Analytical Beam Training for RIS-Assisted Wideband Terahertz Communication

Yuhao Chen, Jingbo Tan, and Linglong Dai
Department of Electronic Engineering, Tsinghua University, Beijing 100084, China
Beijing National Research Center for Information Science and Technology, Beijing 100084, China
E-mail: {chen-yh21, tanjb17}@mails.tsinghua.edu.cn, daill@tsinghua.edu.cn

GLOBECOM 2023 - 2023 IEEE Global Communications Conference | 979-8-3503-1090-0/23/$31.00 ©2023 IEEE | DOI: 10.1109/GLOBECOM54140.2023.10437694

***Abstract*—Terahertz (THz) communication has been considered as one of the promising technologies for future 6G wireless systems. In order to cope with the high path loss in THz systems, reconfigurable intelligent surface (RIS) with low-complexity reflecting elements has been proposed to strengthen signals and improve the spectrum and energy efficiency. In order to acquire accurate direction of the user equippment (UE) to send directional beams, beam training is usually utilized. However, existing beam training frameworks have not taken the wideband beam split effect into consideration, so the beam training accuracy decreases a lot in wideband scenario. To solve the problems mentioned above, we propose an analytical beam training framework in RIS-assisted wideband THz communication systems. Specifically, we firstly propose a power distribution pattern (PDP) based direction estimation scheme, where the exact value of the received power is utilized to analytically calculate the direction. Then, we design the analytical codebook for the proposed framework based on the inherent parameters of the wideband THz system. Simulation results show that the proposed framework can achieve the near-optimal achievable rate performance with a lower beam training overhead.**

***Index Terms*—Terahertz communication, reconfigurable intelligent surface, beam training, wide beam.**

I. Introduction

Terahertz (THz) communication is considered as one of the promising technologies to satisfy the high data rate requirement in future 6G systems due to its capacity to provide tens of GHz bandwidth [1]. However, due to the high frequency of THz band, the THz signals suffer from a severe path loss [2], which limits the coverage of THz signals. Moreover, significant attenuation happens when THz signals are blocked by obstacles such as buildings and trees. Fortunately, reconfigurable intelligent surface (RIS), which is composed of a large number of low-complexity reflecting elements, has been proposed to handle the above two problems [3], [4], which can achieve high spectrum and energy efficiency.

In practical communication systems, the BS and RIS need to know the channel state information (CSI) to realize accurate beamforming. The CSI is usually obtained by channel estimation. However, due to the high overhead needed for accurate channel estimation, the beam training scheme is utilized to realize the near-optimal achievable rate performance with relatively low overhead [5], [6]. Specifically, instead of estimating the entire channel, the beam training scheme only need to utilize directional beams to explore the possible directions in space and choose the best physical direction to transmit signals. So far, several beam training frameworks have been proposed [7]–[11]. For example, in [7], a discrete fourier transform (DFT) based beam training framework was proposed to explore the whole space with DFT codewords. By selecting the direction with maximum received power, the direction of UE was acquired. In order to further reducing the training overhead, in [8]–[10], hierachical codebooks were proposed for massive multiple-input multiple-output (MIMO) systems, where each beam of the upper layer was divided into several narrower beams in the lower layer. By conducting the beam training layer by layer, the best direction can be decided based on the maximum received power at the UE. However, the hierarchical search scheme requires frequent feedback from the user to the BS/RIS, which introduces extra burden to the communication systems. To solve this problem, [11] proposed a multi-directional beams based beam training framework for RIS-assisted massive MIMO systems, which exploited the inherent sparse structure of the BS-RIS-UE channel. By randomly generating the sensing matrix and carry out a few rounds of full-coverage scanning, the best direction lied in the intersection of the generated multi-directional beams. However, in RIS-assisted wideband THz systems, existing beam training frameworks may suffer from severe performance degradation due to the beam split effect in wideband systems [12]. Therefore, it is vital to eliminate the consequence brought by the beam split effect and improve the accuracy of beam training in RIS-assisted wideband systems.

In this paper, we propose an analytical beam training framework for RIS-assisted wideband THz communication systems, where we exploit the power distribution pattern (PDP) of different subcarriers at different physical directions to analytically calculate the direction of UE based on the exact value of the received power. By considering the beam split effect in wideband THz systems during the direction estimation, the accuracy of beam training improves. Specifically, we first elaborate the analysis of the PDP at different physical angles for a certain beam, which is related to the central frequency and the bandwidth of the THz system. Based on the PDP, we propose a PDP-based direction estimation scheme to analytically acquire the direction of UE. Then, we design a codebook for the proposed analytical beam training framework depending on the inherent parameters of the wideband THz systems such as the central frequency and the bandwidth. Finally, simulation results are provided to validate the proposed beam training framework, which reveals that the proposed framework can realize a more accurate beam training and achieve a higher data rate than existing schemes in RIS-assisted wideband THz

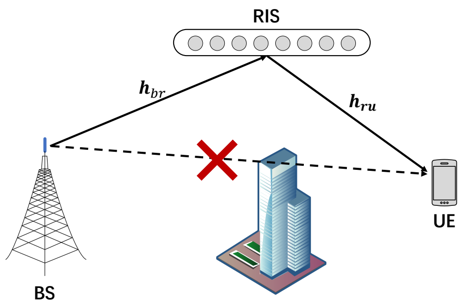


Fig. 1. RIS-assisted wideband THz communication system

systems.

II. System Model and Background

In this section, we firstly introduce the system model of the RIS-assisted wideband THz communication system. Then, the traditional beam training framework based on DFT codebook is reviewed.

A. System Model

We consider a downlink RIS-assisted wideband THz communication system, as illustrated in Fig. 1, where both the BS and the UE have a single antenna, and the RIS is an N-element uniform linear array (ULA). The BS uses orthogonal frequency division multiplexing (OFDM) with M subcarriers to serve the UE. The bandwidth of the system is denoted as B.

We considered the ray-based channel model for wideband THz channel [13]. Specifically, the downlink channel at the m-th subcarrier between the BS and the RIS $\mathbf{h}_{br,m} \in \mathbb{C}^{N\times 1}$ with $m = 1, 2, ..., M$ can be denoted as

$$\mathbf{h}_{br,m} = \sum_{l_1=1}^{L_1} g_{br,m}^{(l_1)} e^{-j\pi\tau_{br,m}^{(l_1)} f_m} \mathbf{a}_N \left(\varphi_m^{(l_1)}\right), \tag{1}$$

where $g_{br,m}^{(l_1)}$ and $\tau_{br,m}^{(l_1)}$ denote the path gain and the time delay of the l_1-th path for $l_1 = 1, 2, ..., L_1$, respectively, f_m denotes the frequency of the m-th subcarrier, which satisfies $f_m = f_c + \frac{B}{M}\left(m - 1 - \frac{M-1}{2}\right)$ with f_c being the central frequency of the system, L_1 denotes the number of paths, and $\mathbf{a}_N \in \mathbb{C}^{N\times 1}$ denotes the array response vector which satisfies

$$\mathbf{a}_N\left(\varphi_m^{(l_1)}\right) = \frac{1}{\sqrt{N}}\left[1, e^{j\pi\varphi_m^{(l_1)}}, e^{2j\pi\varphi_m^{(l_1)}}, ..., e^{(N-1)j\pi\varphi_m^{(l_1)}}\right]^T, \tag{2}$$

where $\varphi_m^{(l_1)}$ denotes the spatial direction of the l_1-th path, which satisfies $\varphi_m^{(l_1)} = \frac{2d}{c} f_m \sin(\gamma^{(l_1)})$ for $l_1 = 1, 2, ..., L_1$ with $\gamma^{(l_1)}$ being the physical direction of the l_1-th path, d is the antenna spacing, c is the speed of light, and λ_c is the wavelength at the central frequency. We define $d = \frac{\lambda_c}{2} = \frac{c}{2f_c}$.

Similar to the definition of $\mathbf{h}_{br,m}$, the downlink channel at the m-th subcarrier between the RIS and the UE $\mathbf{h}_{ru,m} \in \mathbb{C}^{N\times 1}$ with $m = 1, 2, ..., M$ can be denoted as

$$\mathbf{h}_{ru,m} = \sum_{l_2=1}^{L_2} g_{ru,m}^{(l_2)} e^{-j\pi\tau_{ru,m}^{(l_2)} f_m} \mathbf{a}_N\left(\psi_m^{(l_2)}\right), \tag{3}$$

where $g_{ru,m}^{(l_2)}, \tau_{br,m}^{(l_2)}, f_m, L_2, \psi_m^{(l_2)}$ denote the path gain of the l_2-th path, the time delay of the l_2-th path, the frequency of the m-th subcarrier, the number of paths and the spatial angle of the l_2-th path, respectively.

For the RIS, each of the element can re-scatter the incident signal with a certain amount of phase shift. Let θ_n and $\beta_n \in [0, 1]$ denote the phase shift and the amplitude reflection coefficient of the n-th element on the RIS, respectively. The reflecting matrix $\mathbf{\Theta}$ of the RIS can be presented as

$$\mathbf{\Theta} \triangleq \mathrm{diag}\left(\beta_1 e^{j\theta_1}, \beta_2 e^{j\theta_2}, ..., \beta_N e^{j\theta_N}\right), \tag{4}$$

where $\theta_n \in [0, 2\pi)$ and $\beta_n \in [0, 1]$ for $n = 1, 2, ..., N$. Eq. (4) reveals that the antenna elements of RIS are frequency-independent. For simplicity, we assume the phase can be shifted consecutively and $\beta_n = 1$ for all $n \in \{1, 2, ..., N\}$.

Based on the channel model and the reflecting matrix mentioned above, by considering the scenario where the direct link between BS and UE is blocked by obstacles like buildings, the received signal y_m at the m-th subcarrier at the UE side can be presented as

$$y_m = \mathbf{h}_{ru,m}^H \mathbf{\Theta} \mathbf{h}_{br,m} s + n, \tag{5}$$

where $\mathbf{h}_{ru,m}, \mathbf{\Theta}, \mathbf{h}_{br,m}$ denote the channel between RIS and UE, the reflecting matrix of RIS and the channel between BS and RIS, respectively, s denotes the transmitted signal at the BS, and n is the AWGN noise following the distribution $\mathcal{CN}(0, \sigma^2)$ with σ^2 being the noise power.

B. Conventional Beam Training Framework

Conventional beam training framework [7] usually apply the DFT codebook, which can be presented as

$$\begin{aligned}\mathbf{W} &= \left[\mathbf{a}_N(-1), \mathbf{a}_N\left(\tfrac{2-N}{N}\right), \cdots, \mathbf{a}_N\left(\tfrac{N-2}{N}\right)\right] \\ &= \begin{bmatrix} 1 & 1 & \cdots & 1 \\ e^{-j\pi} & e^{j\pi\frac{2-N}{N}} & \cdots & e^{j\pi\frac{N-2}{N}} \\ e^{-2j\pi} & e^{j\pi 2\frac{2-N}{N}} & \cdots & e^{j\pi 2\frac{N-2}{N}} \\ \vdots & \vdots & \ddots & \vdots \\ e^{-Nj\pi} & e^{j\pi N\frac{2-N}{N}} & \cdots & e^{j\pi N\frac{N-2}{N}} \end{bmatrix}.\end{aligned} \tag{6}$$

During the beam training process, each column of $\mathbf{W}$ is applied for beamforming to transmit pilots in different time slots, and the UE can receive a series of pilots with the received power being $\mathbf{p} = [p_0, p_1, \cdots, p_{N-1}]$, where $p_n = |y_n|^2 = \left|\mathbf{h}_{ru,m}^H \mathrm{diag}\left(\mathbf{W}[:, n]\right) \mathbf{h}_{br,m}\right|^2$. Let i denotes the index of the largest receiver power, the best direction $\theta_i = \frac{2i-N}{N}$ can be decided accordingly. Then, the system can transmit data with the beamforming vector being $\mathbf{a}_N\left(\frac{2i-N}{N}\right)$.

The above beam training framework is based on the directional narrow beams achieved by DFT codewords. However, in wideband THz communication systems, considering the frequency-independent property of RIS, there exists beam split effect [12], which means the spatial direction φ_m varies at different subcarriers. This effect will lead to the phenomenon that the beams at different subcarriers are steered to different physical directions. As a consequence, traditional beam training frameworks, if directly applied in wideband systems, will experience a severe performance degradation since the beam is not steered to the intended direction. Therefore, an efficient beam training framework for RIS-assisted wideband THz communication systems is needed.

III. Proposed Analytical Beam Training Framework

In order to estimating the direction of UE accurately in RIS-assisted wideband THz communication systems, in this section, we introduce the proposed analytical beam training framework. Specifically, we first give an overview of our proposed framework. Then, we elaborate the analysis of PDP and propose the PDP-based direction estimation scheme. Finally, based on the PDP-based direction estimation scheme, we design the analytical codebook according to the inherent parameter of the THz system for the proposed framework.

A. Overview of Proposed Analytical Beam Training Framework

In RIS-assisted communication systems, the location of the BS and the RIS are fixed once deployed, so the channel between BS and RIS has a much longer coherence time than that between RIS and UE, which can be treated as quasi-static [14]. Therefore, we suppose the channel $\mathbf{h}_{br,m}$ between BS and RIS is known perfectly by the BS and the purpose of beam training is converted to selecting the best beam direction between the RIS and the UE.

Traditional beam training frameworks only utilize the maximum received power to *choose* the best direction from a series of previously designed directions. However, the beam split effect caused by the large bandwidth makes the received power of a certain beam at different directions in space vary in a particular pattern, which is called the power distribution pattern (PDP) in this paper. If the PDP can be analyzed specifically, the direction of the UE can be derived according to the received power. Following this idea, we can take advantage of the PDP brought by beam split effect in wideband THz systems to make this effect a benefit rather than a drawback. The proposed analytical beam training framework can improve the accuracy of beam training since we consider the beam split effect during the analysis of PDP and analytically calculate the best direction, making the beam training accuracy independent of the angle resolution of the generated beams.

Based on this, the proposed analytical beam training framework can be described as follows. During the beam training stage, in the i-th time slot, the BS transmits training signal with the RIS configuration being $\mathbf{\Theta} = \mathrm{diag}\left(\mathbf{W}_a\left[:, i\right]\right)$, which is selected from the proposed analytical codebook. The UE saves the corresponding received power in each time slot. After the time slots that is allocated for beam training are exhausted, the received power is normalized by the power normalization coefficient. The reason for introducing the power normalization coefficient is that the beam width of different beam pairs differs, which lead to the difference in the array gain since the total transmission energy is fixed. Then, in the calculating stage, the beam pair which has the largest normalized received power $\tilde{\mathbf{p}}$ is chosen, based on which the direction of UE can be directly calculated from the proposed PDP-based direction estimation scheme. Finally, we transmit data according to the estimated direction in the data transmission stage.

In the following two subsections, we will introduce in detail the proposed PDP-based direction estimation scheme and the design of our proposed analytical codebook.

B. Proposed PDP-based Direction Estimation Scheme

We first introduce the method that we apply to generate a wide beam. When generating traditional narrow beams with spatial width $W = 2/N$, the beamforming matrix are designed as $\mathbf{\Theta} = \mathrm{diag}\left(\mathbf{a}_N\left(\theta\right)\right)$, which is inversely proportional to the number of elements on RIS. If the number of elements reduces, the spatial width of the beam will naturally increase. Therefore, we follow the method in [15] and divide the RIS into several sub-arrays to form a wide beam with designed width.

Specifically, we divide the RIS into K sub-arrays, which satisfies

$$\frac{2K}{N_s} \geq \varpi, \tag{7}$$

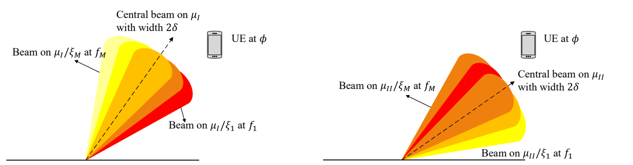
where N_s is the number of antenna elements in each sub-array, and ϖ is the intended width. Since K and N_s are two integers, we have $KN_s \leq N$, which means $K = \lfloor \frac{N}{N_s} \rfloor \geq \frac{N}{N_s} - 1$, so by substituting K with $\frac{N}{N_s} - 1$, we get a sufficient condition for (7) is $N_s \leq \frac{\sqrt{1+2N\varpi}-1}{\varpi}$. In order to get enough array gain to compensate the serious path loss, we need N_s to be as large as possible so the array gain of each sub-array is high enough. Therefore, we increase the N_s until it cannot satisfy (7) to get the largest N_s, and K can be required by $K = \lfloor \frac{N}{N_s} \rfloor$.

After determining K and N_s, we specify the directions of beams generated by different sub-arrays to make them tile on the intended range. Since the direction of the center is μ, and the beam width of each sub-array is $2/N_s$. The direction of the k-th sub-array can be presented as $\nu_k = \mu_c + \frac{2k-K-1}{N_s}, k = 1, 2, ..., K$. Therefore, the reflecting matrix of RIS can be written as

$$\begin{aligned} \mathbf{\Theta}\left[i(k,n), i(k,n)\right] = \frac{1}{\sqrt{N}} e^{j\pi\left[(k-1)N_s+n-1\right]\nu_k} e^{j\epsilon_k}, \\ k = 1, 2, ..., K+1, n = 1, 2, ..., N_s, \end{aligned} \tag{8}$$

where $i(k, n) = (k-1)N_s + n$ denotes the index of the RIS units, and ϵ_k is the phase compensation, which is represented as $\epsilon_k = k\left(\frac{(N_s-1)\pi}{N_s} + \pi\right)$.

For a particular UE at ϕ and a wide beam steered to μ with width 2δ (both in spatial domain), with the subcarrier whose frequency ranging from f_1 to f_M, the directions of beams at



(a) (b)

Fig. 2. Received power of the wide-beam-pair (a) Wide beam I. (b) Wide beam II.

each subcarrier range from $\frac{\mu}{\xi_M}$ to $\frac{\mu}{\xi_1}$, where $\xi_m = f_m/f_c$. As a result, for the m_p-th subcarreier, where m_p satisfies

$$m_p \in \mathcal{M}_p = \left\{ m \mid \tfrac{\mu}{\xi_i} - \delta \le \phi \le \tfrac{\mu}{\xi_m} + \delta, m \in \mathcal{M} \right\}, \quad \mathcal{M} = \{1, 2, ..., M\}, \tag{9}$$

the UE can receive the signal normally. While for the m_n-th subcarrier, which satisfies

$$m_n \in \mathcal{M}_n = \left\{ m \mid \phi < \tfrac{\mu}{\xi_m} - \delta \parallel \phi > \tfrac{\mu}{\xi_m} + \delta, m \in \mathcal{M} \right\}, \quad \mathcal{M} = \{1, 2, ..., M\}, \tag{10}$$

the UE cannot receive the signal normally. Therefore, while μ changes, the received power at UE changes, which carries the information of physical direction of the UE. This property provides us with the possibility to make full use of the exact value of the received power to improve the accuracy and reduce the overhead of beam training in wideband THz communication systems.

We consider a pair of wide beam, as is illustrated in Fig. 2, for wide beam I, subcarriers indexed by $m_{p,I} \in \mathcal{M}_{p,I} = \{m \mid \tilde{m} \le m \le M\}$ can transmit signals normally, while subcarriers indexed by $m_{n,I} \in \mathcal{M}_{n,I} = \{m \mid 1 \le m < \tilde{m}\}$ cannot transmit signals normally. for wide beam II, subcarriers indexed by $m_{n,II} \in \mathcal{M}_{n,II} = \{m \mid 1 \le m < \tilde{m}\}$ can transmit signals normally, while subcarriers indexed by $m_{p,II} \in \mathcal{M}_{p,II} = \{m \mid \tilde{m} \le m \le M\}$ cannot transmit signals normally. This phenomenon results in the difference of the received power corresponding to the two wide beams at the UE. If we design the directions and widths of the beam pair carefully, we can use the received power to analytically derive the physical direction of the UE, which changes the idea of *choosing* to the idea of *calculating*.

We then analyze in detail the received power of the wide-beam-pair. As is illustrated in Fig. 3, we denote the central direction of the wide-beam-pair as $\bar{\mu}$. Since the estimation of the direction is based on the received power, the beam width of each beam pair need to be the same so that their array gain will be the same, which is fair for further calculating. Based on (7), we approximatively set the width of the beam pair as $W = 2\delta = \bar{\mu}B/f_c$. As for the central direction of each wide beam, we want to fully use the channel information carried by each subcarriers. Therefore, for the m-th subcarrier, if it transmits signal normally at beam I, it should not transmit signal at beam II and vice versa. As a result, the difference of the central direction of the two wide beams should equal the beam width, which means $\mu_I = \bar{\mu} - \delta$ and $\mu_{II} = \bar{\mu} + \delta$. With this setting, UE in range $[\bar{\mu} - \delta, \bar{\mu} + \delta]$ is able to receive the signals of both wide beams and the received power can be utilized to calculate the physical direction of this UE.

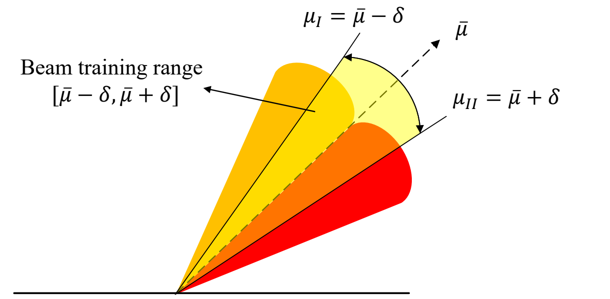


Fig. 3. Designed wide-beam-pair and the corresponding beam training range

For a UE at $\phi \in [\bar{\mu} - \delta, \bar{\mu} + \delta]$, the received power of beam I can be presented as

$$|g_I(\phi)|^2 = \left| \sum_{m=1}^{M} \left[\mathbf{a}_N^H(\phi_m) \boldsymbol{\Theta}(\mu_I) \right] \right|^2 \approx \mathcal{C} \frac{\sin^2\left(\frac{\pi N_s(\phi - \bar{\mu} + \delta)}{2}\right)}{\sin^2\left(\frac{\pi(\phi - \bar{\mu} + \delta)}{2}\right)}, \tag{11}$$

where $\mathcal{C}$ is a constant unrelated to ϕ and μ_I. Similarly, for the same UE at $\phi \in [\bar{\mu} - \delta, \bar{\mu} + \delta]$, the received power of beam II can be presented as

$$|g_{II}(\phi)|^2 \approx \mathcal{C} \frac{\sin^2\left(\frac{\pi N_s(\phi - \bar{\mu} - \delta)}{2}\right)}{\sin^2\left(\frac{\pi(\phi - \bar{\mu} - \delta)}{2}\right)}. \tag{12}$$

Then, we elaborate the proposed PDP-based direction estimation scheme. We introduce the ratio matric χ which is presented as

$$\begin{aligned} \chi &= \frac{|g_I(\phi)|^2 - |g_{II}(\phi)|^2}{|g_I(\phi)|^2 + |g_{II}(\phi)|^2} \\ &= -\frac{\sin(\pi(\phi - \bar{\mu}))\sin(\pi\delta)}{1 - \cos(\pi(\phi - \bar{\mu}))\cos(\pi\delta)}, \end{aligned} \tag{13}$$

where $\phi - \bar{\mu} \in [-\delta, \delta]$. The estimation of direction ϕ can then be acquired by

$$\hat{\phi} = \bar{\mu} - \arcsin\left(\frac{\chi \sin(\pi\delta) - \chi\sqrt{1 - \chi^2}\sin(\pi\delta)\cos(\pi\delta)}{\sin^2(\pi\delta) + \chi^2\cos^2(\pi\delta)} \right). \tag{14}$$

Based on (14), we can estimate the direction ϕ of UE in two time slots.

Compared with traditional beam training schemes based on *choosing* the best direction, the proposed PDP-based direction estimation scheme considers the beam split effect in wideband THz systems and analytically *calculate* the direction of UE based on the PDP of the beam pair, which makes full use of the information that is carried in frequency domain and makes the beam split effect an advantage rather than a drawback. Therefore, the proposed scheme improves the accuracy of beam training since the beam split effect is considered during derivation.

C. Design of the Proposed Analytical Codebook

Based on the above derivation, we will introduce the design of proposed analytical codebook in RIS-assisted wideband THz communication systems. The codebook design can be devided into two steps. The first step is to design the codeword for the directions where beam split effect is very weak, which can be neglected. For this part, using tradtional beam training codebooks is enough. The second step is to design the codeword for the directions where our proposed analytical beam training framework works well. The framework on how to design the codebook is summarized in **Algorithm 1**.

Algorithm 1 Proposed Analytical Codebook Design

Input: Number of units on RIS N; bandwidth B, central frequency f_c; range parameter κ; dividing parameter β

Output: Central directions of the beam pairs $\boldsymbol{\mu}$; estimation range $\boldsymbol{\rho}$; power normalization coefficient $\boldsymbol{\zeta}$; designed analytical codebook $\mathbf{W}$

Initialization:
$\boldsymbol{\mu} = \left[-\frac{1}{N}, \frac{1}{N}\right]$; $\boldsymbol{\rho} = \left[-\frac{1}{N}, \frac{1}{N}\right]$; $\boldsymbol{\zeta} = \left[\frac{2}{N}, \frac{2}{N}\right]$;
while $B\boldsymbol{\mu}[0]/f_c > -\beta$ **do**
    $\boldsymbol{\mu} = \left[\boldsymbol{\mu}[0] - \dfrac{2}{N}, \boldsymbol{\mu}, -\boldsymbol{\mu}[0] + \dfrac{2}{N}\right]$;
    $\boldsymbol{\rho} = \left[\boldsymbol{\rho}[0] - \dfrac{2}{N}, \boldsymbol{\rho}, -\boldsymbol{\rho}[0] + \dfrac{2}{N}\right]$;
    $\boldsymbol{\zeta} = \left[\dfrac{2}{N}, \boldsymbol{\zeta}, \dfrac{2}{N}\right]$;
end while
while $\boldsymbol{\rho}[0] > -1$ **do**
    $\bar{\mu} = -\frac{2\boldsymbol{\rho}[0] f_c}{2f_c - \kappa B}$;
    $\delta = \frac{B}{f_c}\bar{\mu}$;
    $\boldsymbol{\mu} = [-\bar{\mu}, \boldsymbol{\mu}, \bar{\mu}]$;
    $\boldsymbol{\rho} = [-\bar{\mu} - \kappa\delta, \boldsymbol{\rho}, \bar{\mu} + \kappa\delta]$;
    $\boldsymbol{\zeta} = \left[\dfrac{B\bar{\mu}}{f_c}, \boldsymbol{\zeta}, \dfrac{B\bar{\mu}}{f_c}\right]$;
end while
Generate the codebook $\mathbf{W}$ by Eq. (8) based on $\boldsymbol{\mu}$
return $\boldsymbol{\mu}, \boldsymbol{\rho}, \boldsymbol{\zeta}, \mathbf{W}$

Specifically, we need to generate the central direction of the beam pairs $\boldsymbol{\mu}$, the estimation range $\boldsymbol{\rho}$ and the power normalization coefficient $\boldsymbol{\zeta}$ for beam training, based on which we generate the analytical codebook. During the generation process, we first introduce a dividing parameter β to divide the directions where beam split effect can be neglected and the directions where beam split effect is serious. For the first part, we utilize traditional narrow beam to estimate the optimal direction, so the difference of central direction is $2/N$. Traditionally, the direction of narrow beam is equivalent to the estimation range, so $\boldsymbol{\rho}$ is equal to $\boldsymbol{\mu}$ in this part, and the power normalization coefficient $\boldsymbol{\zeta}$ is equal to the width of generated narrow beam. For the second part where the beam split effect is serious, we introduce the range parameter κ. According to (14), for a pair of designed wide beam, although we can theoretically estimate the direction in range $[\bar{\mu} - \delta, \bar{\mu} + \delta]$, the gradient near the boundary is approximately 0, which means a slight error in χ will result in a big mistake in ϕ. In practical communication systems, there exist various kinds of noise and the error in χ is inevitable, so we need to introduce the range parameter $\kappa < 1$ to limit the estimation range in $[\bar{\mu} - \kappa\delta, \bar{\mu} + \kappa\delta]$ to improve the accuracy of estimation. Based on the previous estimation range $\bar{\mu} + \kappa\delta$, the central direction of the next wide beam pair should satisfy

$$\begin{aligned} &\bar{\mu}_{new} - \kappa\frac{B}{2f_c}\bar{\mu}_{new} = \bar{\mu}_{old} + \kappa\frac{B}{2f_c}\bar{\mu}_{old} = -\boldsymbol{\rho}[0], \\ &\qquad\qquad \Longrightarrow \bar{\mu}_{new} = -\frac{2\boldsymbol{\rho}[0] f_c}{2f_c - \kappa B}. \end{aligned} \tag{15}$$

Then the central direction of the next wide beam is determined. We can extend the codebook until the estimation range can cover the whole space. It is worth noting that the array gain of different wide beam is different because the total transmission power is fixed but the width of wide beam varies. Therefore, we need to generate a power normalization coefficient for each wide beam, whose value is in proportion to the beam width. After getting the received power of each codeword $\mathbf{p}$, we get the normalized received power by $\tilde{\mathbf{p}} = \mathbf{p} \odot \boldsymbol{\zeta}$. By choosing the maximum element in $\tilde{\mathbf{p}}$, we can determine which pair of wide beam can estimate the direction of this user. Then the estimated direction $\hat{\phi}$ can be obtained by (14).

IV. Simulation Results

In this section, simulation results are provided to show the achievable rate performance of the proposed analytical beam training framework. The parameters of the RIS-assisted wideband THz communication system are set as: $N_{\rm BS} = 1$, $N_{\rm RIS} = 1024$, $N_{\rm UE} = 1$, $f_c = 100$ GHz, $B = 10$ GHz, the number of subcarriers are set to be 128. We consider that the THz channel is quasi-optical, and the number of paths is set as $L = 1$. The direction of UE is set to satisfy $\phi \sim \mathcal{U}(-\pi/3, \pi/3)$.

Fig. 4 illustrates the achievable rate performance of the proposed beam training framework compared with multi-diretional beam training framework mentioned in Subsection I and traditional exhaustive beam training framework. Here we set the training overhead as 128. The parameter Q in multi-directional beam training framework represents the number of beams sent at each slots. We can observe from Fig. 4 that our proposed analytical beam training framework outperforms the typical frameworks, and it can obtain the near-optimal achievable rate performance compared with the optimal situation

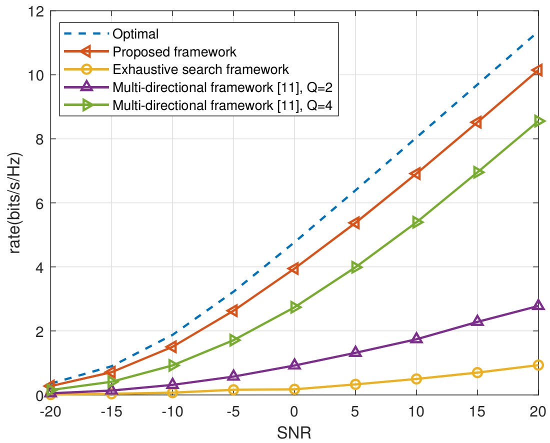


Fig. 4. The achievable rate performance comparison against SNR

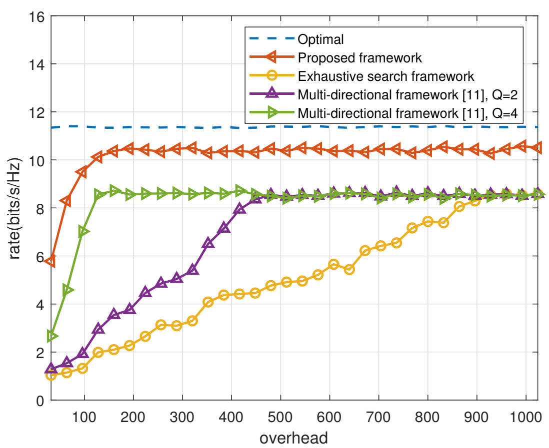


Fig. 5. The achievable sum-rate performance against the beam training overhead

whose direction of UE assumed to be known perfectly by the BS and RIS. In addition, the trational exhaustive search cannot work in such a low overhead. We can also observe that existing beam training scheme such as the exhaustive search framework can barely work in such a low beam training overhead.

To better demonstrate the overhead different frameworks need, Fig. 5 illustrated the achievable rate performance of different frameworks as the training overhead increases. Here we set the SNR is set as 5. The *x-axis* represent the beam training overhead. We can observe from Fig. 5 that proposed analytical beam training framework can achieve near-optimal achievable rate performance after the training overhead is sufficient, and the proposed analytical beam training framework outperforms the existing frameworks. In addition, when the training overhead is limited, the achievable rate performance is far better than existing frameworks. This is because our proposed framework can reduce the training overhead to a large extent when the UE is far from $0°$ since the beam split effect is severe. By scanning the space from $90°/-90°$ to $0°$, we can accurately estimate a large proportion of directions with a very low training overhead. We can also observe from Fig. 5 that the traditional exhaustive search framework can barely work when the number of RIS units is very large.

To sum up, the proposed analytical beam training framework can reach the near-optimal achivable rate with a low training overhead, and is adaptive to future communication systems with a large number of antenna units.

V. Conclusion

In this paper, we investigated the beam training problem in RIS-assisted wideband THz communication systems. To address the beam split problem in wideband communication systems, we proposed the analytical beam training framework, which promoted the traditional idea of *choosing* the best direction to the new idea of *calculating* the best direction. Since the proposed framework considered the beam split effect in wideband communication system and took full advantage of the relative values of the received power, it could improve the beam training accuracy in future communication systems. Simulation results show that the proposed framework can obtain the near-optimal achievable rate performance with a low overhead, which outperforms the existing beam training framework.

Acknowledgement

This work was supported in part by the National Key Research and Development Program of China (Grant No. 2020YFB1805005) and in part by the European Commission through the H2020-MSCA-ITN META WIRELESS Research Project under Grant 956256.

References

[1] T. S. Rappaport, Y. Xing, O. Kanhere, S. Ju, A. Madanayake, S. Mandal, A. Alkhateeb, and G. C. Trichopoulos, "Wireless communications and applications above 100 GHz: Opportunities and challenges for 6G and beyond," *IEEE Access*, vol. 7, pp. 78 729–78 757, Jun. 2019.

[2] C. Han, A. O. Bicen, and I. F. Akyildiz, "Multi-wideband waveform design for distance-adaptive wireless communications in the terahertz band," *IEEE Trans. Signal Process.*, vol. 64, no. 4, pp. 910–922, Feb. 2016.

[3] Q. Wu and R. Zhang, "Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network," *IEEE Commun. Mag.*, vol. 58, no. 1, pp. 106–112, Jan. 2020.

[4] ——, "Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming," *IEEE Trans. Wireless Commun.*, vol. 18, no. 11, pp. 5394–5409, Nov. 2019.

[5] W. Wu, D. Liu, X. Hou, and M. Liu, "Low-complexity beam training for 5G millimeter-wave massive MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 69, no. 1, pp. 361–376, Jan. 2020.

[6] A. Ali, N. González-Prelcic, and R. W. Heath, "Millimeter wave beam-selection using out-of-band spatial information," *IEEE Trans. Wireless Commun.*, vol. 17, no. 2, pp. 1038–1052, Feb. 2018.

[7] J. Suh, C. Kim, W. Sung, J. So, and S. W. Heo, "Construction of a generalized DFT codebook using channel-adaptive parameters," *IEEE Commun. Lett.*, vol. 21, no. 1, pp. 196–199, Jan. 2017.

[8] Z. Xiao, T. He, P. Xia, and X.-G. Xia, "Hierarchical codebook design for beamforming training in millimeter-wave communication," *IEEE Trans. Wireless Commun.*, vol. 15, no. 5, pp. 3380–3392, May 2016.

[9] H. Yu, P. Guan, W. Qu, and Y. Zhao, "An improved beam training scheme under hierarchical codebook," *IEEE Access*, vol. 8, pp. 53 627–53 635, Mar. 2020.

[10] B. Ning, Z. Chen, W. Chen, Y. Du, and J. Fang, "Terahertz multi-user massive MIMO with intelligent reflecting surface: Beam training and hybrid beamforming," *IEEE Trans. Veh. Technol.*, vol. 70, no. 2, pp. 1376–1393, Feb. 2021.

[11] P. Wang, J. Fang, W. Zhang, and H. Li, "Fast beam training and alignment for IRS-assisted millimeter wave/terahertz systems," *IEEE Trans. Wireless Commun.*, pp. 1–1, Apr. 2021.
[12] L. Dai, J. Tan, Z. Chen, and H. Vincent Poor, "Delay-phase precoding for wideband thz massive mimo," *IEEE Trans. Wireless Commun.*, pp. 1–1, 2022.
[13] L. You, X. Gao, G. Y. Li, X.-G. Xia, and N. Ma, "Bdma for millimeter-wave/terahertz massive mimo transmission with per-beam synchronization," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 7, pp. 1550–1563, Jul. 2017.
[14] C. Hu, L. Dai, S. Han, and X. Wang, "Two-timescale channel estimation for reconfigurable intelligent surface aided wireless communications," *IEEE Trans. on Commun.*, vol. 69, no. 11, pp. 7736–7747, Nov. 2021.
[15] L. Zhu, J. Zhang, Z. Xiao, X. Cao, D. O. Wu, and X.-G. Xia, "3-D beamforming for flexible coverage in millimeter-wave UAV communications," *IEEE Wireless Commun. Lett.*, vol. 8, no. 3, pp. 837–840, Jun. 2019.