Efficient Physical Layer Group Key Generation in 5G Wireless Networks

Long Jiao, Pu Wang, Ning Wang, Songlin Chen, Amir Alipour-Fanid, Junqing Le, Kai Zeng
Department of Electrical and Computer Engineering
George Mason University, Fairfax, VA, US 22030
E-mail: {ljiao, pwang20, nwang5, schen41, aalipour, jle36, kzeng2}@gmu.edu

Abstract—We investigate group secret key generation in 5G mmWave Massive MIMO networks and intend to improve the efficiency of channel probing for group key generation. In this paper, a new channel probing strategy for star-topology networks is proposed, which relies on the multiplexing of downlink probing signals, enabled by the hybrid precoding. To improve the group key rates per channel probing, a genetic algorithm (GA) based power allocation algorithm is also proposed. What’s more, to estimate the group key rates based on the probing samples, we propose a scheme for the estimation of group key rates based on the maximum likelihood estimator (MLE). The performance of the proposed scheme on group key rates and bits disagreement ratio (BDR) are provided. The numerical results show that the GA-based downlink channel probing scheme can increase the efficiency of channel probing and have higher group key rates compared with the existing channel probing schemes. When the SNR is 25dB, the key rates of GA-based power allocation scheme are 20% higher than the scheme with the conventional channel probing strategy.

Index Terms—5G, Group key generation, Physical layer security, mmWave, Massive MIMO, Spatial multiplexing.

I. INTRODUCTION

Thanks to the emerging fifth generation (5G) technologies, such as millimeter wave (mmWave), massive MIMO and hybrid precoding, 5G wireless communication are supposed to be the key enabler to satisfy the increasing demand for data service like ultra-reliability communication, massive machine type communication and high data rate [1–3]. To ensure the secure and reliable communication service, the efficient and lightweight security mechanism is desired in the design of 5G networks. To this end, we focus on the security of 5G mmWave wireless networks and consider physical layer key generation in the aforementioned networks.

Different from the traditional Diffie-Hellman (D-H) key exchange mechanism, physical layer key generation mechanisms do not require expensive computation and have the potential to achieve information-theoretic security. Instead of relying on D-H key exchange, physical layer key generation is based on the principle of channel reciprocity and within the channel coherence time, two wireless devices can extract identical secret bits independently by using the sampled sequence from the radio channel between them. In this article, we focus on the physical layer group key generation, which involves multiple wireless nodes and we examine secure group communications among multiple wireless devices by exploiting physical layer information of the radio channel.

Although significant efforts and progress on physical layer group key generation have been made in recent years, there are many roadblocks need to be addressed. For multi-user mmWave wireless networks operating in the time division duplex (TDD) mode [2], to obtain the channel samples, the bi-directional channel probing are required. Currently, the bi-directional channel probing are performed sequentially. In this manner, the larger overhead for channel probing are expected, which can lead to a relatively low group key rate.

To tackle the challenge mentioned above, techniques like multi-user multiplexing in 5G mmWave communication networks can be utilized to improve the efficiency of channel probing. We propose an efficient channel probing scheme for star-topology wireless networks. Our scheme can be extended to the ring-topology and mesh networks with further efforts. After uplink channel estimation, a baseband (BB) precoder is applied in the downlink channel probing, which is enabled by the hybrid precoding [2]. In the downlink channel probing, the channel probing signals of several edge nodes (ENs) are multiplexed in a single slot for channel probing, which enables the central node (CN) to send downlink probing signals to ENs concurrently. For example, for $M$ ENs, in TDD mode, the existing probing schemes need $2M$ channel probing while in our scheme, only $M + 1$ channel probing are required. Such a function is enabled by BB precoder within the hybrid precoder. By allowing to set a precoder at baseband, hybrid precoding provides the design of freedom to have each downlink effective channel approximately orthogonal to other downlink effective channels.

Secondly, to further improve the efficiency of downlink channel probing and increase the group key rate per channel probing, we model the power allocation for group key generation as a multi-objective optimization problem and propose a power allocation algorithm for downlink channel probing. At each algorithm run, GA applies the genetic operator to get the offspring population produced by the parent population. To avoid the algorithm wandering in the infeasible searching region, non-dominant sorting genetic algorithm (NSGA II) [4] is incorporated in the power allocation scheme. The assignment of the transmission power on each downlink stream is determined by NSGA II, which can guide to a spread-out solution set for power allocation parameters and avoid the local optimality.

What’s more, to quantify the performance of the group key rate for the proposed downlink channel probing scheme and the GA-based power allocation algorithm, the group key rate at each round is approximated/estimated via maximum
likelihood estimator (MLE). For instance, at first, the group key rates are decoupled and represented as the combination of several joint entropy. In the second step, parameters defining the distribution of each joint entropy are estimated based on the MLE. Finally, with the estimated parameters, the group key rates are computed. The numerical results show that the group key rates after estimation and the rates based on theoretical rates are computed. The numerical results show that the group key rates after estimation and the rates based on theoretical derivations are in good agreement.

To sum up, regarding the challenges reside in existing group key generation schemes, the main contributions of this paper are summarized as below:

- We propose a downlink channel probing scheme for the star-topology networks, based on hybrid precoding. In future work, this scheme will be extended to ring-topology and mesh networks.
- To improve the group key rates, a GA-based power allocation algorithm is developed in this work to approach the group key rate upper bound in [7]. The time multi-segment algorithms are further developed to achieve or approach the group key rates after estimation and the rates based on theoretical derivations are in good agreement.
- A fast group key rate estimation method based on MLE is proposed.

In section III, we introduce the specifications of our system, the link initialization, uplink, and downlink channel probing. In Section IV, the proposed downlink channel probing scheme is introduced in detail. Specifically, the multiplexing of downlink probing signals, enabled by hybrid precoding, is discussed. The group key generation model as a multigraph, in which the distribution of each joint entropy representing the group key rate.

For the second category, in [10], a secret group key generation scheme was proposed for the star topology using the received signal strength (RSS). Specifically, the channel between an edge node (EN) and a central node (CN) is selected as the reference channel and estimated at first. Then, for each other EN, the CN forwards the difference of two RSSs of the reference channel and the channel linked to that EN. As that EN has the estimation of channel linked to the CN, it can also estimate the reference channel using the received RSS difference. Finally, all nodes can have the reference channel information, and use it to generate the group key.

The existing works [5–7, 10] remain several issues unexplored. At first, the current group key generation schemes did not carefully study the channel probing efficiency. They assume the bi-directional channel probing on each link is done within the channel coherent time [10]. For the short channel coherent time and a large number of group members, improving probing efficiency or obtaining more probing samples in the unit time is desirable. Besides, the estimation of group key rates is unexplored. For instance, for bi-directional physical layer key generation, the key rates, represented by the combination of mutual information, usually need to be inferred/estimated from the channel probing signals by estimating corresponding entropy value [11, 12]. Various toolboxes are developed to compute the mutual information based on bi-directional probing samples, i.e., a clustering-based scheme for the estimation of key rates has been applied in [11]. However, for group key generation, the estimation of group key rate involves multiple nodes and it is challenging to infer the joint distribution of joint entropy representing the group key rate.

II. RELATED WORK

Existing group key generation methods can be broadly classified into two mechanisms. In the first mechanism, every node tries to generate pairwise keys among users using physical layer key generation first, and then generate a group key based on the pairwise keys (e.g., broadcast one of the keys (shortest one) xor’ed with the key associated with each user). In this way, each user can reconstruct the shortest key [5–7]. In the second mechanism, every node tries to conduct channel probing for each pair of users first. In the next step, a user’s reference channel information is selected and shared with other users by sending the channel state difference to other users or broadcasting linear combinations of the collected channel information at each node. Finally, every node generates a group key based on a reference channel or all the channel information [8, 9].

For the first category, a classical strategy for group key generation using the pairwise keys is to utilize tree-based algorithms related to graphs [5, 6]. The basic idea is to treat the group key generation model as a multigraph, in which each pairwise key rate can be viewed as the weight of the edge associated with the corresponding two nodes. Then, a spanning tree can be found in this multigraph, and the group key information can be propagated over this spanning tree by dividing each pairwise key into multiple one-bit segments and transmitting one-time pads of these segments. Simple multi-segment algorithms are further developed to achieve or approach the group key rate upper bound in [7]. The time allocation problem in the channel estimation steps to maximize the group key rates is proposed.

We propose a downlink channel probing scheme for the star-topology networks, based on hybrid precoding. In future work, this scheme will be extended to ring-topology and mesh networks.

A. System Model

We assume all of the nodes operate in the TDD mode and at mmWave frequency band (28GHz). In the scope of this paper, there are $M (M \geq 3)$ nodes, wishing to generate a common secret key through wireless fading channels under a passive eavesdropper.

Due to the limitation of pages, we will start with a star network, which is most commonly seen in practice, such as WLAN and cellular networks. In the network, we have $M$ nodes $M = \{1, 2, ..., M\}$, where $M$ consists of a CN $c \in M$, e.g., base station (BS) or access point (AP), and $M - 1$ ENs. Each node $m \in M \setminus c$. Each EN and CN are equipped with $N_{EN}$ and $N_{CN}$ antennas. What’s more, the number of RF chains at each central node is $N_{RF}$, which affects the maximum number of users that can be simultaneously served by the central node [2]. Moreover, we assume that the reciprocity of mmWave channels between two nodes holds for the downlink and uplink. We adopt the narrow band block fading channel models, as specified in [1, 2, 13], which are constant over multiple channel slots, and change randomly at the beginning of the next block.
Before bi-directional channel estimation, none of the nodes have the prior channel state information (CSI). However, the distributions of CSI are available at each node. For simplicity, the distribution coefficients of channel gains defined in [1, 2, 13] are applied in this work and the developed scheme can be easily extended to other coefficients. Please note that the above assumptions have been widely used in existing works.

B. Channel Probing

1) Link Establishment: According to our design, there is a fast link establishment process between CN and EN for the purpose of angle of arrival (AoA) and angle of departure (AoD) estimation. Here, we only consider the None-line-of-sight (NLOS) paths.

Due to the power constraint at EN, according to [2], we consider the analog precoding scheme which could provide the sufficient antenna gain. Analog combiners utilize the phase shifter at ENs and tries to establish a link with CN c. EN u and CN c need to search over the codebook \( W \) and find the combining vectors \( w_u \in W^E_N \) and \( w_c \in W^C_N \) [2]. The entries in \( W^E_N \) and \( W^C_N \) are normalized to satisfy \( |w_{i,j}^{|E_N}|^2 = N_{E_N}^{-1} \) and \( |w_{c,j}^{|C_N}|^2 = N_{C_N}^{-1} \) with a finite set of possible values, i.e., \( w_{i,j}^{|E_N} = \frac{1}{\sqrt{N_{E_N}}} e^{j\theta_{i,j}} \).

Similar to [2], we assume the number of ENs meets the condition that \( 1 \leq N_r \leq N_{RF} \). For every EN associated with CN c, the CN c serves every EN using a single RF chain. The beamformers \( w_u \) and \( w_c \) try to maximize the received signal by solving the following problem:

\[
\begin{align*}
\max_{w_u, w_c} & \quad |w_u^H H_{uc} w_c|^2, \\
\text{subject to} & \quad w_u \in W^E_N, \quad w_c \in W^C_N,
\end{align*}
\]

where \( \theta_b \) is the beam width of the main lobe, \( M_s \) and \( m_s \) are array gains of the main lobe and side lobe, respectively.

2) Uplink Channel Estimation: After the link initialization stage, the appropriate beamformers have been selected. In the next stage, the uplink channel probing from every EN u to CN c is conducted. Every EN u \( \in M \setminus c \) sends channel probing signals sequentially to CN, which can be expressed as:

\[
y^1_{cu} = w_c^H H_{cu} w_u s_u + w_c^H n_{cu},
\]

where \( n_{cu} \in C^{N_c \times 1} \) is the noise, which satisfies the circular-symmetry complex Gaussian distribution. \( H_{cu} \) is the narrowband mmWave channel matrix. The definition of beamformers \( w_c \) and \( w_u \) will be given later.

After obtaining the uplink channel probing \( y^1_{cu} \), CN c would like to estimate the effective channel \( w_c^H H_{cu} w_u \), which can be performed by the estimators like least squares (LS) or MLE.

3) Downlink Channel Probing: In the downlink channel probing, for every EN u, beamformer \( w_u \) developed in the link establishment stage is applied to receive the signal. Due to the large path loss and high directionality of mmWave signal, it is possible to increase the probing efficiency by sending the downlink channel probing signals concurrently. That is, for every single round of channel probering, CN only needs to perform the channel probing once. For instance, for the CN c, the downlink probing signals are concurrently sent using a beamformer \( f_{cu} \) to \( M - 1 \) ENs, which means the downlink channel probing is performed in a multiplexing manner. Here, \( f_{cu} \) is the hybrid beamformer and it incorporates the analog beamformer and a baseband precoder. The design of the hybrid precoder will be given in section IV-A.

For every EN u, if the interference from other ENs is denoted as \( I_i \), the corrupted downlink channel probing signals, received at EN u, can be represented as:

\[
y^2_{cu} = w_u^H H_{cu} f_{cu} s_u + \sum_{j \in M \setminus \{u,c\}} w_u^H H_{cu} f_{j} s_j + w_u^H n_{cu},
\]

\( I_i \)

In the next section, we will give more details about the design of multiplexing beamformer, which can mitigate the interference \( I_i \). The design of every beamformer \( f_{cu} \) will be given in the next section.

IV. EFFICIENT GROUP KEY GENERATION FOR THE STAR NETWORKS

In this section, we discuss the design of spatial-multiplexing beamformer at first, where the hybrid precoding is adopted. Based on the developed multiplexing beamformer, the group key generation protocols are proposed in section IV-B. In section IV-C, we will give the analytical expression for the group key rates. To further improve the group key rates at the downlink channel probing stage, the power allocation is reformulated into a multi-objective optimization problem in Eq. 28. Finally, in section IV-D, we propose an MLE-based group key rates estimator, which can efficiently estimate the group key rates based on the probing samples.

A. Probing Efficiency Improvement with Hybrid Precoding

As specified in the previous section, CN c has \( N_{RF} \) RF chains. Due to the hardware limitations of Massive MIMO, the analog beamformer combined with the baseband beamformer is usually adopted at mmWave band. The analog beamformer can be regrouped into a single RF precoder \( F_{RF} = [f_1, ..., f_M] \). If we add an extra baseband precoder \( F_{BB} \) at the front end of CN c [1, 2], the combined beamformer and training sequences can be expressed as:

\[
x = F_{RF} F_{BB} s
\]

where vector \( s = [s_1, s_2, ..., s_M]^T \) is the training sequence for channel probing and needs to meet \( E[s^H s] \leq P \). In this way, the received signal at each EN u can be expressed as:

\[
y^2_{cu} = \sum_u w_u^H H_{bu} x + w_u^H n_{cu},
\]
Due to the sparsity of channel $\mathbf{H}$, the solution in Eq. 1 can be satisfied with the matched beamformer [2]. Consequently, the EN $u$ sets $w_u = a_{EN}(\theta_u)$ and CN $c$ takes $v_{cu} = a_{CN}(\phi_{cu})$. Here $(\theta_u)$ and $(\phi_{cu})$ are quantized in the angular space and meet the specifications of the analog beamformers. After gathering the beamforming vectors for the $M-1$ ENs, the RF beamformer of CN $c$ can be represented as $\mathbf{F}_{RF} = [v_{c1}, ..., v_{cM}]$. 

Due to the design freedom provided by baseband precoder $\mathbf{f}_{BB}$ and high directability of mmWave communication, the effective channel of any EN $u \in M \setminus \{c\}$ can be viewed as the combination of antenna gains. Based on the uplink channel estimation, the effective antenna gain vector can accordingly be collected and expressed as $\hat{\mathbf{H}}_u = [\hat{\mathbf{h}}_1/g_{cu}, ..., \hat{\mathbf{h}}_M/g_{cu}]^T$, where $\hat{\mathbf{h}}_M \in C^{1 \times M}$ can be expressed as $\hat{\mathbf{h}}_u = w_{u}^H \mathbf{H}_{cu} \mathbf{F}_{RF}$. Derived from the zero-forcing techniques, the baseband precoder can be adopted here to achieve downlink channel probing concurrently instead of sequentially. We give the definition of the baseband precoder as below:

$$\mathbf{F}_{BB} = \hat{\mathbf{H}}_u (\hat{\mathbf{H}}_u \hat{\mathbf{H}}_u^H)^{-1}, \tag{7}$$

We defined the downlink probing symbols as $\mathbf{s} = [\sqrt{p_1}, ..., \sqrt{p_M}]$, where $\sqrt{p_u}$ is the power loaded on the baseband beamformer of EN $u$. If the effect of the channel estimation is considered, the designed baseband precoder $\mathbf{f}_{BB}$ of EN $j$ is orthogonal to the other effective channels $\hat{\mathbf{h}}_u$, $u \neq j$ and such an effect can be represented as:

$$w_{u}^H \mathbf{H}_{cu} \mathbf{F}_{RF} \mathbf{f}_{BB}^j \approx 0, \quad j \neq u, \quad j, u \in M \setminus \{c\} \tag{8}$$

If we adopt the sparse multi-path mmWave channel model [1, 2, 13], for EN $u$, the uplink and downlink signal can be further simplified into:

$$\begin{align*}
y_{cu}^1 & \approx \sqrt{\rho_{cu}} g_{cu} + \sqrt{p_u} h_{cu}, \\
y_{cu}^2 & \approx \sqrt{\rho_{cu}} g_{cu} + w_{u}^H n_{cu},
\end{align*} \tag{9}$$

where $\rho_{cu}$ is the allocated power at the RF chain and $\rho_{uc}$ is the power for uplink probing signals, $g_{cu}$ is the equivalent channel gain. Eq. 10 indicates that the strongest multipath has been selected in the link initialization stage and the effect of other multipath is filtered by the beamformer. The total power constraint is $\sum_{u=1}^{M} \rho_{cu} \leq P$ and $P$ is the total power budget. $g_{cu}$ is the fading coefficient. If the MLE is applied at both sides, accordingly, the measured channel gain of uplink and downlink can be represented as $h_{cu}^1$ and $h_{cu}^2$, respectively, as below:

$$\begin{align*}
\overline{h}_{c,u} &= \frac{1}{\sqrt{\rho_{uc}}} y_{cu}^1, \\
\overline{h}_{u,c} &= \frac{1}{\sqrt{\rho_{cu}}} y_{cu}^2,
\end{align*} \tag{11}$$

$$\begin{align*}
g_{cu} + G_{uc} h_{cu}^H n_{cu}, \\
g_{cu} + G_{cu} w_{u}^H n_{cu},
\end{align*} \tag{12}$$

Based on Eq. 12, the equivalent channel gain $g_{cu}$ can be estimated.

### B. Group Key Generation Protocols for Star Networks

In section IV-A, the multiplexing of downlink probing signals has been given. The group key generation protocol for mmWave star networks, which consists of 5 steps, will be specified as below:

**Step 1. RF Precoder Design:** Right before the uplink channel probing, CN and every EN need to find the RF precoder that can provide the largest effective gain according to [2]. This problem can be viewed in Eq. 1. This procedure is usually named as beam sweepings.

**Step 2. Uplink Channel Probing and Effective Channel Estimation:** Based on the designed RF beamformer $\mathbf{F}_{RF}$, every EN $u$ sends channel probings sequentially. In the uplink, CN $c$ estimates the effective channels of every ENs in set $M/\{c\}$. The collection of effective antenna gain at CN $c$ can be represented as $\hat{\mathbf{h}}_u$. Based on the effective antenna gain matrix $\hat{\mathbf{H}}_u$, the baseband precoder $\mathbf{F}_{BB}$ is specified.

**Step 3. Downlink Channel Probing and Baseband Precoder Design:** After the precoder design in the last step, at the CN $c$, analog precoder $\mathbf{f}_{RF} = [v_1, ..., v_M]$ and baseband precoder $\mathbf{f}_{BB}$ have been constructed. In this step, CN $c$ needs to perform the downlink channel probing, which results in the received signal at every EN $u$ as Eq. 10.

**Step 4. Reference Channel Selection and Broadcasting:** In this step, CN $c$ randomly picks up EN $j$ from $M \setminus \{c\}$ and the effective channel $\hat{h}_j$ is selected as the reference channel. CN $c$ gets the difference between the reference channel $\hat{h}_j$ and effective channel matrix as $\Delta = \hat{h}_{c,j} - \hat{h}_{c,1},..., \hat{h}_{c,j} - \hat{h}_{c,M}$.

**Step 5. Group Channel Reconstruction:** As claimed in [15], in the process of group key generation, current works usually assume there is a noiseless public channel with infinite capacity to exchange messages among group members. Apparently, such a public channel can be completely accessed by the eavesdropper. Based on the channel difference $\Delta$ broadcasted to every EN $u$ over the public channel, each EN $u$ can retrieve the reference channel $\hat{h}_{c,u}$ based on the estimation of downlink effective channel $\hat{h}_{u,c}$ by solving $\hat{h}_{c,u} - \hat{h}_{c,j}$, where we assume the reciprocity holds and $\hat{h}_{c,u} \approx \hat{h}_{u,c}$.

### C. Group Key Rates and Power Allocation

The efficiency of group key generation can be further improved by appropriate power allocation on the baseband beamformer. That is, we are trying to tune the power of baseband beamformer at CN to increase the group key rates. Utilizing the sleipan-wolf coding as specified in [10], let $R_{sec}^u$ represent the secure key rates at the EN $u$. It can be represented as:

$$R_{star} \triangleq \min_{i \in A_{ij}} \lim_{T \to \infty} I([\tilde{h}_{c,1}, ..., \tilde{h}_{c,M}]; [\Delta, \tilde{h}_{c,j}]) \tag{13}$$

$$R_c \triangleq \lim_{T \to \infty} I([\tilde{h}_{c,1}, ..., \tilde{h}_{c,M}]; [\Delta, Y^c]) \tag{14}$$

$$R_{sec}^u = R_{star} - R_c \tag{15}$$

where $\tilde{h}_{c,j}$ is the uplink reference channel between the CN $c$ and the EN $j$. Considering the channel difference $\Delta$, the secure
key rates $R_{sec}^u$ can be expressed as:

$$R_{sec}^u = R_{star} - R_e$$

$$= I(\mathbf{h}_{c,j}; \mathbf{h}_{u,c}(\Delta))$$

$$= H(\mathbf{h}_{c,j}|\Delta) - H(\mathbf{h}_{c,j}|\mathbf{h}_{u,c}, \Delta)$$

$$= H(\mathbf{h}_{u,c}) + H(\mathbf{h}_{c,j}) - H(\mathbf{h}_{c,j} - H(\mathbf{h}_{u,c}, \Delta))$$

$$= H(\mathbf{h}_{u,c}) + H(\mathbf{h}_{c,j}) - H(\mathbf{h}_{u,c}, \Delta)$$

$$= H(\mathbf{h}_{u,c}) + H(\mathbf{h}_{c,j}) - H(\mathbf{h}_{u,c}, \Delta)$$

$$= H(\mathbf{h}_{u,c}) + H(\mathbf{h}_{c,j}) - H(\mathbf{h}_{u,c}, \Delta)$$

$$+ H(\mathbf{h}_{u,c})$$

(16)

which is the modified group key rates based on [8]. From Eq. 16, we can observe that secure group key rates $R_{sec}^u$ for EN $u$ can be decoupled into the combination of joint entropy $H(\mathbf{h}_{u,c}, \mathbf{h}_{c,j})$ and $H(\mathbf{h}_{u,c}, \mathbf{h}_{c,j})$ and $H(\mathbf{h}_{u,c}, \mathbf{h}_{c,j})$. If we take a term $H(\mathbf{h}_{u,c}, \mathbf{h}_{c,j})$, the joint entropy is defined by the joint distribution of the vector $\mathbf{h}_{u,c}, \mathbf{h}_{c,j}$. Such a vector is the linear combination of estimated effective channel gains and its distribution needs to be derived. The joint distribution of terms in Eq. 16 will be given as follows.

At first, according to [1], we assume the fading coefficient $g_{en}$ follows a circular-symmetric complex gaussian $g_{u} \sim CN(0, \sigma^2_{g_u})$. Noise $n \in CN_N \times 1$ for every EN has a circular-symmetric complex gaussian distribution $n \sim CN(0, \sigma^2_{n})$. We list the distribution of terms in Eq. 16 in the next few paragraphs. The derivation in detail is omitted due to the page limit.

$a)$ : For every single downlink effective channel gain $\mathbf{h}_{u,c}$ and uplink effective channel gain $\mathbf{h}_{u,c}$, based on the distribution of $g_{en}$ and $n$, we can easily find the derived distributions as:

$$\mathbf{h}_{u,c} \sim CN(0, (\sigma^2_{g_u} + |G_{en,c}|^2)||\mathbf{f}_{u,c}|^2 \sigma^2_{n})$$

$$\mathbf{h}_{u,c} \sim CN(0, (\sigma^2_{g_u} + |G_{en,c}|^2)||\mathbf{w}_{u,c}|^2 \sigma^2_{n})$$

(17)

$b)$ : The vector $\{\mathbf{h}_{c,j}, \ldots, \mathbf{h}_{c,j}\}$ contains the negative of uplink channel gains, except for the reference channel gain.

The corresponding distribution can be expressed as:

$$\mathbf{h}_{c,j} \sim CN(0, (\sigma^2_{g_u} + |G_{en,c}|^2)||\mathbf{f}_{c,j}|^2 \sigma^2_{n})$$

$$+ |G_{en,c}|^2||\mathbf{f}_{c,j}|^2 \sigma^2_{n}, \ldots, |G_{en,c}|^2||\mathbf{f}_{c,j}|^2 \sigma^2_{n})$$

(18)

$c)$ : Vector $\mathbf{h}_{c,j}$ contains the effective channel difference between the effective channel gain of any EN $u$ and the reference effective channel gain $\mathbf{h}_{c,j}$. The distribution of vector $\mathbf{h}_{c,j}$ can be represented as:

$$\mathbf{h}_{c,j} \sim CN(0, (\sigma^2_{g_u} + |G_{en,c}|^2)||\mathbf{f}_{c,j}|^2 \sigma^2_{n})$$

$$+(1/(M-1))(|G_{en,c}|^2)||\mathbf{f}_{c,j}|^2 \sigma^2_{n}, u \neq j$$

(19)

d) : The distribution of $\{-\mathbf{h}_{c,1}, \ldots, \mathbf{h}_{c,M}\} \setminus \{-\mathbf{h}_{c,u}, \mathbf{h}_{c,j}\}$ can be derived as:

$$\{-\mathbf{h}_{c,1}, \ldots, \mathbf{h}_{c,M}\} \setminus \{-\mathbf{h}_{c,u}, \mathbf{h}_{c,j}\} \sim CN(0, \sigma^2_{n})$$

$$\text{diag}(\sigma^2_{g_u} + |G_{en,c}|^2)||\mathbf{f}_{c,j}|^2 \sigma^2_{n})$$

(20)

$e)$ : Finally, the distribution of $\{-\mathbf{h}_{c,u}, \mathbf{h}_{u,c}\}$ and $\{\mathbf{h}_{u,c}, \mathbf{h}_{c,j}\}$ can be expressed in Eq. 21 and Eq. 22, in which $c \in C^{M-1}$, whose elements are zeros except for the $u_{th}$ element.

Based on the derivations above, the secure key rates can be expressed as:

$$R_{sec}^u = R_{star} - R_e$$

$$= \log((\pi e)^{M-1}) \text{Cov}(\{-\mathbf{h}_{c,1}, \ldots, \mathbf{h}_{c,M}\} \setminus \{-\mathbf{h}_{c,j}\})$$

$$- \log((\pi e)^{M-2}) \text{Cov}(\{\mathbf{h}_{c,1}, \ldots, \mathbf{h}_{c,M}\} \setminus \{\mathbf{h}_{c,u}\})$$

$$- \log((\pi e)^{M-1}) \text{Cov}(\mathbf{h}_{u,c})$$

$$+ \log((\pi e)^{M}) \text{Cov}(\{\mathbf{h}_{u,c}\})$$

(21)

In order to compute $\text{Cov}(\{\mathbf{h}_{u,c}\})$, the Sylvester’s determinant theorem is adopted, which represents the determinant of covariance matrix in a block-wise form as specified below in Eq. 26. When $D$ is invertible, let $B = c, C = c^T$, and $D = \sigma^2_{g_u} + |G_{en,c}|^2||\mathbf{w}||^2 \sigma^2_{n}$.

$$\text{det}(A \ B) = \text{det}(A) \text{det}(D - CA^{-1}B)$$

(26)

To further simplify and get the analytical expression of Eq. 26, the matrix can be reformulated as:

$$\text{det}(D - CA^{-1}B)$$

$$= \log(\sigma^2_{g_u} + |G_{en,c}|^2||\mathbf{w}||^2 \sigma^2_{n})$$

$$- \frac{1}{\sigma^2_{g_u} + \sigma^2_{g_j} + 2|G_{en,c}|^2||\mathbf{f}_{c,j}|^2 \sigma^2_{n}}$$

Finally, gathering all of derived expressions above, the simplified secure group key rates can be expressed as:

$$R_{sec}^u$$

$$= \log(\sigma^2_{g_u} + |G_{en,c}|^2||\mathbf{f}_{c,j}|^2 \sigma^2_{n}) - x_u$$

$$- \log(|G_{en,c}|^2 \sigma^2_{n}) + \log(\sigma^2_{g_u} + |G_{en,c}|^2||\mathbf{w}||^2 \sigma^2_{n})$$

$$- \frac{1}{\sigma^2_{g_u} + \sigma^2_{g_j} + 2|G_{en,c}|^2||\mathbf{f}_{c,j}|^2 \sigma^2_{n}}$$

(27)

where $x_u = \log(|G_{en,c}|^2||\mathbf{f}_{c,j}|^2 ||\mathbf{w}||^2 \sigma^2_{n})$. The assignment of transmission power $\rho = \{\rho_1, \ldots, \rho_M\}$, for the downlink channel probing, can be performed by adjusting the power on $M$ RF chains under the total power budget constraint at EN $c$, where vector $\rho$ needs to meet the constraint $\sum_{u=1}^{M} \rho_u \leq P$. Due to the fact that, for every EN $u$, the distribution of each equivalent channel gain $g_{en,c}$ is affected by the antenna pattern and the number of multipath, the coefficients of group key rates $R_{sec}^u$ are not the same between different nodes. According to the definition in Eq. 15, the objective function of group key rates can be reformulated into a multi-objective function with the constraints on the power allocation, whose maximum can
be obtained by properly tuning the $\rho$. The objective function is represented as: \[ \min \left\{ \max \{ R_{sec}^1(\rho), \ldots, R_{sec}^M(\rho) \} \right\} \] \] while maximizing the group key rates, the CN $c$ can solve the following problem:

\[
\begin{align*}
\max_{\rho} & \quad [ R_{sec}^1(\rho), R_{sec}^2(\rho), \ldots, R_{sec}^M(\rho) ] \\
\text{subject to} & \quad \sum_{u=1}^M \rho_{cu} \leq P \tag{28}
\end{align*}
\]

We can observe that the objective function in Eq. 28 is a multi-objective function, which usually has a set of solutions and we need to achieve a tradeoff among objectives and constraints. In section V, a GA-based algorithm will be developed for problem 28.

\textbf{D. MLE-based Entropy Estimation}

In this section, we propose a lightweight entropy estimation scheme based on Maximum Likelihood Estimation (MLE), in which the joint entropy in Eq. 27 is derived based on the parameter of joint gaussian distribution, estimated from MLE.

From the distribution derived in Eq. 12, we can observe that if we collect the measurements of effective channels for the EN $u$ in a single vector $\mathbf{h}_{c,u}$, the real/image part for $\mathbf{h}_{c,u}$ which has a circular-symmetric gaussian distribution, is a normal distribution $N(\operatorname{Re}(\mathbf{h}_{c,u}), \operatorname{Re}(\mu), \operatorname{Re}(\Sigma))$ or $N(\operatorname{Im}(\mathbf{h}_{c,u}), \operatorname{Im}(\mu), \operatorname{Im}(\Sigma))$. For instance, for the ease of presentation, we omitted the constant terms in the likelihood $\log N(\operatorname{Re}(\mathbf{h}_{c,u}), \operatorname{Re}(\mu), \operatorname{Re}(\Sigma))$ and terms involving unknowns can be represented as:

\[
\log(\cdot) \triangleq -\frac{1}{2} \log |\Sigma| + \frac{1}{2} (\mathbf{h}_{c,u} - \operatorname{Re}(\mu))^{T} \Sigma (\mathbf{h}_{c,u} - \operatorname{Re}(\mu)), \tag{29}
\]

Here we take the MLE for real part as an example. By setting derivative w.r.t. $\operatorname{Re}(\mu)$ to zero, we can derive the estimated real part of mean $\mu$ as:

\[
\operatorname{Re}(\hat{\mu})_{\text{ML}} = \frac{1}{|\mathbf{h}_{c,u}|} \sum_{i} \operatorname{Re}(\mathbf{h}_{c,u})^i \tag{30}
\]

\[
\operatorname{Re}(\hat{\Sigma})_{\text{ML}} = \frac{1}{|\mathbf{h}_{c,u}|} \sum_{i} (\mathbf{h}_{c,u} - \operatorname{Re}(\hat{\mu})_{\text{ML}})^i \tag{31}
\]

where $|\mathbf{h}_{c,u}|$ indicates the number of elements in vector $\mathbf{h}_{c,u}$. Based on the two estimators above, the joint gaussian distribution of the joint entropy terms listed in Eq. 27 can be estimated accordingly. The numerical results in section VI show that the theoretical group key rates are closely in agreement with the estimation of group key rates.

\textbf{V. GROUP KEY RATES OPTIMIZATION USING GENETIC ALGORITHM}

\textbf{A. Multi-objective Optimization}

In Eq. 28, the power allocation problem for downlink probing is a multi-objective optimization problem, which involves more than one objective function to be optimized. As we all know, the answer for such a problem is a set of solutions, where the goodness of a solution for the multi-objective problem is determined by the dominance. Let’s define the notion of dominance as below:

\textbf{Remark 1.} For two feasible solutions $\mathbf{x}_1$ and $\mathbf{x}_2$, $\mathbf{x}_1$ dominates $\mathbf{x}_2$ if, solution $\mathbf{x}_1$ is no worse than $\mathbf{x}_2$ in all objectives, and solution $\mathbf{x}_1$ is strictly better than $\mathbf{x}_2$ in at least one objective.

One way to find good solutions to multi-objective problems is according to \textit{Pareto optimality}, named after economist Vilfredo Pareto.

1) \textit{Pareto Front}: For the multi-objective problem, maximizing or minimizing a single objective function may be harmful to other objective functions. We are interested in finding solutions that upgrade some objective functions without downgrading anyone else. The movement (upgrading) from the previous solutions to a set of better solutions is called "Pareto improvements". If the current solutions are restricted to a solution set and cannot make Pareto improvements in the next solution updates, the \textit{Pareto optimality} is achieved. Many existing multi-objective optimization algorithms involve the concept of Pareto optimality [4]. Any point belonging to this set is said to be on a front called \textit{Pareto front}.

\[
\{ -h_{u,c}, h_{u,c} \} \sim CN(\mathbf{0}, \begin{bmatrix} \sigma_{g,u}^2 + |G_{u,c}|^2 ||\mathbf{f}(\cdot)||^2 \sigma_n^2 & -\sigma_{g,u}^2 \sigma_{g,u}^2 + |G_{u,c}|^2 ||\mathbf{w}_{hu}||^2 \sigma_n^2 \sigma_n^2 \\ -\sigma_{g,u}^2 & \sigma_{g,u}^2 + |G_{u,c}|^2 ||\mathbf{f}(\cdot)||^2 \sigma_n^2 \end{bmatrix}) \tag{21}
\]

\[
\mathbf{A} = \begin{bmatrix} \sigma_{g,1}^2 + |G_{1,c}|^2 ||\mathbf{f}(\cdot)||^2 \sigma_n^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_{g,M-1}^2 + |G_{M-1,c}|^2 ||\mathbf{f}(\cdot)||^2 \sigma_n^2 \end{bmatrix} + \begin{bmatrix} \sigma_{g,1}^2 & \cdots & \sigma_{g,1}^2 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{bmatrix} \tag{22}
\]

\[
\{ \mathbf{A}, h_{u,c} \} \sim CN(\mathbf{0}, \begin{bmatrix} \mathbf{A} & \mathbf{c} \\ \mathbf{c}^T \sigma_{g,u}^2 + |G_{c,u}|^2 ||\mathbf{w}(\cdot)||^2 \sigma_n^2 \end{bmatrix}) \tag{23}
\]

\[
\mathbf{c}^T = \begin{bmatrix} 0, \ldots, (\sigma_{g,u}^2), \ldots, 0 \end{bmatrix} \tag{24}
\]
2) Existing algorithms: For multi-objective optimization, classical optimization methods tend to convert the multi-objective optimization problem to a single-objective optimization problem, i.e., averaged weighting, which can only emphasize one particular Pareto-optimal solution at a time. To find multiple solutions, classical methods have to be applied many times, hopefully finding a different solution at each simulation run.

B. Non-dominant Sorting Genetic Algorithm for Constrained Optimization

In this paper, to solve the multi-objective defined in Eq. 28, we customize a Genetic Algorithm called Non-dominant Sorting Genetic Algorithm (NSGA II) and modify constraints handling process for NSGA II. In the next few paragraphs, we will give more details about the group key generation scheme based on NSGA II. Right before giving the scheme and the NSGA II in detail, we list several common notions in GA, non-dominant set sorting, crowding distance assignment and reproduction.

1) Non-dominant Set Sorting: NSGA II is based on an operation by sorting the population into several Non-dominated solution sets. According to [4], the Non-dominated solution set can be defined as:

Remark 2. Given a set of solutions, the non-dominated solution set is a set of all the solutions that are not dominated by any member of the solution set.

The non-dominated set of the entire feasible decision space is called the Pareto-optimal set. The boundary defined by the set of all points mapped from the Pareto optimal set is called the Pareto optimal front. In order to approach Pareto optimality front, the overall population, in each round of algorithm, is divided into several fronts \( \mathcal{F} = \cup F_i \).

To sort the non-dominated set, for every solution \( p \) in the population, two quantities are specified: 1) domination count \( n_p \) and 2) \( S_p \), a set containing the solutions dominated by \( p \).

2) Crowding Distance Assignment: As mentioned earlier, it is desired that the obtained solution sets spread widely in the feasible region, along with convergence to the Pareto-optimal set. In NSGA II, crowding distance for every element in the front indicates its distance to near neighbors and guides the selection process toward the uniformly sampled Pareto-optimal front.

- Density Estimation: To get an estimate of the density of solutions surrounding a particular solution in the population, crowding distance serves as an estimate of the perimeter of the cuboid formed by using the nearest neighbors as the vertices. To compute the crowding-distance, the population, at first, needs to be sorted in ascending order of magnitude according to the value of each objective function. Secondly, the boundary solutions for each objective function are assigned an infinite distance value. The rest intermediate solutions are assigned a distance value equal to the absolute normalized difference in the function values of two adjacent solutions.
- Crowded-Comparison Operation: With the crowding distance defined above, a spread-out solution set is selected at various stages of the algorithm by using the crowded-comparison operator \((<)\), with which we hope the solution set can approach to a uniformly spread-out Pareto optimal front. After two steps above (non-dominated sorting and crowding-distance assignment), every entry in the population is assigned two attributes: nondomination rank \( (i_{rank}) \) and crowding distance \( (d_{distance}) \). The crowded-Comparison Operator is defined as:

\[
i < n j \text{ if } \begin{cases} i_{rank} < j_{rank} \\ or \ (i_{rank} = j_{rank} \text{ and } d_{distance} > d_{distance}) \end{cases}\]

By observing operator in Eq. 32, we notice that between two solutions with differing nondomination ranks, the lower (better) rank solution is preferred. Otherwise, if two solutions locate on the same front, then we prefer the solution that is located in a less crowded region.

Algorithm 1 GA-based Group key generation for multi-users mmWave Massive MIMO systems

**Initialization:** Establish links between the CN and EN. The distribution of CSI. Power budget \( P \). SNR for each round. Max generation size \( N_G \)

**Output:** Power allocation result \( \rho \)

For \( t \leq N_G \)

\( R_t = P_t \cup Q_t \)

\( \mathcal{F} = \) non-dominated sorting \( (R_t) \)

\( P_{t+1} = 0 \), and \( i = 1 \)

until\( |P_{t+1}| + |\mathcal{F}_i| \leq N \)

- crowding-distance assignment \( (\mathcal{F}_i) \)
- \( P_{t+1} = P_{t+1} \cup \mathcal{F}_i \)
- \( i = i + 1 \)

Sort \( \mathcal{F}_i \), \(<_n \)

\( P_{t+1} = P_{t+1} \cup \mathcal{F}_i \{1 : (N - |P_{t+1}|)\} \)

\( Q_{t+1} = \) Reproduction \( (P_{t+1}) \)

\( t = t + 1 \)

3) Reproduction: The non-dominated sorting above divided the solutions into several fronts. The entries within the same front can utilize genetic operators like crossover, mutation and selection to produce a new generation (children).

Here we give a brief introduction for them, which have various roles as the genetic operators.

- Crossover: For every single solution \( p \), after the non-dominated sorting, several pieces of solution \( p \) are exchanged with other solutions at the same solution representations. This means the crossover works in a well-searched subspace, and the converged states will remain.
- Mutation: The operation of Mutation usually changes parts of one solution randomly, which increases the diversity of the population and provides a mechanism for escaping from a local optimum. That is, mutation usually leads to a solution outside the current solution subspace.

4) Constraints Handling: If the infeasible solutions violating constraints marginally are placed in the same nondominated level with another solution violating constraints to a large extent, this may cause an algorithm to wander in the infeasible...
search region for more generations before reaching the feasible region through constraint boundaries. In NSGA II, the feasible solutions with large crowding distance are preferred and the infeasible solutions violating constraints will be discarded.

C. GA-based Group Key Generation

We have introduced the key concepts of NSGA II above and gathered all the pieces. In this section, we give the GA-based Group Key Generation algorithm in Algorithm 1.

In the Algorithm 1, a single algorithm run is listed. \( R \) contains the parent population \( P \), and offspring population \( Q \). The elements in \( R \) are classified into several non-dominant set front after the non-dominant sorting. In order to reach a uniformly spreading solution set, crowding distance for each element is computed and serves as the input of crowded comparison. After the crowded comparison sorting, the non-dominant set or the parent set produces the next generation (offspring). Current research has shown that NSGA II has a good performance on non-convex problems and can achieve the approximately uniformly spreading solution set.

VI. NUMERICAL RESULTS

In this section, numerical results are given to illustrate the performance of the group key generation scheme for the star-topology networks. The mmWave MIMO system operates at 28 GHz. We consider mmWave Massive MIMO system adopting analog phase-shifter with multiple radio-frequency (RF) chains. Uniform linear array (ULA) is adopted at CN and EN side with the dimension, \( 16 \times 16 \). For NSGA II algorithm, we set the population size as 400 and the generation size as 600. The mutation probability is set to \( 1/M \). The initial population is randomly generated within the range \((0, P)\).

For random variable \( X = x \), the quantization level is designed according to the cumulative distribution function (CDF) \( F_q(x) \). The range of random variable \( x \) is quantized into \( q \) equally likely regions defined by the quantization boundaries \( x_i = F^{-1}_q(i/q) \) for \( i = 0, 1, \ldots, q \). After quantization, for each \( x \), \( \log_2(q) \) bits are quantized each time.

For the GA based power allocation, the group key rates are given to prove that the minimum of group key rates can be maximized by assigning the power appropriately. For comparison purposes, we provide the performance of uniform power allocation strategy similar to the scheme provided in [10]. In Fig. 1, the blue line represents the GA based power allocation for the group key generation. The line in red indicates the existing channel probing scheme. In both cases, the solid lines in blue/red represent the theoretical group key rates as derived in Eq. 27 and the dotted lines are the estimated group key rates based on Eq. 31. From this figure, it can be observed that the theoretical key rates match the estimated group key rates estimated from the probing samples. Besides, the proposed GA-based power allocation scheme for the group key rates outperforms the case with uniformly distributed powers, and the case with the conventional probing scheme. The gap between GA-based scheme and the existing probing scheme becomes larger as the SNR increases, which indicates the GA-based algorithm has a better performance in the high SNR regime.

In Fig. 2, we provide the group key rates under different number of ENs. The solid line represents the theoretical bound. The dotted line represents the estimated group key rates. It can be observed that the group key rates decrease with the group size. This is due to the fact that the noise accumulated at the channel probing stage leads to the decreasing of group key rates. The important observation is that the group key rates of GA-based scheme and existing probing scheme approach to be same if the group size is larger than 7, which indicates that for the proposed downlink probing scheme, the smaller multiplexing factor is preferred.

Fig. 3 depicts the numerical results of bits disagreement ratio (BDR). The quantization level is set to be 4. From this figure, in the high SNR regime, it can be observed that the GA-based power allocation downlink probing outperforms the scheme that utilizes the uniform power allocation. However, BDR of GA-based power allocation and the existing probing scheme tend to have the same values in the high SNR regime. Based on Fig. 1 and Fig. 3, we can observe that channel probings with the higher energy can reduce the BDR efficiently. However, in a group, the actual group key rates are influenced by the power allocation strategy.

In Fig. 4, if we set the \( SNR = 25dB \), the BDR versus different group size from 3 to 8 are provided. We can observe...
that the BDR of GA-based algorithm and existing channel probing scheme, represented by the bars in blue and red, have lower value compared with the BDR of uniform power allocation scheme. Besides, for the GA-based algorithm and current channel probing scheme, the BDR will grow as the increasing of group size. However, for the uniform power allocation, within a range of the group size between 3 to 5, the BDR is insensitive to the changing of group size. For the group size larger than 5, the BDR of the uniform power allocation scheme will increase accordingly.

VII. CONCLUSION

An efficient channel probing strategy for group key generation in mmWave Massive MIMO networks has been proposed, which is based on the hybrid precoding and the NSGA II algorithm. Besides, a scheme for the group key rates estimation has been developed based on MLE. In the proposed group key generation scheme, a baseband precoder has been applied in the downlink probing, which enables the CN to send downlink channel probing signals to several ENs concurrently. Besides, to further improve the group key rates, the NSGA II is modified to optimize the power allocation for the downlink prober signals, which can reach a spread-out solution region and achieve the Pareto optimality. The estimated group key rates match the theoretical key rates.

ACKNOWLEDGEMENT

This work was supported in part by the Commonwealth Cyber Initiative (CCI) and its Northern Virginia (NOVA) Node, an investment in the advancement of cyber RD, innovation and workforce development. For more information about CCI, visit cyberinitiative.org.

REFERENCES