Sum Secret Key Rate Maximization for TDD Multi-User Massive MIMO Wireless Networks

Guyue Li, *Member, IEEE*, Chen Sun, *Member, IEEE*, Eduard A. Jorswieck, *Fellow, IEEE*, Junqing Zhang, Aiqun Hu, *Member, IEEE*, and You Chen, *Student Member, IEEE*

***Abstract*— Physical-layer key generation (PKG) based on channel reciprocity has recently emerged as a new technique to establish secret keys between devices. Most works focus on pairwise communication scenarios with single or small-scale antennas. However, the fifth generation (5G) wireless communications employ massive multiple-input multiple-output (MIMO) to support multiple users simultaneously, bringing serious overhead of reciprocal channel acquisition. This paper presents a multi-user secret key generation in massive MIMO wireless networks. We provide a beam domain channel model, in which different elements represent the channel gains from different transmit directions to different receive directions. Based on this channel model, we analyze the secret key rate and derive a closed-form expression under independent channel conditions. To maximize the sum secret key rate, we provide the optimal conditions for the Kronecker product of the precoding and receiving matrices and propose an algorithm to generate these matrices with pilot reuse. The proposed optimization design can significantly reduce the pilot overhead of the reciprocal channel state information acquisition. Furthermore, we analyze the security under the channel correlation between user terminals (UTs), and propose a low overhead multi-user secret key generation with non-overlapping beams between UTs. Simulation results demonstrate the near-optimal performance of the proposed precoding and receiving matrices design and the advantages of the non-overlapping beam allocation.**

***Index Terms*— Physical layer security, secret key generation, multi-user massive MIMO, beam domain.**

I. Introduction

PHYSICAL-LAYER key generation (PKG) has emerged as a promising technique to share the symmetric key for cryptographic applications [1]. Based on the reciprocity of the uplink and downlink channels, the communication ends, named Alice and Bob, can establish a pair of common channel information. When the channel information is converted into symmetric keys, Alice and Bob can use them for cryptography to safeguard data communication. The secret keys can be regularly updated, since the channel information varies randomly over time. Furthermore, due to channel decorrelation effect [2], any eavesdropper, named Eve, located half a wavelength away or more from Alice and Bob, will observe an uncorrelated channel information [3]. Thus, Eve cannot infer the secret key based on her own channel observations.

The key distribution is usually handled by the traditional public key cryptography techniques, which however have been facing a number of challenges to be applied in the future networks. Firstly, public key needs to be distributed to different devices in advance, and the key distribution for massive devices is complicated [4]. Secondly, the distributed key for each device usually does not update for a long time which may incur security issues. Thirdly, the public key cryptography may be compromised by the emerging quantum computers in the future [5]. Key generation can thus be a good alternative to complement when the public key cryptography is not suitable.

The PKG process generally contains four stages, namely channel probing, quantization, information reconciliation, and privacy amplification. At the beginning, Alice and Bob alternately transmit pilot signals to obtain correlated channel measurements. Then, they quantize these analog channel measurements into digital bits. Although the uplink and downlink channels are reciprocal, due to the calibration errors in uplink/downlink RF chains, the temporal variation of the channel and the noise, the measurements of the uplink and

Manuscript received January 28, 2020; revised June 12, 2020 and August 6, 2020; accepted September 14, 2020. Date of publication September 24, 2020; date of current version October 7, 2020. The work of Guyue Li was supported in part by the National Natural Science Foundation of China under Grant 61801115 and Grant 61941115, in part by the National Key Research and Development Program of China under Grant 2020YFE0200600, and in part by the Zhishan Youth Scholar Program of SEU under Grant 3209012002A3. The work of Chen Sun was supported in part by the National Natural Science Foundation of China under Grant 61901110 and in part by the Natural Science Foundation of Jiangsu Province under Grant BK20190334. The work of Eduard A. Jorswieck was supported by the German Research Foundation (DFG) under Project JO 801/25-1. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Georges Kaddoum. *(Corresponding author: Chen Sun.)*

Guyue Li is with the School of Cyber Science and Engineering, Southeast University, Nanjing 210096, China, and also with the Purple Mountain Laboratories for Network and Communication Security, Nanjing 210096, China (e-mail: guyuelee@seu.edu.cn).

Chen Sun and Aiqun Hu are with the National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China, and also with the Purple Mountain Laboratories for Network and Communication Security, Nanjing 210096, China (e-mail: sunchen@seu.edu.cn; aqhu@seu.edu.cn).

Eduard A. Jorswieck is with the Institute for Communications Technology, Technische Universitäet Braunschweig, 38106 Braunschweig, Germany (e-mail: Jorswieck@ifn.ing.tu-bs.de).

Junqing Zhang is with the Department of Electrical Engineering and Electronics, University of Liverpool, Liverpool L69 3GJ, U.K. (e-mail: junqing.zhang@liverpool.ac.uk).

You Chen is with the School of Cyber Science and Engineering, Southeast University, Nanjing 210096, China.

Digital Object Identifier 10.1109/TIFS.2020.3026466

downlink channel are not identical but highly correlated. Next, information reconciliation is used to enable Alice and Bob to agree on the same key through error detection protocols or error correction codes [6]. Finally, privacy amplification eliminates any potential information leakage to eavesdroppers.

Based on channel reciprocity, the channel probing stage shares the common random sources between legitimate users to generate the secret keys. Most PKG implementations are realized in the time division duplex (TDD) mode in order to utilize the channel reciprocity. Specifically, Alice and Bob alternately transmit the probing signals to estimate the channel state information (CSI), where the sampling time difference between them should be smaller than the channel coherence time, indicating that the coherence time limits the pilot overhead. As the pilot overhead scales linearly with the number of antennas, single antenna or small-scale multiple-input multiple-output (MIMO) communication systems, which are considered in most of the existing PKG schemes [7], have enough time and frequency resources to obtain the highly correlated CSI for pairwise communication scenarios.

The fifth generation (5G) and beyond communication systems employ massive MIMO technologies to support extremely high throughput and multiple users [8]. However, it is challenging to apply PKG with massive MIMO [9]. As the number of antennas is extremely large in massive MIMO systems, it is impractical for the base station (BS) and the user terminal (UT) to estimate the instantaneous uplink and downlink channel information within the coherence time. In some quasi-static scenarios, the coherence time may be long enough for pairwise channel estimations. However, the channel varies slowly such that the adjacent measurements are highly correlated, which will introduce redundancy and may finally result in failure of key generation. In the previous work, the self-correlation is eliminated by introducing signal preprocessing procedure after channel sounding, such as principal component analysis (PCA) [3]. This procedure also introduces great complexity due to the large data dimension.

Massive MIMO exploits spatial diversity and spatial signatures by allocating different beams/angles of transmitted signal to different directions of users, which enables multi-user communications. Key generation usually occurs between a pair of users. Exploiting massive MIMO for multi-user key generation will not be a straightforward extension from the pairwise key generation. This exploration is currently missing but very important as multi-user secret communications are very demanding.

This paper aims to address the above challenges by generating secret key with multiple users simultaneously in a narrow band massive MIMO system. In particular, we first state the problems for intuitively expanding existing key generation schemes in a multi-user massive MIMO scenario and then propose a new channel dimensionality reduction (CDR) based key generation scheme to address these problems. The main contributions of this paper are listed as follows:

- We propose a novel CDR-based key generation scheme exploiting sparse property of the beam domain channel model. Legitimate users only need to estimate the effective channels at a few dominate beams, therefore the pilot lengths and channel auto-correlations are largely reduced. Furthermore, we derive a closed-form expression of the secret key rate, considering other UTs as non-colluding curious users.
- We present an optimization approach to realize the maximal sum secret key rate under the pilot reuse case where different UTs transmit the identical pilot signals. Specifically, we design the precoding and receiving matrices to reduce the inter-user interference for multi-user communications. The proposed approach reduces the pilot overhead that scales with the number of UTs.
- We provide a security analysis considering the spatial channel correlations between UTs, and reveal that the channel information on the overlapping beams may cause serious information leakage and provide little secret key rate. Therefore, we propose a holistic multi-user secret key generation scheme, where the BS allocates non-overlapping beams to different UTs and multiple UTs can simultaneously generate secret keys with the BS using non-overlapping beams. Numerical results demonstrate the performance improvement of our proposed multi-user secret key generation scheme.

The material in this paper has been partially accepted by IEEE ICC 2020. In our previous work, we focus on the pilot reuse case, and proposed a beam-domain secret key generation approach which can reduce the channel dimension and the pilot overhead in a multi-user massive MIMO system. In this paper, we consider a general multi-user secret key generation framework and provide a general secret key rate, containing both the orthogonal pilot and the reused pilot cases. We prove the optimality of our proposed algorithm in the reused pilot case. Furthermore, we add a security analysis by considering the channel correlation and conduct that the BS employs non-overlapping beams to generate secret keys with different UTs.

We use the following notation throughout the paper: Upper (lower) bold-face letters denote matrices (column vectors); $\mathbf{I}$ denotes the identity matrix while its subscript, if needed, represents its dimensionality. Let $\mathbf{e}_i = [0, 0, \cdots, 0, 1, 0, \cdots, 0]$ denote a unit vector with the ith unit element and $\lambda_i(\mathbf{A})$ represent the ith sorted eigenvalue of matrix $\mathbf{A}$. The superscripts $(\cdot)^H, (\cdot)^T, (\cdot)^*$ stand for the conjugate-transpose, transpose, and conjugate of a matrix, respectively. We use $\mathbb{E}\{\cdot\}$ to denote ensemble expectation and $\mathrm{tr}(\cdot), \det(\cdot)$ to represent matrix trace and determinant operations, respectively. The $\mathrm{vec}(\cdot)$ operator stacks the columns of a matrix into a tall vector, and the symbol $\otimes$ denotes the Kronecker product of two matrices. The inequality $\mathbf{A} \succeq \mathbf{0}$ means that $\mathbf{A}$ is Hermitian positive semi-definite. We use $[\mathbf{A}]_{m,n}$ to denote the (m, n)-th element of matrix $\mathbf{A}$. The covariance matrix of combined random vectors is defined as $\mathcal{R}_{\mathbf{z}_1\mathbf{z}_2\cdots\mathbf{z}_{N_z}} = \mathbb{E}\{\mathbf{z}\mathbf{z}^H\}$, where $\mathbf{z} = \left[\mathbf{z}_1^T, \mathbf{z}_2^T, \cdots, \mathbf{z}_{N_z}^T\right]^T$.

II. Related Work

There are very few papers working on key generation with massive MIMO. The work in [10] employed new channel

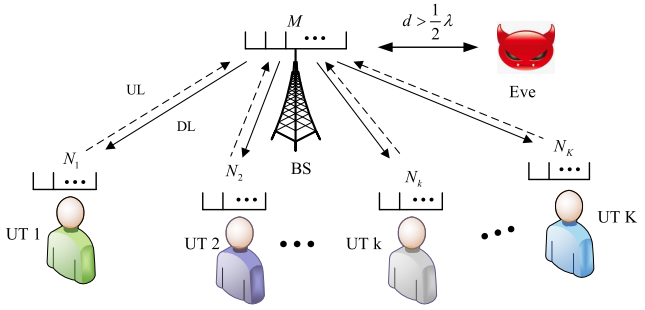


Fig. 1. System model of multi-user secret key generation.

characteristics, e.g., virtual angle of arrival (AoA) and angle of departure (AoD), to generate a shared secret key between two devices. Furthermore, Jiao *et al.* added a small perturbation angle into the AoA of the transmitter as the common randomness to improve the secret key rate constrained by low dynamic of the channel [11]. However, these works only focus on the AoA and AoD to generate the secret key, while the optimal design maximizing the secret key rate is missing.

While pairwise key generation has been extensively investigated, group and multi-user key generation yet receives less attention [1]. Note that, although both schemes have multiple users participating in the key generation process, we distinguish them in this paper. In the group key generation, all of the users share a common secret key. The multi-user key generation studied in this paper, refers to a particular case that each BS-UT pair has a different secret key. Among the existing group key generation protocols [12], [13], the majority of them still perform channel probing in a pairwise manner, resulting in an extremely large overhead and low efficiency. Hence, those works related to PKG among multiple nodes through the optimization of probing rates at individual node pair and channel probing schedule do not scale in this context [14]. In the multi-user key generation, Zhang *et al.* cleverly exploited the multi-user mechanism of OFDMA modulation by assigning non-overlapping subcarriers to different users [15]. However, there is no work exploiting the spatial diversity of massive MIMO to enable multi-user key generation.

Therefore, there is a clear need to investigate key generation with massive MIMO with special consideration to the multi-user applications, in both theoretical analysis on the secret key rate and the design of practical protocols.

III. System Model and Problem Statement

This paper considers a narrow-band star topology network, where a BS simultaneously generates secret keys $\kappa = \{\kappa_1, \kappa_2, \cdots, \kappa_K\}$ with K UTs, as shown in Fig. 1. The BS is equipped with M antennas and the kth UT is equipped with N_k antennas. In the 5G and beyond wireless communications, massive MIMO is the key technology to increase the transmission rate. Under the TDD operation, based on the reciprocal uplink and downlink channels, the BS generates the pairwise key κ_k with the kth UT.

We consider passive eavesdropping here and the active attacks are out of scope in this paper. Specifically, an eavesdropper wants to eavesdrop the secret key in the set of κ based on her own channel observations and all the information exchanged over the public channel. As a general assumption in PKG, we assume that Eve is located at least half wavelength away from BS and all of the UTs, because half wavelength is very close, e.g., 6.25 cm for 2.4GHz and Eve might get detected within a distance below that. Therefore, Eve's channel observations are assumed to be independent of that between the BS and UTs. Besides of Eve, we also consider the potential unintended hearing from other UTs. When two UTs are located close to each other, they may have correlated channel observations. These UTs are treated as curious users, i.e., each of them does not intend to eavesdrop keys of other users. They also do not collude with other UTs or Eve.

The secret keys are extracted from the reciprocal wireless channel information. Thus, we will first introduce the channel model and then state the problems and challenges in multi-user massive MIMO key generation.

A. Channel Model

We consider a geometric channel model with N_P paths. Then, the $N_k \times M$ physical MIMO channel matrix in the downlink associated with the pth path of the kth UT can be expressed as [16]

$$\mathbf{H}_{k,p}^{DL} = \alpha_{k,p}\mathbf{a}_{UT,k}(\theta_{k,p})\mathbf{a}_{BS}^{H}(\varphi_{k,p}), \tag{1}$$

where $\alpha_{k,p}$ is the complex gain of the pth path, $\mathbf{a}_{UT,k}(\theta_{k,p})$ is the UT antenna array response vector with the AoA $\theta_{k,p}$, and $\mathbf{a}_{BS}(\varphi_{k,p})$ is the BS antenna array response vector corresponding to AoD $\varphi_{k,p}$. Specifically, under the uniform linear array (ULA) setup, these vectors are given by

$$\begin{aligned}\mathbf{a}_{UT,k}(\theta_{k,p}) &= \frac{1}{\sqrt{N_k}}\Big[1, e^{-j\frac{2\pi}{\lambda}d\sin(\theta_{k,p})}, \ldots, \\ &\qquad e^{-j(N_k-1)\frac{2\pi}{\lambda}d\sin(\theta_{k,p})}\Big]^T \\ \mathbf{a}_{BS}(\varphi_{k,p}) &= \frac{1}{\sqrt{M}}\Big[1, e^{-j\frac{2\pi}{\lambda}d\sin(\varphi_{k,p})}, \ldots, \\ &\qquad e^{-j(M-1)\frac{2\pi}{\lambda}d\sin(\varphi_{k,p})}\Big]^T,\end{aligned} \tag{2}$$

where λ is the wavelength, and d is the distance between the adjacent antennas. For a narrow-band channel model, the channel response of the kth UT can be expressed as

$$\mathbf{H}_k^{DL} = \sum_{p=1}^{N_P}\mathbf{H}_{k,p}^{DL}. \tag{3}$$

The channel covariance matrices at the BS and the UT sides are defined as

$$\mathbf{R}_{BS,k} = \mathbb{E}\{(\mathbf{H}_k^{DL})^H\mathbf{H}_k^{DL}\}, \tag{4}$$

$$\mathbf{R}_{UT,k} = \mathbb{E}\{\mathbf{H}_k^{DL}(\mathbf{H}_k^{DL})^H\}. \tag{5}$$

In this paper, we assume the uplink and downlink channels are reciprocal, i,e,

$$\mathbf{H}_k^{UL} = (\mathbf{H}_k^{DL})^T, \tag{6}$$

but the channel estimations are affected by noise. The extension to channel non-reciprocity caused by the time difference and hardware imperfection is beyond the scope of this paper.

B. Problem Statement

An intuitive approach is extending existing pairwise PKG approaches via allocating orthogonal pilots among UTs. Firstly, in the downlink channel probing, the BS broadcasts the pilot signal, $\mathbf{S}^{DL} \in \mathbb{C}^{M\times M}$, and the received signal of the kth UT is given by

$$\mathbf{Y}_k^{DL} = \mathbf{H}_k^{DL}\mathbf{S}^{DL} + \mathbf{N}_k^{DL}, \tag{7}$$

where $\mathbf{N}_k^{DL} \in \mathbb{C}^{N_k\times M}$ is the Gaussian noise at UT k. To estimate the perfect CSI, the pilot signal should satisfy the orthogonality, i.e., $\mathbf{S}^{DL}(\mathbf{S}^{DL})^H = \mathbf{I}_M$. By the least square (LS) estimation, UT k estimates his downlink CSI as

$$\mathbf{Z}_k^{DL} = \mathbf{Y}_k^{DL}(\mathbf{S}^{DL})^H = \mathbf{H}_k^{DL} + \mathbf{N}_k^{DL}(\mathbf{S}^{DL})^H. \tag{8}$$

Next, in the uplink channel probing, all the UTs send the pilot signals, $\mathbf{S}_k^{UL} \in \mathbb{C}^{N_k\times N}, k \in \{1, 2, \cdots, K\}$ to the BS simultaneously, where $N = \sum_k N_k$ is the total number of UTs' antennas. The received signal of BS is

$$\mathbf{Y}^{UL} = \sum_{k=1}^{K}\mathbf{H}_k^{UL}\mathbf{S}_k^{UL} + \mathbf{N}^{UL}, \tag{9}$$

where $\mathbf{N}^{UL} \in \mathbb{C}^{N_k\times N}$ is the Gaussian noise at the BS. The BS estimates the uplink CSI of UT k as

$$\mathbf{Z}_k^{UL} = \mathbf{Y}^{UL}(\mathbf{S}_k^{UL})^H = \mathbf{H}_k^{UL}\mathbf{S}_k^{UL}(\mathbf{S}_k^{UL})^H + \mathbf{H}_k^{DL}\sum_{k'\neq k}^{K}\mathbf{S}_{k'}^{UL}(\mathbf{S}_k^{UL})^H + \mathbf{N}^{UL}(\mathbf{S}_k^{UL})^H. \tag{10}$$

To distinguish each UT, the pilot signals of different UTs are designed to satisfy the orthogonality requirement, i.e.,

$$\mathbf{S}_i^{UL}(\mathbf{S}_j^{UL})^H = \begin{cases}\mathbf{I}_{N_i}, & i = j\\ \mathbf{0}_{N_i\times N_j}, & i \neq j.\end{cases} \tag{11}$$

Therefore, (10) can be further reduced to

$$\mathbf{Z}_k^{UL} = \mathbf{H}_k^{UL} + \mathbf{N}^{UL}(\mathbf{S}_k^{UL})^H. \tag{12}$$

When $\mathbf{S}^{DL}$ and $\mathbf{S}_k^{UL}$ are unitary, the noise $\mathbf{N}_k^{DL}(\mathbf{S}^{DL})^H$ and $\mathbf{N}^{UL}(\mathbf{S}_k^{UL})^H$ have the same distribution as $\mathbf{N}_k^{DL}$ and $\mathbf{N}^{UL}$ because Gaussian distribution is isotropic and thereby unitarily invariant.

According to (6), the BS and UT k obtain very similar channel estimations of $\mathbf{Z}_k^{UL} \approx \mathbf{Z}_k^{DL}$. Then, they vectorize the estimations $\mathbf{z}_k^{UL} = \mathrm{vec}(\mathbf{Z}_k^{UL})$ and $\mathbf{z}_k^{DL} = \mathrm{vec}(\mathbf{Z}_k^{DL})$, which are chosen as channel characteristics for key generation. By employing quantization, information reconciliation and privacy amplification, the BS and UT k finally generate consistent secret key κ_k.

However, this intuitive approach has two issues as follows.

1) Define the duration of one round of channel probing as $T_p = T_D + T_U + T_{Switch}$, where T_{Switch} is the switching time from downlink to uplink, T_D and T_U are the pilot transmission time in the downlink and the uplink, respectively. This time needs to be deliberately kept smaller than the channel coherence time, so that BS and UTs can obtain highly correlated CSI in a TDD system. However, in this case, $T_p = (M + \sum_k N_k)\Delta T + T_{Switch}$, where ΔT is the symbol transmission time. To distinguish different antennas of different UTs, the length of uplink pilots scales with the number of antennas N_k as well as the number of UTs K. When M, K and N_k are large, it becomes very challenge to accomplish channel probing within the coherence time.
2) Because of the spatial correlation of the antennas, the elements of $\mathbf{z}_k^{UL}$ and $\mathbf{z}_k^{DL}$ are highly auto-correlated, resulting to long 0s and 1s in the quantized bit sequences. Traditionally, preprocessing approaches, e.g., PCA, are used to reduce the auto-correlation. However, due to the large scale of antennas at both BS and UTs in the future wireless communications, it is complicated to perform PCA algorithm for $\mathbf{z}_k^{UL}$ and $\mathbf{z}_k^{DL}$ with a large dimension of MN_k.

To sum up, the core problems are how to reduce the length of pilots and the high dimension of channel matrix in the multi-user massive MIMO system. Fortunately, literature and field measurements have shown that the beam domain channel matrix reveals the sparse property in typical scenarios [17], [18]. Hence, we propose a new channel dimensionality reduction (CDR)-based key generation scheme to address the above problems.

IV. General CDR-Based Key Generation Scheme

In massive MIMO channels, a few dominant elements contain the most relevant channel information. To reduce the dimensions, we first introduce the beam domain transform and then propose the corresponding key generation scheme. The achievable secret key rate in the proposed scheme is also derived.

A. Beam Domain Transform

Beam domain transform samples the original physical channel by two series of uniformly distributed beams/angles over $[0, 2\pi]$, i.e., transmitting and receiving beams/angles. According to [19], the beam domain channel is

$$\tilde{\mathbf{H}}_k^{DL} = \mathbf{A}_{UT,k}^H\mathbf{H}_k^{DL}\mathbf{A}_{BS}, \tag{13}$$

where

$$\mathbf{A}_{UT,k} = \left[\mathbf{a}_{UT,k}(\theta_1), \mathbf{a}_{UT,k}(\theta_2), \ldots, \mathbf{a}_{UT,k}(\theta_{N_k})\right] \in \mathbb{C}^{N_k\times N_k} \tag{14}$$

and

$$\mathbf{A}_{BS} = [\mathbf{a}_{BS}(\varphi_1), \mathbf{a}_{BS}(\varphi_2), \ldots, \mathbf{a}_{BS}(\varphi_M)] \in \mathbb{C}^{M\times M} \tag{15}$$

are the sampling matrices at the kth UT and the BS, respectively. They satisfy that $\mathbf{A}_{UT,k}^H\mathbf{A}_{UT,k} = \mathbf{I}$, $\mathbf{A}_{BS}^H\mathbf{A}_{BS} = \mathbf{I}$. The (n, m)-th element of $\tilde{\mathbf{H}}_k^{DL}$ represents the channel gains from AoD φ_m to AoA θ_n, where φ_m and θ_n are the mth and nth sample angles, which satisfy that $\sin(\varphi_m) = 2m/M - 1$ and $\sin(\theta_n) = 2n/N_k - 1$ [20]. When the antenna spacing is half wavelength, i.e., $d = \lambda/2$, the matrices $\mathbf{A}_{UT}$ and $\mathbf{A}_{BS}$ become

the unitary discrete Fourier transform (DFT) matrix, defined as [21]

$$[\mathbf{A}_{UT,k}]_{n_1,n_2} = \frac{1}{\sqrt{N_k}} \exp(-j2\pi(n_1-1)(n_2-N_k/2)/N_k)$$
$$[\mathbf{A}_{BS}]_{m_1,m_2} = \frac{1}{\sqrt{M}} \exp(-j2\pi(m_1-1)(m_2-M/2)/M). \tag{16}$$

When the number of antennas tends to infinity, the beam domain channel $\tilde{\mathbf{H}}_k^{DL}$ exhibits the spatial resolution as follows [18].

Proposition 1: When the number of antennas grows to infinity, the beam domain channel $\tilde{\mathbf{H}}_k^{DL}$ tends to $\mathbf{G}_k^{DL}$, i.e., for arbitrary n and m,

$$\lim_{M,N_k\to\infty} [\tilde{\mathbf{H}}_k^{DL} - \mathbf{G}_k^{DL}]_{n,m} = 0, \tag{17}$$

where $\mathbf{G}_k^{DL} \in \mathbb{C}^{N_k\times M}$ is given by

$$[\mathbf{G}_k^{DL}]_{n,m} = \sum_{p=1}^{N_P} \alpha_{k,p}\delta(\theta_{k,p} - \arcsin(2n/N_k - 1)) \times\delta(\varphi_{k,p} - \arcsin(2m/M - 1)). \tag{18}$$

Proof: See Appendix A. ■

Remark 1: From the definition of $\mathbf{G}_k^{DL}$ in (18), for each n and m, there is at most one path p simultaneously satisfying $\theta_{k,p} = \arcsin(2n/N_k - 1)$ and $\varphi_{k,p} = \arcsin(2m/M - 1)$. This indicates that one element in $\mathbf{G}_k^{DL}$ represents channel gains from one AoD $\varphi_{k,p}$ to one AoA $\theta_{k,p}$ and different elements represent channel gains corresponding to different AoAs and AoDs. As there are N_P paths, the number of non-zero entries in matrix $\mathbf{G}_k^{DL}$ is N_P. When the BS and UT k are equipped with a large (but finite) number of antennas, the beam domain channel matrix $\tilde{\mathbf{H}}_{k,p}^{DL}$ can be approximated by $\mathbf{G}_k^{DL}$. In this case, $\tilde{\mathbf{H}}_k^{DL}$ is a very sparse matrix with N_P dominant elements corresponding to the paths. Moreover, these elements become independent with each other as long as these paths are independent.

Define the beam domain channel covariance matrices at the BS and kth UT as

$$\tilde{\mathbf{R}}_{BS,k} = \mathbb{E}\{(\tilde{\mathbf{H}}_k^{DL})^H\tilde{\mathbf{H}}_k^{DL}\} = \mathbf{A}_{BS}^H\mathbf{R}_{BS,k}\mathbf{A}_{BS},$$
$$\tilde{\mathbf{R}}_{UT,k} = \mathbb{E}\{\tilde{\mathbf{H}}_k^{DL}(\tilde{\mathbf{H}}_k^{DL})^H\} = \mathbf{A}_{UT,k}^H\mathbf{R}_{UT,k}\mathbf{A}_{UT,k}, \tag{19}$$

respectively. When the number of antennas grows to infinity, $\tilde{\mathbf{R}}_{BS,k}$ and $\tilde{\mathbf{R}}_{UT,k}$ tend to diagonal matrices with the diagonal elements given by

$$\lim_{M\to\infty} [\tilde{\mathbf{R}}_{BS,k}]_{m,m} - \sum_{p=1}^{N_P} |\alpha_{k,p}|^2\delta(\varphi_{k,p} - \arcsin(2m/M - 1)) = 0,$$
$$\lim_{N_k\to\infty} [\tilde{\mathbf{R}}_{UT,k}]_{n,n} - \sum_{p=1}^{N_P} |\alpha_{k,p}|^2\delta(\theta_{k,p} - \arcsin(2n/N_k - 1)) = 0. \tag{20}$$

The mth diagonal element in $\tilde{\mathbf{R}}_{BS,k}$ represents the channel gains of the mth transmit beam ($\varphi_{k,p} = \arcsin(2m/M - 1)$), and the nth diagonal element in $\tilde{\mathbf{R}}_{UT,k}$ represents the channel gains of the nth receive beam ($\theta_{k,p} = \arcsin(2n/N - 1)$). The beam domain channel covariance matrices also reveal the angular resolution of the channel gains.

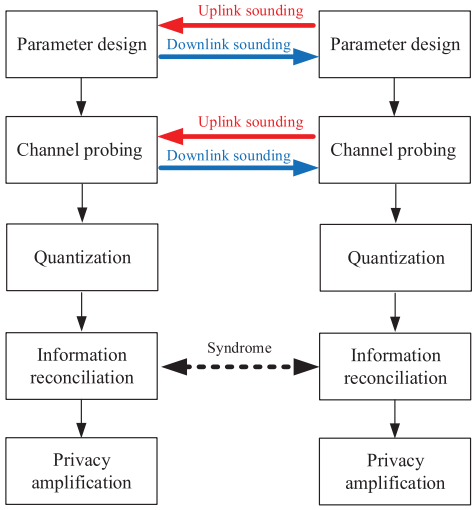


Fig. 2. Flow diagram of CDR based secret key generation scheme.

From the above analysis, one can observe that the representation in the beam domain channel brings the following benefits. Firstly, in the beam domain, the channel matrix reveals the sparse property, i.e., only a few elements contain the most channel information, which reduces the dimension of channel estimation. Secondly, as the number of antennas at the BS and UT increases, the elements of the channel matrix become mutually independent, reducing the redundancy, which is particularly desirable in secret key generation. Thirdly, with a large number of antennas at the BS, the beam domain transform matrix at the BS $\mathbf{A}_{BS}$ is independent of UTs and we can use one identical matrix to transform channel matrices of different UTs into the beam domain. Such property is desirable in multi-user secret key generation.

B. Key Generation Scheme Based on CDR

In the beam domain, only a few elements contain the most channel information, which motivates us to propose a general framework for multi-user secret key generation, as portrayed in Fig. 2. The UT and the BS transmit sounding signals to acquire the statistical CSI. During the parameter design stage, the BS and UTs design precoding and receiving matrices based on the statistical CSI, in order to reduce the dimension of channel estimation. Then, the BS and UTs estimate the effective channel parameters with high correlations. After quantization, the information reconciliation and privacy amplification procedures are used to generate consistent and private secret keys, similarly with the point-to-point secret key generation. Information reconciliation and privacy amplification are thus not particularly designed or optimized in this paper.

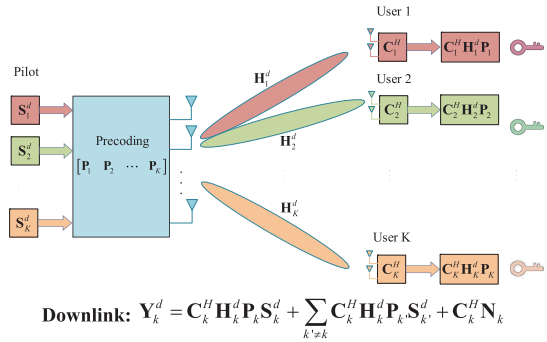
In this paper, we focus on the first two steps, i.e., parameter design and channel probing, which are relatively different from those in the point-to-point secret key generation.

1) *Parameter design:* In this step, BS and UTs design the precoding and receiving matrices according to their statistical CSI. Firstly, in the uplink, each UT employs one antenna to transmit the sounding signals. Then, the BS estimates the covariance matrix and designs the precoding matrix $\mathbf{P}_k$ with equal power allocation $\mathbf{P}_k^H\mathbf{P}_k = \mathbf{I}$. Next, in the downlink, the BS employs the precoding matrix to transmit the downlink sounding signals and UTs estimate the statistical CSI information of the covariance matrix at the UT side and design the receiving matrix $\mathbf{C}_k$ satisfying $\mathbf{C}_k^H\mathbf{C}_k = \mathbf{I}$.
2) *Channel probing:*In this step, BS and UTs probe the channel alternatively and construct the reciprocal channel characteristics with the help of the precoding and receiving matrices. Firstly, the BS transmits the downlink pilot signals by the precoding matrix $\mathbf{P}$ and UTs preprocess the received signals by the matrix $\mathbf{C}^H$ to obtain the reciprocal channel parameters. Next, each UT employs the matrix $\mathbf{C}^*$ to transmit the pilot signals. The BS utilizes the precoding matrix $\mathbf{P}$ to preprocess the received signals and estimate the effective channel.
3) Quantization: After channel probing, using channel quantization alternating (CQA) scheme [22], the BS and each UT quantize the effective channel measurement to generate the initial secret keys.

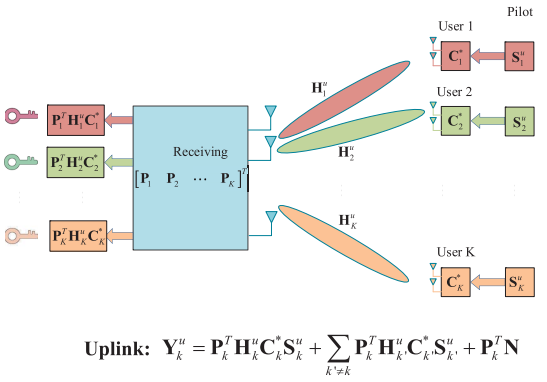
Remark 2: The parameter design is proposed specifically for multi-user massive MIMO secret key generation, which is used to reduce the channel estimation dimension, as well as the inter-user interference. The dimensions of $\mathbf{P}_k$ and $\mathbf{C}_k^H$ are $M \times M_e$ and $N_e \times N_k$ respectively, where M_e and N_e are the reduced dimensions at the BS and UT, which are approximately equal to the number of paths N_P, far smaller than M and N_k. Then, we only need to estimate the effective channel with $N_e \times M_e$, significantly reducing the channel estimation dimension. Moreover, from the analysis of channel characteristics, one can observe that the channel gains of one UT are concentrated within a few beams (directions), which has the potential to separate different UTs by different beams. In addition, the precoding and receiving matrices are determined by the statistical CSI, which can be obtained by some time and frequency resources [18]. Once the parameter design is completed, the BS and UTs can perform multiple channel probing rounds. Each channel probing round, including the uplink and downlink channel sounding, should be completed within one coherence time slot, where the instantaneous CSI keeps constant [23]. Different channel probing rounds operate in different channel coherence time slots, and the instantaneous CSI varies along time, resulting in the variation of the generated secret keys.

The downlink and uplink probing process is illustrated in Fig. 3. Let $\mathbf{P} = [\mathbf{P}_1, \mathbf{P}_2, \cdots, \mathbf{P}_K]$ and $\mathbf{C}^H = [\mathbf{C}_1^H, \mathbf{C}_2^H, \cdots, \mathbf{C}_K^H]$ denote the precoding and receiving matrices in the downlink transmission, respectively. Then, $\mathbf{C}^*$ and $\mathbf{P}^T$ are used as precoding and receiving matrices in the uplink transmission.

Specifically, let $\mathbf{S}_k^{DL} \in \mathbb{C}^{M_e \times T_D}$ denote the downlink pilot from BS to UT k within T_D time slots, which satisfies $\mathbf{S}_k^{DL}(\mathbf{S}_k^{DL})^H = \mathbf{I}$. After multiplying the pilot signals by the



(a) Downlink



(b) Uplink

Fig. 3. Channel probing in secret key generation scheme.

precoding matrix $\mathbf{P}_k$, the BS transmits the summation of all the UTs. Then, UT k multiplies the received signal by the receiving matrix $\mathbf{C}_k^H$, given by

$$\mathbf{Y}_k^{DL} = \mathbf{C}_k^H\mathbf{H}_k^{DL}\mathbf{P}_k\mathbf{S}_k^{DL} + \mathbf{C}_k^H\mathbf{H}_k^{DL}\sum_{k'\neq k}\mathbf{P}_{k'}\mathbf{S}_{k'}^{DL} + \mathbf{C}_k^H\mathbf{N}_k. \tag{21}$$

By the LS estimation, UT k estimates the downlink CSI as

$$\mathbf{Z}_k^{DL} = \mathbf{Y}_k^{DL}(\mathbf{S}_k^{DL})^H = \mathbf{C}_k^H\mathbf{H}_k^{DL}\mathbf{P}_k + \mathbf{C}_k^H\mathbf{H}_k^{DL}\sum_{k'\neq k}\mathbf{P}_{k'}\mathbf{S}_{k'}^{DL}(\mathbf{S}_k^{DL})^H + \mathbf{C}_k^H\mathbf{N}_k(\mathbf{S}_k^{DL})^H. \tag{22}$$

In the uplink transmission, let $\mathbf{S}_k^{UL} \in \mathbb{C}^{N_e \times T_U}$ denote the pilot transmitted by UT k within T_U time slots, satisfying $\mathbf{S}_k^{UL}(\mathbf{S}_k^{UL})^H = \mathbf{I}$. The kth UT transmits pilot signals by the matrix $\mathbf{C}_k^*$, and the BS receives the summation of all the UTs' signals. Then, multiplying by the receiving matrix $\mathbf{P}_k^T$, which is transpose of the precoding matrix in the downlink, the received signal of UT k at the BS can be expressed as

$$\mathbf{Y}_k^{UL} = \mathbf{P}_k^T\mathbf{H}_k^{UL}\mathbf{C}_k^*\mathbf{S}_k^{UL} + \mathbf{P}_k^T\sum_{k'\neq k}\mathbf{H}_{k'}^{UL}\mathbf{C}_{k'}^*\mathbf{S}_{k'}^{UL} + \mathbf{P}_k^T\mathbf{N} \tag{23}$$

where $\mathbf{Y}_k^{UL} \in \mathbb{C}^{M_e \times T_U}$ is the received signals of time length T_U. By employing the LS estimation, the estimated effective channel of UT k can be expressed as

$$\mathbf{Z}_k^{UL} = \mathbf{Y}^{UL}(\mathbf{S}_k^{UL})^H = \mathbf{P}_k^T\mathbf{H}_k^{UL}\mathbf{C}_k^* + \mathbf{P}_k^T\sum_{k'\neq k}\mathbf{H}_{k'}^{UL}\mathbf{C}_{k'}^*\mathbf{S}_{k'}^{UL}(\mathbf{S}_k^{UL})^H + \mathbf{P}_k^T\mathbf{N}(\mathbf{S}_k^{UL})^H. \tag{24}$$

Remark 3: In the uplink and downlink transmissions, the BS and UTs vectorize the estimated effective channel matrices $\mathbf{Z}_k^{UL}$ and $\mathbf{Z}_k^{DL}$ and employ them to generate the secret key, where the reciprocal component between the BS and UT k is $\mathbf{C}_k^H \mathbf{H}_k^{DL} \mathbf{P}_k$ with a small dimension of $N_e \times M_e$. In this way, the dimension of channel characteristics is reduced by $\eta = \frac{M \times N_k}{N_e \times M_e}$ times. The dimensions of M_e and N_e are very small compared with the number of antennas, therefore the reduction η is very significant. In addition, the duration of one round of channel probing T_p is reduced from $(M + \sum_k N_k)\Delta T + T_{Switch}$ to $(M_e + N_e)\Delta T + T_{Switch}$.

Let $\mathbf{S}_{k'k}^{DL} = \mathbf{S}_{k'}^{DL}(\mathbf{S}_k^{DL})^H$ (or $\mathbf{S}_{k'k}^{UL} = (\mathbf{S}_{k'}^{UL}(\mathbf{S}_k^{UL})^H)^T$) represent the covariance matrix of the downlink (or uplink) pilot signals, where $\mathbf{S}_{kk}^{DL} = \mathbf{S}_{kk}^{UL} = \mathbf{I}$. Moreover, as the uplink and downlink channel are reciprocal, for the simplicity of notation, the downlink channel $\mathbf{H}_k^{DL}$ is denoted as $\mathbf{H}_k$, and the uplink channel is $\mathbf{H}_k^{UL} = (\mathbf{H}_k)^T$. Then, the vectorized channel matrices can be expressed as

$$\mathbf{z}_k^{DL} = \mathrm{vec}(\mathbf{Z}_k^{DL}) = \sum_{k'} \left(\left(\mathbf{P}_{k'} \mathbf{S}_{k'k}^{DL} \right)^T \otimes \mathbf{C}_k^H \right) \mathrm{vec}(\mathbf{H}_k) + \mathrm{vec}(\mathbf{C}_k^H \mathbf{N}_k) \tag{25}$$

$$\mathbf{z}_k^{UL} = \mathrm{vec}((\mathbf{Z}_k^{UL})^T) = \sum_{k'} \left(\mathbf{P}_k^T \otimes \mathbf{S}_{k'k}^{UL} \mathbf{C}_{k'} \right) \mathrm{vec}(\mathbf{H}_{k'}) + \mathrm{vec}(\mathbf{N}^T \mathbf{P}_k). \tag{26}$$

Although the uplink and downlink channel of the kth UT are identical, the interference terms between UTs are not reciprocal. Specifically, in the downlink transmission, the interference received at the kth UT is the summation of the transmitted signals of the k'th UT propagating through the channel of the kth UT $\mathbf{H}_k$, while in the uplink transmission, the interference received at the BS for the kth UT is the summation of the transmitted signals of the k'th UT propagating through the channel of different UTs $\mathbf{H}_{k'}$. Thus, the interference will also reduce the agreement.

C. Secret Key Rate

When the BS communicates with one UT, other UTs are potential non-colluding curious users.[1] Under the TDD operation, each UT cannot transmit and receive signals at the same time. The ith UT only has the channel observation in the downlink transmission. Thus, the key rate is the minimum mutual information given other UT's observations. The number of secure bits for the link from the BS to UT k in the mutual information can be expressed as [24]

$$I_k = \min_{i \neq k} I(\mathbf{z}_k^{DL}; \mathbf{z}_k^{UL} | \mathbf{z}_i^{DL}). \tag{27}$$

Remark 4: We assume that the distance between the BS and each UT is several orders of magnitude larger than the wavelength, i.e., there is no UT close to the BS. Then, in the uplink transmission, the channel from one UT to the BS is independent of that from one UT to another UT. Moreover, as UTs transmit signals at the same time and frequency block, they cannot receive signals from other UTs. Thus, the secret key rate is the minimum mutual information between $\mathbf{z}_k^{DL}$ and $\mathbf{z}_k^{UL}$ on the condition of $\mathbf{z}_i^{DL}$.

When the channel estimations of different UTs are independent, the secret key rate degrades to

$$I_k = I(\mathbf{z}_k^{DL}; \mathbf{z}_k^{UL}). \tag{28}$$

In massive MIMO communications, when the beam domain channels of different UTs are non-overlapping, i.e., the channel covariance matrices at the BS are orthogonal, given by

$$\tilde{\mathbf{R}}_{BS,k} \tilde{\mathbf{R}}_{BS,i} = \mathbf{0}, \quad k \neq i, \tag{29}$$

the channel vectors of UT k and UT i are independent. Then, the secret rate $I(\mathbf{z}_k^{DL}; \mathbf{z}_k^{UL} | \mathbf{z}_i^{DL})$ can be degraded as $I(\mathbf{z}_k^{DL}; \mathbf{z}_k^{UL})$ and no secret keys are leaked to potential curious UTs [25].

When the beam domain channels are overlapping, we should consider the information leakage to other UTs. But we can always select non-overlapping beams for different UTs and then the selected channel information is independent. Thus, we can also use (28) to calculate the secret key rate. The overlapping case will be discussed in more detail in Section V-B.

Denote the precoding and receiving matrices in the beam domain as $\tilde{\mathbf{P}}_k = \mathbf{A}_{BS}^H \mathbf{P}_k$ and $\tilde{\mathbf{C}}_k = \mathbf{A}_{UT,k}^H \mathbf{C}_k$, respectively. Let $\mathbf{V}_k = \mathbf{\Lambda}_k^{1/2} \left(\sum_{k'} (\tilde{\mathbf{P}}_{k'} \mathbf{S}_{k'k}^{DL})^T \otimes \tilde{\mathbf{C}}_k^H \right)^H$ and $\mathbf{V}_{kk'} = \mathbf{\Lambda}_{k'}^{1/2} \left(\tilde{\mathbf{P}}_k^T \otimes \mathbf{S}_{k'k}^{UL} \tilde{\mathbf{C}}_{k'}^H \right)^H$, where $\mathbf{\Lambda}_k = \mathbb{E}\{\mathrm{vec}(\tilde{\mathbf{H}}_k)\mathrm{vec}(\tilde{\mathbf{H}}_k)^H\}$ is the full correlation of the beam domain channel. We can compute the secret key rate of UT k as follows.

Theorem 1: When the channels of different UTs become independent, the secret key rate of the kth UT is given by

$$I(\mathbf{z}_k^{DL}; \mathbf{z}_k^{UL}) = -\log\det\left(\mathbf{I} - \mathbf{V}_{kk} \left(\sum_{k'} \mathbf{V}_{kk'}^H \mathbf{V}_{kk'} + \left(\mathbf{P}_k^T \mathbf{P}_k^* \otimes \mathbf{I}_{T_U} \right) \right)^{-1} \mathbf{V}_{kk}^H \times \mathbf{V}_k \left(\mathbf{V}_k^H \mathbf{V}_k + \mathbf{I}_{T_D} \otimes \mathbf{C}_k^H \mathbf{C}_k \right)^{-1} \mathbf{V}_k^H \right). \tag{30}$$

Proof: See Appendix B. ∎

Remark 5: Note that when the number of UTs is one (i.e., there is only one UT), the secret key rate reduced to the single user secret key rate, which is a special case of (30). Moreover, the secret key rate (30) is complicated, which depends on the precoding and receiving matrices, as well as the covariance matrices of pilot signals. To estimate the effective CSI, the pilot signals of one UT should be orthogonal, i.e., $\mathbf{S}_{kk}^{DL} = \mathbf{I}$ and $\mathbf{S}_{kk}^{UL} = \mathbf{I}$ [26]. When the pilot signals between UTs are orthogonal, i.e., $\mathbf{S}_{k'k}^{DL} = \mathbf{0}$ and $\mathbf{S}_{k'k}^{UL} = \mathbf{0}$, there is no interference between UTs. The secret key rate can

[1] This paper focuses on the non-colluding scenario, where no information is shared among the curious users. The colluding case, where curious users share their received signals with each other, can be studied in the future.

be simplified as

$$
\begin{aligned}
&I(\mathbf{z}_k^{DL};\mathbf{z}_k^{UL})\\
&\quad=-\log\det\Big(\mathbf{I}-\mathbf{\Lambda}_k^{1/2}\Big(\tilde{\mathbf{P}}_k^{*}\otimes\tilde{\mathbf{C}}_k\Big)\\
&\qquad\times\Big(\mathbf{I}+\Big(\tilde{\mathbf{P}}_k^{T}\otimes\tilde{\mathbf{C}}_k^{H}\Big)\mathbf{\Lambda}_k\Big(\tilde{\mathbf{P}}_k^{*}\otimes\tilde{\mathbf{C}}_k\Big)\Big)^{-1}\Big(\tilde{\mathbf{P}}_k^{T}\otimes\tilde{\mathbf{C}}_k^{H}\Big)\mathbf{\Lambda}_k\\
&\qquad\times\Big(\tilde{\mathbf{P}}_k^{*}\otimes\tilde{\mathbf{C}}_k\Big)\Big(\mathbf{I}+\Big(\tilde{\mathbf{P}}_k^{T}\otimes\tilde{\mathbf{C}}_k^{H}\Big)\mathbf{\Lambda}_k\Big(\tilde{\mathbf{P}}_k^{*}\otimes\tilde{\mathbf{C}}_k\Big)\Big)^{-1}\\
&\qquad\times\Big(\tilde{\mathbf{P}}_k^{T}\otimes\tilde{\mathbf{C}}_k^{H}\Big)\mathbf{\Lambda}_k^{1/2}\Big).
\end{aligned}
\tag{31}
$$

However, for the orthogonal pilots between UTs, the pilot overhead scales with the number of UTs, which is quite large in multi-user communication systems. In general, due to the short coherent time, employing the orthogonal pilots between users is impractical. Alternatively, pilot signals can be reused between UTs, i.e., $\mathbf{S}_{k'k}^{DL}=\mathbf{I}$ and $\mathbf{S}_{k'k}^{UL}=\mathbf{I}$. Under this condition, there exists inter-user interference and we will design the precoding and receiving matrices to reduce it.

V. Optimization Design With Pilot Reuse

In this section, we consider the CDR-based secret key generation scheme design under the pilot reuse case, where different UTs transmit the identical pilot signals. In this case, we first design the precoding and receiving matrices maximizing the secret key rate and then analyze the security when the channels of different UTs are correlated.

A. Design of Precoding and Receiving Matrices

Under the pilot reuse case, the inter-user interference will affect secret key agreement. Therefore, we need to design the precoding and receiving matrices to maximize the sum secret key rate. Generally, the precoding and receiving matrices contain the transmit directions as well as the transmitted power on each direction. To reduce the interference, we focus on the transmit direction design and consider the equal power allocation of each direction, which can be expressed as

$$
\begin{aligned}
&\max_{\tilde{\mathbf{P}}_k,\tilde{\mathbf{C}}_k}\ R_{\text{sum}}=\sum_k I(\mathbf{z}_k^{DL};\mathbf{z}_k^{UL})\\
&\text{s.t. }\tilde{\mathbf{P}}_k^{H}\tilde{\mathbf{P}}_k=\mathbf{I}\\
&\quad\tilde{\mathbf{C}}_k^{H}\tilde{\mathbf{C}}_k=\mathbf{I},
\end{aligned}
\tag{32}
$$

where the secret key rate is calculated as

$$
\begin{aligned}
&I(\mathbf{z}_k^{DL};\mathbf{z}_k^{UL})\\
&\quad=-\log\det\Big(\mathbf{I}-\mathbf{\Lambda}_k^{1/2}\Big(\tilde{\mathbf{P}}_k^{T}\otimes\tilde{\mathbf{C}}_k^{H}\Big)^{H}\\
&\qquad\times\Big(\mathbf{I}+\sum_{k'}\Big(\tilde{\mathbf{P}}_k^{T}\otimes\tilde{\mathbf{C}}_{k'}^{H}\Big)\mathbf{\Lambda}_{k'}\Big(\tilde{\mathbf{P}}_k^{T}\otimes\tilde{\mathbf{C}}_{k'}^{H}\Big)^{H}\Big)^{-1}\\
&\qquad\times\Big(\tilde{\mathbf{P}}_k^{T}\otimes\tilde{\mathbf{C}}_k^{H}\Big)\mathbf{\Lambda}_k\Big(\sum_{k'}(\tilde{\mathbf{P}}_{k'})^{*}\otimes\tilde{\mathbf{C}}_k\Big)\\
&\qquad\times\Big(\mathbf{I}+\Big(\sum_{k'}(\tilde{\mathbf{P}}_{k'})^{T}\otimes\tilde{\mathbf{C}}_k^{H}\Big)\mathbf{\Lambda}_k\Big(\sum_{k'}(\tilde{\mathbf{P}}_{k'})^{*}\otimes\tilde{\mathbf{C}}_k\Big)\Big)^{-1}\\
&\qquad\times\Big(\sum_{k'}(\tilde{\mathbf{P}}_{k'})^{T}\otimes\tilde{\mathbf{C}}_k^{H}\Big)\mathbf{\Lambda}_k^{1/2}\Big).
\end{aligned}
\tag{33}
$$

Although we consider the sum secret key rate maximization, the key rate differences among UTs are not large. As we consider the equal power allocation for different beams, most UTs can achieve similar secret key rates.

As in the objective function (33), the optimization matrices $\tilde{\mathbf{P}}_k$ and $\tilde{\mathbf{C}}_k$ are involved both inside and outside the matrix inversion operation, the function (33) is not convex on $\tilde{\mathbf{P}}_k$ and $\tilde{\mathbf{C}}_k$, resulting in the non-convex problem (32), which is difficult to solve globally. In order to reduce the computational complexity and lower the pilot overhead, we utilize the interference neutralization approach [25] to mitigate the interference, i.e., for arbitrary matrix $\tilde{\mathbf{C}}_{k'}$ $(k'\neq k)$, the precoding matrix $\tilde{\mathbf{P}}_k$ satisfies

$$
(\tilde{\mathbf{P}}_k^{T}\otimes\tilde{\mathbf{C}}_{k'}^{H})\mathbf{\Lambda}_{k'}=\mathbf{0},\quad k'\neq k.
\tag{34}
$$

This constraint indicates that the precoding matrix $\tilde{\mathbf{P}}_k$ can eliminate the inter-user interference. Note that when the channel beams of different users are non-overlapping, the precoding matrices correspond to different beams, and thus we have

$$
\tilde{\mathbf{P}}_k^{H}\tilde{\mathbf{R}}_{BS,k'}=\mathbf{0},\quad k'\neq k.
\tag{35}
$$

Therefore, the constraint (34) can be easily satisfied.

Under this constraint, problem (32) maximizing the sum secret key rate can be decomposed into the following subproblems maximizing the secret key rate of each user

$$
\begin{aligned}
&\min_{\tilde{\mathbf{P}}_k,\tilde{\mathbf{C}}_k}\ \log\det\Big(\mathbf{I}-\mathbf{\Lambda}_k^{1/2}\Big(\tilde{\mathbf{P}}_k^{*}\otimes\tilde{\mathbf{C}}_k\Big)\Big(\mathbf{I}+\Big(\tilde{\mathbf{P}}_k^{T}\otimes\tilde{\mathbf{C}}_k^{H}\Big)\mathbf{\Lambda}_k\\
&\qquad\times\Big(\tilde{\mathbf{P}}_k^{*}\otimes\tilde{\mathbf{C}}_k\Big)\Big)^{-1}\Big(\tilde{\mathbf{P}}_k^{T}\otimes\tilde{\mathbf{C}}_k^{H}\Big)\mathbf{\Lambda}_k\Big(\tilde{\mathbf{P}}_k^{*}\otimes\tilde{\mathbf{C}}_k\Big)\\
&\qquad\times\Big(\mathbf{I}+\Big(\tilde{\mathbf{P}}_k^{T}\otimes\tilde{\mathbf{C}}_k^{H}\Big)\mathbf{\Lambda}_k\Big(\tilde{\mathbf{P}}_k^{*}\otimes\tilde{\mathbf{C}}_k\Big)\Big)^{-1}\Big(\tilde{\mathbf{P}}_k^{T}\otimes\tilde{\mathbf{C}}_k^{H}\Big)\mathbf{\Lambda}_k^{1/2}\Big)\\
&\text{s.t. }\tilde{\mathbf{P}}_k^{H}\tilde{\mathbf{P}}_k=\mathbf{I}\\
&\quad\tilde{\mathbf{C}}_k^{H}\tilde{\mathbf{C}}_k=\mathbf{I}.
\end{aligned}
\tag{36}
$$

Remark 6: The secret key rate in (36) is equal to that using orthogonal pilots in (31). This means that when the precoding matrix $\tilde{\mathbf{P}}_k$ satisfies condition (34), UTs reusing the identical pilot signals approaches the performance with orthogonal pilot signals. Both schemes can mitigate the inter-user interference. The difference is that orthogonal pilot scheme uses orthogonal pilot signals to separate different UTs, which requires large pilot overhead, while interference neutralization scheme designs the precoding matrices to eliminate the interference, which is independent of pilot signals between UTs.

Note that in problem (36), the secret key rate depends only on the Kronecker product $\tilde{\mathbf{P}}_k^{*}\otimes\tilde{\mathbf{C}}_k$. Define $\mathbf{U}_k=\Big(\tilde{\mathbf{P}}_k^{*}\otimes\tilde{\mathbf{C}}_k\Big)$, which is also a tall unitary matrix. We first consider the matrix $\mathbf{U}_k$ design maximizing the secret key rate, and then, we construct the precoding and receiving matrices satisfying the interference neutralization constraint. The matrix $\mathbf{U}_k$ design problem can be expressed as

$$
\begin{aligned}
&\min_{\mathbf{U}_k}\ \log\det\Big(\mathbf{I}-\Big(\mathbf{\Lambda}_k^{1/2}\mathbf{U}_k\Big(\mathbf{I}+\mathbf{U}_k^{H}\mathbf{\Lambda}_k\mathbf{U}_k\Big)^{-1}\mathbf{U}_k^{H}\mathbf{\Lambda}_k^{1/2}\Big)^{2}\Big)\\
&\text{s.t. }\mathbf{U}_k^{H}\mathbf{U}_k=\mathbf{I}.
\end{aligned}
\tag{37}
$$

The solution of problem (37) is obtained as follows:

Theorem 2: The optimal $\mathbf{U}_k$ maximizing the secret key rate is

$$\mathbf{U}_k = \begin{bmatrix} \mathbf{e}_{\eta_1} & \mathbf{e}_{\eta_2} & \cdots & \mathbf{e}_{\eta_{M_e N_e}} \end{bmatrix}, \tag{38}$$

where $\mathbf{e}_i = [0, 0, \cdots, 0, 1, 0, \cdots, 0]$ is a unit vector with the ith unit element and η_i is the index of the ith sorted eigenvalue of matrix $\mathbf{\Lambda}_k(\mathbf{I} + \mathbf{\Lambda}_k)^{-1}$. The optimal rate is

$$R_k = -\sum_{i=1}^{M_e N_e} \log \lambda_i \left(\mathbf{I} - \mathbf{\Lambda}_k^2 (\mathbf{I} + \mathbf{\Lambda}_k)^{-2} \right). \tag{39}$$

Proof: See Appendix C. ∎

Remark 7: To maximize the secret key rate of the kth UT, the optimal $\mathbf{U}_k$ consists of the unit vectors corresponding to the sorted diagonal elements in $\mathbf{\Lambda}_k$. However, as $\mathbf{U}_k$ has the structure $\tilde{\mathbf{P}}_k^* \otimes \tilde{\mathbf{C}}_k$, in general cases, it cannot satisfy the optimal condition (38). Next, we will employ the channel properties and Theorem 2 to construct the precoding and receiving vectors $\tilde{\mathbf{P}}_k$ and $\tilde{\mathbf{C}}_k$.

As $\mathbf{U}_k$ is consist of vectors $\mathbf{e}_i$, the beam domain precoding and receiving matrices $\tilde{\mathbf{P}}_k$ and $\tilde{\mathbf{C}}_k$ have the similar structure, which only need to select corresponding beams. This indicates that the optimal precoding and receiving matrices $\mathbf{P}_k$ and $\mathbf{C}_k$ are consist of the eigenvectors of the channel covariance matrices, i.e., the precoding matrix $\mathbf{P}_k$ is a sub-matrix of $\mathbf{A}_{BS}$, while the receiving matrix $\mathbf{C}_k$ is a sub-matrix of $\mathbf{A}_{UT,k}$.

Moreover, recalling Proposition 1, as the number of antennas tends to infinity, the beam domain channel matrix $\tilde{\mathbf{H}}_k$ approaches the matrix $\mathbf{G}_k$, where different elements represent the channel gains from different AoDs to different AoAs. The channel gains are concentrated in a few elements in $\mathbf{G}_k$. Specifically, suppose that there are N_P paths, corresponding to N_P AoAs and N_P AoDs. Then, the BS selects the strongest N_P beams, i.e., the precoding matrix $\tilde{\mathbf{P}}_k$ is given by

$$\tilde{\mathbf{P}}_k = \begin{bmatrix} \mathbf{e}_{\eta_{t,k,1}} & \mathbf{e}_{\eta_{t,k,2}} & \cdots & \mathbf{e}_{\eta_{t,k,N_P}} \end{bmatrix} \tag{40}$$

where $\eta_{t,k,1}$ is the index of the sorted eigenvalue of matrix $\mathbf{R}_{BS,k}$. Similarly, UT k selects the strongest N_P receiving directions, i.e., the receiving matrix $\tilde{\mathbf{C}}_k$ is given by

$$\tilde{\mathbf{C}}_k = \begin{bmatrix} \mathbf{e}_{\eta_{r,k,1}} & \mathbf{e}_{\eta_{r,k,2}} & \cdots & \mathbf{e}_{\eta_{r,k,N_P}} \end{bmatrix} \tag{41}$$

where $\eta_{r,k,1}$ is the index of the sorted eigenvalue of matrix $\mathbf{R}_{UT,k}$. The number of paths N_P is relatively small, and M_e and N_e can be chosen equal to the number of paths. Using the precoding and receiving matrices, we can construct $\mathbf{U}_k = \tilde{\mathbf{P}}_k^* \otimes \tilde{\mathbf{C}}_k$ to obtain the N_P^2 elements in $\mathbf{\Lambda}_k$, which contains the channel information of the N_P paths. This approach can significantly reduce the pilot overhead and fit well to massive MIMO channel model and precoding [27].

B. Security Analysis With Overlapping Beams

In the above analysis, we assume that the channel matrices of different UTs are independent. This assumption can be easily satisfied for the non-overlapping case where the channel beams of different UTs are non-overlapping. However, different UTs may have overlapping beams in reality. For example, when two UTs are close to each other, part of their channels may suffer the same propagation paths, resulting in the overlapping beams, i.e.,

$$\tilde{\mathbf{R}}_{BS,k} \tilde{\mathbf{R}}_{BS,k'} \neq \mathbf{0}, \quad k' \neq k. \tag{42}$$

Since the channels of the overlapping beams between $\tilde{\mathbf{R}}_{BS,k}$ and $\tilde{\mathbf{R}}_{BS,k'}$ are highly correlated, the independent assumption does not hold any more. Thus, the information leakage should be considered for the overlapping case design.

Next, we analyze the information leakage ratio for the overlapping case. Note that for the kth UT, the beam domain channel elements in $\tilde{\mathbf{H}}_k$ are statistically independent and thus the secret key rate can be expressed as the summation of the key rate of each beam. Moreover, as UTs are assumed as non-colluding curious users, the information leakage is determined by the UT with the highest correlation. Thus, we focus on the information leakage on one overlapping beam with two UTs as an example.

Suppose that both UT 1 and UT 2 occupy the identical beam b at the BS. Denote the channel gains from beam b at the BS to the dominant beam at UT 1 (or UT 2) as h_1 (or h_2). Assume that both h_1 and h_2 have the unit attenuation power, i.e., $\mathbb{E}\{h_1^2\} = \mathbb{E}\{h_2^2\} = 1$. Define the information leakage ratio as $\gamma = (R_h - R_l)/R_h$, where R_h is the secret key rate with the independent channel assumption and R_l is the key rate when the channels are correlated. Since the correlation reduces the secret key rate, we always have $R_h \geq R_l$. Then, we can calculate γ as follows.

Theorem 3: The information leakage ratio can be expressed as

$$\gamma = 1 - \frac{\log \frac{((1+\sigma^2)^2 - \rho^2)^2}{(1+\sigma^2)(\sigma^6 + 3\sigma^4 - 2\sigma^2\rho^2 + 2\sigma^2)}}{\log \frac{(1+\sigma^2)^2}{\sigma^2(2+\sigma^2)}}, \tag{43}$$

where σ^2 is the noise variance and ρ is the cross channel correlation defined as

$$\rho = \frac{\mathbb{E}\{h_1 h_2\}}{\mathbb{E}\{h_1^2\}\mathbb{E}\{h_2^2\}} = \mathbb{E}\{h_1 h_2\}. \tag{44}$$

Proof: See Appendix D. ∎

Remark 8: The information leakage ratio given by (43) is complicated, depending on the correlation ρ as well as the noise variance σ^2. Next, we will consider a special case. From Appendix D, when ρ is 1 or -1, the information leakage is the highest, which can be calculated as

$$\gamma = \frac{\log \frac{3+2\sigma^2}{\sigma^2(6+11\sigma^2+6\sigma^4+\sigma^6)}}{\log \frac{(1+\sigma^2)^2}{\sigma^2(2+\sigma^2)}}. \tag{45}$$

For high SNRs (low σ^2), as σ^2 tends to 0, the information leakage ratio becomes

$$\lim_{\sigma^2 \to 0} \gamma = 1, \tag{46}$$

which indicates that the secret key rate goes to zero and vanishes.

From the above analysis, one can observe that the channel correlation is mainly caused by the overlapping channel beams. Further, when the channel of the overlapping beams is highly correlated, the information leakage ratio tends to 1.

This result reveals that the correlated overlapping beams provide little secret key rate. Therefore, in the multi-user secret key generation, when two UTs have overlapping channel beams, the BS should allocate *non-overlapping transmitting beams* to different UTs, i.e., the precoding vector $\tilde{\mathbf{P}}_k$ satisfies

$$\tilde{\mathbf{P}}_k \tilde{\mathbf{R}}_{BS,k'} = \mathbf{0}, \quad k' \neq k. \tag{47}$$

This indicates that the allocated transmitting beams for the kth UT are not overlapping with the channel beams of other UTs. Under this condition, the constraint (34) is satisfied, and the channel matrices of the allocated beams for different UTs are independent.

C. A Holistic Parameter Design Algorithm

Combined with the result of above security analysis, we propose a holistic parameter design algorithm as illustrated in Algorithm 1.

Algorithm 1 Parameter Design

Require: $\mathbf{R}_{BS,k}$ and $\mathbf{R}_{UT,k}$
Ensure: $\mathbf{P}_k$ and $\mathbf{C}_k$
1: **At the BS side:**
2: **for** $k = 1 : K$ **do**
3: Calculate the beam domain channel covariance matrix $\tilde{\mathbf{R}}_{BS,k}$ according to (19).
4: Select the strongest non-overlapping beams $\tilde{\mathbf{P}}_k$ according to (40) and (47).
5: Construct the precoding matrix $\mathbf{P}_k = \mathbf{A}_{BS}\tilde{\mathbf{P}}_k$.
6: **end for**
7: **At the UT side:**
8: **for** $k = 1 : K$ **do**
9: Calculate the beam domain channel covariance matrix $\tilde{\mathbf{R}}_{UT,k}$ according to (19).
10: Select the strongest beams $\tilde{\mathbf{C}}_k$ according to (41).
11: Construct the receiving matrix $\mathbf{C}_k = \mathbf{A}_{UT}\tilde{\mathbf{C}}_k$.
12: **end for**

With the help of designed parameters, the BS and UTs can extract reciprocal channel information of the non-overlapping beams, on which the channels of different UTs are independent. Therefore, we can complete the following key generation steps using the same approach as described in SectionIV-B.

It is noteworthy that the design of the matrices $\tilde{\mathbf{P}}_k$ and $\tilde{\mathbf{C}}_k$ depends on the statistical CSI $\tilde{\mathbf{R}}_{BS,k}$ and $\tilde{\mathbf{R}}_{UT,k}$. As the statistical CSI changes on a larger time scale than the instantaneous CSI, it is not necessary to design $\tilde{\mathbf{P}}_k$ and $\tilde{\mathbf{C}}_k$ for each secret key generation round. After designing $\tilde{\mathbf{P}}_k$ and $\tilde{\mathbf{C}}_k$, they can be used to generate secret keys until the statistical CSI changes. Also note that an offline design is possible. Depending on the statistical CSI scenario, we can choose the corresponding pilots.

VI. Numerical Results

In this section, we employ the numerical results to illustrate the performance of secret key generation in multi-user massive MIMO wireless communication systems. A BS, equipped with $M = 128$ antennas, simultaneously communicates with $K = 6$ UTs, each with $N_k = 4$ antennas. Here, we focus on massive antennas at the BS, which significantly affect the performance for the multi-user case. We consider the physical channel model, where there are $N_P = 6$ paths for each UT. We consider a ULA topology at the BS with 0.5λ antenna spacing. The channel is generated according to (1), where the AoDs and AoAs are randomly distributed.

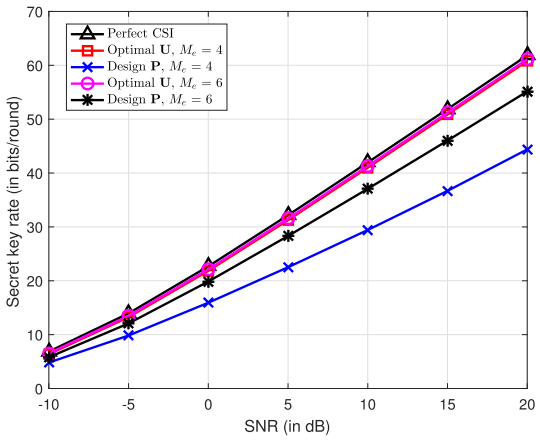


Fig. 4. Secret key rate comparison for one UT.

Fig. 4 presents the secret key rate of single user to confirm that our proposed CDR based secret key generation scheme is also suitable for the single user case. We compare the secret key rate of the designed matrices $\mathbf{U}$ and $\tilde{\mathbf{P}}$ with that of perfect CSI. The perfect CSI provides the complete channel information and achieves the highest secret key rate. Here, we set the matrix $\tilde{\mathbf{C}} = \mathbf{I}$ and consider $M_e = 4$ and $M_e = 6$ cases. From the results, when $M_e = 6$, the secret key rate of optimal $\mathbf{U}$ and the designed $\tilde{\mathbf{P}}$ can approach that of perfect CSI, indicating that by employing the precoding matrix $\tilde{\mathbf{P}}$, the BS and the UT can obtain the almost perfect channel information, significantly reducing the dimension of the channel estimation and the pilot overhead. When $M_e = 4$, the secret key rate approaches that of $M_e = 6$, which contains the most channel power with lower overhead.

Next, we consider the multi-user secret key generation and illustrate an example of multi-user channel gains distribution in the beam domain in Fig. 5. The BS employs the eigenmatrix of the channel $\mathbf{A}_{BS}$ to generate M fixed beams of different directions, where the mth beam is corresponding to the direction $\sin(\varphi_m) = 2m/M - 1$. Then, according to the particular location of the UT, the BS selects a number of beams from the M beams to generate secret key with him.

When six UTs are distributed in different positions, the channel gains of each UT are concentrated within a number of beams (or directions), different UTs occupy non-overlapping channel beams. The attenuation between the adjacent UTs is about 20 dB, significantly reducing inter-user interference. This result indicates that the BS equipped with massive antennas has the potential to achieve multi-user secret key generation.

Fig. 6 compares the secret key rate for different number of users. From the results, we can find that as the number of UTs increases, the secret key rate grows up approximately linearly.

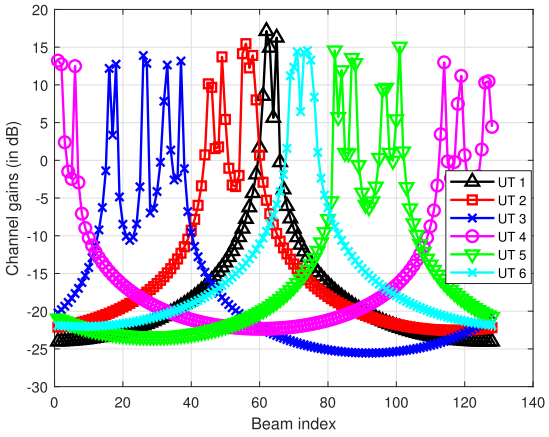


Fig. 5. Multi-user channel gains distribution in the beam domain.

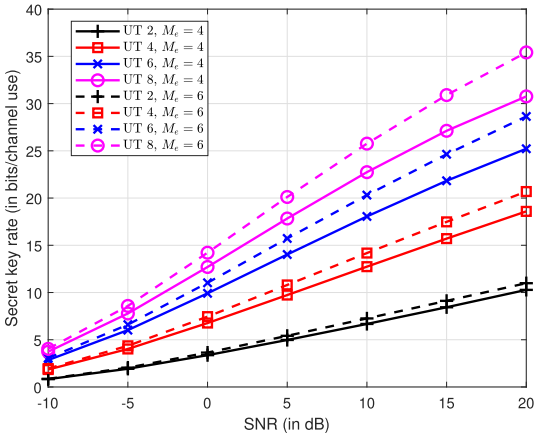


Fig. 6. Secret key rate comparison for different number of UTs.

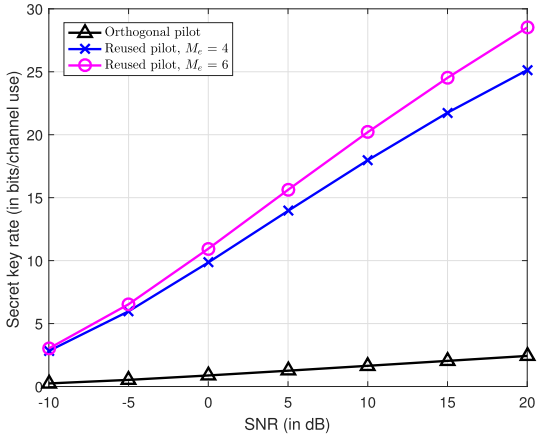


Fig. 7. Unit secret key rate comparison for multiple UTs of orthogonal pilot and reused pilot.

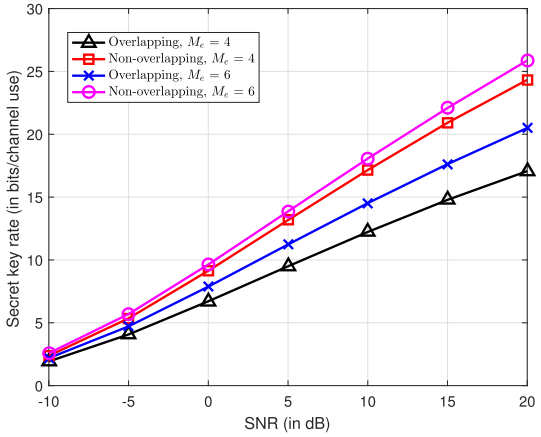


Fig. 8. Unit secret key rate comparison of overlapping and non-overlapping beams.

For example, the key rate of 4 UTs approaches twice than that of 2 UTs. However, as the number of UTs continues increasing, the key rate of each UT becomes lower, due to the interference or information leakage among UTs.

In multi-user secret key generation, the bottleneck is the pilot overhead. Considering the negative effect of pilot overhead, we define the unit secret key rate as

$$R_{\text{unit}} = R_{\text{sum}}/T, \tag{48}$$

where T is the pilot overhead, scaled with the dimension of the effective channel M_e and N_e. As the number of antennas at each UT is 4, we set $N_e = N_k = 4$. Fig. 7 compares the unit secret key rate of reused pilot with $M_e = 4$ and $M_e = 6$ with orthogonal pilot scheme. As the pilot overhead is extremely large for orthogonal pilot scheme, the unit secret key rate suffers serious loss. The reused pilot scheme with $M_e = 6$ achieves the highest rate and the rate of scheme with $M_e = 4$ is close to that of $M_e = 6$.

Fig. 8 compares the unit secret key rate of overlapping and non-overlapping transmitting beam schemes, when the channel beams are overlapping between different UTs. For the overlapping scheme, the BS allocates the strongest transmitting beams for each UT, some of which may be overlapping with other UTs, while for the non-overlapping scheme, the BS allocates the non-overlapping strongest transmitting beams for each UT. For the overlapping transmitting beam scheme, to estimate the channel of overlapping beams for different users, orthogonal pilot is used. Thus, the overhead is a little larger than that of non-overlapping transmitting beam scheme. Here, we do not consider the information leakage of the overlapping beams and only consider the interference between users. We observe that the unit secret key rates of non-overlapping schemes are higher than that of overlapping schemes. The non-overlapping scheme with $M_e = 6$ achieves the highest rate.

Then, we present the information leakage ratio when the channels of different UTs are correlated, as shown in Fig. 9. If the channels of UT 1 and UT 2 are correlated, they may observe similar channel measurements, resulting in the information leakage. When UT 2 is a potential eavesdropper, it can guess part of the key of UT 1, according to its correlated channel measurement. From the result, when the correlation coefficient is 1, the information leakage ratio increases as the SNR grows up. When the correlation coefficient is less than 1, the information leakage ratio increases in the low and middle SNR regions and decreases in the high SNR regions. This is because in the high SNR regions, the BS can obtain the precise channel information and extract the difference between them

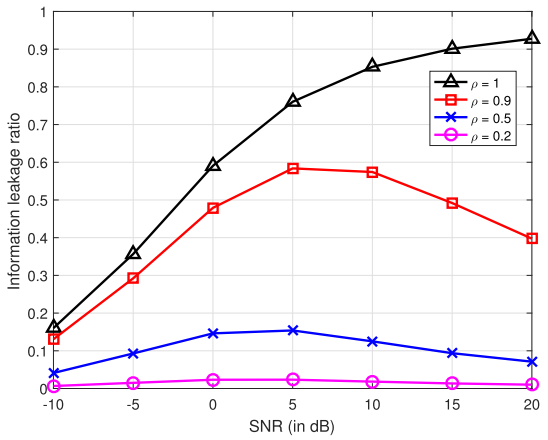


Fig. 9. Information leakage ratio with different correlation coefficients.

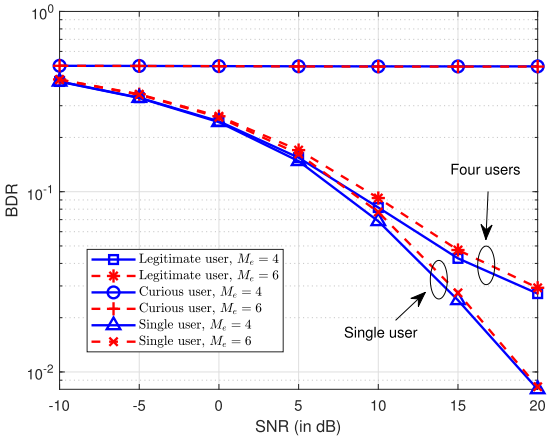


Fig. 10. BDR comparison for legitimate and curious UTs.

to generate the keys. However, the information leakage is still large when the correlation is high.

Next, we evaluate the bit disagreement ratios (BDR) performance of legitimate user and curious user, as shown in Fig. 10. The BDR is defined as the ratio of the number of the disagreement bits to the number of total bits of the initial secret key, which is the quantization result of channel measurement. In the figure, the curve of "legitimate user" refers to the BDR between the BS and each UT, while the curve of "curious user" refers to the BDR of two different UTs, which presents the key disagreement between different UTs. Moreover, we also illustrate the BDR of single user case, which is the best case without inter-user interference. From the results, we can find that the BDR of "legitimate user" approaches that of "single user", which indicates that the BDR performance of our proposed multi-user secret key generation approaches that of single-user key generation. The BDR of "curious user" remains high (about 0.5) for varied SNRs, which means that the quantized channel measurements of different UTs are different, one UT cannot guess the key of other UT based on his observation.

Finally, we evaluate the randoness of the initial key (before privacy amplification) via the National Institute of Standards and Technology (NIST) random test suite [28]. A tested

TABLE I
NIST RANDOM TEST RESULT

	Pass ratio	P-value
Approximate entropy	0.9199	0.4233
Runs	0.9262	0.4590
Ranking	0.9128	0.3810
Longest runs of ones	0.9717	0.3701
Frequency	0.9926	0.5254
FFT	0.9983	0.5800
Block frequency	0.9949	0.5466
Cumulative sums	0.9974	0.4997
Serial	0.9255	0.4474, 0.4903

bit sequence passes a test when the p-value is greater than the threshold, usually chosen as 0.01. We perform 9 NIST statistical tests for 10000 trials, and each initial key has a length of 256 bits. The pass ratios and the averaged p-values are summarized in Table I. For each test, the pass ratio is higher than 90% and the averaged p-value is significantly greater than 0.01. The results reflect a good randomness of the initial key generated via our proposed approach.

VII. CONCLUSION

This paper provided a fundamental design and analysis of the multi-user secret key generation in massive MIMO wireless networks. We provided a beam domain channel model, representing the channel gains from different transmit directions to different receive directions. We derived a closed-form expression of the secret key rate, which depends on the statistical CSI and the precoding and receiving matrices. We provided the optimal conditions for the Kronecker of the precoding the receiving matrices and proposed an algorithm to achieve the maximal sum secret key rate. When the beams of different UTs are non-overlapping, the BS employs several strongest beams of each UT to simultaneously generate secret key. Furthermore, we provided a security analysis by considering the channel correlation between UTs. When the channels of different UTs are correlated, the BS employs the several strongest non-overlapping beams of each UT to generate secret key. Numerical results demonstrate the performance improvement of our proposed multi-user secret key generation scheme. This work focuses on the sum secret key rate maximization, while the power allocation optimization under the fairness constraint among UTs can be further analyzed in the future.

APPENDIX A
PROOF OF PROPOSITION 1

From (13), the (n, m)th element of the beam domain channel $\tilde{\mathbf{H}}_k^{DL}$ can be expressed as

$$\begin{aligned}&[\tilde{\mathbf{H}}_k^{DL}]_{n,m}\\&= \mathbf{a}_{UT,k}(\theta_n)^H \mathbf{H}_k^{DL} \mathbf{a}_{Bs}(\varphi_m)\\&= \sum_p \alpha_{k,p} \mathbf{a}_{UT,k}(\theta_n)^H \mathbf{a}_{UT,k}(\theta_{k,p}) \mathbf{a}_{BS}(\varphi_{k,p})^H \mathbf{a}_{BS}(\varphi_m). \quad (49)\end{aligned}$$

First, we consider the calculation of $\mathbf{a}_{UT,k}(\theta_n)^H \mathbf{a}_{UT,k}(\theta_{k,p})$. As the number of UT antennas tends to infinity, there exists

θ_n equal to $\theta_{k,p}$ ($\theta_n = \theta_{k,p}$), and

$$\mathbf{a}_{UT,k}(\theta_n)^H \mathbf{a}_{UT,k}(\theta_{k,p}) = 1. \tag{50}$$

When θ_n is not equal to $\theta_{k,p}$, we have [29]

$$\lim_{N\to\infty} \mathbf{a}_{UT,k}(\theta_n)^H \mathbf{a}_{UT,k}(\theta_{k,p}) = \lim_{N_k\to\infty} \frac{1}{N_k} \frac{1 - e^{-j\frac{2\pi}{\lambda} d N_k(\sin(\theta_{k,p})-\sin(\theta_n))}}{1 - e^{-j\frac{2\pi}{\lambda} d(\sin(\theta_{k,p})-\sin(\theta_n))}} = 0. \tag{51}$$

Similarly, as the number of BS antennas grows, we have

$$\lim_{M\to\infty} \mathbf{a}_{BS}(\varphi_{k,p})^H \mathbf{a}_{BS}(\varphi_m) = \delta(\varphi_{k,p} - \varphi_m). \tag{52}$$

Thus, the (n, m)th element of $\tilde{\mathbf{H}}_{k,p}$ can be expressed as

$$\lim_{N,M\to\infty} [\tilde{\mathbf{H}}_k^{DL}]_{n,m} - \sum_p \alpha_{k,p}\delta(\theta_k,\, p - \theta_n)\delta(\varphi_{k,p} - \varphi_m) = 0. \tag{53}$$

This completes the proof. ■

Appendix B
Proof of Theorem 1

Assuming zero-mean complex Gaussian random vector for each channel observation $\mathbf{z}_k^{DL}$ or $\mathbf{z}_k^{UL}$, we have [30]

$$\begin{aligned} I(\mathbf{z}_k^{DL}; \mathbf{z}_k^{UL}|\mathbf{z}_i^{DL}) &= H(\mathbf{z}_k^{DL}, \mathbf{z}_i^{DL}) + H(\mathbf{z}_k^{UL}, \mathbf{z}_i^{DL}) \\ &\quad - H(\mathbf{z}_k^{DL}, \mathbf{z}_k^{UL}, \mathbf{z}_i^{DL}) - H(\mathbf{z}_i^{DL}) \\ &= \log \frac{\det(\mathcal{R}_{\mathbf{z}_k^{DL}\mathbf{z}_i^{DL}} \mathcal{R}_{\mathbf{z}_k^{UL}\mathbf{z}_i^{DL}})}{\det(\mathcal{R}_{\mathbf{z}_k^{DL}\mathbf{z}_k^{UL}\mathbf{z}_i^{DL}}) \det(\mathcal{R}_{\mathbf{z}_i^{DL}})}. \end{aligned} \tag{54}$$

Specially, when the channel observations of different UTs are uncorrelated, i.e., the channels of different UTs are independent, the conditional mutual information (54) can be simplified as

$$I(\mathbf{z}_k^{DL}; \mathbf{z}_k^{UL}|\mathbf{z}_i^{DL}) = I(\mathbf{z}_k^{DL}; \mathbf{z}_k^{UL}) = \log \frac{\det(\mathcal{R}_{\mathbf{z}_k^{DL}} \mathcal{R}_{\mathbf{z}_k^{UL}})}{\det(\mathcal{R}_{\mathbf{z}_k^{DL}\mathbf{z}_k^{UL}})}, \tag{55}$$

which only depends on the correlation of uplink and downlink channels.

We will calculate the covariance matrices $\mathcal{R}_{\mathbf{z}_k^{DL}}$, $\mathcal{R}_{\mathbf{z}_k^{UL}}$, and $\mathcal{R}_{\mathbf{z}_k^{DL}\mathbf{z}_k^{UL}}$ to obtain the secret key rate. The matrix $\mathcal{R}_{\mathbf{z}_k^{DL}}$ can be calculated as

$$\begin{aligned} \mathcal{R}_{\mathbf{z}_k^{DL}} &= \mathbb{E}\Bigg\{ \sum_{k'} ((\mathbf{P}_{k'}\mathbf{S}_{k'k}^{DL})^T \otimes \mathbf{C}_k^H)\mathrm{vec}(\mathbf{H}_k)\mathrm{vec}(\mathbf{H}_k)^H \\ &\quad \times \sum_{k'} ((\mathbf{P}_{k'}\mathbf{S}_{k'k}^{DL})^T \otimes \mathbf{C}_k^H)^H \\ &\quad + (\mathbf{I}_{T_D} \otimes \mathbf{C}_k^H)\mathrm{vec}(\mathbf{N}_k)\mathrm{vec}(\mathbf{N}_k)^H (\mathbf{I}_{T_D} \otimes \mathbf{C}_k^H)^H \Bigg\}. \end{aligned} \tag{56}$$

Without loss of generality, we consider the unit covariance matrix of noise, given by

$$\mathbb{E}\Big\{ \mathrm{vec}(\mathbf{N}_k)\mathrm{vec}(\mathbf{N}_k)^H \Big\} = \mathbf{I}_{NM_e}. \tag{57}$$

Recalling $\mathbf{R}_k = (\mathbf{A}_{BS}^* \otimes \mathbf{A}_{UT})\mathbf{\Lambda}_k(\mathbf{A}_{BS}^* \otimes \mathbf{A}_{UT})^H$, we have

$$\begin{aligned} \mathcal{R}_{\mathbf{z}_k^{DL}} &= \sum_{k'} ((\mathbf{P}_{k'}\mathbf{S}_{k'k}^{DL})^T \otimes \mathbf{C}_k^H)\mathbf{R}_k \sum_{k'} ((\mathbf{P}_{k'}\mathbf{S}_{k'k}^{DL})^T \otimes \mathbf{C}_k^H)^H \\ &\quad + (\mathbf{I}_{T_D} \otimes \mathbf{C}_k^H)(\mathbf{I}_{T_D} \otimes \mathbf{C}_k^H)^H \\ &= \sum_{k'} ((\mathbf{A}_{BS}^H \mathbf{P}_{k'} \mathbf{S}_{k'k}^{DL})^T \otimes \mathbf{C}_k^H \mathbf{A}_{UT})\mathbf{\Lambda}_k \\ &\quad \times \sum_{k'} ((\mathbf{A}_{BS}^H \mathbf{P}_{k'} \mathbf{S}_{k'k}^{DL})^T \otimes \mathbf{C}_k^H \mathbf{A}_{UT})^H \\ &\quad + (\mathbf{I}_{T_D} \otimes \mathbf{C}_k^H \mathbf{C}_k). \end{aligned} \tag{58}$$

Let $\tilde{\mathbf{P}}_k = \mathbf{A}_{BS}^H \mathbf{P}_k$ and $\tilde{\mathbf{C}}_k = \mathbf{A}_{UT,k}^H \mathbf{C}_k$. The covariance matrix $\mathcal{R}_{\mathbf{z}_k^{DL}}$ can be rewritten as

$$\mathcal{R}_{\mathbf{z}_k^{DL}} = \Big(\sum_{k'} (\tilde{\mathbf{P}}_{k'}\mathbf{S}_{k'k}^{DL})^T \otimes \tilde{\mathbf{C}}_k^H \Big) \mathbf{\Lambda}_k \Big(\sum_{k'} (\tilde{\mathbf{P}}_{k'}\mathbf{S}_{k'k}^{DL})^T \otimes \tilde{\mathbf{C}}_k^H \Big)^H + (\mathbf{I}_{T_D} \otimes \mathbf{C}_k^H \mathbf{C}_k). \tag{59}$$

Similarly, we can calculate $\mathbf{R}_{\mathbf{z}_k^{UL}}$ as

$$\mathcal{R}_{\mathbf{z}_k^{UL}} = \sum_{k'} \Big(\tilde{\mathbf{P}}_k^T \otimes \mathbf{S}_{k'k}^{UL} \tilde{\mathbf{C}}_{k'}^H \Big) \mathbf{\Lambda}_{k'} \Big(\tilde{\mathbf{P}}_k^T \otimes \mathbf{S}_{k'k}^{UL} \tilde{\mathbf{C}}_{k'}^H \Big)^H + (\mathbf{P}_k^T \mathbf{P}_k^* \otimes \mathbf{I}_{T_U}). \tag{60}$$

Next, we will calculate the determinant of covariance matrix $\mathcal{R}_{\mathbf{z}_k^{DL}\mathbf{z}_k^{UL}}$. Note that the matrix $\mathcal{R}_{\mathbf{z}_k^{DL}\mathbf{z}_k^{UL}}$ can be decomposed as

$$\mathcal{R}_{\mathbf{z}_k^{DL}\mathbf{z}_k^{UL}} = \begin{bmatrix} \mathcal{R}_{\mathbf{z}_k^{DL}} & \mathbf{R}_{\mathbf{z}_k^{DL}\mathbf{z}_k^{UL}} \\ \mathbf{R}_{\mathbf{z}_k^{UL}\mathbf{z}_k^{DL}} & \mathcal{R}_{\mathbf{z}_k^{UL}} \end{bmatrix} \tag{61}$$

where $\mathbf{R}_{\mathbf{z}_k^{DL}\mathbf{z}_k^{UL}}$ represents the covariance of $\mathbf{z}_k^{DL}$ and $\mathbf{z}_k^{UL}$,

$$\begin{aligned} \mathbf{R}_{\mathbf{z}_k^{DL}\mathbf{z}_k^{UL}} &= \mathbb{E}\{\mathbf{z}_k^{DL}(\mathbf{z}_k^{UL})^H\} \\ &= \Big(\sum_{k'} (\tilde{\mathbf{P}}_{k'}\mathbf{S}_{k'k}^{DL})^T \otimes \tilde{\mathbf{C}}_k^H \Big) \mathbf{\Lambda}_k \Big(\tilde{\mathbf{P}}_k^T \otimes \mathbf{S}_{kk}^{UL} \tilde{\mathbf{C}}_k^H \Big)^H. \end{aligned} \tag{62}$$

From the determinant of the block matrix, we have

$$\det(\mathcal{R}_{\mathbf{z}_k^{DL}\mathbf{z}_k^{UL}}) = \det(\mathcal{R}_{\mathbf{z}_k^{DL}}) \det\Big(\mathcal{R}_{\mathbf{z}_k^{UL}} - \mathbf{R}_{\mathbf{z}_k^{UL}\mathbf{z}_k^{DL}} \mathcal{R}_{\mathbf{z}_k^{DL}}^{-1} \mathbf{R}_{\mathbf{z}_k^{DL}\mathbf{z}_k^{UL}} \Big). \tag{63}$$

Hence, the secret key rate can be expressed as

$$\begin{aligned} I(\mathbf{z}_k^{DL}; \mathbf{z}_k^{UL}) &= \log \frac{\det(\mathcal{R}_{\mathbf{z}_k^{UL}})}{\det\Big(\mathcal{R}_{\mathbf{z}_k^{UL}} - \mathbf{R}_{\mathbf{z}_k^{UL}\mathbf{z}_k^{DL}} \mathcal{R}_{\mathbf{z}_k^{DL}}^{-1} \mathbf{R}_{\mathbf{z}_k^{DL}\mathbf{z}_k^{UL}} \Big)} \\ &= -\log\det\Big(\mathbf{I} - \mathcal{R}_{\mathbf{z}_k^{UL}}^{-1} \mathbf{R}_{\mathbf{z}_k^{UL}\mathbf{z}_k^{DL}} \mathcal{R}_{\mathbf{z}_k^{DL}}^{-1} \mathbf{R}_{\mathbf{z}_k^{DL}\mathbf{z}_k^{UL}} \Big). \end{aligned} \tag{64}$$

Let $\mathbf{V}_k = \mathbf{\Lambda}_k^{1/2} \Big(\sum_{k'} (\tilde{\mathbf{P}}_{k'}\mathbf{S}_{k'k}^{DL})^T \otimes \tilde{\mathbf{C}}_k^H \Big)^H$ and $\mathbf{V}_{kk'} = \mathbf{\Lambda}_{k'}^{1/2} \Big(\tilde{\mathbf{P}}_k^T \otimes \mathbf{S}_{k'k}^{UL} \tilde{\mathbf{C}}_{k'}^H \Big)^H$. Then, we can have

$$\begin{aligned} &\mathcal{R}_{\mathbf{z}_k^{UL}}^{-1} \mathbf{R}_{\mathbf{z}_k^{UL}\mathbf{z}_k^{DL}} \mathcal{R}_{\mathbf{z}_k^{DL}}^{-1} \mathbf{R}_{\mathbf{z}_k^{DL}\mathbf{z}_k^{UL}} \\ &\quad = \Bigg(\sum_{k'} \mathbf{V}_{kk'}^H \mathbf{V}_{kk'} + (\mathbf{P}_k^T \mathbf{P}_k^* \otimes \mathbf{I}_{T_U}) \Bigg)^{-1} \\ &\qquad \times \mathbf{V}_{kk}^H \mathbf{V}_k \Big(\mathbf{V}_k^H \mathbf{V}_k + \mathbf{I}_{T_D} \otimes \mathbf{C}_k^H \mathbf{C}_k \Big)^{-1} \mathbf{V}_k^H \mathbf{V}_{kk} \end{aligned} \tag{65}$$

and the secret key rate is given by

$$I(\mathbf{z}_k^{DL};\mathbf{z}_k^{UL}) = -\log\det\left(\mathbf{I}-\mathbf{V}_{kk}\left(\sum_{k'}\mathbf{V}_{kk'}^H\mathbf{V}_{kk'}+(\mathbf{P}_k^T\mathbf{P}_k^*\otimes\mathbf{I}_{T_U})\right)^{-1}\times\mathbf{V}_{kk}^H\mathbf{V}_k\left(\mathbf{V}_k^H\mathbf{V}_k+\mathbf{I}_{T_D}\otimes\mathbf{C}_k^H\mathbf{C}_k\right)^{-1}\mathbf{V}_k^H\right). \quad (66)$$

This completes the proof. ■

Appendix C
Proof of Theorem 2

The sum rate can be expressed as

$$R_k = -\log\det\left(\mathbf{I}-\left(\mathbf{\Lambda}_k^{1/2}\mathbf{U}_k\left(\mathbf{I}+\mathbf{U}_k^H\mathbf{\Lambda}_k\mathbf{U}_k\right)^{-1}\mathbf{U}_k^H\mathbf{\Lambda}_k^{1/2}\right)^2\right). \quad (67)$$

From Eq. 10.55 in [31], we have

$$\begin{aligned}&\mathbf{\Lambda}_k^{1/2}\mathbf{U}_k(\mathbf{I}+\mathbf{U}_k^H\mathbf{\Lambda}_k\mathbf{U}_k)^{-1}\mathbf{U}_k^H\mathbf{\Lambda}_k^{1/2}\\ &\quad=\mathbf{\Lambda}_k^{1/2}\mathbf{U}_k\left(\mathbf{U}_k^H(\mathbf{I}+\mathbf{\Lambda}_k)\mathbf{U}_k\right)^{-1}\mathbf{U}_k^H\mathbf{\Lambda}_k^{1/2}\\ &\quad\preceq\mathbf{\Lambda}_k^{1/2}(\mathbf{I}+\mathbf{\Lambda}_k)^{-1}\mathbf{\Lambda}_k^{1/2}. \quad (68)\end{aligned}$$

Thus, the sorted eigenvalues satisfy

$$\lambda_i(\mathbf{\Lambda}_k^{1/2}\mathbf{U}_k(\mathbf{I}+\mathbf{U}_k^H\mathbf{\Lambda}_k\mathbf{U}_k)^{-1}\mathbf{U}_k^H\mathbf{\Lambda}_k^{1/2}) \le\lambda_i(\mathbf{\Lambda}_k(\mathbf{I}+\mathbf{\Lambda}_k)^{-1}) \quad (69)$$

and we have

$$\lambda_i\left(\mathbf{I}-\left(\mathbf{\Lambda}_k^{1/2}\mathbf{U}_k\left(\mathbf{I}+\mathbf{U}_k^H\mathbf{\Lambda}_k\mathbf{U}_k\right)^{-1}\mathbf{U}_k^H\mathbf{\Lambda}_k^{1/2}\right)^2\right) \ge\lambda_i\left(\mathbf{I}-\mathbf{\Lambda}_k^2(\mathbf{I}+\mathbf{\Lambda}_k)^{-2}\right). \quad (70)$$

Thus,

$$\begin{aligned}R_k&=-\sum_i\log\lambda_i\left(\mathbf{I}-\left(\mathbf{\Lambda}_k\mathbf{U}_k\left(\mathbf{I}+\mathbf{U}_k^H\mathbf{\Lambda}_k\mathbf{U}_k\right)^{-1}\mathbf{U}_k^H\right)^2\right)\\ &\le-\sum_i\log\lambda_i\left(\mathbf{I}-\mathbf{\Lambda}_k^2(\mathbf{I}+\mathbf{\Lambda}_k)^{-2}\right). \quad (71)\end{aligned}$$

The equality holds only when $\mathbf{U}_k$ is consist of the unit vectors with the indices of unit elements corresponding to that of the sorted eigenvalues, i.e.,

$$\mathbf{U}_k=\begin{bmatrix}\mathbf{e}_{\lambda_1} & \mathbf{e}_{\lambda_2} & \cdots & \mathbf{e}_{\lambda_{M_e N_e}}\end{bmatrix} \quad (72)$$

where $\mathbf{e}_i=[0,0,\cdots,0,1,0,\cdots,0]$ is a unit vector with the ith unit element. This completes the proof. ■

Appendix D
Proof of Theorem 3

The secret key rate of UT 1 corresponding to beam b is given by

$$R_{b1}=\log\frac{\det\left(\mathcal{R}_{(h_1+n_1^{DL})(h_2+n_2^{DL})}\mathcal{R}_{(h_1+n_1^{UL})(h_2+n_2^{DL})}\right)}{\det\left(\mathcal{R}_{(h_1+n_1^{DL})(h_1+n_1^{UL})(h_2+n_2^{DL})}\right)\det\left(\mathcal{R}_{(h_2+n_2^{DL})}\right)} \quad (73)$$

where n_1^{UL} and n_1^{DL} (or n_2^{DL}) are the received noise of UT 1 (or UT 2) in the uplink and downlink transmissions. When the variance of the noise is σ^2, we can calculate R_{b1} in closed-form as

$$R_{b1}=\log\frac{((1+\sigma^2)^2-\rho^2)^2}{(1+\sigma^2)(\sigma^6+3\sigma^4-2\sigma^2\rho^2+2\sigma^2)}. \quad (74)$$

Taking the derivative of R_{b1} with respect to ρ, we can obtain

$$\frac{\partial R_{b1}}{\partial\rho}=\frac{-4\rho\sigma^2(1+\sigma^2-\rho)(1+\sigma^2+\rho)(1+\sigma^2-\rho^2)}{(1+\sigma^2)(2\sigma^2+3\sigma^4+\sigma^6-2\sigma^2\rho^2)^2}. \quad (75)$$

As ρ is in the region $[-1,1]$, $(1+\sigma^2-\rho)$, $(1+\sigma^2+\rho)$, and $(1+\sigma^2-\rho^2)$ in the numerator are positive. Then, when ρ is in the region $[0,1]$, the derivative $\frac{\partial R_{b1}}{\partial\rho}\le 0$, which indicates the rate R_{b1} is monotonic decreasing. When ρ is in the region $[-1,0]$, the derivation $\frac{\partial R_{b1}}{\partial\rho}\ge 0$, which indicates the rate R_{b1} is monotonic increasing. Thus, we can have the highest rate R_h with $\rho=0$ as

$$R_h=\log\frac{(1+\sigma^2)^2}{\sigma^2(2+\sigma^2)}, \quad (76)$$

and the lowest rate R_l with $\rho=1$ or $\rho=-1$ as

$$R_l=\log\frac{(2+\sigma^2)^2}{3+4\sigma^2+\sigma^4}. \quad (77)$$

Thus, the information leakage ratio can be calculated as

$$\gamma=1-\frac{\log\frac{((1+\sigma^2)^2-\rho^2)^2}{(1+\sigma^2)(\sigma^6+3\sigma^4-2\sigma^2\rho^2+2\sigma^2)}}{\log\frac{(1+\sigma^2)^2}{\sigma^2(2+\sigma^2)}}. \quad (78)$$

This completes the proof. ■

References

[1] J. Zhang, T. Q. Duong, A. Marshall, and R. Woods, "Key generation from wireless channels: A review," *IEEE Access*, vol. 4, pp. 614–626, Mar. 2016.

[2] C. Chen and M. A. Jensen, "Secret key establishment using temporally and spatially correlated wireless channel coefficients," *IEEE Trans. Mobile Comput.*, vol. 10, no. 2, pp. 205–215, Feb. 2011.

[3] G. Li, A. Hu, J. Zhang, L. Peng, C. Sun, and D. Cao, "High-agreement uncorrelated secret key generation based on principal component analysis preprocessing," *IEEE Trans. Commun.*, vol. 66, no. 7, pp. 3022–3034, Jul. 2018.

[4] J. Zhang, T. Duong, R. Woods, and A. Marshall, "Securing wireless communications of the Internet of Things from the physical layer, an overview," *Entropy*, vol. 19, no. 8, p. 420, Aug. 2017.

[5] C. Cheng, R. Lu, A. Petzoldt, and T. Takagi, "Securing the Internet of Things in a quantum world," *IEEE Commun. Mag.*, vol. 55, no. 2, pp. 116–120, Feb. 2017.

[6] C. Huth, R. Guillaume, T. Strohm, P. Duplys, I. A. Samuel, and T. Güneysu, "Information reconciliation schemes in physical-layer security: A survey," *Comput. Netw.*, vol. 109, pp. 84–104, Nov. 2016.

[7] E. A. Jorswieck, A. Wolf, and S. Engelmann, "Secret key generation from reciprocal spatially correlated MIMO channels," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Atlanta, GA, USA, Dec. 2013, pp. 1–6.

[8] S. A. Busari, K. M. S. Huq, S. Mumtaz, L. Dai, and J. Rodriguez, "Millimeter-wave massive MIMO communication for future wireless systems: A survey," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 2, pp. 836–869, 2nd Quart., 2018.

[9] G. Li, C. Sun, J. Zhang, E. Jorswieck, B. Xiao, and A. Hu, "Physical layer key generation in 5G and beyond wireless communications: Challenges and opportunities," *Entropy*, vol. 21, no. 5, p. 497, May 2019.

[10] L. Jiao, J. Tang, and K. Zeng, "Physical layer key generation using virtual AoA and AoD of mmWave massive MIMO channel," in *Proc. IEEE Conf. Commun. Netw. Secur. (CNS)*, May 2018, pp. 1–9.

[11] L. Jiao, N. Wang, and K. Zeng, "Secret beam: Robust secret key agreement for mmWave massive MIMO 5G communication," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2018, pp. 1–6.

[12] H. Liu, J. Yang, Y. Wang, Y. Chen, and C. E. Koksal, "Group secret key generation via received signal strength: Protocols, achievable rates, and implementation," *IEEE Trans. Mobile Comput.*, vol. 13, no. 12, pp. 2820–2835, Dec. 2014.
[13] C. D. T. Thai, J. Lee, J. Prakash, and T. Q. S. Quek, "Secret group-key generation at physical layer for multi-antenna mesh topology," *IEEE Trans. Inf. Forensics Security*, vol. 14, no. 1, pp. 18–33, Jan. 2019.
[14] R. Jin and K. Zeng, "Physical layer multi-user key generation in wireless networks," *Wireless Netw.*, vol. 24, no. 4, pp. 1043–1054, May 2018.
[15] J. Zhang, M. Ding, D. Lopez-Perez, A. Marshall, and L. Hanzo, "Design of an efficient OFDMA-based multi-user key generation protocol," *IEEE Trans. Veh. Technol.*, vol. 68, no. 9, pp. 8842–8852, Sep. 2019.
[16] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. New York, NY, USA: Cambridge Univ. Press, 2005.
[17] F. Rusek, D. Persson, B. Kiong Lau, E. G. Larsson, T. L. Marzetta, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.
[18] C. Sun, X. Gao, S. Jin, M. Matthaiou, Z. Ding, and C. Xiao, "Beam division multiple access transmission for massive MIMO communications," *IEEE Trans. Commun.*, vol. 63, no. 6, pp. 2170–2184, Jun. 2015.
[19] Y. Wang, W. Xu, H. Zhang, and X. You, "Wideband mmWave channel estimation for hybrid massive MIMO with low-precision ADCs," *IEEE Wireless Commun. Lett.*, vol. 8, no. 1, pp. 285–288, Feb. 2019.
[20] A. M. Sayeed, "Deconstructing multiantenna fading channels," *IEEE Trans. Signal Process.*, vol. 50, no. 10, pp. 2563–2579, Oct. 2002.
[21] L. You, X. Gao, G. Y. Li, X.-G. Xia, and N. Ma, "BDMA for millimeter-wave/terahertz massive MIMO transmission with per-beam synchronization," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 7, pp. 1550–1563, Jul. 2017.
[22] J. W. Wallace, C. C. C. Chen, and M. A. Jensen, "Key generation exploiting MIMO channel evolution: Algorithms and theoretical limits," in *Proc. Eur. Conf. Antennas Propag.*, Berlin, Germany, 2009, pp. 1499–1503.
[23] J. Zhang, G. Li, A. Marshall, A. Hu, and L. Hanzo, "A new frontier for IoT security emerging from three decades of key generation relying on wireless channels," *IEEE Access*, vol. 8, pp. 138406–138446, 2020.
[24] Y. Liang, H. V. Poor, and S. Shamai (Shitz), "Information theoretic security," *Found. Trends Commun. Inf. Theory*, vol. 5, nos. 4–5, pp. 355–580, 2009.
[25] Z. Ho, E. Jorswieck, and S. Engelmann, "Information leakage neutralization for the multi-antenna non-regenerative relay-assisted multi-carrier interference channel," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 9, pp. 1672–1686, Sep. 2013.
[26] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?" *IEEE Trans. Inf. Theory*, vol. 49, no. 4, pp. 951–963, Apr. 2003.
[27] O. E. Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1499–1513, Mar. 2014.
[28] National Institute of Standards and Technology. *NIST SP 800-22: Download Documentation and Software*. Accessed: Jan. 1, 2019. [Online]. Available: https://csrc.nist.gov/projects/random-bit-generation/documentation-and-software
[29] T. Bai and R. W. Heath, "Asymptotic SINR for millimeter wave massive MIMO cellular networks," in *Proc. IEEE 16th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jun. 2015, pp. 620–624.
[30] J. W. Wal and R. K. Sharma, "Automatic secret keys from reciprocal MIMO wireless channels: Measurement and analysis," *IEEE Trans. Inf. Forensics Security*, vol. 5, no. 3, pp. 381–392, Sep. 2010.
[31] G. A. Seber, *A Matrix Handbook for Statisticians*, vol. 15. Hoboken, NJ, USA: Wiley, 2008.

Guyue Li (Member, IEEE) received the B.S. degree in information science and technology and the Ph.D. degree in information security from Southeast University, Nanjing, China, in 2011 and 2017, respectively.

From June 2014 to August 2014, she was a Visiting Student with the Department of Electrical Engineering, Tampere University of Technology, Finland. She is currently an Associate Professor with the School of Cyber Science and Engineering, Southeast University. Her research interests include physical-layer security, secret key generation, radio frequency fingerprint, and link signature.

Chen Sun (Member, IEEE) received the Ph.D. degree in electrical engineering from Southeast University, Nanjing, China, in 2018.

From September 2015 to August 2016, he was a Visiting Student with the Department of Electrical and Computer Engineering, University of California at Davis, Davis, CA, USA. He is currently a Lecturer with the National Mobile Communications Research Laboratory, Southeast University. His research interests include communications and information theory, with emphasis on massive MIMO communications and optical wireless communications.

Eduard A. Jorswieck (Fellow, IEEE) was born in Berlin, Germany, in 1975. From 2008 until 2019, he was the Head of the Chair of Communications Theory and a Full Professor with the Dresden University of Technology (TUD), Germany. He is currently the Managing Director of the Institute of Communications Technology and the Head of the Chair of Communications Systems and a Full Professor with Technische Universitäet Braunschweig, Brunswick, Germany. He has published more than 130 journal articles, 15 book chapters, three monographs, and some 275 conference papers on these topics. His main research interest includes the broad area of communications. In 2006, he received the IEEE Signal Processing Society Best Paper Award. He has been a member of the IEEE SAM Technical Committee since 2015. Since 2017, he has been serving as the Editor-in-Chief of the *EURASIP Journal on Wireless Communications and Networking*. He also serves on the Editorial Board for IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY.

Junqing Zhang received the B.Eng. and M.Eng. degrees in electrical engineering from Tianjin University, China, in 2009 and 2012, respectively, and the Ph.D. degree in electronics and electrical engineering from Queen's University Belfast, U.K., in 2016. From February 2016 to January 2018, he was a Post-Doctoral Research Fellow with Queen's University Belfast. Since February 2018, he has been a Tenure Track Fellow (an Assistant Professor) with the University of Liverpool, U.K. His research interests include Internet of Things, wireless security, physical-layer security, key generation, and radio frequency fingerprinting identification.

Aiqun Hu (Member, IEEE) received the B.Sc.(Eng.), M.Eng.Sc., and Ph.D. degrees from Southeast University in 1987, 1990, and 1993, respectively.

He was invited as a Post-Doctoral Research Fellow with The University of Hong Kong from 1997 to 1998, and a TCT Fellow with Nanyang Technological University in 2006. He has published two books and more than 100 technical articles in wireless communications field. His research interests include data transmission and secure communication technology.

You Chen (Student Member, IEEE) is currently pursuing the bachelor's degree with the School of Cyber Science and Engineering, Southeast University. Her research interest includes physical-layer security.