Graph Neural Network Aided Beamforming for Holographic Millimeter Wave MIMO Systems

Zou Linfu, Pan Zhiwen, Mohammed El-Hajjar, *Senior Member, IEEE*

***Abstract*—Holographic multiple-input multiple-output (HMIMO) systems are considered as one of the potential techniques to meet the demands of next-generation communications by replacing costly and power-hungry devices with sub-half-wavelength antenna elements. However, optimizing the beamforming matrix in the base station (BS) for HMIMO systems is challenging, given the prohibitive overhead of directly estimating the channels between the BS and the user equipment. Instead of following the traditional method of channel estimation and beamforming optimization, in this paper we employ a deep-learning technique to optimize the beamformers at the BS based on a loss function. Specifically, in this paper we introduce a graph neural network (GNN) designed to map the received pilot signals to optimized beamforming matrices and to model interactions among user equipment within the network. The simulation results show that our deep-learning method effectively maximizes the sum-rate objective while using reduced number of pilots than traditional channel estimation and beamforming optimization techniques.**

***Index Terms*—Beamforming, graph neural network, holographic MIMO, millimeter wave.**

I. INTRODUCTION

To support the expected requirements of next-generation wireless networks, holographic multiple-input multiple-output (HMIMO) systems have emerged as a feasible and promising technique that can propagate holographic radios with acceptable power consumption and hardware cost [1]. It is becoming increasingly clear that the fifth-generation wireless communication, or 5G, is now being implemented globally, where massive MIMO (mMIMO) techniques are employed to mitigate high path loss experienced at millimeter-wave (mmWave) frequencies [2]. Forecasts for subsequent wireless network generations anticipate a shift from mmWave towards Terahertz (THz) or even visible light frequencies, giving rise to significant power consumption, expensive fabrication cost, and extensive integration area in mMIMO systems [1]. Consequently, HMIMO systems have emerged as promising techniques to address these challenges, which leverage novel materials such as Graphene or positive-intrinsic-negative (PIN) diodes to optimize the associated cost and enable the manipulation of high-frequency signals [1], [2].

Additionally, HMIMO surfaces are regarded as nearly quasi-continuous spatial apertures with antenna elements separated by less than half the wavelength of the incident mmWave. Therefore, HMIMO surfaces can compensate for the substantial path loss inherent in mmWave propagation by forming sharp beams with weak side-lobes and supporting more antenna elements within a fixed aperture [1], [3].

However, despite the significant progress, several vital technical hurdles persist in HMIMO communication. First, it is necessary to establish appropriate channel estimation approaches to accurately precode and recover the transmitted signal to fully control HMIMO surfaces due to different hardware structures and the approximately continuous aperture [1], which lead to ultra-high computational complexity and overhead. To this end, Cui and Dai [4] considered HMIMO systems with large apertures and the near-field region, where the channel characteristics are affected by both angle and distance. Drawing inspiration from the observed sparsity within the far-field region, a paradigm shift to polar domain sparsity is introduced in [5], leading to the proposition of a polar-domain simultaneous compressed sensing (CS) algorithm tailored for near-field channel estimation. Zhang *et al.* [6] introduced a model-based approach employing dictionary learning and sparse recovery for channel estimation in extremely large MIMO (XL-MIMO) systems, which incorporates neural networks to approximate the transform matrix crucial in the polar-domain CS algorithm. To tackle the non-stationarity in estimating the HMIMO near-field channel, Han *et al.* [7] utilized a model encompassing the last-hop scatters under a spherical wavefront, segmenting the large aperture array into subarrays and applying a CS-based channel estimation method on a subarray level. Additionally, Lu and Dai [8] decomposed the problem of channel estimation for HMIMO systems into Line-of-Sight (LoS) and Non-Line-of-Sight (NLoS) subproblems. Meanwhile, Wei and Dai [9] introduced a hybrid channel model for HMIMO systems, encompassing both near- and far-field scenarios, utilizing the scatterers existing in both fields, where hybrid-field channel estimation was proposed by harnessing the CS algorithm separately within the angular and polar domains.

Furthermore, a deep-learning (DL) framework for hybrid-field channel estimation was proposed in [10], combining a linear and non-linear estimator. Dong *et al.* [11] investigated three distinct Convolution Neural Network (CNN) architectures that leverage the inherent spatial, frequency, and temporal correlations present in mmWave channels to achieve accurate channel estimation. Further Gao *et al.* [12] focused on mmWave massive MIMO systems employing mixed-resolution analog-to-digital converters (ADCs), where they compared traditional channel estimation methods with a deep learning-based neural network approach to demonstrate the improved performance offered by the deep learning method. Li *et al.* [13] investigated channel estimation for communication systems aided by reconfigurable intelligent surface (RIS), where they proposed a method that leverages the statistical properties of the channels that exist among the RIS, users, and access points to improve channel estimation accuracy. In recent efforts to mitigate channel feedback overhead within frequency division duplex (FDD) MIMO systems, several studies proposed autoencoder architectures [14], [15], which can compress the channel state information (CSI) while retaining essential characteristics for accurate signal reconstruction at the receiver. *Although promising, the existing literature for channel estimation in HMIMO systems is mostly based on the idealistic assumption of channel sparsity [16] and considers the substantial training overhead resulting from the deployment of a huge number of antenna elements [1].*

Additionally, the design of HMIMO beamforming is another challenge given the large number of antenna elements and the expected high beam resolution [1]. For this purpose, some research work explored a dynamic metasurface approach, where each antenna element is associated with a tuning weight, which can enable signal propagation along microstrip lines alongside other signals. Under this scenario, Shlezinger *et al.* [17] developed efficient alternating optimization (AO) algorithms aimed at determining the optimal configuration of surface weights. This optimization maximizes the achievable sum rate for uplink and downlink communication scenarios. Later, Wang *et al.* [18] extended the design to wideband channels employing Orthogonal Frequency-Division Multiplexing (OFDM) systems and using low-resolution ADCs, where the weights of the HMIMO beamforming have been jointly optimized along with the digital signal processing, adhering to a predefined bit constraint. Employing a different optimization objective, You *et al.* [19] focused on maximizing energy efficiency within single-cell multi-user mMIMO networks. They approached this by concurrently optimizing the transmit precoding matrices and the activation of weights, leveraging Dinkelbach's transform alongside the alternating optimization (AO) technique. Chen *et al.* [20] addressed the non-convex optimization problem by employing manifold optimization and sequential convex approximation. Xu *et al.* [21] explored the near-field wideband model, addressing the receiver beamforming problem through an iterative algorithm.

On the other hand, holographic principle-based works are another category in which to design the HMIMO beamforming, where the reference waves traverse the substrate, inducing a distinct amplitude and phase at each antenna element. This selective excitation, characterized by specific weights, culminates in generating the desired object wave [3]. Deng *et al.* [3] introduced a novel approach termed holographic-pattern division multiple access (HDMA) for a multi-user beamforming scenario. Their theoretical analysis substantiates that employing a zero-forcing (ZF) precoding technique can potentially attain the system's asymptotic capacity in the context of HDMA. In [22], the authors explored a scenario employing an HMIMO-assisted OFDM downlink communication with a single user, considering frequency-selective channels. The emphasis was on enhancing the achievable rate by concurrently optimizing digital and holographic beamforming, employing an amplitude control optimization algorithm. Continuing along this line, Deng *et al.* [23] proposed an iterative method for the joint optimization of digital and holographic beamforming at the holographic surface. Additionally, the authors of [24] proposed a technique for the joint optimization of beamforming, which is specifically tailored to maximize the sum rates in an HMIMO-assisted downlink multi-user communication system. In addition, Li *et al.* [25] conceived a framework that incorporates an intelligent omni-surface, and conducted a theoretical analysis to assess the ergodic rate performance of the proposed system design. Recent advancements in MIMO systems have explored the structure of overlapped subarrays, movable-antenna arrays, and a unified tensor approach for sensing, as discussed in [26]–[28]. *Although these approaches appear promising compared with traditional mMIMO systems [29], [30], it is crucial to acknowledge that the methods above presume perfect CSI and may ignore the substantial overhead of channel estimation [1]. These considerations underscore the existing limitations within HMIMO beamforming strategies.*

Recently, deep learning has become a powerful tool for wireless communications, showcasing improved performance across diverse domains like channel estimation, detection, and beamforming. Specifically, the authors of [31] and [32] validated the correlation between user position and the CSI, which has been approximated by neural networks drawing upon the universal approximation theorem [33] and then the downlink instantaneous and statistical CSI were obtained through trained neural networks to reduce the overhead. Shen *et al.* [34] considered challenges inherent to radio resource management, including power control and beamforming, which can be effectively transformed into graph optimization problems and proved that graph neural networks (GNNs), including message passing process, can effectively solve the problem and generalize to large scale. Additionally, the authors of [35] and [36] proposed a deep learning-based strategy that is capable of directly optimizing beamformers and the reflective coefficients at the RIS using the received pilots, which signifies a promising leap in optimizing wireless systems through machine learning paradigms. Furthermore, Feng *et al.* [37], [38] presented detection methods that are underpinned by deep learning and aimed at reducing the complexity and overhead for CS-aided multi-dimensional index modulation systems.

Consequently, within the scope of this study, we propose a deep-learning approach to optimize the multi-user HMIMO beamforming problem with implicit channel estimation for mmWave systems. More specifically, we adopt a GNN architecture for capturing the user interactions within the system, which can facilitate the extraction of a mapping function from the received pilot signals to the beamforming vectors at the base station (BS) to maximize the sum rate, while avoiding the conventional two-stage paradigm. First, the communication channels are estimated utilizing the received pilot signals, and then this estimation is employed to address the non-convex optimization problems. We consider an HMIMO-aided multi-user mmWave system, where the user equipment sends uplink pilots to the BS and then the received signals at the BS serve as inputs for the optimized GNN, directly producing downlink beamformers without requiring CSI estimation.

Compared with traditional hybrid analog-digital architectures and existing deep learning-based approaches [39]–[41], our method optimizes the baseband and holographic beam-

TABLE I
NOVELTY COMPARISON WITH THE LITERATURE

	Our paper	[4], [7]–[9]	[6], [10]	[18]	[3]	[23]	[34]	[35]	[39]–[41]
Multiuser	✓				✓	✓	✓	✓	✓
Multiplex	✓				✓	✓		✓	✓
Hybrid method	HMIMO			HMIMO	HDMA	HMIMO		RIS	MIMO
Channel estimation	Implicity	CS	DL					Implicity	Partial CSI
Sub-half wavelength	✓			✓		✓			

formers directly from the received pilot signals. This significantly reduces the overhead typically required for obtaining accurate channel state information, which is a crucial requirement in conventional systems. Furthermore, our model accounts for wave propagation delay and attenuation, while most traditional designs assume idealized propagation conditions. Additionally, we implement more efficient neural networks, where each hidden layer consists of a single-layer fully connected network, with the number of neurons corresponding to antenna elements.

The primary contributions of this paper lie in addressing the conventional challenges associated with CSI estimation and non-convex beamforming problems by employing a GNN-based data-driven approach, thereby optimizing the HMIMO system's performance in mmWave communications. As shown in Table I, while many approaches support multi-user scenarios, the proposed approach focuses on beamforming in Holographic MIMO systems with implicit channel estimation. Furthermore, our work considers the challenges of sub-half wavelength deployments supporting ultra-dense antenna elements within fixed aperture, which can lead to ultra-high computational complexity and overhead. The contributions of this paper can be summarized as follows.

- We consider the beamforming design of HMIMO supporting multiple users, where the BS is equipped with a large number of antenna elements under sub-half-wavelength space operating at mmWave frequency. Explicitly, to compute the optimized beamforming matrix, a GNN architecture is employed to directly map the received pilot signals to the optimized beamformers. The pilots can be utilized more effectively by the proposed machine-learning framework than by the traditional channel estimation-based methods.
- We employ a hybrid holographic beamforming decomposition algorithm to decouple the digital processing and holographic beamforming that can satisfy the hardware constraints of HMIMO systems with digital processing and holographic beamforming.
- Numerical simulations demonstrate that the employed deep-learning framework achieves superior performance over conventional model-based techniques in optimizing the sum rate. Furthermore, the solution addressed by a GNN exhibits robust generalization across a variety of signal-to-noise ratios (SNRs) and users. Additionally, the results show the efficiency of the proposed solution in reducing the high overhead of channel estimation and effectively solving non-convex beamforming problems.

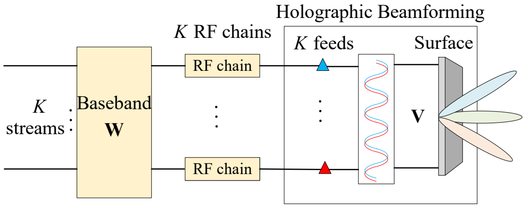

Fig. 1. The holographic beamforming process.

The subsequent sections of this paper are organized as follows. Section II describes the system model, the problem formulation, and the traditional strategy of channel estimation. Then, in Section III, we present the deep-learning framework and GNN architecture tailored for the HMIMO system. Section IV provides simulation results for the proposed scheme in terms of sum-rate objectives. Finally, conclusions are presented in Sec. V.

Notation: $\mathbf{x}$ represents a column vector. $\mathbf{X}$ represents a matrix. $\|\mathbf{X}\|_F$ denotes the Frobenius norm of $\mathbf{X}$, while $\det(\mathbf{X})$ and $|\mathbf{X}|$ denote the determinant of a matrix. $\mathrm{tr}(X)$ denotes the trace. $[\mathbf{X}]_{i,j}$ represents an element in the i-th row and j-th column of a matrix. $\mathbf{X}^\top$ and $\mathbf{X}^\mathrm{H}$ are the transpose and Hermitian transpose of matrices. $\mathbf{X}^{-1}$ denotes the inverse of a square non-singular matrix. $\mathbb{C}^{M\times N}$ denotes the set of $M\times N$ complex matrices. $\mathcal{CN}(0,\sigma^2)$ represents a complex Gaussian distribution following zero mean and σ^2 variance.

II. SYSTEM MODEL AND PROBLEM FORMULATION

In this section, we first introduce the architecture and attributes of an HMIMO surface and then we present the system model for a multi-user mmWave holographic beamforming system. In the proposed model, the users send uplink pilots to the BS, which then computes the optimized holographic beamformer from the received pilots to direct the signals in the desired directions. The sum rate optimization problem is presented, followed by a discussion of the conventional strategy to solve this problem.

A. HMIMO Structure and principle

In HMIMO wireless communication, the transmitter sends signals to the radio frequency (RF) chains responsible for up-converting the baseband signals to the carrier frequency. As shown in Fig. 1, the signals, initially transmitted as high-frequency currents in the RF chain, are converted to

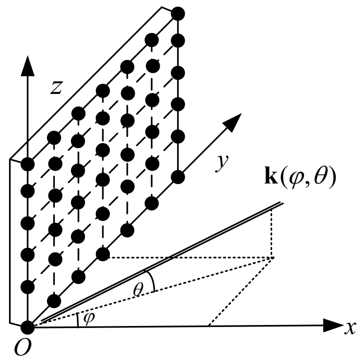

Fig. 2. The Cartesian coordinate.

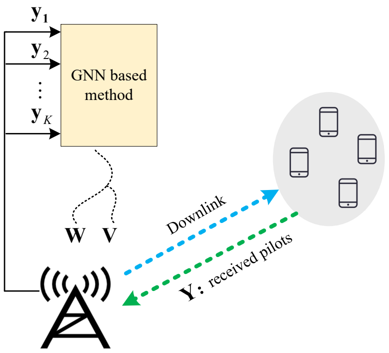

Fig. 3. The considered system model.

electromagnetic waves that propagate across the HMIMO surface. The HMIMO surface subsequently utilizes holographic beamforming to generate beams toward the intended direction.

The HMIMO hardware structure consists of three main parts: feed, substrate, and antenna elements. A detailed explanation of the roles and functions of these components is elaborated below [1]:

- *Feed:* The feeds generate electromagnetic waves originating from the RF chain. These waves are designed to propagate along the surface of the HMIMO system.
- *Substrate:* The substrate functions as a waveguide within the HMIMO system, facilitating the propagation of the reference wave along it. Specifically, the substrate allows the reference waves generated by the feeds to travel to the HMIMO surface directly and leak out to free space after being tuned by the antenna elements.
- *Antennas element:* The antenna elements equipped on the HMIMO surface with sub-wavelength spacing transform the reference waves to object waves to realize the beamforming. Lumped elements, Graphene, or photosensitive devices drive the tuning structure of the HMIMO surfaces.

The main role of the HMIMO is to create the *holographic pattern*, which makes it possible to obtain the object wave by controlling the reference wave. Precisely, the HMIMO structure captures the interference pattern that arises from the interaction between the reference wave and the object wave, following the principles of holographic interference [24].

We adopt a Cartesian coordinate system in which the plane defined by the y and z axes is aligned with the HMIMO surface. In this configuration, the x-axis is perpendicular to the HMIMO surface as shown in Fig. 2. Without loss of generality, we assume that there are $M = M_y \times M_z$ antenna elements alongside the y-axis and z-axis on the HMIMO surface.

The hologram must be designed and effectively constructed using the HMIMO surface during the recording process. At the m-th antennas element, the object wave approaches from the direction specified by the angles (θ, ϕ) and the reference wave activated by the feed k can be expressed as [3]

$$\Psi_o(\mathbf{r}_m, \theta_k, \varphi_k) = e^{-j\mathbf{k}(\theta_k,\varphi_k)\cdot \mathbf{r}_m},$$
$$\Psi_r(\mathbf{r}_m^k, \mathbf{k}^k) = e^{-j\mathbf{k}^k \cdot \mathbf{r}_m^k},$$

where $\mathbf{k}(\theta_k, \varphi_k)$ represents the wave vector for the desired directional propagation towards the k-th user in free space. $\mathbf{r}_m$ denotes the position vector of the m-th radiation element (antenna element) of the array. This vector specifies the location of the antenna element in space relative to a reference point, typically the origin or a specific point on the HMIMO surface. $\mathbf{r}_m^k$ is the distance vector from the k-th feed to the m-th radiation element. $\mathbf{k}^k$ is the propagation vector of the reference wave from the k-th feed. By propagating the object wave from a given direction to the HMIMO surface, the interference wave, expressing a certain hologram, arises from the superposition of the object and reference wave, which can be defined as [3]

$$\Psi_i(\mathbf{r}_m, \theta_k, \varphi_k) = \Psi_o\left(\mathbf{r}_m, \theta_k, \varphi_k\right) \cdot \Psi_r^*(\mathbf{r}_m^k, \mathbf{k}^k). \quad (1)$$

The information contained in Ψ_i, also called a holographic pattern, is recorded by the HMIMO surface.

In construction, an object wave can be generated when the reference wave, once excited, propagates along the surface of the HMIMO structure, where the hologram has been established by (1). Specifically, the wave propagating towards the direction (θ, ϕ) can be expressed as

$$\Psi_i \Psi_r \propto \Psi_o |\Psi_r|^2.$$

B. System Model

We consider a downlink multi-user system assisted by HMIMO, in which a BS with an array of $M = M_y \times M_z$ antenna elements serves K users, each having a single antenna. Here, M_y and M_z represent the number of antenna elements aligned along the y-axis and z-axis, respectively.

Initially, the users transmit uplink pilot signals to the BS, and then the BS computes the optimized holographic beamforming matrix to direct the multi-user signals in the desired directions, as shown in Fig. 3.

Due to the lack of digital processing capability on the HMIMO surface in the downlink stage, the BS is required to manage signal processing at the baseband level. As depicted in Fig. 1, the BS encodes K distinct data streams using baseband processing $\mathbf{W}$. Subsequently, the BS up-converts the processed signals to the carrier frequency, facilitated by the RF chains. Each RF chain sends the up-converted signals to its corresponding feed which transforms the high-frequency current into an electromagnetic wave, termed the reference wave. This reference wave is directed to propagate along the HMIMO surface by utilizing a waveguide structure. The direction of the reference wave at each element is controlled through a holographic beamformer $\mathbf{V}$ to generate the desired beams. Based on the principle of HMIMO, the holographic beamformer $\mathbf{V} \in \mathbb{C}^{M \times K}$ is formed as follows [3]

$$[\mathbf{V}]_{m,k} = e^{-j\mathbf{k}(\varphi_k,\theta_k)\cdot\mathbf{r}_m} \cdot e^{j\mathbf{k}^k\cdot\mathbf{r}_m^k} \cdot e^{-\alpha|\mathbf{r}_m^k|}, \tag{2}$$

where $e^{j\mathbf{k}^k\cdot\mathbf{r}_m^k}$ represents the phase of the reference wave as it propagates from feed k to the m-th radiation element. α and $e^{-\alpha|\mathbf{r}_m^k|}$ denote the attenuation constant and signal loss that occurs as the wave propagates from the feeds to the surface, respectively.[1]

Let $\mathbf{s} \in \mathbb{C}^{L\times 1}$ be the transmitted symbols from the BS to K users satisfying $\mathbb{E}[\mathbf{s}\mathbf{s}^{\mathrm{H}}] = \mathbf{I}$. Therefore, the transmitted signal from the BS can be denoted as $\mathbf{Ws}$. After the holographic beamforming using the HMIMO surface integrated with the BS, the signal received by the k-th user can be expressed as [3]

$$r_k = \mathbf{h}_k^{\mathrm{H}}\mathbf{V}\mathbf{W}\mathbf{s} + n, \tag{3}$$

where $\mathbf{h}_k^{\mathrm{H}} \in \mathbb{C}^{1\times M}$ represents the downlink channel between the k-th user and the BS, and $n \sim \mathcal{CN}(0,\sigma^2)$ denotes the additive white Gaussian noise (AWGN). Assuming channel reciprocity, the downlink transmission channel matrices are merely the transpose of the uplink channel vectors [35]. Consequently, the achievable rate R_k for the k-th user can be computed as

$$R_k = \log\left(1 + \frac{|\mathbf{h}_k^{\mathrm{H}}\mathbf{V}\mathbf{w}_k|^2}{\sum_{i\neq k}|\mathbf{h}_k^{\mathrm{H}}\mathbf{V}\mathbf{w}_i|^2 + \sigma^2}\right), \tag{4}$$

where $\mathbf{w}_k$ is the k-th column of the matrix $\mathbf{W}$ and σ^2 denotes the noise power. $|\mathbf{h}_k^{\mathrm{H}}\mathbf{V}\mathbf{w}_i|^2$ represents the interference signal of the i-th user to the k-th user through the channel $\mathbf{h}_k^H$.

The HMIMO surface can generate multiple beams directed toward the users, which is achieved through the superposition of the holographic pattern represented by the matrices $\mathbf{W}$ and $\mathbf{V}$. The matrix $\mathbf{W}$ typically denotes the baseband precoder, while $\mathbf{V}$ is the holographic beamforming matrix. For each user k, the interference caused by the presence of other users can be mitigated by jointly optimizing the beamformers $\mathbf{W}$ and $\mathbf{V}$ at the BS [3]. The baseband precoder $\mathbf{W}$ and holographic beamformer $\mathbf{V}$ can be jointly optimized to maximize the sum rate $\sum_{k=1}^{K} R_k$.

The knowledge of the channel between users and the BS is needed to optimize the digital precoder and holographic beamformer. To achieve this, an uplink pilot transmission stage is conducted prior to the downlink data transmission stage. In time-division-duplex (TDD) systems, channel estimation, can be acquired by uplink channel estimation due to uplink-downlink channel reciprocity [31]. In frequency-division-duplex (FDD) systems, the CSI at both the transmitter and receiver sides can be obtained by feedback link or by using deep learning methodologies [32].

The strategy we implement for channel estimation is based on the pilot transmission approach in the work of [46]. The strategy involves τ_p samples within each coherence interval specifically for conducting channel estimation based on uplink pilots. This allocation supports a set of τ_p orthogonal pilot sequences, which are distinct. These pilot sequences are then assigned to different users across the network. The simultaneous transmission of the pilots by the users, where the k-th user transmits the pilot sequence $\mathbf{x}_k^{\mathrm{H}} \in \mathbb{C}^{\tau_p}$ satisfying $\mathbf{x}_k^{\mathrm{H}}\mathbf{x}_l = 0$ if $k \neq l$ and $\mathbf{x}_k^{\mathrm{H}}\mathbf{x}_k = \tau_p$, is carefully coordinated. Subsequently, the received pilot signal at the BS represented by the matrix $\mathbf{Y} \in \mathbb{C}^{M\times\tau_p}$ is [46]

$$\mathbf{Y} = \sum_{k=1}^{K}\sqrt{p_u}\mathbf{h}_k\mathbf{x}_k^{\mathrm{H}} + \mathbf{N}, \tag{5}$$

where $\mathbf{N} \in \mathbb{C}^{M\times\tau_p}$ has independent and identically distributed $\mathcal{CN}(0,\sigma^2)$-elements. $\mathbf{h}_k \in \mathbb{C}^{M\times 1}$ denotes the uplink channel between the BS and user k and p_u is the uplink power constraint.

To estimate channel $\mathbf{h}_k$ for user k, the BS exploits orthogonality of the pilots by multiplying $\mathbf{x}_k$ with $\mathbf{Y}$ to obtain [46]

$$\mathbf{y}_k = \mathbf{Y}\mathbf{x}_k = \sqrt{p_u}\tau_p\mathbf{h}_k + \mathbf{N}\mathbf{x}_k. \tag{6}$$

The processed received pilot signal, denoted as $\mathbf{y}_k \in \mathbb{C}^M$, is sufficient to estimate $\mathbf{h}_k$ [46]. Typically, to ensure that the channel $\mathbf{h}_k$ can be recovered successfully, we need at least $\tau_p = K$ pilot symbols. When $\tau_p = K$, one choice of pilot matrix is a discrete Fourier transform (DFT) matrix as recommended in [35]. For comparison purposes, we also examine a more general scenario where $\tau_p \neq K$. Under these circumstances, we initiate by constructing a DFT matrix of size $\tau = \max(\tau_p, K)$, then truncate $\mathbf{x}_k$ to the first τ_p columns or the first K rows.

C. Problem Formulation

The paper aims to maximize the sum rates without explicitly estimating the channel coefficients, given that the channel estimation has a prohibitive overhead requirement. To directly map the pilot signal to the optimized digital precoder and holographic beamformer for sum-rate maximization, we propose to design the optimal digital precoder and holographic beamformer based on the received pilot signal $\mathbf{y}_k$. Specifically, our objective is to address the subsequent optimization problem [3]

$$\begin{aligned} \max_{(\mathbf{W},\mathbf{V})=g(\{\mathbf{y}_k\})} \quad & \sum_{k=1}^{K} R_k, \\ \text{subject to} \quad & ||\mathbf{W}||_F^2 \leq p_d, \\ & 0 \leq |[\mathbf{V}]_{m,k}| \leq \frac{1}{M}, m = 1\ldots M, k = 1\ldots K, \end{aligned} \tag{7}$$

where $0 \leq |[\mathbf{V}]_{m,k}| \leq \frac{1}{M}$ represents the transmitted power spreading to M elements, accounting for the attenuation as the signal propagates to the surface. p_d is the downlink power constraint. $g(\cdot)$ serves as a mapping from the received pilot signals to the digital precoding matrix $\mathbf{W}$ and the holographic beamforming matrix $\mathbf{V}$. In contrast to traditional MIMO systems that depend on complex phase shifting circuits, holographic beamforming is facilitated by meta-material based radiation

---

[1] Note that symbol-level synchronization in MIMO systems is outside the scope of this paper, while this has been addressed in many state-of-the-art works such as [42]–[45].

elements. These elements can be created by incorporating low-power and low-complexity RF switches, such as those made from Graphene. The radiation phase of these elements can be dynamically adjusted by manipulating the bias voltage applied to the RF switches [1].

Addressing problem (7) presents a formidable computational challenge due to its non-convex nature. To address this problem, we suggest representing the mapping function $g(\cdot)$ using a deep neural network and training this network through data-driven learning techniques. This strategy is inspired by the universal approximation theorem, which posits that neural networks with sufficient parameters can approximate a wide range of functions with a high degree of accuracy [33]. Before delving into our proposed solution, it is crucial to outline the traditional approach for solving the optimization problem, which can be considered a benchmark for our proposed solution.

D. Conventional Channel Estimation

This subsection describes the traditional method used to address the challenge (7). This method consists of two sequential stages: an uplink channel estimation phase followed by a downlink sum-rate maximization stage. The problem of maximizing the downlink sum rate, given the estimated channels, has been extensively explored. For example, there are solutions to the downlink sum-rate maximization problem that can be implemented in a fully digital manner [47], or, alternatively, in a hybrid holographic manner [3]. In the following discussion, we concentrate on the phase of uplink pilot transmission and the subsequent channel estimation process.

To estimate the channel $\mathbf{h}_k$ as formulated in equation (6), a variety of estimation techniques can be utilized by tackling the following optimization problem

$$\min_{f(\cdot)} \mathbb{E}\left[\|f(\mathbf{y}_k) - \mathbf{h}_k\|_F^2\right]. \tag{8}$$

However, the optimal solution is computationally expensive. Therefore, we compare a low-complexity least-square (LS) and a higher complexity minimum mean squared error (MMSE) approach with the proposed algorithm. The LS estimator is given by [46]

$$\hat{\mathbf{h}}_k^{\mathrm{LS}} = \frac{1}{\sqrt{p_u \tau_p}} \mathbf{y}_k, \tag{9}$$

while the MMSE estimator is defined as [46]

$$\hat{\mathbf{h}}_k^{\mathrm{MMSE}} = \bar{\mathbf{h}}_k + \sqrt{p_u}\mathbf{R}_k \mathbf{\Psi}_k \left(\mathbf{y}_k - \bar{\mathbf{y}}_k\right), \tag{10}$$

where $\mathbf{\Psi}_k = \left(p_u \tau_p \mathbf{R}_k + \sigma^2 \mathbf{I}_M\right)^{-1}$, the mean $\bar{\mathbf{h}}_k, \bar{\mathbf{y}}_k$ correspond to the LoS component and the received pilot signal, respectively, and $\mathbf{R}_k \in \mathbf{C}^{M \times M}$ represents the covariance matrix that describes the spatial correlation of the NLoS components. It is essential to recognize that the effectiveness of the MMSE estimator relies on having full knowledge of the statistical distributions. In practice, these statistical parameters are often estimated using the sample mean and covariance matrix derived from the observed data points [46].

III. PROPOSED GNN BASED SOLUTION

The traditional strategy for channel estimation is predicated on estimating the elements of $\mathbf{h}_k$ based on the received pilot signals $\mathbf{y}_k$, employing a mean squared error criterion. Nonetheless, the accurate estimation of the channel is not the end goal. Instead, the overarching aim is to optimize network utility as delineated in equation (7). This work proposes a straightforward solution to the problem (7) rather than using explicit channel estimation.

In particular, the objective is to utilize a neural network to approximate the mapping function $g(\cdot)$ as expressed in the optimization problem (7). The pursuit is to adopt a data-driven scheme for training the neural network to imitate the optimal transformation from the acquired pilot signals to the digital precoder and holographic beamformer, thereby facilitating the maximization of the network sum rate. The comprehensive framework is depicted in Fig. 3, where the users first send uplink pilots to the BS, which inputs the received signals to the optimized GNN and obtains the downlink beamformer directly without CSI estimation.

The results presented in [34] elucidate the mathematical equivalence between the GNN architecture and the weighted-minimum-mean-square-error (WMMSE) algorithm. The preference for using $\mathbf{W}_{\text{opt}}$ arises from the inherent limitations of GNNs, especially when dealing with complex optimization problems. GNNs can face challenges such as limited expressive power, as discussed in [48]. Although a fully connected neural network with sufficient width could theoretically approximate this mapping, such an approach would be computationally demanding and resource-intensive, which makes it impractical for real-world implementation. By focusing on outputting $\mathbf{W}_{\text{opt}}$, the complexity of the neural network's task is reduced while avoiding the challenges associated with predicting both $\mathbf{W}$ and $\mathbf{V}$, which can improve the model's efficiency and feasibility. A comparative analysis of the two strategies was conducted, focusing on the sum rate performance, where the findings are presented in Table II. It is important to emphasize that the GNN architectures employed in both strategies were identical, and both were trained using carefully tuned hyperparameters, differing only in the design of the output layer. This setup ensures a fair comparison and highlights the effectiveness of using $\mathbf{W}_{\text{opt}}$ in optimizing the performance while maintaining practical resource constraints. Remarkably, the approach where the two-layer GNN is designed solely to generate the digital precoder $\mathbf{W}_{\text{opt}}$, demonstrated superior performance compared with the strategy that produces both the baseband precoder $\mathbf{W}$ and the holographic precoder $\mathbf{V}$ at SNR= 20 dB with $K = 5$, and $M = 64$. This outcome highlights the efficiency of focusing the GNN on optimizing the $\mathbf{W}_{\text{opt}}$ alone, thus achieving better performance in a more practical manner.

We detail the GNN-based architecture specifically designed for this task in the section.

A. Graphical Representation of Users

Managing interference between users is a crucial aspect of multi-user communications. Hence, the digital precoder and

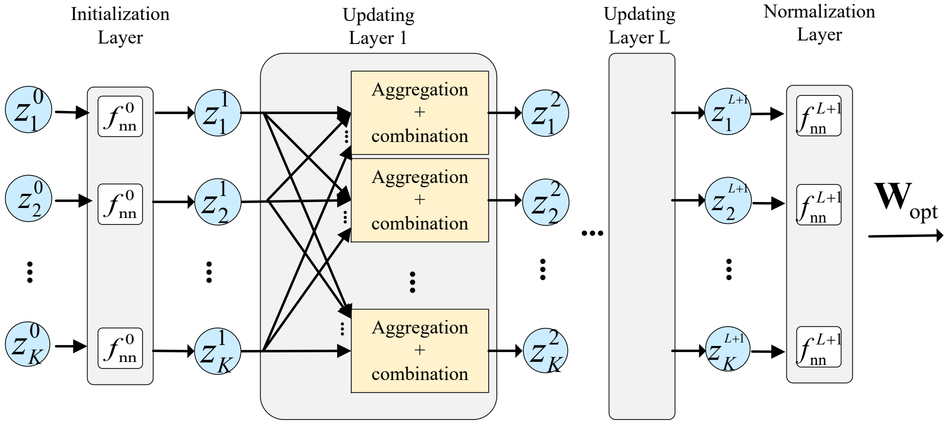

Fig. 4. The proposed GNN architecture.

TABLE II
SUM RATE PERFORMANCE OF DIFFERENT STRATEGIES
($M = 64$).

SNR [dB]	K	$\mathbf{W}_{\text{opt}}$ [bps/Hz]	$\mathbf{W}, \mathbf{V}$ [bps/Hz]
5	5	1.7	1.5
	10	2.2	1.6
15	5	5.9	5.1
	10	6.4	5.2

holographic beamformer must be coordinated to minimize mutual interference. As a result, we put forward the use of a neural network architecture, specifically GNN [49], that is founded on a graph representation of the digital beamformers to handle multi-user interference [34], [35]. The graph includes K nodes corresponding to K users, as depicted in Fig. 4. A representation vector, denoted as $\{\mathbf{z}_k^0\}_{k=1}^K$, is associated with each node, which gets updated in a GNN layer by layer. The primary objective is to encode each node's valuable information in these representation vectors. Following multiple layers, the representation vector associated with each node will contain enough information to support the design of beamforming vectors.

Compared with other neural networks, GNNs offer unique advantages for beamforming in wireless communication systems. One key benefit is their inherent *permutation equivariance* [35], [49]. This property ensures that the GNN architecture remains unaffected by permutations in the user ordering within the network. Consequently, a GNN-based beamforming model trained for a specific number of users can be readily generalized to networks with different user counts as long as the underlying graph structure remains consistent.

Furthermore, GNNs naturally capture the crucial inter-user interactions that influence beamforming decisions [49]. During the message passing phase within the GNN, each user node's update is informed by the information received from its neighbouring user nodes in the graph. This allows the GNN to effectively learn the complex dependencies and spatial relationships between users, improving the beamforming performance.

The GNN decreases model complexity compared with the fully connected neural network (FNN) by adapting to the specific problem structure [49]. In GNNs, graphs are used, with nodes denoting users and edges denoting relationships between them. GNNs are designed to learn a single set of parameters for information processing on the graph. These parameters are applied to all nodes, regardless of the total number of users [35]. In contrast, fully connected neural networks typically have a layer where each user has its own weights and biases. With more users, the number of parameters in this layer rapidly explodes, making it computationally expensive [34]. In the next subsection, we offer an in-depth explanation of the GNN architecture and the training process.

B. *GNN Architecture*

This section outlines the proposed GNN architecture and explains the training process. The primary goal of the GNN is to acquire the graph representation vector $\{\mathbf{z}_k^0\}_{k=1}^K$ by first initializing it, then updating it through L layers that include aggregation and combination modules, and finally normalizing it. The GNN takes the received pilot signal as the input data, with the real and imaginary parts concatenated in a vector form. Then, the initialization layer transforms the input data into a representation vector, denoted as $\{\mathbf{z}_k^1\}_{k=1}^K$, followed by L updating layers which conduct message passing through all nodes to output $\{\mathbf{z}_k^L\}_{k=1}^K$. Finally, a normalization layer generates the beamformer matrix $\mathbf{W}_{\text{opt}}$ via normalization. Fig. 4 depicts the overall architecture.

1) Initialization Layer: The initialization layer receives input features $\{\mathbf{z}_k^0\}_{k=1}^K$ from each user node and generates $\{\mathbf{z}_k^1\}_{k=1}^K$ for the next layer. The input features come from the received pilot signal of each user node and can be expressed as

$$\mathbf{z}_k^0 = [\Re\{\mathbf{y}_k\}^\top, \Im\{\mathbf{y}_k\}^\top]^\top, \tag{11}$$

where $\Re$ and $\Im$ express the real and imaginary components.

With the input feature vector $\mathbf{z}_k^0$, we employ a FNN [35], [49], denoted as f_{nn}^0, to produce $\mathbf{z}_k^1$ for the user nodes, which can be modeled as

$$\mathbf{z}_k^1 = f_{\text{nn}}^0(\mathbf{z}_k^0), \ k = 1, \ldots, K. \tag{12}$$

The representation vectors $\{\mathbf{z}_k^1\}_{k=1}^K$ now contain features related to the user channels and are subsequently forwarded

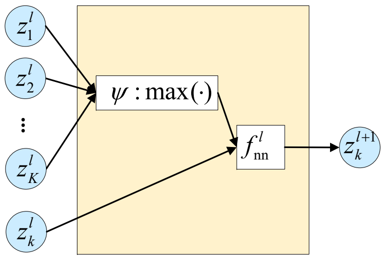

Fig. 5. The architecture of the aggregation process used in each GNN layer.

to the L updating layers of the GNN, aiming to generate the user beamformer $\mathbf{W}_{\text{opt}}$ from $\{\mathbf{z}_k^L\}_{k=1}^K$, where $\mathbf{W}_{\text{opt}}$ is the optimized downlink beamformer which maximizes the system sum rate.

2) Updating Layer: The L updating layers pass messages between nodes to produce the user beamformers [34]. For each updating layer, the output representation vector in the l-th layer, denoted as $\mathbf{z}_k^l$, is formulated by combining its previous representation $\mathbf{z}_k^{l-1}$ and the aggregation of the representations $\mathbf{z}_j^{l-1}$ from its neighbouring nodes. In a typical GNN, the process is given by [49]

$$\mathbf{z}_k^l = f_{\text{combine}}^l\left(\left[\mathbf{z}_k^{l-1}, f_{\text{aggregate}}^l\left(\left\{\mathbf{z}_j^{l-1}\right\}_{j\in\aleph(k)}\right)\right]\right), \tag{13}$$

where $\aleph(k)$ denotes the set of neighboring nodes of the node k, $f_{\text{combine}}(\cdot)$ and $f_{\text{aggregate}}(\cdot)$ are the combining and aggregation function in the l-th layer. Choosing a suitable aggregation function $f_{\text{aggregate}}(\cdot)$ and combining function $f_{\text{combine}}(\cdot)$ in (13) is vitally essential so that the GNN represents and generalizes the features accurately. The combining function f_{combine}^l in (13) can be executed by a FNN to address optimization problems, such as that in (7) [34], while an efficient implementation of $f_{\text{aggregate}}(\cdot)$ has the form [49]

$$f_{\text{aggregate}}^l\left(\left\{\mathbf{z}_j^{l-1}\right\}_{j\in\mathcal{N}(k)}\right) = \psi\left(\left\{\mathbf{z}_j^{l-1}\right\}_{j\in\aleph(k)}\right), \tag{14}$$

where ψ is chosen to be invariant to input permutation and is defined as the element-wise max-pooling function, given by

$$[\psi(\mathbf{z}_1,\ldots,\mathbf{z}_K)]_i = \max\left([\mathbf{z}_1]_i,\ldots,[\mathbf{z}_K]_i\right), \tag{15}$$

which demonstrates strong performance in practical applications and is consistent with the fact that the strongest user dominates the multi-user interference [35], [49].

To conclude, the overall operation of the l-th updating layer can be shown in Fig. 5 and expressed as

$$\mathbf{z}_k^l = f_{\text{nn}}^l\left(\left[\mathbf{z}_k^{l-1}, \psi\left(\left\{\mathbf{z}_j^{l-1}\right\}_{j\in\aleph(k)}\right)\right]\right). \tag{16}$$

The representation vector $\mathbf{z}_k^l$ of the l-th updating layer is generated by passing its previous representation vector of node k, $\mathbf{z}_k^{l-1}$, and then concatenating with the aggregation output of the neighbour nodes to an FNN.

3) Normalization Layer: Following L update layers, the representation vectors $\{\mathbf{z}_k^L\}_{k=1}^K$ are delivered to a normalization layer to produce the beamforming matrix $\mathbf{W}_{\text{opt}} \in \mathbb{C}^{M\times K}$, while satisfying the total power constraint. To this end, we first transmit $\mathbf{z}_k^{L+1}$ to a fully connected layer f_{nn}^{L+1} with $2M$ output units, modeled as [35]

$$\mathbf{z}_k^{L+2} = f_{\text{nn}}^{L+1}\left(\mathbf{z}_k^{L+1}\right), k = 1,\ldots,K, \tag{17}$$

Algorithm 1 Iterative Matrix Decomposition for the HMIMO GNN System.

```
Input: $\mathbf{W}_{\text{opt}}$
Output: $\mathbf{W}, \mathbf{V}$
Initialize $\mathbf{V} = \angle\mathbf{W}_{\text{opt}}$;
repeat
    Update $\mathbf{W} \leftarrow \left(\mathbf{V}^{\text{H}}\mathbf{V}\right)^{-1}\mathbf{V}^{\text{H}}\mathbf{W}_{\text{opt}}$;
    Update $\mathbf{V} \leftarrow \mathbf{W}_{\text{opt}}\mathbf{W}^{\text{H}}\left(\mathbf{W}\mathbf{W}^{\text{H}}\right)^{-1}$;
    Update $\mathbf{V} \leftarrow \frac{1}{\sqrt{M}}\angle\mathbf{V}$;
until Convergence
Update $[\mathbf{V}]_{m,k} \leftarrow [\mathbf{V}]_{m,k} \cdot e^{j\mathbf{k}^k\cdot\mathbf{r}_m^k}$;
```

followed by normalization steps

$$\begin{aligned}
\mathbf{Z}_w &= \left[\mathbf{z}_1^{L+2},\ldots,\mathbf{z}_K^{L+2}\right] \in \mathbb{R}^{2M\times K},\\
\mathbf{Z}_w &= \sqrt{p_d}\,\frac{\boldsymbol{Z}_w}{\|\boldsymbol{Z}_w\|_F},\\
\mathbf{W}_{\text{opt}} &= \boldsymbol{Z}_w(1:M,:) + j\boldsymbol{Z}_w(M+1:2M,:),
\end{aligned} \tag{18}$$

where $\mathbf{Z}(i1 : i2,:)$ denotes the sub-matrix of $\mathbf{Z}$ sliced by collecting from the $i1$-th to the $i2$-th rows of $\mathbf{Z}$. p_d is the downlink power constraint.

It is worth noting that we use the same $f_{\text{nn}}^0, f_{\text{nn}}^l$, and f_{nn}^{L+1} for the progression of node representation vectors across all user nodes, demonstrating the generalization of the GNN in accommodating any number of users. If we modify the number of users in the system, we must also adapt the number of nodes accordingly. The operations learned through equations (12), (16), and (17) are independent of the user number and still work without the need to retrain the neural network.

C. Holographic beamforming decomposition

In the previous sections, we presented the GNN-based HMIMO precoder matrix design to maximize the multi-user sum rate. However, the holographic MIMO system consists of a baseband precoder and holographic beamformer, where the precoder produced by the GNN cannot be used directly to generate a directional beam wave. Therefore, to circumvent these challenges, we aim to decompose the GNN-based HMIMO precoder matrix into a baseband precoder and a holographic beamforming matrix.

The decomposition problem can be formulated as

$$\min_{\mathbf{V},\mathbf{W}} \|\mathbf{W}_{\text{opt}} - \mathbf{V}\mathbf{W}\|_F^2, \tag{19}$$

where $\mathbf{W}_{\text{opt}}$ denotes the optimal precoder matrix produced by the GNN network, and $\mathbf{W}$, and $\mathbf{V}$ express the baseband precoder and holographic beamforming matrix, respectively. Note that as the paper does not focus directly on matrix decomposition, we opt to convert (19) into a convex quadratically-constrained quadratic programming (QCQP) problem, whose solutions to this problem can be found in Algorithm 1 [2].

The overall steps of the GNN-based implicit channel estimation in hybrid holographic beamforming are summarized in Algorithm 2. The CSI is necessary for generating the training

Algorithm 2 Off-line Traning of GNN-based method.

```
Input: Training dataset $\mathcal{D} = \left\{\left(\mathbf{Y}^{(n)}, \mathbf{H}^{(n)}\right)\right\}_{n=1}^{N_{\text{data}}}$, validation data set $\mathcal{V}$
Output: Optimized GNN
Initialize parameters in GNN;
repeat
    for $n = 1 \dots N_{\text{data}}$ do
        Select $\left(\mathbf{Y}^{(n)}, \mathbf{H}^{(n)}\right)$ from $\mathcal{D}$ ;
        Collect GNN output $\mathbf{W}_{\text{opt}}$;
        Compute $\mathbf{W}_{\text{HB}}$ and $\mathbf{W}_{\text{BP}}$ by Algorithm 1;
        Compute train loss using (7);
        Update GNN parameters via optimizer;
    end for
until Non-decrease loss in $\mathcal{V}$
```

samples and evaluating the network sum rate. Once the neural network is trained, CSI is no longer necessary for its operation. It is worth mentioning that the neural network training is completed offline. Therefore, it has no impact on the runtime complexity of the proposed approach.

The neural network adapts its weights throughout the training process to maximize the sum rate in (7) by employing an unsupervised stochastic gradient descent method. The automatic implementation of neural network parameter updates and gradient computations can be achieved using a standard numerical deep learning software package [50]. By conducting end-to-end training, we can jointly optimize the baseband precoder and holographic beamformer, directly leveraging the information obtained from the received pilot signals. In the forthcoming section, the simulation results demonstrate that the proposed deep learning technique can more effectively address problem (7) by requiring fewer pilots to attain equivalent performance compared with the traditional method of separate channel estimation and network maximization approach.

IV. SIMULATION RESULTS AND ANALYSIS

In this section, we evaluate the performance of the proposed GNN-aided algorithm for sum-rate maximization and compare it with the channel estimation-based approach. Specifically, the problem of maximizing the sum rate is formulated as

$$\sum_{k=1}^{K} R_k^{\text{GNN}} = \sum_{k=1}^{K} \log\left(1 + \frac{|\mathbf{h}_k^{\text{H}} \mathbf{w}_{\text{opt}k}|^2}{\sum_{i \neq k} |\mathbf{h}_k^{\text{H}} \mathbf{w}_{\text{opt}i}|^2 + \sigma^2}\right), \quad (20)$$

where $||\mathbf{W}_{\text{opt}}||_F^2 \leq p_d$. Consequently, the loss function is defined as $-\mathbb{E}\left[\sum_1^K R_k^{\text{GNN}}\right]$.

A. Simulation Setting

We consider a multi-user HMIMO system aided by GNN, depicted in Fig. 3, where the system comprises K users and a BS equipped with a uniform rectangular array positioned on the Cartesian plane defined by the coordinates (y, z), characterized by dimensions $M_y \times M_z$. The BS coordinates in meters are $(0, 0, 20)$ along the (x, y, z)-axes. Users are uniformly distributed within a rectangular region $[50, 150] \times [-50, 50]$ in meters on the (x, y)-plane, with a fixed $z = 0$.

Assume that the users are located where a LoS channel exists between the users and BS, so the channel $\mathbf{h}_k$ is Rician fading [35], [46]:

$$\mathbf{h}_k = \beta_k \left(\sqrt{\frac{\kappa}{1+\kappa}} \mathbf{h}_k^{\text{LoS}} + \sqrt{\frac{1}{1+\kappa}} \mathbf{h}_k^{\text{NLoS}}\right), \quad (21)$$

where the LoS and NLoS superscripts represent the line-of-sight and non-line-of-sight components of the channel, respectively, the Rician factor is denoted by κ, and β_k represents the path-loss from user k to the base station.

User locations determine the LoS part of the channel $\mathbf{h}_k$. To be specific, let ϕ_k, θ_k represent the azimuth and elevation angles at which the signal of user k arrives at the HMIMO surface, as depicted in Fig. 2. The antenna elements are sequentially indexed row-wise by $m \in [1, M]$. Consequently, the position of the m-th element relative to the origin, as illustrated in Fig. 2, is [51]

$$\mathbf{u}_m = \left[\begin{array}{ccc} 0, & i(m) d_{\text{H}}, & j(m) d_{\text{V}} \end{array}\right]^{\top}, \quad (22)$$

where $i(m) = \text{mod}(m-1, M_y)$ and $j(m) = \lfloor (m-1)/M_y \rfloor$ are the horizontal and vertical indices of element m, respectively. It should be noted that the mod symbol expresses the modulus operation, and the notation $\lfloor \cdot \rfloor$ represents the floor function, which truncates the argument to the nearest integer less than or equal to the argument. $d = d_H = d_V$ denotes the space between adjacent antenna elements. The HMIMO surface steering vector $\mathbf{a}(\phi, \theta)$ can be expressed as [51]

$$\mathbf{a}(\varphi, \theta) = \left[e^{j\mathbf{k}(\varphi,\theta)^{\top}\mathbf{u}_1}, \dots, e^{\mathbf{j}\mathbf{k}(\varphi,\theta)^{\top}\mathbf{u}_N}\right]^{\top}, \quad (23)$$

where $\mathbf{k}(\varphi, \theta) \in \mathbb{R}^{3 \times 1}$ is the wave vector

$$\mathbf{k}(\varphi, \theta) = \frac{2\pi}{\lambda} [\cos(\theta)\cos(\varphi), \cos(\theta)\sin(\varphi), \sin(\theta)]^{\top}, \quad (24)$$

and

$$\begin{aligned} \sin(\varphi)\cos(\theta) &= \frac{y_k - y^{\text{BS}}}{d_k}, \\ \sin(\theta) &= \frac{z_k - z^{\text{BS}}}{d_k}. \end{aligned} \quad (25)$$

As a result of sub-half-wavelength setup in HMIMO surface, the $\mathbf{h}_k^{\text{NLoS}}$ is modelled as [51]

$$\mathbf{h}_k \sim \mathcal{CN}(\mathbf{0}, \mathbf{R}_k), \quad (26)$$

where the element $[\mathbf{R}_k]_{i,j} = \text{sinc}\left(\frac{2\|\mathbf{u}_i - \mathbf{u}_j\|}{\lambda}\right)$ and $\text{sinc}(x) = \frac{\sin(\pi x)}{\pi x}$.

The channel enters the near-field region when the HMIMO array's diameter expands to be sufficiently large. This condition can be quantified by the Rayleigh distance, which is expressed as $\frac{2D^2}{\lambda}$ [52], where D is the size of the aperture. Assuming the array has an aperture of $D = 0.5$ meters and operates at a central frequency of $f_c = 30$ GHz, the computed Rayleigh distance is roughly 50 meters. However, as elaborated in [52], the effective Rayleigh distance, which is defined from the perspective of the array gain that directly affects the transmission rate, is roughly 15 meters. As a result,

TABLE III
PERFORMANCE COMPARISON OF DIFFERENT NUMBER OF LAYERS (SNR= 20dB, $K = 5$, $M = 64$)

	2 layers	3 layers	4 layers
Sum rate [bps/Hz]	8.7	8.8	8.9

TABLE IV
THE PARAMETERS OF EACH LAYER IN THE GNN.

Name	Size	Activation Function
f_{nn}^0	$2M \times M$	relu
f_{nn}^l	$M \times M$	relu
f_{nn}^{L+1}	$M \times 2M$	/

it is reasonable to model the channel in the far-field region since the user is located 50 meters farther from the BS.

We use a 2-layer GNN, i.e. $L = 2$. A two-layer architecture for the GNN reduces the model complexity while achieving acceptable performance, as shown in Table III. This observation aligns with previous studies' findings [34], [35], showing shallow GNN architectures' high efficiency. The parameters of the FNNs $f_{\text{nn}}^0, f_{\text{nn}}^l$, and f_{nn}^{L+1} with $l = 1, 2$ are summarized in Table IV. We utilize shallow GNN and fewer neurons per layer, which is set based on the number of antennas to reduce the computation complexity for mutual coupling results from the ultra-dense distribution of antenna elements [51]. Simulation results demonstrate that the proposed GNN can attain substantial performance improvements, even when operating with a limited pilot length.

We deploy the designed network using the deep learning library Pytorch [50] and PyTorch Geometric [53]. The training of this neural network is conducted employing the Adam optimization algorithm, starting with a 10^{-4} learning rate. The $10,000$ samples are divided into 6000, 2000, 2000 for training, validation, and testing, respectively. At each training epoch, we transform batch data consisting of 1024 samples to compute the gradients and update the parameters of the neural network. The training process is stopped whenever the loss function on the verification data does not decrease over 10 consecutive epochs. Table V displays the parameters we used in our simulation.

Finally, we compare the performance of the proposed GNN with the following benchmarks:

- *Benchmark 1 - Perfect CSI plus WMMSE:* Given full CSI, the WMMSE technique, as introduced in [47], is implemented to address the problem of maximizing the sum rate.
- *Benchmark 2 - MMSE channel estimation plus WMMSE:* First, we calculate the channels through the MMSE estimation technique outlined in Section III, followed by optimizing the sum rate via the WMMSE method. The essential statistics requisite for the MMSE estimator are derived from the analysis of 10000 channel realizations.
- *Benchmark 3 - LS channel estimation plus WMMSE:* We first estimate the channels using the LS estimator presented in Section III. Subsequently, we proceed to maximize the sum rate by leveraging the WMMSE algorithm.

TABLE V
SIMULATION PARAMETERS.

Name	Value
Frequency f_c	30 GHz
Element space d	$\lambda/4$
Elements number M	64
Rician factor κ	10 dB

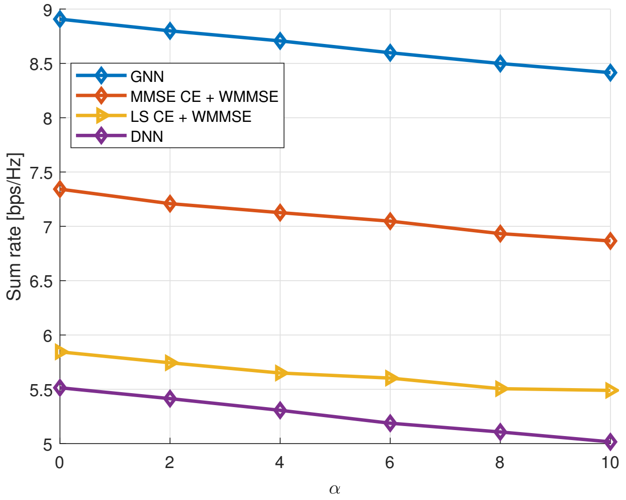

Fig. 6. Comparison of the sum rate performance versus propagation attenuation with $M = 64$, $K = 5$, and SNR $= 20$ dB.

- *Benchmark 4 - DNN:* We deploy a deep neural network that shares a nearly identical architecture with the proposed GNN for optimizing the sum rate, except that the message-passing process in the updating layer is removed to emphasize the superiority of the proposed GNN.
- *Benchmark 5 - Perfect CSI plus SVD:* Given full CSI, the problem of maximizing sum rate is addressed through the application of the Singular Value Decomposition (SVD) algorithm, which breaks down the channel matrix into simpler parts, and by choosing a singular value that is not zero, the signal transmission is established.

B. Simulation Results

We begin by examining the effect of propagation loss as the wave travels from the feed to the surface, where $\alpha = 0$ indicates that no propagation loss occurs. Fig. 6 demonstrates that as α increases, all methods experience a decline in performance. However, the proposed GNN approach continues to outperform the others in terms of sum rate. Given that the materials used to fabricate surfaces exhibit minimal loss, as discussed in [24], [54], we disregard power loss in subsequent simulations. We next evaluate the performance of various methods for sum rate optimization as a function of the antenna spacing, while considering the effects of mutual coupling, which are more significant when the spacing is less than half a wavelength. Fig. 7 illustrates the relationship between sum rate and antenna spacing, where it can be seen that the sum rate

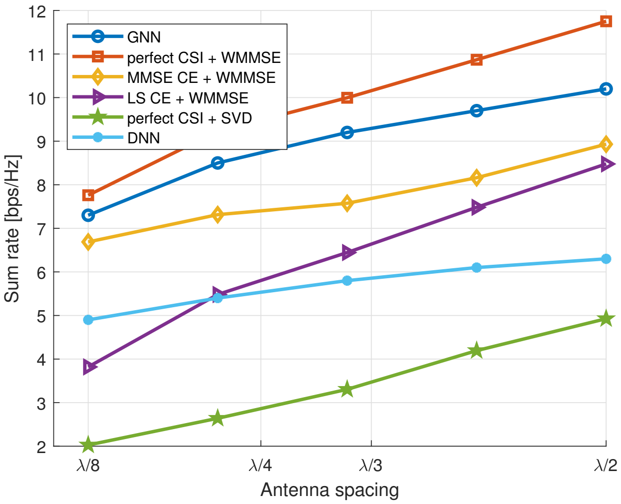

Fig. 7. Comparison of the sum rate performance versus antenna spacing with $M = 64$, $K = 5$, and SNR $= 20$ dB.

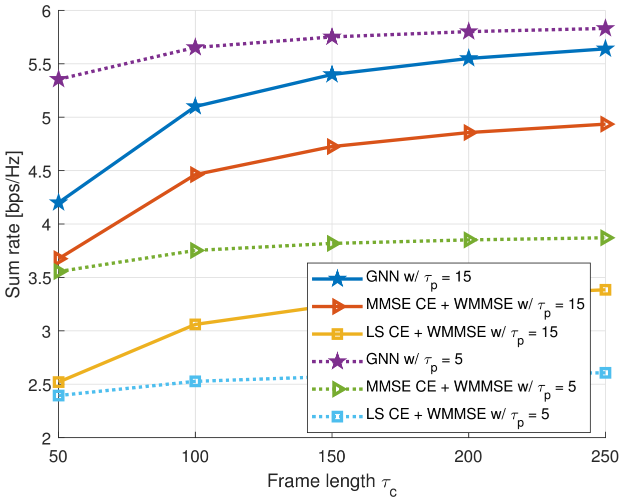

Fig. 9. Comparison of the sum rate performance versus the length of frame with $M = 64$, $K = 5$, and SNR $= 15$ dB.

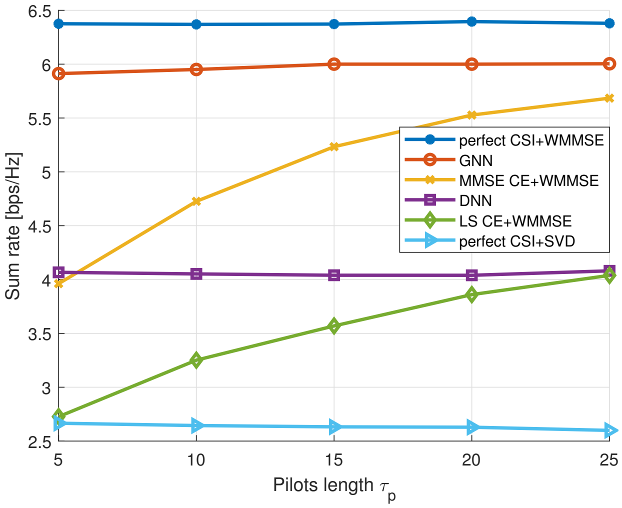

Fig. 8. Comparison of the sum rate performance versus the length of pilots with $M = 64$, $K = 5$, and SNR $= 15$ dB.

experiences an upward trend with increasing antenna spacing. Hence, as shown in Fig. 7, the mutual coupling has a negative impact on the performance of all methods for a fixed number of antennas, where increasing the antenna spacing leads to a larger aperture and reduced mutual coupling. Notably, the proposed GNN-based scheme achieves the highest sum rate across all antenna spacings, except for Benchmark 1, demonstrating its robustness to mutual coupling and effectiveness in optimizing the sum rate. The CE-based methods, represented by Benchmark 2 and Benchmark 3, exhibit moderate performance gains as the antenna spacing increases. Although the DNN-based approach maintains consistent performance across different antenna spacings, it remains inferior to the GNN-based and CE-based methods, indicating its relatively limited optimization effectiveness compared to the GNN-based scheme.

We analyze the effect of the uplink pilot length on the downlink sum rate performance with $M = 64$, $K = 5$, and considering SNR of 15 dB. Fig. 8 shows that our proposed GNN-based algorithm, the conventional perfect CSI plus WMMSE, the SVD, and the DNN strategies have approximately constant downlink sum rate performance as the pilot length increases, while the conventional LS and MMSE channel estimation plus WMMSE strategies experience a gradual increase. Our proposed GNN-based algorithm can achieve performance close to Bechmark 1 with fewer pilots (equal to the number of users) and still better than Bechmark 2 with four times more pilot length, emphasizing the superiority of reducing pilot overhead. The DNN method has a similar trend as the proposed algorithm but worse performance partly because it cannot extract graph structure information between different users or nodes as our proposed algorithm does.

We also investigate the influence of varying frame lengths on sum-rate performance. The first τ_p samples of a frame with total length τ_c are designated for pilot signals, while the remaining $\tau_c - \tau_p$ samples are allocated for data transmission. As illustrated in Fig. 9, all methods experience improved performance with longer frame lengths, as more resources can be devoted to data transmission. It is noteworthy that Benchmarks 2 and 3 benefit from longer pilot duration, whereas the proposed GNN demonstrates superior performance with fewer pilot samples. The sum rate is calculated based on $\tau_c - \tau_p$ excluding the pilot samples.

Next, we evaluate the performance of the HMIMO system when the BS is equipped with different numbers of antennas for $K = 5$, $\tau_p = 5$, and SNR $= 15$ dB. We can observe from Fig. 10 that as the number of antennas increases, the proposed GNN-based algorithm consistently outperforms the benchmarks except for the perfect CSI plus WMMSE algorithm. It is worth noting that the gap becomes more significant as the number of antennas increases. This is because our proposed method can extract channel features and suppress interference efficiently through the message-passing process in the update layers compared with other benchmark methods.

Then, we analyze the impact of SNR and diversely trained parameters on the sum rate, showcasing the generalization capabilities using $M = 64$, $K = 5$, and $\tau_p = 5$. Initially, the

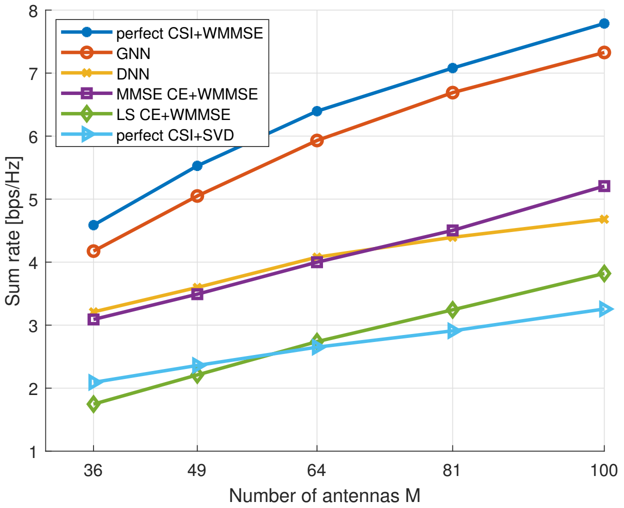

Fig. 10. Comparison of the sum rate performance versus the number of antennas with $K = 5$, $\tau_p = 5$, and SNR $= 15$ dB.

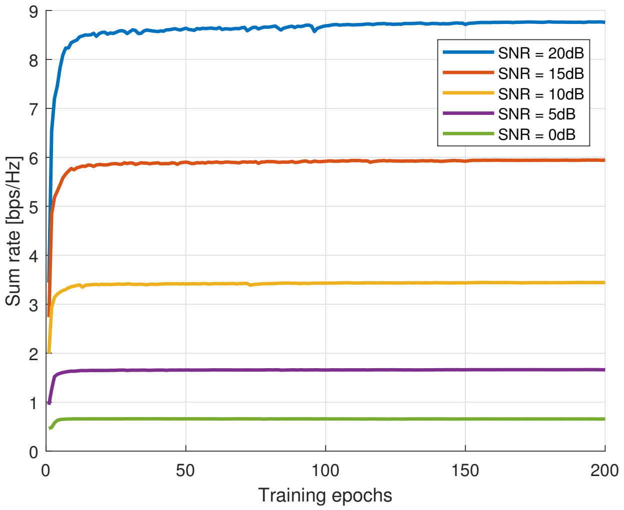

Fig. 13. Comparison of the sum rate performance versus the number of training epochs with $M = 64$, $K = 5$, and $\tau_p = 5$.

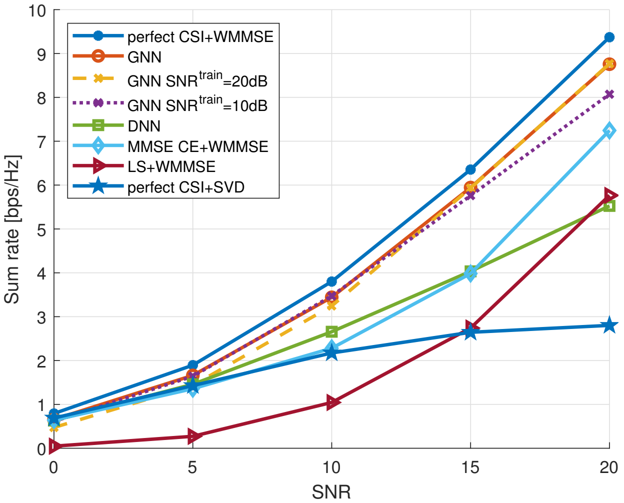

Fig. 11. Comparison of the sum rate performance versus SNR with $M = 64$, $K = 5$, and $\tau_p = 5$.

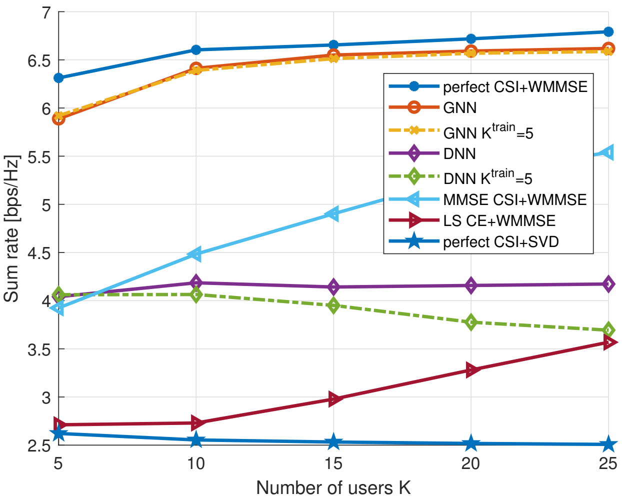

Fig. 12. Comparison of the sum rate performance versus the number of users with $M = 64$, SNR $= 15$ dB, $\tau_p = K$.

performance of the GNN-based algorithm and other benchmarks is plotted under identical train and test parameters. To evaluate generalization, we train two additional GNNs: one at SNR$=$ 10 dB and another at 20 dB. These trained neural networks are then tested under varying SNR conditions. Fig. 11 illustrates the results, where across all strategies, the performance improves with increasing the SNR. Notably, our proposed GNN-based approach consistently outperforms the benchmarks, except for Benchmark 1, across different SNR levels. Moreover, training the GNN at either 10 dB or 20 dB (while testing it at different SNRs) results in minimal performance degradation. This observation indicates the robust generalization capability of the proposed neural network model across various SNR conditions.

Additionally, we assess the system's performance for variable numbers of users, showcasing the GNN's robustness and generalization capability with parameters set to $M = 64$ and SNR$=$ 15 dB. In this simulation setup, the pilot length is equivalent to the number of users, denoted as K. In addition to the standard simulation with equal train and test parameters, we train other GNN and DNN models using $K = 5$ and evaluate their performance across varying numbers of users. The results, depicted in Fig. 12, illustrate notable trends. The proposed GNN-based algorithm demonstrates remarkable generalization across different user counts, consistently outperforming explicit channel estimation benchmarks. The DNN method notably shows a decline in sum rate with an increasing number of users when trained on $K = 5$ users, underscoring the robustness and superiority of our proposed GNN-based approach.

We conclude our analysis by assessing the convergence behavior of the proposed GNN model for $M = 64$, $K = 5$, and $\tau_p = K$. The sum rate plotted against the number of training epochs is presented in Fig. 13. Fig. 13 indicates that around 20 training epochs suffice to attain near-maximum performance across different SNRs. This result highlights the efficiency of the proposed GNN model in achieving near-optimal performance within a reasonably small number of training epochs across diverse SNR conditions.

V. CONCLUSION

Traditional communication system design heavily relies on accurate CSI for optimizing transmission schemes, which is challenging for HMIMO systems due to their large number of antenna elements. This paper introduces an alternative methodology to leverage a trainable neural network to produce beamformers at the BS directly from received pilots, bypassing explicit channel estimation. A versatile GNN architecture deciphers the mapping from the received pilots to per-user beamformers at the BS, enabling direct multi-user sum-rate maximization. Simulation results demonstrate the proposed neural network's efficiency in solving the optimization problem with significantly fewer pilots than conventional methods. Future research directions can encompass hardware imperfections such as wave-guide issues and practical deployment considerations, offering valuable insights into real-world HMIMO implementation.

REFERENCES

[1] T. Gong, P. Gavriilidis, R. Ji, C. Huang, G. C. Alexandropoulos, L. Wei, Z. Zhang, M. Debbah, H. V. Poor, and C. Yuen, "Holographic MIMO Communications: Theoretical Foundations, Enabling Technologies, and Future Directions," *IEEE Commun. Surv. Tutorials*, vol. 26, no. 1, pp. 196–257, 2024.

[2] K. Satyanarayana, M. El-Hajjar, P.-H. Kuo, A. Mourad, and L. Hanzo, "Hybrid Beamforming Design for Full-Duplex Millimeter Wave Communication," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 2, pp. 1394–1404, Feb. 2019.

[3] R. Deng, B. Di, H. Zhang, and L. Song, "HDMA: Holographic-Pattern Division Multiple Access," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 4, pp. 1317–1332, Apr. 2022.

[4] M. Cui and L. Dai, "Channel Estimation for Extremely Large-Scale MIMO: Far-Field or Near-Field?" *IEEE Transactions on Communications*, vol. 70, no. 4, pp. 2663–2677, Apr. 2022.

[5] J. Lee, G.-T. Gil, and Y. H. Lee, "Channel Estimation via Orthogonal Matching Pursuit for Hybrid MIMO Systems in Millimeter Wave Communications," *IEEE Transactions on Communications*, vol. 64, no. 6, pp. 2370–2386, Jun. 2016.

[6] X. Zhang, Z. Wang, H. Zhang, and L. Yang, "Near-Field Channel Estimation for Extremely Large-Scale Array Communications: A Model-Based Deep Learning Approach," *IEEE Communications Letters*, vol. 27, no. 4, pp. 1155–1159, Apr. 2023.

[7] Y. Han, S. Jin, C.-K. Wen, and X. Ma, "Channel Estimation for Extremely Large-Scale Massive MIMO Systems," *IEEE Wireless Communications Letters*, vol. 9, no. 5, pp. 633–637, May 2020.

[8] Y. Lu and L. Dai, "Near-Field Channel Estimation in Mixed LoS/NLoS Environments for Extremely Large-Scale MIMO Systems," *IEEE Transactions on Communications*, vol. 71, no. 6, pp. 3694–3707, Jun. 2023.

[9] X. Wei and L. Dai, "Channel Estimation for Extremely Large-Scale Massive MIMO: Far-Field, Near-Field, or Hybrid-Field?" *IEEE Communications Letters*, vol. 26, no. 1, pp. 177–181, Jan. 2022.

[10] W. Yu, Y. Shen, H. He, X. Yu, J. Zhang, and K. B. Letaief, "Hybrid Far- and Near-Field Channel Estimation for THz Ultra-Massive MIMO via Fixed Point Networks," in *GLOBECOM 2022 - 2022 IEEE Global Communications Conference*, Dec. 2022, pp. 5384–5389.

[11] P. Dong, H. Zhang, G. Y. Li, I. S. Gaspar, and N. NaderiAlizadeh, "Deep CNN-based channel estimation for mmWave massive MIMO systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 13, no. 5, pp. 989–1000, Sep. 2019.

[12] S. Gao, P. Dong, Z. Pan, and G. Y. Li, "Deep learning based channel estimation for massive MIMO with mixed-resolution ADCs," *IEEE Communications Letters*, vol. 23, no. 11, pp. 1989–1993, Nov. 2019.

[13] H. Li, P. Zhiwen, W. Bin, L. Nan, and Y. Xiaohu, "Channel Estimation for Reconfigurable-Intelligent-Surface-Aided Multiuser Communication Systems Exploiting Statistical CSI of Correlated RIS–User Channels," *IEEE Internet of Things Journal*, vol. 11, no. 5, pp. 8871–8881, Mar. 2024.

[14] C.-K. Wen, W.-T. Shih, and S. Jin, "Deep Learning for Massive MIMO CSI Feedback," *IEEE Wireless Communications Letters*, vol. 7, no. 5, pp. 748–751, Oct. 2018.

[15] Y. Jiang, W. Lin, W. Zhao, and C. Wang, "AcsiNet: Attention-Based Deep Learning Network for CSI Prediction in FDD MIMO Systems," *IEEE Wireless Communications Letters*, vol. 12, no. 3, pp. 471–475, Mar. 2023.

[16] R. He, B. Ai, G. Wang, M. Yang, C. Huang, and Z. Zhong, "Wireless Channel Sparsity: Measurement, Analysis, and Exploitation in Estimation," *IEEE Wireless Communications*, vol. 28, no. 4, pp. 113–119, Aug. 2021.

[17] N. Shlezinger, O. Dicker, Y. C. Eldar, I. Yoo, M. F. Imani, and D. R. Smith, "Dynamic Metasurface Antennas for Uplink Massive MIMO Systems," *IEEE Trans. Commun.*, vol. 67, no. 10, pp. 6829–6843, Oct. 2019.

[18] H. Wang, N. Shlezinger, Y. C. Eldar, S. Jin, M. F. Imani, I. Yoo, and D. R. Smith, "Dynamic Metasurface Antennas for MIMO-OFDM Receivers With Bit-Limited ADCs," *IEEE Trans. Commun.*, vol. 69, no. 4, pp. 2643–2659, Apr. 2021.

[19] L. You, J. Xu, G. C. Alexandropoulos, J. Wang, W. Wang, and X. Gao, "Energy Efficiency Maximization of Massive MIMO Communications With Dynamic Metasurface Antennas," *IEEE Trans. Wireless Commun.*, vol. 22, no. 1, pp. 393–407, Jan. 2023.

[20] G. Chen, R. Zhang, H. Zhang, C. Miao, Y. Ma, and W. Wu, "Energy-Efficient Beamforming for Downlink Multi-User Systems with Dynamic Metasurface Antennas," *IEEE Communications Letters*, pp. 1–1, 2024.

[21] J. Xu, L. You, G. C. Alexandropoulos, X. Yi, W. Wang, and X. Gao, "Near-Field Wideband Extremely Large-Scale MIMO Transmissions With Holographic Metasurface-Based Antenna Arrays," *IEEE Transactions on Wireless Communications*, vol. 23, no. 9, pp. 12 054–12 067, Sep. 2024.

[22] B. Di, "Reconfigurable Holographic Metasurface Aided Wideband OFDM Communications Against Beam Squint," *IEEE Trans. Veh. Technol.*, vol. 70, no. 5, pp. 5099–5103, May 2021.

[23] R. Deng, B. Di, H. Zhang, Y. Tan, and L. Song, "Reconfigurable Holographic Surface: Holographic Beamforming for Metasurface-Aided Wireless Communications," *IEEE Trans. Veh. Technol.*, vol. 70, no. 6, pp. 6255–6259, Jun. 2021.

[24] ——, "Reconfigurable Holographic Surface-Enabled Multi-User Wireless Communications: Amplitude-Controlled Holographic Beamforming," *IEEE Trans. Wireless Commun.*, vol. 21, no. 8, pp. 6003–6017, Aug. 2022.

[25] Q. Li, M. El-Hajjar, Y. Sun, I. Hemadeh, Y. Tsai, A. Shojaeifard, and L. Hanzo, "Achievable Rate Analysis of Intelligent Omni-Surface Assisted NOMA Holographic MIMO Systems," *IEEE Transactions on Vehicular Technology*, vol. 73, no. 5, pp. 7400–7405, May 2024.

[26] G. Chen, R. Zhang, H. Ren, W. Yuan, C. Miao, and W. Wu, "Hybrid Beamforming Design with Overlapped Subarrays for Massive MIMO-ISAC Systems," in *GLOBECOM 2023 - 2023 IEEE Global Communications Conference*, Dec. 2023, pp. 528–533.

[27] R. Zhang, L. Cheng, W. Zhang, X. Guan, Y. Cai, W. Wu, and R. Zhang, "Channel Estimation for Movable-Antenna MIMO Systems Via Tensor Decomposition," *IEEE Wireless Communications Letters*, pp. 1–4, 2024.

[28] R. Zhang, L. Cheng, S. Wang, Y. Lou, Y. Gao, W. Wu, and D. W. K. Ng, "Integrated Sensing and Communication With Massive MIMO: A Unified Tensor Approach for Channel and Target Parameter Estimation," *IEEE Transactions on Wireless Communications*, vol. 23, no. 8, pp. 8571–8587, Aug. 2024.

[29] Z. Linfu, P. Zhiwen, A. A. Ahmad, H. Dahrouj, and M. El-Hajjar, "Joint Beamforming and Combining Design for mmWave Integrated Access and Backhaul Networks," *IEEE Open Journal of the Communications Society*, vol. 5, pp. 503–513, 2024.

[30] Q. Li, M. El-Hajjar, I. Hemadeh, A. Shojaeifard, and L. Hanzo, "Low-Overhead Channel Estimation for RIS-Aided Multi-Cell Networks in the Presence of Phase Quantization Errors," *IEEE Transactions on Vehicular Technology*, vol. 73, no. 5, pp. 6626–6641, May 2024.

[31] Z. Linfu, P. Zhiwen, J. Huilin, L. Nan, and Y. Xiaohu, "Deep Learning Based Downlink Channel Covariance Estimation for FDD Massive MIMO Systems," *IEEE Communications Letters*, vol. 25, no. 7, pp. 2275–2279, Jul. 2021.

[32] Y. Yang, F. Gao, G. Y. Li, and M. Jian, "Deep learning-based downlink channel prediction for FDD massive MIMO system," *IEEE Communications Letters*, vol. 23, no. 11, pp. 1994–1998, Nov. 2019.

[33] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," *Neural Networks*, vol. 2, no. 5, pp. 359–366, Jan. 1989.

[34] Y. Shen, Y. Shi, J. Zhang, and K. B. Letaief, “Graph Neural Networks for Scalable Radio Resource Management: Architecture Design and Theoretical Analysis,” *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 1, pp. 101–115, Jan. 2021.

[35] T. Jiang, H. V. Cheng, and W. Yu, “Learning to Reflect and to Beamform for Intelligent Reflecting Surface With Implicit Channel Estimation,” *IEEE J. Sel. Areas Commun.*, vol. 39, no. 7, pp. 1931–1945, Jul. 2021.

[36] Z. Zhang, T. Jiang, and W. Yu, “Learning Based User Scheduling in Reconfigurable Intelligent Surface Assisted Multiuser Downlink,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 16, no. 5, pp. 1026–1039, Aug. 2022.

[37] X. Feng, M. El-Hajjar, C. Xu, and L. Hanzo, “Near-Instantaneously Adaptive Learning-Assisted and Compressed Sensing-Aided Joint Multi-Dimensional Index Modulation,” *IEEE Open Journal of Vehicular Technology*, vol. 4, pp. 893–912, 2023.

[38] ——, “Reduced Complexity Learning-Assisted Joint Channel Estimation and Detection of Compressed Sensing-Aided Multi-Dimensional Index Modulation,” *IEEE Open Journal of Vehicular Technology*, vol. 5, pp. 78–94, 2024.

[39] T. Lin and Y. Zhu, “Beamforming Design for Large-Scale Antenna Arrays Using Deep Learning,” *IEEE Wireless Commun. Lett.*, vol. 9, no. 1, pp. 103–107, Jan. 2020.

[40] Z. Huang, Z. Wang, and S. Chen, “Sub-6GHz Assisted mmWave Hybrid Beamforming with Heterogeneous Graph Neural Network,” *IEEE Trans. Commun.*, pp. 1–1, 2024.

[41] J. Yang, W. Zhu, S. Sun, X. Li, X. Lin, and M. Tao, “Deep Learning for Joint Design of Pilot, Channel Feedback, and Hybrid Beamforming in FDD Massive MIMO-OFDM Systems,” *IEEE Communications Letters*, vol. 28, no. 2, pp. 313–317, Feb. 2024.

[42] J. Liu, K. Mei, X. Zhang, D. McLernon, D. Ma, J. Wei, and S. A. R. Zaidi, “Fine Timing and Frequency Synchronization for MIMO-OFDM: An Extreme Learning Approach,” *IEEE Transactions on Cognitive Communications and Networking*, vol. 8, no. 2, pp. 720–732, Jun. 2022.

[43] L. You, X. Gao, G. Y. Li, X.-G. Xia, and N. Ma, “BDMA for Millimeter-Wave/Terahertz Massive MIMO Transmission With Per-Beam Synchronization,” *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 7, pp. 1550–1563, Jul. 2017.

[44] Y. Feng, H. Shen, W. Lu, N. Zhao, and A. Nallanathan, “Iterative Joint Frequency Synchronization and Channel Estimation for Uplink Massive MIMO,” *IEEE Internet of Things Journal*, vol. 11, no. 17, pp. 28 891–28 905, Sep. 2024.

[45] P. Sabeti, A. Farhang, N. Marchetti, and L. Doyle, “Frequency Synchronization for OFDM-Based Massive MIMO Systems,” *IEEE Transactions on Signal Processing*, vol. 67, no. 11, pp. 2973–2986, Jun. 2019.

[46] Ö. Özdogan, E. Björnson, and E. G. Larsson, “Massive MIMO With Spatially Correlated Rician Fading Channels,” *IEEE Transactions on Communications*, vol. 67, no. 5, pp. 3234–3250, May 2019.

[47] Q. Shi, M. Razaviyayn, Z.-Q. Luo, and C. He, “An Iteratively Weighted MMSE Approach to Distributed Sum-Utility Maximization for a MIMO Interfering Broadcast Channel,” *IEEE Transactions on Signal Processing*, vol. 59, no. 9, pp. 4331–4340, Sep. 2011.

[48] K. Xu, W. Hu, J. Leskovec, and S. Jegelka, “How Powerful are Graph Neural Networks?” in *International Conference on Learning Representations*, Sep. 2018.

[49] V. P. Dwivedi, C. K. Joshi, A. T. Luu, T. Laurent, Y. Bengio, and X. Bresson, “Benchmarking Graph Neural Networks,” *ArXiv*, Dec. 2022.

[50] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Köpf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, and S. Chintala, “PyTorch: An Imperative Style, High-Performance Deep Learning Library,” in *Neural Information Processing Systems*, Dec. 2019.

[51] E. Bjornson and L. Sanguinetti, “Rayleigh Fading Modeling and Channel Hardening for Reconfigurable Intelligent Surfaces,” *IEEE Wireless Commun. Lett.*, vol. 10, no. 4, pp. 830–834, Apr. 2021.

[52] M. Cui, L. Dai, R. Schober, and L. Hanzo, “Near-Field Wideband Beamforming for Extremely Large Antenna Arrays,” *arXiv*, Dec. 2021.

[53] M. Fey and J. E. Lenssen, “Fast Graph Representation Learning with PyTorch Geometric,” *ArXiv*, Mar. 2019.

[54] D. R. Smith, O. Yurduseven, L. P. Mancera, P. Bowen, and N. B. Kundtz, “Analysis of a Waveguide-Fed Metasurface Antenna,” *Phys. Rev. Appl.*, vol. 8, no. 5, p. 054048, Nov. 2017.