Antenna Selection of Cell-Free XL-MIMO Systems with Multi-Agent Reinforcement Learning

Zhilong Liu
School of Electronic and Information Engineering
Beijing Jiaotong University
Beijing, China
zhilongliu@bjtu.edu.cn

Ziheng Liu
School of Electronic and Information Engineering
Beijing Jiaotong University
Beijing, China
ziheng liu01@163.com

Xinjie Wang
65529 Institution
Liaoyang, China
wxj7519@163.com

Jiaxun Li
Intelligent Game and Decision Lab
Beijing, China
lijiaxun@nudt.edu.cn

Jiayi Zhang
School of Electronic and Information Engineering
Beijing Jiaotong University
Beijing, China
zhangjiayi@bjtu.edu.cn

Bo Ai
State Key Laboratory of Rail Traffic Control and Safety
Beijing Jiaotong University
Beijing, China
boai@bjtu.edu.cn

2023 International Conference on Ubiquitous Communication (Ucom) | 979-8-3503-4043-3/23/$31.00 ©2023 IEEE | DOI: 10.1109/UCOM59132.2023.10257633

***Abstract*—Cell-free massive multiple-input multiple-output (mMIMO) and extremely large-scale MIMO (XL-MIMO) have been investigated as promising innovations for next-generation wireless communication systems. In this article, we achieve the combination of the aforementioned technologies and explore such an extended paradigm namely cell-free XL-MIMO. In such systems, high-dimensional matrix operations due to massive antennas consume considerable computational resources. Moreover, with the significant benefits of increasing the number of degrees of freedom, it is necessary to simultaneously reduce power consumption. Therefore, we aim to achieve the antenna selection strategy in cell-free XL-MIMO systems with multi-agent reinforcement learning (MARL). And we investigate the key performance indicators of cell-free XL-MIMO systems, such as spectral efficiency (SE) and energy efficiency (EE). The numerical results and analyses show a significant EE increase and uniform SE coverage in cell-free XL-MIMO systems with MARL methods.**

***Index Terms*—Antenna selection, energy efficiency, extremely large-scale MIMO, multi-agent reinforcement learning, spectral efficiency.**

I. Introduction

In the sixth-generation (6G) wireless communication systems, various advanced technologies are expected to evolve to serve more application scenarios and use cases with unprecedented levels of performance, thus realizing the concept of the Internet of Everything. As the key performance indicators uplift by orders of magnitude in 6G, the massive multiple-input multiple-output (mMIMO) cannot be bound for the above challenges. Therefore, to overcome the capacity bottlenecks of conventional MIMO, emerging technologies such as cell-free mMIMO [1], [2] and extremely large-scale MIMO (XL-MIMO) [3], [4] are proposed, which are promising to shine from blueprint to reality by supporting enormous connections.

In cell-free mMIMO systems, the distribution of massive antennas shortens the communication distance between transceivers. All access points (APs) are connected to a central processing unit (CPU), which enhances the system performance and provides rich macro-diversity gain [1]. Evidently, the architecture of cell-free mMIMO systems effectively promotes the overall quality of service and mitigates inter-cell interference by deploying a large number of geographically distributed APs to serve user equipments (UEs). By contrast, the promising XL-MIMO technology maintains the continuous growth of antennas to promote the channel capacity [5]. Meanwhile, new channel characteristics yield since the transformation of the communication domain from the far-field to the near-field, such as the spherical wave transmission and the non-stationary effect. In previous communication system designs, the electromagnetic wave focused on the planar wave under the angular domain condition. However, the simple Rician [2] and Rayleigh [6] fading channel models are no longer applicable to the near-field channels based on the spherical wavefront assumption. Recently, the location division multiple access has been investigated based on such near-field spherical-wave models. It was shown that UEs located in different locations can be distinguished since the energy can be concentrated on a specific location like the "spotlight effect". Accordingly, the precise requirement of acquiring channel state information (CSI) should be solved first [7].

In fact, the integration of massive antennas on panels of a certain size poses great challenges in the applications of antenna theory, information theory, and electromagnetic propagation theory, thus stimulating the birth of electromagnetic information theory for the near-field communication. In XL-MIMO systems, the power consumption problem should be reexamined in order to reduce construction costs and achieve practical deployment. In light of that not always all antennas serving different UEs operate together, hence, an appropriate antenna selection strategy is of great importance to reduce the power consumption and the computational complexity. In [8], the authors proposed three antenna selection algorithms to improve the energy efficiency (EE) performance in XL-MIMO systems. Apart from [8], it has been investigated using the

This work was supported by the Fundamental Research Funds for the Central Universities under Grant No. 2023YJS001 and ZTE Industry-University-Institute Cooperation Funds under Grant No. HC-CN-20221202003.

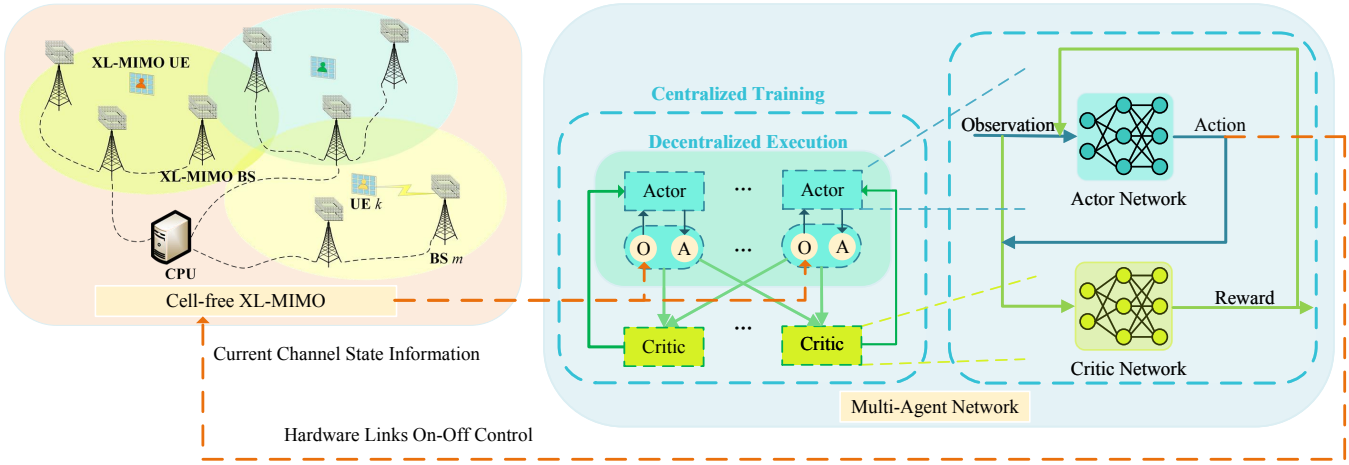

Fig. 1. Illustration of a cell-free XL-MIMO system and the MADDPG algorithm.

gradient boosting decision tree algorithm in millimeter-wave MIMO systems [9].

Actually, the low-complexity algorithm design is worthy of investigation in XL-MIMO systems [10], including the channel model, channel estimation, and heuristic signal processing methods. The low-complexity baseband signal processing algorithms are required to reduce systems' overall computational complexity and energy consumption. Multi-Agent Reinforcement Learning (MARL) has been widely used for decision-making in unmanned aerial vehicles, swarm intelligence, and traffic scheduling. In MIMO systems, this essential tool is preferable for the dynamic environment to allocate communication resources [11]. Motivated by the aforementioned works, we investigate the performance of antenna selection in cell-free XL-MIMO systems using the multi-agent deep deterministic policy gradient (MADDPG) algorithm. The main contributions are summarized as follows:

- Firstly, we investigate a cell-free XL-MIMO system channel model based on the scalar Green's function. More important, we comprehensively analyze and derive the spectral efficiency (SE) and EE expressions to reveal the performance of such a system.
- Secondly, we propose a MADDPG-based antenna selection strategy to reduce system-level power consumption through the Centralized Training and Decentralized Execution (CTDE) architecture.
- Finally, the numerical results show that the MADDPG algorithm can provide uniform service to all UEs and enhance the average EE in cell-free XL-MIMO systems.

II. System Model

In this paper, we investigate a cell-free XL-MIMO system consisting of M BSs and K UEs, where all BSs are connected to a CPU via fronthaul links, as illustrated in Fig. 1 [6]. Each BS and UE are equipped with a uniform planar array (UPA)-based XL-MIMO with $N_r = N_{V_r} N_{H_r}$ and $N_s = N_{V_s} N_{H_s}$ antennas, respectively. The antenna spacing Δ_r, Δ_s at each BS and each UE satisfy less than half of the carrier wavelength, that is $\Delta_r \leqslant \lambda/2$ and $\Delta_s \leqslant \lambda/2$ [12]. For simplicity, the antennas at each BS are indexed row-by-row by $n \in [1, N_r]$, thus the location of BS m with respect to the origin can be denoted as a three-dimension coordinate $\mathbf{r}_m = [r_{m,x}, r_{m,y}, r_{m,z}]^T$. Accordingly, the received signals at all M BSs can be denoted as $\mathbf{a}_r(\mathbf{k}, \mathbf{r}) = [\mathbf{a}_{r,1}(\mathbf{k}, \mathbf{r}), \ldots, \mathbf{a}_{r,M}(\mathbf{k}, \mathbf{r})]$ with the m-th element

$$\mathbf{a}_{r,m}(\mathbf{k}, \mathbf{r}) = [e^{j\mathbf{k}_{r,m}(\varphi,\theta)^T \mathbf{r}_m^{(1)}}, \ldots, e^{j\mathbf{k}_{r,m}(\varphi,\theta)^T \mathbf{r}_m^{(N_r)}}]^T, \quad (1)$$

where $\mathbf{k}_{r,m}(\varphi, \theta) = k[\cos(\theta)\cos(\varphi), \cos(\theta)\sin(\varphi), \sin(\theta)] \in \mathbb{R}^3$ is the receive wave vector with the receive elevation angle θ and the receive azimuth angle φ at BS m. And k is the wavenumber equaling to $2\pi/\lambda$ [13].

Likewise, the antennas at each UE are indexed row-by-row by $n \in [1, N_s]$ and the location of UE k is $\mathbf{s}_k = [s_{k,x}, s_{k,y}, s_{k,z}]^T$. The transmit signals from all K UEs can be denoted as $\mathbf{a}_s(\boldsymbol{\kappa}, \mathbf{s}) = [\mathbf{a}_{s,1}(\boldsymbol{\kappa}, \mathbf{s}), \ldots, \mathbf{a}_{s,K}(\boldsymbol{\kappa}, \mathbf{s})]$ with

$$\mathbf{a}_{s,k}(\boldsymbol{\kappa}, \mathbf{s}) = [e^{j\boldsymbol{\kappa}_{s,k}(\varphi,\theta)^T \mathbf{s}_k^{(1)}}, \ldots, e^{j\boldsymbol{\kappa}_{s,k}(\varphi,\theta)^T \mathbf{s}_k^{(N_s)}}]^T, \quad (2)$$

where $\boldsymbol{\kappa}_{s,k}(\varphi, \theta) = k[\cos(\theta)\cos(\varphi), \cos(\theta)\sin(\varphi), \sin(\theta)] \in \mathbb{R}^3$ is the transmit wave vector at UE k.

A. Channel Model

Based on the investigation above, the corresponding channel coefficient can be denoted as $\mathbf{H}_{mk} = \mathbf{L}_{mk} \odot \mathbf{S}_{mk}$, where $\mathbf{L}_{mk}$ and $\mathbf{S}_{mk}$ represent the large-scale fading (LSF) coefficient and the small-scale fading coefficient, respectively. Following the single-BS single-UE channel model and single-BS multi-UE channel model proposed in [12], [14], the small-scale fading coefficient $\mathbf{S}_{mk} = \sqrt{N_r N_s} \sum_{(\ell_x,\ell_y)\in\varepsilon_r} \sum_{(m_x,m_y)\in\varepsilon_s} s_{mk} \in \mathbb{C}^{N_r \times N_s}$, where s_{mk} can be defined as

$$\mathbf{s}_{mk} = S_a^{(mk)}(\ell_x, \ell_y, m_x, m_y)\mathbf{a}_r(\ell_x, \ell_y, \mathbf{r}^{(m)})\mathbf{a}_s(m_x, m_y, \mathbf{s}^{(k)}). \quad (3)$$

$S_a^{(mk)}(\ell_x, \ell_y, m_x, m_y)$ is the Fourier coefficient with variance $\sigma_{mk}^2(\ell_x, \ell_y, m_x, m_y)$, satisfying

$$S_a^{(mk)}(\ell_x, \ell_y, m_x, m_y) \sim \mathcal{N}_{\mathbb{C}}(0, \sigma_{mk}^2(\ell_x, \ell_y, m_x, m_y)). \quad (4)$$

Furthermore, we denote $d_{mk}^{n_r n_s}$ and $\vartheta_{mk}^{n_r n_s}$ as the distance and the angle between the n_r-th antenna of BS m and the n_s-th antenna of UE k, respectively. Then, the LSF coefficient can be denoted as $\mathbf{L}_{mk} = \lambda\sqrt{G_t F(\vartheta_{mk}^{n_r n_s})}/(4\pi d_{mk}^{n_r n_s}) \in \mathbb{C}^{N_r \times N_s}$ with $F(\vartheta)$ and G_t denoting the normalized power radiation pattern and the antenna gain, respectively.

B. Uplink Transmission

In the uplink, all antennas of UEs simultaneously transmit their data signals to BSs. The received signals $\mathbf{y}_m$ at BS m is

$$\mathbf{y}_m = \sum_{k=1}^{K} \sqrt{p_k}\mathbf{H}_{mk}\mathbf{s}_k + \mathbf{n}_m = \sum_{k=1}^{K} \mathbf{H}_{mk}\mathbf{x}_k + \mathbf{n}_m, \quad (5)$$

where $\mathbf{s}_k = [s_{k,1}, \ldots, s_{k,N_s}]$, $\mathbf{x}_k = [x_{k,1}, \ldots, x_{k,N_s}]^T$, and p_k represent the signal, the transmitted symbol, and the transmit power of each antenna, respectively, satisfying $\mathbf{x}_k = \sqrt{p_k}\mathbf{s}_k$ and $\mathrm{tr}(\mathbf{x}_k\mathbf{x}_k^H) = N_s p_k$. Let $\mathbf{V}_{mk}$ denote the combining matrix designed by BS m for UE k. Then, the local estimation of the transmitted symbol $\mathbf{x}_k$ for UE k at BS m is [15]–[17]

$$\check{\mathbf{x}}_{mk} = \mathbf{V}_{mk}^H\mathbf{H}_{mk}\mathbf{x}_k + \sum_{l=1,l\neq k}^{K} \mathbf{V}_{mk}^H\mathbf{H}_{ml}\mathbf{x}_l + \mathbf{V}_{mk}^H\mathbf{n}_m. \quad (6)$$

III. Spectral Efficiency And Energy Efficiency Analysis

A. Spectral Efficiency Analysis

We notice that the large-scale fading decoding (LSFD) method requires abundant LSF parameters knowledge [18]. However, the acquisition of LSF coefficients grows quadratically with N_r, N_s, M, and K, which is tremendous in cell-free XL-MIMO systems [19]. To overcome this challenge, the CPU can alternatively weight the local processed signal $\check{\mathbf{x}}_{mk}$ by simply taking the average of them across the observations from the M BSs to obtain the final signals as $\hat{\mathbf{x}}_k = \frac{1}{M}\sum_{m=1}^{M}\check{\mathbf{x}}_{mk}$. Based on the above, we derive the uplink achievable SE as the following corollary [20].

Corollary 1: An achievable SE for UE k in cell-free XL-MIMO systems is given by

$$\mathrm{SE}_k = \log_2\left|\mathbf{I}_{N_s} + \mathbf{U}_k^H\mathbf{\Psi}_k^{-1}\mathbf{U}_k\right|, \quad (7)$$

where $\mathbf{U}_k \triangleq \sqrt{p_k}\sum_{m=1}^{M}\mathbb{E}\{\mathbf{V}_{mk}^H\mathbf{H}_{mk}\}$ *and* $\mathbf{\Psi}_k \triangleq \sum_{l=1}^{K}\sum_{m=1}^{M}\sum_{m'=1}^{M} p_l\mathbb{E}\{\mathbf{V}_{mk}^H\mathbf{H}_{ml}\mathbf{H}_{m'l}^H\mathbf{V}_{m'k}\} - \mathbf{U}_k\mathbf{U}_k^H + \sum_{m=1}^{M}\mathbb{E}\{\mathbf{V}_{mk}^H\mathbf{n}_m\mathbf{n}_m^H\mathbf{V}_{mk}\}$.

Note that the equation (7) is suitable for any combining design. With maximum ratio combining, We can derive the closed-form SE expression with $\mathbf{V}_{mk} = \mathbf{H}_{mk}$, which does not require any matrix inversion and has low computational complexity [2], [4].

B. Power Consumption Model and Energy Efficiency

Without loss of generality, we define the circuit power consumption P_c as the sum of all analog devices and digital signal processing in practical cell-free XL-MIMO systems. To be more concrete, the main circuit power consumption consists of five parts:

- The fixed power P_{fix} is a constant that accounts for site cooling, control signaling, load-independent backhaul infrastructure, and baseband processor power.
- The power consumption of the transceiver chains is modeled as $P_{tc} = P_{syn} + |S|P_{bs} + KP_{ue}$, in which P_{syn} is the power of BS local oscillator, $|S|$ is the number of selected antennas, P_{bs} is the circuit components of each BS antenna and P_{ue} is the circuit components of each UE. Note that the power consumption of BSs usually is very high in cell-free XL-MIMO systems, so it is vital to design algorithms to reduce the electricity cost. And P_{bs} accounting for the power-hungry radio-frequency chains is considered as 1 W per antenna.
- The power consumption of the channel encoding and decoding units shall be defined as $P_{c/d} = \sum_{k=1}^{K}\mathrm{SE}_k\left(\mathcal{P}_c + \mathcal{P}_d\right)$, where $\mathcal{P}_c$ and $\mathcal{P}_d$ denote the coding and decoding power densities, respectively.
- The power consumption of signal processing P_s can be quantified by the algorithm computational complexity $\mathcal{C}_{sp}$. The computational efficiency in BS and CPU is defined as $\mathcal{L}_{sp} = 75Gfops/W$. Thus, the power consumption $P_s = B\mathcal{C}_{sp}/\mathcal{L}_{sp}$, where B is the bandwidth.
- The power consumed by fronthaul links between BSs and CPU is modeled as $P_{bh} = \sum_{k=1}^{K}\mathrm{SE}_k\mathcal{P}_{bt}$, where $\mathcal{P}_{bt}$ is the fronthaul traffic power density.

Altogether, the simplified calculation of total power consumption $P_c = P_{fix} + P_{tc} + P_{c/d} + P_s + P_{bh}$ without the cost of accurate channel estimation.

Thus, in cell-free XL-MIMO systems, EE can be defined as the ratio between the number of bits and reliably transmitted power P_t, that is

$$\mathrm{EE} = \frac{B\cdot\sum_{k=1}^{K}\mathrm{SE}_k}{P_c}. \quad (8)$$

IV. MARL-empowered Antenna Selection

In this section, we devise the MARL-based antenna selection scheme for cell-free XL-MIMO wireless communication systems to enhance the system performance.

A. Antenna Selection

In cell-free XL-MIMO systems, as the number of antennas tends to be enormous, the circuit costs and computational complexity of conventional methods based on fully-digital receive arrays will increase dramatically. Additionally, it is unreasonable to activate all radio-frequency links and signal processing units in XL-MIMO. Exploring effective antenna selection techniques is necessary to reduce the number of antennas used in work patterns to enhance the system performance, particularly in energy-constrained environments facing the serious electricity burden and stringent computational complexity requirements [8].

In fact, the mechanism of antenna selection is to select enough antennas from the whole antenna array to reduce hardware costs and energy consumption. By assigning the BS subset of antennas to perform receive combining to each user, the interference power can be reduced. In a complicated dynamic environment, the real-time antenna on-off control

is preferable. However, conventional optimization algorithms have high computational complexity, which makes the solutions unfeasible. Therefore, we propose a novel MARL-based method that overcomes the above shortcomings in the following subsection.

B. MADDPG Antenna Selection Method

In conventional reinforcement learning, *Actor-Critic* methods are widely adopted for the real-time interactions between the environment and agents. By deploying RL methods into a multi-agent environment, we aim to solve the antenna selection problem with it. Owing to the difficulty to realize *Networked Distributed* strategy in practical scenarios, the emergence of CTDE with the simplification of the centralized learning process to an affordable degree is promising. In such a system, each agent consists of an *Actor Network* and a *Critic Network*, which are responsible for assigning actions and updating policies, respectively. Based on the fruitful CTDE system architecture, the use of MARL approaches the optimal joint strategy. In addition, many efficient algorithms have been derived, such as MADDPG, which has been successfully applied to cell-free mMIMO systems in recent years to solve intractable problems.

By assuming the cell-free XL-MIMO system as a multi-agent system, each UE antenna is regarded as an individual agent, selecting antennas from XL-MIMO panels based on the antenna selection criteria. In our work, the agents' actions depend on the current LSF information. All UE antennas independently complete the action allocation based on the local information, while the CPU completes the policy update based on the global information. The specific implemented procedure is summarized by the scheme of Algorithm 1.

Moreover, we describe the proposed antenna selection problem with a MARL tuple $< \mathcal{O}, \mathcal{A}, \mathcal{R}, \mathcal{P}, \gamma >$, where the state space $\mathcal{O} = [o_0, \ldots, o_t, \ldots]$ with the observed state $o_t = [o_{1,t}, \ldots, o_{KN_s,t}]$ and action space $\mathcal{A} = [a_0, \ldots, a_t, \ldots]$ with the assigned action $a_t = [a_{1,t}, \ldots, a_{KN_s,t}]$ at t time slot, depending on the LSF coefficients and the selected antenna indexing scheme, respectively. In the action strategy, '1' denotes selected, and vice versa. Furthermore, $\mathcal{P} : (\mathcal{O}, \mathcal{A}) \rightarrow \mathcal{O}$ is the state transition function, and γ is the discounted factor. The reward functions are $\mathcal{R} = [r_0, \ldots, r_t, \ldots]$ with the reward $r_t = [r_{1,t}, \ldots, r_{KN_s,t}]$ at t time slot. In this multi-agent environment, we model the optimization objective as maximizing the expected reward referring to equation (7) as follows:

$$\max \quad \mathbb{E}[\mathcal{R}] = \sum_{t=t_0}^{T} \gamma^{t-t_0} r_t = \sum_{t=t_0}^{T} \gamma^{t-t_0} \sum_{k=1}^{K} \mathrm{SE}_k^{(t)}, \qquad (9)$$

where t_0 is the current time, and T is the terminal time. Thus, the expected antennas are selected to serve the UEs based on the above objective.

With the CTDE architecture, it combines the current evaluation *Actor Network* θ_π and evaluation *Critic Network* θ_{Q_π} with an additional target *Actor Network* θ'_π and target *Critic Network* $\theta_{Q_{\pi'}}$ for an improved convergence rate. Each agent calculates its own policy gradient of the local *Actor Network* according to the joint state and action information. Also, the objective function for the i-th policy gradient π_i can be designed as $L(\pi_i) = \sum_{o_{i,t}} p_\pi(o_{i,t}) \sum_{a_{i,t}} \pi(a_{i,t}|o_{i,t}) r_{i,t}$. Correspondingly, the i-th reward $r_{i,t}$ is based on the global information a_t and o_t, leading to a centralized global action value $Q_\pi(o_t, a_t)$, which is calculated by the i-th *Critic Network*. The policy gradient of the local *Actor Network* for π_i is $\Delta_{\theta_{\pi_i}} J(\theta_{\pi_i}) = \sum_{a_{i,t}} Q_\pi(o_t, a_t) \Delta_{\theta_{\pi_i}} \pi_i(a_{i,t}|o_{i,t}; \theta_{\pi_i})$.

Besides, the global value $Q_{\pi_i}(o_{i,t}, a_{i,t})$ is calculated by the global *Critic Network*. Then, the mean-squared Bellman error function of the *Critic Network* for the i-th agent is $L(\theta_{Q_\pi}) = \mathbb{E}[(Q_\pi(o_t, a_t) - y_{i,t})^2]$ with the target $y_{i,t} = r_{i,t} + \gamma Q_\pi\Big(o_{t+1}, a_{t+1}|_{a_{t+1} \sim \pi(s_{t+1})}\Big)$. Finally, in order to ensure that the target network tends to be stable in the iterative process, the soft update is carried out in combination with the current network.

Algorithm 1 MADDPG Antenna Selection Algorithm

Input: The LSF information o_t of UE 1, ..., K

1: **while** $i <$ episode Q **do**
2: **for** agent j = 1 to N **do**
3: *Actor Network:* Compute the selected antennas set a_t based on the state information o_t; $o_t \leftarrow$ new state o'_t;
4: Store (o_t, a_t, r_t, o'_t) in replay buffer;
5: *Critic Network:* Sample a random minibatch from the replay buffer. Compute the reward function $\mathcal{R}$ and update *Critic Network* and *Actor Network* by minimizing the loss and policy gradient, respectively.
6: **end for**
7: Update target *Actor* network parameter and *Critic* network parameter for each agent: $\theta_{\pi'_i} \leftarrow \tau\theta_{\pi'_i} + (1-\tau)\theta_{\pi_i}$, $\theta_{Q_{\pi'}} \leftarrow \tau\theta_{Q_{\pi'}} + (1-\tau)\theta_{Q_\pi}$.
8: **end while**

Output: The set of selected antennas S

V. Numerical Results

With the SE and EE performance as well as the MADDPG antenna selection strategy talked above, we present numerical results for cell-free XL-MIMO systems. We adopt the same parameters as shown in [4], where the UEs are located in the near-field of BS. We assume that the fixed circuit power consumption P_{fix} = 18 W, power consumed by the local oscillator of the BS P_{syn} = 2 W, the power required for each BS antenna and UE antenna are P_{bs} = 1 W, P_{ue} = 0.2 W, respectively. The coding power density $\mathcal{P}_c$ = 0.1 W/(Gbit/s), the decoding power density $\mathcal{P}_d$ = 0.8 W/(Gbit/s), and the backhaul traffic power density $\mathcal{P}_{bt}$ = 0.25 W/(Gbit/s). The bandwidth B = 20 MHz. For convenience, we ignore the power consumption of other signal-processing processes.

In Fig. 2, we draw the cumulative distribution function (CDF) of the uplink sum SE under different antenna selection methods. For a fair comparison, we assume the case without antenna selection (No AS) as a benchmark. By introducing MADDPG, each UEs' antenna is regarded as an individual

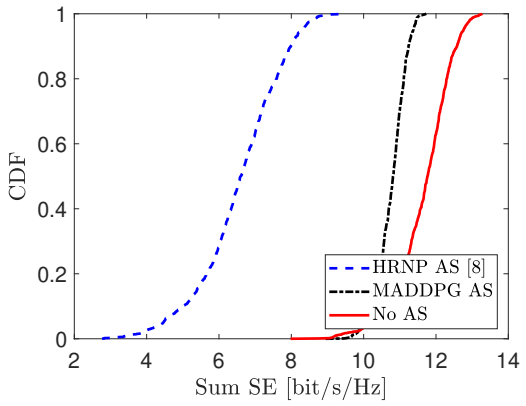

Fig. 2. CDF of the uplink sum SE performance of cell-free XL-MIMO systems under different antenna selection algorithms with single BS equipped with K = 6, $N_r = N_{H_r} \times N_{V_r}$ = 625, $N_s = N_{H_s} \times N_{V_s}$ = 9, and $\Delta_s = \Delta_r = \lambda/3$.

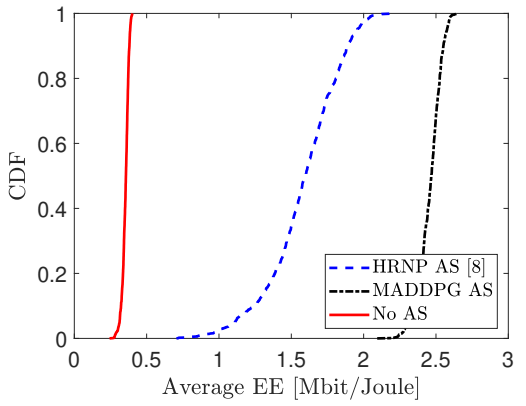

Fig. 3. Average EE of cell-free XL-MIMO systems under different antenna selection schema with K = 6, $N_r = N_{H_r} \times N_{V_r}$ = 625, $N_s = N_{H_s} \times N_{V_s}$ = 9, and $\Delta_s = \Delta_r = \lambda/3$.

agent to select different BS antennas for achieving the maximum SE. The highest received normalized power (HRNP) criterion [8] decreases the system performance with the sacrifice of degrees of freedom. It is noteworthy that the MADDPG method achieves a fairness balance of nearly 10.8 bit/s/Hz, which indicates the merit of cell-free mMIMO systems.

In Fig. 3, we discuss the EE performance of cell-free XL-MIMO systems. It's obvious that the antenna selection strategy boosts EE performance to different degrees. Compared with the baseline, the EE achieves 255%, and 638% improvement with the HRNP criterion and MADDPG algorithm, respectively. This is because the number of active antennas decreases dramatically, saving the huge power consumption. Additionally, the results show that the MADDPG effectively adjusts to the dynamic environment, which is promising to deploy in the distributed network to provide higher EE.

VI. Conclusions

In this paper, we investigated a cell-free XL-MIMO system with multiple UEs to achieve distributed signal processing. We focused on analyzing the SE and EE performance of such a system based on the realistic power consumption model. With the aid of the MADDPG algorithm and the distributed architecture, the SE performance achieves a uniform distribution. In addition, the proposed antenna selection strategy is efficient to improve the EE. In the future, we will focus on the resource allocation of MIMO systems with MARL.

References

[1] J. Zhang, E. Björnson, M. Matthaiou, D. W. K. Ng, H. Yang, and D. J. Love, "Prospective multiple antenna technologies for beyond 5G," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 8, pp. 1637–1660, Aug. 2020.

[2] Z. Liu, J. Zhang, Z. Wang, X. Zhang, H. Xiao, and B. Ai, "Cell-free massive MIMO with mixed-resolution ADCs and I/Q imbalance over Rician spatially correlated channels," *IEEE Trans. Veh. Technol.*, pp. 1–6, early access 2023.

[3] H. Lu and Y. Zeng, "Communicating with extremely large-scale array/surface: Unified modeling and performance analysis," *IEEE Trans. Wireless Commun.*, vol. 21, no. 6, pp. 4039–4053, June 2022.

[4] Z. Liu, Z. Liu, J. Zhang, H. Xiao, B. Ai, and D. W. K. Ng, "Uplink power control for extremely large-scale MIMO with multi-agent reinforcement learning and fuzzy logic," in *Proc. IEEE INFOCOM WKSHPS*, 2023, pp. 1–6.

[5] Z. Wang *et al.*, "Extremely large-scale MIMO: Fundamentals, challenges, solutions, and future directions," *IEEE Wireless Commun.*, pp. 1–9, 2023.

[6] Z. Wang, J. Zhang, B. Ai, C. Yuen, and M. Debbah, "Uplink performance of cell-free massive MIMO with multi-antenna users over jointly-correlated Rayleigh fading channels," *IEEE Trans. Wireless Commun.*, vol. 21, no. 9, pp. 7391–7406, 2022.

[7] Z. Wu and L. Dai, "Location division multiple access for near-field communications," *arXiv:2301.09082*, 2023.

[8] J. C. Marinello, T. Abrão, A. Amiri, E. de Carvalho, and P. Popovski, "Antenna selection for improving energy efficiency in XL-MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 69, no. 11, pp. 13 305–13 318, Nov. 2020.

[9] L. Yang, Q. Zhu, X. Ge, and L. Guo, "Transmit antenna selection for millimeter-wave MIMO system based on GBDT," *J. Commun. Inf. Netw.*, vol. 8, no. 1, pp. 71–79, Mar. 2023.

[10] B. Xu, Z. Wang, H. Xiao, J. Zhang, B. Ai, and D. W. K. Ng, "Low-complexity precoding for extremely large-scale MIMO over non-stationary channels," in *Proc. IEEE ICC*, 2023, pp. 1–6.

[11] X. Chai, H. Gao, J. Sun, X. Su, T. Lv, and J. Zeng, "Reinforcement learning based antenna selection in user-centric massive MIMO," in *Proc. IEEE VTC2020-Spring*, 2020, pp. 1–6.

[12] L. Wei *et al.*, "Multi-user Holographic MIMO surfaces: Channel modeling and spectral efficiency analysis," *IEEE J. Sel. Topics Signal Process.*, vol. 16, no. 5, pp. 1112–1124, Aug. 2022.

[13] H. Lei, Z. Wang, J. Zhang, and B. Ai, "Uplink performance of cell-free extremely large-scale MIMO systems," in *Proc. IEEE ICC*, 2023, pp. 1–6.

[14] A. Pizzo, L. Sanguinetti, and T. L. Marzetta, "Fourier plane-wave series expansion for holographic MIMO communications," *IEEE Trans. Wireless Commun.*, vol. 21, no. 9, pp. 6890–6905, Sep. 2022.

[15] E. Shi, J. Zhang, S. Chen, J. Zheng, Y. Zhang, D. W. Kwan Ng, and B. Ai, "Wireless energy transfer in RIS-aided cell-free massive MIMO systems: Opportunities and challenges," *IEEE Commun. Mag.*, vol. 60, no. 3, pp. 26–32, Mar. 2022.

[16] Y. Zhao, J. Zhang, and B. Ai, "Applications of reconfigurable intelligent surface in smart high speed train communications," *ZTE Commun.*, vol. 27, no. 4, pp. 36–43, Aug. 2021.

[17] Y. Guo, J. Zhang, Z. Lu, and W. Minghui, "Beam tracking and coverage enhancement algorithm for mobile users with intelligent reflecting surface," *ZTE Commun.*, vol. 27, no. 2, pp. 54–59, Apr. 2021.

[18] J. Zhang, S. Chen, Y. Lin, J. Zheng, B. Ai, and L. Hanzo, "Cell-free massive MIMO: A new next-generation paradigm," *IEEE Access*, vol. 7, pp. 99 878–99 888, 2019.

[19] E. Björnson and L. Sanguinetti, "Making cell-free massive MIMO competitive with MMSE processing and centralized implementation," *IEEE Trans. Wireless Commun.*, vol. 19, no. 1, pp. 77–90, Jan. 2020.

[20] Z. Liu, J. Zhang, Z. Wang, B. Ai, and D. W. K. Ng, "Cell-free massive MIMO with low-resolution ADCs and I/Q imbalance over spatially correlated channels," in *Proc. IEEE GLOBECOM*, 2022, pp. 2450–2455.