IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received July 21, 2019, accepted August 1, 2019, date of publication August 6, 2019, date of current version August 21, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2933559

Secure Hybrid Digital and Analog Precoder for mmWave Systems With Low-Resolution DACs and Finite-Quantized Phase Shifters

LING XU[1], LINLIN SUN[1], GUIYANG XIA[1], TINGTING LIU[2], (Member, IEEE), FENG SHU[1], (Member, IEEE), YIJIN ZHANG[1], AND JIANGZHOU WANG[3], (Fellow, IEEE)

[1]School of Electronic and Optical Engineering, Nanjing University of Science and Technology, Nanjing 210094, China
[2]School of Information and Communication Engineering, Nanjing Institute of Technology, Nanjing 211167, China
[3]School of Engineering and Digital Arts, University of Kent, Canterbury CT2 7NT, U.K.

Corresponding author: Feng Shu (shufeng@njust.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61771244, Grant 61472190, Grant 61702258, Grant 61501238, and Grant 61602245.

ABSTRACT Millimeter wave (mmWave) communication has been regarded as one of the most promising technologies for the future generation wireless networks because of its advantages of providing a ultra-wide new spectrum and ultra-high data transmission rate. To reduce the power consumption and circuit cost for mmWave systems, hybrid digital and analog (HDA) architecture is preferred in such a scenario. In this paper, an artificial noise (AN) aided secure HDA beamforming scheme is proposed for mmWave multiple input single output (MISO) system with low resolution digital-to-analog converters (DACs) and finite-quantized phase shifters on radio frequency. The additive quantization noise model for AN aided HDA system is established to make an analysis of the secrecy performance of such systems. With the partial channel knowledge of eavesdropper available, an approximate expression of secrecy rate (SR) is derived. Then using this approximation formula, we propose a two-layer alternatively iterative structure (TLAIS) for optimizing digital precoder (DP) of confidential message (CM), digital AN projection matrix (DANPM) and analog precoder (AP). The inner-layer iteration loop is to design the DP of CMs and DANPM alternatively given a fixed matrix of AP. The outer-layer iteration loop is between digital baseband part and analog part, where the former refers to DP and DANPM, and the latter is AP. Then for a given digital part, we propose a gradient ascent algorithm to find the AP vector. Given a matrix of AP, we make use of general power iteration (GPI) method to compute DP and DANPM. This process is repeated until the terminal condition is reached. Simulation results show that the proposed TLAIS can achieve a better SR performance than existing methods, especially in the high signal-to-noise ratio region.

INDEX TERMS Hybrid digital and analog, mmWave, security, artificial noise, low-resolution digital-to-analog converter (DAC).

I. INTRODUCTION

With the rapid development of wireless communication technologies as well as explosive access of mobile terminals, the demand of wireless data traffic grows exponentially. However, the spectrum of current existing wireless systems is highly congested, which brings a bottleneck for increasing wireless access rate ulteriorly. Millimeter wave (mmWave) communication, whose frequency band ranges from 30GHz to 300GHz, emerges as a promising approach addressing the problem of spectrum congestion [1]–[4].

Nevertheless, although the mmWave communication becomes more and more popular due to its abundant available and under-utilized spectrum, its realistic applications are under many constraints such as the severe free space path loss and rain attenuation [5]. To overcome these challenges, multiple-input-multiple-output (MIMO) is usually adopted to compensate for these shortages. Meanwhile, hybrid digital and analog (HDA) architecture has been employed to further reduce the energy consumption of mmWave

The associate editor coordinating the review of this manuscript and approving it for publication was Antonino Orsino.

MIMO systems [6]–[13]. The design of hybrid precoder was formulated as a problem of sparse signal reconstruction by exploiting the inherent spatial sparse structure of mmWave channel in [6], [7], where orthogonal matching pursuit (OMP) algorithm was used to achieve a comparatively good performance. In [8], the authors proposed three methods of innovative alternative minimization for both fully-connected and partially-connected hybrid combiner in mmWave MIMO systems, which confirms the feasibility and effectiveness of hybrid precoder in mmWave MIMO architecture. In [9], an energy-efficient successive-interference-cancelation-based algorithm with low complexity was proposed for mmWave MIMO systems. Authors in [10] and [11] extended conventional flat-fading mmWave channels to broadband frequency-selective mmWave channels, where the joint analog beamforming for the entire band and respective baseband precoder for each sub-band were required to be designed carefully. Specifically, in [10] the heuristic algorithms in OFDM systems were developed for two scenarios: single-user MIMO and multiple-user multiple-input-single-output scenarios. In [11], the authors developed a novel method by dynamically establishing the structure of sub-arrays. In [12], the hybrid analog and digital (HAD) architecture was used at receiver to make a measurement of direction of arrival (DOA). Here, three low-complexity DOA estimation methods were proposed and the corresponding HAD Cramer-Rao lower bound was also derived. In [13], a robust beamforming method using HAD receive structure was proposed to achieve interference compression.

However, the aforementioned research works mainly focused on HDA mmWave systems without taking security into consideration. Recently, security became a hot issue in many research filed [14]–[18]. Wyner first proposed a discrete, memoryless wiretap channel model to investigate the secrecy of wireless communication [19]. Based on the framework of Wyner's wiretap channel, multiple-antenna technique [20] was applied to enhance security and beamforming has been proven to be secrecy-capacity-achieving under the circumstances that the desired receiver has single antenna and full channel knowledges are available for all terminals [21], [22]. If the transmitter can not obtain the simultaneous channel state information (CSI) of eavesdropper or only know the partial CSI of eavesdropper, artificial noise (AN) was shown to be an effective aiding method to strengthen security [23]–[26]. Recently, some researchers paid a close attention to security in mmWave MIMO channel. Authors in [27] made a systematical investigation of the security performance in HDA mmWave systems under two different CSIs: full and partial. Here, maximum ratio transmission (MRT) method was extended to such scenarios and minimum secrecy outage probability was adopted to design the corresponding hybrid precoder. Considering the sparse features of mmWave channels in [28], a discrete angular domain channel model was proposed to meticulously derive the secure performance for the proposed transmission schemes in slow fading multipath channels. In [29], an AN-aided hybrid precoder was proposed to maximize the lower bound of average secrecy rate (SR).

In practical applications, if low-resolution digital-to-analog converters (DACs) or analog-to-digital converters (ADCs) are adopted, the average power consumption of hybrid transceivers will be greatly reduced. Actually, a high - resolution DACs/ADCs can be power-hungry [30]. In [31], the problem of striking a trade-off between energy efficiency and spectrum efficiency was investigated for analog, hybrid, and digital receivers with low resolution ADCs, respectively. To further reduce energy consumption, hybrid beamforming and digital beamforming with low resolution ADCs were studied in [32]. Here, the authors found that digital beamforming equipped with finite resolution ADCs can achieve a higher achievable rate and can be more energy efficient than hybrid beamforming in the low signal-to-noise ratio (SNR) region. In [33], taking both low resolution DACs and RF losses in hybrid into account in mmWave MIMO system, a quantized hybrid transmitter with additive quantization noise model (AQNM) has been constructed for both fully-connected and partially-connected hybrid architecture, and a lower bound of achievable rate were presented.

However, to the best of our knowledge, there is little literature of making an investigation on how to achieve a security in hybrid mmWave systems with low resolution DACs. Therefore, in our paper, we propose an AN-aided hybrid mmWave transmitter with low resolution DACs and finite-quantized phase shifters. The transmitted signals with the help of AN are first precoded by digital precoder (DP) in baseband, then passing through low-resolution DACs and RF chains before analog precoder (AP). Here, the AQN model is adopted to approximate the quantized signal as a linear output, which simplifies the analysis of further secrecy rate. Since the optimization problem is non-convex and intractable to tackle, an alternate iteration algorithm is resorted to maximize the approximate expression of secrecy rate. Our main contributions are summarized as follows:

1) An AN-aided secure hybrid precoding system model with low-resolution DACs and finite-quantized phase shifters on RF is established. By taking two kinds of quantization errors (QEs) into consideration, the proposed model is completely distinct from conventional non-secure hybrid precoding system without considering QEs and the AQN quantized model is adopted in our system model to approximate the distortion from QE of low-resolution DACs as a linear output. With partial eavesdropping channel knowledge available, an approximate expression for secrecy rate (SR) is derived. This expression converts the original intractable problem into a more easy-to-handle one. Thus, the approximate SR (ASR) expression will significantly simplify the optimization and design of digital precoding (DP) vector of confidential messages, digital AN projection matrix (ANPM) and analog precoding (AP) vector in what follows.

2) A two-layer alternatively iterative structure (TLAIS) is proposed to maximize the ASR by taking quantization errors from both low-resolution DAC and phase shifters into consideration. Given AP vector, the DP vector of confidential messages and ANPM are alternatively computed within an interior iterative loop by making use of general power iterative (GPI) method with the aim to maximize the approximate SR. In particular, by complex Kronecker product manipulation, the problem of maximizing ASR with respect to the optimization matrix ANPM is converted into one with respect to an optimization vector being the vectorization of the corresponding ANPM. Given the DP vector and digital ANPM, a steepest ascent algorithm is used to attain the updated AP vector. The above process is repeated until the terminal condition is reached. Finally, the phases of AP vector is directly taken out as the inputs of the finite quantized phase shifters on RF. More importantly, to reduce the computation complexity, we abstract the non-zero elements in analog part by taking advantage of the sparsity of analog precoder.

The remainder of this paper is organized as follows. Section II presents the system model of AN-aided HDA mmWave multiple input single output (MISO) system with low resolution DACs and finite-quantized phase shifters, and the approximate expression of SR is given in this section. Based on the approximate formula of SR, a TLAIS among DP of confidential message, digital AN projection matrix (DANPM) and AP is proposed in Section III. Performance analysis and simulation evaluations are presented in Section IV. Finally, we draw our conclusions in Section V.

Notations: throughout the paper, matrices, vectors, and scalars are denoted by letters of bold upper case, bold lower case, and lower case, respectively. $(\cdot)^T$, $(\cdot)^*$ and $(\cdot)^H$ denote transpose, conjugate, and conjugate transpose, respectively. $\|\cdot\|_2$ and $\|\cdot\|_F$ denote the l_2 norm of a vector and Frobenius norm of a matrix, respectively. $\mathrm{Tr}(\cdot)$ and $\mathrm{vec}(\cdot)$ are matrix trace and matrix vectorization; $\otimes$ and $\odot$ indicate the Kronecker products and Hadamard products between two matrices, respectively. $\mathrm{diag}(\mathbf{A})$ returns a diagonal matrix consisting of the corresponding diagonal elements of matrix $\mathbf{A}$. $\mathrm{Diagblk}(\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_N)$ returns the block diagonal matrix concatenated of $\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_N$. And $\mathbb{N}[\mathbf{a}]$ returns a vector consisting of the non-zero elements in $\mathbf{a}$. $\mathbf{A}(m, n)$ denotes the element in m^{th} row and n^{th} column, $\mathbf{A}(m_1 : m_2, n_1 : n_2)$ returns a matrix consisting of m_1^{th} to m_2^{th} row and n_1^{th} to n_2^{th} column in $\mathbf{A}$.

II. CHANNEL AND SYSTEM MODELS

A. SYSTEM MODEL

In this section, we consider a system model with HDA transmitter of low resolution DACs as shown in Fig. 1. There are three network nodes: Alice, Bob, and Eve. Working as a transmitter, Alice uses a partially connected hybrid architecture in this paper, where each RF chain is connected to a subset of antennas. Assuming that the number of the transmit antennas

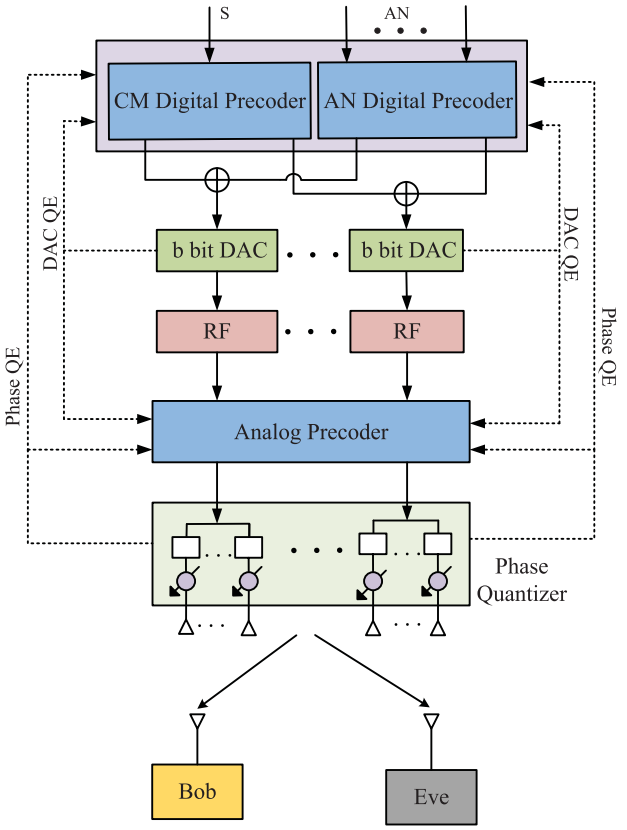


FIGURE 1. System model.

at transmitter is N_T, and the number of RF chains is K, it is clear that $N_T = KT$ with each sub-array having T antennas.

The transmit signal can be expressed as follows:

$$\mathbf{x} = \mathbf{F}_{RF} Q_b(\sqrt{\beta P}\mathbf{f}_{BB} s + \sqrt{(1-\beta)P}\mathbf{T}_{BB}\mathbf{z}), \tag{1}$$

where P is the effective transmitted power, β denotes the power allocation (PA) factor of confidential message, and $1-\beta$ indicates the PA factor of AN. $\mathbf{f}_{BB} \in \mathbb{C}^{K\times 1}$ and $\mathbf{T}_{BB} \in \mathbb{C}^{K\times K}$ represent the digital beamforming vector of confidential message and AN projection matrix, respectively. To satisfy the transmit power constraint, we have $\mathbb{E}\{\mathbf{x}\mathbf{x}^H\} = P_T$. $\mathbf{F}_{RF} \in \mathbb{C}^{N\times K}$ is the analog precoding matrix with the following structure:

$$\begin{aligned}\mathbf{F}_{RF} &= \mathrm{Diagblk}(\mathbf{f}_{RF1}, \cdots, \mathbf{f}_{RFK}) \\ &= \begin{bmatrix} \mathbf{f}_{RF1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{f}_{RF2} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{f}_{RFK}, \end{bmatrix},\end{aligned} \tag{2}$$

where $\mathbf{f}_{RFk}$ denotes a vector defined as

$$\mathbf{f}_{RFk} = [\exp(j\varphi_{k,1}), \exp(j\varphi_{k,2}), \cdots, \exp(j\varphi_{k,M})]^T, \tag{3}$$

where

$$\varphi_{k,i} \in \mathcal{F}_{RF} = \left\{0, \frac{2\pi}{2^{b_{ps}}}, \cdots, \frac{2\pi 2^{b_{ps}-1}}{2^{b_{ps}}}\right\}, \tag{4}$$

$$R_b = \log_2\left(1+\frac{\beta P(1-\eta)^2\mathbf{h}_b\mathbf{F}_{RF}\mathbf{f}_{BB}\mathbf{f}_{BB}^H\mathbf{F}_{RF}^H\mathbf{h}_b^H}{(1-\beta)P(1-\eta)^2\mathbf{h}_b\mathbf{F}_{RF}\mathbf{T}_{BB}\mathbf{T}_{BB}^H\mathbf{F}_{RF}^H\mathbf{h}_b^H+\mathbf{h}_b\mathbf{F}_{RF}\mathbf{R}_{\mathbf{n}_q\mathbf{n}_q}\mathbf{F}_{RF}^H\mathbf{h}_b^H+\sigma^2}\right). \tag{10}$$

$$R_e = \mathbb{E}\left[\log_2\left(1+\frac{\beta P(1-\eta)^2\mathbf{h}_e\mathbf{F}_{RF}\mathbf{f}_{BB}\mathbf{f}_{BB}^H\mathbf{F}_{RF}^H\mathbf{h}_e^H}{(1-\beta)P(1-\eta)^2\mathbf{h}_e\mathbf{F}_{RF}\mathbf{T}_{BB}\mathbf{T}_{BB}^H\mathbf{F}_{RF}^H\mathbf{h}_e^H+\mathbf{h}_e\mathbf{F}_{RF}\mathbf{R}_{\mathbf{n}_q\mathbf{n}_q}\mathbf{F}_{RF}^H\mathbf{h}_e^H+\sigma^2}\right)\right]. \tag{11}$$

where $\mathcal{F}_{RF}$ represents the set of b_{ps}-bit-quantized phases with b_{ps} being the number of quantization bits of phase shifters at RF. $Q_b(\cdot)$ represents the DAC quantization operation. In this paper, we apply the additive quantization noise (AQN) model in [31], [33] to approximate the DACs quantization as a linear output for the simplicity of analysis, which can be formulated as

$$Q_b(\mathbf{u}) \approx (1-\eta)\mathbf{u}+\mathbf{n}_q, \tag{5}$$

where η is defined as the reciprocal of signal-to-quantization-noise ratio. $\mathbf{n}_q$ is the additive quantization noise vector and uncorrelated with the input $\mathbf{u}$, that is, $\mathbb{E}[\mathbf{u}\mathbf{n}_q^H] = \mathbb{E}[\mathbf{n}_q\mathbf{u}^H] = \mathbf{0}$. Therefore, the transmitted signal in (1) can be rewritten as

$$\begin{aligned}\mathbf{x} \approx{}& \sqrt{\beta P}(1-\eta)\mathbf{F}_{RF}\mathbf{f}_{BB}s\\ &+\sqrt{(1-\beta)P}(1-\eta)\mathbf{F}_{RF}\mathbf{T}_{BB}\mathbf{z}+\mathbf{F}_{RF}\mathbf{n}_q.\end{aligned} \tag{6}$$

According to [31], the covariance of the quantization noise can be given by

$$\mathbf{R}_{\mathbf{n}_q\mathbf{n}_q} = \eta(1-\eta)\mathrm{diag}\{\mathbf{R}_{uu}\}, \tag{7}$$

where $\mathbf{R}_{uu} = \beta P\mathbf{f}_{BB}\mathbf{f}_{BB}^H + (1-\beta)P\mathbf{T}_{BB}\mathbf{T}_{BB}^H$. Therefore, the received signal at Bob can be represented as

$$\begin{aligned}y_b ={}& \sqrt{\beta P}(1-\eta)\mathbf{h}_b\mathbf{F}_{RF}\mathbf{f}_{BB}s\\ &+\sqrt{(1-\beta)P}(1-\eta)\mathbf{h}_b\mathbf{F}_{RF}\mathbf{T}_{BB}\mathbf{z}+\mathbf{h}_b\mathbf{F}_{RF}\mathbf{n}_q+n_b.\end{aligned} \tag{8}$$

Similarly, the receive signal at eavesdropper can be given as

$$\begin{aligned}y_e ={}& \sqrt{\beta P}(1-\eta)\mathbf{h}_e\mathbf{F}_{RF}\mathbf{f}_{BB}s\\ &+\sqrt{(1-\beta)P}(1-\eta)\mathbf{h}_e\mathbf{F}_{RF}\mathbf{T}_{BB}\mathbf{z}+\mathbf{h}_e\mathbf{F}_{RF}\mathbf{n}_q+n_e,\end{aligned} \tag{9}$$

where $n_b \sim \mathcal{CN}(0,\sigma_b^2)$ and $n_e \sim \mathcal{CN}(0,\sigma_e^2)$ are additive white Gaussian noise (AWGN) at Bob and Eve, respectively. For the convenience of analysis, we set $\sigma_b^2 = \sigma_e^2$.

Therefore, the achievable rate at Bob and Eve can be formulated in (10) and (11), as shown at the top of this page, respectively. And we can express the optimization problem of maximizing the SR as

$$\begin{aligned}&\max_{\mathbf{F}_{RF},\mathbf{f}_{BB},\mathbf{T}_{BB},\beta} R_s = \max\{0, R_b - R_e\}\\ &\text{subject to } \mathbf{F}_{RF}\in\mathcal{F}_{RF}, \mathbb{E}\{\mathbf{x}\mathbf{x}^H\} = P_T, \quad 0\le\beta\le 1.\end{aligned} \tag{12}$$

Since it is assumed that only the partial knowledge of wiretap channel at Eve is available, we can only obtain the average achievable rate of Eve. Meanwhile, because of partially-connected structure, we have $\mathbf{F}_{RF}^H\mathbf{F}_{RF} = \mathbf{I}_K$. To obey the power constraint, we should have $\|\mathbf{f}_{BB}\|^2 = \|\mathbf{T}_{BB}\|_F^2 = 1$, and $P = \frac{P_T}{(1-\eta)^2+\eta(1-\eta)}$. Therefore, the objective function in optimization problem (12) can be lower bounded as:

$$\begin{aligned}R_s &= R_b - \mathbb{E}[\log_2\{1+\mathrm{SINR}_e\}]\\ &\overset{(a)}{\ge} R_b - \log_2\{1+\mathbb{E}[\mathrm{SINR}_e]\}\\ &= \tilde{R}_s,\end{aligned} \tag{13}$$

where (a) holds due to

$$\mathbb{E}[\log_2(x)] \le \log_2(1+\mathbb{E}[x]). \tag{14}$$

Therefore, replacing the objective function R_s in (12) by $\tilde{R}_s$ forms the following simplified optimization problem

$$\begin{aligned}&\max_{\beta,\mathbf{f}_{BB},\mathbf{T}_{BB},\mathbf{F}_{RF}} \tilde{R}_s = R_b - \log_2\{1+\mathbb{E}[\mathrm{SINR}_e]\}\\ &\text{subject to } \mathbf{F}_{RF}\in\mathcal{F}_{RF}, \quad \|\mathbf{f}_{BB}\|^2 = 1,\ \|\mathbf{T}_{BB}\|_F^2 = 1,\\ &\qquad 0\le\beta\le 1.\end{aligned} \tag{15}$$

B. CHANNEL MODEL

In what follows, a narrow-band clustered mmWave channel model is used with L propagation paths, which can be described as follows:

$$\mathbf{h} = \sqrt{\frac{N_t}{L}}\sum_{l=1}^{L} g_l\mathbf{a}_{t,l}^H(\phi_l), \tag{16}$$

where $\sqrt{\frac{N_t}{L}}$ is the normalized factor, l stands for the index of paths, and L is the number of channel paths. The path gains $g_l \sim \mathcal{CN}(0,1)$ depicts the complex gain of the l^{th} path. $\mathbf{a}_t(\phi_l)$ is the corresponding response vector of transmit antenna array, with ϕ_l denotes the azimuth angles, respectively. For uniform linear array with N elements, the array response vector can be represented as

$$\mathbf{a}(\phi) = \frac{1}{\sqrt{N}}[1, e^{-j\frac{2\pi}{\lambda}dsin(\phi)}, \cdots, e^{-j\frac{2\pi}{\lambda}(N-1)dsin(\phi)}]^T, \tag{17}$$

where λ and d are the wavelength of the signal and distance spacing between the antenna elements,respectively.

Since we consider the mmWave channel with partial channel knowledge of Eve, i.e., the full knowledge of angles of departure (AoD) and the distribution of the gains of mmWave paths α_l. As we can see from (16), the channel can be written more compactly as

$$\mathbf{h} = \sqrt{\frac{N_t}{L}}\mathbf{g}\mathbf{A}_t, \tag{18}$$

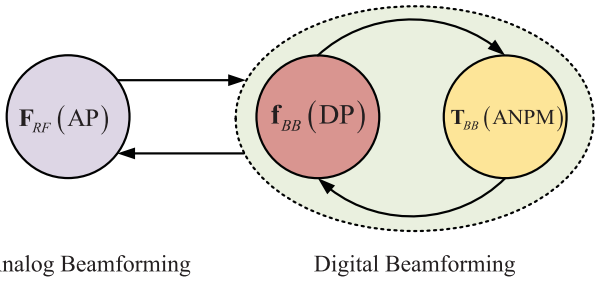


FIGURE 2. Schematic diagram of TLAIS.

with $\mathbf{g} = [g_1, g_2, \cdots, g_L]$ and $\mathbf{A}_t = [\mathbf{a}_{t,1}^H, \mathbf{a}_{t,2}^H, \cdots, \mathbf{a}_{t,L}^H]^T$. Since the each element in $\mathbf{g}$ follows the complex Gaussian distribution with zero mean and unit variance, $\mathbf{g}$ is the i.i.d. complex Gaussian vector.

Therefore, the expectation $\mathbb{E}\{\text{SINR}_e\}$ can be approximated as (19), as shown at the bottom of this page. The approximation in (b) is similarly adopted in [34]–[36].

III. DESIGN AND OPTIMIZATION OF DP AP AND ANPM

In this section, we propose a TLAIS of maximizing the ASR by taking quantization errors from both low-resolution DAC and phase shifters into account as shown in Fig. 2. For the outer loop, given the initial DP vector of confidential messages and ANPM, a steepest ascent algorithm is used to attain the updated AP vector by maximizing the approximate SR (ASR). After completing this, the optimization process turns to the inner loop. Given the AP matrix, the DP vector of confidential messages and ANPM are alternatively attained within an interior iterative loop by using GPI method with the aim to maximize the ASR. The two loops are repeated until their individual terminal conditions are satisfied.

A. DESIGN OF AP MATRIX

Observing the structure of the numerator in (19), we can simplify it to the following expression

$$\begin{aligned}&\text{Tr}[\mathbf{A}_{te,m}\mathbf{F}_{RF}\mathbf{f}_{BB}\mathbf{f}_{BB}^H\mathbf{F}_{RF}^H\mathbf{A}_{te,m}^H]\\&\quad=\text{Tr}[\mathbf{F}_{RF}^H\mathbf{A}_{te,m}^H\mathbf{A}_{te,m}\mathbf{F}_{RF}\mathbf{f}_{BB}\mathbf{f}_{BB}^H]\\&\quad\overset{(c)}{=}\text{vec}(\mathbf{F}_{RF})^H[(\mathbf{f}_{BB}\mathbf{f}_{BB}^H)^T\otimes(\mathbf{A}_{te,m}^H\mathbf{A}_{te,m})]\text{vec}(\mathbf{F}_{RF})\\&\quad=\text{vec}(\mathbf{F}_{RF})^H[\mathbf{F}_{BB}\otimes\mathbf{A}]\text{vec}(\mathbf{F}_{RF}),\end{aligned}\tag{20}$$

where $\mathbf{F}_{BB} = (\mathbf{f}_{BB}\mathbf{f}_{BB}^H)^T \in \mathbb{C}^{K\times K}$ and $\mathbf{A} = \mathbf{A}_{te,m}^H\mathbf{A}_{te,m} \in \mathbb{C}^{N_T\times N_T}$. In the above equation, (c) holds due to the fact that

$$\text{Tr}\left(\mathbf{P}^H\mathbf{Z}\mathbf{P}\mathbf{W}\right)=\text{vec}(\mathbf{P})^H\left(\mathbf{W}^T\otimes\mathbf{Z}\right)\text{vec}(\mathbf{P}).\tag{21}$$

Since the analog beamforming matrix $\mathbf{F}_{RF}$ (20) is block diagonal and most of its elements are zeros, i.e., the dimension of $\mathbf{F}_{RF}$ is $N_T\times K$. In other words, there are N_T non-zero elements and $(K-1)N_T$ zero elements for matrix $\mathbf{F}_{RF}$, so the matrix $\mathbf{F}_{RF}$ can be viewed as a sparse matrix. In order to reduce the computational complexity and dimension in the following, only non-zero elements of $\mathbf{F}_{RF}$ are required to be extracted. Because there are N_T non-zero elements in $\mathbf{F}_{RF}$, let us define

$$\mathbf{d}=\mathbb{N}[\text{vec}(\mathbf{F}_{RF})]=[\mathbf{f}_{RF1},\mathbf{f}_{RF2},\cdots,\mathbf{f}_{RFK}]^T\in\mathbb{C}^{N\times 1},\tag{22}$$

where $\mathbb{N}[]$ represents the mathematic operation of extracting all non-zero elements from $\mathbf{F}_{RF}$. To match the above operation, now, let us extract the corresponding $N\times N$ elements from $NK\times NK$ matrix $(\mathbf{F}_{BB}\otimes\mathbf{A})$ in (20). The rule of extracting method is as follows: matrix $\mathbf{F}_{BB}\otimes\mathbf{A}$ can be partitioned into $K\times K$ or K^2 block matrices with each being $N\times N$ submatrix, each block matrix can be expressed as $\mathbf{F}_{BB}(m,n)\otimes\mathbf{A}$ with $m=1,\cdots,K, n=1,\cdots,K$. Only $M\times M$ elements per block matrix are extracted with the following rule:

$$\begin{aligned}&\Gamma\left((m-1)M+1:mM,(n-1)M+1:nM\right)\\&\quad=[\mathbf{F}_{BB}(m,n)\otimes\mathbf{A}]\left((m-1)M+1:mM,(n-1)M+1:nM\right).\end{aligned}\tag{23}$$

Via the above manipulation, we have a simple form of (20) as follows

$$\text{vec}(\mathbf{F}_{RF})^H[\mathbf{F}_{BB}\otimes\mathbf{A}]\text{vec}(\mathbf{F}_{RF})=\mathbf{d}^H\Gamma\mathbf{d}.\tag{24}$$

$$\begin{aligned}\mathbb{E}\{\text{SINR}_e\}&=\mathbb{E}\left\{\frac{\beta P(1-\eta)^2\frac{N_t}{L}\mathbf{g}_e\mathbf{A}_{te}\mathbf{F}_{RF}\mathbf{f}_{BB}\mathbf{f}_{BB}^H\mathbf{F}_{RF}^H\mathbf{A}_{te}^H\mathbf{g}_e^H}{(1-\beta)P(1-\eta)^2\frac{N_t}{L}\mathbf{g}_e\mathbf{A}_{te}\mathbf{F}_{RF}\mathbf{T}_{BB}\mathbf{T}_{BB}^H\mathbf{F}_{RF}^H\mathbf{A}_{te}^H\mathbf{g}_e^H+\frac{N_t}{L}\mathbf{g}_e\mathbf{A}_{te,m}\mathbf{F}_{RF}\mathbf{R}_{\mathbf{n}_q\mathbf{n}_q}\mathbf{F}_{RF}^H\mathbf{A}_{te}^H\mathbf{g}_e^H+\sigma^2}\right\}\\&\overset{(b)}{\approx}\frac{\mathbb{E}\left\{\beta P(1-\eta)^2\frac{N_t}{L}\mathbf{g}_e\mathbf{A}_{te}\mathbf{F}_{RF}\mathbf{f}_{BB}\mathbf{f}_{BB}^H\mathbf{F}_{RF}^H\mathbf{A}_{te}^H\mathbf{g}_e^H\right\}}{\mathbb{E}\left\{(1-\beta)P(1-\eta)^2\frac{N_t}{L}\mathbf{g}_e\mathbf{A}_{te}\mathbf{F}_{RF}\mathbf{T}_{BB}\mathbf{T}_{BB}^H\mathbf{F}_{RF}^H\mathbf{A}_{te}^H\mathbf{g}_e^H+\frac{N_t}{L}\mathbf{g}_e\mathbf{A}_{te}\mathbf{F}_{RF}\mathbf{R}_{\mathbf{n}_q\mathbf{n}_q}\mathbf{F}_{RF}^H\mathbf{A}_{te}^H\mathbf{g}_e^H+\sigma^2\right\}}\\&=\frac{\beta P(1-\eta)^2\frac{N_t}{L}\mathbb{E}\left\{\text{tr}[\mathbf{g}_e^H\mathbf{g}_{e,m}\mathbf{A}_{te,m}\mathbf{F}_{RF}\mathbf{f}_{BB}\mathbf{f}_{BB}^H\mathbf{F}_{RF}^H\mathbf{A}_{te}^H]\right\}}{(1-\beta)P(1-\eta)^2\frac{N_t}{L}\mathbb{E}\left\{\text{Tr}[\mathbf{g}_e^H\mathbf{g}_e\mathbf{A}_{te}\mathbf{F}_{RF}\mathbf{T}_{BB}\mathbf{T}_{BB}^H\mathbf{F}_{RF}^H\mathbf{A}_{te}^H]+\frac{N_t}{L}\text{Tr}[\mathbf{g}_e^H\mathbf{g}_e\mathbf{A}_{te}\mathbf{F}_{RF}\mathbf{R}_{\mathbf{n}_q\mathbf{n}_q}\mathbf{F}_{RF}^H\mathbf{A}_{te}^H]+\sigma^2\right\}}\\&=\frac{\beta P(1-\eta)^2\frac{N_t}{L}\text{tr}[\mathbb{E}\left\{\mathbf{g}_e^H\mathbf{g}_{e,m}\right\}\mathbf{A}_{te,m}\mathbf{F}_{RF}\mathbf{f}_{BB}\mathbf{f}_{BB}^H\mathbf{F}_{RF}^H\mathbf{A}_{te}^H]}{(1-\beta)P(1-\eta)^2\frac{N_t}{L}\text{tr}[\mathbb{E}\left\{\mathbf{g}_e^H\mathbf{g}_e\right\}\mathbf{A}_{te}\mathbf{F}_{RF}\mathbf{T}_{BB}\mathbf{T}_{BB}^H\mathbf{F}_{RF}^H\mathbf{A}_{te}^H]+\frac{N_t}{L}\text{Tr}[\mathbb{E}\left\{\mathbf{g}_e^H\mathbf{g}_e\right\}\mathbf{A}_{te}\mathbf{F}_{RF}\mathbf{R}_{\mathbf{n}_q\mathbf{n}_q}\mathbf{F}_{RF}^H\mathbf{A}_{te}^H]+\sigma^2}\\&=\frac{\beta P(1-\eta)^2\frac{N_t}{L}\text{tr}[\mathbf{A}_{te}\mathbf{F}_{RF}\mathbf{f}_{BB}\mathbf{f}_{BB}^H\mathbf{F}_{RF}^H\mathbf{A}_{te,m}^H]}{(1-\beta)P(1-\eta)^2\frac{N_t}{L}\text{Tr}[\mathbf{A}_{te}\mathbf{F}_{RF}\mathbf{T}_{BB}\mathbf{T}_{BB}^H\mathbf{F}_{RF}^H\mathbf{A}_{te}^H]+\frac{N_t}{L}\text{Tr}[\mathbf{A}_{te}\mathbf{F}_{RF}\mathbf{R}_{\mathbf{n}_q\mathbf{n}_q}\mathbf{F}_{RF}^H\mathbf{A}_{te}^H]+\sigma^2}.\end{aligned}\tag{19}$$

In the same fashion, (19) can be rewritten in a more concise way as follows:

$$\mathbb{E}\{\text{SINR}_e\} \approx \frac{\mathbf{d}^H\Gamma_{e1}\mathbf{d}}{\mathbf{d}^H\Gamma_{e2}\mathbf{d}+\mathbf{d}^H\Gamma_{e3}\mathbf{d}+\sigma^2} = \frac{\mathbf{d}^H\Gamma_{e1}\mathbf{d}}{\mathbf{d}^H[\Gamma_{e2}+\Gamma_{e3}+\frac{\sigma^2}{K}\mathbf{I}_{Nt}]\mathbf{d}}. \tag{25}$$

Similarly, the achievable rate of Eve is also rewritten as

$$\log_2(1+\mathbb{E}[\text{SINR}_e]) \approx \log_2\left(\frac{\mathbf{d}^H\mathbf{A}_{e,1}\mathbf{d}}{\mathbf{d}^H\mathbf{A}_{e,2}\mathbf{d}}\right), \tag{26}$$

with

$$\mathbf{A}_{e,1} = \Gamma_{e1}+\Gamma_{e2}+\Gamma_{e3}+\frac{\sigma^2}{K}\mathbf{I}_{Nt}, \tag{27}$$

and

$$\mathbf{A}_{e,2} = \Gamma_{e2}+\Gamma_{e3}+\frac{\sigma^2}{K}\mathbf{I}_{Nt}. \tag{28}$$

Similarly, the SINR_b can be also expressed as (29), as shown at the bottom of this page. Therefore, we can further write the achievable rate of Bob as

$$R_b = \log_2\left(\frac{\mathbf{d}^H\mathbf{A}_{b,1}\mathbf{d}}{\mathbf{d}^H\mathbf{A}_{b,2}\mathbf{d}}\right), \tag{30}$$

with

$$\mathbf{A}_{b,1} = \Gamma_{b1}+\Gamma_{b2}+\Gamma_{b3}+\frac{\sigma^2}{K}\mathbf{I}_{Nt}, \tag{31}$$

and

$$\mathbf{A}_{b,2} = \Gamma_{b2}+\Gamma_{b3}+\frac{\sigma^2}{K}\mathbf{I}_{Nt}. \tag{32}$$

Finally, the optimization problem of maximizing $\tilde{R}_s$ (15) can be further recasted as

$$\tilde{R}_s = R_b - \log_2(1+\mathbb{E}[\text{SINR}]) \approx \log_2\left(\frac{\mathbf{d}^H\mathbf{A}_{b,1}\mathbf{d}}{\mathbf{d}^H\mathbf{A}_{b,2}\mathbf{d}}\cdot\frac{\mathbf{d}^H\mathbf{A}_{e,2}\mathbf{d}}{\mathbf{d}^H\mathbf{A}_{e,1}\mathbf{d}}\right). \tag{33}$$

Since the $\tilde{R}_s$ is the non-convex function of $\mathbf{d}$. In particular, all elements in $\mathbf{d}$ have the unit modulus, which satisfies $\mathbf{d}\in\mathcal{F}_{RF}$.

Algorithm 1 Gradient Ascent Algorithm for Analog Precoder

Input:
(1) Initialize $\mathbf{F}_{RF}^{(0)}$ according to (78) and extract non-zero elements in $\mathbf{F}_{RF}^{(0)}$ to initialize $\mathbf{d}^{(0)}$;
(2) Fixed $\mathbf{f}_{BB}$ and $\mathbf{T}_{BB}$, compute $R_s^{(0)}$ with optimal β^0 and set $t=1$, α, threshold value α_{min}, ϵ;
for $\alpha > \alpha_{min}$
1: Compute $\nabla_{\mathbf{d}}^{(t-1)}$ according to (35) and $\mathbf{d}^{(t)} = \mathbf{d}^{(t-1)} + \alpha\nabla_{\mathbf{d}}^{(t-1)}$, reform $\mathbf{F}_{RF}^{(t)} = \frac{1}{\sqrt{M}}\exp\{j\angle(\mathbf{d}^{(t)})\}$;
2: Compute $R_s^{(t)}$ using $\mathbf{T}_{BB}$, $\mathbf{f}_{BB}$ and $\mathbf{F}_{RF}^{(t)}$, search the optimal β^t by 1-D search;
3: **If** $R_s^{(t)} - R_s^{(t-1)} > \epsilon$
$\mathbf{d}^{(t)} = \mathbf{d}^{(t-1)} + \alpha\nabla_{\mathbf{d}}^{(t-1)}$;
else
$\mathbf{d}^{(t)} = \mathbf{d}^{(t-1)}$;
$\alpha = \frac{\alpha}{2}$;
4: $t = t+1$;
end for
Output: $\mathbf{F}_{RF}^{(t)}$.

Therefore, a gradient ascent (GA) method is used to compute the AP matrix. Let us define

$$f(\mathbf{d}) = \frac{\mathbf{d}^H\mathbf{A}_{b,1}\mathbf{d}}{\mathbf{d}^H\mathbf{A}_{b,2}\mathbf{d}}, \quad g(\mathbf{d}) = \frac{\mathbf{d}^H\mathbf{A}_{e,2}\mathbf{d}}{\mathbf{d}^H\mathbf{A}_{e,1}\mathbf{d}}. \tag{34}$$

The gradient of $\tilde{R}_s$ with respect to $\mathbf{d}$ in can be given by

$$\nabla_{\mathbf{d}} = \frac{(f'(\mathbf{d})g(\mathbf{d})+f(\mathbf{d})g'(\mathbf{d}))}{f(\mathbf{d})g(\mathbf{d})\ln 2}, \tag{35}$$

where

$$f'(\mathbf{d}) = \frac{\mathbf{A}_{b,1}^H\mathbf{d}(\mathbf{d}^H\mathbf{A}_{b,2}\mathbf{d}) - (\mathbf{d}^H\mathbf{A}_{b,1}\mathbf{d})\mathbf{A}_{b,2}^H\mathbf{d}}{(\mathbf{d}^H\mathbf{A}_{b,2}\mathbf{d})^2}, \tag{36a}$$

$$g'(\mathbf{d}) = \frac{\mathbf{A}_{e,2}^H\mathbf{d}(\mathbf{d}^H\mathbf{A}_{e,1}\mathbf{d}) - (\mathbf{d}^H\mathbf{A}_{e,2}\mathbf{d})\mathbf{A}_{e,1}^H\mathbf{d}}{(\mathbf{d}^H\mathbf{A}_{e,1}\mathbf{d})^2}. \tag{36b}$$

After obtaining $\nabla_{\mathbf{d}}$, we will renew the value $\mathbf{d}^{(t)}$ of $\mathbf{d}$ by $\mathbf{d}^{(t-1)} + \alpha\nabla_{\mathbf{d}}$ with α being the searching step. The detailed process of GA algorithm proposed is listed in Algorithm 1.

$$\begin{aligned}\text{SINR}_b &= \frac{\beta P(1-\eta)^2\mathbf{h}_b\mathbf{F}_{RF}\mathbf{f}_{BB}\mathbf{f}_{BB}^H\mathbf{F}_{RF}^H\mathbf{h}_b^H}{(1-\beta)P(1-\eta)^2\mathbf{h}_b\mathbf{F}_{RF}\mathbf{T}_{BB}\mathbf{T}_{BB}^H\mathbf{F}_{RF}^H\mathbf{h}_b^H + \mathbf{h}_b\mathbf{F}_{RF}\mathbf{R}_{\mathbf{n}_q\mathbf{n}_q}\mathbf{F}_{RF}^H\mathbf{h}_b^H + \sigma^2} \\ &= \frac{\beta P(1-\eta)^2\text{tr}[\mathbf{F}_{RF}^H\mathbf{h}_b^H\mathbf{h}_b\mathbf{F}_{RF}\mathbf{f}_{BB}\mathbf{f}_{BB}^H]}{(1-\beta)P(1-\eta)^2\text{tr}[\mathbf{F}_{RF}^H\mathbf{h}_b^H\mathbf{h}_b\mathbf{F}_{RF}\mathbf{T}_{BB}\mathbf{T}_{BB}^H] + \text{tr}[\mathbf{F}_{RF}^H\mathbf{h}_b^H\mathbf{h}_b\mathbf{F}_{RF}\mathbf{R}_{\mathbf{n}_q\mathbf{n}_q}] + \sigma^2} \\ &= \frac{\beta P(1-\eta)^2\text{vec}(\mathbf{F}_{RF})^H[\mathbf{F}_{BB}\otimes\mathbf{H}_b]\text{vec}(\mathbf{F}_{RF})}{(1-\beta)P(1-\eta)^2\text{vec}(\mathbf{F}_{RF})^H[(\mathbf{T}_{BB}\mathbf{T}_{BB}^H)^T\otimes\mathbf{H}_b]\text{vec}(\mathbf{F}_{RF}) + \text{vec}(\mathbf{F}_{RF})^H[\mathbf{R}_{nqnq}^T\otimes\mathbf{H}_b]\text{vec}(\mathbf{F}_{RF}) + \sigma^2} \\ &= \frac{\mathbf{d}^H\Gamma_{b1}\mathbf{d}}{\mathbf{d}^H[\Gamma_{b2}+\Gamma_{b3}+\frac{\sigma^2}{K}\mathbf{I}_N t]\mathbf{d}}.\end{aligned} \tag{29}$$

B. DESIGN OF DP VECTOR OF CONFIDENTIAL MESSAGE

In this section, we will design the DP vector $\mathbf{f}_{BB}$ assuming the other two precoders $\mathbf{F}_{RF}$ and $\mathbf{T}_{BB}$ are known in advance. To optimize $\mathbf{f}_{BB}$, the achievable rate of Bob is represented as a function of $\mathbf{f}_{BB}$ as follows

$$\begin{aligned} R_b &= \log_2\left(1+\frac{\xi_1\mathbf{f}_{BB}^H\mathbf{F}_{RF}^H\mathbf{h}_b^H\mathbf{h}_b\mathbf{F}_{RF}\mathbf{f}_{BB}}{\xi_2\mathbf{h}_b\mathbf{F}_{RF}\mathrm{diag}[\mathbf{f}_{BB}\mathbf{f}_{BB}^H]\mathbf{F}_{RF}^H\mathbf{h}_b^H+\gamma_b}\right)\\ &\overset{(d)}{=}\log_2\left(1+\frac{\xi_1\mathbf{f}_{BB}^H\mathbf{F}_{RF}^H\mathbf{h}_b^H\mathbf{h}_b\mathbf{F}_{RF}\mathbf{f}_{BB}}{\xi_2\mathbf{h}_b\mathbf{F}_{RF}[(\mathbf{f}_{BB}\mathbf{f}_{BB}^H)\odot\mathbf{I}_K]\mathbf{F}_{RF}^H\mathbf{h}_b^H+\gamma_b}\right)\\ &\overset{(e)}{=}\log_2\left(1+\frac{\xi_1\mathbf{f}_{BB}^H\tilde{\mathbf{h}}_b^H\tilde{\mathbf{h}}_b\mathbf{f}_{BB}}{\xi_2\mathbf{f}_{BB}^H[\mathbf{I}_K\odot(\tilde{\mathbf{h}}_b^H\tilde{\mathbf{h}}_b)]\mathbf{f}_{BB}+\gamma_b}\right)\\ &=\log_2\left(\frac{\mathbf{f}_{BB}^H\mathbf{Q}_b\mathbf{f}_{BB}}{\mathbf{f}_{BB}^H\mathbf{P}_b\mathbf{f}_{BB}}\right), \end{aligned} \tag{37}$$

where

$$\xi_1 = \beta P(1-\eta)^2, \tag{38}$$
$$\xi_2 = \eta(1-\eta)\beta P, \tag{39}$$
$$\gamma_b = (1-\beta)P(1-\eta)^2\|\mathbf{h}_b\mathbf{F}_{RF}\mathbf{T}_{BB}\|^2 + \eta(1-\eta)(1-\beta)P, \tag{40}$$
$$\tilde{\mathbf{h}}_b = \mathbf{h}_b\mathbf{F}_{RF}, \tag{41}$$
$$\mathbf{Q}_b = \xi_2[\mathbf{I}_K\odot(\tilde{\mathbf{h}}_b^H\tilde{\mathbf{h}}_b)]+\gamma_b\mathbf{I}_K+\xi_1\tilde{\mathbf{h}}_b^H\tilde{\mathbf{h}}_b, \tag{42}$$

and

$$\mathbf{P}_b = \xi_2[\mathbf{I}_K\odot(\tilde{\mathbf{h}}_b^H\tilde{\mathbf{h}}_b)]+\gamma_b\mathbf{I}_K. \tag{43}$$

Equality (d) in (37) is achieved because $\mathrm{diag}[\mathbf{A}]=\mathbf{A}\odot\mathbf{I}$ and (e) holds due to the fact that $\mathrm{Tr}[(\mathbf{A}\odot\mathbf{I})\mathbf{B}]=\mathrm{Tr}[\mathbf{A}(\mathbf{I}\odot\mathbf{B})]$.

Now, we can rewrite $\mathbb{E}\{\mathrm{SINR}_e\}$ as a function of $\mathbf{f}_{BB}$ from (19) as follows:

$$\begin{aligned} \underset{\mathbf{f}_{BB}}{\mathbb{E}[\mathrm{SINR}_e]} &\approx \frac{\lambda_1\mathbf{f}_{BB}^H\mathbf{F}_{RF}^H\mathbf{A}_{te}^H\mathbf{A}_{te}\mathbf{F}_{RF}\mathbf{f}_{BB}}{\lambda_2\mathrm{tr}[\mathbf{A}_{te}\mathbf{F}_{RF}\mathrm{diag}\{\mathbf{f}_{BB}\mathbf{f}_{BB}^H\}\mathbf{F}_{RF}^H\mathbf{A}_{te}^H]+\gamma_e}\\ &=\frac{\lambda_1\mathbf{f}_{BB}^H\tilde{\mathbf{A}}_{te}^H\tilde{\mathbf{A}}_{te}\mathbf{f}_{BB}}{\lambda_2\mathrm{tr}[\tilde{\mathbf{A}}_{te}(\mathbf{f}_{BB}\mathbf{f}_{BB}^H)\odot\mathbf{I}_K\tilde{\mathbf{A}}_{te}^H]+\gamma_e}\\ &=\frac{\lambda_1\mathbf{f}_{BB}^H\tilde{\mathbf{A}}_{te}^H\tilde{\mathbf{A}}_{te}\mathbf{f}_{BB}}{\lambda_2\mathrm{tr}[(\mathbf{f}_{BB}\mathbf{f}_{BB}^H)\{\mathbf{I}_K\odot(\tilde{\mathbf{A}}_{te}^H\tilde{\mathbf{A}}_{te})\}]+\gamma_e}\\ &=\frac{\mathbf{f}_{BB}^H(\lambda_1\tilde{\mathbf{A}}_{te}^H\tilde{\mathbf{A}}_{te})\mathbf{f}_{BB}}{\mathbf{f}_{BB}^H(\lambda_2[\mathbf{I}_K\odot\tilde{\mathbf{A}}_{te}^H\tilde{\mathbf{A}}_{te}]+\gamma_e\mathbf{I}_K)\mathbf{f}_{BB}}, \end{aligned} \tag{44}$$

where

$$\lambda_1 = \beta P(1-\eta)^2N_t/L, \tag{45}$$
$$\lambda_2 = \eta(1-\eta)\beta PN_t/L, \tag{46}$$
$$\gamma_e = (1-\beta)P(1-\eta)^2\frac{N_t}{L}\mathrm{tr}[\mathbf{A}_{te}\mathbf{F}_{RF}\mathbf{T}_{BB}\mathbf{T}_{BB}^H\mathbf{F}_{RF}^H\mathbf{A}_{te}^H] + \frac{N_t}{L}\mathrm{tr}[\mathbf{A}_{te}\mathbf{F}_{RF}\eta(1-\eta)\mathrm{diag}\{(1-\beta)P\mathbf{T}_{BB}\mathbf{T}_{BB}^H\}\times\mathbf{F}_{RF}^H\mathbf{A}_{te}^H]+\sigma^2, \tag{47}$$

and $\tilde{\mathbf{A}}_{te}=\mathbf{A}_{te}\mathbf{F}_{RF}$. Using the above expression of $\mathbb{E}\{\mathrm{SINR}_e\}$, we directly have the average achievable rate at Eve as follows:

$$\log_2(1+\mathbb{E}[\mathrm{SINR}_e])\approx\log_2\left(\frac{\mathbf{f}_{BB}^H\mathbf{Q}_e\mathbf{f}_{BB}}{\mathbf{f}_{BB}^H\mathbf{P}_e\mathbf{f}_{BB}}\right), \tag{48}$$

where

$$\mathbf{Q}_e = \lambda_2[\mathbf{I}_K\odot\tilde{\mathbf{A}}_{te}^H\tilde{\mathbf{A}}_{te}]+\gamma_e\mathbf{I}_K+\lambda_1\tilde{\mathbf{A}}_{te}^H\tilde{\mathbf{A}}_{te}, \tag{49}$$

and

$$\mathbf{P}_e = \lambda_2[\mathbf{I}_K\odot\tilde{\mathbf{A}}_{te}^H\tilde{\mathbf{A}}_{te}]+\gamma_e\mathbf{I}_K. \tag{50}$$

Combining (37) and (48) yields the expression of $\tilde{R}_s$ as

$$\begin{aligned} \tilde{R}_s &= R_b-\log_2(1+\mathbb{E}[\mathrm{SINR}_e])\\ &\approx\log_2\left(\frac{\mathbf{f}_{BB}^H\mathbf{Q}_b\mathbf{f}_{BB}}{\mathbf{f}_{BB}^H\mathbf{P}_b\mathbf{f}_{BB}}\cdot\frac{\mathbf{f}_{BB}^H\mathbf{P}_e\mathbf{f}_{BB}}{\mathbf{f}_{BB}^H\mathbf{Q}_e\mathbf{f}_{BB}}\right). \end{aligned} \tag{51}$$

Observing the form of (51) is the product of fractional quadratic functions, one efficient way to address this problem is GPI algorithm [37]. Therefore, we can obtain the solution to $\mathbf{f}_{BB}$ by resorting to GPI algorithm.

C. DESIGN OF ANPM

In this subsection, we will optimize the digital ANPM by fixing $\mathbf{F}_{RF}$ and $\mathbf{f}_{BB}$. Under this condition, we can rewrite the R_b as a function of $\mathbf{T}_{BB}$ as follows:

$$\begin{aligned} &R_b(\mathbf{T}_{BB})\\ &=\log_2\left(1+\frac{\kappa_b}{\alpha_1\|\tilde{\mathbf{h}}_b\mathbf{T}_{BB}\|^2+\alpha_2\tilde{\mathbf{h}}_b\mathrm{diag}\{\mathbf{T}_{BB}\mathbf{T}_{BB}^H\}\tilde{\mathbf{h}}_b^H+\omega_b}\right)\\ &=\log_2\left(1+\frac{\kappa_b}{\alpha_1\|\tilde{\mathbf{h}}_b\mathbf{T}_{BB}\|^2+\alpha_2\mathrm{tr}\{\mathbf{T}_{BB}^H[\mathbf{I}\odot(\tilde{\mathbf{h}}_b^H\tilde{\mathbf{h}}_b)]\mathbf{T}_{BB}\}+\omega_b}\right)\\ &=\log_2\left(1+\frac{\kappa_b}{\alpha_1\|\tilde{\mathbf{h}}_b\mathbf{T}_{BB}\|^2+\alpha_2\mathrm{tr}\{\mathbf{T}_{BB}^H\mathbf{C}_b\mathbf{T}_{BB}\}+\omega_b}\right)\\ &=\log_2\left(\frac{\mathrm{Tr}\{\mathbf{T}_{BB}^H\mathbf{E}_b\mathbf{T}_{BB}\}}{\mathrm{Tr}\{\mathbf{T}_{BB}^H\mathbf{F}_b\mathbf{T}_{BB}\}}\right)\\ &=\log_2\left(\frac{\mathbf{w}^H(\mathbf{I}\otimes\mathbf{E}_b)\mathbf{w}}{\mathbf{w}^H(\mathbf{I}\otimes\mathbf{F}_b)\mathbf{w}}\right), \end{aligned} \tag{52}$$

where

$$\mathbf{w} = \mathrm{vec}(\mathbf{T}_{BB}), \tag{53}$$
$$\alpha_1 = (1-\beta)P(1-\eta)^2, \tag{54}$$
$$\alpha_2 = \eta(1-\eta)(1-\beta)P, \tag{55}$$
$$\kappa_b = \beta P(1-\eta)^2\mathbf{h}_b\mathbf{F}_{RF}\mathbf{f}_{BB}\mathbf{f}_{BB}^H\mathbf{F}_{RF}^H\mathbf{h}_b^H, \tag{56}$$
$$\omega_b = \eta(1-\eta)\beta P\tilde{\mathbf{h}}_b\mathrm{diag}[\mathbf{f}_{BB}\mathbf{f}_{BB}^H]\tilde{\mathbf{h}}_b^H+\sigma^2, \tag{57}$$
$$\mathbf{C}_b = \mathbf{I}\odot(\tilde{\mathbf{h}}_b^H\tilde{\mathbf{h}}_b), \tag{58}$$
$$\mathbf{E}_b = \alpha_1\tilde{\mathbf{h}}_b^H\tilde{\mathbf{h}}_b+\alpha_2\mathbf{C}_b+\omega_b\mathbf{I}_K+\kappa_b\mathbf{I}_K, \tag{59}$$

and

$$\mathbf{F}_b = \alpha_1\tilde{\mathbf{h}}_b^H\tilde{\mathbf{h}}_b+\alpha_2\mathbf{C}_b+\omega_b\mathbf{I}_K. \tag{60}$$

To obtain the $\tilde{R}_s$, $\mathbb{E}\{\text{SINR}_e\}$ is rewritten as a function of $\mathbf{T}_{BB}$ as follows:

$$\begin{aligned}\mathbb{E}[\text{SINR}_e]&\\ \mathbf{T}_{BB}&\\ &\approx \frac{\kappa_e}{\alpha_3 \text{tr}[\mathbf{T}_{BB}^H \tilde{\mathbf{A}}_{te}^H \tilde{\mathbf{A}}_{te} \mathbf{T}_{BB}] + \alpha_4 \text{tr}[\mathbf{T}_{BB}^H [\mathbf{I}_K \odot (\tilde{\mathbf{A}}_{te}^H \tilde{\mathbf{A}}_{te})] \mathbf{T}_{BB}] + \omega_e}\\ &= \frac{\kappa_e}{\text{tr}[\mathbf{T}_{BB}^H \{\alpha_3 \tilde{\mathbf{A}}_{te}^H \tilde{\mathbf{A}}_{te} + \alpha_4 \mathbf{C}_e + \omega_e \mathbf{I}_K\} \mathbf{T}_{BB}]}\\ &= \frac{\kappa_e}{\text{tr}[\mathbf{T}_{BB}^H \mathbf{F}_e \mathbf{T}_{BB}]}, \end{aligned} \tag{61}$$

where

$$\alpha_3 = (1-\beta)P(1-\eta)^2 \frac{N_t}{L}, \quad \alpha_4 = \eta(1-\eta)(1-\beta)P\frac{N_t}{L}, \tag{62}$$

$$\kappa_e = \beta P (1-\eta)^2 \frac{N_t}{L} \text{tr}[\tilde{\mathbf{A}}_{te} \mathbf{f}_{BB} \mathbf{f}_{BB}^H \tilde{\mathbf{A}}_{te}], \tag{63}$$

and

$$\omega_e = \eta(1-\eta)\beta P \frac{N_t}{L} \text{tr}[\tilde{\mathbf{A}}_{te} \text{diag}(\mathbf{f}_{BB}\mathbf{f}_{BB}^H) \tilde{\mathbf{A}}_{te}^H] + \sigma^2. \tag{64}$$

Therefore, we can obtain

$$\begin{aligned}\log_2(1+\mathbb{E}[\text{SINR}]) &\approx \log_2\left(\frac{\text{tr}[\mathbf{T}_{BB}^H(\mathbf{F}_e + \kappa_e \mathbf{I}_K)\mathbf{T}_{BB}]}{\text{tr}[\mathbf{T}_{BB}^H \mathbf{F}_e \mathbf{T}_{BB}]}\right)\\ &= \log_2\left(\frac{\text{tr}[\mathbf{T}_{BB}^H \mathbf{E}_e \mathbf{T}_{BB}]}{\text{tr}[\mathbf{T}_{BB}^H \mathbf{F}_e \mathbf{T}_{BB}]}\right)\\ &= \log_2\left(\frac{\mathbf{w}^H (\mathbf{I} \otimes \mathbf{E}_e)\mathbf{w}}{\mathbf{w}^H (\mathbf{I} \otimes \mathbf{F}_e)\mathbf{w}}\right), \end{aligned} \tag{65}$$

with

$$\mathbf{C}_e = \mathbf{I}_K \odot (\tilde{\mathbf{A}}_{te}^H \tilde{\mathbf{A}}_{te}), \tag{66}$$

$$\mathbf{E}_e = \alpha_3 \tilde{\mathbf{A}}_{te}^H \tilde{\mathbf{A}}_{te} + \alpha_4 \mathbf{I}_K \odot (\tilde{\mathbf{A}}_{te}^H \tilde{\mathbf{A}}_{te}) + \omega_e \mathbf{I}_K + \kappa_e \mathbf{I}_K, \tag{67}$$

and

$$\mathbf{F}_e = \alpha_3 \tilde{\mathbf{A}}_{te}^H \tilde{\mathbf{A}}_{te} + \alpha_4 \mathbf{I}_K \odot (\tilde{\mathbf{A}}_{te}^H \tilde{\mathbf{A}}_{te}) + \omega_e \mathbf{I}_K. \tag{68}$$

Making use of (52) and (65), the $\tilde{R}_s$ can be reformulated as a function of $\mathbf{w}$ with $\mathbf{F}_{RF}$ and $\mathbf{f}_{BB}$ fixed as follows:

$$\begin{aligned}\tilde{R}_s &= R_b - \log_2(1+\mathbb{E}[\text{SINR}])\\ &\approx \log_2\left(\frac{\mathbf{w}^H (\mathbf{I} \otimes \mathbf{E}_b)\mathbf{w}}{\mathbf{w}^H (\mathbf{I} \otimes \mathbf{F}_b)\mathbf{w}} \cdot \frac{\mathbf{w}^H (\mathbf{I} \otimes \mathbf{F}_e)\mathbf{w}}{\mathbf{w}^H (\mathbf{I} \otimes \mathbf{E}_e)\mathbf{w}}\right), \end{aligned} \tag{69}$$

where $\mathbf{w} = \text{vec}(\mathbf{T}_{BB})$.

Similarly, $\mathbf{w}$ can be solved with GPI, and the corresponding ANPM can be obtained by reverse operation of vector-to-matrix operator.

Until now, we have completed the design of secure hybrid precoders. Our iterative idea can be described as follows: for given AP matrix $\mathbf{F}_{RF}$, the near-optimal DP vector $\mathbf{f}_{BB}$ and ANPM $\mathbf{T}_{BB}$ are iteratively computed; Then given the DP vector $\mathbf{f}_{BB}$ and ANPM $\mathbf{T}_{BB}$, we renew the AP matrix $\mathbf{F}_{RF}$ by GA method. The alternative iterations among $\mathbf{F}_{RF}$, $\mathbf{f}_{BB}$, and $\mathbf{T}_{BB}$ is repeated until a stop criterion satisfies. Here, stop criterion is chosen as follows: $R_s^{(p+1)} - R_s^{(p)} \le \epsilon$ with p being the iteration index. The proposed method is summarized in Algorithm 3.

Algorithm 2 Iterative Method for Digital Precoders

Input:
(1) Fixed $\mathbf{F}_{RF}$, initialize $\mathbf{f}_{BB}^{(0)}$ and $\mathbf{T}_{BB}^{(0)}$ and set $k=0$, threshold value ϵ;
(2) Compute $R_s^{(0)}$ using $\mathbf{f}_{BB}^{(0)}$, $\mathbf{T}_{BB}^{(0)}$ and $\mathbf{F}_{RF}$, search the optimal β^0 by 1-D search;
Repeat
1: Solve $\mathbf{T}_{BB}^{(k+1)}$ using $\mathbf{f}_{BB}^{(k)}$ and $\mathbf{F}_{RF}$ through (69);
2: Solve $\mathbf{f}_{BB}^{(k+1)}$ using $\mathbf{T}_{BB}^{(k)}$ and $\mathbf{F}_{RF}$ through (51);
3: Compute $R_s^{(k+1)}$ using $\mathbf{T}_{BB}^{(k+1)}$, $\mathbf{f}_{BB}^{(k+1)}$ and $\mathbf{F}_{RF}$, search the optimal β^{k+1} by 1-D search;
4: $k = k+1$;
Until $R_s^{(k+1)} - R_s^{(k)} \le \epsilon$;
Output: $\mathbf{f}_{BB}^{(k)}$ and $\mathbf{T}_{BB}^{(k)}$;

Algorithm 3 Proposed TLAIS Algorithm

Input:
(1) Initialize $\mathbf{F}_{RF}^{(0)}$, $\mathbf{f}_{BB}^{(0)}$ and $\mathbf{T}_{BB}^{(0)}$ according to (78), (85) and (86), respectively and set $p=0$, threshold value ϵ;
(2) For $\mathbf{F}_{RF}^{(0)}$, update $\mathbf{f}_{BB}^{(0)}$ and $\mathbf{T}_{BB}^{(0)}$ with Algorithm 2, compute $R_s^{(0)}$ with optimal β^0 and $\mathbf{F}_{RF}^{(0)}$, $\mathbf{f}_{BB}^{(0)}$, $\mathbf{T}_{BB}^{(0)}$;
Repeat
1: For $\mathbf{f}_{BB}^{(p)}$ and $\mathbf{T}_{BB}^{(p)}$, using Algorithm 1 to obtain $\mathbf{F}_{RF}^{(p+1)}$;
2: For $\mathbf{F}_{RF}^{(p+1)}$, using Algorithm 2 to obtain updated $\mathbf{f}_{BB}^{(p+1)}$ and $\mathbf{T}_{BB}^{(p+1)}$;
3: Compute $R_s^{(p+1)}$ using $\mathbf{T}_{BB}^{(p+1)}$, $\mathbf{f}_{BB}^{(p+1)}$ and $\mathbf{F}_{RF}^{(p+1)}$, search the optimal β^{p+1} by 1-D search;
4: $p = p+1$;
Until $R_s^{(p+1)} - R_s^{(p)} \le \epsilon$;
Output: $R_s^{(p)}$.

D. INITIALIZATION OF TLAIS

In the following, we will show how to set the initial values of the proposed TLAIS. In other words, we need to make an initialization for three matrices AP, DP, and ANPM, respectively. Since the optimization problem in (15) is non-convex, it is intractable. To address it, we first consider the scenario without AN, i.e. $\beta = 1$, as in [29]. In this case, the achievable rate of Bob can be expressed as

$$\begin{aligned}R_b &= \log_2\left(1 + \frac{P(1-\eta)^2 \|\mathbf{h}_b \mathbf{F}_{RF} \mathbf{f}_{BB}\|^2}{\eta(1-\eta)P \mathbf{h}_b \mathbf{F}_{RF} \text{diag}\{\mathbf{f}_{BB}\mathbf{f}_{BB}^H\} \mathbf{F}_{RF}^H \mathbf{h}_b^H + \sigma_b^2}\right)\\ &= \log_2\left(1 + \frac{\mu_1 |\mathbf{h}_b \mathbf{F}_{RF} \mathbf{f}_{BB}\|^2}{\mu_2 \mathbf{h}_b \mathbf{F}_{RF} [(\mathbf{f}_{BB}\mathbf{f}_{BB}^H) \odot \mathbf{I}_K] \mathbf{F}_{RF}^H \mathbf{h}_b^H + \sigma_b^2}\right)\\ &= \log_2\left(1 + \frac{\mu_1 \mathbf{f}_{BB}^H \mathbf{F}_{RF}^H \mathbf{h}_b^H \mathbf{h}_b \mathbf{F}_{RF} \mathbf{f}_{BB}}{\mu_2 \mathbf{f}_{BB}^H [\mathbf{I}_K \odot (\mathbf{F}_{RF}^H \mathbf{h}_b^H \mathbf{h}_b \mathbf{F}_{RF})] \mathbf{f}_{BB} + \sigma_b^2}\right), \end{aligned} \tag{70}$$

where $\mu_1 = P(1-\eta)^2$ and $\mu_2 = \eta(1-\eta)P$.

If a fully digital architecture is adopted, i.e., $\mathbf{F}_{RF} = \mathbf{I}_{Nt}$, then (70) can be simplified as

$$R_b = \log_2\left(1 + \frac{\mu_1 \mathbf{f}_{FD}^H \mathbf{h}_b^H \mathbf{h}_b \mathbf{f}_{FD}}{\mu_2 \mathbf{f}_{FD}^H [\mathbf{I}_{N_t} \odot (\mathbf{h}_b^H \mathbf{h}_b)] \mathbf{f}_{FD} + \sigma_b^2}\right) = \log_2\left(\frac{\mathbf{f}_{FD}^H \mathbf{X}_b \mathbf{f}_{FD}}{\mathbf{f}_{FD}^H \mathbf{Y}_b \mathbf{f}_{FD}}\right), \tag{71}$$

where

$$\mathbf{X}_b = \mu_2[\mathbf{I}_{Nt} \odot (\mathbf{h}_b^H \mathbf{h}_b)] + \sigma^2 \mathbf{I}_{N_t} + \mu_1 \mathbf{h}_b^H \mathbf{h}_b, \tag{72}$$

and

$$\mathbf{Y}_b = \mu_2[\mathbf{I}_{Nt} \odot (\mathbf{h}_b^H \mathbf{h}_b)] + \sigma^2 \mathbf{I}_{N_t}. \tag{73}$$

Similarly, the $\log_2(1 + \mathbb{E}[\mathrm{SINR_e}])$ at Eve can be rewritten as

$$\begin{aligned}&\log_2(1 + \mathbb{E}[\mathrm{SINR}_e]) \\ &\overset{(f)}{\approx} \log_2\left(1 + \frac{\mu_3 \mathbf{f}_{BB}^H \mathbf{F}_{RF}^H \mathbf{A}_{te}^H \mathbf{A}_{te} \mathbf{F}_{RF} \mathbf{f}_{BB}}{\mu_4 \mathbf{f}_{BB}^H [\mathbf{I}_K \odot (\mathbf{F}_{RF}^H \mathbf{A}_{te}^H \mathbf{A}_{te} \mathbf{F}_{RF})] \mathbf{f}_{BB} + \sigma^2}\right) \\ &\overset{(g)}{=} \log_2\left(1 + \frac{\mu_3 \mathbf{f}_{FD}^H \mathbf{A}_{te}^H \mathbf{A}_{te} \mathbf{f}_{FD}}{\mu_4 \mathbf{F}_{FD}^H [\mathbf{I}_K \odot (\mathbf{A}_{te}^H \mathbf{A}_{te})] \mathbf{f}_{FD} + \sigma^2}\right) \\ &= \log_2\left(\frac{\mathbf{f}_{FD} \mathbf{X}_e \mathbf{f}_{FD}}{\mathbf{f}_{FD}^H \mathbf{Y}_e \mathbf{f}_{FD}}\right),\end{aligned} \tag{74}$$

where $\mu_3 = P(1-\eta)^2 \frac{N_t}{L}$, $\mu_4 = \eta(1-\eta)P\frac{N_t}{L}$. The approximation (f) in (74) holds because $\beta = 1$ and (g) is true due to $\mathbf{F}_{RF} = \mathbf{I}_{N_t}$. What's more,

$$\mathbf{X}_e = \mu_4[\mathbf{I}_{N_t} \odot (\mathbf{A}_{te}^H \mathbf{A}_{te})] + \sigma^2 \mathbf{I}_{N_t} + \mu_3 \mathbf{A}_{te}^H \mathbf{A}_{te}, \tag{75}$$

and

$$\mathbf{Y}_e = \mu_4[\mathbf{I}_{N_t} \odot (\mathbf{A}_{te}^H \mathbf{A}_{te})] + \sigma^2 \mathbf{I}_{N_t}. \tag{76}$$

Therefore, the approximate expression $\tilde{R}_s$ of SR can be expressed as

$$\tilde{R}_s = R_b - \log_2(1 + \mathbb{E}[\mathrm{SINR}]) \approx \log_2\left(\frac{\mathbf{f}_{FD}^H \mathbf{X}_b \mathbf{f}_{FD}}{\mathbf{f}_{FD}^H \mathbf{Y}_b \mathbf{f}_{FD}} \cdot \frac{\mathbf{f}_{FD}^H \mathbf{Y}_b \mathbf{f}_{FD}}{\mathbf{f}_{FD}^H \mathbf{X}_b \mathbf{f}_{FD}}\right). \tag{77}$$

Therefore, $\mathbf{f}_{FD}$ can be solved with GPI algorithm. After obtaining $\mathbf{f}_{FD}$, we can obtain analog beamforming $\mathbf{F}_{RF}$ as

$$\mathbf{f}_{RF,l} = \frac{1}{\sqrt{T}} \exp\left\{jQ_{bps}[\mathbf{f}_{FD,(l-1)M:lM}]\right\}, \quad l = 1, 2, \cdots, K, \tag{78}$$

where Q_{bps} denotes the operation of quantizing each phase of the $\mathbf{f}_{FD}$ to its nearest value in discrete phase set $\mathcal{F}_{RF}$.

Observing (70), we can assign the initial value of $\mathbf{f}_{BB}$ by rewriting (70) as

$$R_b = \log_2\left(1 + \frac{\mathbf{f}_{BB}^H [\mu_1 \mathbf{F}_{RF}^H \mathbf{h}_b^H \mathbf{h}_b \mathbf{F}_{RF}] \mathbf{f}_{BB}}{\mathbf{f}_{BB}^H \{\mu_2[\mathbf{I}_K \odot (\mathbf{F}_{RF}^H \mathbf{h}_b^H \mathbf{h}_b \mathbf{F}_{RF})] + \sigma^2\} \mathbf{f}_{BB}}\right) = \log_2\left(\frac{\mathbf{f}_{BB}^H \tilde{\mathbf{X}}_b \mathbf{f}_{BB}}{\mathbf{f}_{BB}^H \tilde{\mathbf{Y}}_b \mathbf{f}_{BB}}\right). \tag{79}$$

Similarly, we have the approximate rate at Eve as

$$\log_2(1 + \mathbb{E}[\mathrm{SINR}_e]) \approx \log_2\left(1 + \frac{\mu_3 \mathbf{f}_{BB}^H \mathbf{F}_{RF}^H \mathbf{A}_{te}^H \mathbf{A}_{te} \mathbf{F}_{RF} \mathbf{f}_{BB}}{\mu_4 \mathbf{f}_{BB}^H [\mathbf{I}_K \odot (\mathbf{F}_{RF}^H \mathbf{A}_{te}^H \mathbf{A}_{te} \mathbf{F}_{RF})] \mathbf{f}_{BB} + \sigma^2}\right) = \log_2\left(\frac{\mathbf{f}_{BB}^H \tilde{\mathbf{X}}_e \mathbf{f}_{BB}^H}{\mathbf{f}_{BB}^H \tilde{\mathbf{Y}}_e \mathbf{f}_{BB}}\right), \tag{80}$$

with

$$\tilde{\mathbf{X}}_b = \mu_2[\mathbf{I}_K \odot (\mathbf{F}_{RF}^H \mathbf{h}_b^H \mathbf{h}_b \mathbf{F}_{RF})] + \sigma^2 \mathbf{I}_K + \mu_1 \mathbf{F}_{RF}^H \mathbf{h}_b^H \mathbf{h}_b \mathbf{F}_{RF}, \tag{81}$$

$$\tilde{\mathbf{Y}}_b = \mu_2[\mathbf{I}_K \odot (\mathbf{F}_{RF}^H \mathbf{h}_b^H \mathbf{h}_b \mathbf{F}_{RF})] + \sigma^2 \mathbf{I}_K, \tag{82}$$

$$\tilde{\mathbf{X}}_e = \mu_4[\mathbf{I}_K \odot (\mathbf{F}_{RF}^H \mathbf{A}_{te}^H \mathbf{A}_{te} \mathbf{F}_{RF})] + \sigma^2 \mathbf{I}_K + \mu_3 \mathbf{F}_{RF}^H \mathbf{A}_{te}^H \mathbf{A}_{te} \mathbf{F}_{RF}, \tag{83}$$

and

$$\tilde{\mathbf{X}}_e = \mu_4[\mathbf{I}_K \odot (\mathbf{F}_{RF}^H \mathbf{A}_{te}^H \mathbf{A}_{te} \mathbf{F}_{RF})] + \sigma^2 \mathbf{I}_K. \tag{84}$$

Then the $\tilde{R}_s$ can be rewritten as

$$\tilde{R}_s = R_b - \log_2(1 + \mathbb{E}[\mathrm{SINR}]) \approx \log_2\left(\frac{\mathbf{f}_{BB}^H \tilde{\mathbf{X}}_b \mathbf{f}_{BB}}{\mathbf{f}_{BB}^H \tilde{\mathbf{Y}}_b \mathbf{f}_{BB}} \cdot \frac{\mathbf{f}_{BB}^H \tilde{\mathbf{Y}}_e \mathbf{f}_{BB}}{\mathbf{f}_{BB}^H \tilde{\mathbf{X}}_e \mathbf{f}_{BB}}\right). \tag{85}$$

Therefore, the initial value of $\mathbf{f}_{BB}$ can be obtained by maximizing (85) via GPI scheme.

To set a good initial value of ANPM, it's reasonable to assign the initial value of $\mathbf{T}_{BB}$ by the concept of null-space projection, which forces $\mathbf{T}_{BB}$ in the null space of channel of desired user, thus eliminating the interference of AN to the intended user. The expression of the initial value of $\mathbf{T}_{BB}$ can be given as follows:

$$\mathbf{T}_{BB} = \frac{\mathbf{I}_K - \tilde{\mathbf{h}}_b^H \left(\tilde{\mathbf{h}}_b \tilde{\mathbf{h}}_b^H\right)^{-1} \tilde{\mathbf{h}}_b}{\|\mathbf{I}_K - \tilde{\mathbf{h}}_b^H \left(\tilde{\mathbf{h}}_b \tilde{\mathbf{h}}_b^H\right)^{-1} \tilde{\mathbf{h}}_b\|_F}, \tag{86}$$

where $\tilde{\mathbf{h}}_b = \mathbf{h}_b \mathbf{F}_{RF}$ stands for the equivalent channel of intended user.

Since the beamforming matrices $\mathbf{f}_{BB}$ and $\mathbf{T}_{BB}$ are closely related to the power allocation factor β in accordance with (51) and (69), the reasonable way to find the optimal PA is achieved by the conventional 1-D linear search.

IV. NUMERICAL RESULTS

In this section, we will present the numerical simulations to evaluate the SR performance of our proposed method and compare it with existing methods like AN-aided method in [29]. Here it is particularly pointed out that it's infeasible to directly apply the algorithm without taking low-resolution DACs into account in [29] to our paper. Then, the idea in [29] is modified to suit our model and at the same time choose it as our initial value of our proposed precoders. The second algorithm is the maximum ratio transmission (MRT) algorithm. To guarantee the fairness of performance comparison, the

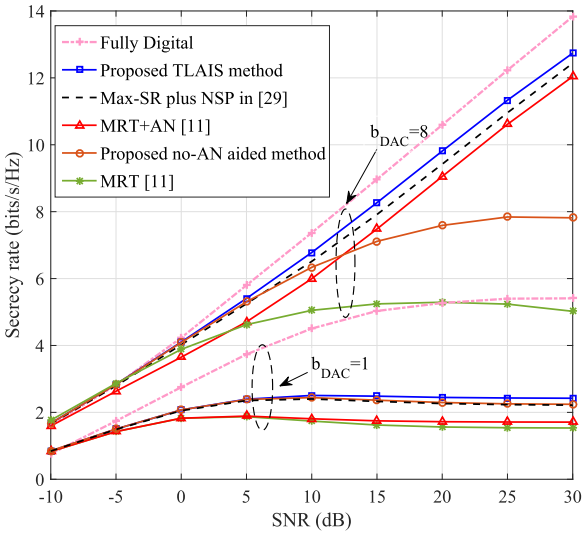


FIGURE 3. SR versus SNR with $b_{DAC} \in \{1, 8\}$.

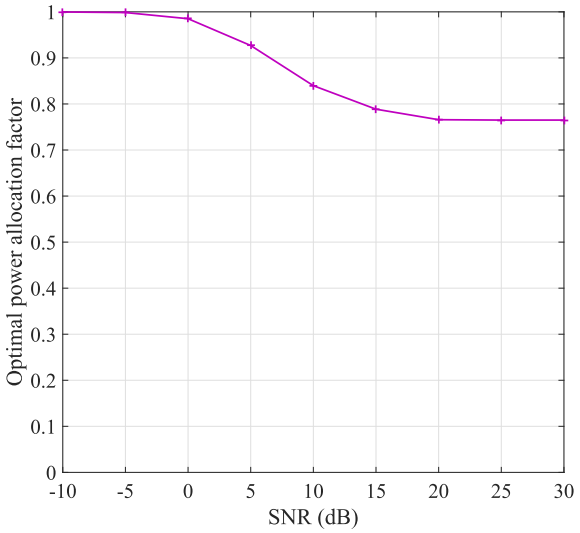


FIGURE 4. Optimal PA factor versus SNR.

same AN power is introduced to MRT. What's more, to reveal the importance of AN, we simultaneously make a comparison of our proposed method without AN-aided together with MRT algorithm. The parameters are chosen as follows. In our simulation experiment, the transmitter at Alice is equipped with 32 transmit antennas with 4 RF chains, and mmWave channel model is also adopted in our paper with $L = 12$ and azimuth angles being a uniform distribution over $[0, 2\pi]$.

Fig. 3 demonstrates the SR performance versus SNR with number b_{DAC} of quantization bits of ADC being 1 and 8. As can be seen from this figure, the proposed TLAIS method performs better than AN-aided method in [29] and MRT plus AN method in terms of SR performance. When $b_{DAC} = 8$, we find an important fact that AN has an important impact on SR performance. Taking the proposed TLAIS as an example, the proposed TLAIS performs much better than the corresponding one without AN. More importantly, with the help of AN, the SR performance of all three methods including our proposed method gradually grows with increasing SNR. Conversely, without the aid of AN, all methods exist the effect of SR performance ceil as SNR increases. The main reason that our proposed method is better than MRT can be explained as follows: unlike our method, the major goal of MRT algorithm is to maximize the receive SNR at Bob and don't concern the receive SNR at Eve. However, our method is to maximize SR. Thus, our proposed method can achieve a better SR performance. When $b_{DAC} = 1$, the SR performance of all methods will tend to be saturated or even declined because the quantization distortion is very significant and it won't drop with the increasing of SNR when the digital-to-analog quantization bit is so small, and in this case, the linear approximation of DAC is imprecise and result in model mismatch.

Fig. 4 shows the optimal PA factor β versus SNR with $b_{DAC} = 8$, where β is to maximize SR by adjusting PA between confidential message and AN. From Fig. 4, it is seen

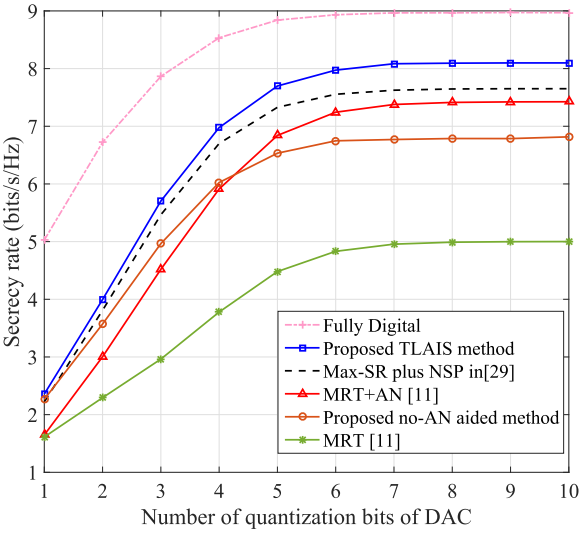


FIGURE 5. SR versus number of quantization bits of DAC with SNR=15dB.

that, in the low SNR region, the optimal PA factor always remains 1, which implies that in this scenario all power should be allocated to transmit confidential message to guarantee the communication quality, and AN has a trivial impact on SR performance. Furthermore, as SNR increases, the optimal PA factor declines gradually. This reveals a fact that more power should be allocated to AN to interfere with eavesdropper, and improves the security of communication. Fig. 4 verifies the effectiveness of AN-aided method especially in the medium and high SNR regions.

Fig. 5 illustrates the SR versus number b_{DAC} of quantization bits of DAC with SNR = 15dB. For the small value of b_{DAC}, the SR is small and its performance loss is significant because the distortion of low-resolution DACs is extremely severe. As b_{DAC} increases, the SR performance improves dramatically, especially when b_{DAC} ranges from 1 to 6 bits.

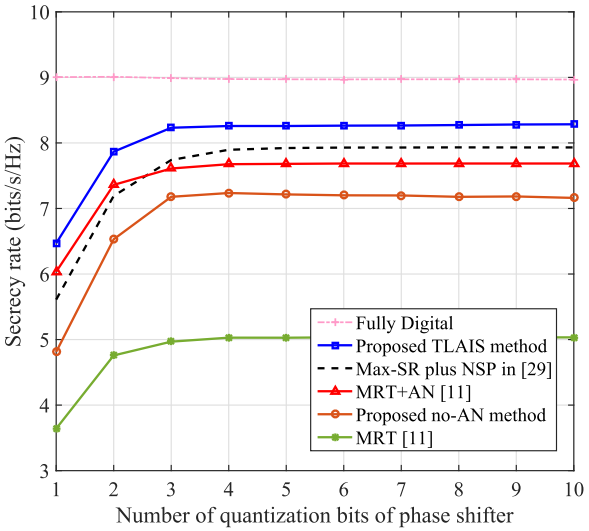


FIGURE 6. SR versus number of quantization bits of phase shifter with SNR=15dB and $b_{DAC} = 8$.

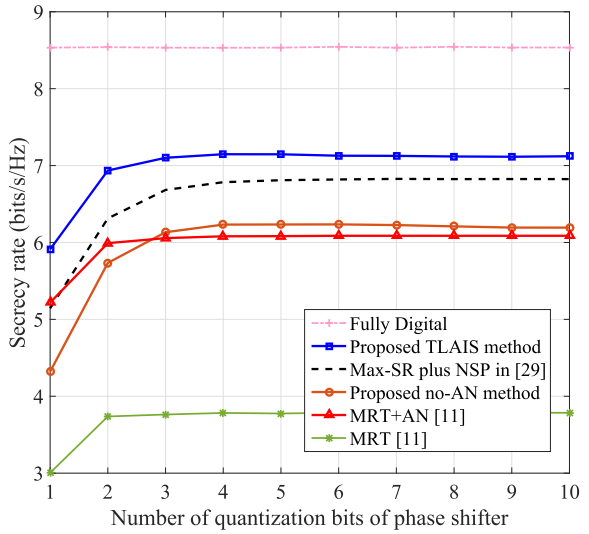


FIGURE 7. SR versus number of quantization bits of phase shifter with SNR=15dB and $b_{DAC} = 4$.

It's noted that the SR performance of five methods all tend to stable when b_{DAC} reach up to 6. In such a situation, the performance loss is negligible. What's more, given a fixed value of b_{DAC}, our proposed method achieves a higher SR performance than other algorithms.

Fig. 6 plots the SR versus number b_{PS} of quantization bits of phase shifters with SNR = 15dB and $b_{DAC} = 8$. The proposed TLAIS performs much better than MRT and MRT with AN in terms of SR for all cases. Additionally, we also find the fact that their SR tends to a flat rate ceil with increase in number of quantization bits of phase shifts. From Fig. 6, it follows that the SR loss due to the effect of quantization error from phase shifters becomes trivial when $b_{PS} \geq 4$.

Fig. 7 plots the SR versus number b_{PS} of quantization bits of phase shifters with SNR = 15dB and $b_{DAC} = 4$. It is noted that Fig. 7 shows the similar performance trend to Fig. 6. In other words, it follows that our proposed no-AN aided method is still better than MRT plus AN method in medium and high values of b_{PS}. This result is consistent with that Fig. 5 at $b_{DAC} = 4$.

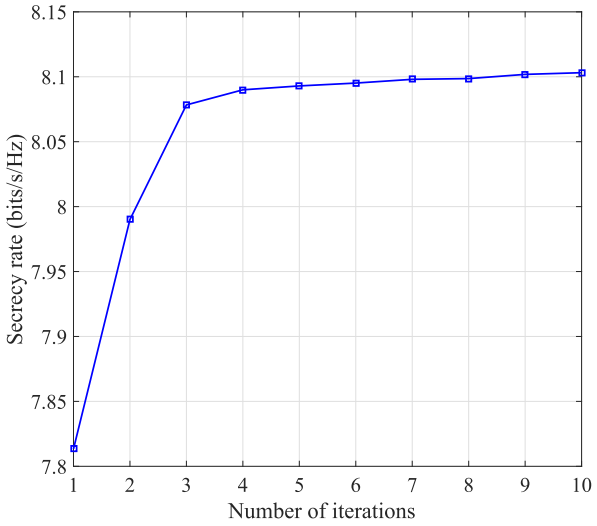


FIGURE 8. Convergence of proposed TLAIS scheme with SNR=15dB and $b_{DAC} = 8$.

Fig. 8 describes the convergence of proposed TLAIS method with SNR = 15dB and $b_{DAC} = 8$. As can be seen from Fig. 8, we observe that the proposed TLAIS method converges rapidly after 5 or 6 iterative times. Additionally, the first and second iterations harvest the largest SR gains, and then it grows slowly during the following iterations.

V. CONCLUSION

In this paper, a TLAIS was proposed for AN-aided hybrid precoding by taking the low-resolution DACs and finite-quantized phase shifters into account. First, we developed a quantized DAC model for AN aided hybrid precoding based on AQN model including the effect of quantization noise. Then, an approximate expression of SR with partial channel knowledge of eavesdropper was also derived. Based on the derived approximate expression, a TLAIS was presented to optimize the design of DP vector, ANPM, and AP matrix by resorting to the principle of GPI. Considering the unit modulus constraints of AP, we adopted a GA algorithm to compute the AP matrix. Furthermore, to reduce the computation complexity, only the non-zero elements in AP matrix were abstracted by taking advantage of the sparsity features of partially-connected hybrid architecture to form an optimization vector. Simulation results indicate that our proposed method can achieve a better SR performance than existing methods such as MRT and Max-SR+NSP.

REFERENCES

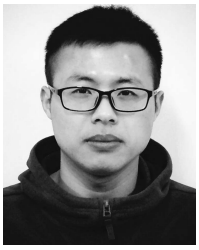
[1] T. Bai, A. Alkhateeb, and R. W. Heath, Jr., "Coverage and capacity of millimeter-wave cellular networks," *IEEE Commun. Mag.*, vol. 52, no. 9, pp. 70–77, Sep. 2014.

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

[3] K. Guan, B. Ai, B. Peng, D. He, G. Li, J. Yang, Z. Zhong, and T. Kürner, "Towards realistic high-speed train channels at 5G millimeter-wave band—Part I: Paradigm, significance analysis, and scenario reconstruction," *IEEE Trans. Veh. Technol.*, vol. 67, no. 10, pp. 9112–9128, Oct. 2018.
[4] Q. Wu, G. Y. Li, W. Chen, D. W. K. Ng, and R. Schober, "An overview of sustainable green 5G networks," *IEEE Wireless Commun.*, vol. 24, no. 4, pp. 72–80, Aug. 2017.
[5] T. S. Rappaport, G. R. Maccartney, M. K. Samimi, and S. Sun, "Wideband millimeter-wave propagation measurements and channel models for future wireless communication system design," *IEEE Trans. Commun.*, vol. 63, no. 9, pp. 3029–3056, Sep. 2015.
[6] O. El Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, Jr., "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1499–1513, Mar. 2014.
[7] Y.-Y. Lee, C.-H. Wang, and Y.-H. Huang, "A hybrid RF/baseband precoding processor based on parallel-index-selection matrix-inversion-bypass simultaneous orthogonal matching pursuit for millimeter wave MIMO systems," *IEEE Trans. Signal Process.*, vol. 63, no. 2, pp. 305–317, Jan. 2015.
[8] X. Yu, J.-C. Shen, J. Zhang, and K. B. Letaief, "Alternating minimization algorithms for hybrid precoding in millimeter wave MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 485–500, Apr. 2016.
[9] X. Gao, L. Dai, S. Han, I. Chih-Lin, and R. W. Heath, Jr., "Energy-efficient hybrid analog and digital precoding for MmWave MIMO systems with large antenna arrays," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 4, pp. 998–1009, Apr. 2016.
[10] F. Sohrabi and W. Yu, "Hybrid analog and digital beamforming for mmWave OFDM large-scale antenna arrays," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 7, pp. 1432–1443, Jul. 2017.
[11] S. Park, A. Alkhateeb, and R. W. Heath, Jr., "Dynamic subarrays for hybrid precoding in wideband mmWave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 5, pp. 2907–2920, May 2017.
[12] F. Shu, Y. Qin, T. Liu, L. Gui, Y. Zhang, J. Li, and Z. Han, "Low-complexity and high-resolution DOA estimation for hybrid analog and digital massive MIMO receive array," *IEEE Trans. Commun.*, vol. 66, no. 6, pp. 2487–2501, Jun. 2018.
[13] L. Sun, Y. Qin, Z. Zhuang, R. Chen, Y. Zhang, J. Lu, F. Shu, and J. Wang, "A robust secure hybrid analog and digital receive beamforming scheme for efficient interference reduction," *IEEE Access*, vol. 7, pp. 22227–22234, Feb. 2019.
[14] L. Zhou, S. Du, H. Zhu, C. Chen, K. Ota, and M. Dong, "Location privacy in usage-based automotive insurance: Attacks and countermeasures," *IEEE Trans. Inf. Forensics Secur.*, vol. 14, no. 1, pp. 196–211, Jan. 2019.
[15] L. Wang, Z. Zhang, M. Dong, L. Wang, Z. Cao, and Y. Yang, "Securing named data networking: Attribute-based encryption and beyond," *IEEE Commun. Mag.*, vol. 56, no. 11, pp. 76–81, Nov. 2018.
[16] J. Wu, M. Dong, K. Ota, J. Li, and Z. Guan, "Big data analysis-based secure cluster management for optimized control plane in software-defined networks," *IEEE Trans. Netw. Service Manag.*, vol. 15, no. 1, pp. 27–38, Mar. 2018.
[17] L. Kuang, L. T. Yang, J. Feng, and M. Dong, "Secure tensor decomposition using fully homomorphic encryption scheme," *IEEE Trans. Cloud Comput.*, vol. 6, no. 3, pp. 868–878, Jul./Sep. 2018.
[18] X. Zhou, Q. Wu, S. Yan, F. Shu, and J. Li, "UAV-enabled secure communications: Joint trajectory and transmit power optimization," *IEEE Trans. Veh. Technol.*, vol. 68, no. 4, pp. 4069–4073, Apr. 2019.
[19] A. D. Wyner, "The wire-tap channel," *Bell Syst. Tech. J.*, vol. 54, no. 8, pp. 1355–1387, Oct. 1975.
[20] X. Chen, D. W. K. Ng, W. H. Gerstacker, and H.-H. Chen, "A survey on multiple-antenna techniques for physical layer security," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 2, pp. 1027–1053, 2nd Quart., 2017.
[21] A. Khisti and G. W. Wornell, "Secure transmission with multiple antennas—Part II: The MIMOME wiretap channel," *IEEE Trans. Inf. Theory*, vol. 56, no. 11, pp. 5515–5532, Jul. 2010.
[22] A. Khisti and G. W. Wornell, "Secure transmission with multiple antennas I: The MISOME wiretap channel," *IEEE Trans. Inf. Theory*, vol. 56, no. 7, pp. 3088–3104, Jul. 2010.
[23] S. Goel and R. Negi, "Guaranteeing secrecy using artificial noise," *IEEE Trans. Wireless Commun.*, vol. 7, no. 6, pp. 2180–2189, Jun. 2008.
[24] Q. Li and W.-K. Ma, "Spatially selective artificial-noise aided transmit optimization for MISO multi-eves secrecy rate maximization," *IEEE Trans. Signal Process.*, vol. 61, no. 10, pp. 2704–2717, May 2013.
[25] R. Negi and S. Goel, "Secret communication using artificial noise," in *Proc. VTC-Fall IEEE 62nd Veh. Technol. Conf.*, vol. 3, Sep. 2005, pp. 1906–1910.
[26] G. Xia, F. Shu, Y. Zhang, J. Wang, S. ten Brink, and J. Speidel, "Antenna selection method of maximizing secrecy rate for green secure spatial modulation," *IEEE Trans. Green Commun. Netw.*, vol. 3, no. 2, pp. 288–301, Jun. 2019.
[27] Y. R. Ramadan, H. Minn, and A. S. Ibrahim, "Hybrid analog–digital precoding design for secrecy mmWave MISO-OFDM systems," *IEEE Trans. Commun.*, vol. 65, no. 11, pp. 5009–5026, Nov. 2017.
[28] Y. Ju, H.-M. Wang, T.-X. Zheng, and Q. Yin, "Secure transmissions in millimeter wave systems," *IEEE Trans. Commun.*, vol. 65, no. 5, pp. 2114–2127, May 2017.
[29] Y. R. Ramadan and H. Minn, "Artificial noise aided hybrid precoding design for secure mmWave MISO systems with partial channel knowledge," *IEEE Signal Process. Lett.*, vol. 24, no. 11, pp. 1729–1733, Nov. 2017.
[30] O. Orhan, E. Erkip, and S. Rangan, "Low power analog-to-digital conversion in millimeter wave systems: Impact of resolution and bandwidth on performance," in *Proc. Inf. Theory Appl. Workshop (ITA)*, Feb. 2015, pp. 191–198.
[31] W. B. Abbas, F. Gomez-Cuba, and M. Zorzi, "Millimeter wave receiver efficiency: A comprehensive comparison of beamforming schemes with low resolution ADCs," *IEEE Trans. Wireless Commun.*, vol. 16, no. 12, pp. 8131–8146, Dec. 2017.
[32] K. Roth and J. A. Nossek, "Achievable rate and energy efficiency of hybrid and digital beamforming receivers with low resolution ADC," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 9, pp. 2056–2068, Sep. 2017.
[33] L. N. Ribeiro, S. Schwarz, M. Rupp, and A. L. F. de Almeida, "Energy efficiency of mmWave massive MIMO precoding with low-resolution DACs," *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 2, pp. 298–312, May 2018.
[34] P.-H. Lin, S.-H. Lai, S.-C. Lin, and H.-J. Su, "On secrecy rate of the generalized artificial-noise assisted secure beamforming for wiretap channels," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 9, pp. 1728–1740, Sep. 2013.
[35] W.-C. Liao, T.-H. Chang, W.-K. Ma, and C.-Y. Chi, "QoS-based transmit beamforming in the presence of eavesdroppers: An optimized artificial-noise-aided approach," *IEEE Trans. Signal Process.*, vol. 59, no. 3, pp. 1202–1216, Mar. 2011.
[36] X. Zhou, J. Li, F. Shu, Q. Wu, Y. Wu, W. Chen, and L. Hanzo, "Secure SWIPT for directional modulation-aided AF relaying networks," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 2, pp. 253–268, Feb. 2019.
[37] N. Lee, H. J. Yang, and J. Chun, "Achievable sum-rate maximizing AF relay beamforming scheme in two-way relay channels," in *Proc. ICC Workshops IEEE Int. Conf. Commun. Workshops*, May 2008, pp. 300–305.

LING XU received the B.S. degree from the Nanjing University of Science and Technology, China, in 2017, where she is currently pursuing the M.S. degree with the School of Electronic and Optical Engineering. Her research interests include wireless communication and mobile networks.

LINLIN SUN received the B.S. and M.S. degrees from the Nanjing University of Science and Technology, Nanjing, China, in 2000 and 2003, respectively, where she is currently pursuing the Ph.D. degree with the School of Electronic and Optical Engineering. She has been an Associate Professor with the Nanjing University of Science and Technology. As the project leader and main performer, she has undertaken and completed more than 30 national, provincial, and ministerial research projects. Her main research interests include satellite communication, millimeter-wave communication, microwave millimeter-wave chip, and solid-state device application.



GUIYANG XIA received the B.S. degree from Huangshan University, Huangshan, China, in 2013, and the M.S. degree from Bohai University, Jinzhou, China, in 2016. He is currently pursuing the Ph.D. degree with the School of Electronic and Optical Engineering, Nanjing University of Science and Technology, China. His research interests include signal processing for wireless communication, physical layer security, and spatial modulation systems.

TINGTING LIU (M'12) received the B.S. degree in communication engineering and the Ph.D. degree in information and communication engineering from the Nanjing University of Science and Technology, Nanjing, China, in 2005 and 2011, respectively. Since 2011, she has been with the School of Communication Engineering, Nanjing Institute of Technology, China. She is currently a Postdoctoral with the Nanjing University of Science and Technology. Her research interests include vehicular fog computing, QoS, game theory, caching-enabled systems, device-to-device networks, and cognitive radio networks.

FENG SHU (M'16) was born in 1973. He received the B.S. degree from the Fuyang Teaching College, Fuyang, China, in 1994, the M.S. degree from Xidian University, Xi'an, China, in 1997, and the Ph.D. degree from Southeast University, Nanjing, China, in 2002. In 2005, he joined the School of Electronic and Optical Engineering, Nanjing University of Science and Technology, Nanjing, where he is currently a Professor, and also a Supervisor of the Ph.D. and graduate students. From 2009 to 2010, he held a Visiting Postdoctoral position at The University of Texas at Dallas. He has published about 300 papers, of which over 200 are in archival journals, including more than 60 papers on the IEEE journals and more than 100 SCI-indexed papers. He holds seven Chinese patents. His research interests include wireless networks, wireless location, and array signal processing. He serves as a TPC member for several international conferences, including the IEEE ICC 2019, the IEEE ICCS 2018/2016, ISAPE 2018, and WCSP 2017/2016/2014. He was awarded with the Mingjiang Chair Professor in Fujian Province. He is currently an Editor of the journal IEEE Access.

YIJIN ZHANG received the B.S. degree from the Nanjing University of Posts and Telecommunications, China, in 2004, the M.S. degree from Southeast University, China, in 2007, and the Ph.D. degree from the Chinese University of Hong Kong, in 2010, all in information engineering. Since 2019, he has been a Professor with the School of Electronic and Optical Engineering, Nanjing University of Science and Technology. His research interests include sequence design and resource allocation in communication networks.

JIANGZHOU WANG (F'17) is currently a Professor and the former Head of the School of Engineering and Digital Arts, University of Kent, U.K. He has published over 300 papers in international journals and conferences and four books in the areas of wireless mobile communications.

Prof. Wang is a Fellow of the Royal Academy of Engineering, U.K. He received the Best Paper Award from the IEEE GLOBECOM2012. He was the IEEE Distinguished Lecturer, from 2013 to 2014. He was the Technical Program Chair of the 2019 IEEE International Conference on Communications (ICC2019), Shanghai, the Executive Chair of the IEEE ICC2015, London, and the Technical Program Chair of the IEEE WCNC2013. He has served as an Editor for a number of international journals, including the IEEE Transactions on Communications, from 1998 to 2013.