SCIENCE CHINA
Information Sciences

• RESEARCH PAPER •

March 2026, Vol. 69, Iss. 3, 132302:1–132302:20
https://doi.org/10.1007/s11432-025-4585-y

Near-field joint spatial-division and multiplexing for XL-MIMO communications

Zhenjun DONG[1], Xinrui LI[1], Yong ZENG[1,2*], Jianhua ZHANG[3], Shi JIN[1] & Tao JIANG[4]

[1]*National Mobile Communications Research Laborator, Southeast University, Nanjing 210096, China*
[2]*Purple Mountain Laboratories, Nanjing 211111, China*
[3]*State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China*
[4]*Research Center of 6G Mobile Communications, School of Cyber Science and Engineering, Huazhong University of Science and Technology, Wuhan 430074, China*

Received 20 March 2025/Revised 21 July 2025/Accepted 9 September 2025/Published online 20 January 2026

Abstract For extremely large-scale multiple-input multiple-output (XL-MIMO) communications, the ultra-large antenna array results in excessive channel estimation overhead. Joint spatial division and multiplexing (JSDM) can alleviate this by using prior knowledge of user groups' channel spatial covariance to design a pre-beamforming matrix. This allows the system to estimate only the low-dimensional equivalent channel after pre-beamforming, rather than the full high-dimensional instantaneous channel, thereby significantly reducing channel estimation overhead. However, the existing JSDM, which relies on the assumptions of a zero-mean channel model, the uniform plane wave (UPW) model, and identical channel statistics for all users within the same group, is no longer suitable for XL-MIMO communications due to new channel characteristics, including near-field non-uniform spherical wave (NUSW) and partial array visibility. Therefore, we propose a near-field JSDM (NF-JSDM) approach for XL-MIMO communications, which leverages the developed near-field statistical channel state information (CSI), including both channel mean and covariance. Compared to the conventional far-field JSDM (FF-JSDM), the proposed NF-JSDM considers a more general non-zero-mean channel model and accommodates variations in statistical CSI among users within the same group. An efficient algorithm for user grouping is also proposed for a more precise division of user groups. Similar to the FF-JSDM, the proposed NF-JSDM also reduces channel estimation overhead via statistical CSI-based pre-beamforming. In contrast, the proposed NF-JSDM further exploits both the first-order and second-order near-field statistical CSI, and introduces a two-stage pre-beamforming scheme that further considers the differences in statistical CSI among intra-group users. Simulation results validate the effectiveness of the proposed user grouping algorithm, and demonstrate the superiority of the proposed NF-JSDM over the conventional FF-JSDM for XL-MIMO communications.

Keywords XL-MIMO communications, NF-JSDM, near-field statistical CSI, NUSW model, partial array visibility, user grouping

Citation Dong Z J, Li X R, Zeng Y, et al. Near-field joint spatial-division and multiplexing for XL-MIMO communications. Sci China Inf Sci, 2026, 69(3): 132302, https://doi.org/10.1007/s11432-025-4585-y

1 Introduction

The evolution of wireless communication systems has been accompanied by the development of advanced antenna technology. Evolving from single-antenna communications in early days, multiple-input multiple-output (MIMO) and massive MIMO have emerged as pivotal physical layer technologies for the fourth-generation (4G) and fifth-generation (5G) mobile communication systems, respectively. Throughout the evolution of antenna technology, the number of antennas has continued to expand, from the typical 8 antennas used in 4G to 64 antennas in 5G [1]. This expansion has provided new spatial degrees of freedom (DoFs) without consuming additional time-frequency resources, thereby improving spectral efficiency. To support the ambitious goals of the sixth-generation (6G) mobile communication networks, such as ultra-high localization accuracy, ultra-high connectivity density, and immersive communication, the scale of antenna arrays in massive MIMO systems is anticipated to be further increased by at least one order of magnitude, termed extremely large-scale MIMO (XL-MIMO) [2–12].

However, increasing the antenna size will cause overwhelming channel estimation overhead, particularly in systems without channel reciprocity, such as frequency division duplexing (FDD) systems. In FDD systems, due to the absence of channel reciprocity, the downlink channel state information (CSI) can only be obtained by the base station (BS) sending pilot sequences to user equipments (UEs), which then subsequently perform channel estimation

* Corresponding author (email: yong_zeng@seu.edu.cn)

and feed the estimated CSI back to the BS. Consequently, the length of the orthogonal pilot sequence is proportional to the number of antennas at the BS side, which results in considerable pilot overhead and feedback overhead.

Existing work on reducing downlink channel estimation overhead primarily includes the compressive sensing (CS)-based methods [13,14], channel knowledge map (CKM) [15,16], and the joint spatial division and multiplexing (JSDM) approach [17,18]. The key to CS-based methods lies in utilizing the sparse characteristic of wireless channels in the angle domain [13] or time domain [14]. Specifically, in CS-based channel estimation, the unknown CSI is represented in sparse form either in the time domain or angle domain, and subsequently, the channel estimation problem is rephrased as the problem of sparse signal recovery. CKM is a channel knowledge database that reflects the essential characteristics of wireless channels [15, 16], which can provide prior environmental information in advance, thereby avoiding repeated online estimation or sensing of static environmental elements and reducing channel estimation overhead. However, the CS-based channel estimation depends on the channel spatial sparsity assumption, which cannot be guaranteed in some practical communication scenarios [19, 20], while CKM requires prior channel information for its construction and relies on UE location during its application.

In particular, the JSDM approach has the great potential in reducing downlink channel estimation overhead [17, 18]. It partitions UEs into multiple groups based on their statistical CSI, by assuming that UEs within the same group have identical statistical CSI. In [17,18], the prior information of user groups' statistical CSI is leveraged to reduce the dimension of the effective channels where instantaneous CSI needs to be estimated. Specifically, the JSDM approach considers two stages of precoding, expressed as $\mathbf{V}=\mathbf{B}\mathbf{P}\in\mathbb{C}^{N\times K}$, where $\mathbf{B}\in\mathbb{C}^{N\times b}$ and $\mathbf{P}\in\mathbb{C}^{b\times K}$ are the first- and second-stage precoding, respectively, with K and N denoting the number of UEs and the number of antennas at the BS end, respectively, and $b\in[K,N]$ being an integer that needs to be determined. The first-stage precoding is designed to eliminate inter-group interference, which depends on the statistical CSI of multiple groups only. The second-stage precoding can be designed as the zero-forcing (ZF) precoding for spatial multiplexing, which depends on the equivalent instantaneous CSI after the first-stage precoding, i.e., $\mathbf{H}^{\text{equ}}=\mathbf{B}^{\text{H}}\mathbf{H}\in\mathbb{C}^{b\times K}$, with $\mathbf{H}\in\mathbb{C}^{N\times K}$ being the channels of K UEs. Consequently, the JSDM approach reduces the requirement for $N\times K$-dimensional complete CSI to $b\times K$-dimensional equivalent CSI, thereby significantly reducing the CSI estimation overhead if $b\ll N$. Note that in [17, 18], the multiple-input single-output (MISO) channel of UE k in group g, denoted by $\mathbf{h}_k\in\mathbb{C}^{N\times 1}$, is modeled as

$$\mathbf{h}_k=\mathbf{\Theta}_g^{\frac{1}{2}}\mathbf{z}_k,\ \ \forall\, k\in\mathbb{K}_g, \tag{1}$$

where $\mathbf{\Theta}_g\in\mathbb{C}^{N\times N}$ denotes the spatial covariance matrix of group g with $\text{rank}\,(\mathbf{\Theta}_g)=r_g$, $\mathbf{z}_k\in\mathbb{C}^{r_g\times 1}\sim\mathcal{N}_{\mathbb{C}}(\mathbf{0},\mathbf{I}_{r_g})$ is a circularly symmetric complex Gaussian (CSCG) random vector with zero mean and covariance matrix $\mathbf{I}_{r_g}$, and $\mathbb{K}_g$ denotes the UE set in group g. By observing the correlation-based stochastic model (CBSM) in (1), it is found that the existing JSDM approach is based on three important assumptions.

- *Zero-Mean Assumption.* The channel mean is equal to zero, i.e., $\mathbb{E}\,[\mathbf{h}_k]=\mathbf{\Theta}_g^{\frac{1}{2}}\mathbb{E}\,[\mathbf{z}_k]=\mathbf{0}$.
- *UPW Assumption.* The channel spatial covariance model is based on the conventional far-field uniform plane wave (UPW) model. Specifically, the spatial covariance model is derived from the far-field UPW-based steering vector, given by $\mathbf{a}_{\text{UPW}}=[1,\mathrm{e}^{-\mathrm{j}\frac{2\pi}{\lambda}d\sin\theta},\ldots,\mathrm{e}^{-\mathrm{j}\frac{2\pi}{\lambda}(N-1)d\sin\theta}]^{\text{T}}\in\mathbb{C}^{N\times 1}$ [21], where d and θ denote the distance between adjacent antenna elements and the signal angle, respectively.
- *Identical Statistical CSI Assumption.* Different UEs in the same group are assumed to have the identical statistical CSI, i.e., $\text{cov}(\mathbf{h}_k)=\text{cov}(\mathbf{h}_{k'})=\mathbf{\Theta}_g,\ \forall\, k,k'\in\mathbb{K}_g$.

Therefore, the existing JSDM method may not be general enough due to *Zero-Mean Assumption* and *Identical Statistical CSI Assumption*. Specifically, *Zero-Mean Assumption* makes the existing JSDM unsuitable for the communication scenarios with mixed line-of-sight (LoS) and non-LoS (NLoS) links. Furthermore, the assumption that UEs in the same group have the identical statistical CSI is rather strong. In addition, the existing JSDM will be inapplicable to XL-MIMO communications, as *UPW Assumption* will no longer hold in XL-MIMO communications. This is because the XL-MIMO communications experience new channel characteristics, i.e., near-field non-uniform spherical wave (NUSW) propagation and partial array visibility properties [2, 22, 23]. Specifically, for the NUSW model, the amplitude of each propagation path across array elements is no longer uniform and the phase variation is generally nonlinear, differing from the UPW model with uniform amplitude and linear phase variation across array elements. Moreover, regarding partial array visibility, scatterers and/or UEs can only see some portions of the large-scale array (XL-array), rather than the whole array. Consequently, in XL-MIMO communications, the statistical CSI model needs to consider the more general near-field NUSW model and partial array visibility, rather than the conventional UPW-based model in the existing far-field JSDM (FF-JSDM). In addition, due to these new channel characteristics, there will be differences in the statistical CSI among adjacent UEs in XL-MIMO communications, even if they are so close to each other that they experience the same set of scatterers. Therefore, *Identical Statistical CSI Assumption* does not hold in XL-MIMO communications. In summary, the existing

FF-JSDM approach is no longer applicable to XL-MIMO communications, which motivates our current work.

In this paper, we propose a near-field JSDM (NF-JSDM) approach for XL-MIMO communications to reduce channel estimation overhead, particularly in downlink systems without channel reciprocity, where CSI acquisition is more critical than in the uplink. Although JSDM was originally proposed for downlink, it has also been extended to uplink scenarios [24,25], where receiver-side beamforming based on user group statistical CSI helps reduce processing complexity and inter-user interference. Therefore, with appropriate modifications, the proposed NF-JSDM can be adapted for uplink communications as well. Since the NF-JSDM requires prior knowledge of statistical CSI, we first develop the near-field first- and second-order statistical CSI models, which are characterized by channel mean and covariance, respectively. Based on these models, we propose a user grouping strategy that exploits the similarity of users' near-field statistical CSI, effectively guiding NF-JSDM precoding design. Building on this, we further develop a generalized NF-JSDM framework for XL-MIMO communications. Several recent studies have investigated key aspects of near-field XL-MIMO communications, including channel modeling, precoding, and user grouping. In terms of channel modeling, Ref. [26] considered the NUSW characteristic but assumed an identical spatial correlation matrix, thereby neglecting spatial correlation among antenna elements, while Ref. [27] proposed a spatial correlation model based on visibility regions but neglected NUSW propagation. In precoding, Ref. [28] developed near-field beam focusing techniques for various antenna architectures, while Ref. [29] introduced triple-polarized precoding for holographic MIMO. However, both approaches rely on full CSI acquisition, incurring significant estimation overhead in near-field XL-MIMO communications. For user grouping, Ref. [7] proposed a strategy to enhance sum-rate and suppress grating lobes, yet without considering spatial non-stationarity. In contrast, our work makes three key contributions: (i) a near-field statistical CSI model that jointly captures NUSW and spatial non-stationarity; (ii) a more accurate user grouping strategy based on the similarity of users' near-field statistical CSI; and (iii) an efficient NF-JSDM framework that significantly reduces channel estimation overhead while ensuring high performance in XL-MIMO communications. Our primary contributions are as follows.

• First, we develop the near-field geometry-based stochastic model (GBSM) for multi-user XL-array communications, considering mixed LoS and NLoS links. The novel near-field GBSM takes into account the NUSW and partial array visibility characteristics, and further considers the spatial consistency by introducing the visibility region (VR) at the UE end. Subsequently, we derive the near-field channel mean and covariance from the developed near-field GBSM. Notably, due to the randomness of partial array visibility, the near-field channel mean does not equal the channel LoS component as in the far-field UPW-based model. In addition, the developed near-field channel covariance no longer exhibits spatial wide-sense stationarity (SWSS), which is different from the conventional far-field model. Furthermore, the near-field spatial correlation function (S-CF) of the LoS component exhibits sparsity as a consequence of partial array visibility. On the other hand, the near-field S-CF of the NLoS component depends on the scattered power distribution characterized by the power location spectrum (PLS), which differs from the far-field one that depends on the power angle spectrum (PAS).

• Second, building on the developed near-field statistical CSI as prior knowledge, we propose the NF-JSDM scheme applicable to XL-MIMO communications. The proposed NF-JSDM method is developed under the assumptions of the non-zero-mean channel, NUSW propagation, and non-identical statistical CSI. To validate that the *Identical Statistical CSI Assumption* does not hold in XL-MIMO communications, we first derive the effective eigenspaces of the developed near-field first- and second-order statistical CSI, and then evaluate the similarity between the effective eigenspaces of different UEs. Simulation results demonstrate that there are significant differences in adjacent users' statistical CSI, even if they experience the same set of scatterers. Based on this observation, we propose an efficient user grouping algorithm for the NF-JSDM scheme, which leverages the similarity of near-field effective eigenspaces. Building on this novel user grouping strategy, we further develop a three-stage NF-JSDM framework. The first- and second-stage precoding depend on the prior knowledge of near-field statistical CSI, while the third stage depends on the instantaneous equivalent CSI obtained after the first two stages of precoding. Specifically, the first-stage precoding is designed to eliminate inter-group interference, while the second stage aims to project the eigen-beamforming along the dominant eigenmodes of each UE's effective channel after the first-stage precoding. Additionally, the third-stage precoding employs ZF precoding for spatial multiplexing. Consequently, in the developed NF-JSDM, only the reduced-dimensional instantaneous equivalent CSI needs to be estimated, thereby significantly reducing CSI estimation overhead.

• Lastly, simulation results validate the proposed user grouping algorithm and the developed NF-JSDM for XL-MIMO communications. The results demonstrate the effectiveness of the proposed algorithm for user grouping in multi-user XL-MIMO communications, as evidenced by comparisons with the exhaustive search algorithm and random grouping algorithm. Furthermore, simulation results demonstrate that the proposed NF-JSDM method is applicable to XL-MIMO communications, as compared to the ZF precoding that utilizes complete CSI and the conventional FF-JSDM method.

Table 1 Mathematical notations.

Symbol	Meaning	Symbol	Meaning	Symbol	Meaning
$\mathbb{C}$	Complex space	$\mathbb{R}$	Real space	$(\cdot)^{\mathrm{T}}$	Transpose
$(\cdot)^{\mathrm{H}}$	Hermitian transpose	$(\cdot)^*$	Complex conjugate	$\|\cdot\|$	Euclidean norm
$\|\cdot\|_F$	Frobenius norm	$\mathbb{E}[\cdot]$	Expectation	$\circ$	Hadamard product
$\lceil\cdot\rceil$	Ceiling operation	$\lfloor\cdot\rfloor$	Floor operation	$\det(\cdot)$	Determinant
$\mathrm{rank}(\cdot)$	Rank	$\mathrm{vec}(\cdot)$	Vectorization	$\mathrm{Span}^{\perp}(\mathbf{X})$	Orthogonal complement
$\mathrm{tr}(\cdot)$	Trace	$\mathrm{Span}(\mathbf{X})$	Column space	$\dim(\mathbf{X})$	Dimension of column space
$\mathrm{cov}(\cdot)$	Covariance	$\mathcal{N}_{\mathbb{C}}(0,\sigma^2)$	CSCG distribution		

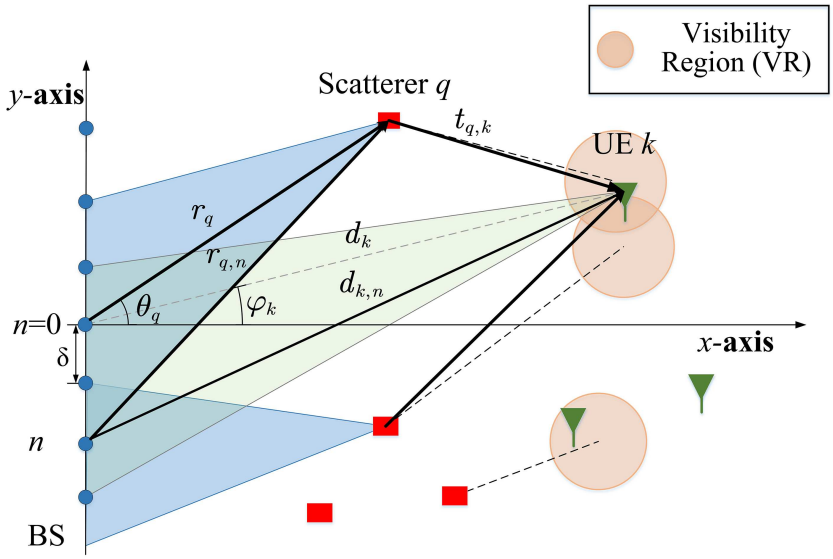

Figure 1 (Color online) A downlink multi-user XL-array communication system in the near-field scattering environment.

The overall structure of this paper is as follows. Section 2 introduces the system model and near-field GBSM. Section 3 derives the near-field channel mean and covariance, which serve as the prior knowledge for NF-JSDM design. Section 4 presents the proposed NF-JSDM in three interrelated parts. First, Subsection 4.1 analyzes the similarity of the effective eigenspace derived from the near-field first- and second-order statistical CSI to verify that the identical statistical CSI assumption does not hold in XL-MIMO communications. Based on this analysis, Subsection 4.2 proposes a refined user grouping algorithm based on eigenspace similarity. Building on the previous analysis and grouping strategy, Subsection 4.3 develops a three-stage NF-JSDM approach. Simulation results are presented in Section 5. Finally, Section 6 concludes the paper. The mathematical notations are summarized in Table 1.

2 System model

As shown in Figure 1, a downlink multi-user MISO (MU-MISO) communication system is considered, where a BS equipped with an XL-array of $N \gg 1$ elements communicates with K single-antenna UEs. Let $\mathbb{K} = \{1, 2, \ldots, K\}$ denote the index set of K UEs. The received signal at UE k, denoted by y_k, is expressed as

$$y_k = \mathbf{h}_k^{\mathrm{H}} \sum_{i \in \mathbb{K}} \mathbf{v}_i a_i + n_k = \mathbf{h}_k^{\mathrm{H}} \mathbf{V}\mathbf{a} + n_k, \quad \forall\, k \in \mathbb{K}, \tag{2}$$

where $\mathbf{h}_k \in \mathbb{C}^{N\times 1}$ is the channel vector between UE k and the BS; $\mathbf{v}_i \in \mathbb{C}^{N\times 1}$ is the precoding vector of UE i with $\|\mathbf{v}_i\| = 1$, and $\mathbf{V} = [\mathbf{v}_i]_{i\in\mathbb{K}} \in \mathbb{C}^{N\times K}$; a_i is the information-bearing symbol of UE i, which satisfies $|a_i|^2 = \frac{P_0}{K}$ and $\mathbb{E}[a_i a_{i'}^*] = 0$, $\forall\, i' \neq i \in \mathbb{K}$, with P_0 being the transmit power and $\mathbf{a} = [a_i]_{i\in\mathbb{K}} \in \mathbb{C}^{K\times 1}$; and $n_k \sim \mathcal{N}_{\mathbb{C}}\left(0, \sigma^2\right)$ is the noise at UE k following independent and identically distributed (i.i.d.) CSCG distribution with variance σ^2. To guarantee effective communication, the channel of K UEs, i.e., $\mathbf{H} = [\mathbf{h}_k]_{k\in\mathbb{K}} \in \mathbb{C}^{N\times K}$, needs to be obtained first. In communication systems without channel reciprocity, the downlink channel estimation overhead is proportional to

the number of antenna elements of the XL-array N. Hence, it is crucial to reduce the channel estimation overhead, especially in XL-MIMO communications.

However, the conventional FF-JSDM method [17, 18], usually used to reduce channel overhead, may not be applicable to XL-MIMO communications, as the aforementioned three assumptions may not hold. In addition, the CBSM in (1) typically adopted by the FF-JSDM method will not be able to characterize XL-MIMO communications effectively. To model the more general near-field CBSM, we consider the non-zero mean channel by utilizing the near-field channel mean and covariance, which is modeled as

$$\mathbf{h}_k = \bar{\mathbf{h}}_k + \mathbf{\Theta}_k^{\frac{1}{2}} \mathbf{z}_k, \ \ \forall\, k \in \mathbb{K}, \tag{3}$$

where $\bar{\mathbf{h}}_k \in \mathbb{C}^{N\times 1}$ and $\mathbf{\Theta}_k \in \mathbb{C}^{N\times N}$ are the near-field channel mean and covariance, respectively, with $\operatorname{rank}(\mathbf{\Theta}_k) = r_k$; $\mathbf{z}_k \in \mathbb{C}^{r_k\times 1}$ is a random vector with $\mathbb{E}[\mathbf{z}_k] = \mathbf{0}$ and $\operatorname{cov}(\mathbf{z}_k) = \mathbf{I}_{r_k}$. Since the NF-JSDM depends on the prior knowledge of statistical CSI, it is essential to first develop the near-field channel mean and spatial covariance models. Consequently, to model the near-field statistical CSI, we first develop the near-field GBSM with mixed LoS and NLoS links in the following.

2.1 Near-field system model

Figure 1 depicts the scenario of multi-user XL-array communications in the near-field scattering environment. The basic XL uniform linear array (XL-ULA) is considered, with the separation between adjacent elements represented by δ. Without loss of generality, the XL-ULA is assumed to lie along the y-axis of the Cartesian coordinate system with its center aligned with the origin. Thus, the distance between the nth element and the origin is $n\delta$, where $n \in \mathbb{N}$ with $\mathbb{N} = \{n : -\lceil \frac{N-1}{2} \rceil \leqslant n \leqslant \lfloor \frac{N-1}{2} \rfloor\}$ denoting the set of XL-array elements. The number of scatterers is denoted as Q. By considering the mixed LoS and NLoS links, the near-field GBSM of UE k is modeled as

$$\mathbf{h}_k = \mathbf{h}_k^{\mathrm{LoS}} + \mathbf{h}_k^{\mathrm{NLoS}}, \tag{4}$$

where $\mathbf{h}_k^{\mathrm{LoS}} \in \mathbb{C}^{N\times 1}$ and $\mathbf{h}_k^{\mathrm{NLoS}} \in \mathbb{C}^{N\times 1}$ are the near-field LoS and NLoS components, respectively.

In our prior work [22], the near-field GBSM $\mathbf{h}_k$ in (4) considered the generic NUSW model and partial array visibility. Within the Rayleigh distance $r_{\mathrm{Rayl}} = \frac{2D^2}{\lambda}$ [30, 31], with D denoting the XL-array aperture and λ being the carrier wavelength, the NUSW model should be considered to accurately capture near-field propagation effects. In the NUSW model, the exact distances of scatterers and/or UEs to the XL-array are used to accurately model the signal amplitude and phase across array elements, rather than their first-order Taylor approximation as in the UPW model. For the partial array visibility, the random binary variables $\boldsymbol{\xi}(e) \in \{0,1\}^{N\times 1}$ are introduced to model the (in)visibility of array elements to a scatterer or UEs, where the symbol 'e' can be replaced by either 'q' or 'k'. Specifically, 'e' = 'q' indicates that the array's (in)visibility to scatterer q for the NLoS links is modeled, whereas 'e' = 'k' indicates that the array's (in)visibility to UE k for the LoS link is modeled. In [22], the partial array visibility was modeled as a two-state Markov process, where the lengths of invisibility and visibility regions over the array axis before transitioning to the other state are modeled as i.i.d. exponential random variables with transition rates $\alpha^{\mathrm{i}}(e)$ (m^{-1}) and $\alpha^{\mathrm{v}}(e)$ (m^{-1}) related to the invisibility state and visibility state, respectively.

While the model in [22] laid a solid foundation for capturing near-field channel characteristics, it was limited to single-user scenarios and did not consider spatial consistency across different user locations. Building upon this foundation, the current work extends the near-field channel model in [22] to multi-user XL-MIMO communication scenarios by further considering the spatial consistency to observe the smooth channel variations as users move. Specifically, the current work introduces the concept of a reference user and incorporates the VR at the user side to model the spatial consistency of the channel. Therefore, the current model can be viewed as a natural and comprehensive extension of the channel model in [22], where we not only adapt the original single-user framework but also expand it to address multi-user scenarios. In particular, the VR at the user end determines the set of scatterers that are visible to any UE k [32], denoted by $\mathbb{Q}_k$, $\forall k \in \mathbb{K}$, which is modeled in the time-space domain to ensure smooth evolution of the channel response as the user moves. Specifically, each VR is defined as an identically sized circular geographical region on the azimuth plane, which determines the visibility of only one scatterer. As shown in Figure 1, if UE k is within a VR, the associated scatterer will be visible to that UE. Meanwhile, if UE k is located within the overlapping area of multiple VRs, it will simultaneously see multiple associated scatterers. Furthermore, the visibility gain is introduced to smoothly control the visibility level of a scatterer as UE k approaches the corresponding VR center [32]. This modeling approach has been validated through extensive indoor and outdoor measurement campaigns [33, 34]. Specifically, Ref. [33] reported 300 MHz outdoor measurements in which the KPowerMeans algorithm and a Kalman filter were used to identify and track scatterer clusters, confirming

the spatial consistency of the model. Additionally, indoor measurements in [34] confirmed that VR radii follow a lognormal distribution, further supporting the modeling validity of the user-side VR concept. The LoS and NLoS components of the near-field GBSM are modeled as follows.

2.2 LoS component of near-field GBSM

The LoS channel coefficient between UE k and the nth array element is given by the following expression, which considers the NUSW and partial array visibility [22]

$$h_{k,n}^{\mathrm{LoS}} = \sqrt{\beta_k^{\mathrm{LoS}}}\xi_n(k)\frac{d_k}{d_{k,n}}\mathrm{e}^{-\mathrm{j}\frac{2\pi}{\lambda}d_{k,n}}, \tag{5}$$

where β_k^{LoS} is the LoS path power of UE k at the reference element $n=0$, defined as $\sqrt{\beta_k^{\mathrm{LoS}}} = \frac{\lambda}{4\pi d_k}$ [22]; $\xi_n(k)$ denotes the (in)visibility of the nth array element to UE k for the LoS link; and $d_{k,n}$ is the distance from UE k to the nth array element, which is expressed as

$$d_{k,n} = \sqrt{d_k^2 + (n\delta)^2 - 2n\delta d_k \sin\varphi_k}, \tag{6}$$

with $d_k = d_{k,0}$ being the distance of UE k to the origin, and $\varphi_k \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ being the angle of UE k with respect to (w.r.t.) the positive x-axis, as shown in Figure 1.

To observe the smooth change of LoS component power across different user positions, i.e., spatial consistency, we introduce the reference UE $k=0$. Thus, $\sqrt{\beta_k^{\mathrm{LoS}}}$ is equivalently expressed as $\sqrt{\beta^{\mathrm{LoS}}}\frac{d}{d_k}$, where $\beta^{\mathrm{LoS}} = \beta_{k=0}^{\mathrm{LoS}} = (\frac{\lambda}{4\pi d})^2$ is the LoS path power of the reference UE with $d = d_{k=0}$. As such, $h_{k,n}^{\mathrm{LoS}}$ in (5) can be equivalently expressed as

$$h_{k,n}^{\mathrm{LoS}} = \sqrt{\beta^{\mathrm{LoS}}}\xi_n(k)\frac{d}{d_{k,n}}\mathrm{e}^{-\mathrm{j}\frac{2\pi}{\lambda}d_{k,n}}. \tag{7}$$

Thus, the LoS component vector $\mathbf{h}_k^{\mathrm{LoS}} \in \mathbb{C}^{N\times 1}$ containing all the N coefficients $h_{k,n}^{\mathrm{LoS}}$ in (7), is given by

$$\mathbf{h}_k^{\mathrm{LoS}} = \sqrt{\beta^{\mathrm{LoS}}}\boldsymbol{\xi}(k) \circ \mathbf{a}_k^{\mathrm{LoS}}, \tag{8}$$

where $\mathbf{a}_k^{\mathrm{LoS}} = \frac{d}{\mathbf{d}_k} \circ \mathrm{e}^{-\mathrm{j}\frac{2\pi}{\lambda}\mathbf{d}_k} \in \mathbb{C}^{N\times 1}$ is the near-field array response vector of UE k with $\mathbf{d}_k = [d_{k,n}]_{n\in\mathbb{N}} \in \mathbb{R}^{N\times 1}$. Compared with the previous work in [22], the LoS component in (8) further enhances multi-user channel modeling. Specifically, since the LoS path power β_k^{LoS} depends on the specific position of user k, the previous model in [22] is unable to characterize the spatial evolution of LoS power across user positions. To overcome this limitation, we further introduce the reference user, which enables the observation and modeling of the LoS power variation with respect to user position. This enhancement improves the channel model's ability to capture spatial consistency in multi-user communication scenarios.

2.3 NLoS component of near-field GBSM

Based on the prior work [22], the NLoS component further considers the VR at the UE end. Thus, the NLoS coefficient between UE k and the nth array element, denoted as $h_{k,n}^{\mathrm{NLoS}}$, can be modeled as

$$h_{k,n}^{\mathrm{NLoS}} = \sqrt{\frac{\beta_k^{\mathrm{NLoS}}}{Q_k}}\sum_{q\in\mathbb{Q}_k} g_{q,k}\xi_n(q)\frac{r_q}{r_{q,n}}\mathrm{e}^{-\mathrm{j}\frac{2\pi}{\lambda}(t_{q,k}+r_{q,n})+\mathrm{j}\omega_q}g_{q,k}^{\mathrm{VR}}, \tag{9}$$

where β_k^{NLoS} is the sum of the NLoS paths' power at the reference array element for UE k; $Q_k = |\mathbb{Q}_k|$ is the number of scatterers visible to UE k; $g_{q,k}$ is a random variable satisfying $\frac{1}{Q_k}\sum_{q\in\mathbb{Q}_k}\mathbb{E}[g_{q,k}^2] = 1$ [22], which corresponds to the signal amplitude at the reference element for UE k, i.e., contributed by scatterer q; $\xi_n(q)$ denotes the (in)visibility of the nth array element to scatterer q for the NLoS link; $r_{q,n}$ is the distance from scatterer q to array element n, which is expressed as

$$r_{q,n} = \sqrt{r_q^2 + (n\delta)^2 - 2n\delta r_q \sin\theta_q}, \tag{10}$$

with $r_q = r_{q,0}$ being the distance of scatterer q to the origin, and $\theta_q \in [-\pi, \pi]$ being the angle of scatterer q w.r.t. the positive x-axis, as shown in Figure 1; $t_{q,k}$ is the distance between UE k and scatterer q; and ω_q denotes the

phase shift due to scatterer q, which is modeled as an i.i.d. random variable with uniform distribution over $[-\pi,\pi)$; $g_{q,k}^{\mathrm{VR}}\in(0,1)$ denotes the visibility gain of scatterer q to UE k, modeled as [32]

$$g_{q,k}^{\mathrm{VR}}=0.5-\arctan\Big(2\sqrt{2}\big(L_{\mathrm{C}}+d_{q,k}^{\mathrm{VR}}-R_{\mathrm{C}}\big)/\sqrt{\lambda L_{\mathrm{C}}}\Big)/\pi, \tag{11}$$

with L_{C} being the width of the transition region, $d_{q,k}^{\mathrm{VR}}$ being the distance from UE k to the center of the VR corresponding to scatterer q, and R_{C} being the VR radius. Particularly, the variables β_k^{NLoS} and $g_{q,k}$ are expressed as $\beta_k^{\mathrm{NLoS}}=\sum_{q\in\mathbb{Q}_k}\frac{\lambda^2\sigma_q}{(4\pi)^3t_{q,k}^2r_q^2}$ and $\sqrt{\frac{\beta_k^{\mathrm{NLoS}}}{Q_k}}g_{q,k}=\frac{\lambda\sqrt{\sigma_q}}{(4\pi)^{3/2}t_{q,k}r_q}$ [22], respectively, where σ_q is the i.i.d. positive random radar cross section of scatterer q [35].

Similar to the LoS component, the reference UE $k=0$ is introduced to model spatial consistency, which is assumed to see all the scatterers, i.e., $|\mathbb{Q}_{k=0}|=Q$. Thus, the NLoS power and the variable $g_q=g_{q,0}$ for UE $k=0$ are expressed as $\beta^{\mathrm{NLoS}}=\sum_{q=1}^{Q}\frac{\lambda^2\sigma_q}{(4\pi)^3t_q^2r_q^2}$ and $\sqrt{\frac{\beta^{\mathrm{NLoS}}}{Q}}g_q=\frac{\lambda\sqrt{\sigma_q}}{(4\pi)^{3/2}t_qr_q}$, respectively, with $t_q=t_{q,0}$. Thus, we have $\sqrt{\frac{\beta_k^{\mathrm{NLoS}}}{Q_k}}g_{q,k}=\sqrt{\frac{\beta^{\mathrm{NLoS}}}{Q}}g_q\frac{t_q}{t_{q,k}}$. As such, $h_{k,n}^{\mathrm{NLoS}}$ in (9) is equivalently expressed as

$$h_{k,n}^{\mathrm{NLoS}}=\sqrt{\frac{\beta^{\mathrm{NLoS}}}{Q}}\sum_{q\in\mathbb{Q}_k}g_qg_{q,k}^{\mathrm{VR}}\xi_n(q)\frac{t_q}{t_{q,k}}\frac{r_q}{r_{q,n}}\mathrm{e}^{-\mathrm{j}\frac{2\pi}{\lambda}(t_{q,k}+r_{q,n})+\mathrm{j}\omega_q}. \tag{12}$$

The NLoS component vector $\mathbf{h}_k^{\mathrm{NLoS}}\in\mathbb{C}^{N\times1}$ containing N coefficients $h_{k,n}^{\mathrm{NLoS}}$ in (12), is written as

$$\mathbf{h}_k^{\mathrm{NLoS}}=\sqrt{\frac{\beta^{\mathrm{NLoS}}}{Q}}\sum_{q\in\mathbb{Q}_k}g_qg_{q,k}^{\mathrm{VR}}\frac{t_q}{t_{q,k}}\mathrm{e}^{-\mathrm{j}\frac{2\pi}{\lambda}t_{q,k}+\mathrm{j}\omega_q}\boldsymbol{\xi}(q)\circ\mathbf{a}_q^{\mathrm{NLoS}}, \tag{13}$$

where $\mathbf{a}_q^{\mathrm{NLoS}}=\frac{r_q}{\mathbf{r}_q}\circ\mathrm{e}^{-\mathrm{j}\frac{2\pi}{\lambda}\mathbf{r}_q}\in\mathbb{C}^{N\times1}$ is the near-field array response vector for scatterer q, with $\mathbf{r}_q=[r_{q,n}]_{n\in\mathbb{N}}\in\mathbb{R}^{N\times1}$ being the distance vector for scatterer q, which contains N distances $r_{q,n}$ in (10).

3 Near-field channel mean and covariance

Based on the developed near-field GBSM in (4), we characterize the near-field channel mean vector and covariance matrix, i.e., $\bar{\mathbf{h}}_k=\mathbb{E}[\mathbf{h}_k]$ and $\boldsymbol{\Theta}_k=\mathbb{E}\big[\mathbf{h}_k(\mathbf{h}_k)^{\mathrm{H}}\big]-\bar{\mathbf{h}}_k\bar{\mathbf{h}}_k^{\mathrm{H}}$.

3.1 Near-field channel mean vector

Since the variable ω_q follows an i.i.d. uniform distribution over $[-\pi,\pi)$, we have $\mathbb{E}\left[\mathrm{e}^{\mathrm{j}\omega_q}\right]=0$. In this case, the expectation of the NLoS component for UE k in (13) is $\mathbb{E}\left[\mathbf{h}_k^{\mathrm{NLoS}}\right]=\mathbf{0}$. Therefore, the near-field channel mean $\bar{\mathbf{h}}_k$ is expressed as

$$\bar{\mathbf{h}}_k=\mathbb{E}\left[\mathbf{h}_k^{\mathrm{LoS}}\right]=\sqrt{\beta}\sqrt{\frac{\mathcal{K}}{\mathcal{K}+1}}\mathbb{E}\big[\boldsymbol{\xi}(k)\big]\circ\mathbf{a}_k^{\mathrm{LoS}}, \tag{14}$$

where $\beta=\beta^{\mathrm{LoS}}+\beta^{\mathrm{NLoS}}$ and $\mathcal{K}=\frac{\beta^{\mathrm{LoS}}}{\beta^{\mathrm{NLoS}}}$ are the received power and K-factor of the reference UE at the reference antenna, respectively. It is worth remarking that the conventional far-field channel mean is equal to its LoS component $\mathbf{h}_k^{\mathrm{UPW\_LoS}}$, which is expressed as [36,37]

$$\bar{\mathbf{h}}_k^{\mathrm{UPW}}=\mathbf{h}_k^{\mathrm{UPW\_LoS}}=\sqrt{\beta_k}\sqrt{\frac{\mathcal{K}_k}{\mathcal{K}_k+1}}\mathbf{a}_k^{\mathrm{UPW\_LoS}}, \tag{15}$$

where β_k and $\mathcal{K}_k$ are the received power and K-factor of UE k at the reference antenna, respectively, and $\mathbf{a}_k^{\mathrm{UPW\_LoS}}=[\mathrm{e}^{-\mathrm{j}\frac{2\pi}{\lambda}(d_k-n\delta\sin\varphi_k)}]_{\forall n\in\mathbb{N}}$ is the far-field array response vector of UE k. By comparing the near- and far-field channel mean, it is found that due to the randomness of partial array visibility, the near-field channel mean is no longer equal to its LoS component as in the far-field model, i.e., $\bar{\mathbf{h}}_k\neq\mathbf{h}_k^{\mathrm{LoS}}$.

By comparing (14) and (15), we observe that the near-field channel mean model can be viewed as a generalization of the conventional far-field model, as described below.

Lemma 1. When $\alpha^{\mathrm{v}}(k)=0$ and $N\delta \ll d_k$, indicating that UE k can observe all elements of the XL-array and that UE k is in the far-field region of the XL-array, we have

$$\bar{\mathbf{h}}_k \approx \bar{\mathbf{h}}_k^{\mathrm{UPW}}. \tag{16}$$

Proof. Please refer to Appendix A.

Lemma 1 demonstrates that for the first-order statistical CSI, the near-field model in (14) generalizes the conventional far-field model in (15), as the former includes the latter as a special case.

3.2 Near-field channel covariance matrix

Since $\mathbb{E}\left[\mathbf{h}_k^{\mathrm{NLoS}}\right]=\mathbf{0}$, the near-field NUSW-based covariance matrix $\boldsymbol{\Theta}_k$ is simplified to

$$\boldsymbol{\Theta}_k = \boldsymbol{\Theta}_k^{\mathrm{LoS}} + \boldsymbol{\Theta}_k^{\mathrm{NLoS}} - \bar{\mathbf{h}}_k\bar{\mathbf{h}}_k^{\mathrm{H}}, \tag{17}$$

where $\boldsymbol{\Theta}_k^{\mathrm{LoS}}=\mathbb{E}\left[\mathbf{h}_k^{\mathrm{LoS}}(\mathbf{h}_k^{\mathrm{LoS}})^{\mathrm{H}}\right]\in\mathbb{C}^{N\times N}$ and $\boldsymbol{\Theta}_k^{\mathrm{NLoS}}=\mathbb{E}\left[\mathbf{h}_k^{\mathrm{NLoS}}(\mathbf{h}_k^{\mathrm{NLoS}})^{\mathrm{H}}\right]\in\mathbb{C}^{N\times N}$ denote the near-field S-CF matrices of LoS and NLoS components, respectively. It is worth noting that the far-field covariance matrix $\boldsymbol{\Theta}_k^{\mathrm{UPW}}$ is equal to the far-field NLoS S-CF $\boldsymbol{\Theta}_k^{\mathrm{UPW\_NLoS}}$, i.e., $\boldsymbol{\Theta}_k^{\mathrm{UPW}}=\boldsymbol{\Theta}_k^{\mathrm{UPW\_NLoS}}$. Next, we characterize the near-field S-CF matrices $\boldsymbol{\Theta}_k^{\mathrm{LoS}}$ and $\boldsymbol{\Theta}_k^{\mathrm{NLoS}}$ separately.

3.2.1 *NUSW-based S-CF of LoS component*

According to the LoS component $\mathbf{h}_k^{\mathrm{LoS}}$ in (8), the near-field LoS S-CF is given by

$$\boldsymbol{\Theta}_k^{\mathrm{LoS}} = \beta\frac{\mathcal{K}}{\mathcal{K}+1}\mathbb{E}\left[\boldsymbol{\xi}(k)\boldsymbol{\xi}^{\mathrm{T}}(k)\right]\circ\left(\mathbf{a}_k^{\mathrm{LoS}}(\mathbf{a}_k^{\mathrm{LoS}})^{\mathrm{H}}\right). \tag{18}$$

Specifically, the element of the matrix $\boldsymbol{\Theta}_k^{\mathrm{LoS}}$ in (18) is

$$\left[\boldsymbol{\Theta}_k^{\mathrm{LoS}}\right]_{n,m} = \beta\frac{\mathcal{K}}{\mathcal{K}+1}\frac{d^2}{d_{k,n}d_{k,m}}\mathrm{e}^{-\mathrm{j}\frac{2\pi}{\lambda}(d_{k,n}-d_{k,m})}\mathbb{E}\left[\xi_n(k)\xi_m(k)\right], \tag{19}$$

where $\mathbb{E}[\xi_n(k)\xi_m(k)]$ is the autocorrelation of the two-state discrete Markov process, $n>m$, which is given by [22]

$$\mathbb{E}\left[\xi_n(k)\xi_m(k)\right] = \left(\frac{\alpha^{\mathrm{i}}(k)}{\alpha(k)}+\frac{\alpha^{\mathrm{v}}(k)}{\alpha(k)}\mathrm{e}^{-\alpha(k)(n-m)\delta}\right)\mathbb{E}\left[\xi_m(k)\right], \tag{20}$$

with $\alpha(k)=\alpha^{\mathrm{i}}(k)+\alpha^{\mathrm{v}}(k)$, $\mathbb{E}[\xi_m(k)]$ being the probability of the mth array element visible to UE k, expressed as $\mathbb{E}[\xi_m(k)]=\frac{\alpha^{\mathrm{i}}(k)}{\alpha(k)}+\frac{\alpha^{\mathrm{v}}(k)}{\alpha(k)}\mathrm{e}^{-\alpha(k)|m-n'(k)|\delta}$ [22]; the parameter $\alpha^{\mathrm{i}}(k)$ denoting the transition rate in the Markov process from the invisibility state to the visibility state, while the parameter $\alpha^{\mathrm{v}}(k)$ denotes the transition rate from the visibility state to the invisibility state; and $n'(k)$ being the center of visible array elements to UE k satisfying $\mathbb{E}[\xi_{n'(k)}(k)]=1$, which is uniformly selected from the array elements. If the reference UE is specified as UE k itself, the near-field LoS S-CF in (19) is rewritten as [22]

$$\left[\boldsymbol{\Theta}_k^{\mathrm{LoS}}\right]_{n,m} = \beta_k\frac{\mathcal{K}_k}{\mathcal{K}_k+1}\frac{d_k^2}{d_{k,n}d_{k,m}}\mathrm{e}^{-\mathrm{j}\frac{2\pi}{\lambda}(d_{k,n}-d_{k,m})}\mathbb{E}\left[\xi_n(k)\xi_m(k)\right]. \tag{21}$$

It is worth mentioning that the far-field LoS S-CF is expressed as $\boldsymbol{\Theta}_k^{\mathrm{UPW\_LoS}}=\bar{\mathbf{h}}_k^{\mathrm{UPW}}\left(\bar{\mathbf{h}}_k^{\mathrm{UPW}}\right)^{\mathrm{H}}$ due to $\mathbf{h}_k^{\mathrm{UPW\_LoS}}=\bar{\mathbf{h}}_k^{\mathrm{UPW}}$. Based on $\bar{\mathbf{h}}_k^{\mathrm{UPW}}$ in (15), the element of matrix $\boldsymbol{\Theta}_k^{\mathrm{UPW\_LoS}}$ is expressed as

$$\left[\boldsymbol{\Theta}_k^{\mathrm{UPW\_LoS}}\right]_{n,m} = \beta_k\frac{\mathcal{K}_k}{\mathcal{K}_k+1}\mathrm{e}^{-\mathrm{j}\frac{2\pi}{\lambda}(m-n)\delta\sin\varphi_k}. \tag{22}$$

By comparing the NUSW- and UPW-based S-CFs of the LoS component in (19) and (22), there are three main differences between them. Firstly, the NUSW-based model no longer exhibits SWSS as the UPW-based model, since $\left[\boldsymbol{\Theta}_k^{\mathrm{LoS}}\right]_{n,m}$ is dependent on the specific locations of the nth and mth array elements rather than their relative distance only. Secondly, the NUSW-based model may exhibit sparsity due to the partial array visibility, compared to the UPW-based model. Thirdly, the developed near-field model exhibits spatial consistency and depends on the UE's location, which differs from the far-field model that is independent of the UE's location.

Lemma 2. When $\alpha^{\mathrm{v}}(k)=0$ and $N\delta \ll d_k$, which means that UE k is guaranteed to see all elements of the XL-array and to be in the far-field region of the XL-array, we have

$$\boldsymbol{\Theta}_k^{\mathrm{LoS}} \approx \boldsymbol{\Theta}_k^{\mathrm{UPW\_LoS}}. \tag{23}$$

Proof. Please refer to Appendix B.

Lemma 2 shows that for the LoS component, the near-field S-CF in (21) generalizes the conventional far-field model in (22), as the former includes the latter as a special case.

3.2.2 *NUSW-based S-CF of NLoS component*

Based on (13), the near-field NUSW-based S-CF of the NLoS component is expressed as

$$\boldsymbol{\Theta}_k^{\mathrm{NLoS}} = \beta \frac{1}{\mathcal{K}+1}\frac{1}{Q}\sum_{q\in\mathbb{Q}_k}\sum_{p\in\mathbb{Q}_k}\mathbb{E}\left[g_q g_p\right] g_{q,k}^{\mathrm{VR}} g_{p,k}^{\mathrm{VR}} \frac{t_q t_p}{t_{q,k}t_{p,k}} \mathrm{e}^{-\mathrm{j}\frac{2\pi}{\lambda}(t_{q,k}-t_{p,k})}\mathbb{E}[\mathrm{e}^{\mathrm{j}\omega_q-\mathrm{j}\omega_p}]\mathbb{E}\left[\boldsymbol{\xi}(q)\boldsymbol{\xi}^{\mathrm{T}}(p)\right]\circ(\mathbf{a}_q^{\mathrm{NLoS}}(\mathbf{a}_p^{\mathrm{NLoS}})^{\mathrm{H}}). \tag{24}$$

With ω_q following i.i.d. uniform distribution over $[-\pi,\pi)$, we have $\mathbb{E}[\mathrm{e}^{\mathrm{j}\omega_q-\mathrm{j}\omega_p}]=\begin{cases}1, & p=q\\ 0, & p\neq q\end{cases}$. Hence, the NUSW-based S-CF of NLoS component in (24) is simplified to

$$\boldsymbol{\Theta}_k^{\mathrm{NLoS}} = \frac{\beta}{\mathcal{K}+1}\frac{1}{Q}\sum_{q\in\mathbb{Q}_k}\mathbb{E}[g_q^2](g_{q,k}^{\mathrm{VR}})^2\frac{t_q^2}{t_{q,k}^2}\mathbb{E}\left[\boldsymbol{\xi}(q)\boldsymbol{\xi}^{\mathrm{T}}(q)\right]\circ\left(\mathbf{a}_q^{\mathrm{NLoS}}(\mathbf{a}_q^{\mathrm{NLoS}})^{\mathrm{H}}\right). \tag{25}$$

Thus, the element of the matrix $\boldsymbol{\Theta}_k^{\mathrm{NLoS}}$ in (25) is given by

$$\left[\boldsymbol{\Theta}_k^{\mathrm{NLoS}}\right]_{n,m} = \beta\frac{1}{\mathcal{K}+1}\frac{1}{Q}\sum_{q\in\mathbb{Q}_k}\mathbb{E}[g_q^2](g_{q,k}^{\mathrm{VR}})^2\frac{t_q^2}{t_{q,k}^2}\frac{r_q^2}{r_{q,n}r_{q,m}}\mathrm{e}^{-\mathrm{j}\frac{2\pi}{\lambda}(r_{q,n}-r_{q,m})}\mathbb{E}[\xi_n(q)\xi_m(q)], \tag{26}$$

where the autocorrelation $\mathbb{E}\left[\xi_n(q)\xi_m(q)\right]$ has the similar expression as in (20), with the center of visible array elements to scatterer q denoted by $n'(q)$, which is chosen uniformly from the array elements.

Similar to the integral-form near-field S-CF in [38], for $Q\to\infty$, $Q^{-1}\mathbb{E}[g_q^2]$ in (26) is interpreted as the infinitesimal power at the reference array element, which is contributed by a differential scatterer located around $\mathbf{s}$. Thus, we have $Q^{-1}\mathbb{E}[g_q^2]=f(\mathbf{s})d\mathbf{s}$, where $f(\mathbf{s})$ represents the probability density function (PDF) of scatterer location $\mathbf{s}\in\mathbf{S}$ for the reference UE, with $\mathbf{S}$ being the random scatterers' support of the reference UE. Hence, for $Q\to\infty$, $\left[\boldsymbol{\Theta}_k^{\mathrm{NLoS}}\right]_{n,m}$ in (26) can be expressed in an integral form as

$$\left[\boldsymbol{\Theta}_k^{\mathrm{NLoS}}\right]_{n,m} = \beta\frac{1}{\mathcal{K}+1}\int_{\mathbf{s}\in\mathbf{S}_k}(g_k^{\mathrm{VR}}(\mathbf{s}))^2\frac{t^2(\mathbf{s})}{t_k^2(\mathbf{s})}\frac{r^2(\mathbf{s})}{r_n(\mathbf{s})r_m(\mathbf{s})}\mathrm{e}^{-\mathrm{j}\frac{2\pi}{\lambda}(r_n(\mathbf{s})-r_m(\mathbf{s}))}\mathbb{E}\left[\xi_n(\mathbf{s})\xi_m(\mathbf{s})\right]f(\mathbf{s})\mathrm{d}\mathbf{s}, \tag{27}$$

where $\mathbf{S}_k$ is the support of random scatterers of UE k; $g_k^{\mathrm{VR}}(\mathbf{s})$ is the visibility gain of scatterer $\mathbf{s}$ for UE k; $t_k(\mathbf{s})$ is the distance between scatterer $\mathbf{s}$ and UE k, with $t(\mathbf{s})=t_0(\mathbf{s})$; $r_n(\mathbf{s})$ is the distance between scatterer $\mathbf{s}$ and the nth array element, with $r(\mathbf{s})=r_0(\mathbf{s})$; $\xi_n(\mathbf{s})$ is the (in)visibility of the nth array element to scatterer $\mathbf{s}$. In particular, if the reference UE is set to UE k itself, $\left[\boldsymbol{\Theta}_k^{\mathrm{NLoS}}\right]_{n,m}$ in (27) is rewritten as

$$\left[\boldsymbol{\Theta}_k^{\mathrm{NLoS}}\right]_{n,m} = \beta_k\frac{1}{\mathcal{K}_k+1}\int_{\mathbf{s}\in\mathbf{S}_k}(g_k^{\mathrm{VR}}(\mathbf{s}))^2\frac{r^2(\mathbf{s})}{r_n(\mathbf{s})r_m(\mathbf{s})}\mathrm{e}^{-\mathrm{j}\frac{2\pi}{\lambda}(r_n(\mathbf{s})-r_m(\mathbf{s}))}\mathbb{E}\left[\xi_n(\mathbf{s})\xi_m(\mathbf{s})\right]f_k(\mathbf{s})\mathrm{d}\mathbf{s}, \tag{28}$$

where $f_k(\mathbf{s})$ is the PDF for UE k with $f_k(\mathbf{s})d\mathbf{s}=Q_k^{-1}\mathbb{E}[g_{q,k}^2]$, $\forall\mathbf{s}\in\mathbf{S}_k$. As the integration in (28) needs to be performed for all (n,m) antenna element pairs, the computational complexity of constructing the matrix $\boldsymbol{\Theta}_k^{\mathrm{NLoS}}$ scales as $O(N^2S_k)$, where $S_k=|\mathbf{S}_k|$ denotes the number of discrete scatterers for user k used in the numerical integration.

It is worth noting that the conventional far-field UPW-based S-CF of the NLoS component is given by [39]

$$\left[\boldsymbol{\Theta}_k^{\mathrm{UPW\_NLoS}}\right]_{n,m} = \frac{\beta_k}{\mathcal{K}_k+1}\int_{\theta_k^{\min}}^{\theta_k^{\max}}\mathrm{e}^{-\mathrm{j}\frac{2\pi}{\lambda}(m-n)\delta\sin\theta}f_k(\theta)\mathrm{d}\theta, \tag{29}$$

where $\theta \in [\theta_k^{\min}, \theta_k^{\max}]$ is the angle of arrival (AoA) of scatterers for UE k, with $\theta_k^{\min}$ and $\theta_k^{\max}$ being the minimum and maximum AoAs of scatterers for UE k, respectively, and $f_k(\theta)$ is the PAS of UE k.

For the NLoS component, a comparison between the developed near-field S-CF in (27) and the conventional far-field model in (29) reveals three important differences. First, the near-field S-CF depends on the PLS characterized by AoA and distances of scatterers, whereas the far-field S-CF depends on the PAS characterized only by the scatterers' AoAs. Second, compared to the far-field model, the near-field model no longer exhibits SWSS, since $\left[\mathbf{\Theta}_k^{\mathrm{NLoS}}\right]_{n,m}$ in (28) depends on the specific array indices n and m, rather than the array index difference $m-n$. Lastly, the developed near-field model exhibits spatial consistency and depends on the UE's location, while the far-field model is independent of the UE's location.

Lemma 3. When $g_k^{\mathrm{VR}}(\mathbf{s}) = 1$, $\alpha^{\mathrm{v}}(\mathbf{s}) = 0$, $N\delta \ll r(\mathbf{s})$, $\forall \mathbf{s} \in \mathbf{S}_k$, which means that UE k is in the VR center of scatterer $\mathbf{s}$, the probability of array visible to scatterer $\mathbf{s}$ is 1, and scatterers are in the far-field region of the XL-array, we have

$$\mathbf{\Theta}_k^{\mathrm{NLoS}} \approx \mathbf{\Theta}_k^{\mathrm{UPW\_NLoS}}. \tag{30}$$

Proof. Please refer to Appendix C.

From Lemma 3, it is found that the developed near-field NLoS S-CF in (28) is more general than the far-field one in (29), since the former includes the latter as a special case.

According to Lemmas 1–3, we observe that the near-field channel covariance model is a generalization of the conventional far-field model, as shown below.

Remark 1. When $g_k^{\mathrm{VR}}(\mathbf{s}) = 1$, $\alpha^{\mathrm{v}}(\mathbf{s}) = 0$, and $N\delta \ll r(\mathbf{s})$, $\forall \mathbf{s} \in \mathbf{S}_k$, and $\alpha^{\mathrm{v}}(k) = 0$ and $N\delta \ll d_k$, which means that UE k is in the VR center of scatterer $\mathbf{s}$, the probability of the array visible to both scatterer $\mathbf{s}$ and UE k is 1, and scatterers and UE k are in the far-field region of the XL-array, we have

$$\mathbf{\Theta}_k \approx \mathbf{\Theta}_k^{\mathrm{UPW}}. \tag{31}$$

Remark 1 demonstrates that the near-field second-order statistical CSI model in (17) is more general than the conventional far-field model, as it is applicable to mixed far- and near-field communications.

4 Near-field joint spatial-division and multiplexing

The developed near-field first- and second-order statistical CSI in Section 3 serves as crucial prior information in the NF-JSDM approach. Before designing the NF-JSDM approach, we will conduct an analysis to assess the validity of *Identical Statistical CSI Assumption* in XL-MIMO communications. This analysis focuses on comparing the similarity of statistical CSI's eigenspaces among different UEs, ultimately revealing that *Identical Statistical CSI Assumption* does not hold in XL-MIMO communications.

4.1 The similarity of channel eigenspace

In the communications scenarios with mixed LoS and NLoS links, the channel eigenspace cannot be simply equal to the eigenspace derived from the eigenvalue decomposition (EVD) of the channel covariance matrix, as in the FF-JSDM method. This distinction arises from the fact that in purely NLoS communication scenarios, statistical CSI typically involves only the second-order statistics of the channel, i.e., the covariance matrix, whose EVD directly yields the channel eigenspace. However, in wireless communication scenarios with mixed LoS and NLoS links, both the first-order and second-order statistics must be considered. As a result, the channel eigenspace in such scenarios is determined not only by the covariance matrix but also by the channel mean vector. Therefore, before analyzing channel eigenspace similarity, it is necessary to discuss the effective eigenspace formed by the joint first- and second-order statistical CSI.

First, the eigenspace of the second-order statistical CSI for UE k, denoted by $\left\{\mathbf{U}_k, \mathbf{\Lambda}_k\right\}$, is obtained through the EVD of the spatial covariance matrix $\mathbf{\Theta}_k$ directly, i.e., $\mathbf{\Theta}_k = \mathbf{U}_k\mathbf{\Lambda}_k\mathbf{U}_k^{\mathrm{H}}$ with $\mathbf{U}_k \in \mathbb{C}^{N\times r_k}$ and $\mathbf{\Lambda}_k \in \mathbb{R}^{r_k\times r_k}$. Then, for the first-order statistical CSI, the eigenspace can be regarded as the orthogonal component of the channel mean vector $\bar{\mathbf{h}}_k$ in the normal direction of the subspace spanned by the eigenspace $\mathbf{\Theta}_k^{\frac{1}{2}} = \mathbf{U}_k\mathbf{\Lambda}_k^{\frac{1}{2}}$, denoted by $\bar{\mathbf{h}}_k^{\perp} \in \mathbb{C}^{N\times 1}$, and is given by

$$\bar{\mathbf{h}}_k^{\perp} = \bar{\mathbf{h}}_k - \bar{\mathbf{h}}_k^{\parallel}, \tag{32}$$

where $\bar{\mathbf{h}}_k^{\|} \in \mathbb{C}^{N\times 1}$ denotes the projection component of $\bar{\mathbf{h}}_k$ onto the subspace $\mathbf{\Theta}_k^{\frac{1}{2}}$, which is expressed as

$$\bar{\mathbf{h}}_k^{\|} = \mathbf{\Theta}_k^{\frac{1}{2}} \underbrace{\left((\mathbf{\Theta}_k^{\frac{1}{2}})^{\mathrm{H}} \mathbf{\Theta}_k^{\frac{1}{2}}\right)^{-1} \left(\mathbf{\Theta}_k^{\frac{1}{2}}\right)^{\mathrm{H}} \bar{\mathbf{h}}_k}_{\mathbf{x}_k}, \tag{33}$$

where $\mathbf{x}_k$ is the linear operator between $\mathbf{\Theta}_k^{\frac{1}{2}}$ and $\bar{\mathbf{h}}_k^{\|}$. Thus, $\mathbf{h}_k$ in (3) can be equivalently expressed as

$$\mathbf{h}_k = \bar{\mathbf{h}}_k^{\perp} + \mathbf{\Theta}_k^{\frac{1}{2}}(\mathbf{x}_k + \mathbf{z}_k) = (\mathbf{\Theta}_k^{\mathrm{eff}})^{\frac{1}{2}} \mathbf{z}_k^{\mathrm{eff}}, \tag{34}$$

where $\mathbf{\Theta}_k^{\mathrm{eff}} = \mathbf{U}_k^{\mathrm{eff}} \mathbf{\Lambda}_k^{\mathrm{eff}} (\mathbf{U}_k^{\mathrm{eff}})^{\mathrm{H}} \in \mathbb{C}^{N\times N}$ is the effective eigenmatrix of UE k, with $\mathbf{U}_k^{\mathrm{eff}} \in \mathbb{C}^{N\times r_k^{\mathrm{eff}}}$ and $\mathbf{\Lambda}_k^{\mathrm{eff}} \in \mathbb{R}^{r_k^{\mathrm{eff}}\times r_k^{\mathrm{eff}}}$ representing the effective unitary eigenvector matrix and effective diagonal eigenvalue matrix, respectively, and $r_k^{\mathrm{eff}} = \mathrm{rank}(\mathbf{\Theta}_k^{\mathrm{eff}})$; and $\mathbf{z}_k^{\mathrm{eff}} \in \mathbb{C}^{r_k^{\mathrm{eff}}\times 1}$ is the effective random vector of UE k. The expressions of the effective eigenspace, denoted by $\{\mathbf{U}_k^{\mathrm{eff}}, \mathbf{\Lambda}_k^{\mathrm{eff}}\}$, and the effective random vector $\mathbf{z}_k^{\mathrm{eff}}$ need to be discussed separately when $\bar{\mathbf{h}}_k^{\perp} = \mathbf{0}$ or $\bar{\mathbf{h}}_k^{\perp} \neq \mathbf{0}$. Specifically, if $\bar{\mathbf{h}}_k^{\perp} = \mathbf{0}$, the effective eigenspace and the effective random vector are respectively expressed as

$$\{\mathbf{U}_k^{\mathrm{eff}}, \mathbf{\Lambda}_k^{\mathrm{eff}}\} = \{\mathbf{U}_k, \mathbf{\Lambda}_k\} \text{ and } \mathbf{z}_k^{\mathrm{eff}} = \mathbf{x}_k + \mathbf{z}_k,$$

with $r_k^{\mathrm{eff}} = r_k$, $\mathbb{E}\left[\mathbf{z}_k^{\mathrm{eff}}\right] = \mathbf{x}_k$, and $\mathrm{cov}\left(\mathbf{z}_k^{\mathrm{eff}}\right) = \mathbf{I}_{r_k}$; otherwise, $\{\mathbf{U}_k^{\mathrm{eff}}, \mathbf{\Lambda}_k^{\mathrm{eff}}\}$ and $\mathbf{z}_k^{\mathrm{eff}}$ are respectively expressed as

$$\mathbf{U}_k^{\mathrm{eff}} = \left[\bar{\mathbf{h}}_k^{\perp}/\|\bar{\mathbf{h}}_k^{\perp}\|_F, \mathbf{U}_k\right], \mathbf{\Lambda}_k^{\mathrm{eff}} = \mathrm{diag}\big(\|\bar{\mathbf{h}}_k^{\perp}\|_F^2, \mathrm{diag}(\mathbf{\Lambda}_k)\big), \text{ and } \mathbf{z}_k^{\mathrm{eff}} = [1; \mathbf{x}_k + \mathbf{z}_k],$$

with $r_k^{\mathrm{eff}} = r_k + 1$, $\mathbb{E}\left[\mathbf{z}_k^{\mathrm{eff}}\right] = [1; \mathbf{x}_k]$, and $\mathrm{cov}\left(\mathbf{z}_k^{\mathrm{eff}}\right) = \mathrm{diag}\left(0, \mathbf{1}_{r_k}\right)$. To analyze the similarity between effective eigenspaces of UEs k and k', we introduce the variable $\Upsilon\big((\mathbf{\Theta}_k^{\mathrm{eff}})^{\frac{1}{2}}, (\mathbf{\Theta}_{k'}^{\mathrm{eff}})^{\frac{1}{2}}\big) \in [0,1]$, defined as

$$\Upsilon\big((\mathbf{\Theta}_k^{\mathrm{eff}})^{\frac{1}{2}}, (\mathbf{\Theta}_{k'}^{\mathrm{eff}})^{\frac{1}{2}}\big) = \frac{\sum_{i=1}^{r_k^{\mathrm{eff}}} \sum_{j=1}^{r_{k'}^{\mathrm{eff}}} \left| \left(\left[(\mathbf{\Theta}_k^{\mathrm{eff}})^{\frac{1}{2}}\right]_{:,i} \right)^{\mathrm{H}} \left[(\mathbf{\Theta}_{k'}^{\mathrm{eff}})^{\frac{1}{2}}\right]_{:,j} \right|^2}{\left\|(\mathbf{\Theta}_k^{\mathrm{eff}})^{\frac{1}{2}}\right\|_F^2 \left\|(\mathbf{\Theta}_{k'}^{\mathrm{eff}})^{\frac{1}{2}}\right\|_F^2}, \tag{35}$$

where the case of $\Upsilon\big((\mathbf{\Theta}_k^{\mathrm{eff}})^{\frac{1}{2}}, (\mathbf{\Theta}_{k'}^{\mathrm{eff}})^{\frac{1}{2}}\big) = 1$ means that UEs k and k' have the same effective eigenspace. The similarity in (35) can be equivalently expressed as

$$\Upsilon\big((\mathbf{\Theta}_k^{\mathrm{eff}})^{\frac{1}{2}}, (\mathbf{\Theta}_{k'}^{\mathrm{eff}})^{\frac{1}{2}}\big) = \frac{\left\| \left((\mathbf{\Theta}_k^{\mathrm{eff}})^{\frac{1}{2}}\right)^{\mathrm{H}} \left(\mathbf{\Theta}_{k'}^{\mathrm{eff}}\right)^{\frac{1}{2}} \right\|_F^2}{\left\|\left(\mathbf{\Theta}_k^{\mathrm{eff}}\right)^{\frac{1}{2}}\right\|_F^2 \left\|\left(\mathbf{\Theta}_{k'}^{\mathrm{eff}}\right)^{\frac{1}{2}}\right\|_F^2}.$$

According to the Cauchy-Schwarz inequality, we have $\left\|\left((\mathbf{\Theta}_k^{\mathrm{eff}})^{\frac{1}{2}}\right)^{\mathrm{H}} \left(\mathbf{\Theta}_{k'}^{\mathrm{eff}}\right)^{\frac{1}{2}}\right\|_F^2 \leqslant \left\|(\mathbf{\Theta}_k^{\mathrm{eff}})^{\frac{1}{2}}\right\|_F^2 \left\|\left(\mathbf{\Theta}_{k'}^{\mathrm{eff}}\right)^{\frac{1}{2}}\right\|_F^2$. Therefore, it is evident that the similarity $\Upsilon\big((\mathbf{\Theta}_k^{\mathrm{eff}})^{\frac{1}{2}}, (\mathbf{\Theta}_{k'}^{\mathrm{eff}})^{\frac{1}{2}}\big)$ is bounded between 0 and 1. The overall computational complexity for $\Upsilon\big((\mathbf{\Theta}_k^{\mathrm{eff}})^{\frac{1}{2}}, (\mathbf{\Theta}_{k'}^{\mathrm{eff}})^{\frac{1}{2}}\big)$ in (35) is $O(Nr_{k'}^{\mathrm{eff}} r_k^{\mathrm{eff}} + Nr_{k'}^{\mathrm{eff}} + Nr_k^{\mathrm{eff}})$, which accounts for $r_{k'}^{\mathrm{eff}} r_k^{\mathrm{eff}}$ inner products between N-dimensional complex vectors, each requiring $O(N)$ operations, as well as the Frobenius norms of two matrices with dimensions $N \times r_{k'}^{\mathrm{eff}}$ and $N \times r_k^{\mathrm{eff}}$, respectively.

Next, we provide numerical results of the effective eigenspaces' similarity among different UEs, based on the developed near-field and conventional far-field statistical CSI models. The near-field channel mean and covariance are defined in (14) and (17), respectively, while the far-field ones are defined in (15) and (29), respectively. The carrier frequency and the number of XL-array elements are set to 3.5 GHz and 512, respectively. It can be calculated that the XL-array size equals $D = (N-1)\delta$ =21.9 m. According to the definition of the Rayleigh distance, the near-field region of the XL-array is determined to be within 11191 m. The received power of any UE k, as well as that of the reference UE is set to 1 for normalization, i.e., $\beta = \beta_k = 1$, $\forall k$. The K-factor of any UE k in the far-field model is set to be the same as that of the reference UE for simplicity, i.e., $\mathcal{K}_k = \mathcal{K}$, $\forall k$. In addition, for the partial array visibility characteristic, the transition rates $\alpha^{\mathrm{i}}(q)$ and $\alpha^{\mathrm{v}}(q)$ of any scatterer q are set to be identical for simplicity, while for UEs, the parameter $\alpha^{\mathrm{i}}(k)$ is set to be the same for all UEs and $\alpha^{\mathrm{v}}(k)$ is set inversely proportional to the distance d_k to ensure that UE k farther away from the XL-array is more likely to see the entire array, i.e., $\alpha^{\mathrm{v}}(k) \propto \frac{1}{d_k}$. The specific values of transition rates for scatterers and UEs are set according to the measurements [40]. Furthermore, each VR's center is assumed to coincide with its associated scatterer for simplicity, and the parameters of the VR model refer to the parameters of the COST 273 macrocell scenario [41].

Table 2 The simulation parameters.

Parameter	Value	Parameter	Value
Carrier frequency	3.5 GHz	Wavelength λ	0.086 m
XL-array elements N	512	Element spacing δ	$\frac{\lambda}{2}$ m
Reference UE location	$[65, 50]^{\mathrm{T}}$ m	VR radius R_{C}	100 m
UE k location region	$x \in [20, 220]$ m, $y \in [-100, 100]$ m	UE k sub-region	$x \in [55, 75]$ m, $y \in [40, 60]$ m
Number of scatterers	20	Transition region width L_{C}	20 m
Partial visibility of scatterer q	$\alpha^{\mathrm{v}}(q) = 0.9$ m^{-1}, $\alpha^{\mathrm{i}}(q) = 0.1$ m^{-1}, $n'(q)$: random	Partial visibility of UE k	$\alpha^{\mathrm{v}}(k) = \frac{18}{d_k}$ m^{-1}, $\alpha^{\mathrm{i}}(k) = 0.1$ m^{-1}, $n'(k) = -\lceil \frac{N-1}{2} \rceil$
Scatterer region	$x \in [10, 16]$ m, $y \in [5, 25]$ m	–	–

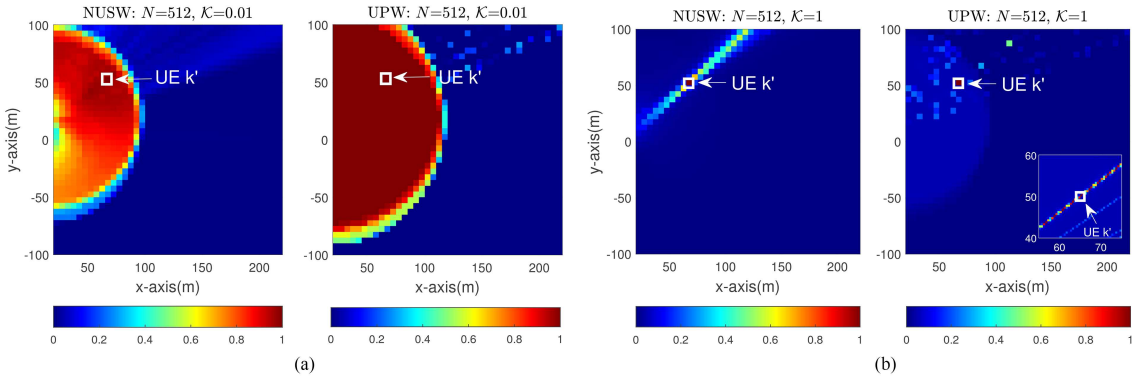

Figure 2 (Color online) $\Upsilon\left((\boldsymbol{\Theta}_k^{\mathrm{eff}})^{\frac{1}{2}}, (\boldsymbol{\Theta}_{k'}^{\mathrm{eff}})^{\frac{1}{2}}\right)$ based on the near- and far-field statistical CSI models for versus different values of the K-factor. (a) $\mathcal{K} = 0.01$; (b) $\mathcal{K} = 1$.

The PAS of any UE k for the far-field model and the PLS of the reference UE for the near-field model are assumed to be uniform. The detailed parameters are defined in Table 2.

Figure 2 presents the numerical results of $\Upsilon\big((\boldsymbol{\Theta}_k^{\mathrm{eff}})^{\frac{1}{2}}, (\boldsymbol{\Theta}_{k'}^{\mathrm{eff}})^{\frac{1}{2}}\big)$ in (35) based on near- and far-field models versus different values of the K-factor, where $\mathcal{K} = 0.01$ and $\mathcal{K} = 1$ are considered. As shown in Figure 2, the location of UE k' is set to be fixed and the same as that of the reference UE, while UE k moves within a square with a step size of 5 m along both the x- and y-axes, as detailed in Table 2. It is observed from Figure 2(a) that when adjacent UEs are so close to each other that they experience the same set of scatterers, there are significant differences in their effective eigenspaces for the near-field model. In contrast, for the far-field model, the effective eigenspaces of adjacent UEs are identical. This is expected since the developed near-field model takes into account the partial array visibility characteristic and non-uniform received power across array elements. In addition, Figure 2(b) demonstrates that for the near-field model, UEs located within the adjacent angle range of the fixed-position UE k's angle $\varphi_{k'}$ exhibit a high degree of similarity in their effective eigenspace. For the UPW-based model, by observing the sub-region specified in Table 2, which considers a more refined step size of 0.5 m to achieve higher angular resolution, it is found that only UEs in the direction of $\varphi_{k'}$ have high similarity in their effective eigenspaces, which is attributed to the high spatial resolution provided by the XL-array. These observations demonstrate that *Identical Statistical CSI Assumption* is no longer valid for XL-array communications. Consequently, compared to the FF-JSDM method, the NF-JSDM requires a more refined user grouping strategy.

To facilitate a clearer understanding of the proposed NF-JSDM approach, we briefly summarize the three key assumptions adopted in the NF-JSDM method, which are reflected in the near-field CBSM expression in (3), as follows.

- *Non-zero Mean Assumption.* The channel mean is no longer equal to zero, i.e., $\mathbb{E}[\mathbf{h}_k] = \bar{\mathbf{h}}_k \neq \mathbf{0}$.
- *NUSW Assumption.* The developed near-field first- and second-order statistical CSI model in (3) captures the near-field NUSW propagation characteristic, which differs from the far-field UPW assumption in the conventional FF-JSDM method.
- *Non-identical Statistical CSI Assumption.* Simulation results reveal that even adjacent UEs exhibit significantly different effective eigenspaces, resulting in heterogeneous effective eigenspaces among UEs, i.e., $\left(\boldsymbol{\Theta}_k^{\mathrm{eff}}\right) \neq \left(\boldsymbol{\Theta}_{k'}^{\mathrm{eff}}\right)$,

$\forall\, k, k' \in \mathbb{K}_g$.

4.2 User grouping for near-field XL-MIMO communications

The similarity analysis of channel eigenspaces in Subsection 4.1 reveals the necessity for an effective user grouping scheme in multi-user XL-MIMO communications. Specifically, UEs that have highly similar effective eigenspaces among each other will be grouped together. Let $G < K$ denote the number of user groups, and the group set is denoted by $\mathbb{G} = \{1, 2, \ldots, G\}$. A binary variable matrix $\mathbf{X} = [x_{g,k}]_{g\in\mathbb{G}, k\in\mathbb{K}} \in \{0,1\}^{G\times K}$ is introduced, where

$$x_{g,k} = \begin{cases} 1, & \text{UE } k \text{ is partitioned into group } g, \\ 0, & \text{otherwise.} \end{cases} \tag{36}$$

Accordingly, the UE set of group g is denoted by $\mathbb{K}_g = \{k : x_{g,k} = 1, \forall\, k \in \mathbb{K}\}$, with $|\mathbb{K}_g| = K_g$. The similarity of effective eigenspaces among UEs in group g is defined as

$$\Upsilon_{\text{group}}(\mathbb{K}_g) = \frac{\sum_{i=1}^{K_g}\sum_{j=i}^{K_g} \Upsilon(\mathbb{K}_g(i), \mathbb{K}_g(j))}{(K_g+1)/2}, \tag{37}$$

which is normalized by $(K_g+1)/2$ to ensure that it falls within the interval $[1, K_g]$. Then, the similarity sum among all the UEs all the UEs is defined as

$$\Upsilon_{\text{sum}}(\mathbf{X}) = \sum_{g\in\mathbb{G}} \Upsilon_{\text{group}}(\mathbb{K}_g). \tag{38}$$

By optimizing the user grouping index matrix $\mathbf{X}$, the problem of maximizing the sum similarity $\Upsilon_{\text{sum}}(\mathbf{X})$ can be formulated as

$$\begin{aligned} \Upsilon_{\max} = &\max_{\mathbf{X}} \ \Upsilon_{\text{sum}}(\mathbf{X}) \\ \text{s.t.} \quad & C1: \sum_{g\in\mathbb{G}} x_{g,k} = 1, \quad \forall\, k \in \mathbb{K}, \\ & C2: \sum_{k\in\mathbb{K}} x_{g,k} \geqslant 1, \quad \forall\, g \in \mathbb{G}, \\ & C3: x_{g,k} \in \{0,1\}, \quad \forall\, k \in \mathbb{K}, \quad g \in \mathbb{G}, \end{aligned} \tag{39}$$

where constraint $C1$ ensures that each UE is partitioned into only one group, and constraint $C2$ specifies that each group must contain at least one UE.

Problem (39) is a mixed-integer non-convex optimization problem, which is challenging to obtain its globally optimal solution efficiently. Therefore, we propose an efficient greedy algorithm, as summarized in Algorithm 1, which comprises two stages. In Stage 1, G UEs with the least mutual similarity in their effective eigenspaces are first selected, and each is assigned to a distinct group. Then, in Stage 2, the remaining $K-G$ UEs are sequentially assigned to the group whose eigenspace is most similar to theirs. To improve clarity, a more detailed explanation of the algorithm's procedure has also been included. Specifically, we first initialize the assigned UE index set denoted by $\mathbb{T}_{\text{ass}}$, as an empty set, and the unassigned UE index set denoted by $\mathbb{T}_{\text{una}}$, as the complete set of UEs from 1 to K. The UE set for each group g is initialized as empty, i.e., $\mathbb{K}_g = \emptyset$, $\forall\, g$. During Stage 1, if no UE has yet been assigned to any group, a random UE is selected from the unassigned set $\mathbb{T}_{\text{una}}$ to serve as the first group representative, and is regarded as the optimal UE k^*, as shown in Step 4. Otherwise, for each unassigned UE $k \in \mathbb{T}_{\text{una}}$, we evaluate the similarity metric $\Upsilon_{\text{group}}(\mathbb{T}_{\text{ass}} \cup \{k\})$ in (37), as shown in Step 8. The UE that minimizes this similarity is then selected as the optimal UE k^* and assigned to a new group, as described in Steps 9 and 10. This process repeats until one UE has been assigned to each of the G groups. In Stage 2, the remaining $K-G$ UEs are assigned to groups one by one. For each UE, a candidate is first randomly selected from the unassigned set $\mathbb{T}_{\text{una}}$, as described in Step 17. Then, the sum similarity Υ_{sum} in (38) between this UE and the users already assigned to each group is evaluated, as shown in Step 20. The UE is subsequently assigned to the group that yields the maximum sum similarity metric, as indicated in Steps 21 and 22. This process is repeated until all the remaining $K-G$ UEs have been grouped.

Unlike exhaustive search-based grouping, which incurs prohibitive complexity with $\sum_{i=1}^{G}(-1)^i \frac{(G-i)^K}{i!(G-i)!}$ calculations, i.e., the second-kind Stirling number [42], our proposed greedy algorithm significantly reduces the computational complexity, achieving a complexity order of $O(GK)$. Furthermore, by grouping users based on the similarity of their statistical CSI eigenspaces, the proposed Algorithm 1 enables a more refined user grouping scheme for the following NF-JSDM framework.

Algorithm 1 An efficient user grouping algorithm.

Initialize $\mathbb{T}_{\text{una}} = \{1, 2, \ldots, K\}$, $\mathbb{T}_{\text{ass}} = \emptyset$, and $\mathbb{K}_g = \emptyset$, $\forall\, g$;
For $g \in \mathbb{G}$
  If $\mathbb{T}_{\text{ass}} = \emptyset$
    Randomly pick a UE k^* from $\mathbb{T}_{\text{una}}$;
  Else
    Let $\Upsilon_{\min} = +\infty$;
    For $k \in \mathbb{T}_{\text{una}}$
      Calculate $\Upsilon_{\text{group}'} = \Upsilon_{\text{group}}(\mathbb{T}_{\text{ass}} \cup \{k\})$;
      If $\Upsilon_{\min} > \Upsilon_{\text{group}'}$
        Update $\Upsilon_{\min} = \Upsilon_{\text{group}'}$ and $k^* = k$;
      End
    End
  End
  Update $\mathbb{K}_g = \{k^*\}$, $\mathbb{T}_{\text{una}} = \mathbb{T}_{\text{una}}/\{k^*\}$ and $\mathbb{T}_{\text{ass}} = \mathbb{T}_{\text{ass}} \cup \{k^*\}$;
End
While $\mathbb{T}_{\text{una}} \neq \emptyset$ **do**
  Randomly pick a UE k from $\mathbb{T}_{\text{una}}$;
  Let $\Upsilon_{\max} = -\infty$;
  For $g \in \mathbb{G}$
    Calculate $\Upsilon_{\text{sum}} = \sum_{i \in \mathbb{G}, i \neq g} \Upsilon_{\text{group}}(\mathbb{K}_i) + \Upsilon_{\text{group}}(\mathbb{K}_g \cup \{k\})$;
    If $\Upsilon_{\max} < \Upsilon_{\text{sum}}$
      Update $\Upsilon_{\max} = \Upsilon_{\text{sum}}$ and $g^* = g$;
    End
  End
  Update $\mathbb{K}_{g^*} = \mathbb{K}_{g^*} \cup \{k\}$ and $\mathbb{T}_{\text{una}} = \mathbb{T}_{\text{una}}/\{k\}$;
End
Output $\Upsilon_{\max}$ and $\mathbb{K}_g$, $\forall g \in \mathbb{G}$.

4.3 Three-stage NF-JSDM

According to the proposed user grouping scheme in Section 4.2, we propose the NF-JSDM for XL-MIMO communications. The NF-JSDM method takes into account the general near-field NUSW-based channel with non-zero mean, as well as the differences in statistical CSI among UEs within the same group. First, based on (2), the received signal of group g, denoted by $\mathbf{y}_g = [y_k]_{k \in \mathbb{K}_g} \in \mathbb{C}^{K_g \times 1}$, is expressed as

$$\mathbf{y}_g = \mathbf{H}_g^{\mathrm{H}} \mathbf{V} \mathbf{a} + \mathbf{n}_g = \mathbf{H}_g^{\mathrm{H}} \mathbf{V}_g \mathbf{a}_g + \underbrace{\sum_{g' \neq g} \mathbf{H}_g^{\mathrm{H}} \mathbf{V}_{g'} \mathbf{a}_{g'}}_{\text{inter-group interference}} + \mathbf{n}_g, \tag{40}$$

where $\mathbf{H}_g = [\mathbf{h}_k]_{k \in \mathbb{K}_g} \in \mathbb{C}^{N \times K_g}$ is the channel matrix of group g, with $\mathbf{h}_k$ defined in (34); $\mathbf{V}_g = [\mathbf{v}_k]_{k \in \mathbb{K}_g} \in \mathbb{C}^{N \times K_g}$ is the precoding matrix of group g; $\mathbf{a}_g = [a_k]_{k \in \mathbb{K}_g} \in \mathbb{C}^{K_g \times 1}$ is the symbol vector of group g; and $\mathbf{n}_g = [n_k]_{k \in \mathbb{K}_g} \in \mathbb{C}^{K_g \times 1}$ is the noise vector at group g, $\forall\, g \in \mathbb{G}$.

Since *Identical Statistical CSI Assumption* is invalid in XL-MIMO communications, UEs in the same group are considered to have similar but not identical effective eigenspaces, i.e., $\boldsymbol{\Theta}_k^{\text{eff}} \approx \boldsymbol{\Theta}_{k'}^{\text{eff}}, \forall\, k, k' \in \mathbb{K}_g$ and $k \neq k'$. Therefore, unlike the two-stage FF-JSDM, the NF-JSDM is designed as a three-stage precoding with an additional stage for aligning the dominant channel directions of each UE, which is expressed as

$$\mathbf{V}_g = \mathbf{B}_g \mathbf{W}_g \mathbf{P}_g, \tag{41}$$

where $\mathbf{B}_g \in \mathbb{C}^{N \times b_g}$ is the first-stage precoding for group g, with b_g being an integer variable that needs to be determined; $\mathbf{W}_g = [\mathbf{W}_k]_{k \in \mathbb{K}_g} \in \mathbb{C}^{b_g \times w_g}$ is the second-stage precoding for group g, with $\mathbf{W}_k \in \mathbb{C}^{b_g \times w_k}$ being the second-stage precoding of UE k, and w_k denoting the number of dominant eigenmodes of UE k after the first-stage precoding, satisfying $w_g = \sum_{k \in \mathbb{K}_g} w_k$; $\mathbf{P}_g = [\mathbf{p}_k]_{k \in \mathbb{K}_g} \in \mathbb{C}^{w_g \times K_g}$ is the third-stage precoding for group g, with $\mathbf{p}_k \in \mathbb{C}^{w_g \times 1}$ being the third-stage precoding of UE k, satisfying $w_g \geqslant K_g$.

The first-stage precoding $\mathbf{B}_g$ depends on the statistical CSI, and is intended to eliminate inter-group interference, so that

$$\mathbf{H}_g^{\mathrm{H}} \mathbf{B}_{g'} \approx \mathbf{0}, \quad \forall\, g' \neq g. \tag{42}$$

Moreover, the second-stage precoding $\mathbf{W}_k$ also leverages the statistical CSI, and aims to align eigen-beamforming along the dominant eigenmodes of UE k's equivalent channel after precoding $\mathbf{B}_g$, denoted by $\mathbf{h}_k^{\text{equ}} = \mathbf{B}_g^{\mathrm{H}} \mathbf{h}_k \in \mathbb{C}^{b_g \times 1}$, $\forall k \in \mathbb{K}_g$. Furthermore, the third-stage precoding $\mathbf{P}_g$ is designed to exploit the spatial multiplexing of group g, relying on the instantaneous effective channel after the first two stages, denoted by $\mathbf{H}_g^{\text{ins}} = \mathbf{W}_g^{\mathrm{H}} \mathbf{B}_g^{\mathrm{H}} \mathbf{H}_g \in \mathbb{C}^{w_g \times K_g}$.

Thus, the instantaneous effective CSI required for NF-JSDM is denoted by $\mathbf{H}^{\text{ins}} = \text{diag}\left(\left[\mathbf{H}_g^{\text{ins}}\right]_{\forall g}\right) \in \mathbb{C}^{W\times K}$, with $W = \sum_{g=1}^{G} w_g$ and $W \ll N$. The specific design of the developed three-stage NF-JSDM is as follows.

4.3.1 *First-stage precoding*

The inter-group interference can be eliminated exactly when $\text{Span}(\mathbf{U}_g^{\text{eff}}) \subseteq \text{Span}^{\perp}(\{\mathbf{U}_{g'}^{\text{eff}} : \forall\, g' \neq g\})$, where $\mathbf{U}_g^{\text{eff}} \in \mathbb{C}^{N\times r_g^{\text{eff}}}$ is the maximal linearly independent system of the matrix $\left[\mathbf{U}_k^{\text{eff}}\right]_{k\in\mathbb{K}_g}$ with $r_g^{\text{eff}} = \text{rank}\left(\left[\mathbf{U}_k^{\text{eff}}\right]_{k\in\mathbb{K}_g}\right)$, $\forall\, g$. Thus, the spatial DoF of group g, denoted by S_g, is required to satisfy

$$S_g = \dim\big(\text{Span}(\mathbf{U}_g^{\text{eff}}) \cap \text{Span}^{\perp}(\{\mathbf{U}_{g'}^{\text{eff}} : \forall\, g' \neq g\})\big) \geqslant K_g. \tag{43}$$

Due to the difference in effective eigenspaces among adjacent UEs in XL-MIMO communications, even the UEs that experience the same scatterer set may be divided into multiple groups. In this case, the spatial DoF of these groups will be limited. For example, consider a scenario where UEs experiencing the same scatterer set are divided into two groups, namely $g = 1$ and $g = 2$. Consequently, the number of effective eigenspaces shared by the two groups is greater than 0, i.e., $\dim\big(\text{Span}(\mathbf{U}_{g=1}^{\text{eff}}) \cap \text{Span}(\mathbf{U}_{g=2}^{\text{eff}})\big) > 0$. Hence, when the exact inter-group interference elimination is enforced, the spatial DoF of group $g = 1$ is $S_{g=1} = r_{g=1}^{\text{eff}} - \dim\big(\text{Span}(\mathbf{U}_{g=1}^{\text{eff}}) \cap \text{Span}(\mathbf{U}_{g=2}^{\text{eff}})\big)$, which is limited by the number of shared eigenspaces. To relax the spatial DoF limitation, we first determine the set of groups with similar effective eigenspaces to group g, denoted by $\mathbb{T}_g$, according to the following criteria

$$\mathbb{T}_g = \left\{g' : \Upsilon\left(\mathbf{U}_g^{\text{eff}}, \mathbf{U}_{g'}^{\text{eff}}\right) \in (\varepsilon_{\min}, \varepsilon_{\max}), \forall\, g' \neq g\right\}, \tag{44}$$

where the interval $(\varepsilon_{\min}, \varepsilon_{\max})$ is the judgment threshold. The case of $\Upsilon\left(\mathbf{U}_g^{\text{eff}}, \mathbf{U}_{g'}^{\text{eff}}\right) \geqslant \varepsilon_{\max}$ is considered as excessive grouping, implying that groups g and g' share approximately identical eigenspaces.

Subsequently, we utilize the approximate inter-group interference elimination to design the first-stage precoding. This involves selecting $r_{g'}^{\text{eff}*}$ eigenmodes from the effective eigenspaces of group g' that are least similar to the effective eigenspaces of group g, denoted by $\mathbf{U}_{g'}^{\text{eff}*} \in \mathbb{C}^{N\times r_{g'}^{\text{eff}*}}$, $\forall\, g' \in \mathbb{T}_g$. Thus, the constraint in (43) is relaxed to

$$S_g^{*} = \dim\big(\text{Span}(\mathbf{U}_g^{\text{eff}}) \cap \text{Span}^{\perp}\big(\mathbf{U}_{\mathbb{T}_g}^{\text{eff}*} \cup \mathbf{U}_{\mathbb{T}_g'}^{\text{eff}}\big)\big) \geqslant K_g, \tag{45}$$

where $\mathbf{U}_{\mathbb{T}_g}^{\text{eff}*} = \left[\mathbf{U}_{g'}^{\text{eff}*}\right]_{g'\in\mathbb{T}_g} \in \mathbb{C}^{N\times r_{\mathbb{T}_g}^{\text{eff}*}}$ and $\mathbf{U}_{\mathbb{T}_g'}^{\text{eff}} = \left[\mathbf{U}_{g'}^{\text{eff}}\right]_{g'\in\mathbb{T}_g'} \in \mathbb{C}^{N\times r_{\mathbb{T}_g'}^{\text{eff}}}$, with $r_{\mathbb{T}_g}^{\text{eff}*} = \sum_{g'\in\mathbb{T}_g} r_{g'}^{\text{eff}*}$, $r_{\mathbb{T}_g'}^{\text{eff}} = \sum_{g'\in\mathbb{T}_g'} r_{g'}^{\text{eff}}$, and $\mathbb{T}_g' = \{g' : \forall g' \neq g \text{ and } g' \notin \mathbb{T}_g\}$ denoting the set of groups excluding group g and group set $\mathbb{T}_g$. The selection of parameter $r_{g'}^{\text{eff}\star}$ depends on the tolerable signal power outside the subspace spanned by the corresponding dominant eigenmodes, $\forall\, g' \in \mathbb{T}_g$.

First, we define the eigenmodes composed of groups except group g, denoted as $\mathbf{\Xi}_{-g}$, which are formulated as $\mathbf{\Xi}_{-g} = \left[\mathbf{U}_{\mathbb{T}_g}^{\text{eff}*}, \mathbf{U}_{\mathbb{T}_g'}^{\text{eff}}\right] \in \mathbb{C}^{N\times\left(r_{\mathbb{T}_g}^{\text{eff}*} + r_{\mathbb{T}_g'}^{\text{eff}}\right)}$. To eliminate inter-group interference, $\mathbf{B}_g$ needs to lie in the null space of the matrix $\mathbf{\Xi}_{-g}$. The singular value decomposition (SVD) of $\mathbf{\Xi}_{-g}$ is given by

$$\mathbf{\Xi}_{-g} = \left[\mathbf{E}_{-g}^{\text{non-null}}, \mathbf{E}_{-g}^{\text{null}}\right] \mathbf{\Sigma}_{-g} \mathbf{F}_{-g}^{\text{H}}, \tag{46}$$

where $\mathbf{E}_{-g}^{\text{non-null}} \in \mathbb{C}^{N\times\text{rank}(\mathbf{\Xi}_{-g})}$ represents the non-null space spanned by the non-zero singular values of $\mathbf{\Xi}_{-g}$, and $\mathbf{E}_{-g}^{\text{null}} \in \mathbb{C}^{N\times(N-\text{rank}(\mathbf{\Xi}_{-g}))}$ corresponds to the desired null space spanned by the zero singular values. As a result, the precoding $\mathbf{B}_g$ is expressed as follows, with $b_g = N - \text{rank}\left(\mathbf{\Xi}_{-g}\right)$

$$\mathbf{B}_g = \mathbf{E}_{-g}^{\text{null}}. \tag{47}$$

4.3.2 *Second-stage precoding*

After the first-stage precoding, the received signal $\mathbf{y}_g$ in (40) is expressed as

$$\mathbf{y}_g \approx \mathbf{H}_g^{\text{H}} \mathbf{V}_g \mathbf{a}_g + \mathbf{n}_g. \tag{48}$$

The second-stage precoding $\mathbf{W}_k$ is obtained by projecting the eigen-beamforming along the dominant eigenmodes of the equivalent channel's correlation matrix, i.e., $\mathbf{C}_k = \mathbb{E}\left[\mathbf{h}_k^{\text{equ}}(\mathbf{h}_k^{\text{equ}})^{\text{H}}\right] \in \mathbb{C}^{b_g\times b_g}$, $\forall\, k \in \mathbb{K}_g$. Thus, the SVD of the correlation matrix $\mathbf{C}_k$ is given by

$$\mathbf{C}_k = \mathbf{B}_g^{\text{H}} \mathbf{U}_k^{\text{eff}} (\mathbf{\Lambda}_k^{\text{eff}})^{\frac{1}{2}} \mathbb{E}\left[\mathbf{z}_k^{\text{eff}} (\mathbf{z}_k^{\text{eff}})^{\text{H}}\right] (\mathbf{\Lambda}_k^{\text{eff}})^{\frac{1}{2}} (\mathbf{U}_k^{\text{eff}})^{\text{H}} \mathbf{B}_g = \hat{\mathbf{U}}_k \hat{\mathbf{\Lambda}}_k \hat{\mathbf{U}}_k^{\text{H}}, \tag{49}$$

Table 3 Location regions of UEs and scatterer clusters.

UE index	Location region (m)	Cluster index	Location region (m)
1–5	$x \in [20, 25]$, $y \in [20, 25]$	2	$x \in [80, 90]$, $y \in [-110, -100]$
6–10	$x \in [45, 50]$, $y \in [55, 60]$		
11–15	$x \in [70, 85]$, $y \in [-90, -80]$	3	$x \in [220, 230]$, $y \in [0, 10]$
16–20	$x \in [150, 160]$, $y \in [0, 15]$		

where $\hat{\mathbf{U}}_k = \left[\hat{\mathbf{U}}_k^{\text{non-null}}, \hat{\mathbf{U}}_k^{\text{null}}\right] \in \mathbb{C}^{b_g \times b_g}$ and $\hat{\mathbf{\Lambda}}_k \in \mathbb{C}^{b_g \times b_g}$ denote the unitary eigenvector matrix and the diagonal eigenvalue matrix of $\mathbf{C}_k$, respectively, with $\hat{\mathbf{U}}_k^{\text{non-null}} \in \mathbb{C}^{b_g \times w_k}$ being the non-null space, which contains w_k dominant eigenmodes, and $w_k = \text{rank}(\mathbf{C}_k) \leqslant \min\{b_g, r_k^{\text{eff}}\}$. Thus, $\mathbf{W}_k$ is expressed as

$$\mathbf{W}_k = \hat{\mathbf{U}}_k^{\text{non-null}}, \ \forall k \in \mathbb{K}_g. \tag{50}$$

For the NF-JSDM, the number of equivalent eigenspaces of UE k equals $\text{rank}(\mathbb{E}\left[(\mathbf{h}_k^{\text{H}}\mathbf{B}_g\mathbf{W}_k)^{\text{H}}\mathbf{h}_k^{\text{H}}\mathbf{B}_g\mathbf{W}_k\right])$ after the first two stages of precoding, whereas for the conventional FF-JSDM, that number equals $\text{rank}(\mathbb{E}\left[(\mathbf{h}_k^{\text{H}}\mathbf{B}_g)^{\text{H}}\mathbf{h}_k^{\text{H}}\mathbf{B}_g\right])$ with this $\mathbf{B}_g$ representing the first-stage precoding of the FF-JSDM.

4.3.3 *Third-stage precoding*

The ZF precoding is introduced as the third-stage precoding $\mathbf{P}_g$ for spatial multiplexing, given by

$$\mathbf{P}_g = \mathbf{H}_g^{\text{ins}}\left(\left(\mathbf{H}_g^{\text{ins}}\right)^{\text{H}}\mathbf{H}_g^{\text{ins}}\right)^{-1}. \tag{51}$$

Therefore, the precoding of UE k in group g is given by

$$\mathbf{v}_k = \frac{\mathbf{B}_g\mathbf{W}_g\mathbf{p}_k}{\|\mathbf{B}_g\mathbf{W}_g\mathbf{p}_k\|}, \ k \in \mathbb{K}_g. \tag{52}$$

Subsequently, the signal-to-interference-plus-noise ratio (SINR) of UE k is expressed as

$$\gamma_k = \frac{\rho|\mathbf{h}_k^{\text{H}}\mathbf{v}_k|^2}{1 + \rho\sum_{i\in\mathbb{K}/\{k\}}|\mathbf{h}_k^{\text{H}}\mathbf{v}_i|^2}, \tag{53}$$

where $\rho = \frac{P_0}{K\sigma^2}$ is the transmit signal-to-noise ratio (SNR). The sum rate of K UEs in bps/Hz is

$$R = \sum_{k\in\mathbb{K}}\log_2(1 + \gamma_k). \tag{54}$$

5 Simulation

Unless otherwise stated, the simulation parameters are defined in Table 2, and $\mathcal{K} = 0.01$ is considered. Besides the scatterer cluster considered in Table 2, namely cluster 1, two additional clusters are taken into account, whose location regions are defined in Table 3. The number of UEs is set to $K = 20$, and their randomly distributed locations are also defined in Table 3. The positions of UEs and scatterers in Table 3 are carefully designed to ensure that all users and scatterers are located within the near-field region of the XL-array, thereby capturing key channel characteristics of practical near-field communication environments. It is noted that an eigenvalue smaller than 0.1% of the total sum of eigenvalues is considered to be negligible. The judgment threshold $(\varepsilon_{\min}, \varepsilon_{\max})$ is set to $(0.2, 0.9)$. In addition, the number of selected eigenmodes $r_{g'}^{\text{eff}\star}$ for any group g' is set to be equal, where $\forall\, g' \in \mathbb{T}_g$ and $\forall\, g \in \mathbb{G}$.

Figure 3 shows the relationship between the number of groups G and the maximum sum similarity $\Upsilon_{\max}$ in (39) obtained through the proposed Algorithm 1, the exhaustive searching algorithm and the random grouping algorithm, where the developed near-field model is considered. To ensure computational feasibility and a fair performance comparison, the number of UEs is uniformly set to $K = 7$ for the proposed Algorithm 1, exhaustive search, and random grouping, since the computational complexity of exhaustive search becomes prohibitive for larger K. The number of random realizations for UE locations is set to 30. From Figure 3, it can be observed that the proposed Algorithm 1 exhibits a maximum performance loss of less than 7% when $G = 3$, compared to the

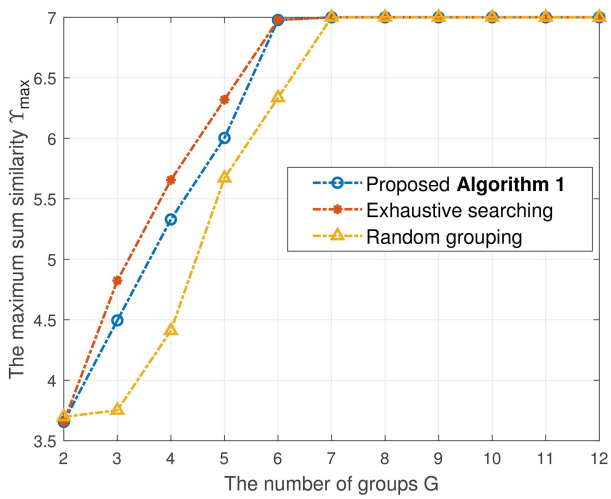

Figure 3 (Color online) The number of groups G versus the maximum sum similarity $\Upsilon_{\max}$ obtained through the proposed Algorithm 1, exhaustive searching, and random grouping, where the developed near-field model is considered and $K = 7$.

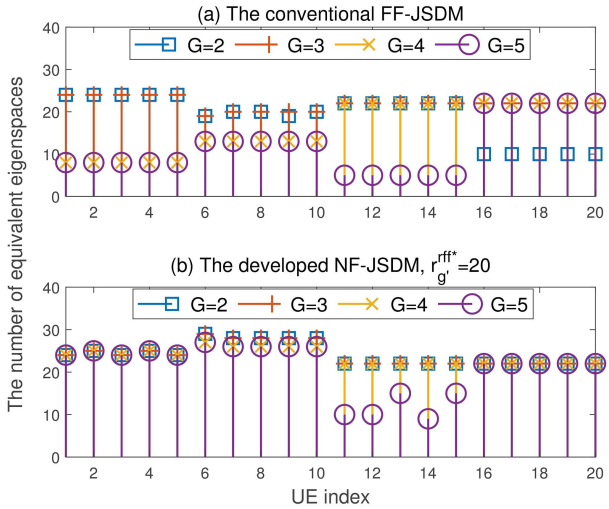

Figure 4 (Color online) The number of equivalent eigenspaces for each UE (a) the conventional FF-JSDM and (b) the proposed NF-JSDM with $r_{g'}^{\text{eff}\star} = 20$, $\forall g' \in \mathbb{T}_g$ and $\forall g$, where the ground-truth near-field channel model is considered and $G = 2$, 3, 4, 5 are considered.

exhaustive searching algorithm. Although the random grouping method has the lowest computational complexity, it incurs the highest performance loss. It is also observed that for the three methods, $\Upsilon_{\max}$ increases as the number of groups G grows when G is less than K. This is expected, as the larger G is, the less likely it is that UEs with significantly different statistical CSI will be grouped together. These results validate that the proposed Algorithm 1 is effective for user grouping in XL-MIMO communications.

Figure 4 shows the number of equivalent eigenspaces for each UE by considering the conventional FF-JSDM in Figure 4(a) and the proposed NF-JSDM with $r_{g'}^{\text{eff}\star}$ set to 20 in Figure 4(b), $\forall\, g' \in \mathbb{T}_g$ and $\forall\, g$, where the ground-truth near-field channel model is considered and different numbers of user groups ($G = 2$, 3, 4, 5) are considered based on Algorithm 1. It is observed from Figure 4 that in the FF-JSDM scheme, all UEs achieve the maximum equivalent eigenspaces only when $G = 3$, whereas in the proposed NF-JSDM, this maximum is achieved for $G = 2$, 3, and 4. This is because when $G = 2$, this limitation in FF-JSDM arises due to the assumption of identical statistical CSI within the same group, which prevents effective alignment of channel eigenspaces for UEs experiencing different scatterer clusters but assigned to the same group. On the other hand, when $G = 4$ or 5, the reduced equivalent eigenspaces in FF-JSDM result from excessive grouping, which leads to a decrease in the spatial DoFs available to each group. In addition, as shown in Figure 4, when $G = 3$, there is a significant difference in the number of equivalent eigenspaces for UEs 6–10 between the FF- and NF-JSDM methods. This is because UEs 1–10 are grouped together, and the FF-JSDM fails to align the channel eigenspaces of UEs 6–10 due to the assumption of identical statistical CSI. Furthermore, Figure 4 shows that the number of equivalent eigenspaces for UEs 11–15 in the developed NF-JSDM decreases when $G = 5$. This is because UEs 11–15, which experience similar scattering environments, are partitioned into two separate groups, thereby reducing the spatial DoFs available to these two groups. These results verify the assumption of non-identical statistical CSI in the NF-JSDM and highlight the importance of choosing an appropriate number of user groups in the developed NF-JSDM method.

Figure 5 gives a comparison of the sum rate obtained through the ZF precoding utilizing full CSI, the conventional FF-JSDM [17, 18], and the proposed NF-JSDM with $r_{g'}^{\text{eff}\star} = 20$, $\forall\, g' \in \mathbb{T}_g$ and $\forall\, g \in \mathbb{G}$, under different numbers of user groups based on Algorithm 1, where both the far- and near-field channel models are considered. It is observed from Figure 5 that in the far-field scenario, when $G = 2$, the sum rate of the NF-JSDM is higher than that of the FF-JSDM, and when $G \geqslant 3$, the curves of the proposed NF-JSDM and FF-JSDM perfectly overlap. Especially when $G = 3$, the FF- and NF-JSDM can achieve approximately the same performance as that of the ZF precoding. This indicates that the proposed NF-JSDM is applicable to the far-field communications, and its performance is superior to that of the FF-JSDM when the number of groups is insufficient. Additionally, Figure 5 shows that in the near-field scenario, the NF-JSDM can achieve approximately the same performance as ZF when $G = 4$, whereas there is a notable performance gap between the FF-JSDM and ZF precoding. The performance loss of the FF-JSDM is caused by the reduction of equivalent eigenspaces, which corresponds to the results in Figure 4. Compared to ZF precoding, the proposed NF-JSDM can achieve comparable performance while significantly reducing the channel

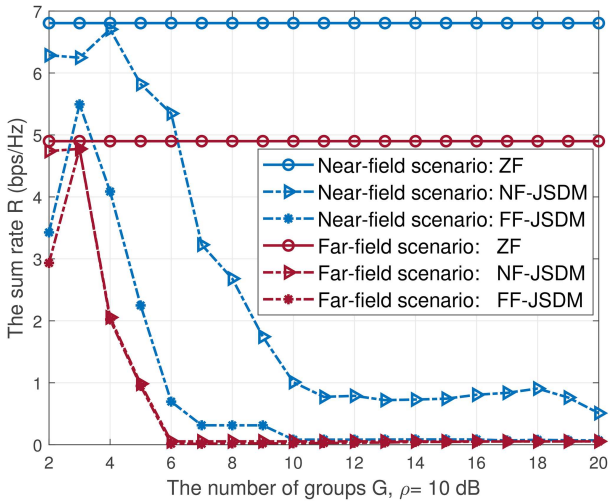

Figure 5 (Color online) The sum rate R versus the different numbers of user groups under the ZF precoding, the conventional FF-JSDM [17,18], and the proposed NF-JSDM with $r_{g'}^{\text{eff}\star} = 20$, $\forall\, g' \in \mathbb{T}_g$ and $\forall\, g \in \mathbb{G}$, where both the near- and far-field channel models are considered.

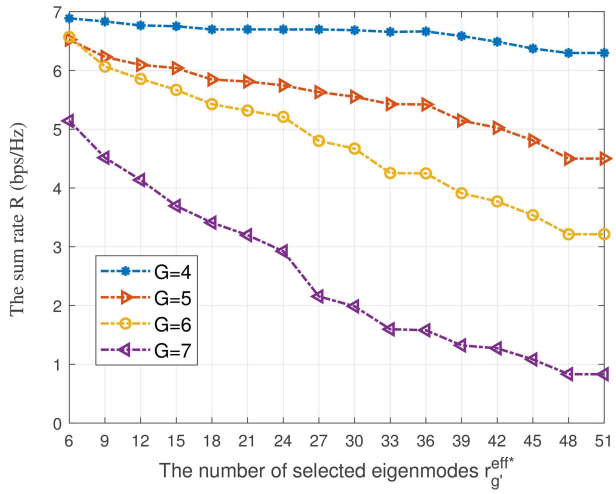

Figure 6 (Color online) The sum rate versus the number of selected eigenmodes $r_{g'}^{\text{eff}\star}$ in the proposed NF-JSDM, $\forall\, g' \in \mathbb{T}_g$ and $\forall g \in \mathbb{G}$, where the ground-truth near-field channel model is considered and $G = 4, 5, 6, 7$ are considered.

estimation overhead, as it reduces the required instantaneous CSI dimensionality from $N \times K$ to $W \times K$, where $W \ll N$. As shown in Figure 5, the sum rates of both NF-JSDM and FF-JSDM first increase and then decrease with the number of groups G. When G is small, users with similar eigenspaces tend to be grouped together, and eliminating intra-group interference for any user k may reduce its available spatial DoF, leading to performance loss. As G becomes too large, users with similar eigenmodes may be assigned to different groups, limiting the spatial DoF available for inter-group interference suppression, which can also be observed in Figure 4. These observations demonstrate that the proposed NF-JSDM is applicable to XL-MIMO communications, and it is more general than the conventional FF-JSDM.

Figure 6 shows the sum rate versus the number of selected eigenmodes $r_{g'}^{\text{eff}\star}$ in the proposed NF-JSDM, $\forall\, g' \in \mathbb{T}_g$, $\forall\, g \in \mathbb{G}$, where the ground-truth near-field channel model is considered and different numbers of user groups ($G = 4, 5, 6, 7$) are considered based on Algorithm 1. The variable $r_{g'}^{\text{eff}\star}$ ranges from 6 to 51 with a step size of 3. It is observed from Figure 6 that the sum rates decrease as the number of selected eigenmodes increases. This is because the increase in eigenmodes reduces the DoF available to group g for the first-stage precoding design, thereby increasing inter-group interference and ultimately leading to a reduced sum rate. It is worth noting, however, that the reduction in DoF also results in a lower dimension for the first-stage precoding matrix, which helps to reduce computational complexity. This highlights a trade-off between performance and complexity in selecting the number of eigenmodes. Moreover, Figure 6 also shows that, for a given number of selected eigenmodes $r_{g'}^{\text{eff}\star}$, increasing the number of user groups leads to a more rapid decline in the sum rate. This is because excessive grouping imposes stricter spatial DoF constraints on each group, thereby reducing the equivalent eigenspaces available to more UEs. The above results offer useful guidelines for determining the number of eigenmodes to maximize the sum rate. In particular, by selecting the most appropriate number of user groups, it is possible to choose the largest feasible number of eigenmodes $r_{g'}^{\text{eff}\star}$ within an acceptable performance degradation threshold, to effectively reduce computational complexity.

6 Conclusion

This paper investigated the NF-JSDM for multi-user XL-MIMO communications with mixed LoS and NLoS links, which is based on the proposed user grouping algorithm according to the similarity of the developed near-field statistical CSI. In the developed NF-JSDM, the prior knowledge of near-field statistical CSI is modeled by taking into account the NUSW, partial array visibility, and spatial consistency. Based on the similarity of the effective eigenspace in near-field statistical CSI, an efficient algorithm was proposed to achieve precise user grouping. According to the proposed user grouping scheme, the more general NF-JSDM was developed, which is not limited to the assumption of zero channel mean, and considers the differences in statistical CSI of UEs within the same group

and near-field channel characteristics. Simulation results validated the effectiveness of the proposed user grouping algorithm, and demonstrated that the developed NF-JSDM is applicable to XL-array communications. In the future, we will further extend the proposed NF-JSDM precoding framework to support multi-stream transmission per user and adapt it to hybrid digital-analog architectures, which are more practical for XL-MIMO implementations.

Acknowledgements This work was supported by Natural Science Foundation for Distinguished Young Scholars of Jiangsu Province (Grant No. BK20240070) and Fundamental Research Funds for the Central Universities of China (Grant Nos. 2242022k60004, 2242023K5003).

References

1 Heath Jr R W, Lozano A. Foundations of MIMO Communication. Cambridge: Cambridge University Press, 2018
2 Lu H, Zeng Y, You C, et al. A tutorial on near-field XL-MIMO communications toward 6G. IEEE Commun Surv Tutorials, 2024, 26: 2213–2257
3 Zhang J, Bjornson E, Matthaiou M, et al. Prospective multiple antenna technologies for beyond 5G. IEEE J Sel Areas Commun, 2020, 38: 1637–1660
4 Carvalho E D, Ali A, Amiri A, et al. Non-stationarities in extra-large-scale massive MIMO. IEEE Wireless Commun, 2020, 27: 74–80
5 Dong Z, Li X, Zeng Y. Characterizing and utilizing near-field spatial correlation for XL-MIMO communication. IEEE Trans Commun, 2024, 72: 7922–7937
6 Wang H, Feng C, Zeng Y, et al. Enhancing spatial multiplexing and interference suppression for near-and far-field communications with sparse MIMO. ArXiv:2408.01956
7 Li X, Dong Z, Zeng Y, et al. Multi-user modular XL-MIMO communications: near-field beam focusing pattern and user grouping. IEEE Trans Wireless Commun, 2024, 23: 13766–13781
8 Jeon J, Lee G, Ibrahim A A I, et al. MIMO evolution toward 6G: modular massive MIMO in low-frequency bands. IEEE Commun Mag, 2021, 59: 52–58
9 Cui M, Dai L. Channel estimation for extremely large-scale MIMO: far-field or near-field? IEEE Trans Commun, 2022, 70: 2663–2677
10 Wang Z, Zhang J, Du H, et al. A tutorial on extremely large-scale MIMO for 6G: fundamentals, signal processing, and applications. IEEE Commun Surv Tutorials, 2024, 26: 1560–1605
11 Yang S, Peng Y, Lyu W, et al. Near-field channel estimation for extremely large-scale Terahertz communications. Sci China Inf Sci, 2024, 67: 192302
12 Zhang N, Zhang J, Xing C, et al. Intelligent secure near-field communication. Sci China Inf Sci, 2024, 67: 199302
13 Rao X, Lau V K. Distributed compressive CSIT estimation and feedback for FDD multi-user massive MIMO systems. IEEE Trans Signal Process, 2014, 62: 3261–3271
14 Qi C, Huang Y, Jin S, et al. Sparse channel estimation based on compressed sensing for massive MIMO systems. In: Proceedings of the Proc IEEE ICC, London, 2015. 4558–4563
15 Zeng Y, Xu X. Toward environment-aware 6G communications via channel knowledge map. IEEE Wireless Commun, 2021, 28: 84–91
16 Zeng Y, Chen J, Xu J, et al. A tutorial on environment-aware communications via channel knowledge map for 6G. IEEE Commun Surv Tutorials, 2024, 26: 1478–1519
17 Adhikary A, Junyoung Nam A, Jae-Young Ahn A, et al. Joint spatial division and multiplexing—the large-scale array regime. IEEE Trans Inform Theory, 2013, 59: 6441–6463
18 Adhikary A, Al Safadi E, Samimi M K, et al. Joint spatial division and multiplexing for mm-wave channels. IEEE J Sel Areas Commun, 2014, 32: 1239–1255
19 Liu X, Zhang J, Tang P, et al. Channel sparsity variation and model-based analysis on 6, 26, and 105 GHz measurements. IEEE Trans Veh Technol, 2024, 73: 9387–9397
20 Jiang T, Zhang J, Shafi M, et al. The comparative study of s-v model between 3.5 and 28 GHz in indoor and outdoor scenarios. IEEE Trans Veh Technol, 2020, 69: 2351–2364
21 Abdi A, Kaveh M. A space-time correlation model for multielement antenna systems in mobile fading channels. IEEE J Sel Areas Commun, 2002, 20: 550–560
22 Dong Z, Li X, Zeng Y, et al. Near-field spatial correlation for multi-path XL-array communications with partial visibility. In: Proceedings of IEEE GlobeCom, Kuala Lumpur, Malaysia, 2023. 1525–1530
23 Lu H, Zeng Y. Communicating with extremely large-scale array/surface: unified modeling and performance analysis. IEEE Trans Wireless Commun, 2021, 21: 4039–4053
24 Bayraktar M, Guvensen G M. An efficient interference-aware constrained massive MIMO beamforming for mm-wave JSDM. IEEE Access, 2021, 9: 87877–87897
25 Bayraktar M, Guvensen G M. An efficient constrained mm-wave hybrid massive MIMO beamforming for JSDM based NOMA. In: Proceedings of the ICC 2021-IEEE International Conference on Communications, 2021. 1–6
26 Amiri A, Manchón C N, De Carvalho E. A message passing based receiver for extra-large scale MIMO. In: Proceedings of the Proc IEEE International Workshop on CAMSAP, 2019. 564–568
27 Ali A, Carvalho E D, Heath R W. Linear receivers in non-stationary massive MIMO channels with visibility regions. IEEE Wireless Commun Lett, 2019, 8: 885–888
28 Zhang H, Shlezinger N, Guidi F, et al. Beam focusing for near-field multiuser MIMO communications. IEEE Trans Wireless Commun, 2022, 21: 7476–7490
29 Wei L, Huang C, Alexandropoulos G C, et al. Tri-polarized holographic MIMO surfaces for near-field communications: channel modeling and precoding design. IEEE Trans Wireless Commun, 2023, 22: 8828–8842
30 Selvan K T, Janaswamy R. Fraunhofer and Fresnel distances: unified derivation for aperture antennas. IEEE Antennas Propag Mag, 2017, 59: 12–15
31 Yang J, Zeng Y, Jin S, et al. Communication and localization with extremely large lens antenna array. IEEE Trans Wireless Commun, 2021, 20: 3031–3048
32 Verdone R, Zanella A. Pervasive Mobile and Ambient Wireless Communications: COST Action 2100. London: Springer, 2012
33 Zhu M, Eriksson G, Tufvesson F. The COST 2100 channel model: parameterization and validation based on outdoor MIMO measurements at 300 MHz. IEEE Trans Wireless Commun, 2013, 12: 888–897
34 Flordelis J, Li X, Edfors O, et al. Massive MIMO extensions to the COST 2100 channel model: modeling and validation. IEEE Trans Wireless Commun, 2019, 19: 380–394
35 Skolnik M I. Introduction to Radar Systems. New York: McGraw-Hill1980, 1980
36 Björnson E, Sanguinetti L. Utility-based precoding optimization framework for large intelligent surfaces. In: Proceedings of the 53rd Asilomar Conference on Signals, Systems, and Computers, 2019. 863–867
37 Ertel R B, Cardieri P, Sowerby K W, et al. Overview of spatial channel models for antenna array communication systems. IEEE Pers Commun, 1998, 5: 10–22
38 Dong Z, Zeng Y. Near-field spatial correlation for extremely large-scale array communications. IEEE Commun Lett, 2022, 26: 1534–1538
39 Liu A, Lau V K N. Two-stage subspace constrained precoding in massive MIMO cellular systems. IEEE Trans Wireless Commun, 2015, 14: 3271–3279

40 Gao X, Tufvesson F, Edfors O. Massive MIMO channels-Measurements and models. In: Proceedings of the Proc ACSSC, Pacific Grove, 2013. 280–284
41 Correia L M. Mobile Broadband Multimedia Networks: Techniques, Models and Tools for 4G. San Diego: Academic Press, 2010
42 Rennie B C, Dobson A J. On stirling numbers of the second kind. J Combin Theory, 1969, 7: 116–121

Appendix A

When $\alpha^{\mathrm{v}}(k)=0$, we have $\mathbb{E}[\xi_n(k)]=1$, $\forall\, n$. Then, by setting reference UE as UE k itself, $\left[\bar{\mathbf{h}}_k\right]_n$ in (14) reduces to

$$\left[\bar{\mathbf{h}}_k\right]_n=\beta_k\frac{\mathcal{K}_k}{\mathcal{K}_k+1}\frac{d_k}{d_{k,n}}\mathrm{e}^{-\mathrm{j}\frac{2\pi}{\lambda}d_{k,n}}. \tag{A1}$$

When $N\delta\ll d_k$, the distance $d_{k,n}$ in (6) is approximated as

$$d_{k,n}\approx d_k\sqrt{1-2n\delta\sin\varphi_k/d_k}\overset{(a)}{\approx} d_k-n\delta\sin\varphi_k, \tag{A2}$$

where (a) holds based on the first-order Taylor approximation. By substituting (A2) into (A1), $\left[\bar{\mathbf{h}}_k\right]_n$ reduces to

$$\left[\bar{\mathbf{h}}_k\right]_n=\beta_k\mathcal{K}_k/(\mathcal{K}_k+1)\cdot\mathrm{e}^{-\mathrm{j}\frac{2\pi}{\lambda}(d_k-n\delta\sin\varphi_k)}. \tag{A3}$$

The proof of Lemma 1 is thus completed.

Appendix B

When $\alpha^{\mathrm{v}}(k)=0$, we have $\mathbb{E}[\xi_n(k)\xi_m(k)]=1$, $\forall n,m$. Then, $\left[\mathbf{\Theta}_k^{\mathrm{LoS}}\right]_{n,m}$ in (21) reduces to

$$\left[\mathbf{\Theta}_k^{\mathrm{LoS}}\right]_{n,m}=\beta_k\frac{\mathcal{K}_k}{\mathcal{K}_k+1}\frac{d_k^2}{d_{k,n}d_{k,m}}\mathrm{e}^{-\mathrm{j}\frac{2\pi}{\lambda}(d_{k,n}-d_{k,m})}. \tag{B1}$$

When $N\delta\ll d_k$, $d_{k,n}$ in (6) can be approximated as (A2). By substituting (A2) into (B1), $\left[\mathbf{\Theta}_k^{\mathrm{LoS}}\right]_{n,m}$ reduces to

$$\left[\mathbf{\Theta}_k^{\mathrm{LoS}}\right]_{n,m}=\beta_k\mathcal{K}_k/(\mathcal{K}_k+1)\cdot\mathrm{e}^{-\mathrm{j}\frac{2\pi}{\lambda}(m-n)\delta\sin\varphi_k}. \tag{B2}$$

The proof of Lemma 2 is thus completed.

Appendix C

When $g_k^{\mathrm{VR}}(\mathbf{s})=1$, $\left[\mathbf{\Theta}_k^{\mathrm{NLoS}}\right]_{n,m}$ in (28) is simplified to

$$\left[\mathbf{\Theta}_k^{\mathrm{NLoS}}\right]_{n,m}=\beta_k\frac{1}{\mathcal{K}_k+1}\int_{\mathbf{s}\in\mathbf{S}_k}\frac{r^2(\mathbf{s})}{r_n(\mathbf{s})r_m(\mathbf{s})}\mathrm{e}^{-\mathrm{j}\frac{2\pi}{\lambda}(r_n(\mathbf{s})-r_m(\mathbf{s}))}\mathbb{E}\left[\xi_n(\mathbf{s})\xi_m(\mathbf{s})\right]f_k(\mathbf{s})\mathrm{d}\mathbf{s}. \tag{C1}$$

When $\alpha^{\mathrm{i}}(\mathbf{s})=0$, $\forall\,\mathbf{s}\in\mathbf{S}_k$, we get $\mathbb{E}[\xi_n(\mathbf{s})\xi_m(\mathbf{s})]=1$, $\forall\, n,m$. Therefore, $\left[\mathbf{\Theta}_k^{\mathrm{NLoS}}\right]_{n,m}$ in (C1) is further simplified to

$$\left[\mathbf{\Theta}_k^{\mathrm{NLoS}}\right]_{n,m}=\beta_k\frac{1}{\mathcal{K}_k+1}\int_{\mathbf{s}\in\mathbf{S}_k}\frac{r^2(\mathbf{s})}{r_n(\mathbf{s})r_m(\mathbf{s})}\mathrm{e}^{-\mathrm{j}\frac{2\pi}{\lambda}(r_n(\mathbf{s})-r_m(\mathbf{s}))}f_k(\mathbf{s})\mathrm{d}\mathbf{s}. \tag{C2}$$

When $N\delta\ll r(\mathbf{s})$, $\forall\mathbf{s}\in\mathbf{S}_k$, the distance $r_n(\mathbf{s})$ in (C2) is approximated as

$$r_n(\mathbf{s})\approx r(\mathbf{s})\sqrt{1-2n\delta\sin\theta(\mathbf{s})/r(\mathbf{s})}\overset{(a)}{\approx} r(\mathbf{s})-n\delta\sin\theta(\mathbf{s}), \tag{C3}$$

where (a) holds based on the first-order Taylor approximation, and $\theta(\mathbf{s})$ is the AoA of scatterer $\mathbf{s}$ w.r.t. the positive x-axis. By substituting (C3) into (C2), $\left[\mathbf{\Theta}_k^{\mathrm{NLoS}}\right]_{n,m}$ reduces to

$$\left[\mathbf{\Theta}_k^{\mathrm{NLoS}}\right]_{n,m}\approx\beta_k\frac{1}{\mathcal{K}_k+1}\int_{\mathbf{s}\in\mathbf{S}_k}\mathrm{e}^{-\mathrm{j}\frac{2\pi}{\lambda}(m-n)\delta\sin\theta(\mathbf{s})}f_k(\mathbf{s})\mathrm{d}\mathbf{s}. \tag{C4}$$

As a result, the integral function in (C4) which is only related to $\theta(\mathbf{s})$, can be expressed as

$$\left[\mathbf{\Theta}_k^{\mathrm{NLoS}}\right]_{n,m}\approx\beta_k\frac{1}{\mathcal{K}_k+1}\int_{\theta_k^{\min}}^{\theta_k^{\max}}\mathrm{e}^{-\mathrm{j}\frac{2\pi}{\lambda}(m-n)\delta\sin\theta}f_k(\theta)\mathrm{d}\theta. \tag{C5}$$

Thus, the proof of Lemma 3 is completed.