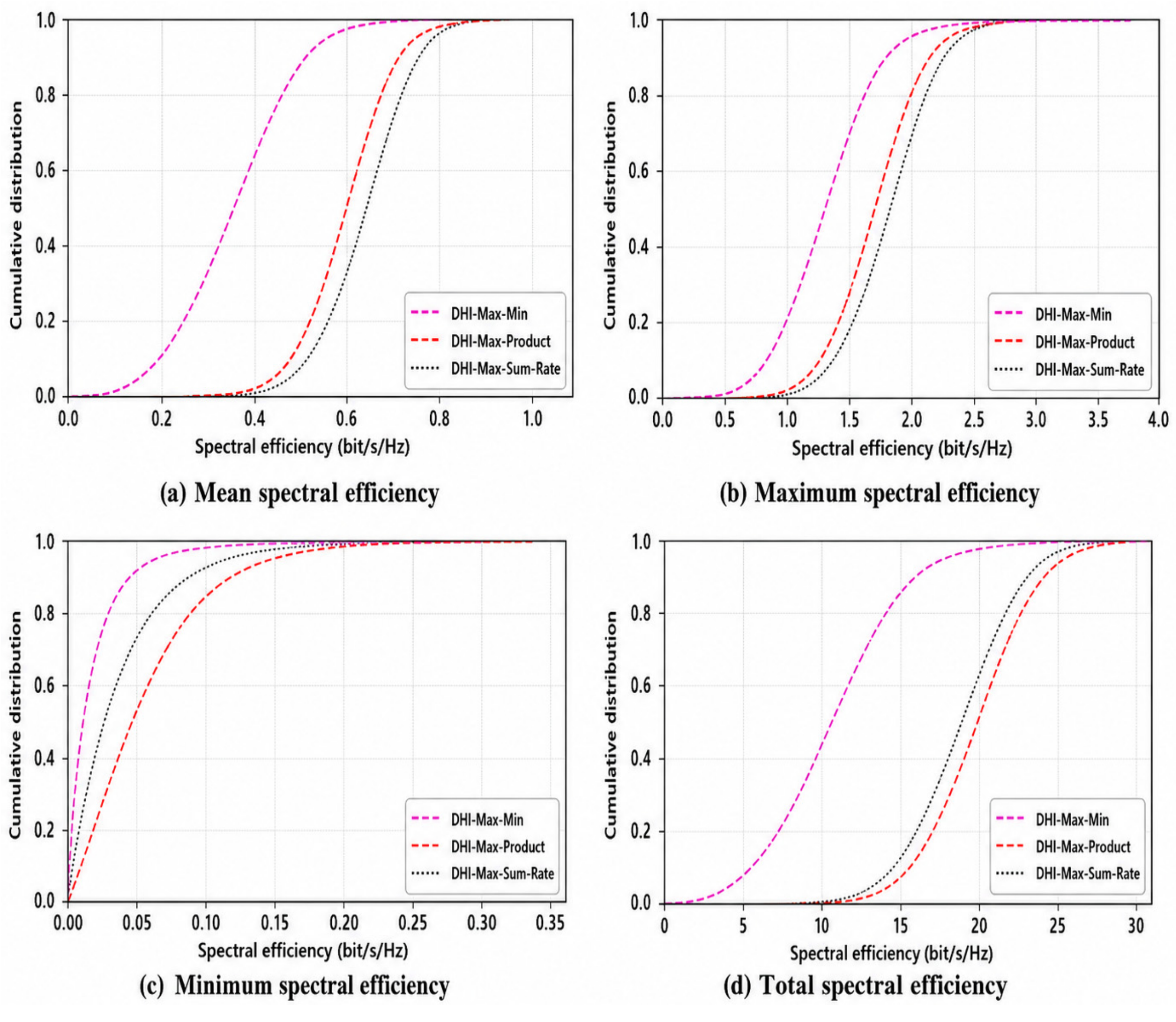

Figure 5. Cumulative distribution of spectral efficiency in the DHI model for cell-free massive MIMO systems.

Table 8 provides a numerical interpretation of the CDF. For the mean spectral-efficiency distribution, DHI-Max-Sum-Rate achieves the highest median value of 0.658 bits/s/Hz, while DHI-Max-Product achieves a very close median value of 0.655 bits/s/Hz. Both strategies clearly outperform DHI-Max-Min, which achieves a median mean spectral efficiency of 0.418 bits/s/Hz. A similar trend is observed for the total spectral-efficiency distribution, where the median values are 13.364, 20.963, and 21.043 bits/s/Hz for DHI-Max-Min, DHI-Max-Product, and DHI-Max-Sum-Rate, respectively. These results confirm that DHI-Max-Sum-Rate provides the strongest throughput-oriented behavior, while DHI-

Max-Product provides nearly comparable spectral-efficiency performance. However, in the minimum-SE distribution, DHI-Max-Product achieves the strongest 10th, 50th, and 90th percentile values, indicating better lower-tail user protection. Therefore, Figure 5 demonstrates a clear trade-off: DHI-Max-Sum-Rate is preferable for maximizing spectral efficiency; DHI-Max-Product provides the most balanced spectral efficiency behavior. In contrast, you should view DHI-Max-Min primarily as a tool for fairness rather than spectral efficiency maximization.

Table 8. Percentile-based summary of spectral-efficiency CDF results.

Metric	Strategy	10th Percentile	50th Percentile	90th Percentile
Mean SE	DHI-Max-Min	0.264	0.418	0.558
Mean SE	DHI-Max-Product	0.510	0.655	0.771
Mean SE	DHI-Max-Sum-Rate	0.517	0.658	0.771
Maximum SE	DHI-Max-Min	1.135	1.601	2.162
Maximum SE	DHI-Max-Product	1.407	1.795	2.273
Maximum SE	DHI-Max-Sum-Rate	1.440	1.813	2.283
Minimum SE	DHI-Max-Min	0.001	0.010	0.043
Minimum SE	DHI-Max-Product	0.004	0.024	0.086
Minimum SE	DHI-Max-Sum-Rate	0.000	0.000	0.053
Total SE	DHI-Max-Min	8.452	13.364	17.859
Total SE	DHI-Max-Product	16.309	20.963	24.667
Total SE	DHI-Max-Sum-Rate	16.541	21.043	24.681

Figure 6 presents the probability density distribution of spectral efficiency under the DHI-Max-Min, DHI-Max-Product, and DHI-Max-Sum-Rate strategies. Unlike the cumulative distribution in Figure 5, which shows the overall ordering of the achieved spectral-efficiency values, the density distribution in Figure 6 provides a clearer view of how frequently different spectral-efficiency levels occur among the UEs.

The DHI-Max-Min strategy clusters spectral efficiency values tightly around a moderate range. This indicates that the fairness-oriented strategy reduces large differences among users and supports a more stable service distribution. This behavior is useful in uplink CF-mMIMO scenarios where maintaining balanced user service is more important than maximizing only the strongest links.

In contrast, DHI-Max-Product and DHI-Max-Sum-Rate show broader density distributions extending toward higher spectral-efficiency values. This means that these strategies create more high-rate transmission opportunities, especially for users with stronger channel conditions or more favorable interference levels. Among them, DHI-Max-Sum-Rate exhibits the strongest throughput-oriented behavior because it directly emphasizes the maximization of total spectral efficiency. However, the wider distribution also indicates greater variation among users, which confirms the expected trade-off between throughput maximization and service uniformity.

Therefore, Figure 6 supports the conclusion obtained from Figure 5: DHI-Max-Min provides a more balanced spectral-efficiency distribution, DHI-Max-Sum-Rate improves high-rate performance, and DHI-Max-Product offers an intermediate operating behavior between fairness and throughput.

Figure 7 compares the SINR performance of the evaluated DHI-based uplink power-control strategies. The SINR results show how each strategy manages the relationship between useful received signal power and inter-user interference. This metric is important because improving uplink power allocation does not only require increasing transmit power; it also requires controlling the interference introduced to other UEs in the CF-mMIMO network.

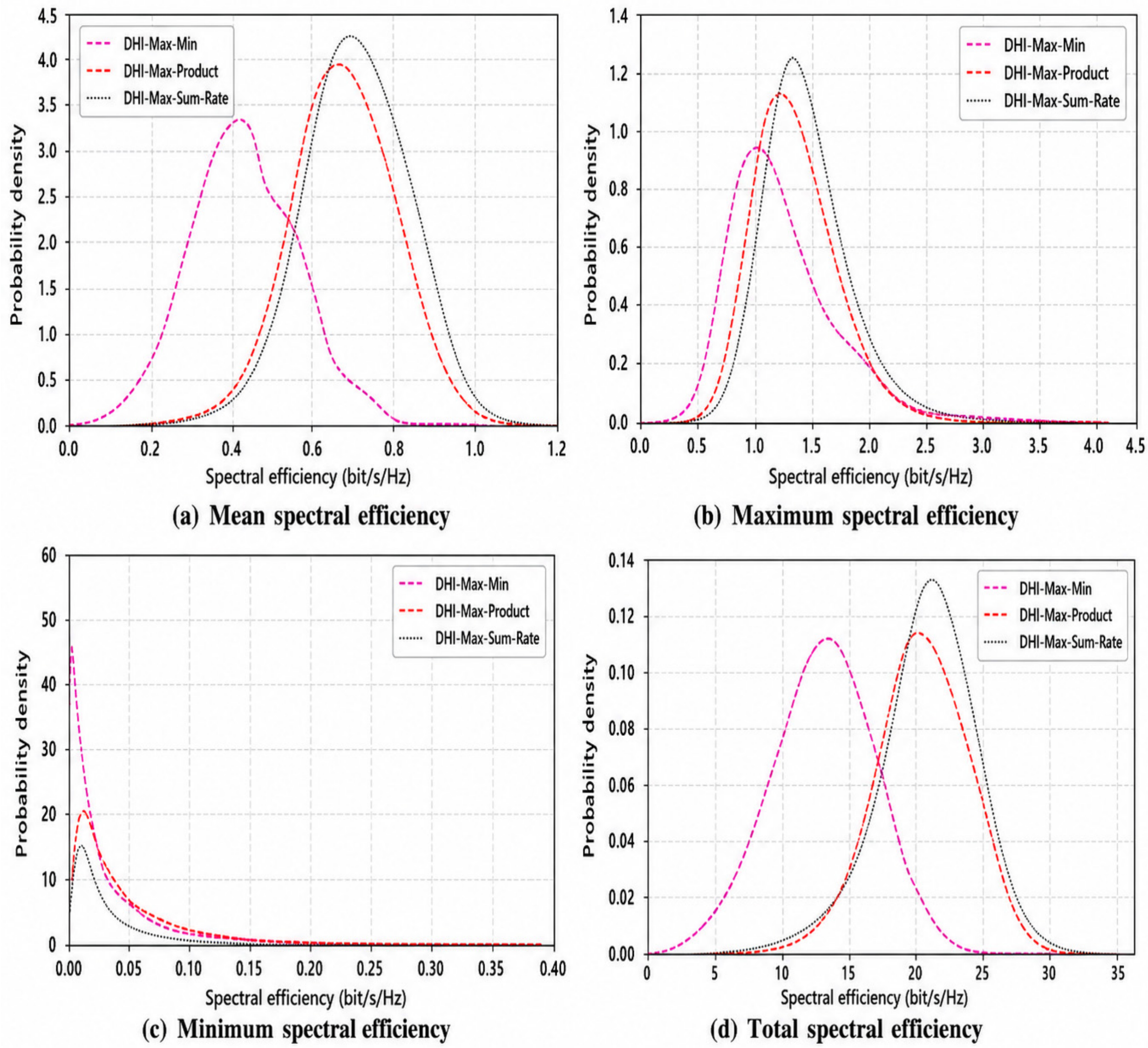

Figure 6. Probability density distribution of spectral efficiency in the DHI model for cell-free massive MIMO systems.

The DHI-Max-Min strategy provides more stable SINR behavior because it avoids allocating excessive power only to users with favorable channel conditions. This supports a more balanced signal-quality distribution among UEs, which is consistent with its fairness-oriented objective. However, because the method prioritizes weaker users, it may not always achieve the highest SINR values for users with stronger channels.

DHI-Max-Product and DHI-Max-Sum-Rate show stronger SINR-oriented behavior. The max-product strategy improves the collective SINR distribution by attempting to enhance the signal quality across users, while the max-sum-rate strategy tends to favor power allocations that increase the achievable rate of users with better channel conditions. Therefore, these two strategies can achieve higher SINR values than DHI-Max-Min, but they may also produce greater variation among users.

Overall, Figure 7 demonstrates that SINR improvement and fairness do not always increase together. DHI-Max-Min ensures stable and balanced SINR, while DHI-Max-Product and DHI-Max-Sum-Rate are better choices for maximizing signal quality and throughput. This confirms the role of the proposed framework in analyzing different uplink operating priorities under the same CF-mMIMO environment.

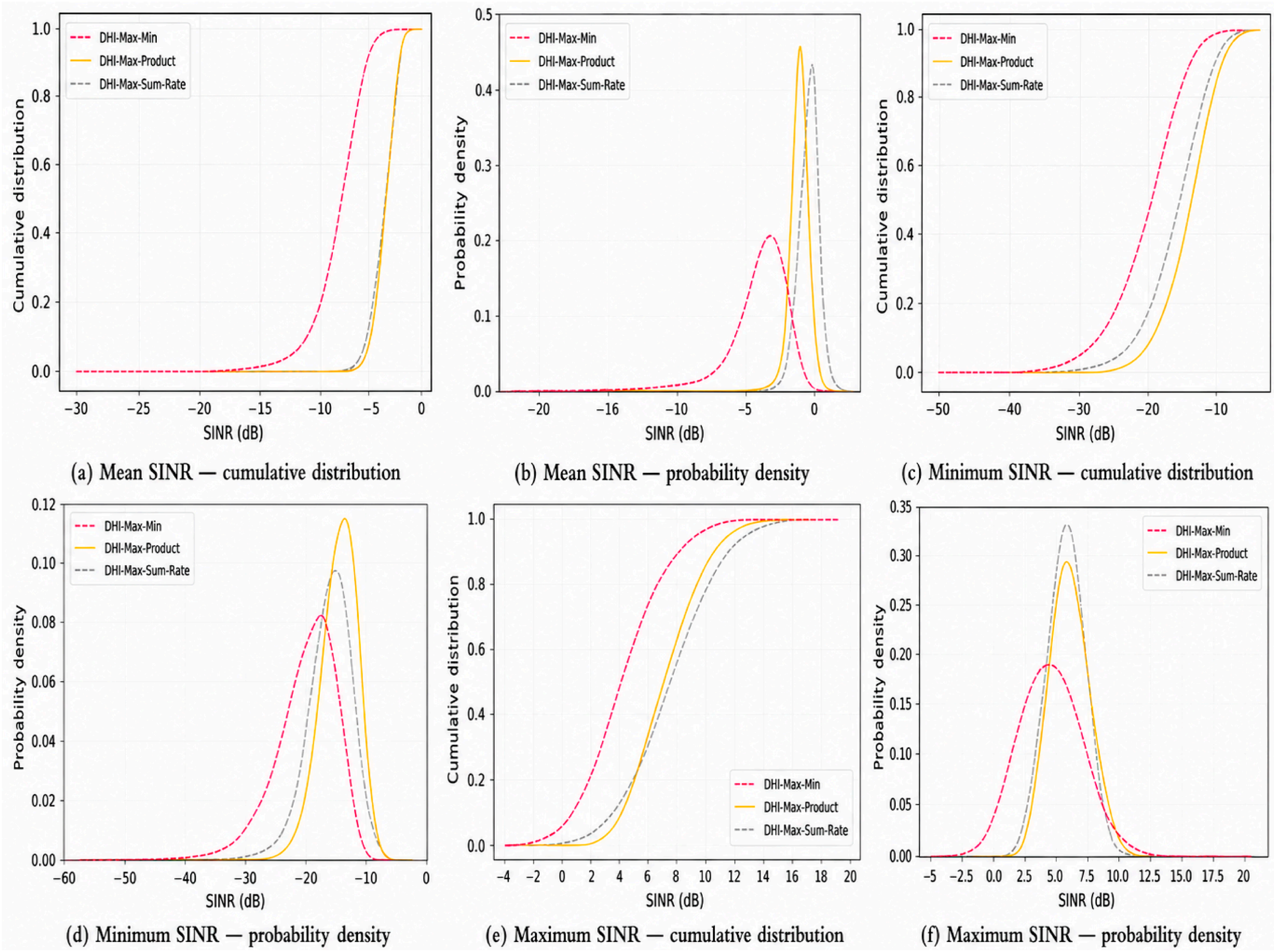

Figure 7. Cumulative distribution and probability density comparison of SINR for power-control strategies in DHI model.

To clarify the comparison methodology, SEPFM and SRM are used as fairness-oriented benchmark methods from [9]. The comparison is based on Jain's fairness index because this metric directly measures the uniformity of user service and allows the fairness behavior of the proposed DHI-based strategies to be positioned against existing fairness-aware approaches. The proposed DHI strategies were evaluated under the same AP-scale settings shown in Figure 8, namely (M = 60) and (M = 120), while the SEPFM and SRM values were used as benchmark reference values from [9]. Therefore, the purpose of this comparison is to evaluate the relative fairness behavior of the proposed strategies against representative literature methods, rather than to claim strict equivalence in all simulation parameters.

Figure 8 shows that DHI-Max-Min achieves the highest Jain's fairness index among the proposed strategies and benchmark methods. At (M = 60), DHI-Max-Min reaches 0.930, which is higher than SEPFM and SRM, with fairness values of 0.920 and 0.915, respectively. When the number of APs increases to (M = 120), DHI-Max-Min further improves to 0.989, slightly outperforming SEPFM and SRM, which achieve 0.985 and 0.975, respectively. DHI-Max-Product also provides competitive fairness performance, reaching 0.917 at (M = 60) and 0.965 at (M = 120), while DHI-Max-Sum-Rate produces lower fairness values because its main objective is throughput maximization rather than user-service balance. These results confirm the expected trade-off among the proposed strategies: While DHI-Max-Min prioritizes user fairness, DHI-Max-Product provides a balanced middle ground, and DHI-Max-Sum-Rate targets spectral efficiency improvements.

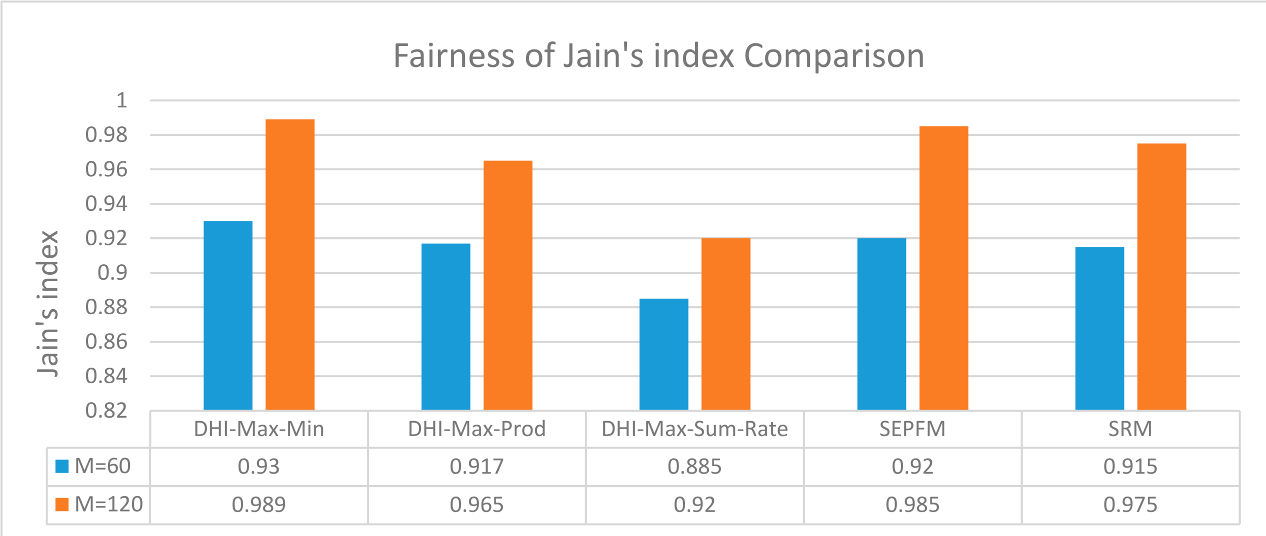

Figure 8. Fairness comparison (Jain's index) of DHI-based strategies and [9] at different AP densities.

Figure 9 presents the CDF of user rates for the proposed DHI-based strategies compared with the benchmark method in [9]. By illustrating the probability of user rates falling below a threshold, the CDF curves capture both worst-case and best-case user performance. The results show that DHI-Max-Sum-Rate shifts the distribution toward higher values, validating its throughput-oriented nature. The DHI-Max-Product strategy provides a closely comparable user-rate distribution while maintaining a more balanced performance profile. In contrast, DHI-Max-Min produces a more conservative user-rate distribution because its main objective is to protect weak users and improve service uniformity rather than maximize total throughput. Compared with [9], the proposed DHI-based strategies demonstrate competitive user-rate performance while offering objective-specific flexibility for different uplink power-control requirements.

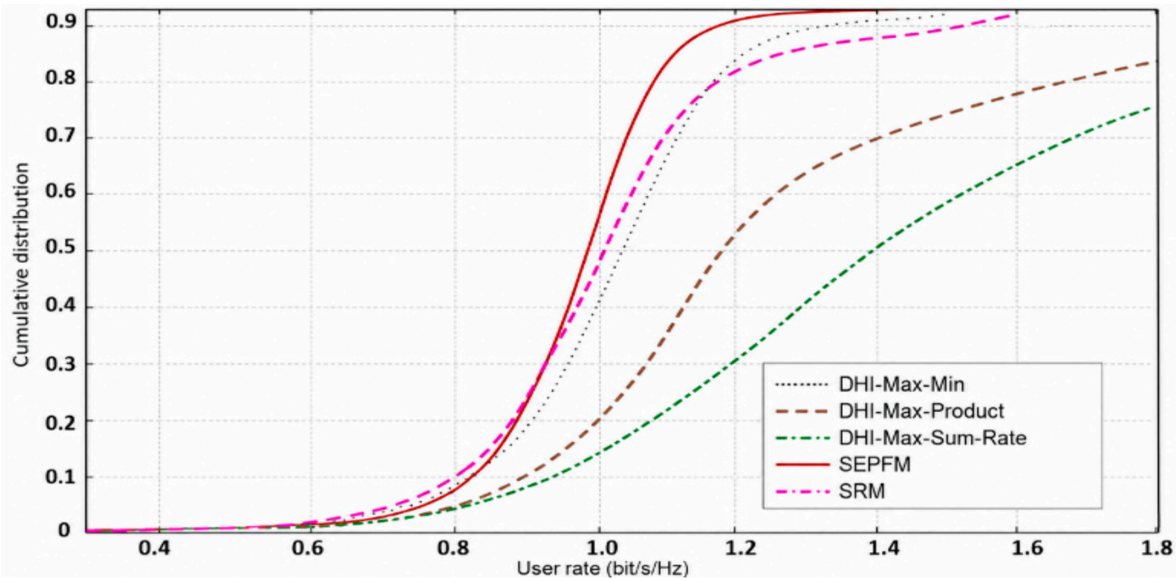

Figure 9. CDF of user rates for DHI-based strategies compared with [9].

Previous DRL-based power-control studies also show the relevance of learning-based methods for CF-mMIMO power allocation. Zhao et al. reported that DQN- and DDPG-based dynamic power allocation achieved higher sum spectral efficiency than WMMSE and required only about 0.1% of the WMMSE execution time [23]. Rahmani et al. proposed a TD3-based uplink power-control method that uses large-scale fading coefficients as input and addresses the trade-off between sum rate and fairness in uplink CF-mMIMO [25]. More recently, Zhang et al. developed a PPO-based multi-agent DRL method for uplink power control with mobile users and showed improved sum-rate performance, satisfied-user behavior, convergence speed, and lower complexity compared with selected benchmark schemes [27]. However, the numerical values reported in these studies are not directly comparable with the present results because they use different system configurations, channel models, mobility assumptions, baselines, and performance metrics. Therefore, the comparison is used to position the proposed DHI framework within the DRL-based CF-mMIMO literature rather than to claim a direct numerical superiority. Unlike fixed-objective DRL methods, the proposed DHI framework evaluates multiple objective-specific operating modes, including fairness-oriented, SINR-oriented, and throughput-oriented uplink power control under a common simulation environment.

The computational-complexity comparison is summarized in Table 9. The results clearly show that the proposed DHI-based framework requires substantially less execution time than the conventional optimization-based method for all three strategies. For the Max-Min strategy, the execution time is reduced from 1.43 s to 0.6922 s, corresponding to an absolute reduction of 0.7378 s and an improvement of approximately 51.6%. For the Max-Product strategy, the runtime decreases from 1.14 s to 0.1853 s, which corresponds to a reduction of 0.9547 s or about 83.7%. Similarly, for the Max-Sum-Rate strategy, the execution time decreases from 1.26 s to 0.1892 s, yielding a reduction of 1.0708 s or approximately 85%.

Table 9. Computational complexity comparison between the proposed DHI-based strategies and the optimization-based method [20].

Strategy	DHI-Based Method	Optimization-Based Method [20]
Max-Min	0.6922	1.43
Max-Product	0.1853	1.14
Max-Sum-Rate	0.1892	1.26

The overall results demonstrate that no single strategy dominates all performance metrics. DHI-Max-Min works best to balance user service and protect weaker UEs, as shown by its higher Jain's fairness index and uniform power and spectral efficiency distributions. In contrast, DHI-Max-Sum-Rate maximizes total throughput and boosts user rates, though it increases performance disparities between users. DHI-Max-Product provides an intermediate operating point by improving the collective SINR behavior while maintaining a more balanced distribution than pure sum-rate maximization. The results show that dynamic uplink power control in CF-mMIMO networks involves a clear trade-off between fairness, SINR, spectral efficiency, and processing complexity. Therefore, the selection of the most suitable strategy should depend on the target network objective, whether the priority is balanced user service, signal-quality improvement, throughput maximization, or fast online decision-making.

To further support the statistical reliability of the spectral-efficiency results, Table 10 reports a snapshot-based statistical summary computed. An analysis of the mean, standard deviation, and 95% confidence interval for each strategy shows that DHI-Max-Sum-Rate leads. It achieves the highest average mean SE and total SE at 0.649 and 20.774 bits/s/Hz, respectively. DHI-Max-Product provides very close performance, with mean SE and total SE values of 0.647 and 20.699 bits/s/Hz, confirming its balanced spectral-efficiency behavior. In contrast, DHI-Max-Min produces lower mean and total SE values because its objective is mainly fairness-oriented rather than throughput-oriented. The narrow 95% confidence intervals indicate that the estimated average values are stable over the evaluated simulation snapshots, while the standard deviation reflects the variability caused by changing channel and mobility conditions.

Table 10. Statistical summary of spectral-efficiency results over available simulation snapshots.

Strategy	Mean SE (Mean $\pm$ Std)	95% CI for Mean SE	Total SE (Mean $\pm$ Std)	95% CI for Total SE
DHI-Max-Min	0.414 ± 0.115	[0.412, 0.417]	13.258 ± 3.674	[13.186, 13.330]
DHI-Max-Product	0.647 ± 0.104	[0.645, 0.649]	20.699 ± 3.318	[20.633, 20.764]
DHI-Max-Sum-Rate	0.649 ± 0.101	[0.647, 0.651]	20.774 ± 3.237	[20.711, 20.838]

10. Conclusions

This paper presents a Deep Hybrid Intelligent framework for dynamic uplink power control in cell-free massive MIMO networks. Using Soft Actor-Critic learning, it handles continuous power allocation under mobility and time-varying channel conditions. We compared max-min, max-product, and max-sum-rate strategies to assess the trade-offs between user fairness, SINR, spectral efficiency, and computational load.

The simulation results confirmed that the evaluated strategies provide different but complementary performance behaviors. DHI-Max-Min achieved the strongest fairness performance, reaching a Jain's fairness index of 0.989 at 120 APs. This result confirms its suitability for protecting weaker UEs and improving service balance in dense CF-mMIMO deployments. DHI-Max-Sum-Rate provided stronger SINR, spectral-efficiency, and user-rate performance, making it more suitable for throughput-oriented operation, although it produced greater variation among users. DHI-Max-Product offered an intermediate operating point by improving collective SINR behavior while maintaining a more balanced performance profile than pure sum-rate maximization.

The execution-time results further demonstrated the computational advantage of the evaluated framework. Compared with conventional optimization-based implementations, the reported execution-time reductions were 51.6%, 83.7%, and 85.0% for the max-min, max-product, and max-sum-rate strategies, respectively. These findings show that shifting part of the computational burden toward learning-based decision-making can support faster online power-control operation in dynamic uplink CF-mMIMO environments.

Overall, the results confirm that uplink power control in CF-mMIMO systems involves a clear trade-off among fairness, SINR, spectral efficiency, and computational complexity. Therefore, the most suitable strategy depends on the target network objective. Fairness-oriented operation is preferable when service uniformity and weak-user protection are required, while sum-rate-oriented operation is more suitable when maximizing total throughput is the main priority. The max-product strategy can be considered when a compromise between signal quality and balanced user service is needed.

Despite these findings, this study has several limitations. First, the evaluation was conducted through simulation, and practical implementation factors such as fronthaul capacity, synchronization errors, signaling overhead, hardware impairments, and energy efficiency were not fully considered. Another practical limitation is related to hardware-imposed SINR ceilings. The present simulation computes SINR based on the adopted channel, interference, and noise model without explicitly imposing an error-vector-magnitude (EVM) constraint. In practical radio-frequency transceivers, transmitter and receiver impairments introduce residual distortion, which can limit the maximum achievable SINR even under favorable channel conditions. For example, an EVM level of 1% corresponds approximately to a maximum effective SINR of $\frac{1}{\mathrm{EVM}^2} = 10^4$, or about 40 dB. Therefore, extremely high simulated SINR values should be interpreted as idealized upper-bound behavior rather than guaranteed practical receiver performance. Future work will incorporate EVM-aware

SINR capping and hardware-impairment models to evaluate how robust the proposed DHI is under more realistic implementation constraints. Second, the study focused on uplink power control only; therefore, joint uplink–downlink optimization remains an important extension. Third, the benchmark comparison was limited to the selected optimization-based strategies, and broader comparison with additional classical and deep reinforcement learning baselines would further strengthen the evaluation. Finally, the current framework does not explicitly optimize latency, energy efficiency, or robustness under irregular traffic conditions.

Additional limitations are related to energy efficiency, convergence behavior, and sample efficiency. The present study evaluates uplink power control mainly through SINR, spectral efficiency, fairness, user rate, and execution time, but it does not explicitly optimize energy efficiency in bits/Joule. Moreover, the SAC training process was not analyzed using detailed convergence indicators such as reward curves, actor loss, critic loss, or entropy evolution. The sample efficiency of SAC was also not fully examined, although this aspect is important because DRL methods may require many environment interactions during offline training. Future work will therefore consider energy-aware reward design, convergence analysis, and sample-efficiency evaluation under different network sizes, mobility patterns, and channel conditions.

Future work will extend the proposed framework toward joint uplink and downlink power control, energy-efficient power allocation, fronthaul-aware optimization, and more realistic deployment scenarios. Further investigation may also include multi-agent learning, transfer learning, and online adaptation mechanisms to improve robustness under different user densities, mobility patterns, and traffic conditions.

Author Contributions: Conceptualization, H.A.J. and M.F.A.R.; methodology, H.A.J.; software, H.A.J.; validation, H.A.J., M.F.A.R., F.H. and S.M.; formal analysis, H.A.J.; investigation, H.A.J.; resources, M.F.A.R., F.H. and S.M.; data curation, H.A.J.; writing—original draft preparation, H.A.J.; writing—review and editing, M.F.A.R., F.H. and S.M.; visualization, H.A.J.; supervision, M.F.A.R., F.H. and S.M.; project administration, M.F.A.R.; funding acquisition, not applicable. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data supporting the findings of this study were generated through simulation experiments conducted by the authors. The simulation parameters and methodological details are provided within the article. Additional data may be made available from the corresponding author upon reasonable request.

Acknowledgments: The authors acknowledge the Department of Computer and Communication Systems Engineering, Faculty of Engineering, Universiti Putra Malaysia, for its academic support and provision of research facilities during the preparation of this work.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Patel, S.H. Dynamic Spectrum Sharing And Management Using Drone-Based Platforms For Next-generation Wireless Networks. *Preprints* **2024**, 2024120709. [CrossRef]
2. Mohammadi, M.; Mobini, Z.; Ngo, H.Q.; Matthaiou, M. Next-Generation Multiple Access With Cell-Free Massive Mimo. *Proc. IEEE* **2024**, *112*, 1372–1420. [CrossRef]
3. Demir, Ü.T.; Bjørnson, E.; Sanguinetti, L. Foundations of User-Centric Cell-Free Massive Mimo. *Found. Trends® Signal Process.* **2021**, *14*, 162–472. [CrossRef]
4. Pesantez Diaz, F.; Estevez, C. Performance Evaluation Of Cf-Mmimo Wireless Systems Using Dynamic Mode De-composition. *Telecom* **2024**, *5*, 846–891. [CrossRef]

5. Umar, M.M.; Garko, A.B.; Suru, H.U.; Sarki, M.N. A Systematic Review Of Resource Allocation For 5g Networks: Challenges, Methods And Future Research Directions. *Niger. J. Technol.* **2026**, *45*, 155–177. [CrossRef]
6. Silva, M.; Santos, J.; Curado, M. Dynamic Resource Allocation For Cell-Free Massive Mimo Networks With Critical Services Support. *IEEE Internet Things J.* **2025**, *12*, 32072–32085. [CrossRef]
7. Fu, J.; Mobini, Z.; Ngo, H.Q.; Zhu, P.; Matthaiou, W. mmse-Based Processing In Cell-Free Massive Mimo Systems. *IEEE Wirel. Commun. Lett.* **2024**, *14*, 330–334. [CrossRef]
8. Lai, P.; He, Q.; Cui, G.; Chen, F.; Grundy, J.; Abdelrazek, M.; Hosking, J.; Yang, Y. Cost-Effective User Allocation In 5G Noma-Based Mobile Edge Computing Systems. *IEEE Trans. Mob. Comput.* **2021**, *21*, 4263–4278. [CrossRef]
9. Quoc Pham, V.; Hoang Kha, H.; Ty Khanh, L. On Proportional Fairness Of Uplink Spectral Efficiency In Cell-Free Massive Mimo Systems. *Int. J. Commun. Syst.* **2020**, *33*, e4534. [CrossRef]
10. Hao, C.; Vu, T.T.; Ngo, H.Q.; Dao, M.N.; Dang, X.; Wang, C.; Matthaiou, M. Joint User Association And Power Control For Cell-Free Massive Mimo. *IEEE Internet Things J.* **2024**, *11*, 15823–15841. [CrossRef]
11. Zewail, I.; Shokair, M.; Ghallab, R.; Zayed, M.M. Interference-aware power allocation in GFDM- and OFDM-based cognitive radio networks using hybrid PSO-SA optimization. *Wirel. Pers. Commun.* **2026**, *advance online publication*. [CrossRef]
12. Rajapaksha, N.; Manosha, K.B.S.; Rajatheva, N.; Latva-Aho, M. Deep Learning-Based Power Control For Cell-Free Massive Mimo Networks. In *Proceedings of the ICC 2021-IEEE International Conference on Communications*; IEEE: Montreal, QC, Canada, 2021; pp. 1–7.
13. Zhao, Y.; Niemegeers, I.G.; De Groot, S.H. Deep Q-Network Based Dynamic Power Allocation For Cell-Free Massive Mimo. In *Proceedings of the 2021 IEEE 26th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)*; IEEE: Porto, Portugal, 2021; pp. 1–7.
14. Rajamallaiah, A.; Naresh, S.V.K.; Raghuvamsi, Y.; Manmadharao, S.; Bingi, K.; Guerrero, J.M. Deep Reinforcement Learning For Power Converter Control: A Comprehensive Review Of Applications And Challenges. *IEEE Open J. Power Electron.* **2025**, *6*, 1769–1802. [CrossRef]
15. Zhao, Y.; Niemegeers, I.G.; De Groot, S.M.H. Dynamic Power Allocation For Cell-Free Massive Mimo: Deep Rein-forcement Learning Methods. *IEEE Access* **2021**, *9*, 102953–102965. [CrossRef]
16. Farooq, M.; Ngo, H.Q. Mirror Prox Algorithm For Large-Scale Cell-Free Massive Mimo Uplink Power Control. *IEEE Commun. Lett.* **2022**, *26*, 2994–2998. [CrossRef]
17. Mohammadzadeh, S.; Ghourtani, M.R.; Cumanan, K.; Burr, A.; Xiao, P. Pilot And Data Power Control For Scalable Uplink Cell-Free Massive Mimo. *IEEE Open J. Commun. Soc.* **2025**, *6*, 10829–10844. [CrossRef]
18. Rahmani, M.; Bashar, M.; Dehghani, M.J.; Xiao, P.; Tafazolli, R.; Debbah, M. Deep Reinforcement Learning-Based Power Allocation In Uplink Cell-Free Massive Mimo. In *Proceedings of the 2022 IEEE Wireless Communications and Networking Conference (WCNC)*; IEEE: Austin, TX, USA, 2022; pp. 459–464.
19. Diarra, D.; Absaloms, H.O.; Langat, P.K. Optimizing Uplink Power Control For Energy Efficiency In Mmwave User-Centric Cell-Free Massive Mimo With Deep Reinforcement Learning. *J. Eng. Appl. Sci.* **2026**, *73*, 11. [CrossRef]
20. Zhang, Y.; Zhang, J.; Jin, Y.; Buzzi, S.; Ai, B. Deep Learning-Based Power Control For Uplink Cell-Free Massive Mimo Systems. In *Proceedings of the GLOBECOM 2021–2021 IEEE Global Communications Conference*; IEEE: Madrid, Spain, 2021; pp. 1–6.
21. Alvi, N.M.; Alvi, W.M.; Zhou, X.; Li, J.; Wei, Y. Constrained Soft Actor–Critic For Joint Computation Offloading And Resource Allocation In UAV-Assisted Edge Computing. *Sensors* **2026**, *26*, 1149. [CrossRef] [PubMed]
22. Mendoza, C.F.; Kaneko, M.; Rupp, M.; Schwarz, S. Accelerated Deep Reinforcement Learning For Uplink Power Control In A Dynamic Cell-Free Massive Mimo network. *IEEE Wirel. Commun. Lett.* **2024**, *13*, 1710–1714. [CrossRef]
23. D'ANdrea, C.; Zappone, A.; Buzzi, S.; Debbah, M. Uplink Power Control In Cell-Free Massive Mimo Via Deep Learning. In *Proceedings of the 2019 IEEE 8th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*; IEEE: Le Gosier, Guadeloupe, 2019; pp. 554–558.
24. Rahmani, M.; Bashar, M.; Dehghani, M.J.; Akbari, A.; Xiao, P.; Tafazolli, R.; Debbah, M. Deep Reinforcement Learning-Based Sum Rate Fairness Trade-Off For Cell-Free Mmimo. *IEEE Trans. Veh. Technol.* **2022**, *72*, 6039–6055. [CrossRef]
25. Zhang, X.; Kaneko, M.; Lui, J.C.S.; Ji, Y. Multi-Agent Deep Reinforcement Learning-Based Uplink Power Control In Cell-Free Massive Mimo With Mobile Users. *IEEE Trans. Veh. Technol.* **2025**, *74*, 17796–17811. [CrossRef]
26. Tankaria, H.; Sugimoto, S. A Regularized Limited Memory Bfgs Method For Large-Scale Unconstrained Optimization And Its Efficient Implementations. *Comput. Optim. Appl.* **2021**, *82*, 61–88. [CrossRef]
27. Rahmani, M.; Bashar, M.; Dehghani, M.J.; Xiao, P.; Tafazolli, R.; Debbah, M. Learning-Based Downlink Power Allocation In Cell-Free Massive Mimo Systems. *IEEE Trans. Wirel. Commun.* **2022**, *22*, 174–188. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.