

LLM-Empowered Near-Field Communications for Low-Altitude Economy

Zhuo Xu, Tianyue Zheng, and Linglong Dai, *Fellow, IEEE*

Abstract—The low-altitude economy (LAE) has recently received widespread attention from both academia and industry. To facilitate and support the successful implementation of the LAE, we fortunately find that the LAE and near-field communications in extremely large-scale MIMO (XL-MIMO) systems are a natural combination. Specifically, the LAE can utilize the near-field beamfocusing characteristic to accurately focus the beam energy to the positions of different unmanned aerial vehicles, and utilize the new distance dimension to further enhance the entire spectrum efficiency. However, most existing works on near-field communications only consider the ideal scenario in a horizontal plane and how to efficiently achieve near-field communications for LAE is still a blank in the literature and faces several challenges. To fill in this blank, inspired by the powerful large language models (LLM) which can act as a general wireless communications optimization solver, in this paper, we first apply LLM to solve the spectrum efficiency maximization problem of near-field communications for LAE. Specifically, our proposed LLM-based scheme can accurately distinguish far-field and near-field users and achieve joint optimization of precoding and power allocation through elaborately designing adapters and finetuning the pretrained GPT-2. Simulation results substantiate the efficacy and excellence of our proposed scheme compared to the existing benchmark schemes.

Index Terms—The low-altitude economy (LAE), large language models (LLM), near-field communications.

I. INTRODUCTION

In recent years, the low-altitude economy (LAE) has attracted significant attention from both industry and academia across several countries [1]–[4]. In LAE, several flying equipment such as unmanned aerial vehicles (UAVs) are employed to foster various applications, such as urban transportation, logistics, agriculture, and tourism¹ [5]. From the perspective of wireless communications, the LAE networks utilize UAVs to meet different communication task requirements, where the airspace offers greater freedom of movement compared to terrestrial networks. To ensure the successful implementation

of LAE, the stable and safe operation of UAVs is particularly important. Specifically, the UAVs require seamless wireless communication connections and accurate trajectory planning and tracking.

To facilitate and support the successful implementation of the LAE, extremely large-scale MIMO (XL-MIMO) has been considered as one potential key technology [6]–[9]. Different from massive MIMO systems, XL-MIMO deploys extremely large-scale antenna arrays (ELAA), which could achieve higher spatial resolution and multiplexing gain. Besides, in XL-MIMO systems, with the increasing number of antennas at base station (BS), the near-field region is enlarged and the near-field channel should be accurately modeled by the spherical-wave model rather than the planar-wave model applied in far-field. For instance, the near-field region of a ELAA with 256 antennas at 30 GHz is approximately 326.5 meters [10], which aligns with practical urban cell sizes. The near-field channel is related to both angle and distance, and it presents additional focusing ability in distance domain, which can concentrate the beam energy at specific locations like a flashlight [11]. Consequently, the spherical-wave model based near-field communications will bring great opportunities for XL-MIMO systems.

Fortunately, we find that the LAE and near-field communications in XL-MIMO systems match naturally. Compared to ground users, the UAVs are closer to BS antenna arrays with practical height, therefore they are more probable to be within the near-field region and thus benefit from near-field communications. Specifically, the LAE can utilize the near-field beamfocusing characteristic to accurately focus the beam energy to the positions of different UAVs, thereby mitigating their interference [12], [13]. Besides, unlike the classical far-field spatial division multiple access (SDMA), the near-field location division multiple access (LDMA) can serve UAVs at identical angle but different distances simultaneously, which enhances the entire spectrum efficiency of the LAE network through extra distance dimension [14].

However, most existing works on near-field communications only consider the ideal scenario in a horizontal plane for ground equipments. While precoding and power allocation in near-field communications have been studied, the LAE scenario introduces unique challenges not fully addressed in prior work as the scenarios and models of near-field communication for LAE become quite complex. Consequently, how to efficiently solve the spectrum efficiency maximization problem of near-field communications for LAE is still a blank in the literature. Specifically, as the near-field multi-user communications for LAE requires the joint optimization of precoding and power allocation for both UAV and ground users, more parameters need to be optimized. Besides, the transition of the system

This work was supported in part by the National Science Fund for Distinguished Young Scholars (Grant No. 62325106), in part by the National Key Research and Development Program of China (Grant No. 2023YFB3811503), and in part by the National Natural Science Foundation of China (Grant No.62031019). (Corresponding author: Linglong Dai.)

Z. Xu, T. Zheng, and L. Dai are with the Department of Electronic Engineering, Tsinghua University, Beijing 100084, China, and also with the State Key Laboratory of Space Network and Communications, Tsinghua University, Beijing 100084, China (e-mails: xz23@mails.tsinghua.edu.cn; zhengty22@mails.tsinghua.edu.cn; daill@tsinghua.edu.cn). L. Dai is also with Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA 02139, United States (Email: daill@mit.edu). Z. Xu and T. Zheng should be considered co-first authors as they contributed equally to the project.

¹It should be noted that in addition to UAVs, there are other types of flying equipment such as electric vertical take-off and landing aircraft (eVTOL) used in LAE. For simplicity, UAVs are uniformly used instead in this paper.

model in LAE correspondingly leads to different near-field regions for UAVs and ground users, where the far-field, near-field, and far-field regions appear in sequence with the distance from the BS increasing. Thus, it is necessary to distinguish between far and near-field users of both UAVs and ground users and group them for precoding and power allocation optimization, which is however not easy to solve.

Inspired by the recent great progress of large language models (LLM) which has the potential to act as a general optimization solver for diverse optimization problems in wireless communications [15]–[21], in this paper, we first apply LLM to solve the spectrum efficiency maximization problem of near-field communications for LAE. Our contributions are summarized as follows²:

- **New application scenario:** It is pointed out that the LAE and near-field communications can match ingeniously. The UAVs in LAE can utilize the near-field beamfocusing characteristics to achieve precise beam energy focusing, which could focus energy at their locations based on both angle θ and distance r . Besides, they can utilize exploit extra distance domain resources to further improve spatial multiplexing and spectrum efficiency. Furthermore, they are closer to BS antenna arrays with practical height, thus enabling more UAVs to benefit from near-field communications. To the best of our knowledge, this paper is the first to investigate on near-field communications for LAE.
- **New system model:** Owing to the new application scenario of near-field communications for LAE, unlike the existing works on near-field communications, new system model is adopted, where the height and tilting angle of the BS antennas are considered. Further, as the new system model has led to changes in near-field region of the horizontal plane, we analyze and propose a concept called the effective near-field region to redefine it, where the horizontal plane is sequentially divided into far-field, near-field, and far-field with the distance from the BS increasing.
- **New technology:** To solve the challenging spectrum efficiency maximization problem of near-field communications for LAE, we apply the novel and powerful LLM to achieve the joint optimization of precoding and power allocation. With powerful reasoning and inference capabilities, LLM excels to tackle the complex, non-convex optimization problem, leveraging their scalability and adaptability to achieve superior performance. Specifically, we propose the LLM-empowered near-field multi-user communications scheme, which can jointly distinguish far-field and near-field users and design multi-user precoding matrix. With elaborately designed adapters, by finetuning the pretrained GPT-2, the proposed method achieves near-optimal performance.

It should be emphasized that the three contribution points are closely related: we first introduce a **new application scenario** for LAE, exploiting its synergy with near-field com-

munications; this motivates a **new XL-MIMO system model** tailored to LAE's spatial characteristics, incorporating BS height and tilting angle; and to tackle the unique optimization challenges posed by this scenario and model, we propose a **new LLM-based technology** to maximize spectral efficiency efficiently.

The rest of this paper is organized as follows. Section II introduces the system model. In Section III, the effective near-field region of horizontal plane is analyzed. The proposed LLM-empowered near-field communications scheme is discussed in Section IV. Simulation results and conclusions are presented in Section V and Section VI.

Notation: $\mathbb{C}$ denotes the set of complex numbers; Upper-case and lower-case boldface letters represent matrices and vectors; $(\cdot)^{-1}$, $(\cdot)^T$, $(\cdot)^H$ denote the inverse, transpose, and conjugate transpose, respectively; $|\cdot|$ denotes the absolute operator; $\mathcal{CN}(\mu, \Sigma)$ denotes the Gaussian distribution with mean μ and covariance Σ ; $\mathbf{I}$ denotes an identity matrix.

II. SYSTEM MODEL

A downlink XL-MIMO communication system is considered, where a BS deployed with a N -element uniform linear array (ULA) serves K single-antenna users. It should be pointed out that different from simple XL-MIMO system modeling in the literature, we adopt more practical scenarios, where the height and tilting angle of the BS antenna array are considered. For simplicity, we adopt a simplified cartesian coordinate model, i.e., the x - z plane in the x - y - z coordinate. We emphasize that the simplified model in x - z plane differs from conventional model in x - y plane. The model in x - z plane involves the BS height h_B and tilting angle θ_{tit} , which is more practical for urban deployments and should be carefully considered for LAE's near-field communications. It is also the main difference of the system model in the new scenarios. Thus, in this paper we simplify analysis by focusing on horizontal distance and vertical height, critical for LAE's near-field communications, while ignoring y -direction variation for concise expression. In contrast, the conventional x - y plane model fail to consider the effect of BS height h_B and tilting angle θ_{tit} . For clear comparison, we also ignore the y plane for conventional model, which is illustrated in Fig. 1.

As illustrated in Fig. 1, we denote $(x_k, 0)$, (x'_k, h'_k) and $(0, h_B)$ as the coordinate of the ground users, UAV, and the center of the ULA, respectively³. In addition, θ_{tit} denotes the boresight angle of the BS, which is also called as tilting angle [22]. Besides, θ_k and θ'_k denote the vertical angle of the ground users and UAV, i.e., $\theta_k = \tan^{-1}(\frac{h_B}{x_k})$ and $\theta'_k = \tan^{-1}(\frac{h'_k - h_B}{x_k})$. Without loss of generality, we use (x_k, h_k) to represent the coordinate of user k uniformly and the vertical angle of user k is $\theta_k = \tan^{-1}(\frac{|h_k - h_B|}{x_k})$.

Let $\mathbf{h}_k \in \mathbb{C}^{N \times 1}$ denotes the downlink channel of user k , then its received signal could be expressed as

$$y_k = \mathbf{h}_k^H \mathbf{W} \mathbf{p}_s + n, \quad (1)$$

³For simplicity, we adopt a quasi-static environment where UAVs and ground users' positions are fixed during the channel coherence time, as is typical for precoding and power allocation in near-field communications.

²Simulation codes are provided to reproduce the results in this paper: <http://oa.ee.tsinghua.edu.cn/dailong/publications/publications.html>.

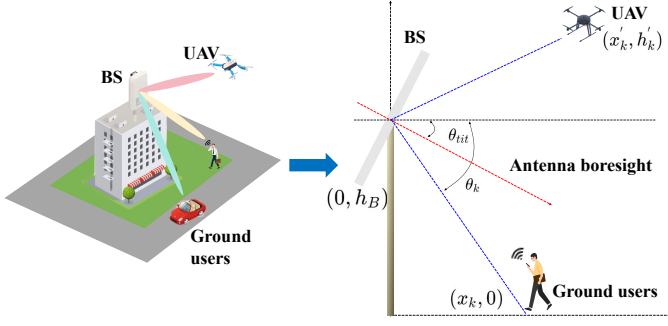


Fig. 1. Illustration of the system model and channel model.

where $\mathbf{W} = [\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_K] \in \mathbb{C}^{N \times K}$ denotes the transmit precoding matrix, $\mathbf{P} = \text{diag}\{\sqrt{P_1}, \sqrt{P_2}, \dots, \sqrt{P_K}\} \in \mathbb{C}^{K \times K}$ denotes the power allocation matrix satisfying $\sum_{k=1}^K P_k \leq P$, P denotes the maximum transmit power, $\mathbf{s}$ denotes the power-normalized transmitted signal with constraint $\mathbb{E}[\mathbf{s}\mathbf{s}^H] = \mathbf{I}$ and n denotes the received noise following $\mathcal{CN}(0, \sigma^2)$, where σ^2 denotes the variance of the noise.

Generally, the channel model could be separated into far and near-field model according to the electromagnetic wave propagation characteristics. The Rayleigh distance is usually considered as the boundary, which is defined as $R = \frac{2D^2}{\lambda}$, where D denotes the array aperture and λ represents the carrier wavelength [23]. In classical MIMO systems, the number of the array elements is not large and the Rayleigh distance is negligible, therefore planar-wave propagation model is applied to model the far-field channel. The Saleh-Valenzuela model is widely considered, and the far-field channel $\mathbf{h}_k^{\text{far}}$ could be represented as

$$\mathbf{h}_k^{\text{far}} = \sqrt{N}\alpha_0\mathbf{a}(\theta_0) + \sqrt{\frac{N}{L}} \sum_{l=1}^L \alpha_l\mathbf{a}(\theta_l), \quad (2)$$

where $\alpha_0, \theta_0, \alpha_l, \theta_l, L$ denote the complex gain and the angle-of-departure (AoD) of the line-of-sight (LoS) path, the complex gain and the AoD of the non-line-of-sight (NLoS) paths, and the total number of the NLoS paths, respectively.

For the ULA, the beam steering vector $\mathbf{a}(\theta)$ could be represented as

$$\mathbf{a}(\theta) = \frac{1}{\sqrt{N}} [1, e^{j\pi \sin \theta}, \dots, e^{j(N-1)\pi \sin \theta}]^T, \quad (3)$$

where $\theta \in [-\pi/2, \pi/2]$ denotes the physical direction.

In XL-MIMO systems, as the number of BS antennas increases, the near-field region becomes larger accordingly and the spherical-wave propagation model should be used to characterize the near-field channel $\mathbf{h}_k^{\text{near}}$ as [10]:

$$\mathbf{h}_k^{\text{near}} = \sqrt{N}\alpha_0\mathbf{b}(\theta_0, r_0) + \sqrt{\frac{N}{L}} \sum_{l=1}^L \alpha_l\mathbf{b}(\theta_l, r_l), \quad (4)$$

Moreover, $\mathbf{b}(\theta, r)$ denotes the near-field beam steering vector, where $\theta = \theta_k - \theta_{tit}$ in our model. Unlike far-field beam steering vector focusing the beam energy towards specific directions, the near-field beam steering vector could focus the beam energy on specific locations, which is also called the

near-field beamfocusing vector [11]. For ULA, the near-field beamfocusing vector $\mathbf{b}(\theta, r)$ could be represented as

$$\mathbf{b}(\theta, r) = \frac{1}{\sqrt{N}} [e^{-j\frac{2\pi}{\lambda}(r_0-r)}, \dots, e^{-j\frac{2\pi}{\lambda}(r_{N-1}-r)}]^T, \quad (5)$$

where r_n and r denote the distance between the user and the n -th element and center of the BS antenna, respectively. The r_n could be expressed as

$$\begin{aligned} r_n &= \sqrt{r^2 - 2ndr \sin \theta + n^2d^2} \\ &\stackrel{(a)}{\approx} r - nd \sin \theta + \frac{n^2d^2 \cos^2 \theta}{2r}, \end{aligned} \quad (6)$$

where approximation (a) is the Fresnel approximation, which is derived by $\sqrt{1+x} = 1 + \frac{x}{2} - \frac{x^2}{8} + \mathcal{O}(x^3)$. It can be obtained from (4) and (5) that the near-field channel is determined by both the angle and distance. Unlike the simple XL-MIMO system modeling in [14], the near-field region of horizontal plane becomes different in our adopted practical model, which is analyzed in the following section.

III. EFFECTIVE NEAR-FIELD REGION

A concept called effective near-field region (ENFR) is proposed in this section to define the near-field region of horizontal plane in our adopted practical XL-MIMO system model. Specifically, like the definition in [24] and [25], we define the ENFR through the beamforming gain loss. In the ENFR, the beamforming gain loss adopting far-field beamforming vectors is lower than the predefined threshold Δ , i.e. $1 - |\mathbf{b}(\theta, r)^H \mathbf{a}(\theta)| \geq \Delta$, where $\mathbf{a}(\theta) = \frac{1}{\sqrt{N}} [1, e^{j\pi \sin \theta}, \dots, e^{j(N-1)\pi \sin \theta}]^T$ denotes the far-field beamforming vector for the ULA. Therefore, the ENFR can be defined through the following lemma.

Lemma 1. For our adopted practical XL-MIMO system model discussed in Section II, the ENFR could be expressed as

$$I_{\text{ENFR}} = \left[\frac{h_B}{\tan \theta_k^-}, \frac{h_B}{\tan \theta_k^+} \right], \quad (7)$$

where θ_k^- and θ_k^+ are two solutions for $\sin \theta_k \cos^2(\theta_k - \theta_{tit}) = \frac{2h_B\beta_\Delta^2\lambda}{N^2d^2}$. β_Δ is the solution for $|G(\beta_\Delta)| = 1 - \Delta$, where $|G(\beta)| = |\int_0^\beta e^{-j\frac{1}{2}\pi t^2} dt|/\beta$.

Proof. First, we define $\mu(\theta, r) = |\mathbf{b}(\theta, r)^H \mathbf{a}(\theta)|$, which could be further expressed as

$$\begin{aligned} \mu(\theta, r) &= \left| \frac{1}{N} \sum_{n=-(N-1)/2}^{(N-1)/2} e^{j\pi n^2 \frac{d^2 \cos^2 \theta}{\lambda r}} \right| \\ &= |F(x)|, \end{aligned} \quad (8)$$

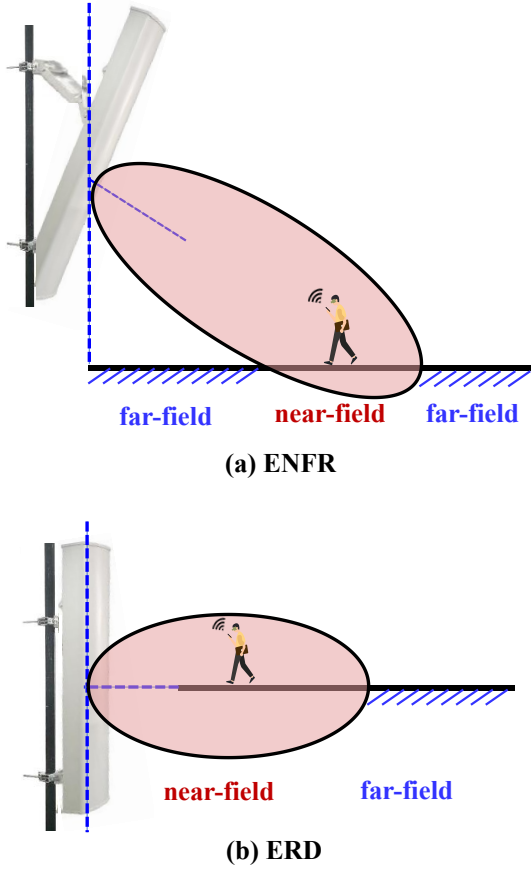


Fig. 2. Comparison between the ENFR and ERD: (a) ENFR; (b) ERD

where $x = \frac{d^2 \cos^2 \theta}{\lambda r}$. Besides, $F(x)$ could be expressed as

$$\begin{aligned} F(x) &= \left| \frac{1}{N} \sum_{n=-(N-1)/2}^{(N-1)/2} e^{j\pi n^2 x} \right| \\ &\approx \left| \frac{1}{N} \int_{-N/2}^{N/2} e^{j\pi n^2 x} dn \right| \\ &\approx \frac{2}{\sqrt{2xN}} \int_0^{\sqrt{2xN}/2} e^{j\frac{\pi}{2}t^2} dt \\ &= G(\beta), \end{aligned} \quad (9)$$

where $\beta = \frac{\sqrt{2xN}}{2} = \sqrt{\frac{N^2 d^2 \cos^2 \theta}{2\lambda r}}$. Thus, to satisfy $1 - |\mathbf{b}(\theta, r)^H \mathbf{a}(\theta)| \geq \Delta$, $\beta \geq \beta_\Delta$ is needed, where $|G(\beta_\Delta)| = 1 - \Delta$.

Therefore, we can substitute $\tan \theta_k = \frac{h_B}{x_k}$ and $r = \sqrt{x_k^2 + h_B^2}$ into $\beta_\Delta = \sqrt{\frac{N^2 d^2 \cos^2 \theta}{2\lambda r}}$, then we can get $\sin \theta_k \cos^2(\theta_k - \theta_{tit}) = \frac{2h_B \beta_\Delta^2 \lambda}{N^2 d^2}$. By solving it, we can get θ_k^- and θ_k^+ , then (7) can be obtained and the proof is completed. $\square$

From **Lemma 1**, we can get that unlike existing works distinguishing the far and near-field regions according to effective Rayleigh distance (ERD), the ENFR in our adopted practical model divides the entire space into three regions [25]. Specifically, the comparison between the ERD and ENFR is

illustrated in Fig. 2. It is shown that the horizontal plane is sequentially divided into far-field, near-field, and far-field with the horizontal distance from the BS increasing. When the ground users are located in the ENFR, they can be considered as near-field users that can benefit from near-field beamfocusing.

Due to our consideration the height and tilting angle of the BS antenna and the ENFR of horizontal plane, the multi-user spectrum efficiency maximization problem in XL-MIMO systems will be more complex. Incorporating BS height and tilting angle introduces challenges absent in ideal horizontal plane models, such as complex near-field boundaries varying with θ_{tit} and h_B , necessitating adaptive classification and precoding for LAE's user distribution. Thus, how to apply LLM for empowering near-field multi-user communications in XL-MIMO systems is a critical problem, which is analyzed in the following section.

IV. LLM-EMPOWERED NEAR-FIELD MULTI-USER COMMUNICATIONS

In this section, the spectrum efficiency maximization problem of near-field multi-user communications is first formulated. Then, we illustrate the proposed model, as presented in Fig. 3. Specifically, the proposed model can jointly distinguish between far and near-field and design multi-user precoding matrix, which is elaborated respectively as follows. Next, we conclude the advantages of the proposed LLM-based scheme over other conventional solvers.

A. Problem Formulation

Based on (1), the signal-to-interference-plus-noise ratio (SINR) of user k could be represented as

$$\text{SINR}_k = \frac{P_k |\mathbf{h}_k^H \mathbf{w}_k|^2}{\sum_{j \neq k} P_j |\mathbf{h}_k^H \mathbf{w}_j|^2 + \sigma^2}. \quad (10)$$

Then, the achievable rate of the k -th user is:

$$R_k = \log_2(1 + \text{SINR}_k). \quad (11)$$

Thus, the spectrum efficiency maximization problem of near-field multi-user communications could be expressed as:

$$\begin{aligned} \max_{\{\mathbf{W}, \mathbf{P}\}} \quad & \sum_k R_k = \sum_k \log_2(1 + \text{SINR}_k) \\ \text{s.t.} \quad & C_1 : \sum_{k=1}^K P_k \leq P, \\ & C_2 : P_k \geq 0, \\ & C_3 : \alpha_N \leq \alpha_c, \\ & C_4 : R_k \geq R_{\min}, \\ & C_5 : \|\mathbf{w}_k\|^2 = 1. \end{aligned} \quad (12)$$

where α_c denotes the pre-set constant within $[0, 1]$ and α_N denotes the power allocation factor for the near-field users, i.e., the transmit power allocated to near-field users is $P_N = \alpha_N P \leq \alpha_c P$. Besides, $R_{\min}$ denotes the minimum data rate for each user. The constraint C_1 , C_2 and C_3 are the limitations of transmit power. The primary purpose of C_3 to enforce a flexible power allocation mechanism that accounts for the distinct propagation properties of near-field and far-field regions

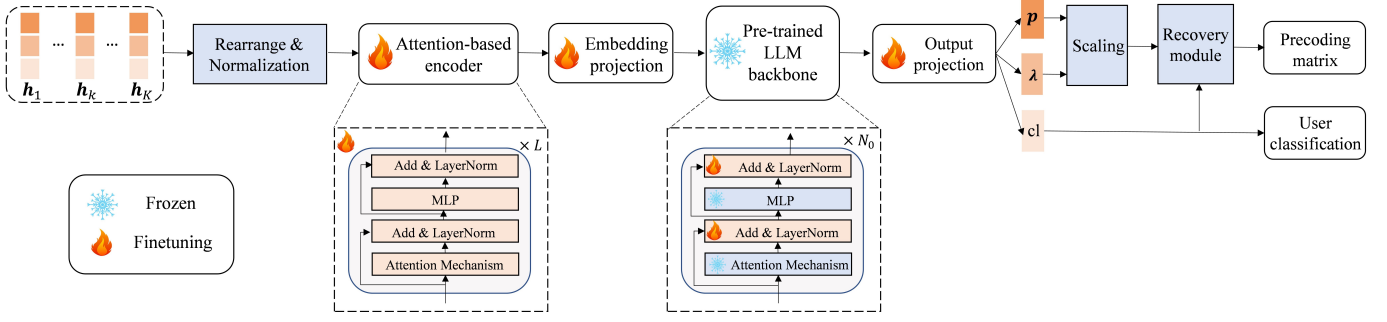


Fig. 3. The model framework of proposed LLM-empowered near-field multi-user precoding in LAE.

in our practical XL-MIMO system model. C_3 allows the LLM to dynamically adjust power allocation based on the number of users in each region and their channel conditions, ensuring efficient use of the total power budget while maximizing overall spectral efficiency. Without C_3 , optimization could disproportionately favor the one group (usually the near-field users), affecting the fairness and the ability of the system to serve a mixed user population effectively. The constraint C_4 represents the rate of every user every user should exceed the minimum rate $R_{\min}$. The constraint C_5 are the normalized constraint.

However, the problem (12) is non-convex and hard to get the globally optimal solution, as the practical XL-MIMO system model is considered and the constraint C_4 is non-convex. To solve it, a LLM-based scheme is proposed in the following subsection, which can distinguish far and near-field users and achieve joint optimization of precoding and power allocation.

B. Distinguishing Between Far and Near-Field Users

It should be emphasized that distinguishing between far and near-field users is necessary [26]. If all users are identified as far-field users, they will face serious spectrum efficiency performance loss as the planar-wave based far-field model becomes inaccurate in near-field region [11]. By contrast, if all users are identified as near-field users, since the extra distance dimension of near-field model, its corresponding computational and storage overhead is unacceptable and unnecessary in practical systems. Besides, due to the difficulty in obtaining the precise user distance, it is infeasible to distinguish them by directly comparing their distance with the ERD.

It should be noted that the user classification directly influences the precoding process by enabling tailored strategies for each user group. For near-field users, it applies the near-field beamfocusing vector, which accounts for both angle and distance to maximize SNR through precise energy focusing. For far-field users, it employs far-field beamsteering, optimizing directional energy distribution. This dual approach, combined with power allocation weighted by α_N in constraint C_3 , ensures efficient resource use and interference mitigation, leveraging the distinct spherical-wave and planar-wave propagation characteristics of near-field and far-field regions, respectively.

Thus, in this subsection, we present how to achieve user classification according to the proposed framework. The com-

plex channel of the K users are concatenated to form the input of the network, denoted as $\mathbf{H} = [\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_K] \in \mathbb{C}^{N \times K}$. To facilitate network processing and convergence, the complex channel $\mathbf{H}$ is rearranged in real matrix as $\mathbf{X}_{\text{in}} \in \mathbb{R}^{K \times 2N}$. Then we perform batch normalization for the input $\mathbf{X}_{\text{in}}$ as $\frac{\mathbf{X}_{\text{in}} - \mu}{\sigma}$, where μ and σ represent the mean value and standard deviation of a batch of corresponding input data. The normalization process can effectively facilitate network training and convergence. Then, an attention-based encoder is implemented to capture the relation among the users and extract preliminary features before LLM. The encoder comprises of $L = 3$ blocks of trainable transformer decoder, as shown in Fig. 3. The structure of each block includes a multi-head self-attention module and a multilayer perceptron (MLP) module. The normalized input as $\mathbf{X}_{\text{norm}}$ is sequentially processed by the multi-head self-attention module and MLP module. Then the output of the encoder can be written as

$$\mathbf{X}_{\text{en}} = \text{Encoder}(\mathbf{X}_{\text{norm}}), \quad (13)$$

where $\text{Encoder}(\cdot)$ represents the attention-based encoder. Acquiring the encoded input $\mathbf{X}_{\text{en}}$, an embedding projection module is applied to linearly project $\mathbf{X}_{\text{en}}$ to align the hidden dimensions with the backbone model, yielding $\mathbf{X}_{\text{emb}} \in \mathbb{R}^{K \times d}$, where d is the hidden dimension of LLM backbone.

The preprocessed multi-user channel then serves as the input of LLM backbone:

$$\mathbf{X}_{\text{LLM}} = \text{LLM}(\mathbf{X}_{\text{emb}}), \quad (14)$$

where $\text{LLM}(\cdot)$ denotes backbone networks of the LLM. Without loss of generality, the smallest version of GPT-2 [27] with feature dimension $d = 768$ is chosen as the LLM backbone in this work. It should be noted that, in the proposed method, the GPT-2 backbone can be flexibly replaced with other LLMs, such as Llama [28] and Qwen [29]. The rationale for the selection of LLM backbone is the trade-off between computational complexity and performance. The backbone of GPT-2 is also composed of stacked transformer decoders, as shown in Fig. 3. During the training process, only addition, layer normalization layers are fine-tuned for adapting the LLM to the specific task while self-attention and Multilayer Perceptron (MLP) layers are frozen to retain universal knowledge [30]. Finally, the output projection module is designed to convert the output

features of the LLM into the final user classification results:

$$\mathbf{X}_{\text{out}} = \text{Sigmoid}(\text{Linear}(\mathbf{X}_{\text{LLM}})), \quad (15)$$

where $\text{Linear}(\cdot)$ is the linear projection and $\text{Sigmoid}(\cdot)$ converts the output within the range $[0, 1]$. Hence, the output of user identification is obtained as $\hat{\mathbf{X}}_{\text{cl}} = \mathbf{X}_{\text{out}}[:, 0] \in \mathbb{R}^{K \times 1}$, where $\hat{\mathbf{X}}_{\text{cl}}$ indicates whether the user is located in far-field or near-field region.

In the training process, the ground truth of user class is available, which is denoted as $\mathbf{X}_{\text{cl}}$. The mean square error (MSE) is adopted as the loss function to minimize the classification error, i.e.,

$$\text{Loss}_{\text{cl}} = \|\mathbf{X}_{\text{cl}} - \hat{\mathbf{X}}_{\text{cl}}\|_2^2, \quad (16)$$

where $\|\cdot\|$ is the l_2 norm⁴.

C. Proposed LLM-based Multi-user Precoding

In this subsection, we elaborate on LLM-empowered near-field multi-user precoding for LAE. As proven in [31], the optimal downlink beamforming vectors for (12) without the constraint C_3, C_4 follows the structure as

$$\mathbf{w}_k^* = \frac{(\mathbf{I}_N + \sum_{k=1}^K \frac{\lambda_k}{\sigma^2} \mathbf{h}_k \mathbf{h}_k^H)^{-1} \mathbf{h}_k}{\|(\mathbf{I}_N + \sum_{k=1}^K \frac{\lambda_k}{\sigma^2} \mathbf{h}_k \mathbf{h}_k^H)^{-1} \mathbf{h}_k\|_2}, \quad \forall k, \quad (17)$$

where λ_k is a positive parameter and $\sum_{k=1}^K \lambda_k = P$. Given this knowledge, we only require to learn $\lambda = [\lambda_1, \lambda_2, \dots, \lambda_K]$ rather than the entire high-dimension matrix $\mathbf{W}$, to obtain the normalized precoding vector. Therefore, for multi-user precoding, we focus on learning the key features λ and power allocation vector $\mathbf{p} = [P_1, P_2, \dots, P_K]$ with specific designs to satisfy the constraint.

For traditional methods, WMMSE algorithm is widely adopted to estimate the parameters. However, WMMSE inherently converges to local optimal solution, resulting suboptimal performance. Secondly, the iterative nature of WMMSE introduce prohibitive execution delays in real-time deployments. To address these challenges, we introduce LLM for multi-user precoding in this work.

The main structure of the proposed network is shared with that for user classification illustrated in the last subsection, and thus we skip the repeated part here. After the output projection module, λ and $\mathbf{p}$ is obtained as $\lambda = \mathbf{X}_{\text{out}}[:, 1] \in \mathbb{R}^{K \times 1}$ and $\mathbf{p} = \mathbf{X}_{\text{out}}[:, 2] \in \mathbb{R}^{K \times 1}$. Then $\mathbf{p}$ and λ is scaled to meet the constraint $\sum_{k=1}^K \lambda_k = \sum_{k=1}^K P_k = P$. We further check whether the constraint in C_3 is satisfied; if not, $\mathbf{p}$ and λ is rescaled according to C_3 . Obtaining the normalized $\hat{\mathbf{p}}$ and $\hat{\lambda}$, the recovery module is applied to acquire the precoding matrix based on (17) and power allocation vector.

For multi-user precoding, we directly adopt the opposite number of sum rate $-\sum_k R_k$ as the loss function in an unsupervised learning way. Besides, to meet the requirement in constraint C_4 , we add an additional penalty loss to ensure the minimal rate is larger than $R_{\min}$. Denote the computed

rate of the K users as $\mathbf{R} = [R_1, R_2, \dots, R_K]$, and thereby the penalty loss can be written as

$$\text{penal} = \|\max\{R_{\min} - \mathbf{R}, \mathbf{0}\}\|_1, \quad (18)$$

where $\|\cdot\|_1$ is the l_1 loss of the vector. Thus the entire loss function of the multi-user precoding is

$$\text{Loss}_{\text{pre}} = -\sum_k R_k + \gamma_1 \text{penal}, \quad (19)$$

where γ_1 presents the trade-off between sum rate performance and the penalty for user fairness constraint. Note that a larger γ_1 indicates stricter enforcement of user fairness constraints. we set $\gamma_1 = 10$ in this work, which sufficiently enforces user fairness constraints, ensuring all user rates exceed required thresholds. Therefore, the final loss function of the entire network is

$$\text{Loss} = \gamma_2 \text{Loss}_{\text{cl}} + \text{Loss}_{\text{pre}}, \quad (20)$$

where γ_2 balances the performance of the two subtasks (user classification and multi-user precoding). A larger γ_2 prioritizes classification performance, while a smaller γ_2 emphasizes precoding optimization. Based on the experimental results, $\gamma_2 = 5$ achieves a favorable balance between the two tasks, yielding near-optimal performance for both simultaneously.

It should be noted that the derivation of the ENFR in Section III contributes to the optimization problem (12) by enabling precise classification of near-field and far-field users, shaping constraint C_3 and informing tailored precoding strategies. ENFR's replacement of ERD impacts spectral efficiency by accurately identifying near-field users for beamfocusing. The ENFR ensures optimal power allocation and beamforming, improving the solution's effectiveness.

D. Advantages of LLM-based Scheme Over Conventional Solvers

While traditional optimization solvers (e.g., convex relaxation, gradient descent) or other data-driven methods (e.g., CNNs) may also be able to address the above optimization problem, our LLM-based scheme offers distinct advantages:

a) *Scalability and Complexity Handling*: Equipped with large parameter size, one of the most notable advantages of LLMs is their powerful fitting ability to handle complex optimization problem. For near-field communication problems in LAE, LLMs, such as GPT-2, efficiently process high-dimensional, variable-length data (e.g., channel matrix $\mathbf{H}$). Traditional solvers, in contrast, struggle with the non-convex optimization problem and computational cost as dimensions grow. The performance of traditional data-driven methods may also degrade with more complex problem and high-dimensional channel, since the relatively small model size limits their ability to extract efficient and global features from the data.

b) *Adaptability to Dynamics*: Fine-tuning with adapters allows LLMs to adapt quickly to LAE's changing conditions, unlike traditional static solvers or less flexible neural networks. This is critical for real-time optimization may enable more efficient and versatile deployment.

⁴We train the classification model with MSE loss and cross-entropy, respectively, and find the network trained with MSE loss outperforms that trained with cross-entropy.

c) *Generalization Capability*: Leveraging pretrained knowledge, LLMs exhibit excellent generalization ability for multi-task tasks and scenarios, approximating near-optimal solutions for non-convex problems across tasks and scenarios. In this work, it can achieve near-optimal performance for both power allocation and multi-user precoding tasks, as illustrated in Section V. On the other hand, the traditional data-drive methods exhibit poor generalization ability and lack multitasking capability, requiring retraining when the CSI distribution changes.

So far, we have analyzed our proposed LLM-based scheme for near-field multi-user communications. Simulation results are presented to substantiate the efficacy and superiority of proposed LLM-based scheme in the following section.

V. SIMULATION RESULTS

A. Simulation Setups

In this section, simulation results are presented to verify the performance of our proposed scheme. Specifically, a downlink XL-MIMO system is considered, where $N = 256$, $h_B = 15$ m and $\theta_{tit} = 5^\circ$. Besides, we set the carrier frequency is 30 GHz and the antenna spacing is $d = \frac{\lambda}{2} = 0.5$ cm. For the user distribution, they are distributed randomly, where the range of x_k and h_k are $[0, 200$ m] and $[0, 30$ m]. The maximal user number is set as 10, while the noise power is $\sigma^2 = 0.01$. For multi-user precoding tasks, we assume that we have got the perfect channel state information (CSI). In practice, CSI estimation errors can affect the system's performance, and in practice, several near-field channel estimation schemes have been proposed to achieve near-perfect CSI [10]. For example, a representative near-field channel estimation scheme was proposed in [10], where it fully utilizes polar-domain sparsity of the near-field channel to achieve compressed sensing-based estimation, achieving high accuracy with low pilot overhead.

A training dataset comprising 8,000 samples, a validation dataset comprising 1,000 samples and a testing dataset comprising 1,000 samples are constructed respectively, according to the channel model in [10], [32]. For the hyper-parameters in network training, we set number of training epoch as 500, the batch size as 100, and learning rate as 0.0001. In this work, we utilize Adam optimizer for model training with $\text{betas} = (0.9, 0.999)$, $\text{weight_decay} = 0.0001$. Besides, all the training and inference of the proposed model is conducted on an NVIDIA GeForce RTX 4090 24GB GPU.

In addition, some benchmark comparison schemes are considered as follows: (1) Capacity [33], where dirty-paper coding is utilized to achieve the capacity of the multiple-antenna Gaussian broadcast channel; (2) CNN [31], where a CNN-based scheme for the optimization of downlink beamforming is adopted. We further improve the scheme to achieve joint user classification and multi-user precoding; (3) Transformer, where we employ the sequence-to-sequence mapping capability of Transformer to implement user classification and multi-user near-field precoding simultaneously. (4) Near-field NOMA [34], where non-orthogonal multiple access (NOMA) scheme is utilized in near-field communications with a dynamic power allocation algorithm; (5) Near-field LDMA [14], where the near-field polar-domain analog codebook and zero-forcing (ZF) digital precoding scheme with equal power

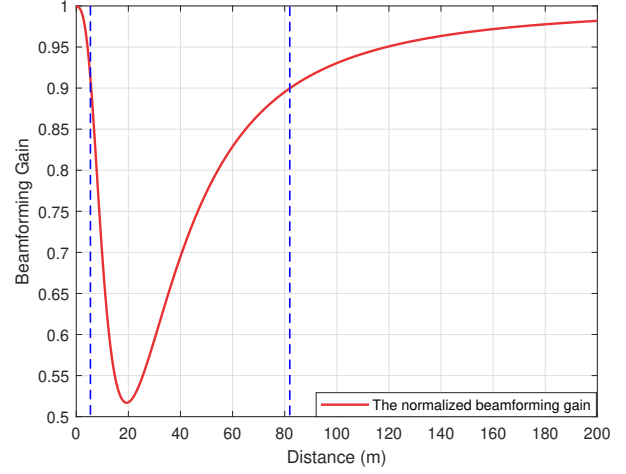


Fig. 4. The normalized beamforming gain with far-field beamforming vector in our adopted model.

allocation are applied; (6) Far-field SDMA, where the DFT codebook is adopted unlike the near-field LDMA.

B. Performance Analysis

First, the normalized beamforming gain with far-field beamforming vector in our adopted model is illustrated in Fig. 4, where the predefined threshold is $\Delta = 0.1$ and the dashed blue line is plotted to show it. It should be noted that here $\Delta = 0.1$ is selected as the normalized beamforming gain loss threshold, representing a 10% degradation. This value balances the extent of the near-field region with performance, aligning with typical thresholds in near-field XL-MIMO studies [7]. It can be shown as the distance increases, the normalized beamforming gain exhibits a pattern of initially declining followed by a rise.. In other words, for the adopted 3D XL-MIMO system, the ENFR of horizontal plane is located between two far-field regions, which verifies the accuracy of **Lemma 1**.

Then to analyze convergence rate of neural network training, we depict the training loss, as well as validation loss, of the proposed model over training epochs in Fig. 5. We can observe that the overall trend of training loss declines with the increase of training epoch. When the training epoch reaches about 30, the training loss function gradually converges. To mitigate overfitting, we utilize a validation set for model selection. Specifically, the final model is chosen based on the lowest validation loss, ensuring robustness and generalization. For example, as illustrated in Fig. 5, the model in Epoch 304 achieves the lowest validation loss and is selected for testing.

Besides, the results of different schemes for distinguishing between far and near-field users is presented in Table I. It can be demonstrated that our proposed LLM-based scheme exceeds classical CNN-based scheme and transformer-based scheme, and achieves near-optimal classification accuracy for distinguishing between far and near-field users in high SNR regime. To further assess the robustness of proposed LLM-based classifier, we evaluate the classification accuracy when only imperfect CSI with noise can be obtained. The test

TABLE I
CLASSIFICATION ACCURACY VS. SNR.

Schemes	0 dB	5 dB	10 dB	15 dB	20 dB
Proposed scheme	0.9478	0.9852	0.9904	0.9909	0.9918
CNN [31]	0.8030	0.8221	0.8281	0.8289	0.8291
Transformer	0.9214	0.9436	0.9579	0.9631	0.9732

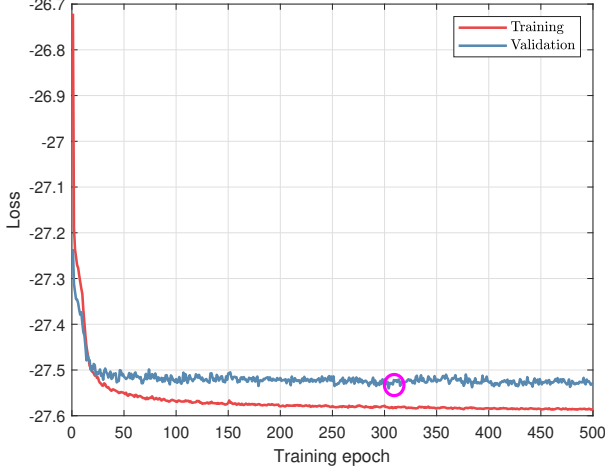


Fig. 5. Training loss and validation loss against training epoch.

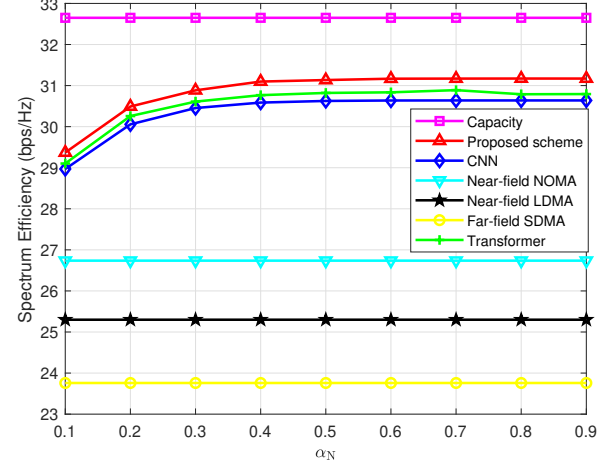


Fig. 7. Spectrum efficiency against α_N .

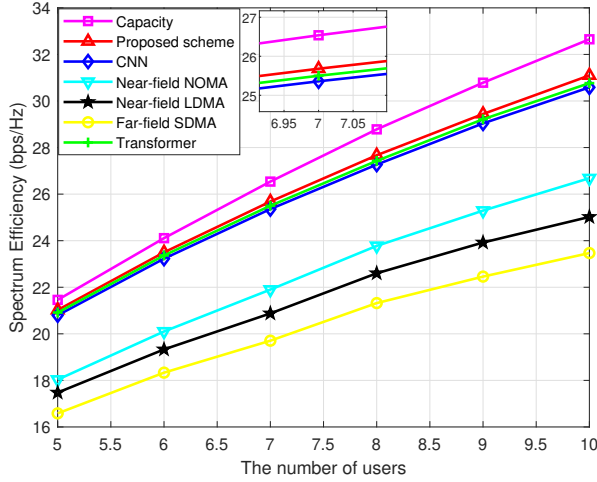


Fig. 6. Spectrum efficiency against K .

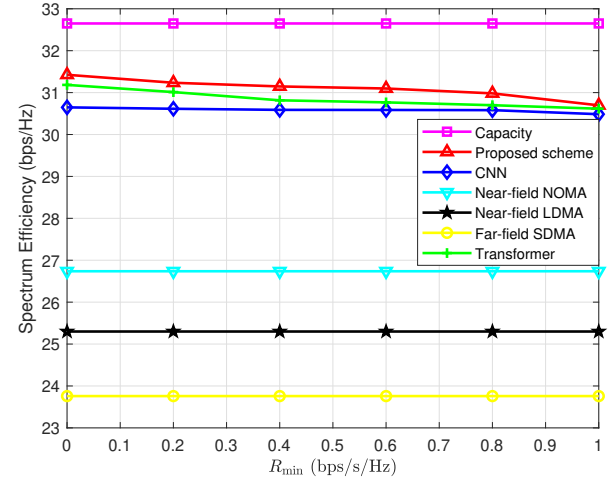


Fig. 8. Spectrum efficiency against $R_{\min}$.

SNR ranges from 0 dB to 20 dB. The proposed method can obtain classification accuracy over 99% if SNR is larger than 10 dB, achieving 17% higher accuracy compared to CNN-based approach. When the SNR is relatively low, 0 dB for instance, the classification accuracy is still about 95%, indicating the strong robustness of LLM-based scheme.

Furthermore, performance comparison of our proposed scheme with the benchmark schemes from multiple perspectives are shown in detail from Fig. 6 to Fig. 9. Fig. 6 illustrate the spectrum efficiency performance against user number K ,

which increases from 5 to 10. The minimum data rate is $R_{\min} = 0.6$ bps/s/Hz, the maximal ratio of power allocated to near-field users $\alpha_c = 0.4$, and transmit power and noise power is set as $P = 0$ dBW, $\sigma^2 = -20$ dBW, respectively. As the user number increases, the spectrum efficiency increases with further exploitation of multiplexing gain. However, this trend does not extend indefinitely. Beyond $K \approx N$, efficiency would saturate due to limited spatial degrees of freedom and increased interference, aligning with Shannon's channel capacity limits [35]. As shown in Fig. 6, considering both the angle and

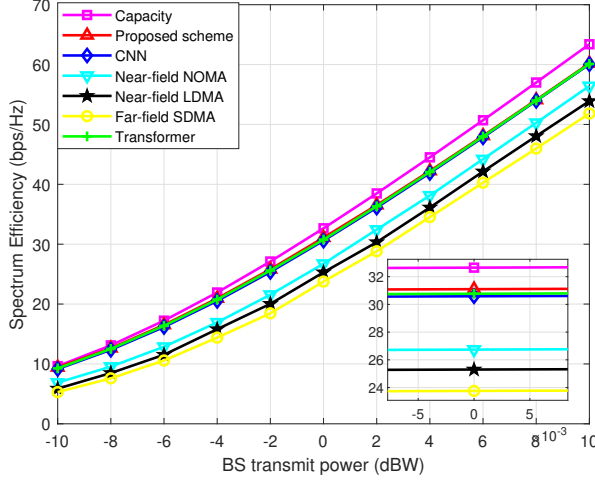


Fig. 9. Spectrum efficiency against P .

distance information, near-field LDMA and near-field NOMA-based methods achieve higher spectrum efficiency than far-field SDMA considering only angle information. Besides, the deep learning-based methods further improve the performance, and the proposed method outperforms other baselines for all user numbers. This verifies the potential of LLM in solving the complex near-field multi-user precoding problem for LAE. It is worth noting that, transformers and transformer-based LLMs, inherently possess the ability to process variable-length sequences, thus our proposed scheme can be directly applied to different number of users without requiring any modifications. In contrast, for traditional CNN-based methods, to accommodate varying numbers of users, input data under different user numbers must be zero-padded to match the shape of the maximum user number.

Fig. 7 presents the impact of maximal near-field power ratio on spectrum efficiency. We assume the user number as $K = 10$, while the other simulation settings are the same as those in Fig. 6. Since the majority of existing methods, including near-field LDMA, near-field NOMA and far-field SDMA scheme, fail to consider this factor, we only present the spectrum efficiency performance without constraint C_3 in (12), i.e. $\alpha_N = 1$, which is shown as horizontal lines. Then we mainly focus on the comparison of the proposed scheme with CNN-based scheme and transformer-based scheme. As α_N increases from 0 to 0.5, the spectral efficiency grows with the relaxation of constraints on α_N . This is primarily because stricter power allocation constraints for near-field users limit the solution space available to the network, thereby degrading the spectral efficiency. When α_N is greater than 0.5, the spectral efficiency tends to remain constant, indicating that α_N is not a dominant factor affecting performance at this point. As illustrated in Fig. 7, thanks to LLMs' superior feature extraction and robustness for complex scenarios, proposed LLM-based method achieves higher spectrum efficiency performance than CNN-based approach and transformer-based approach.

In addition, we evaluate the impact of the minimum rate

$R_{\min}$ to the spectrum efficiency performance, as illustrated in Fig. 8. The minimum rate $R_{\min}$ changes from 0 to 1 bps/s/Hz, where $R_{\min} = 0$ implies that no constraint for $R_{\min}$ is imposed. The user number is set as $K = 10$, and other parameters remain the same as Fig. 6. Similar to Fig. 7, traditional methods mainly lack constraints on the minimal rate to guarantee the fairness among users, and they exhibit horizontal performance lines with $R_{\min} = 0$. With the increase of $R_{\min}$, the spectrum efficiency performance gradually drops with more emphasis on user fairness, rather than merely on spectrum efficiency. Therefore, a proper $R_{\min}$ can be selected to strike a balance between performance and fairness, and in this work we set $R_{\min} = 0.6$ bps/s/Hz for other simulations. LLMs acquire the ability to handle complex optimization problems with constraints on power allocation and user fairness, that are difficult for traditional optimization methods, thereby achieving satisfying performance for near-field multi-user precoding in LAE.

In Fig. 9 we compare the spectrum efficiency performance under different BS transmit powers, which grows from -10 dBW to 10 dBW. Besides, we assume $R_{\min} = 0.6$ bps/s/Hz, $\alpha_c = 0.4$, $\sigma^2 = -20$ dBW and $K = 10$. As presented in Fig. 9, the proposed LLM-based scheme achieves near-optimal spectrum efficiency performance and outperforms other benchmark schemes for the entire transmit power range. Owing to the increasing size of network, LLM-based method exhibits superior optimization and generalization capabilities and outperforms other deep learning-based methods in terms of performance. Moreover, AI-based methods achieve higher spectrum efficiency than traditional codebook-based methods (i.e. near-field NOMA, near-field LDMA as well as far-field SDMA-based methods) due to their feature extraction capabilities and increased degrees of freedom in solution space.

In summary, based on the analysis from Fig. 6 to Fig. 9, it can be concluded that our proposed scheme can demonstrate excellent performance under various parameter settings, which verifies the strong robustness and generalization ability of the proposed LLM-based scheme.

Moreover, we evaluate the performance of our proposed LLM-based scheme across a range of γ_2 values. As mentioned before, γ_2 balances the performance of the two subtasks. The simulation result for different values of γ_2 is illustrated in Table. III, where γ_2 increases from 0.2 to 30. The following simulation result is averaged for different SNRs from 5 dB to 25 dB, and we set $R_{\min} = 0.6$ bps/s/Hz, $\alpha_c = 0.4$, $\sigma^2 = -20$ dBW and $K = 8$. As shown in the table, with the increase of γ_2 , the classification performance progressively improves until reaching the model's maximum attainable performance, after which it remains stable. Conversely, the sum rate of the precoding task initially remains stable but gradually diminishes with further increases in γ_2 . According to the simulation results, setting γ_2 within [5, 10] is proper.

Finally, We compare the model training and inference time, as well as the computational complexity of the proposed method with other DL-based baselines to assess the difficulty of deploying the model in practical scenarios, as shown in Table II. All experiments are conducted on the same machine with batch size 100. The CNN-based method, with the smallest

TABLE II
NETWORK PARAMETERS (TRAINING PARAMETERS/TOTAL PARAMETERS) AND THE TRAINING/INTERFERENCE TIME PER BATCH

Metric	CNN	Transformer	Proposed scheme
Network parameters (1×10^6)	0.92/0.92	13.15/13.15	5.95/129.56
Training time (ms)	13.81	70.91	59.92
Inference time (ms)	11.71	62.80	52.56

TABLE III
COMPARISONS OF PERFORMANCE WITH DIFFERENT γ_2 .

γ_2	0.2	0.5	1	2	5	10	15	20	30
sum rate	27.877	27.860	27.876	27.881	27.876	27.875	27.607	27.657	27.462
accuracy	0.961	0.969	0.982	0.989	0.992	0.991	0.992	0.992	0.991

model size, achieves the fastest training and inference time. However, it cannot achieve satisfactory performance for both user classification and multi-user precoding. While our LLM-based scheme incurs higher computational complexity than the CNN-based approach, its superior performance, including accurate user classification accuracy and enhanced spectrum efficiency, justifies this trade-off between performance and cost. These gains are critical for 6G near-field XL-MIMO systems, where high data rates and reliability are paramount. Techniques such as antenna selection or hardware acceleration can further optimize complexity, ensuring practical deployability. Besides, it is worth noting that the training and inference time of the proposed method is even shorter than that of the Transformer, although the total parameter size is much larger than Transformer. It is mainly due to the inference acceleration specific to the GPT model. Therefore, the proposed method is a promising method to deploy in practical communication networks.

VI. CONCLUSIONS

In this paper, we first apply LLM to solve the spectrum efficiency maximization problem of near-field communications for LAE. By elaborately designing adapters and finetuning the pretrained GPT-2, our proposed LLM-based scheme can accurately distinguish far-field and near-field users and achieve joint optimization of precoding and power allocation. Simulation results substantiate the efficiency of the proposed scheme, which it could demonstrate excellent performance under various parameter settings. For the future research, how to apply LLM to solve other physical layer communications for near-field communications in LAE will be a promising problem. For example, LLM-empowered near-field sensing, integrated sensing and communications (ISAC) and distributed beamforming for LAE may be critical future research directions [36]–[39].

REFERENCES

- [1] J. Wan, H. Ren, C. Pan, Z. Zhang, S. Gao, Y. Yu, and C. Wang, "Sensing capacity for integrated sensing and communication systems in low-altitude economy," *arXiv:2411.06983*, 2024.
- [2] J. Tang, Y. Yu, C. Pan, H. Ren, D. Wang, J. Wang, and X. You, "Co-operative ISAC-empowered low-altitude economy," *arXiv:2412.20371*, 2024.
- [3] X. Ye, Y. Mao, X. Yu, S. Sun, L. Fu, and J. Xu, "Integrated sensing and communications for low-altitude economy: A deep reinforcement learning approach," *arXiv:2412.04074*, 2024.
- [4] Z. Li, Z. Gao, K. Wang, Y. Mei, C. Zhu, L. Chen, X. Wu, and D. Niyato, "Unauthorized UAV countermeasure for low-altitude economy: Joint communications and jamming based on MIMO cellular systems," *IEEE Internet Things J.*, 2024.
- [5] Y. Zeng, Q. Wu, and R. Zhang, "Accessing from the sky: A tutorial on UAV communications for 5G and beyond," *Proc. IEEE*, vol. 107, no. 12, pp. 2327–2375, Dec. 2019.
- [6] Z. Wang, J. Zhang, H. Du, D. Niyato, S. Cui, B. Ai, M. Debbah, K. B. Letaief, and H. V. Poor, "A tutorial on extremely large-scale MIMO for 6G: Fundamentals, signal processing, and applications," *IEEE Commun. Surv. Tutor.*, 2024.
- [7] H. Lu, Y. Zeng, C. You, Y. Han, J. Zhang, Z. Wang, Z. Dong, S. Jin, C.-X. Wang, T. Jiang *et al.*, "A tutorial on near-field XL-MIMO communications towards 6G," *IEEE Commun. Surv. Tutor.*, 2024.
- [8] K. Chen, C. Qi, J. Huang, O. A. Dobre, and G. Y. Li, "Near-field communications for extremely large-scale MIMO: A beamspace perspective," *IEEE Commun. Mag.*, 2025.
- [9] C. You, Y. Cai, Y. Liu, M. Di Renzo, T. M. Duman, A. Yener, and A. L. Swindlehurst, "Next generation advanced transceiver technologies for 6G," *IEEE J. Sel. Areas Commun.*, vol. 43, no. 3, pp. 582–627, Mar. 2025.
- [10] M. Cui and L. Dai, "Channel estimation for extremely large-scale MIMO: Far-field or near-field?" *IEEE Trans. Commun.*, vol. 70, no. 4, pp. 2663–2677, Apr. 2022.
- [11] H. Zhang, N. Shlezinger, F. Guidi, D. Dardari, M. F. Imani, and Y. C. Eldar, "Beam focusing for near-field multiuser MIMO communications," *IEEE Trans. Wireless Commun.*, vol. 21, no. 9, pp. 7476–7490, Sep. 2022.
- [12] Z. Wang, X. Mu, and Y. Liu, "Beamfocusing optimization for near-field wideband multi-user communications," *IEEE Trans. Commun.*, vol. 73, no. 1, pp. 555–572, Jan. 2025.
- [13] Y. Xu, W. Li, C. Huang, C. Zhu, Z. Yang, J. Yang, J. He, Z. Zhang, and M. Debbah, "Hashing beam training for integrated ground-air-space wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 42, no. 12, pp. 3477–3489, Dec. 2024.
- [14] Z. Wu and L. Dai, "Multiple access for near-field communications: SDMA or LDMA?" *IEEE J. Sel. Areas Commun.*, vol. 41, no. 6, pp. 1918–1935, Jun. 2023.
- [15] F. Jiang, Y. Peng, L. Dong, K. Wang, K. Yang, C. Pan, D. Niyato, and O. A. Dobre, "Large language model enhanced multi-agent systems for 6G communications," *IEEE Wireless Commun.*, vol. 31, no. 6, pp. 48–55, Dec. 2024.
- [16] H. Li, M. Xiao, K. Wang, D. I. Kim, and M. Debbah, "Large language model based multi-objective optimization for integrated sensing and communications in UAV networks," *arXiv:2410.05062*, 2024.
- [17] W. Lee and J. Park, "LLM-empowered resource allocation in wireless communications systems," *arXiv:2408.02944*, 2024.
- [18] J. Tong, J. Shao, Q. Wu, W. Guo, Z. Li, Z. Lin, and J. Zhang, "Wirelessagent: Large language model agents for intelligent wireless networks," *arXiv:2409.07964*, 2024.
- [19] J. Shao, J. Tong, Q. Wu, W. Guo, Z. Li, Z. Lin, and J. Zhang,

- “Wirelessllm: Empowering large language models towards wireless intelligence,” *arXiv:2405.17053*, 2024.
- [20] Y. Cui, J. Guo, C.-K. Wen, S. Jin, and E. Tong, “Exploring the potential of large language models for massive MIMO CSI feedback,” *arXiv:2501.10630*, 2025.
- [21] J. Guo, Y. Cui, C.-K. Wen, and S. Jin, “Prompt-enabled large AI models for CSI feedback,” *arXiv:2501.10629*, 2025.
- [22] W. Lee, S.-R. Lee, H.-B. Kong, S. Lee, and I. Lee, “Downlink vertical beamforming designs for active antenna systems,” *IEEE Trans. Commun.*, vol. 62, no. 6, pp. 1897–1907, Jun. 2014.
- [23] J. Sherman, “Properties of focused apertures in the fresnel region,” *IRE Trans. Antennas Propag.*, vol. 10, no. 4, pp. 399–408, Jul. 1962.
- [24] Z. Wu, M. Cui, and L. Dai, “Enabling more users to benefit from near-field communications: From linear to circular array,” *IEEE Trans. Wireless Commun.*, vol. 23, no. 4, pp. 3735–3748, Apr. 2024.
- [25] M. Cui and L. Dai, “Near-field wideband beamforming for extremely large antenna arrays,” *IEEE Trans. Wireless Commun.*, vol. 23, no. 10, pp. 13 110–13 124, Oct. 2024.
- [26] Y. Zhang, C. You, L. Chen, and B. Zheng, “Mixed near-and far-field communications for extremely large-scale array: An interference perspective,” *IEEE Commun. Lett.*, vol. 27, no. 9, pp. 2496–2500, Sep. 2023.
- [27] R. K. Alec, J. Wu, R. Child, D. Luan, D. Amodei, and I. Sutskever, “Language models are unsupervised multitask learners,” *OpenAI blog*, vol. 1, no. 8, p. 9, 2019.
- [28] H. Touvron, L. Martin, K. R. Stone, P. Albert, A. Almahairi, Y. Babaei, N. Bashlykov, S. Batra, P. Bhargava, and *et al.* Shrutu Bhosale, “Llama 2: Open foundation and fine-tuned chat models,” *arXiv:2307.09288*, 2023.
- [29] A. Yang, B. Yang, B. Zhang, B. Hui, B. Zheng, B. Yu, and C. L. *et al.*, “Qwen2.5 technical report,” *arXiv:2412.15115*, 2024.
- [30] T. Zheng and L. Dai, “Large language model enabled multi-task physical layer network,” *arXiv:2412.20772*, 2024.
- [31] W. Xia, G. Zheng, Y. Zhu, J. Zhang, J. Wang, and A. P. Petropulu, “A deep learning framework for optimization of MISO downlink beamforming,” *IEEE Trans. Commun.*, vol. 68, no. 3, pp. 1866–1880, 2020.
- [32] Y. Lu and L. Dai, “Near-field channel estimation in mixed LoS/NLoS environments for extremely large-scale MIMO systems,” *IEEE Trans. Commun.*, vol. 71, no. 6, pp. 3694–3707, Jun. 2023.
- [33] N. Jindal, W. Rhee, S. Vishwanath, S. A. Jafar, and A. Goldsmith, “Sum power iterative water-filling for multi-antenna Gaussian broadcast channels,” *IEEE Trans. Inform. Theory*, vol. 51, no. 4, pp. 1570–1580, Apr. 2005.
- [34] Z. Xu, Z. Wu, and L. Dai, “How to enhance spectrum efficiency for near-field communications: From LDMA to NOMA?” *submitted to IEEE Trans. Commun.*, 2024.
- [35] E. Björnson, E. G. Larsson, and M. Debbah, “Massive MIMO for maximal spectral efficiency: How many users and pilots should be allocated?” *IEEE Trans. Wireless Commun.*, vol. 15, no. 2, pp. 1293–1308, Feb. 2016.
- [36] W. Yu, H. He, S. Song, J. Zhang, L. Dai, L. Zheng, and K. B. Letaief, “Ai and deep learning for thz ultra-massive MIMO: From model-driven approaches to foundation models,” *arXiv:2412.09839*, 2024.
- [37] F. You, H. Du, K. Huang, and A. Jamalipour, “JPPPO: Joint power and prompt optimization for accelerated large language model services,” *arXiv:2411.18010*, 2024.
- [38] W. Yu, Y. Ma, H. He, S. Song, J. Zhang, and K. B. Letaief, “Deep learning for near-field XL-MIMO transceiver design: Principles and techniques,” *IEEE Commun. Mag.*, vol. 63, no. 1, pp. 52–58, Jan. 2025.
- [39] F. You, H. Du, K. Huang, and A. Jamalipour, “Network-aided efficient large language model services with denoising-inspired prompt compression,” *arXiv:2412.03621*, 2024.



Zhuo Xu (Graduate Student Member, IEEE) received the B.E. degree in communication engineering from the Harbin Institute of Technology, Harbin, China, in 2023. He is currently pursuing the M.S. degree with the Department of Electronic Engineering, Tsinghua University, Beijing, China. His research interests include extremely large-scale MIMO (XL-MIMO), near-field MIMO communications, and AI for communications. He received the National Scholarship in 2020 and 2022.



Tianyue Zheng (Graduate Student Member, IEEE) received the B.E. degree in information engineering from Southeast University, Nanjing, China, in 2022. She is currently pursuing the Ph.D. degree in the Department of Electronic Engineering, Tsinghua University, Beijing, China. Her research interests include extremely large scale MIMO (XL-MIMO), CSI acquisition and AI for communications. She has received the National Scholarship in 2019 and the Excellent Student of Jiangsu Province in 2021.



Linglong Dai (Fellow, IEEE) received the B.S. degree from Zhejiang University, Hangzhou, China, in 2003, the M.S. degree from the China Academy of Telecommunications Technology, Beijing, China, in 2006, and the Ph.D. degree from Tsinghua University, Beijing, in 2011. From 2011 to 2013, he was a Post-Doctoral Researcher with the Department of Electronic Engineering, Tsinghua University, where he was an Assistant Professor from 2013 to 2016, an Associate Professor from 2016 to 2022, and has been a Professor since 2022. His current research

interests include massive MIMO, reconfigurable intelligent surface (RIS), millimeter-wave and Terahertz communications, near-field communications, machine learning for wireless communications, and electromagnetic information theory.

He has coauthored the book *MmWave Massive MIMO: A Paradigm for 5G* (Academic Press, 2016). He has authored or coauthored over 100 IEEE journal papers and over 60 IEEE conference papers. He also holds over 20 granted patents. He has received five IEEE Best Paper Awards at the IEEE ICC 2013, the IEEE ICC 2014, the IEEE ICC 2017, the IEEE VTC 2017-Fall, the IEEE ICC 2018, and the IEEE GLOBECOM 2023. He has also received the Tsinghua University Outstanding Ph.D. Graduate Award in 2011, the Beijing Excellent Doctoral Dissertation Award in 2012, the China National Excellent Doctoral Dissertation Nomination Award in 2013, the URSI Young Scientist Award in 2014, the IEEE Transactions on Broadcasting Best Paper Award in 2015, the Electronics Letters Best Paper Award in 2016, the National Natural Science Foundation of China for Outstanding Young Scholars in 2017, the IEEE ComSoc Asia-Pacific Outstanding Young Researcher Award in 2017, the IEEE ComSoc Asia-Pacific Outstanding Paper Award in 2018, the China Communications Best Paper Award in 2019, the IEEE Access Best Multimedia Award in 2020, the IEEE Communications Society Leonard G. Abraham Prize in 2020, the IEEE ComSoc Stephen O. Rice Prize in 2022, the IEEE ICC Best Demo Award in 2022, and the National Science Foundation for Distinguished Young Scholars in 2023. He was listed as a Highly Cited Researcher by Clarivate Analytics from 2020 to 2023. He was elevated as an IEEE Fellow in 2022.