# Multi-User Key Rate Optimization for Near-Field Extremely Large-Scale Antenna Array Communications

Tianyu Lu, *Member, IEEE,* Liquan Chen, *Senior Member, IEEE,* Junqing Zhang, *Senior Member, IEEE,* and Trung Q. Duong, *Fellow, IEEE*

***Abstract*—Extremely large-scale antenna arrays (ELAA) require near-field spherical wave modeling due to the substantial increase in the number of antennas, which introduces new spatial dimensions to physical layer key generation (PLKG). We investigate multi-user PLKG in near-field environments, where a base station with an ELAA simultaneously generates secret keys with multiple users. We derive an analytical expression for the key rate (KR). By utilizing spatial dimensions of distance and angle in near-field environments, we apply eigenvalue decomposition and singular value decomposition to design precoding matrices to reduce interference among user equipments (UEs) and extract uncorrelated subchannels. Given that the KR is non-convex, we approximate it and optimize the precoding matrix to increase the KR. After precoding design, the KR depends on the transmit power allocated to the subchannels. Two optimization problems are formulated to further optimize transmit power allocation. The first problem focuses on maximizing the sum KR. We apply the Lagrange multiplier method to determine the optimal power allocation variables by searching the Lagrange multiplier. To reduce computational complexity, a supervised feedforward neural network (FNN) is designed to capture the relationship between the power allocation variables and the Lagrange multiplier. The second optimization problem focuses on KR fairness. By introducing a slack variable that is smaller than the KRs of all users, we use the CVX toolbox to find optimal power allocation variables that maximize this slack variable. To further reduce complexity, the Lagrange multiplier method offers an analytical solution for power allocation variables in terms of Lagrange multipliers determined by the slack variable in the high-power case. We employ a bisection algorithm to find the slack variable. Furthermore, we propose an FNN to map transmit power to the slack variable. Simulations demonstrate that the proposed methods efficiently leverage near-field effects for multi-user PLKG, reducing pilot overhead.**

***Index Terms*—Extremely large-scale antenna array, deep learning, multi-user key generation, near-field communications.**

This research is supported by the National Natural Science Foundation of China (No. U22B2026) and the National Key Research and Development Program of China (No. 2020YFE0200600). The work of J. Zhang was supported by the UK EPSRC under grant ID EP/V027697/1. The work of T. Q. Duong was supported in part by the Canada Excellence Research Chair (CERC) Program CERC-2022-00109 and in part by the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant Program RGPIN-2025-04941. (*Corresponding author: L. Chen*)

T. Lu and L. Chen are with the School of Cyber Science and Engineering, Southeast University, Nanjing, 210096, China. (e-mail: effronlu@seu.edu.cn; lqchen@seu.edu.cn)

L. Chen is also with the Purple Mountain Laboratories for Network and Communication Security, Nanjing, 211111, China.

J. Zhang is with the Department of Electrical Engineering and Electronics, University of Liverpool, Liverpool, L69 3GJ, United Kingdom. (email: junqing.zhang@liverpool.ac.uk)

T. Q. Duong is with the Faculty of Engineering and Applied Science, Memorial University, St. John's, NL A1C 5S7, Canada, and with the School of Electronics, Electrical Engineering and Computer Science, Queen's University Belfast, Belfast, U.K, (e-mail: tduong@mun.ca).

## I. Introduction

THE sixth generation (6G) wireless communications will expand application scenarios with high data rate, low latency, and massive connections [1]. With the increase of coverage and network heterogeneity, 6G security and privacy are crucial [2]. Traditional cryptographic schemes consisting of symmetric encryption and public key cryptography (PKC) are not often inapplicable due to key distribution and computational complexity. PKC necessitates the presence of public key infrastructure, which is not always effective for securing device-to-device (D2D) communication [3]. Besides, PKC based on computational security may not apply to lightweight devices and is threatened by quantum computers with powerful computational power [4]. Symmetric encryption is an alternative solution for securing D2D communication, but key distribution is challenging. Pre-distributed keys can address the issue, but they are rarely refreshed and lack scalability [5]. This is particularly problematic in dynamic environments where devices frequently join and leave the network [6].

Physical layer key generation (PLKG) exploits the randomness of the wireless channel for Alice and Bob to generate secret keys. PLKG emerges as a promising technique to offer quantum-resistant secret keys for lightweight devices [7]. Since the wireless channels vary with time, Alice and Bob leverage this natural random source for key generation. Relying on channel reciprocity, Alice and Bob share nearly the same channel characteristics within a short channel snapshot. Besides, the property of spatial correlation makes the wireless channels of eavesdroppers uncorrelated with that of Alice and Bob over distances of the order of a few wavelengths [3].

In 6G communications, the extremely large-scale antenna array (ELAA) is emerging as a promising approach to enhance spectral efficiency and spatial resolution [8]. Since ELAA significantly scales up the number of antennas, the radiated electromagnetic waves need to be modeled by near-field spherical waves, where the signals received by different antennas experience different distances and angles from the transmitter. The near-field spherical waves enable the base station (BS) to concentrate signal energy on both distance and angles, introducing a new distance dimension to PLKG systems. Even if an eavesdropper is positioned between the user and the

BS and shares the same angle as the user, the near-field propagation model allows the user to gain an advantage over the eavesdropper through the distance dimension [9], [10].

As the number of users in PLKG systems increases, the issue of pilot overhead becomes more significant when users generate secret keys with the BS sequentially in a point-to-point manner. To address this, multi-user key generation has recently been explored [11], [12], where users simultaneously generate secret keys with the BS. The orthogonal frequency-division multiple access (OFDMA) modulation is employed to allocate non-overlapping subcarriers to different users, allowing each user to independently generate secret keys with the BS from their respective subcarriers [11]. Furthermore, as the number of antennas in multiple-input multiple-output (MIMO) systems increases, the far-field channel can be transformed into the beam-domain channel, where the beams of each user correspond to different angles of arrival (AoAs) at the BS. Thus, each user can exploit the randomness of beams associated with different AoAs [12]. However, the scheme of [12] is based on far-field channels, where the planar wavefront allows the BS to steer energy toward specific users by exploiting their distinct AoAs. If users share the same AoA, the spatial diversity may not be utilized for PLKG. In contrast, under the spherical-wave channel model, the near-field effect introduces spatial degrees of freedom in both distance and angle. This enables users located at different angles or distances to simultaneously generate secret keys with the BS [13].

Precoding design has recently been explored to enable multiple users to extract randomness from their channels [12], [14]. In massive MIMO networks, a hybrid precoding scheme was developed to mitigate interference from other users [14]. Specifically, users sequentially send pilots to the BS for uplink channel probing, resulting in pilot overhead that scales linearly with the number of users. To allow the BS to probe the channel for all users simultaneously, the sparsity of the far-field beam-domain channel has been leveraged. Precoding matrices are aligned with non-overlapping beams, effectively reducing both interference and pilot overhead [12]. Furthermore, this scheme exploited the randomness from the beams of a user with different AoAs in a single channel probing, thereby providing more random sources. However, this scheme assumed equal power allocation for precoding matrices, leaving the power allocation problem unaddressed. In far-field environments, large antenna arrays lead to asymptotically orthogonal steering vectors across angles, allowing effective angular-alignment-based precoding for multi-user key generation [12]. However, near-field channels exhibit spatially varying path gains, angles, and distances across the array, making angular alignment insufficient. In this case, relying solely on angular alignment may introduce correlation between measurements of paths. Therefore, precoding strategies must jointly exploit both angle and distance diversity to suppress inter-path interference.

The power allocation algorithm for precoding matrices in multi-user key generation has primarily focused on maximizing the sum key rate (KR) for all users within a constrained transmit power budget [12], [14], [15]. The KR quantifies the maximum number of secret keys that can be generated for a user. Due to the non-convex nature of the power allocation problem, a genetic algorithm was proposed to optimize the sum KR [14]. Subsequently, the sum-KR maximization problem was approximated using convex optimization theory to simplify the solution [15]. However, these existing approaches often rely on iterative optimization algorithms, which are computationally complex and difficult to implement in practical scenarios. Therefore, there is a need to develop a simpler algorithm for effective power allocation. Moreover, focusing solely on maximizing the sum KR can lead to the user's KR being sacrificed with the poorest channel quality. In ELAA near-field systems, the distance dimension can be utilized to enable both near and far users to generate secret keys simultaneously. However, if the far user has significantly poor channel quality, the transmit power allocation must consider the sum KR of all users. Therefore, it is essential to prioritize the KR fairness by maximizing the minimum KR in the system, rather than merely maximizing the sum KR.

To facilitate multi-user key generation in ELAA systems, this paper exploits the spatial dimensions of angles and dimensions in near-field channels to design precoding matrices. Furthermore, the power allocation variables associated with these precoding matrices are optimized to maximize the sum KR and address KR fairness. A supervised feedforward neural network (FNN) is trained to determine the optimal power allocation variables. Our main contributions are summarized as follows:

- We investigate the multi-user channel probing scheme in near-field ELAA systems to relieve pilot overhead, where users can simultaneously generate secret keys with the BS. We apply a precoding scheme based on singular value decomposition (SVD) and eigenvalue decomposition (EVD) to mitigate interference from other users by leveraging the angle and distance dimensions of the near-field effect and covert spatial-correlated channels to subchannels. Since the KR is non-convex, we derive the approximate expression of the KR. We optimize the precoding matrix to maximize the KR.
- We investigate maximizing the sum KR through subchannel power allocation. We utilize the Lagrange multiplier method to find the analytical expression of the optimal power allocation variables based on the Lagrange multiplier. We then gather a large dataset of samples pairing transmit power with the corresponding Lagrange multiplier using a water-filling algorithm. To reduce computational complexity, we train a supervised FNN to learn the relationship between these variables. This allows us to directly predict the optimal power allocation variables based on the Lagrange multiplier.
- We further investigate the KR fairness. First, we introduce a slack variable that is smaller than all KRs and use a standard convex optimization solver to maximize the slack variable. Furthermore, by approximating the KRs in high-power scenarios, we derive analytical expressions for the power allocation variables as functions of the Lagrange multipliers and for the Lagrange multipliers as functions of the slack variable. A bisection search algorithm is employed to determine the slack variable.

Additionally, we collect data on various transmit powers and corresponding slack variables from the solver and train a supervised FNN to learn about their relationship. Compared to iterative optimization algorithms, the FNN can directly predict power allocation variables using these analytical expressions.

- Simulations show that our proposed FNN-based power allocation schemes outperform equal allocation schemes in terms of transmit power, the number of antennas, distances and angles. Besides, even when all users share the same angles, the spatial dimension of distance can be exploited to enable simultaneous multi-user key generation in near-field ELAA systems.

The remainder of this paper is organized as follows: Section II presents the multi-user key generation in near-field ELAA systems. A multi-user channel probing scheme is proposed in Section III. The analytical expression of the KR is further derived. In Section VI, we apply the SVD and EVD methods to design the precoding matrices which are transferred into power allocation variables. In Section V, we utilize the Lagrange multiplier method to maximize the sum KR, where the optimal power allocation variables are determined by the Lagrange multiplier. We train a FNN to learn the relationship between the transmit power and the Lagrange multiplier. In Section VI, we use the FNN-based power allocation scheme to address KR fairness. Section VII presents the simulation results, followed by the conclusions in Section VIII.

*Notations:* Boldface lowercase letters and boldface uppercase letters denote vectors and matrices, respectively. $\mathrm{diag}(\cdot)$ forms a diagonal matrix out of its vector argument. $(\cdot)^T$, $(\cdot)^H$, $(\cdot)^{-1}$ and $(\cdot)^*$ denote the transpose, conjugate transpose, inverse, and conjugate, respectively. $\mathbb{C}^{m\times n}$ is the complex space of a $m\times n$ matrix. $\mathbf{I}_N$ denotes the $N\times N$ identity matrix. $\mathcal{CN}(\mu,\sigma^2)$ denotes the circularly symmetric complex Gaussian distribution with mean $\mu$ and variance $\sigma^2$. $\mathbb{E}\{\cdot\}$ denotes the statistical expectation. $I(\cdot)$ denotes mutual information. $h(\cdot)$ is the differenctial entropy. The matrix with all elements equal to zero is denoted as $\mathbf{0}$. Table I summarizes the notations used throughout the paper.

TABLE I
SYSTEM PARAMETERS

| Notation | Definition |
|---|---|
| $N$ | Number of antennas at the base station |
| $K$ | Number of user equipments |
| $r_{b_k}, r_{b_k,l}$ | Distance from the $k$-th UE and its $l$-th path to the reference antenna |
| $\psi_{b_k}$, $\psi_{b_k,l}$ | Azimuth angle of the $k$-th UE and its $l$-th propagation path |
| $\beta_{b_k,l}$ | Complex path gain of the $l$-th path of the $k$-th UE |
| $L_k$ | Number of multipath components for the $k$-th UE |
| $\mathbf{P}_k$ | Precoding matrix for the $k$-th UE |
| $d_R$ | Rayleigh distance |
| $\mathbf{h}_k$ | Near-field channel vector of the $k$-th UE |
| $\mathbf{b}\left(\psi_{b_k,l}, r_{b_k,l}\right)$ | Array response vector |
| $\mathbf{c}(\psi_{b_k,l}, r_{b_k,l})$ | Gain correlation vector |
| $\mathbf{R}_k$ | Channel covariance matrix of the $k$-th UE |
| $\mathbf{S}_d, \mathbf{s}_u$ | Downlink and uplink packets |
| $L$, $Q$ | Lengths of downlink and uplink packets |
| $N_s$ | Symbol vector length |
| $\sigma_b^2$, $\sigma_a^2$ | Noise power at the UE and the BS, respectively |
| $\widehat{\sigma}_b^2$, $\widehat{\sigma}_a^2$ | Estimated noise power at the UE and the BS |
| $P_b$, $P_a$ | Transmit power at the UE and the BS |
| $\mathbf{B}_k$ | Aggregated channel covaraince matrix of the $k$-th UE |
| $\widetilde{\mathbf{V}}_k^{(0)}$ | Interferernce mitigation matrix of the $k$-th UE |
| $D_k$ | Number of extracted subchannels for the $k$-th UE |
| $\mathbf{P}_{e,k}$ | Effective precoding matrix of the $k$-th UE |
| $\mathbf{W}_k$ | Subchannel extraction matrix of the $k$-th UE |
| $\mathbf{R}_{hk}$ | Channel covariance matrix after applying $\widetilde{\mathbf{V}}_k^{(0)}$ |
| $\mathbf{U}_{hk}$ | Channel correlation matrix of the $k$-th UE |
| $\mathbf{\Lambda}_{hk}$ | Channel eigenvalue matrix of the $k$-th UE |
| $\mathbf{P}_{uk}$ | Uplink precoding matrix of the $k$-th UE |
| $\mathbf{U}_k$, $\mathbf{V}_k$ | Direction control matrices of the $k$-th UE |
| $\mathbf{\Lambda}_k$ | Power allocation matrix of the $k$-th UE |
| $\lambda_{hk,j}$ | Channel eigenvalue |
| $\lambda_{k,j}$ | Power allocation variable |
| $x_{k,j}$ | Power allocation variable normalized by $\lambda_{hk,j}$ |
| $I_k$ | KR of the $k$-th UE |
| $C_k$ | Upper bound on the KR of the $k$-th UE |
| $\mu$ | Lagrange multiplier for sum KR maximization |
| $\mu_k$ | Lagrange multiplier for KR fairness of the $k$-th UE |
| $t$ | Slack variable introduced for KR fairness optimization |

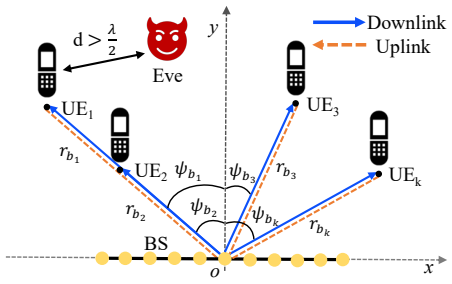



Fig. 1. Channel model.

## II. SYSTEM MODEL

### A. System Overview

Figure 1 presents the setup for a multi-user key generation system in near-field scenarios, where a BS is equipped with a $N$-antenna ELAA and $K$ pieces of user equipment (UE) are fitted with a single antenna. The BS controls the precoding matrices for generating secret keys with $K$ UEs simultaneously, which will be elaborated in Section IV. Two power allocation strategies are investigated considering maximizing the sum KR and ensuring KR fairness in Section V and VI, respectively. Since eavesdroppers locate half wavelength from the BS and UEs, the channels of eavesdroppers are uncorrelated with that of BS and UEs [12].

As illustrated in Fig. 2, the multi-user key generation protocol consists of four steps: channel probing, quantization, information reconciliation, and privacy amplification. In the channel probing, the UEs and BS transmit pilots to each other and estimate the reciprocal channels. In the quantization, the UEs and BS transfer the channel measurements into binary sequences. Information reconciliation algorithms wipe off the discrepancies between the quantized sequences due to the noise. In the privacy amplification, algorithms are leveraged to remove the key bits leaked to eavesdroppers. Finally, the $k$-th UE and the BS agree on a secret key that cannot be obtained by other $K-1$ UEs. Our paper concentrates on the channel probing step, which will be elaborated in Section III.

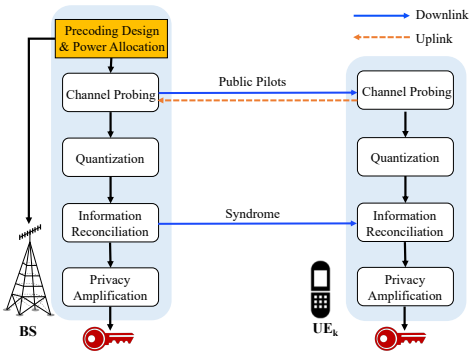



Fig. 2. System model.

### B. Device Configuration

Figure 1 presents a two-dimensional coordinate system, where the BS is situated along the $x$-axis and UEs are deployed on the positive side of the $y$-axis. The BS is modelled as a uniform linear array (ULA) with $N$-antennas. The coordinate of the reference antenna is $(0,0)$ and the coordinate of the $n$-th antenna is $(\delta_n d, 0)$, where $d$ is the antenna spacing and $\delta_n = n - \frac{N-1}{2}$ for $n = 0, 1, \ldots, N-1$. The coordinate of the $k$-th UE is $\left(r_{b_k}\sin\psi_{b_k}, r_{b_k}\sqrt{1-\sin^2\psi_{b_k}}\right)$, where $r_{b_k}$ is the distance from the $k$-th UE to the reference antenna and $\psi_{b_k}$ is the azimuth angle. The BS controls $K$ precoding matrices, $\{\mathbf{P}_k | \mathbf{P}_k \in \mathbb{C}^{N\times N_s}, k = 1, \ldots, K\}$, for channel probing with the $k$-th UE, where $N_s$ is the symbol vector length.

### C. Near-Field Channel Model

The radiation field of electromagnetic waves is divided into two regions: the radiative near-field region and the far-field region [16]. The Rayleigh distance $d_R = \frac{2D^2}{\lambda}$ is commonly used to distinguish between these two regions [17], where $\lambda$ is the wavelength and $D$ represents the array aperture. When the array aperture of a ULA is defined as $D = (N-1)d$ and $d = \lambda/2$, the Rayleigh distance of a ULA is $d_R = \frac{1}{2}(N-1)^2\lambda$. For example, at a carrier frequency of 28 GHz and with 100 antennas, any UE within a distance of 52.5 meters from the BS is considered to be in its near-field region.

If a UE is located in the near-field region of this array, the channel has to be modelled as a near-field channel with spherical-wave assumptions. According to [18], [19], the near-field channel[1] from the $k$-th UE to BS is modelled as

$$\mathbf{h}_k = \sqrt{\frac{N}{L_k}} \sum_{l=1}^{L_k} \beta_{b_k,l}\mathbf{c}\left(\psi_{b_k,l}, r_{b_k,l}\right) \odot \mathbf{b}(\psi_{b_k,l}, r_{b_k,l}), \quad (1)$$

where $\mathbf{h}_k \in \mathbb{C}^{N\times 1}$, $\beta_{b_k,l}$ is the channel gain of the $l$-th path, $\psi_{b_k,l}$ denotes the azimuth angle of the $l$-th scatter, $r_{b_k,l}$ is the distance from the $l$-th scatter to the reference antenna of the ULA, and $L_k$ is the number of scatters.

[1]Our paper considers a stationary near-field channel model, where the entire antenna array is assumed to be visible to the UE, consistent with the approach in [8]. In contrast, non-stationary near-field channels account for the fact that different portions of an ELAA may be either visible or invisible to a given UE, resulting in a fundamentally different channel modeling [17].

The near-field channel is composed of array response vectors and gain correlation vectors. The array response vector represents the actual phase variations between all receive antennas and the reference antenna, which is given by

$$\mathbf{b}\left(\psi_{b_k,l}, r_{b_k,l}\right) = \frac{1}{\sqrt{N}}\Big[e^{-j\frac{2\pi}{\lambda}\left(r^{(0)}_{b_k,l} - r_{b_k,l}\right)}, \ldots, e^{-j\frac{2\pi}{\lambda}\left(r^{(n)}_{b_k,l} - r_{b_k,l}\right)}, \ldots, e^{-j\frac{2\pi}{\lambda}\left(r^{(N-1)}_{b_k,l} - r_{b_k,l}\right)}\Big]^T, \quad (2)$$

where $r^{(n)}_{b_k,l} = \sqrt{r^2_{b_k,l} + d^2\delta_n^2 - 2r_{b_k,l}\sin\psi_{b_k,l}d\delta_n}$ is the distance from the $l$-th scatter to the $n$-th antenna and $r_{b_k,l}$ denotes the distance from the $l$-th scatter to the reference antenna. The gain correlation vector $\mathbf{c}(\psi_{b_k,l}, r_{b_k,l})$ describes the actual amplitude variations between all receive antennas and the reference antenna, which is given by

$$\mathbf{c}(\psi_{b_k,l}, r_{b_k,l}) = \left[\frac{r_{b_k,l}}{r^{(0)}_{b_k,l}}, \frac{r_{b_k,l}}{r^{(1)}_{b_k,l}}, \ldots, \frac{r_{b_k,l}}{r^{(N-1)}_{b_k,l}}\right]^T. \quad (3)$$

### D. Correlation Modeling

In our work, the near-field channel $\mathbf{h}_k$ is modeled as a complex Gaussian random vector, which is a commonly adopted assumption in rich-scattering propagation environments. As shown in [8], the near-field channel follows the distribution $\mathbf{h}_k \sim \mathcal{CN}(\mathbf{0}, \mathbf{R}_k)$, where $\mathbf{R}_k \in \mathbb{C}^{N\times N}$ denotes the channel covariance matrix (CCM) that characterizes the long-term spatial correlation among the $N$ antennas at the BS.

The CCM of the near-field channel $\mathbf{h}_k$ is expressed as

$$\mathbf{R}_k = \mathbb{E}\left\{\mathbf{h}_k\mathbf{h}_k^H\right\} \in \mathbb{C}^{N\times N}. \quad (4)$$

Let $h_{k,n}$ denote the $n$-th element of $\mathbf{h}_k$. The spatial correlation between the $n$-th and $m$-th antennas is given by $[\mathbf{R}_k]_{n,m} = \mathbb{E}\{h_{k,n}h^*_{k,m}\}$, which can be expressed as

$$[\mathbf{R}_k]_{n,m} = \frac{1}{L_k}\sum_{q=1}^{L_k} \frac{r^2_{b_k,q}\mathbb{E}\{|\beta_{b_k,q}|^2\}e^{-j\frac{2\pi}{\lambda}\left(r^{(n)}_{b_k,q} - r^{(m)}_{b_k,q}\right)}}{r^{(n)}_{b_k,q}r^{(m)}_{b_k,q}}. \quad (5)$$

The near-field CCM in (5) is determined by the scatterer distribution in terms of both angular and distance domains. The angles and distances may follow different distributions depending on the propagation environment. However, obtaining a closed-form expression for the resulting CCM is generally intractable. For detailed analysis, we refer the reader to [18]. Here, we assume that the channel vector $\mathbf{h}_k$ follows a complex Gaussian distribution. The CCM $\mathbf{R}_k$ is numerically estimated as $\mathbf{R}_k = \mathbb{E}\{\mathbf{h}_k\mathbf{h}_k^H\}$ via Monte Carlo simulations using a large number of channel realizations.

## III. Multi-User Channel Probing in Near-Field Communications

To extract secrets keys from the near-field channel $\mathbf{h}_k$ in (1), the $k$-th UE follows a ping-pong pattern by receiving a downlink packet from the BS and subsequently transmitting an uplink packet to the BS to measure $\mathbf{h}_k$. If the $K$ UEs probe their channels sequentially, the pilot overhead increases linearly with the number of UEs. In this section, a multi-user channel probing scheme is proposed to reduce pilot

overhead. That is, the $K$ UEs simultaneously probe their uplink channels, while the BS probes the downlink channels in parallel. Therefore, the pilot length does not increase with the number of users.

The key challenge in multi-user channel probing is the interference caused by multiple users sharing the same uplink and downlink packets. The main approach is to design precoding matrices that mitigate interference from other UEs. The design of these precoding matrices, $\{\mathbf{P}_k\}$, will be discussed in detail in Section IV. This section focuses on the multi-user channel probing process, followed by the derivation of the analytical expression for the KR based on the measurements.

### A. Downlink Multi-User Channel Probing

The BS broadcasts the same downlink packet $\mathbf{S}_d \in \mathbb{C}^{L\times N_s}$ to $K$ UEs, where $\mathbf{S}_d^H\mathbf{S}_d = L\mathbf{I}_{N_s}$ and $L$ is the length of the downlink packet. Specifically, $L$ should satisfy $L \geq N_s$ to ensure that the $k$-th UE can estimate a measurement vector with $N_s$ dimensions. To enable the $k$-th UE to extract secret keys from its measurement, the downlink packet is multiplied by the respective precoding matrix $\mathbf{P}_k$. Thus, the sum of precoded downlink packets is represented as $\sum_{k'=1}^{K}\mathbf{P}_{k'}\mathbf{S}_d^H$.

When the sum of precoded downlink packets is broadcasted to $K$ UEs, the $k$-th UE gets the received signal as follows:

$$\mathbf{y}_{b_k} = \mathbf{h}_k^H\mathbf{P}_k\mathbf{S}_d^H + \underbrace{\mathbf{h}_k^H\sum\nolimits_{k'\neq k}\mathbf{P}_{k'}\mathbf{S}_d^H}_{\text{interference}} + \mathbf{n}_{b_k}, \tag{6}$$

where $\mathbf{y}_{b_k} \in \mathbb{C}^{1\times L}$, $\mathbf{n}_{b_k} \in \mathbb{C}^{1\times L}$ is the noise vector, $\mathbf{n}_{b_k} \sim \mathcal{CN}\left(0, \sigma_b^2\mathbf{I}\right)$ and $\sigma_b^2$ is the noise power at the UE. As shown in (6), the interference from precoding matrices of other UEs will influence the $k$-th UE. Thus, the design of $\{\mathbf{P}_k\}$ should mitigate interference among UEs so that the $K$ UE can share the same downlink packet to relieve pilot overhead.

By the least squares (LS) estimation, the $k$-th UE measures its downlink near-field channel as

$$\begin{aligned}\bar{\mathbf{z}}_{b_k} &= \mathbf{h}_k^H\sum\nolimits_{k'=1}^{K}\mathbf{P}_{k'}\mathbf{S}_d^H\mathbf{S}_d\left(\mathbf{S}_d^H\mathbf{S}_d\right)^{-1} + \mathbf{n}_{b_k}\mathbf{S}_d\left(\mathbf{S}_d^H\mathbf{S}_d\right)^{-1}\\ &= \mathbf{h}_k^H\mathbf{P}_k + \mathbf{h}_k^H\sum\nolimits_{k'\neq k}\mathbf{P}_{k'} + \bar{\mathbf{n}}_{b_k},\end{aligned} \tag{7}$$

where $\bar{\mathbf{z}}_{b_k} \in \mathbb{C}^{1\times N_s}$, $\bar{\mathbf{n}}_{b_k} \in \mathbb{C}^{1\times N_s}$ is the noise vector after LS estimation, $\bar{\mathbf{n}}_{b_k} = \frac{1}{\sqrt{L}}\mathbf{n}_{b_k}\tilde{\mathbf{S}}_d$, and $\tilde{\mathbf{S}}_d = \frac{1}{\sqrt{L}}\mathbf{S}_d$ is the normalized result of $\mathbf{S}_d$ with $\tilde{\mathbf{S}}_d^H\tilde{\mathbf{S}}_d = \mathbf{I}_{N_s}$.

We compute the Hermitian transpose of the above equation, which is given by

$$\mathbf{z}_{b_k} = \mathbf{P}_k^H\mathbf{h}_k + \sum\nolimits_{k'\neq k}\mathbf{P}_{k'}^H\mathbf{h}_k + \widehat{\mathbf{n}}_{b_k}, \tag{8}$$

where $\mathbf{z}_{b_k} \in \mathbb{C}^{N_s\times 1}$, $\widehat{\mathbf{n}}_{b_k} = \bar{\mathbf{n}}_{b_k}^H$, $\widehat{\mathbf{n}}_{b_k} \sim \mathcal{CN}\left(0, \widehat{\sigma}_b^2\mathbf{I}_{N_s}\right)$, and $\widehat{\sigma}_b^2 = \frac{\sigma_b^2}{L}$ is the estimation noise power at the UE.

### B. Uplink Multi-User Channel Probing

UEs send uplink packets, $\mathbf{s}_u = [s_1, \ldots, s_Q]^T \in \mathbb{C}^{Q\times 1}$, where $Q$ is the length of the uplink packets, $\mathbf{s}_u^H\mathbf{s}_u = QP_b$ and $P_b$ is the uplink transmit power. The BS receives the signal as

$$\mathbf{Y}_a = \sum\nolimits_{k=1}^{K}\mathbf{h}_k\mathbf{s}_u^H + \mathbf{N}_a, \tag{9}$$

where $\mathbf{Y}_a \in \mathbb{C}^{N\times Q}$ and $\mathbf{N}_a \in \mathbb{C}^{N\times Q}$ is the uplink noise matrix. Each entry of $\mathbf{N}_a$ is a Gaussian noise with power $\sigma_a^2$.

By the LS channel estimation, the BS measures the combined channels of $K$ UEs, which is given by

$$\bar{\mathbf{z}}_a = \mathbf{Y}_a\mathbf{s}_u\left(\mathbf{s}_u^H\mathbf{s}_u\right)^{-1} = \sum\nolimits_{k=1}^{K}\mathbf{h}_k + \bar{\mathbf{n}}_a, \tag{10}$$

where $\bar{\mathbf{z}}_a \in \mathbb{C}^{N\times 1}$, $\bar{\mathbf{n}}_a = \mathbf{N}_a\mathbf{s}_u\left(\mathbf{s}_u^H\mathbf{s}_u\right)^{-1} \in \mathbb{C}^{N\times 1}$ is the noise vector after LS estimation and $\bar{\mathbf{n}}_a \sim \mathcal{CN}\left(0, \frac{\sigma_a^2}{QP_b}\mathbf{I}_N\right)$.

To acquire the measurement of the $k$-th UE, BS applies the precoding matrix $\mathbf{P}_k$ to $\bar{\mathbf{z}}_a$, which is given by

$$\mathbf{z}_{a_k} = \mathbf{P}_k^H\sum\nolimits_{k=1}^{K}\mathbf{h}_k + \widehat{\mathbf{n}}_{a_k}, \tag{11}$$

where $\mathbf{z}_{a_k} \in \mathbb{C}^{N_s\times 1}$, $\widehat{\mathbf{n}}_{a_k} = \mathbf{P}_k^H\bar{\mathbf{n}}_a \sim \mathcal{CN}\left(0, \widehat{\sigma}_a^2\mathbf{P}_k^H\mathbf{P}_k\right)$ is the estimation noise after precoding and $\widehat{\sigma}_a^2 = \frac{\sigma_a^2}{QP_b}$ is the estimation noise power at the BS.

The uplink pilot length $Q$ can be set to any value satisfying $Q \geq 1$. Regardless of $Q$, the BS, equipped with $N$ antennas, can estimate $N$ channel coefficients during uplink channel probing. By applying the precoding matrix $\mathbf{P}_k \in \mathbb{C}^{N\times N_s}$ to the measurement vector $\bar{\mathbf{z}}_a \in \mathbb{C}^{N\times 1}$, the BS obtains an $N_s$-dimensional precoded measurements $\mathbf{z}_{a_k} \in \mathbb{C}^{N_s\times 1}$.

### C. Key Rate

The KR denotes the maximum number of secret keys that can be extracted from the near-field channel $\mathbf{h}_k$ by using the measurements of the BS and the $k$-th UE, i.e., $\mathbf{z}_{a_k}$ and $\mathbf{z}_{b_k}$. When eavesdroppers are positioned half a wavelength away from the UEs, their channels become uncorrelated with those of the UEs. As a result, the KR is equivalent to the mutual information between $\mathbf{z}_{a_k}$ and $\mathbf{z}_{b_k}$ [20], i.e., $I(\mathbf{z}_{a_k}; \mathbf{z}_{b_k})$. The half-wavelength condition for eavesdropping conditions has been investigated from both theoretical and experimental aspects. Theoretical analysis and simulations using various channel correlation models have confirmed the validity of the half-wavelength assumption in environments with rich scattering [21]. Additionally, eavesdropping threats were investigated using testbeds with IEEE 802.11 PHY protocol [22]. The results showed that, in environments with strong multipath propagation, eavesdroppers' channels exhibit minimal correlation with those of legitimate users.

Since $\mathbf{z}_{a_k}$ and $\mathbf{z}_{b_k}$ follow a Gaussian distribution, the mutual information $I(\mathbf{z}_{a_k}; \mathbf{z}_{b_k})$ is calculated using the differential entropy of Gaussian random variables. Thus, we define $\mathbf{R}_{a_k} = \mathbb{E}\{\mathbf{z}_{a_k}\mathbf{z}_{a_k}^H\}$ and $\mathbf{R}_{b_k} = \mathbb{E}\{\mathbf{z}_{b_k}\mathbf{z}_{b_k}^H\}$ as the covariance matrices of $\mathbf{z}_{a_k}$ and $\mathbf{z}_{b_k}$, respectively. Besides, we define $\mathbf{R}_{a_kb_k} = \mathbb{E}\{\mathbf{z}_{a_k}\mathbf{z}_{b_k}^H\}$ and $\mathbf{R}_{b_ka_k} = \mathbb{E}\{\mathbf{z}_{b_k}\mathbf{z}_{a_k}^H\}$ as the cross-covariance matrices of $\mathbf{z}_{a_k}$ and $\mathbf{z}_{b_k}$.

When the channels of different UEs become independent, we substitute $\mathbf{z}_{a_k}$ in (11) and $\mathbf{z}_{b_k}$ in (8) to the covariance and

cross-covariance matrices and calculate them as

$$\begin{aligned}\mathbf{R}_{a_k} &= \mathbf{P}_k^H \left(\sum\nolimits_{k'} \mathbf{R}_{k'}\right) \mathbf{P}_k + \widehat{\sigma}_a^2 \mathbf{P}_k^H \mathbf{P}_k,\\ \mathbf{R}_{b_k} &= \left(\sum\nolimits_{k'} \mathbf{P}_{k'}^H\right) \mathbf{R}_k \left(\sum\nolimits_{k'} \mathbf{P}_{k'}\right) + \widehat{\sigma}_b^2 \mathbf{I}_{N_s},\\ \mathbf{R}_{a_k b_k} &= \mathbf{P}_k^H \mathbf{R}_k \left(\sum\nolimits_{k'} \mathbf{P}_{k'}\right),\\ \mathbf{R}_{b_k a_k} &= \left(\sum\nolimits_{k'} \mathbf{P}_{k'}^H\right) \mathbf{R}_k \mathbf{P}_k. \end{aligned} \tag{12}$$

According to the differential entropy of Gaussian random variables [20], the KR of the $k$-th UE is expressed as

$$\begin{aligned} I_k &= I(\mathbf{z}_{a_k}; \mathbf{z}_{b_k}) = h(\mathbf{z}_{a_k}) + h(\mathbf{z}_{b_k}) - h(\mathbf{z}_{a_k}, \mathbf{z}_{b_k}) \\ &= -\log_2 \left( \left| \mathbf{I}_{N_s} - \mathbf{R}_{a_k b_k} \mathbf{R}_{b_k}^{-1} \mathbf{R}_{b_k a_k} \mathbf{R}_{a_k}^{-1} \right| \right). \end{aligned} \tag{13}$$

By substituting (12) to (13), we derive the KR in terms of $\{\mathbf{P}_k\}$, which is shown in (14) at the top of the next page. The KR is affected by precoding matrices $\{\mathbf{P}_k\}$. Next, we will discuss the design of $\{\mathbf{P}_k\}$ to ensure that interference from other UEs does not influence the KR. Besides, the KR depends on the CCMs $\{\mathbf{R}_k\}$ in (4) and this information is utilized in precoding design. The method for estimating the CCM of the near-field channel is provided in [23].

## IV. Precoding Design for Interference Mitigation and Uncorrelated Subchannels Extraction

The multi-user PLKG scheme reduces the pilot overhead by allowing $K$ UEs to simultaneously probe $\mathbf{h}_k$ in the downlink while the BS concurrently probes $\mathbf{h}_k$ in the uplink, resulting in interference between the UEs. To address this, the design of the $k$-th precoding matrix $\mathbf{P}_k$ should minimize the interference from other UEs to the KR of the $k$-th UE in (14). Besides, since the dimension of the near-field channel $\mathbf{h}_k$ is $N$, both the BS and the $k$-th UE can probe $\mathbf{h}_k$ to obtain a measurement of dimension $N$ for generating secret keys. However, as noted in [8], there is a spatial correlation between channel coefficients from antennas of the ELAA. Thus, the $\mathbf{P}_k$ should also reduce the correlation between measurements. Next, we apply the SVD and EVD methods to design $\mathbf{P}_k$ for interference mitigation and to convert the channels into uncorrelated subchannels. With this precoding design, the analytical expression for the KR in (14) is simplified, enabling further optimization of $\mathbf{P}_k$ to enhance the KR.

Note that the precoding design process, which involves EVD and SVD operations as well as key rate optimization, is executed solely at the BS, which possesses sufficient computational resources. On the user side, the procedure remains lightweight, involving only LS channel estimation.

### A. Precoding Design to Mitigate Interference

The near-field channels propagate from the $k$-UE to the BS through paths with specific angles and distances. Due to the different coordinates of the UEs, the angle and distance ranges of paths differ for each UE. The CCM of the $k$-th UEs $\mathbf{R}_k$ in (4) can be decomposed to a channel subspace of rank $n_k < N$ to capture the spatial correlation of $\mathbf{h}_k$ caused by the specific range of angles and distances. Therefore, we can find a matrix $\mathbf{L}_k$ with rank $(N - n_k)$ such that $\mathbf{R}_{k'}\mathbf{L}_k = 0$. The $\mathbf{L}_k$ is referred to as the null space of $\mathbf{R}_k$. If the precoding matrix of the $k$-th UE $\mathbf{P}_k$ aligns with the null space of the CCM of the $k'$-th UE, we have

$$\mathbf{R}_{k'}\mathbf{P}_k = 0, \text{ for } k' \neq k. \tag{15}$$

Based on (15), interference between UEs can be mitigated. In near-field scenarios, each user's channel exhibits a richer spatial structure due to variations in both angles and distances. This spatial diversity reduces the overlap between the subspaces of their CCMs. Consequently, it becomes feasible to design precoding matrices that align with the subspace of the intended user's channel while simultaneously residing in the null space of the CCMs of unintended users, thereby minimizing interference to other users. In far-field environments, the proposed method remains applicable if users experience paths with distinct angles, resulting in non-overlapping CCM subspaces.

*1) Performing SVD on CCMs of Other UEs to Find Null Space:* Based on [24], we apply the block diagonalization method to find the null space of $\mathbf{R}_{k'}$ for $k' \neq k$. We define the aggregated CCM as $\mathbf{B}_k = [\mathbf{R}_1; \ldots; \mathbf{R}_{k-1}; \mathbf{R}_{k+1}; \ldots; \mathbf{R}_K] \in \mathbb{C}^{N(K-1)\times N}$. Based on SVD, $\mathbf{B}_k$ is transferred as

$$\mathbf{B}_k = \widetilde{\mathbf{U}}_k \begin{bmatrix} \widetilde{\mathbf{\Lambda}}_k & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} [\widetilde{\mathbf{V}}_k^{(1)} \ \widetilde{\mathbf{V}}_k^{(0)}]^H, \tag{16}$$

where $\widetilde{\mathbf{U}}_k$ is a $N(K-1) \times N(K-1)$ unitary matrix, $\widetilde{\mathbf{\Lambda}}_k$ is a $r_k \times r_k$ diagonal matrix with non-negative singular values, $\widetilde{\mathbf{V}}_k^{(1)} \in \mathbb{C}^{N\times r_k}$ and $\widetilde{\mathbf{V}}_k^{(0)} \in \mathbb{C}^{N\times D_k}$ contains vectors corresponding to the non-zero and zero singular values, respectively. $D_k = N - r_k$ is the dimension of the null space. $\widetilde{\mathbf{V}}_k^{(0)}$ constitutes an orthogonal basis for the null space of $\mathbf{B}_k$.

*2) Align the Precoding Matrix with the Null Space:* When we acquire $\widetilde{\mathbf{V}}_k^{(0)}$ from (16), the precoding matrix $\mathbf{P}_k$ can be constructed as $[\mathbf{P}_{e,k}, \mathbf{0}_k]$, where $\mathbf{P}_{e,k} \in \mathbb{C}^{N\times D_k}$ is the effective precoding matrix of the $k$-th UE, and $\mathbf{0}_k$ is a zero matrix of size $N \times (N_s - D_k)$. Specifically, $\mathbf{P}_{e,k}$ consists of the first $D_k$ columns of $\mathbf{P}_k$, and is designed as $\mathbf{P}_{e,k} = \widetilde{\mathbf{V}}_k^{(0)}\mathbf{W}_k$ to ensure its alignment with the null space of the CCMs of other UEs. Here, $\widetilde{\mathbf{V}}_k^{(0)}$ serves to mitigate inter-user interference, while $\mathbf{W}_k \in \mathbb{C}^{D_k\times D_k}$ further refines the precoding to extract uncorrelated subchannels. The design of $\mathbf{W}_k$ will be discussed in detail in Section IV-B.

Based on the precoding structure $\mathbf{P}_{e,k} = \widetilde{\mathbf{V}}_k^{(0)}\mathbf{W}_k$, the requirement in (15) can be satisfied. The $k$-th UE selects the first $D_k$ dimensions of the measurements for key generation. As a result, the covariance and cross-covariance matrices in (12) can be rewritten using the reduced $D_k$-dimensional observations instead of the full $N_s$-dimensional ones. Consequently, the KR in (13) can be simplified as

$$\begin{aligned} I_k = -\log_2 \Big( \Big| \mathbf{I}_{D_k} - \mathbf{P}_{e,k}^H \mathbf{R}_k \mathbf{P}_{e,k} \left(\mathbf{P}_{e,k}^H \mathbf{R}_k \mathbf{P}_{e,k} + \widehat{\sigma}_b^2 \mathbf{I}_{D_k}\right)^{-1} \\ \times \mathbf{P}_{e,k}^H \mathbf{R}_k \mathbf{P}_{e,k} \left(\mathbf{P}_{e,k}^H \mathbf{R}_k \mathbf{P}_{e,k} + \widehat{\sigma}_a^2 \mathbf{P}_{e,k}^H \mathbf{P}_{e,k}\right)^{-1} \Big| \Big). \end{aligned} \tag{17}$$

$$I_k = -\log_2\left(\left|\mathbf{I}_{N_s} - \mathbf{P}_k^H\mathbf{R}_k\left(\sum\nolimits_{k'}\mathbf{P}_{k'}\right)\times\left(\left(\sum\nolimits_{k'}\mathbf{P}_{k'}^H\right)\mathbf{R}_k\left(\sum\nolimits_{k'}\mathbf{P}_{k'}\right)+\widehat{\sigma}_b^2\mathbf{I}_{N_s}\right)^{-1}\times\left(\sum\nolimits_{k'}\mathbf{P}_{k'}\right)\mathbf{R}_k\mathbf{P}_k^H\right.$$
$$\left.\times\left(\mathbf{P}_k^H\left(\sum\nolimits_{k'}\mathbf{R}_{k'}\right)\mathbf{P}_k+\widehat{\sigma}_a^2\mathbf{P}_k^H\mathbf{P}_k\right)^{-1}\right|\right). \quad (14)$$

Furthermore, we express the KR in terms of $\mathbf{W}_k$ as

$$I_k = -\log_2\left(\left|\mathbf{I}_{D_k} - \mathbf{W}_k^H\mathbf{R}_{hk}\mathbf{W}_k\left(\mathbf{W}_k^H\mathbf{R}_{hk}\mathbf{W}_k+\widehat{\sigma}_b^2\mathbf{I}_{D_k}\right)^{-1}\right.\right.$$
$$\left.\left.\times\,\mathbf{W}_k^H\mathbf{R}_{hk}\mathbf{W}_k\left(\mathbf{W}_k^H\mathbf{R}_{hk}\mathbf{W}_k+\widehat{\sigma}_a^2\mathbf{W}_k^H\mathbf{W}_k\right)^{-1}\right|\right), \quad (18)$$

where $\mathbf{R}_{hk}\in\mathbb{C}^{D_k\times D_k}$ represents the covariance matrix of channels after applying $\widetilde{\mathbf{V}}_k^{(0)}$, which is given by

$$\mathbf{R}_{hk} = \widetilde{\mathbf{V}}_k^{(0)H}\mathbf{R}_k\widetilde{\mathbf{V}}_k^{(0)}. \quad (19)$$

Since the dimension of subchannels for the $k$-th UE is $D_k$, the symbol length $N_s$ must satisfy $N_s \geq \max_k\{D_k\}$ to ensure that both the $k$-th UE and the BS can extract $D_k$ subchannels during downlink and uplink channel probing, respectively. In particular, since each UE has a single antenna, the downlink packet length $L$ must satisfy $L \geq N_s \geq \max_k\{D_k\}$ to enable the $k$-th UE to obtain a measurement of length $N_s$ and extract $D_k$ subchannels.

### *B. Precoding Design to Extract Uncorrelated Subchannels*

Given $\mathbf{R}_{hk}$ in (19), the KR in (18) depends on $\mathbf{W}_k$. Since there exists spatial correlation, $\mathbf{W}_k$ is applied to convert channels to uncorrelated subchannels. The KR with respect to $\mathbf{W}_k$ in (18) is non-convex due to the presence of matrix inverses and determinants, which is challenging to optimise. To simplify the KR, we gradually decompose $\mathbf{R}_{hk}$ and $\mathbf{W}_k$.

*1) Obtaining the Channel Correlation Matrix and Channel Eigenvalue Matrix:* $\mathbf{R}_{hk}$ in (19) consists of the channel correlation matrix and channel eigenvalue matrix. To find these two matrices, we do the EVD to $\mathbf{R}_{hk}$ as

$$\mathbf{R}_{hk} = \mathbf{R}_{hk}^{1/2}\mathbf{R}_{hk}^{H/2} = \mathbf{U}_{hk}\mathbf{\Lambda}_{hk}\mathbf{U}_{hk}^H, \quad (20)$$

where $\mathbf{R}_{hk}^{H/2} = \mathbf{\Lambda}_{hk}^{1/2}\mathbf{U}_{hk}^H$. $\mathbf{\Lambda}_{hk} = \text{diag}([\lambda_{hk,1},\ldots,\lambda_{hk,D_k}]) \in \mathbb{C}^{D_k\times D_k}$ is the channel eigenvalue matrix, where $\lambda_{hk,j}$ is the channel eigenvalue of the $j$-th subchannel, $\sum_{j=1}^{D_k}\lambda_{hk,j} = P_{hk}D_k$ and $P_{hk}$ is the channel variance of the $k$-th UE. $\mathbf{U}_{hk}\in\mathbb{C}^{D_k\times D_k}$ is the channel correlation matrix, where the $j$-th column describes the direction of the $j$-th subchannel.

*2) Obtaining the Spatial Decorrelation Matrix and Eigenvalue Normalization Matrix:* As shown in (18), we can design $\mathbf{W}_k$ to control $\mathbf{W}_k^H\mathbf{R}_{hk}\mathbf{W}_k$ to increase the KR. Since $\mathbf{R}_{hk}$ can be decomposed as $\mathbf{R}_{hk}^{1/2}\mathbf{R}_{hk}^{H/2}$ in (20), we have

$$\mathbf{W}_k^H\mathbf{R}_{hk}\mathbf{W}_k = (\mathbf{R}_{hk}^{H/2}\mathbf{W}_k)^H(\mathbf{R}_{hk}^{H/2}\mathbf{W}_k) = \mathbf{G}_k^H\mathbf{G}_k, \quad (21)$$

where $\mathbf{G}_k = \mathbf{R}_{hk}^{H/2}\mathbf{W}_k \in \mathbb{C}^{D_k\times D_k}$ is a new matrix variable optimized to increase the KR. Since $\mathbf{R}_{hk}$ is decomposed to $\mathbf{U}_{hk}$ and $\mathbf{\Lambda}_{hk}$ in (20), we decompose $\mathbf{W}_k$ as

$$\mathbf{W}_k = (\mathbf{R}_{hk}^{H/2})^{-1}\mathbf{G}_k = \mathbf{U}_{hk}\mathbf{\Lambda}_{hk}^{-1/2}\mathbf{G}_k. \quad (22)$$

We design the $\mathbf{U}_{hk}$ to mitigate the spatial correlation to acquire uncorrelated subchannels. We use the $\mathbf{\Lambda}_{hk}^{-1/2}$ to normalize the channel eigenvalues. We can control the direction and power for estimating subchannels through $\mathbf{G}_k$.

*3) Obtaining the Direction Control Matrix and Power Allocation Matrix:* To allocate transmit power and direct the transmit power toward specific spatial directions to estimate subchannels, we resort to the following singular value decomposition to convert $\mathbf{G}_k$ as

$$\mathbf{G}_k = \mathbf{V}_k\mathbf{\Lambda}_k\mathbf{U}_k^H, \quad (23)$$

where $\mathbf{V}_k\in\mathbb{C}^{D_k\times D_k}$, $\mathbf{U}_k\in\mathbb{C}^{D_k\times D_k}$ and $\mathbf{\Lambda}_k\in\mathbb{C}^{D_k\times D_k}$. Substituting (23) to (22), we further decompose $\mathbf{W}_k$ as

$$\mathbf{W}_k = (\mathbf{R}_{hk}^{H/2})^{-1}\mathbf{V}_k\mathbf{\Lambda}_k\mathbf{U}_k^H = \mathbf{U}_{hk}\mathbf{\Lambda}_{hk}^{-1/2}\mathbf{V}_k\mathbf{\Lambda}_k\mathbf{U}_k^H. \quad (24)$$

After applying the spatial decorrelation matrix $\mathbf{U}_{hk}$ and eigenvalue normalization matrix $\mathbf{\Lambda}_{hk}^{-1/2}$ to channels, $\mathbf{V}_k$ and $\mathbf{U}_k$ controls the signal direction to estimate $D_k$ subchannels while $\mathbf{\Lambda}_k = \text{diag}\left(\left[\sqrt{\lambda_{k,1}},\ldots,\sqrt{\lambda_{k,D_k}}\right]\right)$ controls the power allocated on each direction. $\lambda_{k,j}$ is the transmit power allocated to estimate the $j$-th subchannel of the $k$-th UE.

When $\mathbf{W}_k$ is decomposed as (24), we have

$$\mathbf{W}_k^H\mathbf{R}_{hk}\mathbf{W}_k = \mathbf{U}_k\mathbf{\Lambda}_k\mathbf{V}_k^H\mathbf{V}_k\mathbf{\Lambda}_k\mathbf{U}_k^H \overset{(a)}{=} \mathbf{U}_k\mathbf{\Lambda}_k^2\mathbf{U}_k^H. \quad (25)$$

Equation (a) holds because $\mathbf{V}_k$ is unitary, i.e., $\mathbf{V}_k^H\mathbf{V}_k = \mathbf{I}_{D_k}$. Also, based on (24), we have

$$\mathbf{W}_k^H\mathbf{W}_k = \mathbf{U}_k\mathbf{\Lambda}_k\mathbf{V}_k^H\mathbf{\Lambda}_{hk}^{-1}\mathbf{V}_k\mathbf{\Lambda}_k\mathbf{U}_k^H. \quad (26)$$

Substituting (25) and (26) into (18), the KR is simplified as

$$I_k = -\log_2\left(\left|\mathbf{I}_{D_k} - \mathbf{U}_k\mathbf{\Lambda}_k^2\mathbf{U}_k^H\left(\mathbf{U}_k\mathbf{\Lambda}_k^2\mathbf{U}_k^H+\widehat{\sigma}_b^2\mathbf{I}_{D_k}\right)^{-1}\right.\right.$$
$$\left.\left.\mathbf{U}_k\mathbf{\Lambda}_k^2\mathbf{U}_k^H\left(\mathbf{U}_k\mathbf{\Lambda}_k^2\mathbf{U}_k^H+\underbrace{\widehat{\sigma}_a^2\mathbf{U}_k\mathbf{\Lambda}_k\mathbf{V}_k^H\mathbf{\Lambda}_{hk}^{-1}\mathbf{V}_k\mathbf{\Lambda}_k\mathbf{U}_k^H}_{\mathbf{G}_{n,a_k}}\right)^{-1}\right|\right). \quad (27)$$

According to interference mitigation matrix $\widetilde{\mathbf{V}}_k^{(0)}$ in (16) and subchannel extraction matrix $\mathbf{W}_k$ in (24), the effective precoding matrix $\mathbf{P}_{e,k} = \widetilde{\mathbf{V}}_k^{(0)}\mathbf{W}_k$ can be derived from

$$\mathbf{P}_{e,k} = \widetilde{\mathbf{V}}_k^{(0)}\mathbf{W}_k = \widetilde{\mathbf{V}}_k^{(0)}\mathbf{U}_{hk}\mathbf{\Lambda}_{hk}^{-1/2}\mathbf{V}_k\mathbf{\Lambda}_k\mathbf{U}_k^H. \quad (28)$$

### *C. Optimizing the Direction Control Matrix to Estimate Subchannels to Increase the Key Rate*

It is still hard to directly optimize the KR in (27) since it is affected by three matrix variables $\mathbf{V}_k$, $\mathbf{U}_k$ and $\mathbf{\Lambda}_k$. Besides, $\mathbf{\Lambda}_k$ affects the $\mathbf{G}_{n,a_k}$ of the KR in (27), which

means $\mathbf{\Lambda}_k$ influences the noise covariance matrix at the BS. This makes it hard to decouple the $\mathbf{V}_k$, $\mathbf{U}_k$, and $\mathbf{\Lambda}_k$ and separately optimize them to increase the KR. Therefore, we approximate the analytical expression of the KR in (27). The approximate expression can be separately optimized in terms of the direction control matrices $\mathbf{V}_k$ and $\mathbf{U}_k$ and power allocation matrix $\mathbf{\Lambda}_k$. We first optimize the $\mathbf{V}_k$ and $\mathbf{U}_k$ to increase the KR. The optimisation for $\mathbf{\Lambda}_k$ is elaborated in Section V and VI.

*1) Deriving the Approximation Expression of the Key Rate:* Since the uplink precoding is applied after LS estimation, it alters the noise covariance matrix at the BS, which in turn affects the KR. We next analyze how this transformation impacts the noise structure and the resulting KR expression. The BS measures the near-field channel in the uplink as $\bar{\mathbf{z}}_a$ as in Eq. (10), with the noise covariance matrix given by $\mathbf{R}_{n,a} = \widehat{\sigma}_a^2\mathbf{I}_N$. The $\mathbf{P}_k$ is then applied to the measurement, yielding $\mathbf{z}_{a_k} = \mathbf{P}_k^H\bar{\mathbf{z}}_a$. Since the subchannel dimension for the $k$-th UE is $D_k$, we extract the first $D_k$ entries of $\mathbf{z}_{a_k}$ as the effective measurement vector, denoted as $\mathbf{z}_{a_k,e} = \mathbf{P}_{e,k}^H\bar{\mathbf{z}}_a$. Accordingly, the $\mathbf{P}_{e,k}$ modifies the noise vector as $\mathbf{P}_{e,k}^H\bar{\mathbf{n}}_a$, leading to a noise covariance matrix observed at the BS after precoding $\mathbf{G}_{n,a_k} = \mathbf{P}_{e,k}^H\mathbf{R}_{n,a}\mathbf{P}_{e,k} = \widehat{\sigma}_a^2\mathbf{P}_{e,k}^H\mathbf{P}_{e,k}$. After applying the SVD to $\mathbf{P}_{e,k}$ as in Eq. (28), the KR expression in Eq. (27) can be simplified, where $\mathbf{G}_{n,a_k} = \widehat{\sigma}_a^2\mathbf{U}_k\mathbf{\Lambda}_k\mathbf{V}_k^H\mathbf{\Lambda}_{hk}^{-1}\mathbf{V}_k\mathbf{\Lambda}_k\mathbf{U}_k^H$. Therefore, the power allocation matrix $\mathbf{\Lambda}_k$ affects the noise covariance matrix at the BS after precoding. Consequently, the matrices $\mathbf{U}_k$, $\mathbf{\Lambda}_k$, and $\mathbf{V}_k$ appear in a coupled form inside matrix inverse operations, making the KR expression analytically intractable.

To simplify the KR in Eq. (27), we design the uplink precoding matrix as

$$\mathbf{P}_{u,k} = \widetilde{\mathbf{V}}_k^{(0)}\mathbf{U}_{hk}, \tag{29}$$

which effectively mitigates interference and extracts uncorrelated channel components. Notably, the simplified precoding matrix $\mathbf{P}_{u,k}$ used in the uplink channel probing phase does not include the direction control matrices $\mathbf{V}_k$, $\mathbf{U}_k$, or the power allocation matrix $\mathbf{\Lambda}_k$, due to an approximation in the objective function. These matrices, however, are incorporated in the downlink probing phase to optimize the KR.

To facilitate the key rate approximation, we compare the channel measurements obtained using the simplified uplink precoding matrix $\mathbf{P}_{u,k}$ with those obtained using $\mathbf{P}_{e,k}$. With $\mathbf{P}_{e,k}$, the effective uplink measurement at the BS is given by

$$\mathbf{z}_{a_k,e} = \mathbf{P}_{e,k}^H\mathbf{h}_k + \mathbf{P}_{e,k}^H\sum\nolimits_{k'\neq k}\mathbf{h}_{k'} + \mathbf{P}_{e,k}^H\bar{\mathbf{n}}_a, \tag{30}$$

and the corresponding downlink measurement at the UE is

$$\mathbf{z}_{b_k,e} = \mathbf{P}_{e,k}^H\mathbf{h}_k + \sum\nolimits_{k'\neq k}\mathbf{P}_{e,k'}^H\mathbf{h}_k + \widehat{\mathbf{n}}_{b_k,e}. \tag{31}$$

By replacing $\mathbf{P}_{e,k}$ with $\mathbf{P}_{u,k}$ in the uplink probing phase, the resulting approximate measurement becomes

$$\widetilde{\mathbf{z}}_{a_k} = \mathbf{P}_{u,k}^H\mathbf{h}_k + \mathbf{P}_{u,k}^H\sum\nolimits_{k'\neq k}\mathbf{h}_{k'} + \mathbf{P}_{u,k}^H\bar{\mathbf{n}}_a. \tag{32}$$

Combining Eq. (28) and Eq. (29), we obtain $\mathbf{P}_{e,k} = \mathbf{P}_{u,k}\mathbf{X}$, where $\mathbf{X} = \mathbf{\Lambda}_{hk}^{-1/2}\mathbf{V}_k\mathbf{\Lambda}_k\mathbf{U}_k^H$.

Based on the above analysis, we have $\mathbf{z}_{a_k,e} = \mathbf{X}^H\widetilde{\mathbf{z}}_{a_k}$ which is a function of $\widetilde{\mathbf{z}}_{a_k}$ without adding any new noise or randomness. In addition, both $\widetilde{\mathbf{z}}_{a_k}$ and $\mathbf{z}_{b_k,e}$ are noisy observations of the same channel $\mathbf{h}_k$. Therefore, once $\widetilde{\mathbf{z}}_{a_k}$ is known, $\mathbf{z}_{b_k,e}$ and $\mathbf{z}_{a_k,e}$ become conditionally independent, which establishes the Markov chain $\mathbf{z}_{b_k,e} \to \widetilde{\mathbf{z}}_{a_k} \to \mathbf{z}_{a_k,e}$ [15].

Based on Theorem 2 in [15], the data-processing inequality can be used to find the approximation of the KR, given by

$$C_k = I(\widetilde{\mathbf{z}}_{a_k};\mathbf{z}_{b_k,e}) \geq I(\mathbf{z}_{a_k,e};\mathbf{z}_{b_k,e}) = I_k. \tag{33}$$

The approximation of the KR $I(\widetilde{\mathbf{z}}_{a_k};\mathbf{z}_{b_k,e})$ is always greater than the KR in Eq. (27).

To calculate the KR, the covariance matrices of uplink and downlink measurements $\widetilde{\mathbf{z}}_{a_k}$, $\mathbf{z}_{b_k,e}$ can be derived from

$$\begin{aligned}
\widetilde{\mathbf{R}}_{a_k} &= \mathbf{P}_{u,k}^H\mathbf{R}_k\mathbf{P}_{u,k} + \widehat{\sigma}_a^2\mathbf{P}_{u,k}^H\mathbf{P}_{u,k}\\
&= \mathbf{U}_{hk}^H\mathbf{R}_{hk}\mathbf{U}_{kh} + \widehat{\sigma}_a^2\mathbf{U}_{hk}^H\mathbf{U}_{hk} = \mathbf{\Lambda}_{hk} + \widehat{\sigma}_a^2\mathbf{I}_{D_k},\\
\mathbf{R}_{b_k} &= \mathbf{U}_k\mathbf{\Lambda}_k^2\mathbf{U}_k^H + \widehat{\sigma}_b^2\mathbf{I}_{D_k},\\
\widetilde{\mathbf{R}}_{a_kb_k} &= \mathbf{P}_{u,k}^H\mathbf{R}_k\mathbf{P}_k = \mathbf{U}_{hk}^H\mathbf{R}_{hk}\mathbf{W}_k = \mathbf{\Lambda}_{hk}^{1/2}\mathbf{V}_k\mathbf{\Lambda}_k\mathbf{U}_k^H,\\
\widetilde{\mathbf{R}}_{b_ka_k} &= \mathbf{P}_k^H\mathbf{R}_k\mathbf{P}_{u,k} = \mathbf{U}_k\mathbf{\Lambda}_k\mathbf{V}_k^H\mathbf{\Lambda}_{hk}^{1/2}.
\end{aligned} \tag{34}$$

Based on (34), $C_k$ is given by

$$C_k = -\log_2\left(\left|\mathbf{I}_{D_k} - \widetilde{\mathbf{R}}_{a_kb_k}\mathbf{R}_{b_k}^{-1}\widetilde{\mathbf{R}}_{b_ka_k}\widetilde{\mathbf{R}}_{a_k}^{-1}\right|\right), \tag{35}$$

which is further expanded at the top of the next page in (36).

*2) Optimizing the Direction Control Matrix to Increase the Key Rate:* According to Woodbury matrix identity, i.e. $(\mathbf{I}+\mathbf{UV})^{-1} = \mathbf{I}-\mathbf{U}(\mathbf{I}+\mathbf{VU})^{-1}\mathbf{V}$, we simplified (36) to (37), which is shown at the top of the next page. Since $\mathbf{U}_k$ does not affect the $C_k$, we set it as $\mathbf{I}_{D_k}$. Since $\mathbf{V}_k$ is variable, maximizing the KR in (38) can be simplified by minimizing the second item $C_{2,k}$ which is expanded in (40). According to Corollary 1 in [25], the direction control matrix $\mathbf{V}_k$ is set as the $\mathbf{I}_{D_k}$ to minimize the $C_{2,k}$.

When $\mathbf{U}_k = \mathbf{I}_{D_k}$ and $\mathbf{V}_k = \mathbf{I}_{D_k}$, $\mathbf{P}_{e,k}$ in (28) can be further derived as

$$\mathbf{P}_{e,k} = \widetilde{\mathbf{V}}_k^{(0)}\mathbf{W}_k = \underbrace{\widetilde{\mathbf{V}}_k^{(0)}\mathbf{U}_{hk}}_{\mathbf{U}_{sk}}\underbrace{\mathbf{\Lambda}_{hk}^{-1/2}\mathbf{\Lambda}_k}_{\mathbf{\Lambda}_{sk}}, \tag{39}$$

We find that the $\mathbf{V}_k = \mathbf{I}_{D_k}$ and $\mathbf{U}_k = \mathbf{I}_{D_k}$ means the BS does not change the direction of the signal to estimate the subchannels and will optimize the KR. This agrees with intuition since $\mathbf{U}_{hk}$ decorrelates the channels and the phase change induced by $\mathbf{U}_k$ and $\mathbf{V}_k$ will bring correlation between subchannels. In (39), all other matrices are constants, while $\mathbf{\Lambda}_k = \text{diag}\left(\left[\sqrt{\lambda_{k,1}},\ldots,\sqrt{\lambda_{k,D_k}}\right]\right)$ serves as a variable matrix that can be further optimized in the next step.

Once we obtain $\mathbf{P}_{e,k}$ from (39), it can be seen that $\mathbf{P}_{e,k}$ is determined by $\widetilde{\mathbf{V}}_k^{(0)}$, $\mathbf{U}_{hk}$, $\mathbf{\Lambda}_{hk}^{-1/2}$, and $\mathbf{\Lambda}_k$. Specifically, $\widetilde{\mathbf{V}}_k^{(0)}$ mitigates inter-user interference, $\mathbf{U}_{hk}$ decorrelates spatially correlated channels to extract uncorrelated subchannels, $\mathbf{\Lambda}_{hk}^{-1/2}$ normalizes the channel eigenvalues, and $\mathbf{\Lambda}_k$ controls the power allocation across different spatial directions.

It is worth emphasizing that this precoding structure is fundamentally different from conventional ZF-based designs aimed at minimizing channel estimation error or maximizing

$$C_k = \log_2\left(\frac{\left|\mathbf{\Lambda}_{hk} + \frac{\sigma_a^2}{QP_b}\mathbf{I}_{D_k}\right|}{\left|\frac{\sigma_a^2}{QP_b}\mathbf{I}_{D_k} + \mathbf{\Lambda}_{hk}^{1/2}\left(\mathbf{I}_{D_k} - \mathbf{V}_k\mathbf{\Lambda}_k\mathbf{U}_k^H\frac{L}{\sigma_b^2}\left(\frac{L}{\sigma_b^2}\mathbf{U}_k\mathbf{\Lambda}_k\mathbf{V}_k^H\mathbf{V}_k\mathbf{\Lambda}_k\mathbf{U}_k + \mathbf{I}_{D_k}\right)^{-1}\mathbf{U}_k\mathbf{\Lambda}_k\mathbf{V}_k^H\right)\mathbf{\Lambda}_{hk}^{1/2}\right|}\right) \tag{36}$$

$$= \log_2\left(\frac{\left|\mathbf{I}_{D_k} + \frac{\sigma_a^2}{QP_b}\mathbf{\Lambda}_{hk}^{-1}\right|}{\left|\frac{\sigma_a^2}{QP_b}\mathbf{\Lambda}_{hk}^{-1} + \left(\frac{L}{\sigma_b^2}\mathbf{V}_k\mathbf{\Lambda}_k\mathbf{\Lambda}_k\mathbf{V}_k^H + \mathbf{I}_{D_k}\right)^{-1}\right|}\right) \tag{37}$$

$$= \log_2\left(\left|\mathbf{I}_{D_k} + \frac{\sigma_a^2}{QP_b}\mathbf{\Lambda}_{hk}^{-1}\right|\right) - \log_2\left(\left|\frac{\sigma_a^2}{QP_b}\mathbf{\Lambda}_{hk}^{-1} + \left(\frac{L}{\sigma_b^2}\mathbf{V}_k\mathbf{\Lambda}_k\mathbf{\Lambda}_k\mathbf{V}_k^H + \mathbf{I}_{D_k}\right)^{-1}\right|\right) = C_{1,k} - C_{2,k}. \tag{38}$$

channel capacity. While both approaches suppress interference, our scheme is specifically tailored to maximize the sum KR under the key generation framework. In particular, the decorrelation and normalization steps transform the spatially correlated channel into a set of uncorrelated subchannels. The subsequent power allocation further improves the sum KR. This security-oriented design distinguishes our method from conventional precoding strategies that focus on communication performance.

Once we obtain $\mathbf{P}_{e,k}$ from (39), we can discuss its practical implementation in a real system where the BS is connected to an ELAA. In (39), the $j$-th column of $\mathbf{U}_{sk} = \widetilde{\mathbf{V}}_k^{(0)}\mathbf{U}_{hk} \in \mathbb{C}^{N\times D_k}$ is a normalized beamforming vector $\bar{\mathbf{u}}_{k,j}$ that controls the phases of the signal at each antenna element for the $j$-th symbol. By aligning the $\mathbf{U}_{sk}$ with the eigenvectors of the channel and the null space of other UEs, $\mathbf{U}_{sk}$ ensures that the signal is directed along the optimal directions to estimate subchannels while minimizes interference from other UEs. To control the power for estimating subchannels, the $j$-th diagonal element of $\mathbf{\Lambda}_{sk} = \mathbf{\Lambda}_{hk}^{-1/2}\mathbf{\Lambda}_k \in \mathbb{C}^{D_k\times D_k}$ in (39), denoted as $a_{k,j}$, manages the power allocation across different antennas by adjusting the magnitude of the elements in the beamforming vector $\mathbf{u}_{k,j} = a_{k,j}\bar{\mathbf{u}}_{k,j}$. Since the BS applies precoding for all $K$ UEs simultaneously, $\sum_k \mathbf{u}_{k,j}$ is the combined beamforming vector for $j$-th symbol in an uplink/downlink packet.

Since $\mathbf{V}_k = \mathbf{I}_{D_k}$, the KR in (38) can be simplified as

$$C_k = \sum_{j=1}^{D_k}\log_2\left(1 + \frac{\widehat{\sigma}_a^{-2}\widehat{\sigma}_b^{-2}\lambda_{hk,j}\lambda_{k,j}}{\widehat{\sigma}_b^{-2}\lambda_{k,j} + \widehat{\sigma}_a^{-2}\lambda_{hk,j} + 1}\right). \tag{41}$$

For clarity of expression, we normalize $\lambda_{k,j}$ by $x_{k,j} = \lambda_{k,j}/\lambda_{hk,j}$. $\{x_{k,j}\}$ are a set of power allocation variables to be optimized. Thus, the KR in (41) can be simplified as

$$C_k = \sum_{j=1}^{D_k}\log_2\left(1 + \frac{x_{k,j}/m_{a,k,j}/m_{b,k,j}}{x_{k,j}/m_{b,k,j} + 1/m_{a,k,j} + 1}\right), \tag{42}$$

where $m_{a,k,j} = \frac{\sigma_a^2}{QP_b\lambda_{hk,j}}$ and $m_{b,k,j} = \frac{\sigma_b^2}{L\lambda_{hk,j}}$. The $j$-th sub-function of $C_k$ in (42) represents the KR achievable by the BS and the $k$-th UE through estimation of the $j$-th subchannel. Thus, optimizing $\mathbf{\Lambda}_k$ to increase the KR is equivalent to optimizing a set of scalar variables $\{x_{k,j}\}$.

$\mathbf{\Lambda}_k$ should let $\{\mathbf{P}_{e,k}\}$ meet the requirement that the transmit power at the BS is not greater than $P_a$. Based on (39), the transmit power at the BS $\sum_{k=1}^K \|\mathbf{P}_{e,k}\|_F^2$ can be expanded as

$$\sum_{k=1}^K \|\mathbf{P}_{e,k}\|_F^2 = \sum_{k=1}^K \mathrm{Tr}\left(\mathbf{P}_{e,k}\mathbf{P}_{e,k}^H\right) = \sum_{k=1}^K \mathrm{Tr}\left(\mathbf{\Lambda}_k^2\mathbf{\Lambda}_{hk}^{-1}\right)$$
$$= \sum_{k=1}^K\sum_{j=1}^{D_k}\frac{\lambda_{k,j}}{\lambda_{hk,j}} = \sum_{k=1}^K\sum_{j=1}^{D_k} x_{k,j}. \tag{43}$$

Substituting (43) to the sum power constraint $\sum_{k=1}^K \|\mathbf{P}_{e,k}\|_F^2 \le P_a$, the constraint can be reformulated as:

$$\sum_{k=1}^K \|\mathbf{P}_{e,k}\|_F^2 = \sum_{k=1}^K\sum_{j=1}^{D_k} x_{k,j} \le P_a. \tag{44}$$

### D. Summary

In this section, $\mathbf{P}_{e,k}$ is decomposed in (39) to mitigate interference and convert channels into uncorrelated subchannels. Consequently, the KR of the $k$-th UE is simplified in (42) which is determined by $\{x_{k,j}\}$. $x_{k,j}$ controls the transmit power allocated to the $j$-the subchannel of the $k$-th UE. Considering the sum power constraint for $\{x_{k,j}\}$ in (44), we will develop two power allocation algorithms to further increase the KR. In Section V, we optimize $\{x_{k,j}\}$ for maximizing the sum KR of $K$ UEs. In Section VI, we optimize $\{x_{k,j}\}$ to solve the problem of the KR fairness.

## V. Power Allocation for Subchannels to Maximize the Sum Key Rate

To maximize the sum KR of $K$ UEs, we optimize the power allocation variables $\{x_{k,j}\}$ for the subchannels, where the KR of the $k$-th UE $C_k$ is simplified as shown in (42). The transmit power $P_a$ that is distributed among the $K$ UEs adheres to the sum power constraint in (44). Therefore, the optimization problem (P1) is formulated as follows:

$$\text{(P1)} \quad \max_{x_{k,1},\ldots,x_{k,D_k}} \sum_{k=1}^K C_k$$
$$\text{s.t.} \sum_{k=1}^K\sum_{j=1}^{D_k} x_{k,j} \le P_a,\ x_{k,j} \ge 0. \tag{45}$$

We obtain the optimal $x_{k,j}$ from (P1) and calculate $\lambda_{k,j} = x_{k,j}\lambda_{hk,j}$. Once we substitute $\lambda_{k,j}$ to $\mathbf{\Lambda}_k$, the $\mathbf{P}_{e,k}$ can be obtained from (39).

### A. Water-Filling Algorithm for Power Allocation

The (P1) can be solved by the Lagrange multiplier method. According to [20], when the objective function is concave over a convex set, the solution obtained from the Lagrange multiplier method can achieve the global maximum. First of all, we

$$C_{2,k} = \log_2\left(\left|\frac{\sigma_a^2}{QP_b}\mathbf{\Lambda}_{hk}^{-1}\left(\frac{L}{\sigma_b^2}\mathbf{V}_k\mathbf{\Lambda}_k\mathbf{\Lambda}_k\mathbf{V}_k^H + \mathbf{I}_{D_k}\right) + \mathbf{I}_{D_k}\right|\right) - \log_2\left(\left|\frac{L}{\sigma_b^2}\mathbf{V}_k\mathbf{\Lambda}_k\mathbf{\Lambda}_k\mathbf{V}_k^H + \mathbf{I}_{D_k}\right|\right)$$
$$= \log_2\left(\left|\frac{\sigma_a^2}{QP_b}\mathbf{\Lambda}_{hk}^{-1}\right|\right) + \log_2\left(\left|\left(\frac{L}{\sigma_b^2}\mathbf{V}_k\mathbf{\Lambda}_k\mathbf{\Lambda}_k\mathbf{V}_k^H + \mathbf{I}_{D_k}\right) + \frac{QP_b}{\sigma_a^2}\mathbf{\Lambda}_{hk}\right|\right) - \log_2\left(\left|\frac{L}{\sigma_b^2}\mathbf{\Lambda}_k\mathbf{\Lambda}_k + \mathbf{I}_{D_k}\right|\right), \quad (40)$$

discuss whether the objective function of (P1), $f = \sum_{k=1}^{K} C_k$, is concave. The second-order partial derivatives of $f$ are given at the top of the next page.

Based on (46), the second-order partial derivatives of $f$ are $\frac{\partial^2 f}{\partial x_{k,j}^2} \leq 0$, which indicates the $N_D = \sum_{k=1}^{K} D_k$ subfunctions are concave. By the Karush–Kuhn–Tucker (KKT) conditions, the solution of (P1) can be obtained by the water-filling algorithm [26]. Given the water-filling level $\mu$, the KKT conditions are given by

$$\frac{\partial f}{\partial x_{k,j}} = \mu,\ \mu\left(\sum\nolimits_{k=1}^{K}\sum\nolimits_{j=1}^{D_k} x_{k,j} - P_a\right) = 0,$$
$$\sum\nolimits_{k=1}^{K}\sum\nolimits_{j=1}^{D_k} x_{k,j} \leq P_a,\ x_{k,j} \geq 0. \quad (47)$$

Based on the first-order partial derivative of $f$, we express $x_{k,j}$ in terms of $\mu$ in (48), which is given at the top of the next page. By substituting $x_{k,j}$ to $\sum_{k=1}^{K}\sum_{j=1}^{D_k} x_{k,j} = P_a$, a water-filling algorithm in [27] can be applied to find the Lagrange multiplier $\mu$. Based on $\mu$, the $x_{k,j}$ can be derived from (48).

### B. Deep Learning-based Power Allocation

The water-filling algorithm provides optimal power allocation variables for (P1). However, it relies on the bisection method to search for the Lagrange multiplier $\mu$ under the power constraint, which requires many iterations. A closed-form solution for $\mu$ in terms of $P_a$ is not available. In contrast, without needing iterative optimization, a deep neural network can directly predict the Lagrange multiplier $\mu$. The computationally intensive training is performed offline, enabling fast and efficient inference during the online stage.

Previous studies, such as [28] and [27], have applied unsupervised deep learning for power allocation by learning the relationship between input parameters and power allocation variables, with the loss function defined as the sum KR. However, the number of neurons in the output layer grows linearly with the number of power allocation variables. We find that the power allocation variables $\{x_{k,j}\}$ have a closed-form expression in terms of $\mu$. By directly learning the relationship between the $\{x_{k,j}\}$ and $\mu$, we can reduce the complexity of the neural network. Given $m_{a,k,j}$ and $m_{b,k,j}$, this allows $\{x_{k,j}\}$ to be computed directly from the predicted $\mu$ in (48).

We introduce a supervised feedforward neural network (FNN) to learn the relationship between $P_a$ and $\mu$, which we refer to as a key generation Lagrange multiplier network (KGLM-Net). Given $\mu$, the transmit power allocated for the $j$-th subchannel of the $k$-th UE $x_{k,j}$ can be derived in (48). Subsequently, the precoding matrices can be obtained from (24), which is given by $\mathbf{W}_k = \mathbf{U}_{hk}\mathbf{\Lambda}_{hk}^{-1/2}\mathbf{\Lambda}_k$, where $\mathbf{\Lambda}_k = \text{diag}\left(\left[\sqrt{\lambda_{k,1}}, \sqrt{\lambda_{k,2}}, \ldots, \sqrt{\lambda_{k,D_k}}\right]\right)$ and $\lambda_{k,j} = \lambda_{hk,j}x_{k,j}$. Once $\mathbf{W}_k$ is acquired, $\mathbf{P}_{e,k}$ can be derived from (39).

*1) Network Structure:* As illustrated in Fig. 3, the KGLM-Net constitutes three types of layers, i.e., input, hidden and output layers. The information flow is from the input neuron, through the hidden neuron and to the output neuron. The Levenberg-Marquardt optimizer is used for training feedforward neural networks [29]. According to Levenberg-Marquardt optimization, weight and bias values are updated to minimize the objective function, i.e. mean square error (MSE) to fit a neural network [30].

*2) Data Generation:* To train the proposed KGLM-Net, we construct a dataset consisting of sample pairs $(P_a, \mu)$ generated via the water-filling algorithm described in [27]. Each transmit power $P_a$ is uniformly sampled from the interval $[p_{\min}, p_{\max}]$, and the corresponding Lagrange multiplier $\mu$ is computed by solving problem (P1) using the water-filling solution. A total of $T_D = 200$ samples are generated and randomly split into 140 training samples, 30 validation samples, and 30 test samples.

*3) Data Preprocessing:* To facilitate efficient and stable training, both input and output values are normalized to the range $[0, 1]$. Specifically, the $i$-th transmit power sample $P_a(i)$ is normalized as

$$P_a(i) = \frac{P_a(i) - \min\{P_a(j)\}}{\max\{P_a(j)\} - \min\{P_a(j)\}}, \quad (49)$$

where $j \in 1, \ldots, T_D$. Besides, the $i$-th sample of the Lagrange multiplier $\mu(i)$ is normalized as

$$\mu(i) = \frac{\mu(i) - \min\{\mu(j)\}}{\max\{\mu(j)\} - \min\{\mu(j)\}}. \quad (50)$$

*4) Training and Inference:* During the training phase, KGLM-Net updates its parameters through supervised learning, aiming to minimize the MSE, which is given by

$$Loss = \frac{1}{N_m}\sum\nolimits_{n=1}^{N_m}\left|\widetilde{\mu}_n - \mu_n\right|^2, \quad (51)$$

where $N_m$ is the number of training samples, $\widetilde{\mu}_n$ is the value calculated in the $n$-th training, and $\mu_n$ is the actual Lagrange multiplier. The training is finished when MSE on the validation set does not show improvement over 6 epochs or iteration epochs reach the maximum 1000.

In the online inference phase, the BS feeds input transmit power into the KGLM-Net to obtain the Lagrange multiplier $\hat{\mu}$. Given $\hat{\mu}$, the power allocation variables $\{x_{k,j}\}$ are obtained according to (48). Accordingly, $\mathbf{P}_{e,k}$ can be derived from (39).

## VI. Power Allocation for Subchannel to Ensure Key-Rate Fairness

In Section V, we focus on optimising $\{x_{k,j}\}$ to maximize the sum KR, derive $\mathbf{\Lambda}_k$ and use it to construct the precoding matrix $\mathbf{P}_{e,k}$ in (39). However, if a UE is located too far from

$$\frac{\partial^2 f}{\partial x_{k,j}^2} = \frac{-2m_{a,k,j}m_{b,k,j}x_{k,j} - m_{b,k,j}^2(2m_{a,k,j}+1)}{\ln 2\left(m_{a,k,j}x_{k,j}^2 + m_{b,k,j}(2m_{a,k,j}+1)x_{k,j} + m_{b,k,j}^2(m_{a,k,j}+1)\right)^2}, \tag{46}$$

$$x_{k,j} = \frac{\sqrt{m_{b,k,j}^2(2m_{a,k,j}+1)^2 - 4m_{a,k,j}m_{b,k,j}(m_{b,k,j}(m_{a,k,j}+1) - \frac{1}{\ln 2\mu})}}{2m_{a,k,j}} + \frac{-m_{b,k,j}(2m_{a,k,j}+1)}{2m_{a,k,j}}. \tag{48}$$

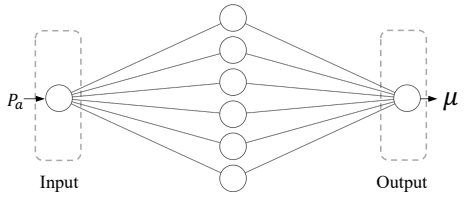



Fig. 3. KGLM-Net.

the BS in a near-field multi-user system, its channel quality deteriorates. As a result, more transmit power is allocated to UEs with better channel conditions, sacrificing the KR of the distant UE in order to maximize the sum KR. In this section, we will address KR fairness by increasing the minimum KR among the $K$ UEs as much as possible, a problem known as minimum-KR maximization.

Under total power constraint, we optimize the power allocation variables $\{x_{k,j}\}$ to increase the minimum KR. The optimization problem (P2) is formulated as follows:

$$\text{(P2)} \max_{\{x_{k,j}\}} \min_{k} C_k \quad \text{s.t.} \sum\nolimits_{k=1}^{K}\sum\nolimits_{j=1}^{D_k} x_{k,j} \le P_a,\ x_{k,j} \ge 0. \tag{52}$$

### A. CVX-based Power Allocation

To solve the structure of the max-min in the objective function, we introduce a slack variable $t$, which is constrained to be smaller than or equal to each KR $C_k$. The optimization problem (P2) can then be reformulated as follows:

$$\text{(P3)} \max_{\{x_{k,j}\}} t \quad \text{s.t. } C_1 \ge t, C_2 \ge t, \ldots, C_K \ge t, \sum\nolimits_{k=1}^{K}\sum\nolimits_{j=1}^{D_k} x_{k,j} \le P_a,\ x_{k,j} \ge 0. \tag{53}$$

As discussed in Section V, the KR $C_k$ in (42) is concave. Thus, the (P3) can be solved by the convex optimization toolbox CVX [31], which is a Matlab-based modelling system for convex optimization. Since it is based on an iterative optimization algorithm, it is computationally complex. This raises the question of whether a closed-form solution exists between the transmit power $P_a$ and the power allocation variables $\{x_{k,j}\}$ in the minimum-KR maximization problem.

### B. Water Filling Algorithm for Power Allocation

The CVX-based power allocation method based on iterative optimization algorithms optimizes (P3) to find the optimal power allocation variables $\{x_{k,j}\}$. To reduce computational complexity, we will explore whether there is a closed-form solution of $\{x_{k,j}\}$ in terms of $P_a$, simplifying the search process for $\{x_{k,j}\}$. First, we introduce the Lagrange multipliers $\mu_k$ and transfer (P3) to KKT conditions for obtaining optimal $\{x_{k,j}\}$. Second, to derive the analytical expressions for determining the optimal $\{x_{k,j}\}$ in terms of $\mu_k$, we approximate the KR in the high-power case. We derive the analytical expression for $\mu_k$ as a function of $t$. Third, we use the bisection algorithm to search $t$. If $t$ is obtained, the optimal $\{x_{k,j}\}$ can be obtained using the analytical expressions.

*1) Lagrange Multiplier Method:* As analyzed in Section V, the objective function $C_k$ of the (P3) is concave. As noted in [26], the Lagrange multiplier method can be applied to find the optimal $\{x_{k,j}\}$ in terms of $\mu_k$ and $t$. We derive the corresponding KKT conditions as

$$g_{k,j}(x_{k,j}) = \frac{\partial C_k}{\partial x_{k,j}} = \mu_k,\ C_k = t, \tag{54}$$

$$\mu_k\left(\sum\nolimits_{k=1}^{K}\sum\nolimits_{j=1}^{D_k} x_{k,j} - P_a\right) = 0, \tag{55}$$

$$\sum\nolimits_{k=1}^{K}\sum\nolimits_{j=1}^{D_k} x_{k,j} \le P_a,\ x_{k,j} \ge 0. \tag{56}$$

As shown in (54), when the power allocation is optimal, $g_{k,1}(x_{k,1}) = ...g_{k,j}(x_{k,j})... = g_{k,D_k}(x_{k,D_k}) = \mu_k$, which means the increasing rates of the transmit power assigned to the $j$-th channel are the same for the $k$-th UE [26]. The condition (54) implies that the KRs of all UEs are equal when (P3) reaches its optimal solution, which is noted in [26].

Since there are log functions and fractions in $C_k$ in (42), the KKT conditions cannot be directly transferred to find the analytical expression of $\{x_{k,j}\}$ in terms of $\mu_k$ and the $t$. Thus, we approximate the KR $C_k$ in the high-power case.

*2) High-Power Case:* If the transmit power at the UE $P_b$ is high, $m_{a,k,j} = \frac{\sigma_a^2}{QP_b\lambda_{h_{k,j}}}$ tends to be zeros. The objective function $C_k$ in (42) can be approximated by

$$C_k = \sum\nolimits_{j=1}^{D_k}\log_2\left(1 + \frac{x_{k,j}}{m_{a,k,j}x_{k,j} + m_{b,k,j} + m_{a,k,j}m_{b,k,j}}\right) \approx \sum\nolimits_{j=1}^{D_k}\log_2\left(1 + \frac{x_{k,j}}{m_{b,k,j}}\right). \tag{57}$$

The first-order derivative of (57) is given by

$$g_{k,j}(x_{k,j}) = \frac{\partial R_k}{\partial x_{k,j}} = \frac{1}{\ln 2\,(m_{b,k,j} + x_{k,j})} = \mu_k. \tag{58}$$

**Algorithm 1** Bisection Algorithm

**Input:** $P_a$, $t_{max}$, $t_{min}$, $D_k$, $m_{b,k,j}$, $\epsilon$;
**Output:** $t$.
1: Set $t = (t_{min} + t_{max})/2$;
2: Calculate $x_{k,j}$ according to (61);
3: **repeat**
4: **if** $\sum_{k=1}^{K}\sum_{j=1}^{D_k} x_{k,j} < P_a$ **then**
5: $t_{min} = t$;
6: **else**
7: $t_{max} = t$;
8: **end if**
9: Set $t = (t_{min} + t_{max})/2$;
10: Calculate $x_{k,j}$ according to (61);
11: **until** $|\sum_{k=1}^{K}\sum_{j=1}^{D_k} x_{k,j} - P_a| \leq \epsilon$.

Based on (54), we get the $x_{k,j}$ in a function of $\mu_k$ as

$$x_{k,j} = \frac{1}{\ln 2\mu_k} - m_{b,k,j}. \tag{59}$$

Substituting (57) and (59) into (54), we get the relationship between $\mu_k$ and $t$, which is given by

$$\mu_k = \left(\ln 2 \times 2^{\frac{t+\sum_{j=1}^{D_k}\log_2\left(m_{b,k,j}\right)}{D_k}}\right)^{-1}. \tag{60}$$

Substituting (59) and (60) to (56), we have

$$\sum_{k=1}^{K}\sum_{j=1}^{D_k} 2^{\frac{t+\sum_{j=1}^{D_k}\log_2\left(m_{b,k,j}\right)}{D_k}} - m_{b,k,j} = P_a. \tag{61}$$

*3) Bisection Method:* Based on (61), we design a bisection algorithm to find the $t$, given in **Algorithm 1**. Once the variable $t$ is calculated from **Algorithm 1**, we obtain $\mu_k$ from $t$ based on (60). Accordingly, we get optimal $x_{k,j}$ from $\mu_k$ based on (59). Compared to the CVX toolbox, the bisection algorithm relieves the computational complexity by only searching the slack variable $t$.

### C. Deep Learning-based Power Allocation

The water-filling algorithm uses the bisection algorithm to search $t$ so that the $\mu_k$ can be derived from $t$ in (60) and $\{x_{k,j}\}$ from $\mu_k$ in (59). Compared to the CVX-based algorithm, the water-filling algorithm reduces the complexity since it only requires the search of $t$. However, the bisection algorithm still needs to search $t$ to find the optimal power allocation variables when a total transmit power $P_a$ is provided. To reduce computational complexity, we propose leveraging a FNN to learn the mapping between the transmit power $P_a$ and the parameter $t$, enabling the direct prediction of the optimal allocation variables $\{x_{k,j}\}$ based on the analytical expressions in (60) and (59).

*1) Network Structure:* We introduce the FNN to learn the relationship between the transmit power $P_a$ and the slack variable $t$. The FNN architecture consists of a single input representing $P_a$, one hidden layer with 20 neurons, and one output neuron representing $t$.

*2) Date Generation:* To train the FNN, we generate 200 pairs of transmit power $P_a(i)$ and corresponding slack variable $t(i)$ by solving problem (P3) using the CVX toolbox. The dataset is randomly divided into 140 samples for training, 30 for testing, and 30 for validation.

*3) Date Preprocessing:* The normalization of the $i$-th transmit power sample $P_a(i)$ follows the same procedure as in (49). The $i$-th slack variable is normalized as

$$t(i) = \frac{t(i) - \min\{t(j)\}}{\max\{t(j)\} - \min\{t(j)\}}. \tag{62}$$

*4) Training and Inference:* In the offline training, the loss function is defined as MSE. The training stops either when the validation MSE fails to improve for 6 epochs or when a maximum of 1000 epochs is reached.

During the online inference phase, the trained FNN predicts the slack variable $t$ given the input transmit power. Therefore, $\mu_k$ is derived from $t$ as shown in equation (60). Accordingly, once $\mu_k$ is obtained, $x_{k,j}$ is determined by equation (59). The equations based on (60) and (59) are based on the approximation approach employed in high-power regimes of $P_b$. If $P_b$ is not sufficiently large, the CVX-based power allocation method in Sec. VI-A can be used.

## VII. Numerical Results

This section presents numerical results that demonstrate the efficiency of the proposed near-field key generation schemes.

### A. Simulation Setup

The BS is positioned along the $x$-axis, with its reference antenna located at $(0, 0)$. The antennas are spaced at a distance of $d$ that is equal to half of the wavelength, i.e., $d = \lambda/2$. There are four UEs and the $k$-th UE, $k \in \{1, 2, 3, 4\}$, is situated at coordinate $\left(r_{b_k}\sin(\psi_{b_k}), r_{b_k}\sqrt{1-\sin^2(\psi_{b_k})}\right)$, where $r_{b_k}$ is the distance from the $k$-th UE to the reference antenna and $\psi_{b,k}$ is the angle between the $k$-th BS-UE link and $y$-axis. The carrier frequency is set as $f_c = 3.5$ GHz. The noise powers are set to $\sigma_0^2 = \sigma_a^2 = \sigma_b^2 = -114$ dBm [32].

The channel gain is set as $\beta_{b_k,l} \sim \mathcal{CN}(0, \alpha_{b_k,l})$, where $\alpha_{b_k,l}$ is the distance-dependent path-loss effect. The path-loss effect from the $k$-the UE to the BS over the $l$-th path is $\alpha_{b_k,l} = \alpha_0\left(\frac{r_{b_k}}{r_0}\right)^{-\epsilon_{b_k}}$, where $\epsilon_{b_k} = 3.5$ is the path-loss exponent and $\alpha_0 = -30$ dB denotes the path-loss effect at $r_0 = 1$ m [33].

### B. Considered Algorithms

The comparison and proposed algorithms for sum-KR maximization are as follows:

1) **Equal Power Allocation (EPA):** This scheme equally divides $P_a$ to $\{x_{k,j}\}$. Thus, we calculate $x_{k,j}$ as $x_{k,j} = P_a/N_D$, where $N_D = \sum_k D_k$. Futhermore, we have $\lambda_{k,j} = \lambda_{hk,j}P_a/N_D$, where $\lambda_{k,j}$ is the $j$-th diagonal element of $\mathbf{\Lambda}_k$. Given $\mathbf{\Lambda}_k$, $\mathbf{P}_{e,k}$ in (39) can be obtained. The KR of the $k$-th UE $I_k$ is calculated in (17).
2) **Water Filling-based Power Allocation (WFPA):** To maximize the sum KR, $\{x_{k,j}\}$ is optimized from (P1).

In Section V-A, the closed-form expression of $\{x_{k,j}\}$ is derived in terms of $\mu$ in (48). The water-filling algorithm is applied to find $\mu$, as shown in Algorithm 1 in [27]. Based on $\mathbf{P}_{e,k}$ in (39), $I_k$ is calculated from (17).

3) **FNN-based Power Allocation (FNNPA):** A supervised FNN is trained to learn the relationship between $\mu$ and $P_a$, which is described in Section V-B. The dataset of samples is collected from the WFPA scheme. When FNN predicts $\mu$ from a $P_a$, $\{x_{k,j}\}$ can be derived from (48). Based on $\{x_{k,j}\}$, $\mathbf{P}_{e,k}$ is designed in (39) and $I_k$ is calculated in (17).

The proposed algorithms for KR fairness are as follows:

1) **CVX-based Power Allocation for KR Fairness (CVXPAKF):** Considering KR fairness in near-field scenarios, the BS applies the power allocation method to maximize the minimum KR in Section VI-A. Since the objective function is concave, the (P3) is solved by the CVX toolbox to find the optimal $\{x_{k,j}\}$. Based on $\mathbf{P}_{e,k}$, $I_k$ is calculated from (17).
2) **Water Filling-based Power Allocation for KR Fairness (WFPAKF):** By introducing a slack variable $t$ and Lagrange multipliers $\{\mu_k\}$, the (P3) considering KR fairness is reformulated to a water-filling problem in Section VI-B. To simplify the KR, it is approximated in (57) when $P_b$ is high. Therefore, $x_{k,j}$ is obtained from $\mu_k$ in (59). Based on (60), the Lagrange multiplier $\mu_k$ is determined by $t$. A bisection algorithm in **Algorithm 1** is proposed to find $t$. Based on $t$, the upper bound on the KR is calculated from (57).
3) **FNN-based Power Allocation for KR Fairness (FNNPAKF):** A FNN is designed to learn the relationship between $t$ and $P_a$ in high-power case, described in Section VI-C. The dataset of samples is collected from the CVXPAKF scheme. Once $t$ is predicted by the FNN given the $P_a$, the $\{x_{k,j}\}$ are determined in terms of $\mu_k$ from (59), and $\mu_k$ is expressed in terms of $t$ as shown in (60). Based on $\{x_{k,j}\}$, $\mathbf{P}_{e,k}$ is designed in (39) and $I_k$ is calculated in (17).

### C. Results of FNN

We evaluate the training and test performance of the proposed FNN.

Figure 4 plots the MSE between the target and predicted Lagrange multipliers on the training and validation sets versus the number of training epochs under different numbers of UEs ($K = 2, 4, 6, 8$). The training process is terminated based on an early stopping criterion: if the MSE on the validation set does not improve for 6 consecutive epochs, training is halted to prevent overfitting. The steadily decreasing MSE curves indicate that the training process converges effectively.

To evaluate the model performance, we take the case of 4 UEs as an example. As shown in Fig. 5, the data points from both the training and test sets lie closely along the ideal line $\widehat{\mu} = \mu$, indicating a strong consistency between the predicted and target Lagrange multipliers. To quantitatively assess the consistency between the predicted and target Lagrange multipliers, we compute the Pearson correlation coefficients for both the training and test sets. The results are 0.999998 for the training set and 0.999992 for the test set, confirming that the FNN can effectively predict the Lagrange multiplier.

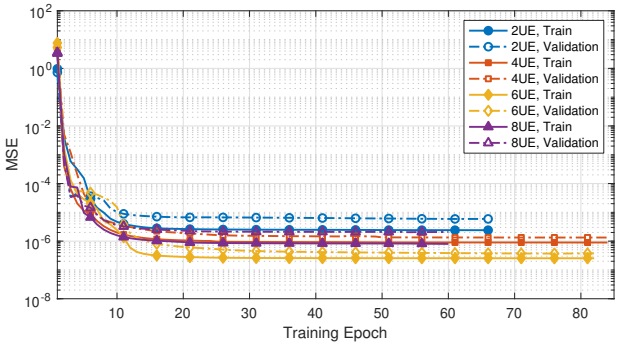
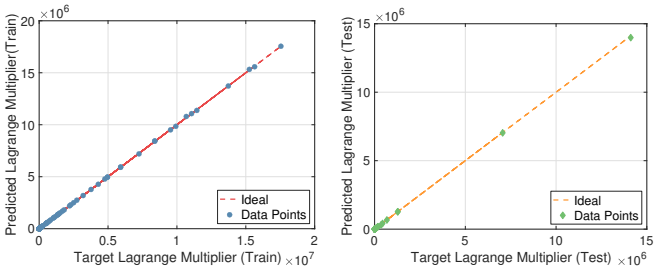



Fig. 4. MSE versus training epochs.




(a) (b)

Fig. 5. (a) $\widehat{\mu}$ versus $\mu$ on training data (140 samples). (b) $\widehat{\mu}$ versus $\mu$ on test data (30 samples).

### D. Results of KR

We evaluated the KR against transmit power, the number of antennas, the distances, and the angles.

Figure 6 shows the sum KR of 4 UEs versus transmit power $P_a$. As the transmit power $P_a$ increases, the channel estimation errors are reduced, resulting in an increase in the sum KR. In near-field propagation environments where UEs share different angles $\psi_{b_k}$ and distances with respect to the reference antenna $r_{b_k}$, our proposed precoding method in Section IV enables UEs to generate secret keys with the BS simultaneously. The WFPA scheme outperforms the EPA scheme since the WFPA scheme allocates more power to those subchannels with better channel quality. The FNNPA scheme almost matches the WFPA scheme, which indicates the proposed FNN learned the relationship between the transmit power $P_a$ and $\mu$. Based on $\mu$, power allocation variables $\{x_{k,j}\}$ can be derive from (48). Additionally, the EPA scheme converges to the WFPA and the FNNPA schemes as $P_a$ increases. At high transmit power levels, the optimal power allocation tends to become equal power allocation.

Figure 6 also shows the sum KR of 4UEs versus transmit power $P_a$ considering KR fairness, namely improving the minimum KR in near-field scenarios. In the CVXPAKF scheme, the sum KR is smaller than that of the WFPA scheme because the power allocation based on maximizing the minimum KR sacrifices the KRs of nearby UEs to enhance the KRs of distant

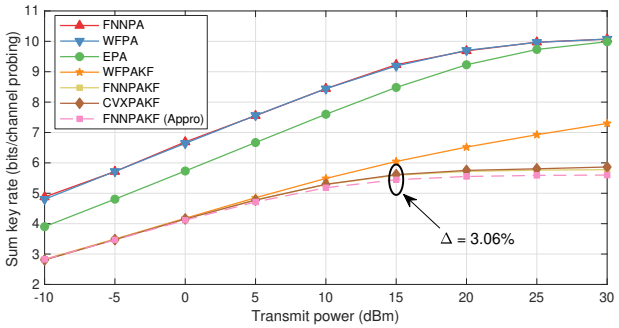



Fig. 6. Sum key rate versus transmit power ($P_b = 20$ dBm, $N = 256$, $r_{b_1} = 20$ m, $r_{b_2} = 40$ m, $r_{b_3} = 80$ m, $r_{b_4} = 100$ m, $\psi_{b_1} = -0.5$ radian, $\psi_{b_2} = 0.5$ radian, $\psi_{b_3} = 1$ radian, and $\psi_{b_4} = -1$ radian).

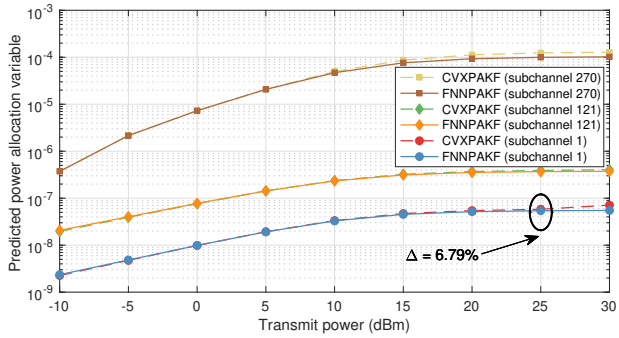



Fig. 7. Predicted $x_{k,j}$ versus transmit power ($P_b = 20$ dBm, $N = 256$, $r_{b_1} = 20$ m, $r_{b_2} = 40$ m, $r_{b_3} = 80$ m, $r_{b_4} = 100$ m, $\psi_{b_1} = -0.5$ radian, $\psi_{b_2} = 0.5$ radian, $\psi_{b_3} = 1$ radian, and $\psi_{b_4} = -1$ radian).

users. The WFPAKF scheme nearly matches the CVXPAKF scheme when $P_a$ is low and gradually deviates as $P_a$ increases. At low $P_a$, $P_b$ is relatively high, and the KR can be approximated by (57). Moreover, the WFPAKF scheme uses a simple bisection search algorithm, which reduces the complexity compared to the CVXPAKF scheme. The FNNPAKF scheme closely matches the performance of the CVXPAKF scheme, indicating that the proposed FNN accurately predicts the $t$ based on the input $P_a$. Under the optimal power allocation for key rate fairness, the four UEs achieve the same key rate, each equal to the predicted $t$. Once $t$ is obtained, the variables $\mu_k$ and $x_{k,j}$ can be computed using equations (60) and (59), respectively. By substituting the predicted $x_{k,j}$ into the original KR expression, the resulting KR can be evaluated. At $P_a = 15$ dBm, the relative error between the FNNPAKF (Appro) and CVXPAKF schemes is $3.06\%$, while the average relative error across all considered power levels is $2.28\%$. When $P_a \leq 5$ dBm, the KR values of both schemes are nearly identical, since the value of $P_b$ remains sufficiently large, ensuring the validity of the approximation. As $P_a$ increases, the error introduced by the approximation becomes more significant due to the growing term $m_{a,k,j}x_{k,j}$ in the denominator in equation (57), resulting in a visible deviation between the two curves.

Figure 7 illustrates the predicted power allocation variables as a function of the transmit power $P_a$. For subchannels indexed by 270, 121, and 1, the average relative errors across different $P_a$ values are $8.6\%$, $3.4\%$, and $5.2\%$, respectively. The prediction aligns more closely with the target values when $P_a$ is small and $P_b$ is large, as a higher power $P_b$ causes $m_{a,k,j}$ to approach zero, thereby improving the approximation. However, as $P_a$ increases, the allocated power $x_{k,j}$ becomes larger, making the term $m_{a,k,j}x_{k,j}$ in the denominator of equation (57) increasingly significant. This degrades the accuracy of the approximation used in the prediction.

In Fig. 8, we study the sum KR of 4 UEs versus the number of antennas $N$ at the BS. Increasing $N$ improves the resolution of angles and distances, aiding in distinguishing UEs. As a result, more near-field channels that do not overlap with those of other UEs become available, leading to an increase in the sum KR. Additionally, the length of the downlink packets scales linearly with $N_s = \max_k\{D_k\}$, where $D_k$ is the maximum number of subchannels for the $K$ UEs. As $N$

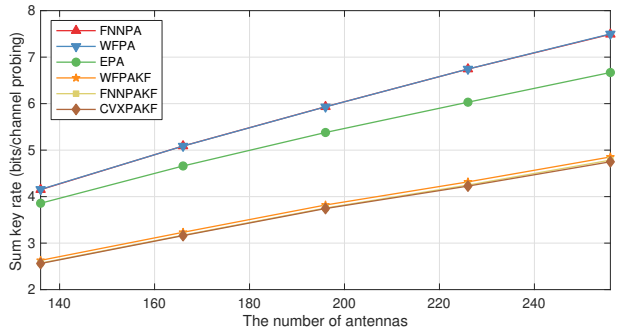



Fig. 8. Sum key rate versus the number of antennas ($P_a = 5$ dBm, $P_b = 20$ dBm, $r_{b_1} = 20$ m, $r_{b_2} = 40$ m, $r_{b_3} = 80$ m, $r_{b_4} = 100$ m, $\psi_{b_1} = -0.5$ radian, $\psi_{b_2} = 0.5$ radian, $\psi_{b_3} = 1$ radian, and $\psi_{b_4} = -1$ radian).

increases, $D_k$ also increases because more antennas allow the BS to use SVD to identify additional spatial dimensions for the $k$-th UE to extract more subchannels. Thus, increasing $N$ improves channel estimation accuracy, leading to a higher sum KR. The WFPA and the FNNPA schemes outperform the EPA scheme with equal power allocation. As for KR fairness, the WFPAKF scheme nearly matches the CVXPAKF and the FNNPAKF schemes when $P_a = 5$ dBm and $P_b = 20$ dBm, which validates the effectiveness of **Algorithm 1**.

Figure 9 shows the KR $I_4$ of the fourth UE as its distance from the BS increases. The fourth UE is selected for this analysis because it is the farthest from the BS in previous figures. We set the fourth UE to share the same angle as the first UE, with $\psi_{b_4} = \psi_{b_1} = -0.5$ radian. By varying the distance of the fourth UE from near to far, we evaluate whether the KR fairness algorithms improve the KR for UEs with poorer channel quality. When the first UE is located with the distance $r_{b_1} = 10$ m, the distance of the fourth UE is defined as $r_{b_4} = (r_{b_1} + \Delta r)$ m, where $\Delta r$ m represents the change in distance relative to $r_{b_1}$ m. The $I_4$ initially increases, reaches a peak, and then decreases. Despite the poorer channel quality at greater distances for the fourth UE, the increased distance between the first and fourth UEs allows the BS to find additional spatial dimensions for extracting more subchannels. Thus, the $I_4$ increases and then decreases as the effect from channel quality dominates when $r_{b_4}$ is large. In far-field key generation [12], if another UE shares the same angle with

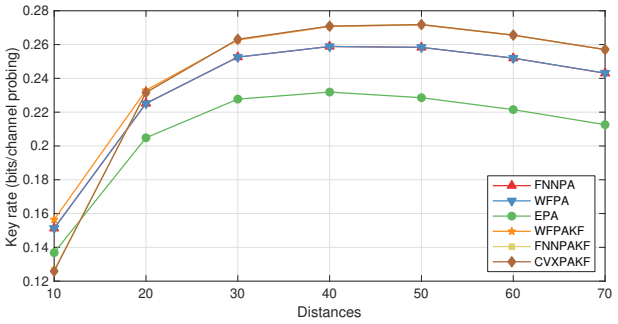



Fig. 9. The key rate of the fourth UE versus the distances ($P_a = 5$ dBm, $P_b = 5$ dBm, $N = 256$, $r_{b_1} = 20$ m, $r_{b_2} = 40$ m, $r_{b_3} = 80$ m, $\psi_{b_1} = -0.5$ radian, $\psi_{b_2} = 0.5$ radian, $\psi_{b_3} = 1$ radian and $\psi_{b_4} = -0.5$ radian).

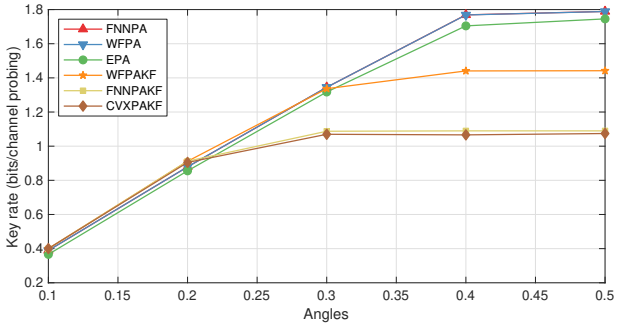



Fig. 10. The key rate of the fourth UE versus the angles ($P_a = 5$ dBm, $P_b = 5$ dBm, $N = 256$, $r_{b_1} = 20$ m, $r_{b_2} = 40$ m, $r_{b_3} = 80$ m, $r_{b_4} = 20$ m, $\psi_{b_1} = -0.5$ radian, $\psi_{b_2} = 0.5$ radian, and $\psi_{b_3} = 1$ radian).

the fourth UE but is located closer to the BS, the distance dimension can not be utilized for key generation, even when a larger number of antennas is employed to distinguish between the angles of the two UEs. Moreover, the KR of the CVXPAKF is initially lower than that of the WFPA and the FNNPA schemes and then becomes greater than theirs. This occurs because when the fourth UE is close to the BS, it is a stronger UE among the four UEs, so its KR is sacrificed for the weaker UE. However, when the fourth UE is far from the BS, it becomes the weaker UE, the proposed algorithms considering KR fairness allocate more power to the fourth UE.

Figure 10 investigates the KR of the fourth UE $I_4$ versus the angles. The distances of the first UE and the fourth UE are set the same as $r_{b_1} = r_{b_2} = 20$ m. The angle of the fourth UE is defined as $\psi_{b_4} = (\psi_{b_1} + \Delta\psi)$ radian, where $\Delta\psi$ radian represents the variation of angle relative to the first UE. The dimension of the angles can be leveraged for the design of precoding matrices based on EVD and SVD when two UEs have the same distances with respect to the reference antenna. As the angles between these two UEs increase, the KR of the WFPA and FNNPA schemes increases and gradually stabilizes. This is due to the greater separation in angles, which enables the BS to identify more spatial dimensions for the fourth UE, allowing it to exploit additional subchannels for key generation. However, when the difference in angles becomes sufficiently large, near-field channels of the first and the fourth UE are uncorrelated, resulting in no further improvement in KR. In terms of KR fairness, the KR of the CVXPAKF exceeds that of the WFPA and the FNNPA schemes and then becomes lower than theirs. When the angle difference between the fourth and first UEs is small, the fourth UE is considered weaker due to the limited spatial dimensions available for extracting subchannels. Consequently, the proposed algorithms for KR fairness allocate more power to this weaker UE. As the angle difference increases, the fourth UE becomes stronger since its distance from the BS is shorter compared to the second and third UEs. Consequently, its transmit power may be reduced to benefit UEs with poorer channel quality.

## VIII. Conclusion

This paper studied the multi-user PLKG in near-field propagation scenarios. We designed a precoding scheme based on EVD and SVD methods and exploited the spatial dimensions of distances and angles in near-field scenarios to enable multi-user PLKG. We derived an analytical expression for the KR. Since the KR is non-convex, we approximated it and used EVD and SVD to convert precoding matrices into power allocation variables. To maximize the sum KR, we applied the Lagrange multiplier method to derive an analytical expression for the power allocation variables in terms of the Lagrange multiplier. Additionally, we designed a supervised FNN to learn the relationship between transmit power and the Lagrange multiplier. This approach allowed us to shift complex computations to offline training, thereby reducing the computational complexity during online inference. To address KR fairness, we employed the CXV toolbox to solve the optimization problem to improve the minimum KR. Moreover, we used the FNN to output power allocation variables specifically for KR fairness in high-power scenarios. Numerical results confirmed that even when two UEs share the same angles, the distance dimension in near-field scenarios can be effectively utilized for multi-user PLKG.

## References

[1] C.-X. Wang, X. You, X. Gao *et al.*, "On the road to 6G: Visions, requirements, key technologies and testbeds," *IEEE Commun. Surv. Tutor.*, vol. 25, no. 2, pp. 905 – 974, Feb. 2023.

[2] V.-l. Nguyen, P.-c. Lin, B.-c. Cheng *et al.*, "Security and privacy for 6G : A survey on prospective technologies and challenges," *IEEE Commun. Surveys Tuts.*, vol. 23, no. 4, pp. 2384–2428, Fourth Quarter 2021.

[3] J. Zhang, G. Li, A. Marshall *et al.*, "A new frontier for IoT security emerging from three decades of key generation relying on wireless channels," *IEEE Access*, vol. 8, pp. 138 406–138 446, Aug. 2020.

[4] G. Li, P. Staat, H. Li *et al.*, "RIS-Jamming: Breaking key consistency in channel reciprocity-based key generation," *IEEE Trans. Inf. Forensics Secur.*, vol. 19, pp. 5090–5105, Apr. 2024.

[5] E. Ronen, C. O'Flynn, A. Shamir *et al.*, "IoT goes nuclear: Creating a ZigBee chain reaction," in *Proc. IEEE Symposium on Security and Privacy*, San Jose, CA, USA, May 2017, pp. 195–212.

[6] H. Yang, Z. Li, C. Luo *et al.*, "InaudibleKey2.0: Deep learning-empowered mobile device pairing protocol based on inaudible acoustic signals," *IEEE/ACM Trans. Netw.*, no. 1, pp. 1–15, Jun. 2024.

[7] A. Chorti, A. N. Barreto, S. Köpsell *et al.*, "Context-aware security for 6G wireless: The role of physical layer security," *IEEE Commun. Standards Mag.*, vol. 6, no. 1, pp. 102–108, Mar. 2022.

[8] H. Lu, Y. Zeng, C. You *et al.*, "A tutorial on near-field XL-MIMO communications towards 6G," *IEEE Commun. Surv. Tutor.*, vol. 26, no. 4, pp. 2213–2257, Fourthquarter 2024.

[9] Z. Zhang, Y. Liu, Z. Wang *et al.*, "Physical layer security in near-field communications," *IEEE Trans. Veh. Technol.*, pp. 1–6, 2024.

[10] G. J. Anaya-Lopez, J. P. Gonzalez-Coma, and F. J. Lopez-Martinez, "Leakage subspace precoding and scheduling for physical layer security in multi-user XL-MIMO systems," *IEEE Communications Letters*, vol. 27, no. 2, pp. 467–471, Feb. 2023.

[11] J. Zhang, M. Ding, D. Lopez-Perez *et al.*, "Design of an efficient OFDMA-based multi-user key generation protocol," *IEEE Trans. Veh. Technol.*, vol. 68, no. 9, pp. 8842–8852, Sep. 2019.

[12] G. Li, C. Sun, E. A. Jorswieck *et al.*, "Sum secret key rate maximization for TDD multi-user massive MIMO wireless networks," *IEEE Trans. Inf. Forensics Security*, vol. 16, pp. 968–982, Sep. 2021.

[13] Z. Wu and L. Dai, "Multiple access for near-field communications: SDMA or LDMA?" *IEEE J. Sel. Areas Commun.*, vol. 41, no. 6, pp. 1918–1935, May 2023.

[14] L. Jiao, P. Wang, N. Wang *et al.*, "Efficient physical layer group key generation in 5G wireless networks," in *Proc. IEEE Conf. Commun. Netw. Secur. (CNS)*, Avignon, France, Aug. 2020, pp. 1–9.

[15] C. Sun and G. Li, "Power allocation and beam scheduling for multi-user massive MIMO secret key generation," *IEEE Access*, vol. 8, pp. 164 580–164 592, Sep. 2020.

[16] M. Cui and L. Dai, "Channel estimation for extremely large-scale MIMO: Far-Field or near-field?" *IEEE Trans. Commun.*, vol. 70, no. 4, pp. 2663–2677, Apr. 2022.

[17] H. Zhang, N. Shlezinger, F. Guidi *et al.*, "Beam focusing for multi-user MIMO communications," *IEEE Trans. Commun.*, vol. 21, no. 9, pp. 7476–7490, Sep. 2022.

[18] Z. Dong and Y. Zeng, "Near-field spatial correlation for extremely large-scale array communications," *IEEE Commun. Lett.*, vol. 26, no. 7, pp. 1534–1538, Apr. 2022.

[19] Z. Lu, Y. Han, S. Jin *et al.*, "Near-field localization and channel reconstruction for ELAA systems," *IEEE Trans. Wireless Commun.*, pp. 1–1, 2023.

[20] B. T. Quist and M. A. Jensen, "Optimal channel estimation in beamformed systems for common-randomness-based secret key establishment," *IEEE Trans. Inf. Forensics Security*, vol. 8, no. 7, pp. 558–559, Jul. 2013.

[21] X. He, H. Dai, W. Shen *et al.*, "Toward proper guard zones for link signature," *IEEE Trans. Wireless Commun.*, vol. 15, no. 3, pp. 2104–2117, Mar. 2016.

[22] J. Zhang, R. Woods, T. Q. Duong *et al.*, "Experimental study on key generation for physical layer security in wireless communications," *IEEE Access*, vol. 4, pp. 4464–4477, Aug. 2016.

[23] Y. Pan, C. Pan, S. Jin *et al.*, "RIS-aided near-field localization and channel estimation for the terahertz system," *IEEE J. Sel. Top. Signal Process.*, vol. 17, no. 4, pp. 878–892, Jul. 2023.

[24] S. Sigdel and W. A. Krzymień, "Simplified fair scheduling and antenna selection algorithms for multiuser MIMO orthogonal space-division multiplexing downlink," *IEEE Trans. Veh. Technol.*, vol. 58, no. 3, pp. 1329–1344, Mar. 2009.

[25] Q. Zhu and Y. Hua, "Optimal pilots for maximal capacity of secret key generation," in *Proc. IEEE GLOBECOM*, Hawaii, USA, Dec. 2019, pp. 1–6.

[26] C. Xing, Y. Jing, S. Wang *et al.*, "New viewpoint and algorithms for water-filling solutions in wireless communications," *IEEE Trans. Signal Process.*, vol. 68, pp. 1618–1634, Feb. 2020.

[27] T. Lu, L. Chen, J. Zhang *et al.*, "Reconfigurable intelligent surface-assisted key generation for millimetre-wave multi-user systems," *IEEE Trans. Inf. Forensics Secur.*, vol. 19, pp. 5373–5388, 2024.

[28] C. Chen, J. Zhang, T. Lu *et al.*, "Secret key generation for IRS-assisted multi-antenna systems: A machine learning-based approach," *IEEE Trans. Inf. Forensics Security*, vol. 19, pp. 1086–1098, Apr. 2024.

[29] L. Jiao, G. Sun, J. Le *et al.*, "Machine learning-assisted wireless PHY key generation with reconfigurable intelligent surfaces," in *Proc. ACM WiseML 2021*, Abu Dhabi, UAE, Jun. 2021, p. 61–66.

[30] H. Zhang, T. Zhou, T. Xu *et al.*, "FNN-based prediction of wireless channel with atmospheric duct," in *Proc. IEEE International Conference on Communications*, Montreal, Canada, Aug. 2021, pp. 1–6.

[31] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1," http://cvxr.com/cvx, Mar. 2014.

[32] G. Lee, H. Lee, J. Oh *et al.*, "Channel estimation for reconfigurable intelligent surface with a few active elements," *IEEE Trans. Veh. Technol.*, vol. 72, no. 6, pp. 8170–8174, Jun. 2023.

[33] S. Zhang and R. Zhang, "Capacity characterization for intelligent reflecting surface aided MIMO communication," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 8, pp. 1823–1838, Aug. 2020.


**Tianyu Lu** (Member, IEEE) received the B.S. and M.S. degrees in Electrical and Computer Engineering from Nanjing University of Posts and Telecommunications, in 2016 and 2019, respectively. He received the Ph.D. degree in Cyber Science and Engineering from Southeast University. He is currently a Postdoctoral Research Fellow with the Centre for Wireless Innovation (CWI), Queen's University Belfast. His research interests include physical layer security, key generation, and secure communications.

**Liquan Chen** (Senior Member, IEEE) received the Ph.D. degree from Southeast University, Nanjing, China, in 2005. He was a Post-Doctoral Researcher with Southeast University from 2005 to 2007, where he was an Associate Professor from 2008 to 2018. He was a Visiting Scholar with the National University of Singapore, Singapore, from 2011 to 2012. He is currently a Professor with the School of Cyber Science and Engineering, Southeast University. His research interests include information security, cryptography, and network security protocol.

**Junqing Zhang** (Senior Member, IEEE) received B.Eng and M.Eng degrees in Electrical Engineering from Tianjin University, China in 2009 and 2012, respectively, and a Ph.D. degree in Electronics and Electrical Engineering from Queen's University Belfast, UK in 2016. From Feb. 2016 to Jan. 2018, he was a Postdoctoral Research Fellow at Queen's University Belfast. From Feb. 2018 to Oct. 2022, he was a Tenure Track Fellow and then a Lecturer (Assistant Professor) at the University of Liverpool, UK. Since Oct. 2022, he has been a Senior Lecturer (Associate Professor) at the University of Liverpool. His research interests include the Internet of Things, wireless security, physical layer security, key generation, radio frequency fingerprint identification, and wireless sensing. Dr. Zhang is a co-recipient of the IEEE WCNC 2025 Best Workshop Paper Award. He is a Senior Area Editor of IEEE Transactions on Information Forensics and Security and an Associate Editor of IEEE Transactions on Mobile Computing.

**Trung Q. Duong** (Fellow, IEEE) is a Canada Excellence Research Chair (CERC) and a Full Professor at Memorial University, Canada. He is also an adjunct professor at Queen's University Belfast, UK, a visiting professor at Kyung Hee University, South Korea, and an adjunct professor at Duy Tan University, Vietnam. His current research interests include wireless communications, quantum machine learning, and quantum optimization.

He is the Editor-in-Chief of IEEE Communications Surveys & Tutorials and an IEEE ComSoc Distinguished Lecturer. He has received the two prestigious awards from the Royal Academy of Engineering (RAEng): RAEng Research Chair and the RAEng Research Fellow. He is the recipient of the prestigious Newton Prize 2017. He is a Fellow of the Engineering Institute of Canada (EIC), the Canadian Academy of Engineering (CAE), the Institution of Engineering and Technology (IET), and Asia-Pacific Artificial Intelligence Association (AAIA).