Distance-Aware Precoding for Near-Field Capacity Improvement in XL-MIMO

Zidong Wu, Mingyao Cui, Zijian Zhang, and Linglong Dai
Beijing National Research Center for Information Science and Technology (BNRist)
Department of Electronic Engineering, Tsinghua University, Beijing 100084, China
Email: {wuzd19, cmy20, zhangzj20}@mails.tsinghua.edu.cn, daill@tsinghua.edu.cn

2022 IEEE 95th Vehicular Technology Conference (VTC2022-Spring) | 978-1-6654-8243-1/22/$31.00 ©2022 IEEE | DOI: 10.1109/VTC2022-Spring54318.2022.9860745

***Abstract*—Extremely large-scale MIMO (XL-MIMO) communication is a promising technology to improve the capacity for future 6G networks. With a very large number of antennas, the near-field property of XL-MIMO systems becomes significant. Unlike the classical far-field line-of-sight (LoS) channel with only one available data stream, significantly increased degrees of freedom (DoFs) are available in the near-field LoS channel. However, limited by the small number of radio frequency (RF) chains, the existing hybrid precoding architecture widely used for 5G is not able to fully utilize the extra DoFs in the near-field region. In this paper, to exploit the near-field effect as a new possibility for capacity improvement, the distance-aware precoding (DAP) architecture is developed, where each RF chain can be flexibly configured to active or inactive according to the distance-related DoFs. Moreover, based on the developed DAP architecture, a DAP algorithm is proposed to optimize the number of activated RF chains and precoding matrices to match the increased DoFs. Finally, simulation results verify that, the proposed DAP scheme can efficiently utilize the extra DoFs in the near-field region to improve the spectrum efficiency.**

***Index Terms*—Extremely large-scale MIMO, hybrid precoding, near-field, spectrum efficiency**

I. Introduction

Due to the explosively growing demand for communication capacity, spatial multiplexing is regarded as a key technology to significantly increase the spectrum efficiency. By exploiting the spatial degrees of freedom (DoFs) with spatial multiplexing, massive multiple-input multiple-output (MIMO) in 5G is able to increase the capacity by orders of magnitude. To further exploit the promising potential of spatial DoFs, traditional 5G massive MIMO is evolving to extremely large-scale MIMO (XL-MIMO) for future 6G communications [1].

The evolution from massive MIMO to XL-MIMO not only implies a significant increase in the number of antennas, but also leads to a fundamental change in the characteristics of the electromagnetic field. Specifically, since the number of antennas of massive MIMO systems is usually not very large, the classical *far-field* MIMO channel model is widely adopted, which is based on planar wave assumptions [2]. By contrast, with the dramatically increasing number of antennas in XL-MIMO communications, receivers may lie in the *near-field* region of the transmitter, which is defined by the area within the Rayleigh distance [3]. Unlike the classical *far-field* channel based on planar wave assumptions, the *near-field* channel should be modeled based on spherical wave assumptions in XL-MIMO communications, leading to significant near-field effect [4]. Thus, traditional far-field transmission techniques may suffer a severe performance loss due to the mismatch between *far-field* assumptions and *near-field* channel property.

Most of the existing works on near-field communications focus on overcoming the deterioration caused by the near-field effect. To be specific, initial works [5], [6] studied the near-field beamforming techniques by designing the direction- and distance-dependent beams in narrow-band communications. To achieve near-field beamforming in wideband systems, by exploiting the varied spherical wavefronts at different frequencies, a phase-delay focusing method was proposed in [4] to realize near-field beamforming. To acquire high gain beamforming, channel estimation is essential for XL-MIMO communications. Several works have also been studied to overcome the performance loss of channel estimation in the near-field region. In [7], a channel estimation method was proposed to capture the sparsity in the joint direction-distance (polar) domain caused by spherical waves. We can find that, existing works on near-field communications indicate that the near-field effect only induces performance degradation.

Nevertheless, the recently developed electromagnetic information theory (EIT) indicates that, wireless communications can also benefit from the near-field effect [8]. Specifically, for far-field channels, due to the high attenuation of millimeter-wave and terahertz bands, the rank-one line-of-sight (LoS) path becomes predominant in XL-MIMO, resulting in limited spatial multiplexing gain. Compared with the far-field LoS path with *planar* wavefronts determined by a single spatial angle, the near-field LoS path with *spherical* wavefronts contains a range of angles, which brings significantly increased spatial DoFs (or transmission modes) [9]. Owing to the increased spatial DoFs, it is expected that the spectrum efficiency can be naturally enhanced in the near-field region.

Unfortunately, the widely considered hybrid precoding architecture for XL-MIMO can hardly benefit from the increased DoFs provided by the near-field LoS path. Due to the limited DoFs of the far-field MIMO channel, the reduced number of RF chains for the classical hybrid precoding is still larger than the spatial DoFs. Thus, the classical hybrid precoding with a much reduced number of RF chains can fully utilize the spatial DoFs to achieve the near-optimal spectrum efficiency in the far-field region [10]. By contrast, this reduced number of RF chains is much smaller than the increased DoFs in the near-field region, resulting in very limited data streams. That is to say, the classical hybrid precoding architecture can hardly utilize the increased spatial DoFs in the near-field region. To the best of our knowledge, this important problem of how to

exploit the extra spatial DoFs brought by the near-field effect to improve the capacity has not been studied yet.

To fill in this gap, in this paper, the theoretical analysis of DoFs of near-field channel is provided, and the precoding architecture for near-field MIMO communications is investigated. First, we introduce the eigenproblem in the electromagnetic theory to investigate the increased DoFs of the near-field LoS channel. Then, we propose a distance-aware precoding (DAP) architecture, which provides a new possibility to improve the capacity by utilizing the near-field effect. The key idea is to adaptively adjust the number of RF chains to match the distance-related DoFs in the near-field region. Based on the DAP architecture, a DAP algorithm is proposed to optimize the spectrum efficiency. Finally, simulations verify that, the proposed DAP scheme can efficiently utilize the extra DoFs to improve the spectrum efficiency in the near-field region.

Notations: $\mathbb{C}$ denotes the set of complex numbers; $[\cdot]^{-1}$, $[\cdot]^T$, $[\cdot]^H$ and $\mathrm{diag}(\cdot)$ denote the inverse, transpose, conjugate-transpose and diagonal operations, respectively; $\|\cdot\|_F$ denotes the Frobenius norm of the matrix; $|\cdot|$ and $\angle[\cdot]$ denotes the norm and the angle of its complex argument; $\mathbf{I}_N$ is an $N\times N$ identity matrix; $\mathbf{1}_L$ denotes an L-length vector with all elements being 1.

II. System Model

In this paper, we consider a single-user XL-MIMO communication system. The transmitter and receiver are equipped with N_t and N_r antennas, respectively. To enable the transmission of N_s data streams, the transmitter and receiver contain N_t^RF and N_r^RF RF chains, satisfying $N_\mathrm{s}\le N_\mathrm{t}^\mathrm{RF}\le N_\mathrm{t}$ and $N_\mathrm{s}\le N_\mathrm{r}^\mathrm{RF}\le N_\mathrm{r}$. We assume that the number of RF chains equals the number of data streams for both transmitter and receiver. Thus, the $N_\mathrm{s}\times 1$ symbol vector $\mathbf{s}$ is followed by a $N_\mathrm{s}\times N_\mathrm{s}$ digital precoder $\mathbf{F}_\mathrm{D}$ and $N_\mathrm{t}\times N_\mathrm{s}$ analog precoder $\mathbf{F}_\mathrm{A}$. The received signal is expressed as

$$\mathbf{y}=\mathbf{H}\mathbf{F}_\mathrm{A}\mathbf{F}_\mathrm{D}\mathbf{s}+\mathbf{n}, \tag{1}$$

where $\mathbf{y}$ denotes the $N_\mathrm{r}\times 1$ received vector, $\mathbf{H}$ is the $N_\mathrm{r}\times N_\mathrm{t}$ channel matrix, and $\mathbf{n}$ represents the noise vector $\mathbf{n}=[n_1,n_2,\cdots,n_{N_\mathrm{r}}]$, whose entities are mutually independent and follow the complex Gaussian distribution $\mathcal{CN}(0,\sigma_\mathrm{n}^2)$.

Due to the high attenuation of millimeter-wave and terahertz bands, LoS path dominants in the XL-MIMO channel. The classical far-field channel can be approximated as

$$\mathbf{H}\approx\alpha_\mathrm{LoS}\Lambda_\mathrm{r}(\phi_\mathrm{LoS}^\mathrm{r})\Lambda_\mathrm{t}(\phi_\mathrm{LoS}^\mathrm{t})\mathbf{a}_\mathrm{r}(\phi_\mathrm{LoS}^\mathrm{r})\mathbf{a}_\mathrm{t}^H(\phi_\mathrm{LoS}^\mathrm{t}), \tag{2}$$

where α_LoS is the gain of the LoS path, $\phi_\mathrm{LoS}^\mathrm{t}$ and $\phi_\mathrm{LoS}^\mathrm{r}$ are the azimuth angles of departure and arrival (AoDs/AoAs) for LoS path, respectively. The functions $\Lambda_\mathrm{t}(\phi_\mathrm{LoS}^\mathrm{t})$ and $\Lambda_\mathrm{r}(\phi_\mathrm{LoS}^\mathrm{r})$ represent the antenna element gain for the transmitter and receiver at the corresponding angle, respectively. The antenna array response vectors $\mathbf{a}_\mathrm{t}$ and $\mathbf{a}_\mathrm{r}$ are introduced to capture the spatial correlation characteristics for MIMO communications.

However, with the extremely large number of antennas in both transmitter and receiver, wireless communications may lie in the near-field region. Then the far-field LoS channel in (2) is no longer accurate to describe the spatial relations. As shown in Fig. 1, based on the geometric free space assumption [3], the near-field LoS channel can be modeled as

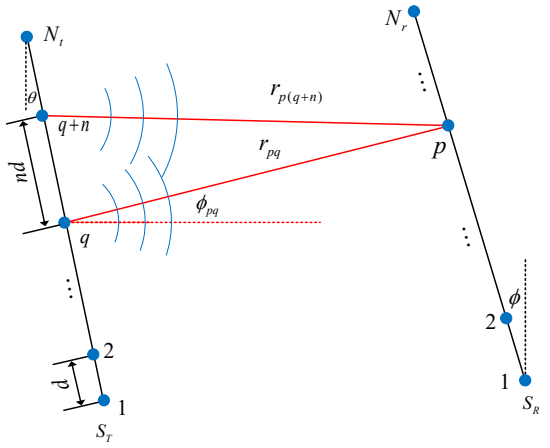

Fig. 1. The near-field geometric channel model.

$$\mathbf{H}=\begin{bmatrix}\alpha_{11}e^{-j\frac{2\pi}{\lambda}r_{11}} & \cdots & \alpha_{1N_\mathrm{t}}e^{-j\frac{2\pi}{\lambda}r_{1N_\mathrm{t}}}\\ \vdots & \ddots & \vdots\\ \alpha_{N_\mathrm{r}1}e^{-j\frac{2\pi}{\lambda}r_{N_\mathrm{r}1}} & \cdots & \alpha_{N_\mathrm{r}N_\mathrm{t}}e^{-j\frac{2\pi}{\lambda}r_{N_\mathrm{r}N_\mathrm{t}}}\end{bmatrix}, \tag{3}$$

where r_{pq} and α_{pq} denotes the distance and normalized path gain between p^{th} receiver antenna and q^{th} transmitter antenna. It is worth noting that ULA is considered here for simplicity. It is straightforward to deploy the more realistic planar arrays.

Remark 1: It can be proved that, the far-field channel model in (2) is a particular case of the near-field channel in (3) when the array apertures $(N_\mathrm{r}-1)d$ and $(N_\mathrm{t}-1)d$ can be negligible compared with the communication distance r.

Moreover, it can be obtained that the rank of the far-field LoS channel in (2) is only 1, i.e., has very limited DoFs. Therefore, with much fewer RF chains, the classical hybrid precoding architecture can fully utilize the limited spatial DoFs to achieve the near-optimal spectrum efficiency. By contrast, the rank of the near-field LoS channel in (3) can be huge, which indicates a significantly increased multiplexing gain compared with the far-field LoS channel.

As illustrated in [11], the DoFs of *parallel* continuous linear arrays are researched by incorporating the prolate spheroidal wave functions (PSWFs). The singular values of the near-field LoS channel can be approximated with the eigenvalues of PSWFs. And there is a critical value after which the eigenvalues of PSWFs (or singular values of the channel) fall off exponentially, which reflects the DoFs of the near-field LoS channel. We generalize the results of DoFs in [11] to a pair of *non-parallel* positioned linear arrays as

$$N_\mathrm{DoF}(r)\approx\frac{(N_\mathrm{t}-1)(N_\mathrm{r}-1)d^2\cos\theta\cos\phi}{\lambda r}, \tag{4}$$

where d, λ and r denotes the element spacing, carrier wavelength and transmission distance, respectively. The arrays are positioned as in Fig. 1. The angle between the array and vertical line is denoted by θ and ϕ for transmitter and receiver, respectively. As the distance decreases, the DoFs significantly improves, which enables the communication systems to transmit more data streams for capacity improvement.

To verify the effectiveness of the estimation method utilizing the eigenvalues of PSWFs, the comparison of the calculated singular values through singular value decomposition (SVD) and estimated singular values with PSWFs for parallel arrays

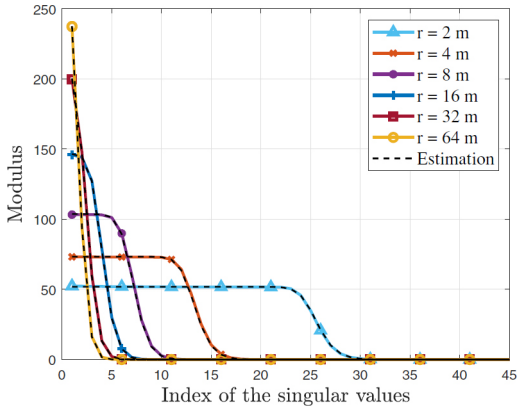

Fig. 2. Calculated and estimated singular values for a pair of parallel 256 ULA. Colored solid lines represent the calculated singular values based on SVD on different distances, while black dashed lines are derived with eigenvalues of PSWFs with equal power for corresponding distances.

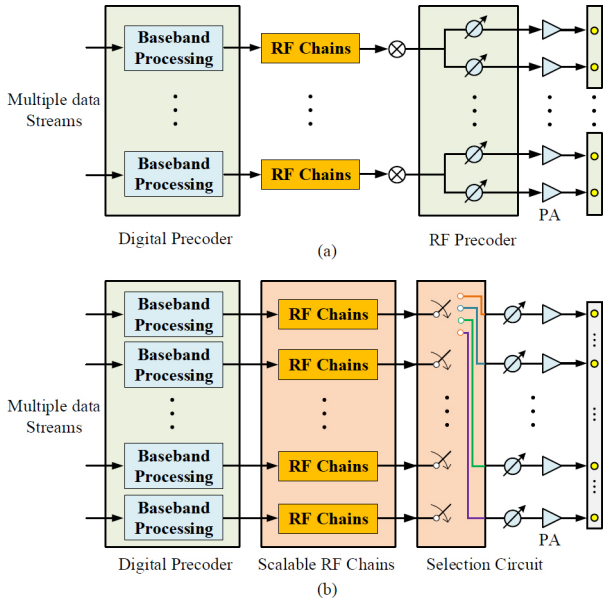

Fig. 3. Comparison of the precoding architectures: (a) Classical sub-connected hybrid precoding architecture; (b) Proposed DAP architecture.

is provided in Fig. 2. It indicates that the singular values can be accurately estimated without complicated SVD when the size of channel matrix $\mathbf{H}$ becomes huge.

III. Proposed DAP Architecture

With the small DoFs of the far-field LoS channel, the classical far-field hybrid precoding with a reduced number of RF chains could achieve near-optimal spectrum efficiency. However, with the increased DoFs of the near-field LoS channel, classical hybrid precoding with limited RF chains is unable to utilize the extra DoFs to approach the optimal spectrum efficiency. To our best knowledge, an architecture that can flexibly change the number of RF chains to match the increased DoFs in the near-field region is still a blank.

To efficiently utilize the distance-related DoFs of the near-field LoS channel, we develop a DAP architecture that can flexibly control the number of activated RF chains to match the varying DoFs as shown in Fig. 3.

Specifically, the sub-connected architecture is adopted in the proposed architecture, in which each antenna element only connects to one RF chain. To enable a variable number of RF chains, extra RF chains are equipped in the transmitter and receiver and each RF chain can be configured to active or inactive. In addition, each antenna element can dynamically select one arbitrary RF chain through a selection circuit. The RF chains without any connection can be regarded as inactive, thus enabling the control of the number of RF chains.

Based on the proposed DAP structure, the received signal in equation (1) can be rewritten as

$$\mathbf{y} = \mathbf{H}\mathbf{F}_{\mathrm{A}}\mathbf{F}_{\mathrm{S}}\mathbf{F}_{\mathrm{D}}\mathbf{s} + \mathbf{n}, \tag{5}$$

where the $N_{\mathrm{t}} \times N_{\mathrm{s}}$ selection matrix $\mathbf{F}_{\mathrm{S}}$ represents the adjustable selection matrix between array elements and activated RF chains. Each entity of $\mathbf{F}_{\mathrm{S}}$ is assigned either 0 or 1. Since each antenna element can only select one RF chain, each row of $\mathbf{F}_{\mathrm{S}}$ contains one element that equals 1. Besides, since the analog precoder is implemented by analog phase shifters, $\mathbf{F}_{\mathrm{A}}$ has an extra constant norm constraint. Different from the classical hybrid precoding architecture, the introduction of selection matrix changes the size of analog precoder, which is a $N_{\mathrm{t}} \times N_{\mathrm{t}}$ diagonal matrix with constant norm constraints as

$$\mathcal{F} \triangleq \left\{ \mathrm{diag}(\widetilde{\theta}_1, \cdots, \widetilde{\theta}_{N_{\mathrm{t}}}) \;\middle|\; \widetilde{\theta}_i \in \mathbb{C},\ \left|\widetilde{\theta}_i\right| = 1,\ \forall i \right\}. \tag{6}$$

When considering the precoding algorithm to maximize the spectrum efficiency, we can assume that an optimal combiner is employed at the receiver side [2]. Thus, the spectrum efficiency can be written as

$$R = \log_2 \left(\left| \mathbf{I} + \frac{1}{\sigma_{\mathrm{n}}^2} \mathbf{H}\mathbf{F}_{\mathrm{A}}\mathbf{F}_{\mathrm{S}}\mathbf{F}_{\mathrm{D}}\mathbf{F}_{\mathrm{D}}^H\mathbf{F}_{\mathrm{S}}^H\mathbf{F}_{\mathrm{A}}^H\mathbf{H}^H \right| \right). \tag{7}$$

IV. Proposed DAP Algorithm

A. Overview of the DAP Algorithm

Based on the proposed DAP structure, the key problem is to determine the number of data streams N_{s}. As N_{s} increases, the proposed DAP architecture approaches the fully-digital precoding architecture, which is considered to be ideal in terms of spectrum efficiency but with high power consumption. Similar to the hybrid precoding scheme in the far-field region, since the DoFs can be well estimated, a reduced number of RF chains can be deployed to match the DoFs in the near-field region to achieve the near-optimal spectrum efficiency. With fixed N_{s}, the optimization can be formulated as

$$\begin{aligned} \mathcal{P}: \max_{\mathbf{F}_{\mathrm{A}},\mathbf{F}_{\mathrm{S}},\mathbf{F}_{\mathrm{D}}} \ & R(\mathbf{F}_{\mathrm{A}},\mathbf{F}_{\mathrm{S}},\mathbf{F}_{\mathrm{D}}) \\ \text{s.t. } & C_1: \|\mathbf{F}_{\mathrm{A}}\mathbf{F}_{\mathrm{S}}\mathbf{F}_{\mathrm{D}}\|_F^2 \le P_{\mathrm{tot}} \\ & C_2: \mathbf{F}_{\mathrm{A}} \in \mathcal{F} \\ & C_3: (\mathbf{F}_{\mathrm{S}})_{ij} \in \{0,1\},\ \forall i,j \\ & C_4: \mathrm{diag}(\mathbf{F}_{\mathrm{S}}\mathbf{F}_{\mathrm{S}}^H) = \mathbf{1}_{N_t}. \end{aligned} \tag{8}$$

The precoding algorithm can be summarized into three stages. First, the number of data streams as well as the number of RF chains N_{s} is determined according to the DoFs of the near-field LoS channel. Then, we determine the selection matrix $\mathbf{F}_{\mathrm{S}}$ by optimizing the partitioning pattern of arrays, which aims to choose the best pattern of subarrays that

maximizes the spectrum efficiency. Finally, we can obtain the corresponding analog precoder $\mathbf{F}_{\mathrm{A}}$ and digital precoder $\mathbf{F}_{\mathrm{D}}$.

B. Estimation of Available Data Streams N_{s}

Since the eigenvalues of PSWFs offer an accurate approximation of the singular values of the near-field LoS channel, the singular values of the channel can be obtained when the transmission distance r can be acquired. Then, we could adopt a water-filling process to obtain the DoFs. Precisely, N_{DoF} is equal to the number of non-zero power allocated to the sub-channels in the water-filling process. After obtaining N_{DoF}, similar to the classical hybrid precoding architectures, we set $N_{\mathrm{s}} = N_{\mathrm{DoF}}$ to fully utilize the spatial multiplexing gain for capacity improvement at minimum energy consumption.

Remark 2: The choice of $N_{\mathrm{s}} = N_{\mathrm{DoF}}$ is not optimal, since increased number of RF chains could further approach fully-digital precoding. Yet $N_{\mathrm{s}} \geq N_{\mathrm{DoF}}$ is an essential condition to guarantee the full utilization of spatial multiplexing gain.

C. Optimization of Digital Precoder $\mathbf{F}_{\mathrm{D}}$

With a fixed number of data streams N_{s}, to simplify the formulation in (8), we define the effective channel $\mathbf{H}_{\mathrm{e}} \triangleq \mathbf{H}\mathbf{F}_{\mathrm{A}}\mathbf{F}_{\mathrm{S}}$. By assuming the SVD of effective channel $\mathbf{H}_{\mathrm{e}} = \mathbf{U}_{\mathrm{e}}\mathbf{\Sigma}_{\mathrm{e}}\mathbf{V}_{\mathrm{e}}^{H}$, we can obtain the solution of digital precoder

$$\mathbf{F}_{\mathrm{D}}^{\mathrm{opt}} = \mathbf{V}_{\mathrm{e}}\mathbf{\Gamma}, \tag{9}$$

where $\mathbf{\Gamma} = \mathrm{diag}(\sqrt{p_1}, \sqrt{p_2} \cdots, \sqrt{p_{N_s}})$ denotes the diagonal power allocation matrix. Then the spectrum efficiency is rewritten as

$$R = \sum_{i=1}^{N_{\mathrm{s}}} \log_2\left(1 + \frac{\lambda_i^2(\mathbf{H}\mathbf{F}_{\mathrm{A}}\mathbf{F}_{\mathrm{S}})p_i}{\sigma_{\mathrm{n}}^2}\right), \tag{10}$$

where $\lambda_i^2(\cdot)$ denotes the i^{th} singular value of the matrix. The elements of $\mathbf{\Gamma}$ can be calculated through the water-filling process with the power constraint $\sum_i p_i = P_{\mathrm{tot}}$.

D. Optimization of Selection Matrix $\mathbf{F}_{\mathrm{S}}$

We denote the set of array indices that connect to the i^{th} RF chain as $\mathcal{S}_i = \left\{n_{i1}, n_{i2}, \cdots, n_{|\mathcal{S}_i|}\right\}$, where $|\mathcal{S}_i|$ denotes the cardinal of $\mathcal{S}_i$. The selection matrix $\mathbf{F}_{\mathrm{S}}$ represents the linking pattern of the RF chains and arrays and can be easily obtained with $\mathcal{S}_i$. According to Jensen's inequality, the spectrum efficiency can be further rewritten as

$$\begin{aligned} R &= \sum_{i=1}^{N_{\mathrm{s}}} \log_2\left(1 + \frac{\lambda_i^2(\mathbf{H}_{\mathrm{e}}\mathbf{F}_{\mathrm{S}})p_i}{\sigma_{\mathrm{n}}^2}\right) \\ &\leq N_{\mathrm{s}} \log_2\left(1 + \frac{1}{N_{\mathrm{s}}}\sum_{i=1}^{N_{\mathrm{s}}} \frac{\lambda_i^2(\mathbf{H}_{\mathrm{e}}\mathbf{F}_{\mathrm{S}})p_i}{\sigma_{\mathrm{n}}^2}\right). \end{aligned} \tag{11}$$

As illustrated in Fig. 2, the singular values of the near-field LoS channel are close to each other if N_{s} satisfies $N_{\mathrm{s}} \leq N_{\mathrm{DoF}}$. Thus the approximation is tight for near-field XL-MIMO systems. Then, the optimization of selection matrix $\mathbf{F}_{\mathrm{S}}$ and analog precoder $\mathbf{F}_{\mathrm{A}}$ can be simplified as

$$(\mathbf{F}_{\mathrm{A}}^{\mathrm{opt}}, \mathbf{F}_{\mathrm{S}}^{\mathrm{opt}}) = \underset{\mathbf{F}_{\mathrm{A}}, \mathbf{F}_{\mathrm{S}}}{\mathrm{argmax}} \sum_{i=1}^{N_{\mathrm{s}}} \lambda_i^2(\mathbf{H}\mathbf{F}_{\mathrm{A}}\mathbf{F}_{\mathrm{S}}). \tag{12}$$

Utilizing the feature of $\mathbf{F}_{\mathrm{S}}$, $\mathbf{H}\mathbf{F}_{\mathrm{A}}\mathbf{F}_{\mathrm{S}}$ can be rewritten as

$$\mathbf{H}\mathbf{F}_{\mathrm{A}}\mathbf{F}_{\mathrm{S}} = \mathbf{H}\mathbf{F}_{\mathrm{A}}\mathbf{P}_{\mathrm{S}}\widetilde{\mathbf{F}}_{\mathrm{S}} = \mathbf{H}\mathbf{P}_{\mathrm{S}}\widetilde{\mathbf{F}}_{\mathrm{A}}\widetilde{\mathbf{F}}_{\mathrm{S}}, \tag{13}$$

where $\mathbf{P}_{\mathrm{S}}$ denotes a permutation matrix, $\widetilde{\mathbf{F}}_{\mathrm{A}}$ denotes the diagonal analog precoder with a permutation operation of $\mathbf{F}_{\mathrm{A}}$ and $\widetilde{\mathbf{F}}_{\mathrm{S}}$ follows a block diagonal pattern of

$$\widetilde{\mathbf{F}}_{\mathrm{S}} = \begin{bmatrix} \mathbf{1}_{\mathcal{S}_1} & 0 & \cdots & 0 \\ 0 & \mathbf{1}_{\mathcal{S}_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{1}_{\mathcal{S}_{N_{\mathrm{s}}}} \end{bmatrix}, \tag{14}$$

where $\mathbf{1}_{\mathcal{S}_i}$ denotes a vector with $|\mathcal{S}_i|$ elements equal to 1. The equation (13) is derived from the property of the permutation matrix $\mathbf{P}_{\mathrm{S}}$, whose order can be switched with the diagonal analog precoder $\mathbf{F}_{\mathrm{A}}$ with an extra permutation operation.

Due to the block diagonal feature of $\widetilde{\mathbf{F}}_{\mathrm{S}}$, $\widetilde{\mathbf{H}} = \mathbf{H}\mathbf{P}_{\mathrm{S}}$ can be classified by columns as submatrices $\widetilde{\mathbf{H}} = [\mathbf{H}_{\mathcal{S}_1}, \cdots, \mathbf{H}_{\mathcal{S}_{\mathrm{Ns}}}]$. Therefore, the optimization object in (12) can be rewritten as

$$\begin{aligned} &\max \sum_{i=1}^{N_{\mathrm{s}}} \lambda_i^2(\mathbf{H}\mathbf{F}_{\mathrm{A}}\mathbf{F}_{\mathrm{S}}) \\ &= \max \left\| \mathbf{H}\mathbf{P}_{\mathrm{S}}\widetilde{\mathbf{F}}_{\mathrm{A}}\widetilde{\mathbf{F}}_{\mathrm{S}} \right\|_F^2 \\ &= \max \left\| \left[\mathbf{H}_{\mathcal{S}_1}\mathbf{f}_1, \cdots, \mathbf{H}_{\mathcal{S}_{N_{\mathrm{s}}}}\mathbf{f}_{N_{\mathrm{s}}}\right] \right\|_F^2 \\ &\overset{(a)}{\approx} \sum_{i=1}^{N_{\mathrm{s}}} \lambda_1(\mathbf{H}_{\mathcal{S}_i}^H \mathbf{H}_{\mathcal{S}_i}), \end{aligned} \tag{15}$$

where $\mathbf{f}_i$ denotes the non-zero elements of the i^{th} colomn of $\widetilde{\mathbf{F}}_{\mathrm{A}}\widetilde{\mathbf{F}}_{\mathrm{S}}$. The approximation (a) is derived by omiting the constant modulus constraints of $\mathbf{F}_{\mathrm{A}}$. Therefore, the optimization object can be formulated as the summation of the largest singular value of submatrices. Since the complexity of exhaustive searching to find the sets $\mathcal{S}$ is unaffordable, we propose a low-complexity algorithm to optimize the selection matrix $\mathbf{F}_{\mathrm{S}}$ by utilizing normalized Minkowski ℓ_1-norm as a approximation of the largest singular value as

$$\lambda_1(\mathbf{R}_{\mathcal{S}}) \approx \hat{\lambda}_1(\mathbf{R}, \mathcal{S}) = \frac{1}{|\mathcal{S}|}\sum_{i \in \mathcal{S}}\sum_{j \in \mathcal{S}} |[\mathbf{R}_{i,j}]|. \tag{16}$$

Therefore, we could perform a low-complex greedy searching process to maximize the average ℓ_1-norm for each sets $\mathcal{S}_i$. The proposed algorithm is summarized in **Algorithm 1**.

Different from dynamic subarray design algorithms in [12] and [13], we aim to obtain subarrays with similar scales. As illustrated in Fig. 2, since the singular values of the near-field LoS channel are close to each other, the scale of subarrays should be similar to generate equal beamforming gain. First, we initialize all sets $\mathcal{S}$ with uniformly distributed antennas. Then, we add the antennas to sets one by one by selecting the j^{th} antenna with the largest absolute value in which j^{th} antenna is not added yet. If the cardinal of sets exceeds the maximum scale N_{bound}, we remove the m^{th} antenna with the least contribution and add it to the set with the largest improvement. Finally, we check each set to remove the least contributor, which aims to eliminate the influence of the manual initialization.

E. Optimization of Analog Precoder $\mathbf{F}_{\mathrm{A}}$

As illustrated in (15). The optimal analog precoder corresponding to the i^{th} subarray can be formulated as

$$\angle(\mathbf{f}_{\mathrm{A},\mathcal{S}_i}^{\mathrm{opt}}) = \angle(\mathbf{v}_{\mathcal{S}_i}), \tag{17}$$

Algorithm 1 Near-Field Subarray Partitioning Alogorithm.

Input: Channel $\mathbf{H}$, N_{s}, N_{bound} and N_{t}.
Output: $\mathcal{S}_1, \mathcal{S}_2, \cdots, \mathcal{S}_{N_{\mathrm{s}}}$

$\mathbf{R} = \mathbf{H}^H\mathbf{H}$, $\mathcal{S}_{\mathrm{sel}} = \varnothing$, $n_{\mathrm{group}} = \lfloor \frac{N_{\mathrm{t}}}{N_{\mathrm{s}}} \rfloor$
Initialize $\mathcal{S}_i = \{i \cdot n_{\mathrm{group}}\}$, $\mathcal{S}_{\mathrm{sel}} \leftarrow \mathcal{S}_{\mathrm{sel}} \cup \{i \cdot n_{\mathrm{group}}\}$, for $i = 1, 2, \cdots, N_{\mathrm{s}}$
for $k = 1 : N_{\mathrm{t}} - N_{\mathrm{s}}$ **do**
  $\{i_k, j_k\} = \underset{i \in \mathcal{S}_{\mathrm{sel}}, j \notin \mathcal{S}_{\mathrm{sel}}}{\operatorname{argmax}} |[\mathbf{R}]_{i,j}|$
  $\hat{r} = \underset{r \in \{1,\cdots,N_{\mathrm{s}}\}}{\operatorname{argmax}} \hat{\lambda}_1(\mathbf{R}, \mathcal{S}_r \cup \{j_k\}) - \hat{\lambda}_1(\mathbf{R}, \mathcal{S}_r)$
  $\mathcal{S}_{\mathrm{sel}} \leftarrow \mathcal{S}_{\mathrm{sel}} \cup j_k$, $\mathcal{S}_{\hat{r}} \leftarrow \mathcal{S}_{\hat{r}} \cup j_k$
  if $|\mathcal{S}_{\hat{r}}| \geq N_{\mathrm{bound}}$ **then**
    $\hat{m} = \underset{m \in \mathcal{S}_{\hat{r}}}{\operatorname{argmin}} \sum_{n \in \mathcal{S}_{\hat{r}}} |\mathbf{R}_{m,n}|$
    $\hat{r}' = \underset{r' \neq \hat{r}}{\operatorname{argmax}} \hat{\lambda}_1(\mathbf{R}, \mathcal{S}_{r'} \cup \hat{m}) - \hat{\lambda}_1(\mathbf{R}, \mathcal{S}_{r'})$
    $\mathcal{S}_{\hat{r}} \leftarrow \mathcal{S}_{\hat{r}} \backslash \hat{m}$, $\mathcal{S}_{\hat{r}'} \leftarrow \mathcal{S}_{\hat{r}'} \cup \hat{m}$
  end if
end for
for $l = 1 : N_{\mathrm{s}}$ **do**
  $\hat{m} = \underset{m \in \mathcal{S}_l}{\operatorname{argmin}} \sum_{n \in \mathcal{S}_l} |\mathbf{R}_{m,n}|$
  $\hat{r}' = \underset{r'}{\operatorname{argmax}} \hat{\lambda}_1(\mathbf{R}, \mathcal{S}_{r'} \cup \hat{m}) - \hat{\lambda}_1(\mathbf{R}, \mathcal{S}_{r'})$
  $\mathcal{S}_{\hat{r}} \leftarrow \mathcal{S}_{\hat{r}}^{r'} \backslash \hat{m}$, $\mathcal{S}_{\hat{r}'} \leftarrow \mathcal{S}_{\hat{r}'} \cup \hat{m}$
end for

where $\mathbf{v}_{\mathcal{S}_i}$ is the singular vector corresponding to the largest singular value of $\mathbf{H}_{\mathcal{S}_i}$. Therefore, each array $\mathbf{f}_{\mathrm{A},\mathcal{S}_i}^{\mathrm{opt}}$ can be determined satisfying the constraints of the norms. Finally, the formulation of analog precoder can be written as

$$\mathbf{F}_{\mathrm{A}}^{\mathrm{opt}} = \mathbf{P}_{\mathrm{S}} \mathrm{diag}\left(\left[(\mathbf{f}_{\mathrm{A},\mathcal{S}_1}^{\mathrm{opt}})^T, \cdots, (\mathbf{f}_{\mathrm{A},\mathcal{S}_{\mathrm{Ns}}}^{\mathrm{opt}})^T\right]\right) \mathbf{P}_{\mathrm{S}}^T. \quad (18)$$

V. Simulation Results

We consider a single-user XL-MIMO communication scenario, where the transmitter and receiver are equipped with ULA with 256 elements spacing of half wavelength. The carrier wave is set to 100 GHz. We assume that the large-scale channel fading in (3) can be ignored through perfect power control.

The spectrum efficiency for the proposed DAP architecture over different distances is shown in Fig. 4. The SNR $= P_{\mathrm{tot}}/\sigma_n^2$ is set to 30 dB. The baselines include fully-digital precoding, fully-connected hybrid precoding with 8, 4 RF chains [14], and sub-connected hybrid precoding with 8, 4 RF chains [15]. When the transmitter-receiver distance is small, since the proposed architecture can benefit from the extra DoFs in the near-field region, the proposed scheme can achieve about 40% increase in the spectrum efficiency compared to classical hybrid precoding schemes with 8 RF chains.

VI. Conclusions

In this paper, the significantly increased DoFs in the near-field region are theoretically analyzed. To utilize the increased spatial DoFs, a DAP architecture and the corresponding precoding algorithm are proposed for XL-MIMO communication scenarios, in which the number of activated RF chains can be flexibly adjusted to match the increased DoFs in the near-field region. Through simulations, the proposed DAP scheme can efficiently utilize the increased DoFs to improve the spectrum efficiency. Different from alleviating problems brought by the near-field effect, this paper has revealed a novel method to utilize the near-field property for capacity improvement.

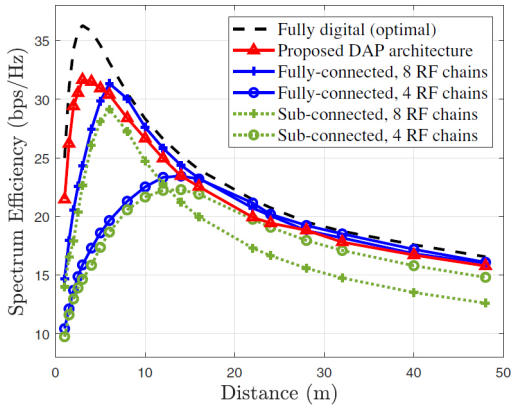

Fig. 4. The comparison of spectrum efficiency over different distances.

References

[1] E. Björnson, L. Sanguinetti, H. Wymeersch, J. Hoydis, and T. L. Marzetta, "Massive MIMO is a reality—what is next?: Five promising research directions for antenna arrays," *Digit. Signal Process*, vol. 94, pp. 3–20, Nov. 2019.

[2] O. Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1499–1513, Jan. 2014.

[3] J. Sherman, "Properties of focused apertures in the fresnel region," *IEEE Trans. Antennas Propag.*, vol. 10, no. 4, pp. 399–408, Jul. 1962.

[4] M. Cui, L. Dai, R. Schober, and L. Hanzo, "Near-field wide-band beamforming for extremely large antenna array," *arXiv preprint arXiv:2109.10054*, Sep. 2021.

[5] D. Headland, Y. Monnai, D. Abbott, C. Fumeaux, and W. Withayachumnankul, "Tutorial: Terahertz beamforming, from concepts to realizations," *APL Photon.*, vol. 3, no. 5, 2018, Art. No. 051101.

[6] H. Zhang, N. Shlezinger, F. Guidi, D. Dardari, M. F. Imani, and Y. C. Eldar, "Beam focusing for near-field multi-user MIMO communications," *arXiv preprint arXiv:2105.13087*, May 2021.

[7] Y. Han, S. Jin, C. Wen, and X. Ma, "Channel estimation for extremely large-scale massive mimo systems," *IEEE Commun. Lett.*, vol. 9, no. 5, pp. 633–637, May 2020.

[8] Z. Zhang and L. Dai, "Continuous-aperture MIMO for electromagnetic information theory," *arXiv preprint arXiv:2111.08630*, Nov. 2021.

[9] N. Decarli and D. Dardari, "Communication modes with large intelligent surfaces in the near field," *arXiv preprint arXiv:2108.10569*, Aug. 2021.

[10] R. W. Heath, N. González-Prelcic, S. Rangan, W. Roh, and A. M. Sayeed, "An overview of signal processing techniques for millimeter wave mimo systems," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 436–453, Apr. 2016.

[11] D. A. B. Miller, "Communicating with waves between volumes: evaluating orthogonal spatial channels and limits on coupling strengths," *Appl. Opt.*, vol. 39, no. 11, pp. 1681–1699, Apr. 2000.

[12] S. Park, A. Alkhateeb, and R. W. Heath, "Dynamic subarrays for hybrid precoding in wideband mmwave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 5, pp. 2907–2920, May 2017.

[13] K. Xu, F. C. Zheng, P. Cao, H. Xu, and X. Zhu, "A low complexity greedy algorithm for dynamic subarrays in mmwave MIMO systems," in *2019 IEEE 90th Vehicular Technology Conference (VTC2019-Fall)*, Sep. 2019, pp. 1–5.

[14] X. Yu, J. Z. J. Shen, and K. B. Letaief, "Alternating minimization algorithms for hybrid precoding in millimeter wave MIMO systems," *IEEE J. Sel. Areas Commun.*, vol. 10, no. 3, pp. 485–500, Apr. 2016.

[15] X. Gao, L. Dai, S. Han, C.-L. I, and R. W. Heath, "Energy-efficient hybrid analog and digital precoding for mmwave MIMO systems with large antenna arrays," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 4, pp. 998–1009, Apr. 2016.