Fully-Decoupled RAN for Feedback-Free Multi-Base Station Transmission in MIMO-OFDM System

Yunting Xu, Zongxi Liu, Bo Qian, *Member, IEEE*, Hongyang Du, *Member, IEEE*, Jiacheng Chen, Jiawen Kang, *Senior Member, IEEE*, Haibo Zhou, *Senior Member, IEEE*, and Dusit Niyato, *Fellow, IEEE*

***Abstract*— Coordinated multi-base station (BS) transmission has emerged as a fundamental access technology to augment network capability and improve spectrum efficiency. However, the computation-intensive feedback of channel state information (CSI) poses significant challenges in determining physical-layer parameters for coordinated BSs. In this paper, we investigate a feedback-free mechanism that leverages fixed precoding matrix indicator (PMI), rank indicator (RI), and channel quality indicator (CQI) for coordinated BS transmission over a fully-decoupled radio access network (FD-RAN). Aiming to maximize user equipment (UE) throughput without CSI feedback, we calculate an optimal feedback-free parameter across spatial, frequency, and time domains only through UE geolocations. First, to determine MIMO transmission layer and precoding strategy in the spatial domain, we introduce a hierarchical reinforcement learning (HRL) framework to jointly select PMI and RI for coordinated BSs. Subsequently, for designing a more fine-grained subband transmission, transformer module is employed to capture the subcarrier correlations within OFDM symbols. Finally, given the unpredictable channel variations, we leverage a diffusion model to generate representative channel for fixed PMI, RI, and CQI over time-varied networks. Simulations demonstrate that 2 BSs feedback-free transmission can enhance 13% throughput compared with 1 BS CLSM transmission, which provides a design principle for next-generation transceiver technologies.**

***Index Terms*— 6G, fully-decoupled RAN, feedback-free transmission, generative AI, diffusion model.**

Received 15 March 2024; revised 31 August 2024; accepted 7 November 2024. Date of publication 20 January 2025; date of current version 27 February 2025. This work was supported in part by the National Natural Science Foundation Original Exploration Project of China under Grant 62250004; in part by the National Natural Science Foundation of China under Grant 62271244; in part by the Major Key Project of Peng Cheng Laboratory (PCL) under Grant PCL2024A03; in part by the Natural Science Fund for Distinguished Young Scholars of Jiangsu Province under Grant BK20220067; in part by the High-Level Innovation and Entrepreneurship Talent Introduction Program Team of Jiangsu Province under Grant JSSCTD202202; in part by the National Research Foundation, Singapore, and the Infocomm Media Development Authority under its Future Communications Research and Development Program; in part by the Defence Science Organization (DSO) National Laboratories under the AI Singapore Program under Grant FCP-NTU-RG-2022-010 and Grant FCP-ASTAR-TG-2022-003; in part by Singapore Ministry of Education (MOE) Tier 1 under Grant RG87/22; and in part by the NTU Centre for Computational Technologies in Finance (NTU-CCTF). *(Corresponding author: Haibo Zhou.)*

Yunting Xu, Zongxi Liu, and Haibo Zhou are with the School of Electronic Science and Engineering, Nanjing University, Nanjing 210023, China (e-mail: yuntingxu@smail.nju.edu.cn; zongxiliu@smail.nju.edu.cn; haibozhou@nju.edu.cn).

Bo Qian is with the Information Systems Architecture Science Research Division, National Institute of Informatics, Tokyo 101-8430, Japan (e-mail: boqian@ieee.org).

Hongyang Du is with the Department of Electrical and Electronic Engineering, The University of Hong Kong, Hong Kong (e-mail: duhy@eee.hku.hk).

Jiacheng Chen is with the Department of Strategic and Advanced Interdisciplinary Research, Peng Cheng Laboratory (PCL), Shenzhen 518000, China (e-mail: chenjch02@pcl.ac.cn).

Jiawen Kang is with the Automation of School, Guangdong University of Technology, Guangzhou 510006, China (e-mail: kavinkang@gdut.edu.cn).

Dusit Niyato is with the School of Computer Science and Engineering, Nanyang Technological University, Singapore 639798 (e-mail: dniyato@ntu.edu.sg).

Digital Object Identifier 10.1109/JSAC.2025.3531577

I. Introduction

With the ever-increasing demand for higher data rates and improved spectral efficiency in wireless networks, coordinated transmission among multi-base station (BS) has evolved as an innovative paradigm to support user equipment (UE) service for next-generation wireless networks (6G) [1], [2], [3], [4]. One of the state-of-the-art technologies in wireless coordination, namely co-frequency diversity employed by multi-BS transmission, has demonstrated significant potential in optimizing spectrum utilization and augmenting network capacity [5]. Nevertheless, the fundamental property of the performance gain brought by co-frequency diversity is critically dependent on the feedback of channel state information (CSI), especially for multiple-input multiple-output (MIMO) systems [6]. In current fifth-generation (5G) wireless networks, the real-time CSI is acquired by transmitting reference signals from the BS to the UE, enabling the calculation of physical-layer parameters, including precoding matrix indicator (PMI), MIMO rank indicator (RI), and channel quality indicator (CQI), for orthogonal frequency-division multiplexing (OFDM) symbol transmission [7]. The PMI, RI, and CQI will be subsequently fed back to the BS and be dynamically adjusted across different subframes. Consequently, when confronted with the joint determination of physical-layer parameters for multi-BS, the intensive computation and communication of real-time CSI feedback has posed significant challenges for the exploitation of co-frequency diversity [8].

Targeting at reducing CSI feedback overhead while guaranteeing the wireless data rate, extensive network optimization technologies, e.g., sparse [9] and quantized [10] CSI feedback, have been investigated in the past few years. However, these approaches cannot essentially reduce the complexity and frequency for the joint calculation of physical-layer parameters involving multi-BS [11]. Recently, an evolutionary

fully-decoupled radio access network (FD-RAN) architecture, which employs a feedback-free mechanism for physical-layer transmission, has been proposed to enable the extensible and flexible design for multi-BS transmission [12], [13], [14]. Instead of relying on the frequent feedback of CSI to calculate physical-layer parameters, the feedback-free mechanism in FD-RAN leverages fixed PMI, RI, and CQI across all transmission periods based on the geolocation information of UE's position. The fixed physical-layer parameters will be precalculated with geolocation-related statistical channel data, thus avoiding the intensive real-time CSI processing. This innovative transmission design has been regarded as a promising resolution for multi-BS cooperation [15].

Nevertheless, although the feedback-free mechanism of FD-RAN simplifies the frequent CSI feedback process, the determination of time-invariant physical-layer parameters for each UE location remains a significant challenge. Given the dynamic and stochastic characteristics of the wireless channel across different subframes, directly utilizing geolocation-based statistical channel data may result in unsatisfactory performance [16]. However, channels at one location exhibit high correlations in the spatial, frequency, and time domains. With the recent revolutionary development of generative artificial intelligence (GAI) technologies, cutting-edge GAI algorithms such as the diffusion model and variational autoencoder (VAE) have shown considerable potential in accurately capturing the correlation within dynamically changing environments [17], [18]. Consequently, through learning the distribution of the statistical channel data across the spatial, frequency, and time domains, GAI is capable of generating a representative channel that adequately adapts to a wide range of time-varied wireless conditions [19], [20]. The generated representative channel can be exploited to obtain the time-invariant physical-layer parameters for coordinated multi-BS transmission, thus providing an effective solution for the feedback-free mechanism of the FD-RAN architecture.

Based on the aforementioned investigation, we summarize the challenges and motivations of utilizing the innovative FD-RAN architecture for multi-BS transmission as follows

- Communication overhead: To eliminate the necessity for frequent real-time CSI feedback between the UE and the coordinated BSs, FD-RAN consistently employs the time-invariant PMI, RI, and CQI across all subframes, thus significantly reducing the communication overhead for coordinated multi-BS transmission.
- Computation overhead: The complexity of the joint calculation for physical-layer parameters grows exponentially with the number of coordinated BSs, making it computationally prohibitive for each subframe transmission. FD-RAN precalculates the time-invariant physical-layer parameters based on statistical channel data, thereby substantially reducing the intensive computation overhead.
- Dynamic wireless environment: The stochastic characteristic of the changing wireless channel results in unsatisfactory performance for the fixed parameter transmission of FD-RAN. Through learning the distribution of statistical channel data across different subframes, GAI can generate a representative channel to tremendously enhance the performance of FD-RAN feedback-free transmission.

To this end, this paper targets at enhancing the feedback-free transmission performance of FD-RAN for coordinated multi-BS transmission in MIMO-OFDM system. We formulate a problem that aims to maximize the UE throughput, with an optimization of physical-layer parameters within the spatial, frequency, and time domains based GAI technologies. Specifically, the PMI will be tailored to each subcarrier in an OFDM symbol to enable more fine-grained subband transmission, the RI will decide the joint transmit data layer for all coordinated BSs, and the CQI will select a modulation and coding scheme (MCS) to determine the overall throughput of the coordinated multi-BS transmission. The detailed process of the FD-RAN feedback-free mechanism includes: (1) Leverage location-based statistical channel data to precalculate the PMI, RI, and CQI for coordinated BSs; and (2) Consistently apply these physical-layer parameters across all transmission periods. The main contributions of this paper are highlighted as follows

- We design a hierarchical reinforcement learning (HRL) framework to jointly optimize PMI and RI on co-frequency transmission. The higher-tier of HRL will determine the MIMO spatial layer and the lower-tier will select the precoding matrix for all coordinated BSs, which provides an efficient way to handle the complicated precoding problem in the spatial domain.
- We employ the transformer encoder with a scaled dot-product attention module to capture the inherent correlations of subcarriers within an OFDM symbol, which offers a more fine-grained subband transmission in the frequency domain compared with OLSM[1] mechanism.
- Given the continuous and unpredictable channel variations in the time domain, we exploit a conditional diffusion model, one type of GAI, to generate a representative channel for calculating the MSC, which will be effective for a wide range of dynamic time-varied conditions.
- Extensive simulations have been conducted in the Vienna 5G Link Level Simulator [22] with a complete physical-layer module. The results have verified that a feedback-free transmission of 2 BS can enhance the UE throughput by 13% compared to 1 BS CLSM[2] and reach 94% performance of a heuristic codebook iteration (HCI) method, which provides a design principle for the development of next-generation wireless networks.

The rest of this paper is structured as follows. Section II introduces the related work. Section III presents the system model. Section IV describes the feedback-free resolution and the simulation results are given in Section V. Finally, we conclude the contributions in Section VI.

[1]Open-loop spatial multiplexing (OLSM) corresponds to the transmission mode (TM) 3 in Long Term Evolution (LTE) [21]. OLSM provides the feedback information using a predefined PMI sequence for each subcarrier in an OFDM symbol, along with an RI and a CQI.

[2]Closed-loop spatial multiplexing (CLSM) corresponds to the TM 4 in LTE [21]. CLSM provides the feedback information using a PMI for all subcarriers in an OFDM symbol, along with an RI and a CQI.

II. Related Work

The field of coordinated multi-BS transmission has witnessed substantial development in recent decades owing to its potential to enhance spectrum utilization and augment network capacity [23]. State-of-the-art techniques such as coordinated multi-point (CoMP) transmission and cell-free massive MIMO leverage the co-frequency diversity to enhance signal strength, thereby improving UE throughput and ensuring more reliable connectivity across extensive networks [24]. However, the current physical-layer transmission mode relies on the UE periodical measurement report of CSI feedback, which is computation-intensive and results in high communication overhead due to the exponential increased coordination complexity with the number of BSs. To alleviate this situation, considerable research works have been conducted utilizing limited CSI feedback technologies. Chen et al. [9] leveraged sparsity-inducing mean-squared error (MSE) minimization for cell-free transmission. Silva et al. [10] proposed a low-complexity BS block compression method to relieve the fronthaul burden in distributed MIMO system. These technologies have been demonstrated to achieve nearly equivalent spectral efficiency to the transmission attained with perfect CSI, while significantly reducing the overhead of CSI feedback.

To further address the challenges posed by real-time computational complexity in CSI feedback procedure, the mechanism of physical-layer transmission without any CSI information has received extensive attention. Bletsas et al. [25] explored a zero-feedback scheme wherein carrier synchronization is abandoned among distributed transmitters and verified the possibility of deploying the zero-feedback scheme in low-cost sensor networks. Hanna et al. [26] proposed a destination-feedback-free approach that leverages coarse radio mobility and destination information for distributed beamforming process. Simulations have demonstrated an effective signal to noise ratio (SNR) improvement of 9 dB in the beamforming direction, achieved through a broad placement of four radio units. In addition, the UE location based feedback-free mechanism has been widely adopted to optimize the network performance. Chikha et al. [27] formulated a radio environment map (REM) for inter-cell coordination in massive-MIMO systems. Wang et al. [28] proposed a location-aware beamforming method in cell-free millimeter wave scenario. Furthermore, Yu et al. [12] leveraged UE location to obtain a fixed PMI, RI, and CQI in coordination BS transmission. which provides an efficient way to simplify operations and enhance coordination efficiency.

Apart from the flexibility benefits that feedback-free transmission affords in BS coordination, there remains significant potential for performance improvement through the employment of advanced AI technologies with geolocation information [29], [30], [31]. Ji et al. [32] proposed a multi-agent deep reinforcement learning method with rate splitting multiple access for beamforming configuration. Xu et al. [33] introduced a reservoir computing structure-based network (RC-net) to realize the rank and link adaptation for MIMO-OFDM Detection. Recently, generative AI has shown enormous potential in the optimization of physical-layer performance [34], [35]. Ye et al. [36] used a generative adversarial network (GAN) based channel agnostic communication to realize an efficient physical-layer transmission. Liu et al. [19] leveraged the VAE to generate a representative precoding matrix with a location based feedback-free mechanism. Kim et al. [37] exploited a diffusion model for end-to-end (E2E) channel coding design, which demonstrates extensive capability in creating an effective solution over complicated time-varied channels and achieves significant throughput enhancement.

III. System Model

In this section, we present the wireless channel model and the transmission process of multi-downlink BSs (DBSs) coordination over FD-RAN architecture.

A. Channel Model

Consider that the UE is equipped with N_r receive antennas, and each DBS has N_t transmit antennas, we use the clustered delay line (CDL) defined in 3GPP 38.901 [38] to model the channel impulse response (CIR). As given in Eq. (1), as shown at the bottom of the page, the time-variant CIR $h_{r,t}(\tau)$ of the r-th receive and t-th transmit antenna is formulated by adding the line of sight (LoS) channel $h_{r,t}^{LoS}(\tau, \Delta\tau_{LoS})$ to the non-line of sight (NLoS) components $h_{r,t,c,p}(\tau, \Delta\tau_{c,p})$, where τ denotes the time instant, $\Delta\tau_{LoS}$ and $\Delta\tau_{c,p}$ are the time delay of LoS and NLoS paths, respectively, $c \in [1, C]$ describes the modeled channel clusters, and $p \in [1, N_{p_c}]$ is the subpath of cluster c. The subpath p in the same cluster will share similar angle of arrival (AoA) and angle of departure (AoD), having the angle gains of G_{AoA} and G_{AoD}, respectively. The DBS transmit antenna is considered to be a dual-polarized uniform planar array (UPA), which consists of N_{t_1} horizontal and N_{t_2} vertical antennas. The transmit antenna number can be calculated as $N_t = 2N_{t_1}N_{t_2}$ and the antenna gain is denoted as G_t. The UE received antenna is considered to be a single-polarized uniform linear array (ULA) with an antenna gain of G_r. Consequently, the channel gains of G_{LoS} and $G_{c,p}$ are expressed as

$$G_{LoS/c,p} = G_r G_{AoA} G_t G_{AoD}. \tag{2}$$

Besides, the Ricean K-factor K_R in Eq. (1) describes the scaled power ratio of LoS path and NLoS paths. ϕ_{LoS} and

$$\begin{aligned} h_{r,t}(\tau) &= \sqrt{\frac{K_R}{K_R+1}} h_{r,t}^{LoS}(\tau, \Delta\tau_{LoS}) + \sqrt{\frac{1}{K_R+1}} \sum_{c=1}^{C} \sum_{p=1}^{N_{p_c}} h_{r,t,c,p}(\tau, \Delta\tau_{c,p}) \\ &= \sqrt{\frac{K_R}{K_R+1}} G_{LoS} e^{j(\phi_{LoS} + 2\pi f_c \Delta\tau_{LoS})} \delta(\tau - \Delta\tau_{LoS}) + \sqrt{\frac{1}{K_R+1}} \sum_{c=1}^{C} \sum_{p=1}^{N_{p_c}} G_{c,p} e^{j(\phi_{c,p} + 2\pi f_c \Delta\tau_{c,p})} \delta(\tau - \Delta\tau_{c,p}). \end{aligned} \tag{1}$$

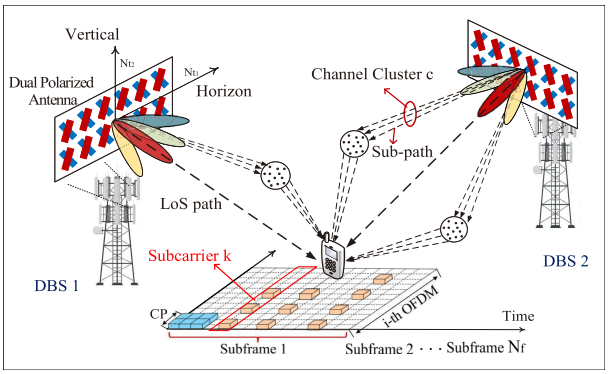


Fig. 1. Coordinated multi-DBS transmission in FD-RAN.

$\phi_{c,p}$ are the initial phases of the channel paths, f_c is the carrier signal frequency, and $\delta(\cdot)$ is the Dirac delta function to model the unit channel impulse that includes the time delay effect.

B. Coordinated Downlink Data Transmission

As shown in Fig. 1, different DBSs in FD-RAN will coordinately transmit the same signal to a UE and the coordinated DBS set is defined as $\mathcal{D} = \{1, \cdots, D\}$. For OFDM symbol i, the UE received signal in time domain $y^i_{r,t}(\tau)$ is expressed as

$$y^i_{r,t}(\tau) = \sum_{d=1}^{D} \alpha_d h^i_{d,r,t}(\tau) * x^i(\tau) + \mathrm{n}^i(\tau), \tag{3}$$

where α_d and $h^i_{d,r,t}(\tau)$ denote the pathloss and the CIR between DBS d and the UE, respectively, $x^i(\tau)$ and $\mathrm{n}^i(\tau)$ are the transmit symbol and received noise in time domain, and $*$ denotes the convolution operation. Since OFDM is typically transmitted in frequency domain, $h^i_{d,r,t}(\tau)$, $x^i(\tau)$ and $\mathrm{n}^i(\tau)$ will be sampled at a sampling rate of $N_s\Delta f$ to obtain the discrete sequences $h^i_{d,r,t}[n]$, $x^i[n]$, $\mathrm{n}^i[n]$, and $y^i_{r,t}[n]$, $n = 0, \cdots, N_s - 1$, where N_s denotes the sampling number and Δf is the subcarrier spacing. Given that different DBSs transmit the same OFDM symbol to the UE, the signals from different DBSs can be looked upon as multi-paths transmission. Therefore, the inter-symbol interference (ISI) caused by multi-paths effect can be effectively mitigated using the cyclic prefix (CP) technique. Consider the delay of all paths from DBS $\forall d \in \mathcal{D}$ to be less than the duration of CP, for a specific subcarrier k, the discrete UE received signal $y^i_r[k]$ at antenna r is calculated as the sum of discrete signal $y^i_{r,t}[k]$ from all transmit antennas as follows:

$$y^i_r[k] = \sum_{t=1}^{N_t} y^i_{r,t}[k] = \sum_{t=1}^{N_t}\sum_{d=1}^{D} h^i_{d,r,t}[k]x^i[k] + \mathrm{n}^i[k], \tag{4}$$

where $y^i_{r,t}[k]$, $h^i_{d,r,t}[k]$, $x^i[k]$, and $\mathrm{n}^i[k]$ are the discrete Fourier transform (DFT) of $y^i_{r,t}[n]$, $\alpha_d h^i_{d,r,t}[n]$, $x^i[n]$, and $\mathrm{n}^i[n]$, respectively. An example of $\mathrm{DFT}\{x^i[n]\}$ is given as follows:

$$x^i[k] = \mathrm{DFT}\{x^i[n]\} = \sum_{n=0}^{N_s-1} x^i[n] e^{-\mathrm{j}\frac{2\pi}{N_s}nk}. \tag{5}$$

The detailed derivation of Eq. (4) is provided in Appendix A.

Remark 1: The normal CP duration t_{cp} is defined as 5.21 μs in 3GPP 38.211 [39] for subcarrier spacing Δ_f = 15 KHz. The difference distance of multi-paths Δd is calculated by multiplying the CP duration with the speed of light v, i.e., $\Delta d = t_{cp} \times v$, to possess the value of 1563 m. Consequently, the ISI caused by different channel paths of DBS $\forall d \in \mathcal{D}$ can be eliminated once their propagation distance difference is less than Δd (1563 m), which satisfies almost all situations under the coverage range of current 5G BSs.

Consider the MIMO transmission with multiple layer data, let $\mathbf{x}^i_k \in \mathbb{C}^{N_\ell \times 1}$ be the N_ℓ layer transmit data and $\mathbf{y}^i_k \in \mathbb{C}^{N_r \times 1}$ be the received signal at subcarrier k defined as follows:

$$\mathbf{x}^i_k = \left[x^i_1[k], \cdots, x^i_{N_l}[k]\right]^{\mathrm{T}}, \tag{6}$$

$$\mathbf{y}^i_k = \left[y^i_1[k], \cdots, y^i_{N_r}[k]\right]^{\mathrm{T}}. \tag{7}$$

Then, Eq. (4) is transformed into

$$\mathbf{y}^i_k = \sum_{d\in\mathcal{D}} \left(\mathbf{H}^i_{d,k}\mathbf{W}_{d,k}\right)\mathbf{x}^i_k + \mathbf{n}^i_k, \tag{8}$$

where $\mathbf{W}_{d,k} \in \mathbb{C}^{N_t \times N_\ell}$ is the precoding matrix of DBS d to map the N_ℓ-th layer data $\mathbf{x}^i_k$ into N_t transmit antenna ports. $\mathbf{H}^i_{d,k} \in \mathbb{C}^{N_r \times N_t}$ is the channel matrix with its r-th row and t-th column $\mathbf{H}^i_{d,k}(r,t) = h^i_{d,r,t}[k]$. $\mathbf{n}^i_k \sim \mathcal{CN}\left(0, \sigma^2\mathbf{I}_{N_r}\right)$ is the noise with variance σ^2. The transmit layer N_ℓ should not exceed the least number of transmit and receive antennas as

$$1 \le \max\left(N_\ell\right) \le \min\left(N_t, N_r\right). \tag{9}$$

C. Channel Equalization

After UE receives the signal $\mathbf{y}^i_k$ from all DBS, a channel equalization process will be performed using zero forcing (ZF) or minimum mean square error (MMSE) equalizer. We use $\mathbf{H}^i_k\mathbf{W}_k \in \mathbb{C}^{N_r \times N_\ell}$ to denote the effective channel matrix as

$$\mathbf{H}^i_k\mathbf{W}_k = \sum_{d=1}^{D}(\mathbf{H}^i_{d,k}\mathbf{W}_{d,k}), \tag{10}$$

where $\mathbf{H}^i_k \in \mathbb{C}^{N_r \times DN_t}$ and $\mathbf{W}_k \in \mathbb{C}^{DN_t \times N_\ell}$ are defined as

$$\mathbf{H}^i_k = \left[\mathbf{H}^i_{1,k}, \cdots, \mathbf{H}^i_{D,k}\right], \tag{11}$$

$$\mathbf{W}_k = \left[\mathbf{W}^{\mathrm{T}}_{1,k}, \cdots, \mathbf{W}^{\mathrm{T}}_{D,k}\right]^{\mathrm{T}}. \tag{12}$$

The equalization process will reverse the channel effects, thus recovering the received signal to the original transmitted signal. We consider the ZF equalizer that uses the pseudo inverse of effective channel matrix $\mathbf{H}^i_k\mathbf{W}_k$ as

$$\mathbf{Z}^i_k = \left(\left(\mathbf{H}^i_k\mathbf{W}_k\right)^{\mathrm{H}}\mathbf{H}^i_k\mathbf{W}_k\right)^{-1}\left(\mathbf{H}^i_k\mathbf{W}_k\right)^{\mathrm{H}}, \tag{13}$$

where $(\cdot)^{\mathrm{H}}$ is the Hermitian transpose and $\mathbf{Z}^i_k \in \mathbb{C}^{N_\ell \times N_r}$ is the ZF equalizer for subcarrier k in i-th OFDM symbol. After filtering the received signal $\mathbf{y}^i_k$ with $\mathbf{Z}^i_k$, the post-equalization symbol vector $\widetilde{\mathbf{y}}^i_k \in \mathbb{C}^{N_\ell \times 1}$ can be calculated as

$$\begin{aligned}\widetilde{\mathbf{y}}^i_k = \mathbf{Z}^i_k\mathbf{y}^i_k &= \mathbf{Z}^i_k\left(\sum_{d=1}^{D}\left(\mathbf{H}^i_{d,k}\mathbf{W}_{d,k}\right)\mathbf{x}^i_k + \mathbf{n}^i_k\right)\\ &= \underbrace{\mathbf{Z}^i_k\mathbf{H}^i_k\mathbf{W}_k}_{\mathbf{E}^i_k \in \mathbb{C}^{N_\ell \times N_\ell}}\mathbf{x}^i_k + \underbrace{\mathbf{Z}^i_k\mathbf{n}^i_k}_{\mathbb{C}^{N_\ell \times 1}},\end{aligned} \tag{14}$$

where $\mathbf{E}_k^i = \mathbf{Z}_k^i \mathbf{H}_k^i \mathbf{W}_k$ recovers the received signal for N_ℓ transmit data layers. Based on Eq. (14), we obtain the post-equalization signal-to-interference-plus-noise ratio $\mathrm{SINR}_{k,l}^i$ for layer $l \in \{1, \cdots, N_\ell\}$ in the following expression

$$\mathrm{SINR}_{k,l}^i = \frac{\left|\mathbf{E}_k^i(l,l)\right|^2}{\sum_{l' \neq l} \left|\mathbf{E}_k^i(l,l')\right|^2 + \sigma^2 \sum_r \left|\mathbf{Z}_k^i(l,r)\right|^2}, \quad (15)$$

where $\mathbf{E}_k^i(l,l')$ is the l-th row and the l'-th column of $\mathbf{E}_k^i$. Through the mitigation of the effective channel effects, post-equalization SINR offers a more accurate representation for the received signal quality after the channel equalization process.

D. Problem Formulation

To measure the influence of precoding matrix $\mathbf{W}_{d,k}$, $\forall d \in \mathcal{D}$ for coordinated DBSs transmission, the definition mutual information M_k^i based on post-equalization SINR is introduced in this paper, which is calculated as the sum rate of all spatial layers based on the channel capacity of Shannon Theory [40]:

$$\begin{aligned} M_k^i &= \sum_{l=1}^{N_\ell} \log_2\left(1 + \mathrm{SINR}_{k,l}^i\left(\mathbf{W}_k\right)\right) \\ &= \sum_{l=1}^{N_\ell} \log_2\left(1 + \frac{\left|\mathbf{E}_k^i(l,l)\right|^2}{\sum_{l' \neq l} \left|\mathbf{E}_k^i(l,l')\right|^2 + \sigma^2 \sum_r \left|\mathbf{Z}_k^i(l,r)\right|^2}\right). \end{aligned} \quad (16)$$

The basic idea to choose the optimal precoding matrix is to find $\mathbf{W}_k^* = [\mathbf{W}_{1,k}^{*\mathrm{T}}, \cdots, \mathbf{W}_{D,k}^{*\mathrm{T}}]^{\mathrm{T}}$ that maximizes the mutual information over the total system bandwidth and subframe duration, i.e., over subcarrier- $(1, \ldots, K)$ and temporal-range of OFDM- $(1, \ldots, I)$, as can be expressed as follows:

$$\begin{aligned} \mathbf{W}_k^* &= \arg \max_{\mathbf{W}_k \in \mathcal{W}} \sum_{k=1}^{K} \sum_{i=1}^{I} M_k^i\left(\mathbf{W}_k\right) \\ &= \arg \max_{\mathbf{W}_k \in \mathcal{W}} \sum_{k=1}^{K} \sum_{i=1}^{I} \sum_{l=1}^{N_\ell} \log_2\left(1 + \mathrm{SINR}_{k,l}^i\right), \end{aligned} \quad (17)$$

where $\mathcal{W} = \{\mathcal{W}_l\}_{l=1}^L$ is the predefined L layer precoding matrices with $\mathcal{W}_l$ being the l-th layer precoding matrix set. If a precoding matrix $\mathbf{W}_{d,k} \in \mathcal{W}_l$, the transmit layer is determined by $\mathrm{rank}(\mathbf{W}_{d,k}) = l$. Therefore, once $\mathbf{W}_k^*$ is given, we can obtain the optimal transmit layer $N_\ell{}^*$ for DBS $d \in \mathcal{D}$

$$N_\ell{}^* = l = \mathrm{rank}\left(\mathbf{W}_{d,k}^*\right), \quad \forall d \in \mathcal{D}, \quad (18)$$

and $\mathrm{rank}(\mathbf{W}_{1,k}) = \mathrm{rank}(\mathbf{W}_{2,k}) = \cdots = \mathrm{rank}(\mathbf{W}_{D,k})$ since DBS $d \in \mathcal{D}$ will transmit the same OFDM symbols and possesses the same transmit layer.

For CQI selection of transmissions from multi-DBS, we follow the single BS-UE connection strategy that averages the post-equalization SINR over the frequency-time resources of interest. Effective SINR mapping (ESM) methods map several post-equalization SINR values into an equivalent $\mathrm{SNR}_{\mathrm{eff}}$ value of an additive white gaussian noise (AWGN) channel [41]. The equivalent AWGN channel has similar block error rate (BLER) to the original OFDM system. A general mathematical expression of ESM is given by

$$\mathrm{SNR}_{\mathrm{eff}} = \beta \mathrm{f}^{-1}\left(\frac{1}{N^{\mathrm{pe}}} \sum_{\mathrm{pe}=1}^{N^{\mathrm{pe}}} \mathrm{f}\left(\frac{\mathrm{SINR}^{\mathrm{pe}}}{\beta}\right)\right), \quad (19)$$

where the value of β is predefined in 3GPP 38.214 [42] for different MCSs. f can be the exponential function for exponential effective SINR mapping (EESM) method, and N^{pe} is the total number of post-equalization values $\mathrm{SINR}^{\mathrm{pe}}$. The result of $\mathrm{SNR}_{\mathrm{eff}}$ corresponds to the CQI that specifies the modulation scheme and coding rate.

The UE throughput is jointly decided by PMI, RI, and CQI. Instead of dynamically updating these parameters through real-time CSI feedback, the feedback-free scheme will employ fixed PMI, RI, and CQI across all subframes. The objective is to find the optimal PMI, RI, and CQI of each location to maximize the UE throughput. Consequently, we can maximize the expectation of mutual information to obtain $\mathbf{W}_k^*$, i.e.,

$$\max_{\mathbf{W}_k} \quad \mathbb{E}\left[\sum_{k=1}^{K} \sum_{i=1}^{I} M_k^i\left(\mathbf{W}_k\right)\right], \quad (20)$$

Then, $\mathrm{SINR}_{k,l}^i\left(\mathbf{W}_k^*\right)$ will be taken into Eq. (19) to calculate the optimal CQI. However, the expectation of mutual information is difficult to obtain given the continuous channel variations over time. Therefore, we leverage the Monte Carlo method [19] to approximate the expectation with N_f subframe channels and the objective function is reformulated as

$$\max_{N_\ell, \mathbf{W}_{d,k}} \frac{1}{N_f} \sum_{f=1}^{N_f} \sum_{k=1}^{K} \sum_{i=1}^{I} \sum_{l=1}^{N_\ell} \log_2\left(1 + \mathrm{SINR}_{k,l}^{i\ (f)}\right) \quad (21a)$$

$$\text{s.t.} \quad 1 \leq \max\left(N_\ell\right) \leq \min\left(N_t, N_r\right), \quad (21b)$$

$$\mathbf{W}_{d,k} \in \mathcal{W}_{N_\ell}, \quad N_\ell \in \mathcal{L}, \ \forall k \in \mathcal{K}, \ \forall d \in \mathcal{D}, \quad (21c)$$

$$\mathrm{rank}\left(\mathbf{W}_{d,k}\right) = N_\ell, \ N_\ell \in \mathcal{L}, \ \forall k \in \mathcal{K}, \ \forall d \in \mathcal{D}, \quad (21d)$$

$$\left\|\mathbf{W}_{d,k}\right\|_F = 1, \quad \forall k \in \mathcal{K}, \ \forall d \in \mathcal{D}. \quad (21e)$$

Eq. (21a) aims to maximize the average N_f frame mutual information, i.e., $\frac{1}{N_f}\sum_{f=1}^{N_f}[\sum_{k=1}^{K}\sum_{i=1}^{I} M_k^i(\mathbf{W}_k)]$, thereby replacing the expectation operation $\mathbb{E}[\cdot]$ in Eq. (20). Constraint (21b) bounds the minimum and maximum number of transmit layers. Contraints (21c) and (21d) ensure that $\mathbf{W}_{d,k}$ of all DBS $d \in \mathcal{D}$ belong to the N_ℓ-th layer precoding matrix set with $\mathcal{L} = \{1, \cdots, L\}$. $\|\cdot\|_F$ in constraint (21e) is the Frobenius norm that guarantees the precoding matrix is scaled with unit power. Subsequently, the optimal precoding matrix $\mathbf{W}_{d,k}^*$ obtained by problem (21) will be used in the calculation of $\mathrm{SNR}_{\mathrm{eff}}$ using the EESM method, which transforms the Eq. (19) as

$$\mathrm{SNR}_{\mathrm{eff}} = -\beta \ln\left(\frac{1}{N_f K I N_l} \sum_{f=1}^{N_f} \sum_{k=1}^{K} \sum_{i=1}^{I} \sum_{l=1}^{N_\ell} \exp\left(-\frac{\mathrm{SINR}_{k,l}^{i\ (f)}}{\beta}\right)\right) \quad (22)$$

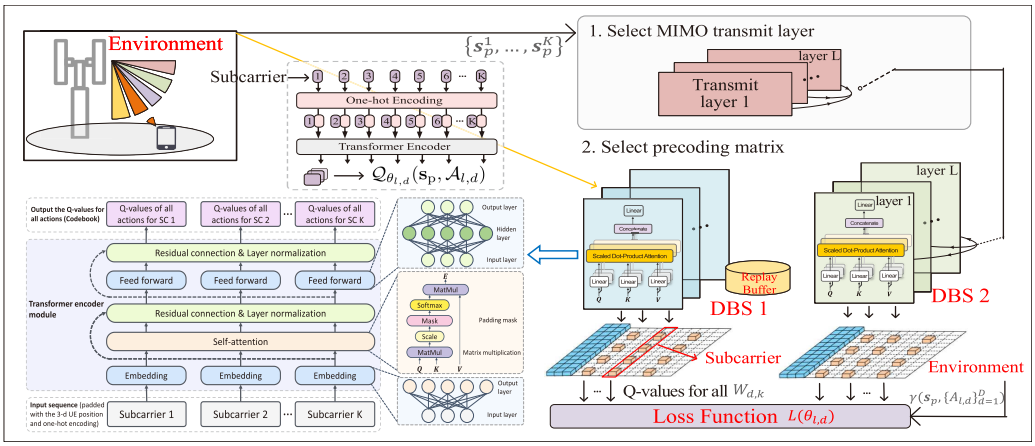


Fig. 2. Transformer encoder based HRL framework for codebook selection by using geolocation information.

where $N_f K I N_l$ is the total number of the post-equalization SINR values. The final CQI is obtained by looking up the SNR to CQI Mapping_Table [22] as

$$\text{Mapping\_Table}\left(\text{SNR}_{\text{eff}}\right) \longrightarrow \text{CQI}. \tag{23}$$

IV. Proposed Feedback-Free Solution

In this section, we first propose a transformer encoder based HRL algorithm to obtain optimal precoding matrices under a specific subframe. The higher tier of HRL network will select the option for transmit layers and the lower tier will sequentially output the $\mathcal{Q}$-values for all precoding matrices. Subsequently, a conditional diffusion model is employed to generate a representative channel $\boldsymbol{h}_k^r,\ \forall k \in \mathcal{K}$ for the original N_f subframe channel matrices, transforming problem (21) into

$$\max_{N_\ell, \mathbf{W}_{d,k}} \quad \sum_{k=1}^{K}\sum_{i=1}^{I}\sum_{l=1}^{N_\ell} \log_2\left(1+\text{SINR}_{k,l}^{i}(\boldsymbol{h}_k^r)\right) \tag{25a}$$

$$\text{s.t.} \quad (21a)-(21e). \tag{25b}$$

Then, a transmit layer N_ℓ and a precoding matrix $\mathbf{W}_{d,k}$ can be obtained by solving problem (25) based on the output of the proposed HRL network and the generated representative channel. The obtained transmit layer N_ℓ and the precoding matrix $\mathbf{W}_{d,k}$ will be fixed over the N_f subframe channels, thus providing a feasible solution to problem (21). In addition, the generated representative channel will be used to calculate the post-equalization SINR to obtain the optimal CQI based on Eq. (22) and Eq. (23).

A. Transformer Encoder Based HRL Algorithm

Fig. 2 presents the proposed HRL framework to train a codebook selection strategy based on geolocation information. We define the training state, action, and reward as follows.

1) State: The feedback-free transmission mechanism will select physical layer parameters only through geolocation information. Therefore, the three-dimensional coordinates p_x, p_y, and p_z of the UE position p are used as the state information. Besides, in order to model the feature of subcarrier, the $\text{onehot}(k) \in \mathbb{C}^{1\times K}$ encoding for subcarrier k is introduced, with its k'-th element expressed as

$$\text{onehot}(k)_{k'} = \begin{cases} 1 & \text{if } k'=k \\ 0 & \text{else} \end{cases}, \ k'=1,\cdots,K. \tag{26}$$

Consequently, the network state is formulated as

$$\mathbf{s}_{\text{p}}^k = \left[\text{p}_x,\ \text{p}_y,\ \text{p}_z,\ \text{onehot}\,(k)\right]^{\text{T}}, \quad \forall k \in \mathcal{K}. \tag{27}$$

2) Action: In the HRL framework, we exploit two-tier actions to jointly optimize the RI and PMI. Each DBS is modeled as an agent to choose their own actions. For agent d, if the selected transmit layer is $l \in \mathcal{L}$, we define the higher tier action for DBS d as $\mathcal{A}_d^{high} = l$. The maximum value of $\mathcal{A}_d^{high}$ is limited to $\min\left(N_t, N_r\right)$, ensuring that the transmit layer does not exceed the least number of transmit and receive antennas in constraint (21b). Based on the transmit layer l, each DBS agent will choose the precoding matrix $\mathbf{W}_{d,k}$ for the low tier actions $\mathcal{A}_d^{low}$ from the l-th layer precoding set $\mathcal{W}_l$ for all subcarrier $k \in \mathcal{K}$. Therefore, we formulate the HRL action $\{\mathcal{A}_{l,d}\}_{d=1}^{D}$ for all DBS agents with l transmit layers as

$$\{\mathcal{A}_{l,d}\}_{d=1}^{D} = \begin{cases} \mathcal{A}_d^{high} = l,\ l \in \mathcal{L},\ \forall d \in \mathcal{D}, \\ \mathcal{A}_d^{low} = \{\underbrace{\mathbf{W}_{1,k},\cdots,\mathbf{W}_{D,k}}_{\mathbf{W}_{d,k}\in\mathcal{W}_l,\ \forall d\in\mathcal{D}}\}. \end{cases} \tag{28}$$

Remark 2: We leverage the predefined Type I codebook $\mathcal{W}_{\text{TypeI}} = \{\mathcal{W}_l\}_{l=1}^{L}$ for the HRL action. 3GPP 38.214 [42] defined the layer 1 to layer 8 precoding matrix for Type I codebook. The number of precoding matrices in the set of $\mathcal{W}_l$ is denoted as $N_{\mathcal{W}_l}$. In this paper, we leverage

$$\mathbf{v}_{q_{t_1},n_{t_1}} = \mathbf{r}_{q_{t_1}} \times \mathbf{v}_{n_{t_1}} = \text{diag}\left(\left[1\ \text{e}^{\text{j}\frac{2\pi q_{t_1}\cdot 1}{N_{t_1}O_{t_1}}}\ \text{e}^{\text{j}\frac{2\pi q_{t_1}\cdot 2}{N_{t_1}O_{t_1}}}\ \cdots\ \text{e}^{\text{j}\frac{2\pi q_{t_1}\cdot(N_{t_1}-1)}{N_1O_{t_1}}}\right]\right) \times \left[1\ \text{e}^{\text{j}\frac{2\pi n_{t_1}\cdot 1}{N_{t_1}}}\ \text{e}^{\text{j}\frac{2\pi n_{t_1}\cdot 2}{N_{t_1}}}\ \cdots\ \text{e}^{\text{j}\frac{2\pi n_{t_1}\cdot(N_{t_1}-1)}{N_{t_1}}}\right]^{\text{T}}$$

$$= \left[1\ \text{e}^{\text{j}\frac{2\pi(n_{t_1}O_{t_1}+q_{t_1})\cdot 1}{N_{t_1}O_{t_1}}}\ \text{e}^{\text{j}\frac{2\pi(n_{t_1}O_{t_1}+q_{t_1})\cdot 2}{N_{t_1}O_{t_1}}}\ \cdots\ \text{e}^{\text{j}\frac{2\pi(n_{t_1}O_{t_1}+q_{t_1})\cdot(N_{t_1}-1)}{N_{t_1}O_{t_1}}}\right]^{\text{T}}, q_{t_1}=0,1,\cdots,O_{t_1}-1, n_{t_1}=0,1,\cdots,N_{t_1}-1. \tag{24}$$

the multi-agent mechanism that enables each DBS agent to independently generate its actions, thereby avoiding the exponential dimensionality $\mathcal{O}(N_{\mathcal{W}_l}^D)$ for the action space of D DBSs. The basic idea of precoding matrix component is to use an over-sampled rotation factor $\mathbf{r}$ with coefficient q to refine the DFT beam $\mathbf{v}$. We give an example of DFT beam $\mathbf{v}_{n_{t_1}} = \left[1\ \mathrm{e}^{\mathrm{j}\frac{2\pi n_{t_1}\cdot 1}{N_{t_1}}}\cdots \mathrm{e}^{\mathrm{j}\frac{2\pi n_{t_1}\cdot(N_{t_1}-1)}{N_{t_1}}}\right]^{\mathrm{T}} \in \mathbb{C}^{N_{t_1}\times 1}, n_{t_1} = 0,1,\cdots,N_{t_1}-1$ for N_{t_1} UPA horizonal antennas. The over-sampled value is denoted as O_{t_1} and the rotation factor $\mathbf{r}_{q_{t_1}} \in \mathbb{C}^{N_{t_1}\times N_{t_1}}$ with $q_{t_1} = 0,1,\cdots,O_{t_1}-1$ is given in Eq. (24), as shown at the bottom of the previous page. Based on the DFT beam basis, we provide the formulation of Type I codebook in Appendix B.

3) Reward: When the UE position p is obtained from the environment, we employ the encoder module of transformer to sequentially input the states of all subcarriers, i.e., $\mathbf{s}_p^k, \forall k \in \mathcal{K}$, which is indicated by $\mathbf{s}_\mathrm{p} = \{\mathbf{s}_\mathrm{p}^1, \cdots, \mathbf{s}_\mathrm{p}^K\}$. Subsequently, for a specific $\mathbf{s}_\mathrm{p}$, we define the reward function for action $\{\mathcal{A}_{l,d}\}_{d=1}^D$ as the sum of mutual information as

$$\gamma(\mathbf{s}_\mathrm{p}, \{\mathcal{A}_{l,d}\}_{d=1}^D) = \sum_{k=1}^{K}\sum_{i=1}^{I}\sum_{l=1}^{N_\ell} \log_2(1+\mathrm{SINR}_{k,l}^i). \quad (29)$$

Therefore, maximizing the reward will maximize the objective function (21a) when $N_f = 1$. Meanwhile, all DBS agents will use the higher tier action of transmit layer l to choose a precoding matrix from $\mathcal{W}_l$, thus ensuring compliance with constraints (21b) and (21d). The precoding matrices in Type I codebook are normalized with unit power, thereby guaranteeing the satisfaction of constraint (21e).

In the encoder module of the transformer, for transmit layer l, each DBS agent d will output $\mathcal{Q}_{\theta_{l,d}}(\mathbf{s}_\mathrm{p}, \mathcal{A}_{l,d})$ for for all $N_{\mathcal{W}_l}$ precoding matrices in precoding set $\mathcal{W}_l$ as

$$\mathcal{Q}_{\theta_{l,d}}(\mathbf{s}_\mathrm{p}, \mathcal{A}_{l,d}) = \left[\mathcal{Q}_{\theta_{l,d}}(\mathbf{s}_\mathrm{p}, \mathcal{A}_{l,d}^1), \cdots, \mathcal{Q}_{\theta_{l,d}}(\mathbf{s}_\mathrm{p}, \mathcal{A}_{l,d}^{N_{\mathcal{W}_l}})\right]^{\mathrm{T}} \quad (30)$$

where $\theta_{l,d}$ denotes the layer l network parameters of DBS agent d. To determine the higher tier action for each training epoch, we calculate the sum of the maximum $\mathcal{Q}_{\theta_{l,d}}(\mathbf{s}_\mathrm{p}, \mathcal{A}_{l,d})$ of all D agents, i.e., $\sum_{d=1}^D \max_{\mathcal{A}_{l,d}} \mathcal{Q}_{\theta_{l,d}}(\mathbf{s}_\mathrm{p}, \mathcal{A}_{l,d})$, which serves as a score indicator to identify the best transmit layer l. The higher tier action $\mathcal{A}_d^{high}$ is finally obtained with a probability of $1-\epsilon_{\mathcal{A}^{high}}$ to choose the best layer, and with a probability of $\epsilon_{\mathcal{A}^{high}}$ to randomly select from $\mathcal{L}$ as follows:

$$\mathcal{A}_d^{high} = \begin{cases} \arg\max\limits_{l\in\mathcal{L}} \left(\sum\limits_{d=1}^{D} \max\limits_{\mathcal{A}_{l,d}} \mathcal{Q}_{\theta_{l,d}}(\mathbf{s}_\mathrm{p}, \mathcal{A}_{l,d})\right), \\ \qquad\qquad with\ \ prob. = 1-\epsilon_{\mathcal{A}^{high}}, \\ l \in \mathcal{L}, \qquad with\ \ prob. = \epsilon_{\mathcal{A}^{high}}. \end{cases} \quad (31)$$

The probability $\epsilon_{\mathcal{A}^{high}}$ in Eq. (31) will linear decays with $\Delta\epsilon_{\mathcal{A}^{high}}$ to guarantee that a better transmit layer is selected during the training process. In the transformer encoder module, the scaled dot-product attention with multi-heads [43] is employed to capture the inherent correlations between the subcarriers of an OFDM symbol. The state of each subcarrier will be first embedded into a tensor of d_m features as follows:

$$\mathbf{s}_\mathrm{p}^{emb} = \mathrm{embedding}\left([\mathbf{s}_\mathrm{p}^1, \cdots, \mathbf{s}_\mathrm{p}^K]^{\mathrm{T}}\right), \quad (32)$$

through a fully connected (FC) neural network. Consider that the head number is N_H, we define the components of query $\boldsymbol{Q}_{l,d}^{h_i}$, key $\boldsymbol{K}_{l,d}^{h_i}$, and value $\boldsymbol{V}_{l,d}^{h_i}$ in the scaled dot-product attention for the h_i-th head of layer l and agent d as

$$\begin{cases} \boldsymbol{Q}_{l,d}^{h_i} = \mathbf{s}_\mathrm{p}^{emb}\mathbf{X}_{l,d}^{\boldsymbol{Q}_{h_i}} \\ \boldsymbol{K}_{l,d}^{h_i} = \mathbf{s}_\mathrm{p}^{emb}\mathbf{X}_{l,d}^{\boldsymbol{K}_{h_i}}, l\in\mathcal{L}, d\in\mathcal{D}, h_i\in[1,N_H], \\ \boldsymbol{V}_{l,d}^{h_i} = \mathbf{s}_\mathrm{p}^{emb}\mathbf{X}_{l,d}^{\boldsymbol{V}_{h_i}} \end{cases} \quad (33)$$

where $\mathbf{X}_{l,d}^{\{\boldsymbol{Q}/\boldsymbol{K}/\boldsymbol{V}\}_{h_i}}$ denotes the h_i-th head network parameter with $\frac{d_m}{N_H}$ features for the query, key, and value components, respectively. The output of the h_i-th attention is calculated by

$$\mathcal{T}_{l,d}^{h_i} = \mathrm{softmax}\left(\frac{\boldsymbol{Q}_{l,d}^{h_i}(\boldsymbol{K}_{l,d}^{h_i})^{\mathrm{T}}}{\sqrt{d_m/N_H}}\right)\boldsymbol{V}_{l,d}^{h_i}, \quad (34)$$

in which $\mathcal{T}_{l,d}^{h_i}$ possesses the feature number $\frac{d_m}{N_H}$ and operation $\mathrm{softmax}$ is to utilize the exponential function to normalize the output vector. The outputs of N_H heads will be concatenated into a tensor with d_m features, which will be fed forward through residual block, layer normalization, and FC layer to obtained $\mathcal{Q}_{\theta_{l,d}}(\mathbf{s}_\mathrm{p}, \mathcal{A}_{l,d})$ for precoding matrices in $\mathcal{W}_l$. For simplicity, we use $\mathcal{R}(\cdot)$ to indicate these processes as

$$\mathcal{Q}_{\theta_{l,d}}(\mathbf{s}_\mathrm{p}, \mathcal{A}_{l,d}) = \mathcal{R}\left([\mathcal{T}_{l,d}^1, \cdots, \mathcal{T}_{l,d}^{N_H}]^{\mathrm{T}}\right). \quad (35)$$

Algorithm 1 Transformer Encoder Based HRL Algorithm.

Input: Three-dimensional coordinates p_x, p_y, and p_z.
Output: The $\mathcal{Q}_{\theta_{l,d}}(\mathbf{s}_\mathrm{p}, \mathcal{A}_{l,d})$ values of precoding matrix.
for *Episode* $e = 1, 2, \cdots, N_E$ **do**
  Arbitrarily obtain a position p from the environment;
  Obtain $\mathbf{s}_\mathrm{p}^k$ with onehot encoding based on Eq. (27);
  Obtain the state of all subcarriers $\mathbf{s}_\mathrm{p} = \{\mathbf{s}_\mathrm{p}^1, \cdots, \mathbf{s}_\mathrm{p}^K\}$;
  if *prob.* $= 1-\epsilon_{\mathcal{A}^{high}}$ **then**
    $\mathcal{A}_d^{high} = \arg\max_{l\in\mathcal{L}} \left(\sum_{d=1}^D \max_{\mathcal{A}_{l,d}} \mathcal{Q}_{\theta_{l,d}}(\mathbf{s}_\mathrm{p}, \mathcal{A}_{l,d})\right)$;
  else
    $\mathcal{A}_d^{high} = l \in \{1,\ 2,\ \cdots, L\}$;
  end
  for *DBS agent* $\forall d \in \mathcal{D}$ **do**
    Compute $\mathcal{Q}_{\theta_{l,d}}(\mathbf{s}_\mathrm{p}, \mathcal{A}_{l,d})$ using Eqs. (32) - (35);
    Obtain current $\mathcal{A}_{l,d}$ based on ϵ-greedy strategy;
    Obtain $\gamma(\mathbf{s}_\mathrm{p}, \{\mathcal{A}_{l,d}\}_{d=1}^D)$ based on Eq. (29);
    Store $[\mathbf{s}_\mathrm{p}, \mathcal{A}_{l,d}, \gamma(\mathbf{s}_\mathrm{p}, \{\mathcal{A}_{l,d}\}_{d=1}^D)]$ in replay buffer;
    if $e \geq$ *batch size* N_e **then**
      Update l layer network using Eqs. (36)-(37)
    end
  end
  Update $\epsilon_{\mathcal{A}^{high}}$ with $\epsilon_{\mathcal{A}^{high}} \leftarrow \epsilon_{\mathcal{A}^{high}} - \Delta\epsilon_{\mathcal{A}^{high}}$
end
Utilization: Obtain optimal MIMO transmit layer l^* and precoding matrix $\mathbf{W}_{d,k}^*$ based on $\max_{\mathcal{A}_{l,d}} \mathcal{Q}_{\theta_{l,d}}(\mathbf{s}_\mathrm{p}, \mathcal{A}_{l,d})$.

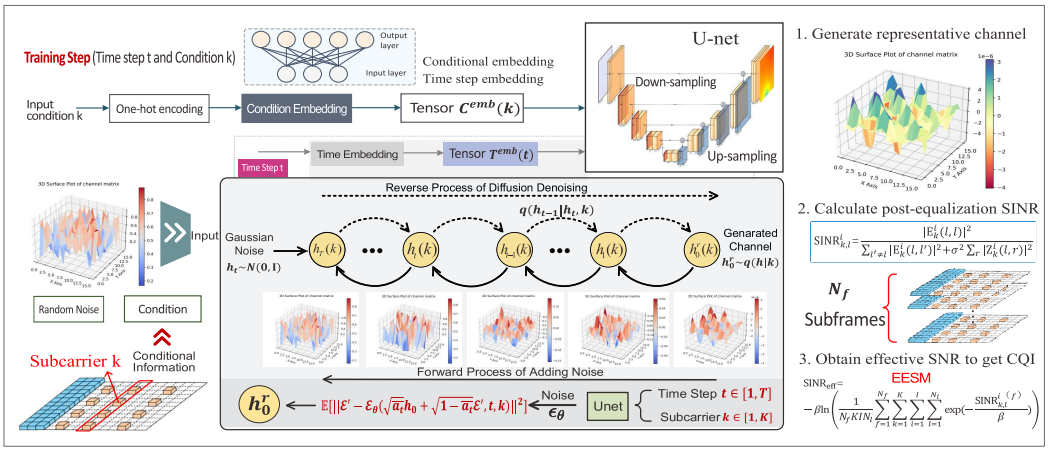


Fig. 3. Conditional diffusion model for representative channel generation with the subcarrier information.

Denote the number of training episode as N_E, for episode e, current action $\mathcal{A}_{l,d}^e$ will be selected by an ϵ-greedy strategy, i.e, with probability ϵ to choose an action obtained from $\max_{\mathcal{A}_{l,d}} \mathcal{Q}_{\theta_{l,d}}(\mathbf{s}_\mathrm{p}, \mathcal{A}_{l,d})$ and with probability $1-\epsilon$ to randomly choose any action. $\mathcal{A}_{l,d}^e$ will be taken into the calculation of the current reward $\gamma(\mathbf{s}_\mathrm{p}, \{\mathcal{A}_{l,d}^e\}_{d=1}^D)$ based on Eq. (29) and will be stored in a reply buffer for batch-size training. Let the number of batch-size be N_e, the loss function for the $\{[\mathbf{s}_\mathrm{e}, \mathcal{A}_{l,d}^e, \gamma(\mathbf{s}_\mathrm{p}, \{\mathcal{A}_{l,d}^e\}_{d=1}^D)]\}_{e=1}^{N_e}$ tuples is expressed as

$$\mathrm{L}(\theta_{l,d}) = \frac{1}{2N_e}\sum_{e=1}^{N_e}\left\|\mathcal{Q}_{\theta_{l,d}}(\mathbf{s}_e, \mathcal{A}_{l,d}^e)) - \gamma(\mathbf{s}_e, \{\mathcal{A}_{l,d}^e\}_{d=1}^D)\right\|_2. \tag{36}$$

According to the loss function, the l-th layer network parameters of agent $\forall d \in \mathcal{D}$ will be updated using gradient descend approach with a learning rate of α, which is formulated as

$$\theta_{l,d} \leftarrow \theta_{l,d} - \alpha\nabla_{\theta_{l,d}} L(\theta_{l,d}), \quad l \in \mathcal{L}, \ \forall d \in \mathcal{D}. \tag{37}$$

The detailed procedures of the proposed transformer encoder based HRL algorithm is given in **Algorithm 1**.

B. Conditional Diffusion Model for Channel Generation

The optimization problem Eq. (21) involves the calculation of mutual information over N_f subframe channels, which is computationally prohibitive for traditional model-based or rule-based methods. To adapt to the N_f time-varied channels, we employ conditional diffusion model to generate a representative channel for **Algorithm 1**, thus reducing the long-term computation overhead throughout the transmission process.

In the diffusion model depicted in Fig. 3, the original channel $\boldsymbol{h}_0$ is sampled from its data distribution $\boldsymbol{h}_0 \sim q(\boldsymbol{h}|k)$. The label of subcarrier k is used as the conditioning information to provide more fine-grained channel representation in frequency domain. A forward process wherein random noise will be gradually added to the time-varied channel data is modeled as a Markov chain and the noise is assumed to obey the Gaussian distribution with a variance of $b_\mathrm{t}, \mathrm{t} \in [0, T]$. Therefore, the forward diffusion process is defined as

$$\begin{aligned} q(\boldsymbol{h}_\mathrm{t}|\boldsymbol{h}_{\mathrm{t}-1}, k) &= q(\boldsymbol{h}_\mathrm{t}|\boldsymbol{h}_{\mathrm{t}-1}) \\ &= \mathcal{N}\left(\boldsymbol{h}_\mathrm{t}; \sqrt{1-b_\mathrm{t}}\boldsymbol{h}_{\mathrm{t}-1}, b_\mathrm{t}\mathbf{I}\right), \end{aligned} \tag{38}$$

where the condition k is proved to be an independent variable [44]. Consequently, the posterior probability from the original data $\boldsymbol{h}_0$ to the final $\boldsymbol{h}_\mathrm{T}$ can be formulated by

$$\begin{aligned} q(\boldsymbol{h}_{1:\mathrm{T}}|\boldsymbol{h}_0, k) &= \prod_{\mathrm{t}=1}^{T} q(\boldsymbol{h}_\mathrm{t}|\boldsymbol{h}_{\mathrm{t}-1}, k) \\ &= \prod_{\mathrm{t}=1}^{T} \mathcal{N}\left(\boldsymbol{h}_\mathrm{t}; \sqrt{1-b_\mathrm{t}}\boldsymbol{h}_{\mathrm{t}-1}, b_\mathrm{t}\mathbf{I}\right). \end{aligned} \tag{39}$$

The forward noise adding process can be directly obtained given an arbitrary time step t with the following closed-form

$$\begin{aligned} \boldsymbol{h}_\mathrm{t} &= \sqrt{1-b_\mathrm{t}}\boldsymbol{h}_{\mathrm{t}-1} + \sqrt{b_\mathrm{t}}\boldsymbol{\varepsilon}_{\mathrm{t}-1} \\ &\overset{a_\mathrm{t}:=1-b_\mathrm{t}}{=} \sqrt{a_\mathrm{t}}\boldsymbol{h}_{\mathrm{t}-1} + \sqrt{1-a_\mathrm{t}}\boldsymbol{\varepsilon}_{\mathrm{t}-1} \\ &\overset{(1)}{=} \sqrt{a_\mathrm{t}a_{\mathrm{t}-1}}\boldsymbol{h}_{\mathrm{t}-2} + \sqrt{1-a_\mathrm{t}a_{\mathrm{t}-1}}\boldsymbol{\varepsilon}_{\mathrm{t}-2} \\ &\overset{\bar{a}_\mathrm{t}:=\prod_{\mathrm{t}'=0}^{\mathrm{t}} a_{\mathrm{t}'}}{=} \sqrt{\bar{a}_\mathrm{t}}\boldsymbol{h}_0 + \sqrt{1-\bar{a}_\mathrm{t}}\boldsymbol{\varepsilon}_0, \end{aligned} \tag{40}$$

wherein $\boldsymbol{\varepsilon}_{\mathrm{t}=0:\mathrm{T}} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ and $\overset{(1)}{=}$ is obtained by

$$\sqrt{a_\mathrm{t} - a_\mathrm{t}a_{\mathrm{t}-1}}\boldsymbol{\varepsilon}_{\mathrm{t}-2} + \sqrt{1-a_\mathrm{t}}\boldsymbol{\varepsilon}_{\mathrm{t}-1} = \sqrt{1-a_\mathrm{t}a_{\mathrm{t}-1}}\boldsymbol{\varepsilon}_{\mathrm{t}-2}, \tag{41}$$

which is upheld based on the sum property of two Gaussian-distributed variables as

$$\mathcal{N}\left(0, \sigma_1^2\mathbf{I}\right) + \mathcal{N}\left(0, \sigma_2^2\mathbf{I}\right) \sim \mathcal{N}\left(0, \left(\sigma_1^2 + \sigma_2^2\right)\mathbf{I}\right). \tag{42}$$

Therefore, the channel $\boldsymbol{h}_\mathrm{t}$ with added noise at time step $t \in [1, \mathrm{T}]$ can be obtained by the following distribution

$$\begin{aligned} \boldsymbol{h}_\mathrm{t} \sim q(\boldsymbol{h}_\mathrm{t}|\boldsymbol{h}_0, k) &= q(\boldsymbol{h}_\mathrm{t}|\boldsymbol{h}_0) \\ &= \mathcal{N}\left(\boldsymbol{h}_\mathrm{t}; \sqrt{\bar{a}_\mathrm{t}}\boldsymbol{h}_0, (1-\bar{a}_t)\mathbf{I}\right). \end{aligned} \tag{43}$$

After T time steps of the noise addition process, the channel $\boldsymbol{h}_\mathrm{T}$ will follow an isotropically independent Gaussian distribution. In consequence, we can sample a data point from $\mathcal{N}(\mathbf{0}, \mathbf{I})$ and execute a reverse diffusion process to obtain a representative channel from the original channel data distribution as $\boldsymbol{h}_k^r \sim q(\boldsymbol{h}|k)$. The reverse process is tractable when $\boldsymbol{h}_0$ is known, which can be calculated as

$$\begin{aligned} q(\boldsymbol{h}_{\mathrm{t}-1}|\boldsymbol{h}_\mathrm{t}, \boldsymbol{h}_0, k) &= q(\boldsymbol{h}_{\mathrm{t}-1}|\boldsymbol{h}_\mathrm{t}, \boldsymbol{h}_0) \\ &\overset{(2)}{=} q(\boldsymbol{h}_\mathrm{t}|\boldsymbol{h}_{\mathrm{t}-1}) \frac{q(\boldsymbol{h}_{\mathrm{t}-1}|\boldsymbol{h}_0)}{q(\boldsymbol{h}_\mathrm{t}|\boldsymbol{h}_0)} \\ &\overset{(3)}{=} \mathcal{N}\left(\boldsymbol{h}_{\mathrm{t}-1}; \tilde{\boldsymbol{\mu}}_\mathrm{t}(\boldsymbol{h}_\mathrm{t}, \boldsymbol{h}_0, k), \tilde{b}_\mathrm{t}\mathbf{I}\right), \end{aligned} \tag{44}$$

Algorithm 2 Conditional Diffusion Algorithm.

Input: The original channel data $\boldsymbol{h}_0$ of subcarrier k.
Ouput: A representative channel $\boldsymbol{h}_k^r \sim q(\boldsymbol{h}|k)$.
repeat
  Obtain an original channel $\boldsymbol{h}_0 \sim q(\boldsymbol{h}|k)$,
  Obtain $\mathrm{t} \sim \mathrm{Uniform}(\{1, \ldots, \mathrm{T}\})$,
  Obtain Gaussian noise $\boldsymbol{\varepsilon}' \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$,
  Obtain $\boldsymbol{\varepsilon}_\theta\left(\sqrt{\bar{a}_\mathrm{t}}\boldsymbol{h}_0 + \sqrt{1-\bar{a}_\mathrm{t}}\boldsymbol{\varepsilon}', \mathrm{t}, k\right)$ through Unet,
  Take gradient descent step with $\nabla_\theta \left\|\boldsymbol{\varepsilon}' - \boldsymbol{\varepsilon}_\theta\left(\sqrt{\bar{a}_\mathrm{t}}\boldsymbol{h}_0 + \sqrt{1-\bar{a}_\mathrm{t}}\boldsymbol{\varepsilon}', \mathrm{t}, k\right)\right\|^2$
until *The noise predicted network* $\boldsymbol{\varepsilon}_\theta$ *converges*;
Obtain $\boldsymbol{h}_\mathrm{T} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$,
for *sampling step* $\mathrm{t} = \mathrm{T}, \cdots, 1$ **do**
  Obtain $\mathbf{z} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ if $\mathrm{t} > 1$, else $\mathbf{z} = \mathbf{0}$,
  Obtain $\boldsymbol{h}_{\mathrm{t}-1} = \frac{1}{\sqrt{a_\mathrm{t}}}\left(\boldsymbol{h}_\mathrm{t} - \frac{b_\mathrm{t}}{\sqrt{1-\bar{a}_\mathrm{t}}}\boldsymbol{\varepsilon}_\theta\left(\boldsymbol{h}_\mathrm{t}, \mathrm{t}, k\right)\right) + b_\mathrm{t}\mathbf{z}$
end
Utilization: Generate representative channel for $\forall k \in \mathcal{K}$.

wherein $\overset{(2)}{=}$ is based on the Bayes' rule and $\tilde{\boldsymbol{\mu}}_\mathrm{t}(\boldsymbol{h}_\mathrm{t}, \boldsymbol{h}_0)$ in $\overset{(3)}{=}$ is calculated by

$$\begin{aligned}\tilde{\boldsymbol{\mu}}_\mathrm{t}\left(\boldsymbol{h}_\mathrm{t}, \boldsymbol{h}_0, k\right) &= \frac{\sqrt{\bar{a}_{\mathrm{t}-1}}b_t}{1-\bar{a}_\mathrm{t}}\boldsymbol{h}_0 + \frac{\sqrt{a_\mathrm{t}}\left(1-\bar{a}_{\mathrm{t}-1}\right)}{1-\bar{a}_\mathrm{t}}\boldsymbol{h}_\mathrm{t} \\ &\overset{(4)}{=} \frac{1}{\sqrt{a_\mathrm{t}}}\left(\boldsymbol{h}_\mathrm{t} - \frac{b_\mathrm{t}}{\sqrt{1-\bar{a}_\mathrm{t}}}\boldsymbol{\varepsilon}\left(\boldsymbol{h}_\mathrm{t}, \mathrm{t}, k\right)\right), \end{aligned} \tag{45}$$

where $\overset{(4)}{=}$ is derived by reparameterizing $\boldsymbol{h}_\mathrm{t}$ using Eq. (43) as $\boldsymbol{h}_\mathrm{t}\left(\boldsymbol{h}_0, \boldsymbol{\varepsilon}\right) = \sqrt{\bar{a}_\mathrm{t}}\boldsymbol{h}_0 + \sqrt{1-\bar{a}_\mathrm{t}}\boldsymbol{\varepsilon}(\boldsymbol{h}_\mathrm{t}, \mathrm{t}, k)$ ($\boldsymbol{\varepsilon}(\boldsymbol{h}_\mathrm{t}, \mathrm{t}, k) \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ is simplified as $\boldsymbol{\varepsilon}$). Therefore, we have

$$\boldsymbol{h}_0 = \frac{1}{\sqrt{\bar{a}_\mathrm{t}}}\left(\boldsymbol{h}_\mathrm{t}\left(\boldsymbol{h}_0, \boldsymbol{\varepsilon}\right) - \sqrt{1-\bar{a}_\mathrm{t}}\boldsymbol{\varepsilon}\right), \tag{46}$$

and thus the result of $\tilde{\boldsymbol{\mu}}_\mathrm{t}(\boldsymbol{h}_\mathrm{t}, \boldsymbol{h}_0, k)$ in Eq. (45) can be obtained. Besides, $\tilde{b}_\mathrm{t}$ in Eq. (44) is formulated as

$$\tilde{b}_\mathrm{t} := \frac{1-\bar{a}_{\mathrm{t}-1}}{1-\bar{a}_\mathrm{t}}b_\mathrm{t} \approx b_\mathrm{t}. \tag{47}$$

The details for the derivation of $\overset{(3)}{=}$ can be found in [45]. As a consequence, the training process of diffusion is to exploit a parameterized model ρ_θ to approximate the reverse process

$$\rho_\theta\left(\boldsymbol{h}_{\mathrm{t}-1}|\boldsymbol{h}_\mathrm{t}, k\right) = \mathcal{N}\left(\boldsymbol{h}_{\mathrm{t}-1}; \boldsymbol{\mu}_\theta\left(\boldsymbol{h}_\mathrm{t}, \mathrm{t}, k\right), \sigma_\theta^2\left(\boldsymbol{h}_\mathrm{t}, \mathrm{t}, k\right)\mathbf{I}\right), \tag{48}$$

the mean value $\boldsymbol{\mu}_\theta\left(\boldsymbol{h}_\mathrm{t}, \mathrm{t}, k\right)$ of which is constructed as

$$\boldsymbol{\mu}_\theta\left(\boldsymbol{h}_\mathrm{t}, \mathrm{t}, k\right) = \frac{1}{\sqrt{a_\mathrm{t}}}\left(\boldsymbol{h}_\mathrm{t} - \frac{b_\mathrm{t}}{\sqrt{1-\bar{a}_\mathrm{t}}}\boldsymbol{\varepsilon}_\theta\left(\boldsymbol{h}_\mathrm{t}, \mathrm{t}, k\right)\right), \tag{49}$$

with $\boldsymbol{\varepsilon}_\theta(\boldsymbol{h}_\mathrm{t}, \mathrm{t}, k)$ denoting a Gaussian noise computed by the noise predicted network $\boldsymbol{\varepsilon}_\theta$. The variance of the predicted noise is assumed to possess a fixed value and is formulated as

$$\sigma_\theta^2\left(\boldsymbol{h}_\mathrm{t}, \mathrm{t}, k\right) = b_\mathrm{t}. \tag{50}$$

Therefore, the reverse process can be executed by sampling $\boldsymbol{h}_{\mathrm{t}-1} \sim \rho_\theta(\boldsymbol{h}_{\mathrm{t}-1}|\boldsymbol{h}_\mathrm{t})$ with the reparameterization technique as

$$\begin{aligned}\boldsymbol{h}_{\mathrm{t}-1} = &\frac{1}{\sqrt{a_\mathrm{t}}}\left(\boldsymbol{h}_\mathrm{t} - \frac{b_\mathrm{t}}{\sqrt{1-\bar{a}_\mathrm{t}}}\boldsymbol{\varepsilon}_\theta\left(\boldsymbol{h}_\mathrm{t}, \mathrm{t}, k\right)\right) \\ &+ b_\mathrm{t}\mathbf{z}, \quad \mathbf{z} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}). \end{aligned} \tag{51}$$

The optimization for the noise predicted network $\boldsymbol{\varepsilon}_\theta$ is to use the variational bound on the negative log likelihood of $\rho_\theta(\boldsymbol{h}_0|k)$ as the loss function

$$\mathbb{E}\left[-\log \rho_\theta\left(\boldsymbol{h}_0|k\right)\right] \leq \mathbb{E}_q\left[-\log \frac{\rho_\theta\left(\boldsymbol{h}_{\mathrm{T}:0}|k\right)}{q\left(\boldsymbol{h}_{\mathrm{T}:1}|\boldsymbol{h}_0, k\right)}\right] =: \mathrm{L}\left(\boldsymbol{\varepsilon}_\theta\right), \tag{52}$$

where $\rho_\theta(\boldsymbol{h}_{\mathrm{T}:0}|k)$ is defined as

$$\rho_\theta\left(\boldsymbol{h}_{\mathrm{T}:0}|k\right) := \rho_\theta\left(\boldsymbol{h}_\mathrm{T}|k\right)\prod_{\mathrm{t}=\mathrm{T}}^{1}\rho_\theta\left(\boldsymbol{h}_{\mathrm{t}-1}|\boldsymbol{h}_\mathrm{t}, k\right). \tag{53}$$

Through effective verification in [45], a simpler loss function based on a variant of the variational bound is given by

$$L'(\boldsymbol{\varepsilon}_\theta) := \mathbb{E}_{\mathrm{t}, k, \boldsymbol{h}_0, \boldsymbol{\varepsilon}'}\left[\left\|\boldsymbol{\varepsilon}' - \boldsymbol{\varepsilon}_\theta\left(\sqrt{\bar{a}_\mathrm{t}}\boldsymbol{h}_0 + \sqrt{1-\bar{a}_\mathrm{t}}\boldsymbol{\varepsilon}', \mathrm{t}, k\right)\right\|^2\right], \tag{54}$$

in which $\boldsymbol{\varepsilon}' \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ is the sampling Gaussian noise. The network of $\boldsymbol{\varepsilon}_\theta$ exploits a U-Net framework that consists of down-sampling $\mathcal{S}_d$ and up-sampling $\mathcal{S}_u$ processes to predict the noise. The down-sampling module leverages convolutional layers followed by max pooling operation to reduce the spatial dimension and increase the feature number. The up-sampling modules uses transposed convolutions to inversely increase the spatial dimension and reduce the feature representation. In U-Net, the label of conditional subcarrier k will first perform the Onehot encoding and be embedded with FC layers as

$$C^{emb}(k) = \mathrm{embedding}\left(\mathrm{Onehot}\left(k\right)\right), \quad \forall k \in \mathcal{K}. \tag{55}$$

The time step t will also be embedded to combine with $C^{emb}(k)$ for refining the output of the up-sampling process in U-Net, which is expressed as

$$\mathrm{T}^{emb}(\mathrm{t}) = \mathrm{embedding}\left(\mathrm{t}\right), \quad \forall \mathrm{t} \in \{1, \cdots, \mathrm{T}\}, \tag{56}$$

with the combination operation defined as

$$\mathcal{S}_u^{j+1} = \left(C^{emb}(k) \times \mathcal{S}_d^j + \mathrm{T}^{emb}(t)\right) \oplus \mathcal{S}_u^j, \tag{57}$$

where $\oplus$ indicates the concatenation operator and j is the j-th module in the U-net. Once the training of the noise predicted network converges $\boldsymbol{\varepsilon}_\theta$, we can generate representative channels for all subcarriers $\forall k \in \mathcal{K}$. Through conditioning on subcarrier k, the channel dimensionality is significantly reduced, which alleviates memory usage and speeds up calculations in each step, enabling a more efficient channel generation process for the diffusion model. The detailed procedures of the proposed conditional diffusion model is given in **Algorithm 2**.

Finally, we give the overall feedback-free mechanism for multi-DBS transmission in **Algorithm 3**. The fixed transmission parameters of PMI, RI, and CQI are obtained based on UE positions using **Algorithm 1** and **Algorithm 2**. The proposed conditional diffusion network will first generate a representative channel matrix from N_f time-varied channels. The N_f channels are no needed to be stored after the diffusion model has been trained to convergence, thereby avoiding excessive consumption of cache resources. Subsequently, the generated representative channel is employed to attain the transmit layer (RI) and precoding matrix (PMI)

Algorithm 3 Feedback-Free Transmission Mechanism.

Input: The original channel data and UE position.
Ouput: The fixed physical layer transmission parameters.
for *Each position* p **do**
  Generate a representative channel $\mathbf{H}_k^r, \forall k \in \mathcal{K}$ based on **Algorithm 2**,
  Use the representative channel in **Algorithm 1** to obtain transmit layer N_l and precoding matrix $\mathbf{W}_{d,k}$,
  Calculate the post-equalized SINR based on Eq. (15),
  Calculate $\mathrm{SNR}_{\mathrm{eff}}$ using EESM method with $\mathrm{SNR}_{\mathrm{eff}} = \beta \mathrm{f}^{-1}\left(\frac{1}{N^{\mathrm{pe}}}\sum_{\mathrm{pe}=1}^{N^{\mathrm{pe}}} \mathrm{f}\left(\frac{\mathrm{SINR}^{\mathrm{pe}}}{\beta}\right)\right)$,
  $\mathrm{Mapping\_Table}\left(\mathrm{SNR}_{\mathrm{eff}}\right) \longrightarrow \mathrm{CQI}$
end
Utilization: Fix PMI, RI, and CQI for all time periods.

TABLE I
PARAMETERS SETTINGS

Parameters	Values	Parameters	Values
DBS transmit antenna N_t	16	UE receive antenna N_r	4
DBS dual-polarized	8×1	UE single-polarization	4×1
Number of DBS D	2	Number of UE position	1138
Number of subcarrier K	72	Number of OFDM symbol	14
Center frequency	3.5 GHz	Coding scheme	$LDPC$
Subframe number N_f	100	CQI value range	1 - 15
Subframe duration	1 ms	Transmit layer N_l	1 - 4

through the transformer encoder based HRL framework, which will be used to calculated the post-equalization SINR with Eq. (15). By leveraging the EESM method, the post-equalization SINR of N_f subframes is equivalent to a value of $\mathrm{SNR}_{\mathrm{eff}}$, through which we can obtain the CQI value from a lookup $\mathrm{Mapping\_Table}$ as expressed in Eq. (23).

V. SIMULATION RESULTS

In this section, we conduct simulations based on the 5G Link Level Simulator [22] to evaluate the performance of the proposed transformed encoder based HRL algorithm and the conditional diffusion model. We use the Outdoor 1 scenario provided by DeepMIMO [46] as shown in Fig. 4. The 3D ray-tracing technology and CDL channel model in MATLAB are utilized to simulate the time-varied channels operating at 3.5 GHz center frequency. We consider a simulation setting consisting of 2 DBS and 1138 UE positions, the position coordinates of which are illustrated in Fig. 5. The DBS equips 8×1 dual-polarized transmit antennas with $N_t = 16$, and a UE has 4×1 singel-polarized receive antennas with $N_r = 4$. A transmission subframe includs 72 subcarriers and 14 OFDM symbols, with a duration of 0.1 ms. We utilize 100 subframes to simulate the time-varied channels, which possess a total time period of 10 ms and are assumed to approximate a wide-range channel variation. Besides, this paper employs LDPC channel coding and CQI values ranging from 1 to 15, covering code efficiencies from QPSK at 0.1523 to 64QAM at 5.5547. Based on the complete physical-layer module, we can obtain the UE throughput with the amount of successfully transmitted data per unit of time, measured in megabits per second (Mbit/s) within this paper. The calculation for the UE throughput is associated with the bit error rate (BER), which can be expressed as $Throughput = (1 - BER) \times$ *number of transmitted bits per second*. The detailed parameter settings are presented in Table I.

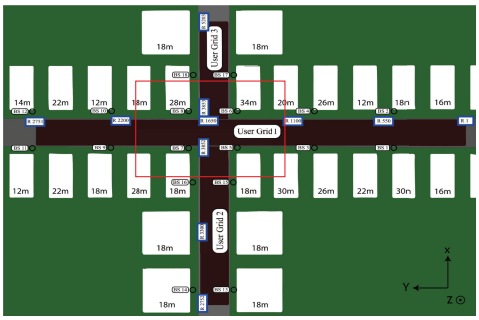


Fig. 4. DeepMIMO 3.5 GHz 3D ray-tracing scenario.

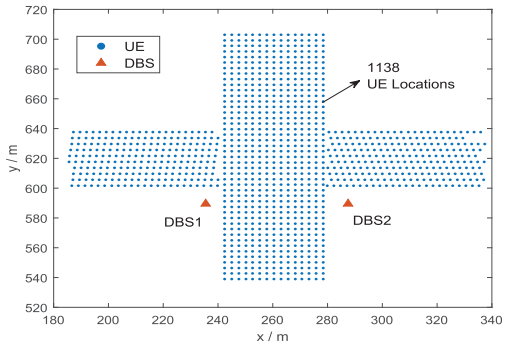


Fig. 5. The coordinates of 2 DBS and 1138 UE.

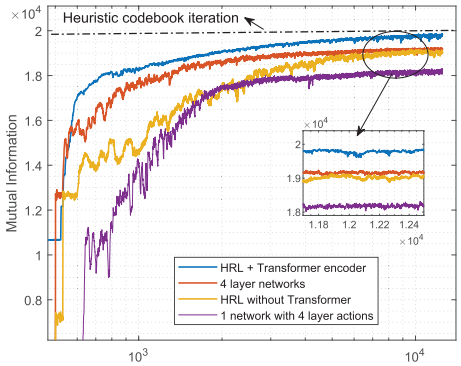


Fig. 6. The training process of HRL.

A. UE Throughput Performance in 1-Th Subframe

We first validate the performance of transformer encoder based HRL (T-HRL) algorithm in one specific subframe. Fig.6 shows the training process of precoding matrix selection for 2 DBSs. Since the minimum transmit and receive antennas is 4, the training action will select the precoding matrix from 1 to 4 layer precoding set $\{\mathcal{W}_l\}_{l=1}^4$. If using a centralized network to output the $\mathcal{Q}$-values for all 4 layer actions, the large action size will make the network difficult to converge thus resulting in the worst training performance. The proposed HRL outperforms the 1 network training but is slightly worse than using 4 independent networks for the 4 layer precoding actions. This is due to the essential property that HRL algorithm is a trade-off between the 1 network and 4 networks training. For the proposed T-HRL algorithm, transformer encoder sequentially outputs the actions for all subcarriers within an OFDM symbol, which is a more fine-grained subband transmission scheme compared with the other

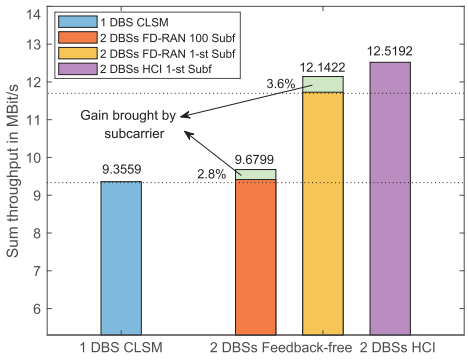


Fig. 7. Throughput of the 1-st subframe.

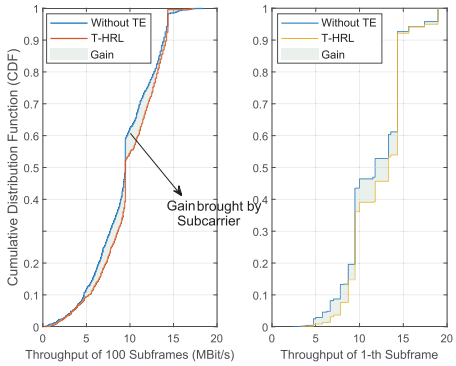


Fig. 8. CDF plot for the 1-st subframe.

three training mechanisms. Therefore, the T-HRL possesses the best training performance.

Based on the training results of the T-HRL algorithm, we obtain the PMI, RI, and CQI in one subframe, which is used to validate the feedback-free mechanism of FD-RAN. A heuristic codebook iteration (HCL) method that traverses all precoding matrices with a linear complexity is used as the performance benchmark. HCI can be looked upon as a feedback mechanism, thereby possessing the best performance. Note that the detailed process of the HCI method is provided in Appendix C. Another comparative method is the CLSM mechanism, which is currently employed as an optimal transmission scheme in 5G wireless networks. However, its application to multi-BS transmissions is computationally prohibitive due to its exponential complexity in iterating the codebook for all BSs. As shown in Fig.7, the feedback-free mechanism in FD-RAN, when utilizing the T-HRL training, achieves up to 97% UE throughput of the HCI method and significantly enhances the performance of the CLSM transmission. This result demonstrates the efficiency of feedback-free transmission within a single subframe. Furthermore, we present the performance gain for the fine-grained subband transmission scheme, enabled by the transformer encoder, with the sum mean UE throughput in Fig.7 and the cumulative distribution function (CDF) in Fig.8. Nevertheless, when utilizing the PMI, RI, and CQI calculated in one subframe consistently across all 100 subframe transmission periods, the 2 DBSs feedback-free mechanism possesses almost the same throughput as 1 DBS CLSM mechanism, which indicates a need for physical-layer parameter optimization in time domain.

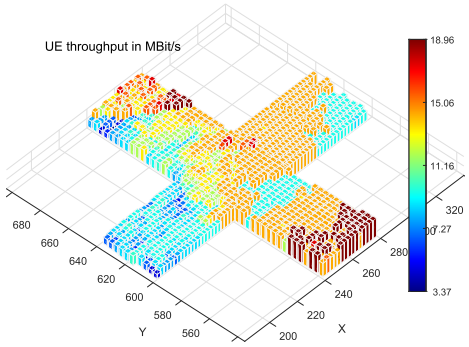


Fig. 9. The heatmap of UE throughput based on location.

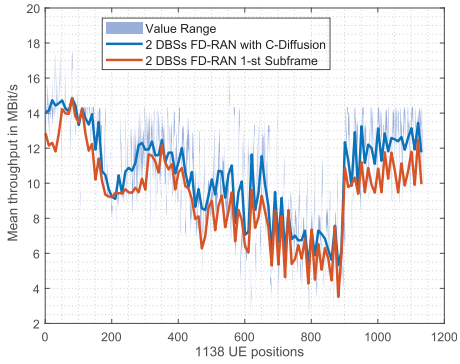


Fig. 10. Throughput enhancement based on 1-st subframe parameters.

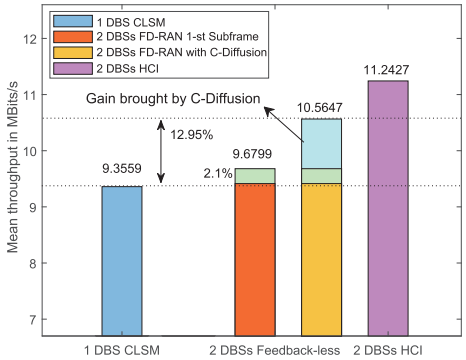


Fig. 11. Comparison of the throughput between FD-RAN, CLSM, and HCI.

B. Performance Over 100 Subframes

We further evaluate the performance of conditional diffusion model in adapting to dynamically changing channels. The U-net structure down-samples the channel matrices in the process of $(16, 16) \rightarrow (7, 7) \rightarrow (1, 1)$, where $(16, 16)$ is obtained by 2 DBSs, 16 transmit antennas, 4 receive antennas, and 2 elements of the complex channel values. We sample a representative channel generated by the conditional diffusion model based on the original 100 subframe channels. Fig. 9 illustrates the heatmap of the UE throughput at 1138 locations using the PMI, RI, and CQI calculated by the representative channel. The heatmap roughly demonstrates a certain correlation between the throughput and the geographical location. Fig. 10 presents the performance comparison with the transmission using the parameters calculated in one subframe. To better elaborate the benefits brought by diffusion model, we give the sum mean throughput of the 1138 UE poistions in Fig. 11. The simulation result shows that the 2 DBSs feedback-free transmission based on the generated representative channel

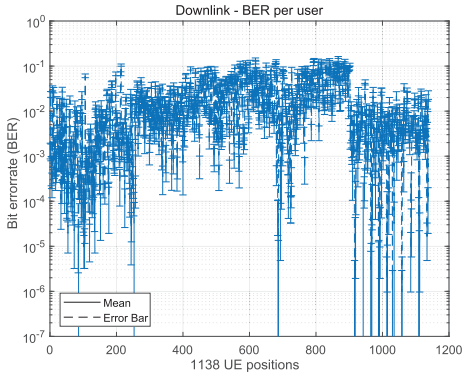


Fig. 12. The bit error rate of 1138 UE positions in FD-RAN.

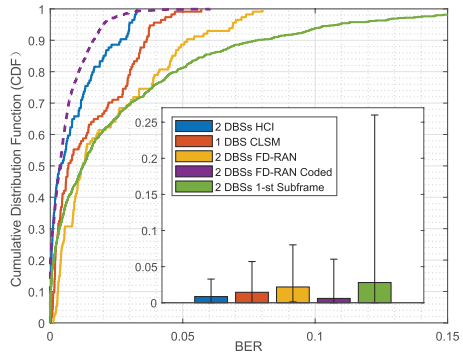


Fig. 13. The CDF and mean value of bit error rate in FD-RAN, CLSM, and HCI.

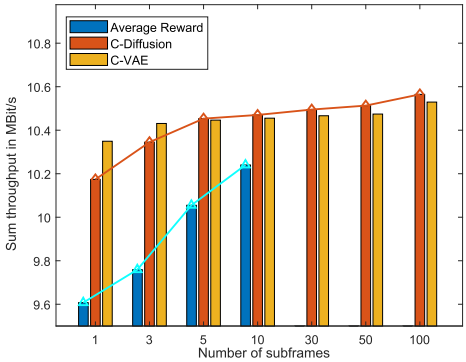


Fig. 14. UE throughput of GAI methods with different subframe numbers.

can enhance the UE throughput by 13% compared to 1 DBS CLSM and reach 94% performance of the HCI method.

After performing a complete 2 DBSs coordinated physical-layer transmission, the BERs of the feedback-free mechanism at the 1138 UE locations are given in Fig. 12. Most of the BERs are distributed between 10^{-2} and 10^{-4}, which is an acceptable range for a wireless communication system. Besides, the CDF and mean BER values are depicted in Fig. 13. The 2 DBSs HCI and 1 DBS CLSM mechanism possess the minimum BER values since they are operated on a feedback mechanism. The BER of the feedback-free mechanism has been significantly reduced after channel decoding and is less than the transmission with one subframe parameters, which further validates the efficiency of the diffusion model.

Finally, we compare the feedback-free transmission performance with an averaged reward method and a conditional VAE approach in [19]. The detailed procedures of VAE for representative channel generation are provided in Appendix D. The averaged reward method uses the average N_f frame mutual information as the reward in the T-HRL training to adapt to N_f subframe channels. The simulation results in Fig. 14 validate the effectiveness of this method, however, it can only be used within 10 subframe due to the exhausted computation capacity in the fine-grained subcarrier transmission. The conditional VAE can find a best channel representation in the latent space, thus performing better than the conditional diffusion model in the case of only one generated channel. As the number of generated channel increases, the conditional diffusion model possesses the highest UE throughput, which demonstrates extensive capability in reconstructing the representative channel for complicated time-varying channels.

VI. Conclusion

In this paper, we have investigated the feedback-free mechanism in FD-RAN for coordinated multi-BS transmission through UE geolocation. The optimization for fixed physical-layer transmission parameters have been formulated to maximize the UE throughput within the spatial, frequency, and time domains, respectively. Considering the stochastic and unpredictable channel variations of wireless network, generative AI technique has been employed to create feasible resolutions for modulation and coding scheme under the dynamic time-varied environment. Simulation results based on a versatile link-level simulator have demonstrated the significant effects of the feedback-free mechanism in facilitating the efficient and flexible transmission of multi-BS coordination, which offers a guiding principle for the design of next-generation transceiver technologies. In our future work, we will further investigate the implementations of both single BS and multi-BS connection to improve the overall network performance.

Appendix A
Derivation of Equation (4)

We first prove the following equation for transmit antenna t, once the signals from different DBSs are looked as multipaths and the ISI effect is eliminated during the CP duration

$$y^i_{r,t}[k] = \sum_{d\in\mathcal{D}} h^i_{d,r,t}[k]x^i[k] + \mathrm{n}^i[k],\ 0 \le k \le N_s - 1. \quad (58)$$

By substituting $y^i_{r,t}(\tau)$, $h^i_{d,r,t}(\tau)$, $x^i(\tau)$ and $\mathrm{n}^i(\tau)$ with discrete time sequences $y^i_{r,t}[n]$, $h^i_{d,r,t}[n]$, $x^i[n]$, and $\mathrm{n}^i[n]$, Eq. (3) is transformed into

$$y^i_{r,t}[n] = \sum_{d\in\mathcal{D}} \alpha_d h^i_{d,r,t}[n] * x^i[n] + \mathrm{n}^i[n], \quad (59)$$

in which $0 \le n \le N_s - 1$. Omit the pathloss α_d and perform $\mathrm{DFT}\{y^i_{r,t}[n]\}$, then we can obtain $y^i_{r,t}[k]$ for subcarrier k

$$\begin{aligned} y^i_{r,t}[k] &= \sum_{n=0}^{N_s-1} y^i_{r,t}[n]\mathrm{e}^{-\mathrm{j}\frac{2\pi}{N_s}nk} \\ &= \sum_{n=0}^{N_s-1}\left\{\sum_{m=0}^{\infty}\sum_{d=1}^{D} h^i_{d,r,t}[m]x^i[n-m] + \mathrm{n}^i[n]\right\} \\ &\quad \mathrm{e}^{-\mathrm{j}\frac{2\pi}{N_s}nk} \\ &= \left\{\sum_{m=0}^{\infty}\sum_{d=1}^{D} h^i_{d,r,t}[m]\sum_{n=0}^{N_s-1} x^i[n-m]\right\}\mathrm{e}^{-\mathrm{j}\frac{2\pi}{N_s}nk} \\ &\quad + \mathrm{n}^i[k] \end{aligned}$$

$$= \sum_{m=0}^{\infty}\sum_{d=1}^{D} h_{d,r,t}^{i}[m]\mathrm{e}^{-\mathrm{j}\frac{2\pi}{N_s}mk}\sum_{n=0}^{N_s-1} x^{i}[n-m] \mathrm{e}^{-\mathrm{j}\frac{2\pi}{N_s}(n-m)k} + \mathrm{n}^{i}[k]. \tag{60}$$

Let $j = n - m$, we have

$$\sum_{n=0}^{N_s-1} x^{i}[n-m]\mathrm{e}^{-\mathrm{j}\frac{2\pi}{N_s}(n-m)k} = \sum_{j=-m}^{N_s-1-m} x^{i}[j]\mathrm{e}^{-\mathrm{j}\frac{2\pi}{N_s}jk}. \tag{61}$$

Since CP is a copy of the end part of an OFDM symbol, $x^{i}[j]\mathrm{e}^{-\mathrm{j}\frac{2\pi}{N_s}jk}$ is periodic with the period of N_s once m is less than the CP duration, which means that the delay of different channel paths are less than CP duration. Based on the property of the periodic sequence, we have the following expression

$$\sum_{j=-m}^{N_s-1-m} x^{i}[j]\mathrm{e}^{-\mathrm{j}\frac{2\pi}{N_s}jk} = \sum_{j=0}^{N_s-1} x^{i}[j]\mathrm{e}^{-\mathrm{j}\frac{2\pi}{N_s}jk} = x^{i}[k]. \tag{62}$$

Consequently, Eq. (60) can be converted to

$$\begin{aligned} y_{r,t}^{i}[k] &= \sum_{m=0}^{\infty}\sum_{d=1}^{D} h_{d,r,t}^{i}[m]\mathrm{e}^{-\mathrm{j}\frac{2\pi}{N_s}mk}x^{i}[k] + \mathrm{n}^{i}[k] \\ &= \sum_{d=1}^{D}\sum_{m=0}^{N_s-1} h_{d,r,t}^{i}[m]\mathrm{e}^{-\mathrm{j}\frac{2\pi}{N_s}mk}x^{i}[k] + \mathrm{n}^{i}[k] \\ &= \sum_{d\in\mathcal{D}} h_{d,r,t}^{i}[k]x^{i}[k] + \mathrm{n}^{i}[k]. \end{aligned} \tag{63}$$

Finally, by the sum of discrete signals from all transmit antennas, we can obtain the received signal $y_{r}^{i}[k] = \sum_{t=1}^{N_t} y_{r,t}^{i}[k]$ in frequency domain for subcarrier k and received antenna r.

APPENDIX B
TYPE I CODEBOOK MATRIX

Let $n_1 = n_{t_1}O_{t_1} + q_{t_1}$, DFT beam $\mathbf{v}_{q_{t_1},n_{t_1}}$ of the N_{t_1} horizontal antennas in Eq. (24) can be transformed into

$$\mathbf{v}_{n_1} = \left[1\ \mathrm{e}^{\mathrm{j}\frac{2\pi n_1\cdot 1}{N_{t_1}O_{t_1}}}\ \mathrm{e}^{\mathrm{j}\frac{2\pi n_1\cdot 2}{N_{t_1}O_{t_1}}}\ \cdots\ \mathrm{e}^{\mathrm{j}\frac{2\pi n_1\cdot(N_{t_1}-1)}{N_{t_1}O_{t_1}}}\right]^{\mathrm{T}}, \quad n_1 = 0, 1, \cdots, O_{t_1}N_{t_1} - 1. \tag{64}$$

We can obtain the DFT beam $\mathbf{v}_{n_2}, n_2 = 0, 1, \cdots, O_{t_2}N_{t_2} - 1$ in the same form for the N_{t_2} vertical antennas with an over-sampled value O_{t_2}. A precoding matrix $\mathbf{W}$ is basically composed by wide-band $\mathbf{W^1}$ and sub-band $\mathbf{W^2}$ properties as

$$\mathbf{W} = \mathbf{W^1}\mathbf{W^2} = \begin{bmatrix} \mathcal{B}_{n_1,n_2} & 0 \\ 0 & \mathcal{B}_{n_1,n_2} \end{bmatrix} \mathbf{W^2}, \tag{65}$$

where $\mathcal{B}_{n_1,n_2}$ denotes the Kronecker product of $\mathbf{v}_{n_1}$ and $\mathbf{v}_{n_2}$

$$\mathcal{B}_{n_1,n_2} = \mathbf{v}_{n_1} \otimes \mathbf{v}_{n_2}, \tag{66}$$

and $\mathbf{W^2}$ is designed for dynamic beam selection under different antenna polarizations. For layer 1 precoding matrices, we have $\mathbf{W^2} = \frac{1}{\sqrt{2N_{t_1}N_{t_2}}}[\mathbf{e}, \varphi_{n_3}\mathbf{e}]^{\mathrm{T}}$ with co-phasing φ_{n_3} as

$$\mathbf{W}_{n_1,n_2,n_3}^{(1)} = \frac{1}{\sqrt{2N_{t_1}N_{t_2}}}\begin{bmatrix} \mathcal{B}_{n_1,n_2} \\ \varphi_{n_3}\mathcal{B}_{n_1,n_2} \end{bmatrix}, \tag{67}$$

$$\varphi_{n_3} = e^{j\pi\frac{n_3}{2}}, \quad n_3 = 0,\ 1,\ 2,\ 3. \tag{68}$$

We also give the expression of layer 4 precoding matrix as follows. The details can be obtained in 3GPP 38.214 [42].

$$\mathbf{W}_{n_1,n_1',n_2,n_2',n_3}^{(4)} = \frac{1}{\sqrt{4\cdot 2N_{t_1}N_{t_2}}} \begin{bmatrix} \mathcal{B}_{n_1,n_2} & \mathcal{B}_{n_1',n_2'} & \mathcal{B}_{n_1,n_2} & \mathcal{B}_{n_1',n_2'} \\ \varphi_{n_3}\mathcal{B}_{n_1,n_2} & \varphi_{n_3}\mathcal{B}_{n_1',n_2'} & -\varphi_{n_3}\mathcal{B}_{n_1,n_2} & -\varphi_{n_3}\mathcal{B}_{n_1',n_2'} \end{bmatrix}. \tag{69}$$

APPENDIX C
HEURISTIC CODEBOOK ITERATION METHOD

The HCI method first uses a precoding matrix of DN_t transmit antennas $\mathbf{W}_k' \in \mathbb{C}^{DN_t}$ to obtain the physical-layer transmission parameters. Then, Eq. (6) can be reformulated as

$$\mathbf{y}_k = [\mathbf{H}_{1,k},\ \cdots,\ \mathbf{H}_{D,k}] \underbrace{\begin{bmatrix} \mathbf{W}_{1,k} \\ \vdots \\ \mathbf{W}_{D,k} \end{bmatrix}}_{\mathbf{W}_k' \in \mathbb{C}^{DN_t\times N_l}} \mathbf{x}_k + \mathbf{n}_k. \tag{70}$$

Subsequently, we can obtain an optimal precoding matrix $\mathbf{W}_k'^{*}$ by maximizing the following mutual information

$$\max_{\mathbf{W}_k'\in\mathcal{W}^{DN_t}} \mathbb{E}\left[\sum_{k=1}^{K}\sum_{i=1}^{I} M_k^{i}\left(\mathbf{W}_k'\right)\right], \tag{71}$$

and the transmit layer for all DBS $d \in \mathcal{D}$ is attained by

$$N_l{}^{*} = \mathrm{rank}\left(\mathbf{W}_k'^{*}\right). \tag{72}$$

However, we cannot direct obtain $\mathbf{W}_{d,k}, \forall d \in \mathcal{D}$ since $\mathbf{W}_{d,k}$ belongs to the precoding set of N_t transmit antennas, which is different from $\mathbf{W}_k' \in \mathbb{C}^{DN_t}$. Therefore, we propose the following heuristic codebook iteration procedure for obtain the optimal precoding matrix of $\mathbf{W}_{d,k}^{*}$

$$\mathbf{H}_{1,k} \xrightarrow[HCI]{} \mathbf{W}_{1,k}^{*} \xrightarrow[HCI]{[\mathbf{H}_{1,k},\mathbf{H}_{2,k}];\mathbf{W}_{1,k}^{*}} \mathbf{W}_{2,k}^{*} \rightarrow \cdots \xrightarrow[HCI]{[\mathbf{H}_{1,k},\cdots,\mathbf{H}_{D,k}];\mathbf{W}_{1,k}^{*},\cdots,\mathbf{W}_{(D-1),k}^{*}} \mathbf{W}_{D,k}^{*} \Longrightarrow \{\mathbf{W}_{d,k}{}^{*}\}_{d=1}^{D}, \tag{73}$$

where the operation $\xrightarrow[HCI]{}$ can be expressed as

$$\mathbf{W}_{d,k}^{*} = \max_{\mathbf{W}_{d,k}} \mathbb{E}\left[\sum_{k=1}^{K}\sum_{i=1}^{I} M_k^{i}\left(\mathbf{W}_{1,k},\cdots,\mathbf{W}_{d-1,k}\right)\right]. \tag{74}$$

Finally, the CQI is obtained based on Eq. (22) - Eq. (23).

APPENDIX D
CONDITIONAL VAE FOR CHANNEL GENERATION

The conditional VAE first compresses the original N_f time-varying channels into N_f latent variables using an encoder module. These latent variables are assumed to follow Gaussian distributions with their mean values and variances denoted by $\boldsymbol{\mu}_f$ and $\boldsymbol{\sigma}_f^2$, $f = 1, \cdots, N_f$, respectively. Subsequently, the conditional VAE will exploit a decoder module to recover N_f channel matrices based on the N_f latent variables. With regard to the objective to generate a representative channel, a representative latent variable is selected to be the input of

the VAE decoder module. The representative latent variable can be decided with its mean value and variance closest to the average mean value $\overline{\boldsymbol{\mu}}$ and variance $\overline{\boldsymbol{\sigma}}^2$ over N_f latent variables, which is expressed by

$$\underset{f\in\{1,\cdots,N_F\}}{\arg\min} \Big(\left(\boldsymbol{\mu}_f - \overline{\boldsymbol{\mu}}\right)^{\mathrm{T}} \left(\boldsymbol{\mu}_f - \overline{\boldsymbol{\mu}}\right) + \left(\boldsymbol{\sigma}_f^2 - \overline{\boldsymbol{\sigma}}^2\right)^{\mathrm{T}} \left(\boldsymbol{\sigma}_f^2 - \overline{\boldsymbol{\sigma}}^2\right) \Big). \quad (75)$$

Consequently, a representative latent variable can be sampled based on the selected mean value and variance to generate a representative channel by the VAE decoder for the original N_f time-varying channels.

REFERENCES

[1] Q. Wu, S. Zhang, B. Zheng, C. You, and R. Zhang, "Intelligent reflecting surface-aided wireless communications: A tutorial," *IEEE Trans. Commun.*, vol. 69, no. 5, pp. 3313–3351, May 2021.

[2] J. Zhao et al., "Fully-decoupled radio access networks: A resilient uplink base stations cooperative reception framework," *IEEE Trans. Wireless Commun.*, vol. 22, no. 8, pp. 5096–5110, Aug. 2023.

[3] J. Zheng, J. Zhang, E. Björnson, and B. Ai, "Impact of channel aging on cell-free massive MIMO over spatially correlated channels," *IEEE Trans. Wireless Commun.*, vol. 20, no. 10, pp. 6451–6466, Oct. 2021.

[4] J. Xue, K. Yu, T. Zhang, H. Zhou, L. Zhao, and X. Shen, "Cooperative deep reinforcement learning enabled power allocation for packet duplication URLLC in multi-connectivity vehicular networks," *IEEE Trans. Mobile Comput.*, vol. 23, no. 8, pp. 8143–8157, Aug. 2024.

[5] R. Zhang, K. Xiong, Y. Lu, B. Gao, P. Fan, and K. B. Letaief, "Joint coordinated beamforming and power splitting ratio optimization in MU-MISO SWIPT-enabled HetNets: A multi-agent DDQN-based approach," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 2, pp. 677–693, Feb. 2022.

[6] B. Qian et al., "Enabling fully-decoupled radio access with elastic resource allocation," *IEEE Trans. Cognit. Commun. Netw.*, vol. 9, no. 4, pp. 1025–1040, Aug. 2023.

[7] *NR; Multiplexing and Channel Coding*, Standard (TS) 38.212, Version 17.2.0, 3GPP, 3rd Generation Partnership Project, 2022.

[8] C. You, B. Zheng, and R. Zhang, "Channel estimation and passive beamforming for intelligent reflecting surface: Discrete phase shift and progressive refinement," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2604–2620, Nov. 2020.

[9] S. Chen, J. Zhang, E. Björnson, Ö. T. Demir, and B. Ai, "Energy-efficient cell-free massive MIMO through sparse large-scale fading processing," *IEEE Trans. Wireless Commun.*, vol. 22, no. 12, pp. 9374–9389, Dec. 2023.

[10] M. Silva, L. Ramalho, I. Almeida, E. Medeiros, and A. Klautau, "CSI compression for distributed-MIMO with centralized precoding and power allocation," *IEEE Commun. Lett.*, vol. 27, no. 6, pp. 1535–1539, Jun. 2023.

[11] R. Zhang, K. Xiong, Y. Lu, D. W. K. Ng, P. Fan, and K. B. Letaief, "SWIPT-enabled cell-free massive MIMO-NOMA networks: A machine learning-based approach," *IEEE Trans. Wireless Commun.*, vol. 23, no. 7, pp. 6701–6718, Jul. 2024.

[12] Q. Yu et al., "A fully-decoupled RAN architecture for 6G inspired by neurotransmission," *J. Commun. Inf. Netw.*, vol. 4, no. 4, pp. 15–23, Dec. 2019.

[13] T. Zhang et al., "Handover-free multi-connectivity mobility management for downlink FD-RAN: A hierarchical DRL based approach," *IEEE Trans. Cognit. Commun. Netw.*, early access, Aug. 30, 2024, doi: 10.1109/TCCN.2024.3452639.

[14] Z. Liu, M. Liwang, S. Hosseinalipour, H. Dai, Z. Gao, and L. Huang, "RFID: Towards low latency and reliable DAG task scheduling over dynamic vehicular clouds," *IEEE Trans. Veh. Technol.*, vol. 72, no. 9, pp. 12139–12153, Sep. 2023.

[15] Y. Xu, B. Qian, K. Yu, T. Ma, L. Zhao, and H. Zhou, "Federated learning over fully-decoupled RAN architecture for two-tier computing acceleration," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 3, pp. 789–801, Mar. 2023.

[16] C. You and R. Zhang, "Hybrid offline-online design for UAV-enabled data harvesting in probabilistic LoS channels," *IEEE Trans. Wireless Commun.*, vol. 19, no. 6, pp. 3753–3768, Jun. 2020.

[17] Z. Liu et al., "Leveraging self-supervised learning for MIMO-OFDM channel representation and generation," 2024, *arXiv:2407.07702*.

[18] J. Wang et al., "A unified framework for guiding generative AI with wireless perception in resource constrained mobile edge networks," *IEEE Trans. Mobile Comput.*, vol. 23, no. 11, pp. 10344–10360, Nov. 2024.

[19] J. Liu, J. Chen, Z. Liu, and H. Zhou, "Enabling feedback-free MIMO transmission for FD-RAN: A data-driven approach," 2023, *arXiv:2311.14329*.

[20] X. Hou, J. Wang, C. Jiang, Z. Meng, J. Chen, and Y. Ren, "Efficient federated learning for metaverse via dynamic user selection, gradient quantization and resource allocation," *IEEE J. Sel. Areas Commun.*, vol. 42, no. 4, pp. 850–866, Apr. 2024.

[21] *Evolved Universal Terrestrial Radio Access (E-UTRA); Physical Layer Procedures*, Standard 38.212, Version 17.2.0, 3GPP, 3rd Generation Partnership Project, 2022.

[22] S. Pratschner et al., "Versatile mobile communications simulation: The Vienna 5G link level simulator," *EURASIP J. Wireless Commun. Netw.*, vol. 2018, no. 1, p. 226, Dec. 2018.

[23] B. Zheng, C. You, W. Mei, and R. Zhang, "A survey on channel estimation and practical passive beamforming design for intelligent reflecting surface aided wireless communications," *IEEE Commun. Surveys Tuts.*, vol. 24, no. 2, pp. 1035–1071, 2nd Quart., 2022.

[24] F. Qamar, K. B. Dimyati, M. N. Hindia, K. A. B. Noordin, and A. M. Al-Samman, "A comprehensive review on coordinated multi-point operation for LTE–A," *Comput. Netw.*, vol. 123, pp. 19–37, Aug. 2017.

[25] A. Bletsas, A. Lippman, and J. N. Sahalos, "Simple, zero-feedback, distributed beamforming with unsynchronized carriers," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 7, pp. 1046–1054, Sep. 2010.

[26] S. Hanna, E. Krijestorac, and D. Cabric, "Destination-feedback free distributed transmit beamforming using guided directionality," *IEEE Trans. Mobile Comput.*, vol. 22, no. 10, pp. 5858–5869, Oct. 2023.

[27] W. B. Chikha, M. Masson, Z. Altman, and S. B. Jemaa, "Radio environment map based inter-cell interference coordination for massive-MIMO systems," *IEEE Trans. Mobile Comput.*, vol. 23, no. 1, pp. 785–796, Jan. 2024.

[28] S. Wang et al., "Location-aware coordinated beamforming for high mobility cell-free mmWave systems," *IEEE Wireless Commun. Lett.*, vol. 13, no. 3, pp. 899–903, Mar. 2024.

[29] H. Du et al., "Enabling AI-generated content services in wireless edge networks," *IEEE Wireless Commun.*, vol. 31, no. 3, pp. 226–234, Jun. 2024.

[30] Z. Liu, Y. Zhao, S. Hosseinalipour, Z. Gao, L. Huang, and H. Dai, "TDRA: A truthful dynamic reverse auction for DAG task scheduling over vehicular clouds," *IEEE Trans. Veh. Technol.*, vol. 73, no. 3, pp. 4337–4351, Mar. 2024.

[31] J. Wang et al., "Generative AI for integrated sensing and communication: Insights from the physical layer perspective," *IEEE Wireless Commun.*, vol. 31, no. 5, pp. 246–255, Oct. 2024.

[32] J. Ji, L. Cai, K. Zhu, and D. Niyato, "Decoupled association with rate splitting multiple access in UAV-assisted cellular networks using multi-agent deep reinforcement learning," *IEEE Trans. Mobile Comput.*, vol. 23, no. 3, pp. 2186–2201, Mar. 2024.

[33] J. Xu, Z. Zhou, L. Li, L. Zheng, and L. Liu, "RC-struct: A structure-based neural network approach for MIMO-OFDM detection," *IEEE Trans. Wireless Commun.*, vol. 21, no. 9, pp. 7181–7193, Sep. 2022.

[34] M. Xu et al., "Unleashing the power of edge-cloud generative AI in mobile networks: A survey of AIGC services," *IEEE Commun. Surveys Tuts.*, vol. 26, no. 2, pp. 1127–1170, 2nd Quart., 2024.

[35] H. Du et al., "Enhancing deep reinforcement learning: A tutorial on generative diffusion models in network optimization," 2023, *arXiv:2308.05384*.

[36] H. Ye, G. Y. Li, B. F. Juang, and K. Sivanesan, "Channel agnostic end-to-end learning based communication systems with conditional GAN," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Dec. 2018, pp. 1–5.

[37] M. Kim, R. Fritschek, and R. F. Schaefer, "Learning end-to-end channel coding with diffusion models," in *Proc. 26th Int. ITG Workshop Smart Antennas 13th Conf. Syst. Commun. Coding (WSA SCC)*, Feb. 2023, pp. 1–6.

[38] *Study on Channel Model for Frequencies From 0.5 To 100 GHz*, Standard (TR) 38.901, Version 17.1.0, 3GPP, 3rd Generation Partnership Project, 2022.

[39] *NR; Physical Channels and Modulation*, Standard (TS) 38.211, Version 18.1.0, 3GPP, 3rd Generation Partnership Project, 2024.

[40] S. Schwarz, C. Mehlführer, and M. Rupp, "Calculation of the spatial preprocessing and link adaption feedback for 3GPP UMTS/LTE," in *Proc. Wireless Adv.*, Jun. 2010, pp. 1–6.

[41] X. He, K. Niu, Z. He, and J. Lin, "Link layer abstraction in MIMO-OFDM system," in *Proc. Int. Workshop Cross Layer Design*, Sep. 2007, pp. 41–44.
[42] *NR; Physical Layer Procedures for Data*, Standard (TS) 38.214, Version 18.1.0, 3GPP, 3rd Generation Partnership Project, 2024.
[43] A. Vaswani et al., "Attention is all you need," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 30, Jun. 2017, pp. 5998–6008.
[44] J. Wang et al., "Generative artificial intelligence assisted wireless sensing: Human flow detection in practical communication environments," *IEEE J. Sel. Areas Commun.*, vol. 42, no. 10, pp. 2737–2753, Oct. 2024.
[45] J. Ho, A. N. Jain, and P. Abbeel, "Denoising diffusion probabilistic models," in *Proc. Adv. Neural Inf. Process. Syst.*, Jan. 2020, pp. 6840–6851.
[46] A. Alkhateeb, "DeepMIMO: A generic deep learning dataset for millimeter wave and massive MIMO applications," in *Proc. Inf. Theory Appl. Workshop (ITA)*, Feb. 2019, pp. 1–8.

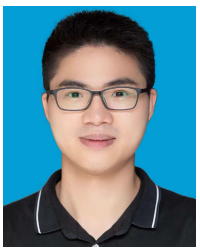
Yunting Xu received the B.S. degree in communication engineering from Nanjing University, Nanjing, China, in 2017, where he is currently pursuing the Ph.D. degree with the School of Electronic Science and Engineering. His current research interests include B5G/6G networks, vehicular ad hoc networks, and cybertwin-enabled dynamic resource management and optimization in the field of emerging wireless networks.

Zongxi Liu received the B.S. degree in electrical engineering and automation from Nanjing Normal University, Nanjing, China, in 2022. He is currently pursuing the M.S. degree in communication engineering with Nanjing University, Nanjing. His research interests include transmission schemes for B5G/6G networks and machine learning for wireless communications.

Bo Qian (Member, IEEE) received the B.S. and M.S. degrees in statistics from Sichuan University, Chengdu, China, in 2015 and 2018, respectively, and the Ph.D. degree in information and communication engineering from Nanjing University, Nanjing, China, in 2022. From 2022 to 2024, he was a Post-Doctoral Fellow with the Peng Cheng Laboratory, Shenzhen, China. He is currently the NII Project Researcher of the National Institute of Informatics, Tokyo, Japan. His research interests include resource management in B5G/6G networks, SAGIN, and VANET. He was a recipient of the Best Paper Award from IEEE VTC2020-Fall.

Hongyang Du (Member, IEEE) received the B.Eng. degree from the School of Electronic and Information Engineering, Beijing Jiaotong University, Beijing, and the Ph.D. degree from the Interdisciplinary Graduate Program, College of Computing and Data Science, Energy Research Institute @ NTU, Nanyang Technological University, Singapore. He is currently an Assistant Professor with the Department of Electrical and Electronic Engineering, The University of Hong Kong. His research interests include edge intelligence, generative AI, semantic communications, and network management. He was a recipient of the IEEE Daniel E. Noble Fellowship Award from the IEEE Vehicular Technology Society in 2022, the IEEE Signal Processing Society Scholarship from the IEEE Signal Processing Society in 2023, Singapore Data Science Consortium (SDSC) Dissertation Research Fellowship in 2023, and the NTU Graduate College's Research Excellence Award in 2024. He served as the Editor-in-Chief Assistant for IEEE COMMUNICATIONS SURVEYS AND TUTORIALS from 2022 to 2024, an Editor for IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, and the Guest Editor for *IEEE Vehicular Technology Magazine*. He was recognized as an Exemplary Reviewer of IEEE TRANSACTIONS ON COMMUNICATIONS and IEEE COMMUNICATIONS LETTERS.

Jiacheng Chen received the Ph.D. degree in information and communications engineering from Shanghai Jiao Tong University, Shanghai, China, in 2018. From 2015 to 2016, he was a Visiting Scholar with the BBCR Group, University of Waterloo, Canada. Currently, he is an Assistant Researcher with Peng Cheng Laboratory, Shenzhen, China. His research interests include fully decoupled radio access network technologies. He has served as the Guest Editor for IEEE INTERNET OF THINGS JOURNAL and the Workshop Co-Chair for IEEE/CIC ICCC from 2021 to 2024. He was a recipient of the Journal of Communications and Information Networks (JCIN) Best Paper Award in 2016 and the Chinese Institute of Electronics (CIE) Outstanding Scientific Paper in the Field of Electronic Information in 2020.



Jiawen Kang (Senior Member, IEEE) received the M.S. and Ph.D. degrees from Guangdong University of Technology, China, in 2015 and 2018, respectively. He was a Post-Doctoral Researcher with Nanyang Technological University, Singapore, from 2018 to 2021. He is currently a Full Professor with Guangdong University of Technology. His main research interests include blockchain, metaverse, and AIGC in wireless communications and networking.

Haibo Zhou (Senior Member, IEEE) received the Ph.D. degree in information and communication engineering from Shanghai Jiao Tong University, Shanghai, China, in 2014. From 2014 to 2017, he was a Post-Doctoral Fellow with the Broadband Communications Research Group, Department of Electrical and Computer Engineering, University of Waterloo. He is currently a Full Professor with the School of Electronic Science and Engineering, Nanjing University, Nanjing, China. His research interests include resource management and protocol design in B5G/6G networks, vehicular ad hoc networks, and space-air-ground integrated networks. He was a recipient of the 2019 IEEE ComSoc AsiaPacific Outstanding Young Researcher Award, the 2023 IEEE ComSoc WTC Outstanding Young Researcher Award, the 2023–2024 IEEE ComSoc Distinguished Lecturer, and the 2023–2025 IEEE VTS Distinguished Lecturer. He served as the Track/Symposium Co-Chair for IEEE/CIC ICCC 2019, IEEE VTC-Fall 2020, IEEE VTC-Fall 2021, WCSP 2022, IEEE GLOBECOM 2022, IEEE ICC 2024, and IEEE GLOBECOM 2024. He is an Associate Editor of IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, IEEE INTERNET OF THINGS JOURNAL, *IEEE Network Magazine*, and *Journal of Communications and Information Networks*.

Dusit Niyato (Fellow, IEEE) received the B.Eng. degree from the King Mongkut's Institute of Technology Ladkrabang (KMITL), Thailand, and the Ph.D. degree in electrical and computer engineering from the University of Manitoba, Canada. He is currently a Professor with the College of Computing and Data Science, Nanyang Technological University, Singapore. His research interests include mobile generative AI, edge intelligence, decentralized machine learning, and incentive mechanism design.