Nanoscale Reconfigurable Intelligent Surface Design and Performance Analysis for Terahertz Communications

Xinying Ma, Zhi Chen, Senior Member, IEEE, and Chongwen Huang, Member, IEEE

Abstract—Terahertz (THz) communications have been envisioned as a promising enabler to provide ultra-high data transmission for sixth generation (6G) wireless networks. To tackle the blockage vulnerability brought by severe attenuation and poor diffraction of THz waves, a nanoscale reconfigurable intelligent surface (NRIS) is developed to smartly manipulate the propagation directions of incident THz waves. In this paper, the electric properties of the graphene are investigated by revealing the relationship between conductivity and applied voltages, and then an efficient hardware structure of electrically-controlled NRIS is designed based on Fabry-Perot resonance model. Particularly, the phase response of NRIS can be programmed up to 306.82 degrees. To analyze the hardware performance, we jointly design the passive and active beamforming for NRIS aided THz communication system. Particularly, an adaptive gradient descent (A-GD) algorithm is developed to optimize the phase shift matrix of NRIS by dynamically updating the step size during the iterative process. Finally, numerical results demonstrate the effectiveness of our designed hardware architecture as well as the developed algorithm.

Index Terms—Terahertz (THz) communications, nanoscale reconfigurable intelligent surface (NRIS), hybrid beamforming, adaptive gradient descent (A-GD).

I. INTRODUCTION

With the continuous explosion growth of data traffic in wireless communications, sixth generation (6G) communication networks are expected to meet a great deal of pressing requirements in the near future [1]. To meet the bandwidth-hungry applications (e.g., on-chip communication, virtual reality, kiosk downloading service, nanoscale localization and nanoscale biomedical communication), terahertz (THz) frequency band (0.1–10 THz) has been regarded as a prospective alternative to provide large spectrum bandwidth and support ultra-high data transmission for 6G communication networks [2]. However, there are still some imperative challenges existing in THz communications. On the one hand, due to the high path attenuation and strong molecular absorption effect experienced by THz waves, the transmission distance of THz communications is limited within a small area, and thus is not applicable for the practical communication scenarios [4], [5]. On the other hand, THz waves at such a high frequency band undergo extremely poor diffraction and are easily blocked by the obstacles. To tackle this issue, the concept of nanoscale reconfigurable intelligent surface (NRIS) is newly proposed to mitigate blockage vulnerability and improve coverage capability [6], [7], [8]. To be specific, the NRIS, which consists of a large number of passive reflecting elements is a kind of physical meta-surface and belongs to a special case of conventional RIS except for the extremely small physical size. Each reflecting element is capable of adjusting the phase shifts by using a smart central processor [9]. In addition, NRIS is passive and lacks the active radio frequency (RF) chains, and thus is more energy-efficient compared with existing active devices, such as amplify-and-forward relaying [10] and massive multiple-input multiple-output (MIMO). Therefore, the combination of NRIS and THz communications is worthy of further exploration.

Deploying NRISs in the THz communication system is essential, but some challenges also emerge accordingly. To realize reliable THz communications, the channel state information (CSI) acquisition is the primary mission before the data transmission begins. Different from conventional communication systems with active devices, the main difficulty of channel estimation in NRIS-enabled THz systems is that these reflecting elements are unable to execute the signal processing. Prominently, by leveraging the sparse features of THz MIMO channel, the work in [11] converts the channel estimation problem into the sparse signal recovery problem, and a low complexity compressed sensing based channel estimation scheme is developed to realize the efficient signal reconstruction. Once the CSI is acquired at the base station (BS) side or the mobile station (MS) side, the passive and active beamforming can be jointly designed for NRIS-enabled MIMO systems. For instance, the work in [12] optimizes the transmit beamforming at the BS and the passive beamforming at the NRIS for the purpose of minimizing the total transmit power at the BS. To reap the benefits of multiple NRISs, the authors aim to solve a new cooperative multibeam multi-hop routing design problem and maximize the minimum
received signal power among all users [13], [14]. Besides, the authors investigate the joint design of digital beamforming at the BS and analog beamforming at NRISs for the NRIS-enabled THz MIMO systems from the perspective of deep reinforcement learning [15]. In addition, other software research directions are also investigated extensively, such as energy efficiency optimization [8], [16], data rate maximization [17], [18], [19], secure communication [20]. With regard to the hardware design of NRIS, many existing works pay attention to testing NRIS characteristics and implementing NRIS prototypes at THz frequency bands [21], [22], [23], but the algorithm design is not actually taken into consideration. Apart from these aforementioned research interests, the joint hardware design and hybrid beamforming for the NRIS-empowered THz MIMO communication system is still treated as an open problem.

In order to compensate for the research gap, the joint hardware design and hybrid beamforming optimization for the NRIS-enabled THz MIMO system is presented in this paper, which is the first attempt to practically combine the hardware characteristics of the NRIS and the software design together. In particular, a novel graphene-based hardware structure of NRIS with a wide phase response range and a desired reflecting amplitude is designed. Considering the hardware properties of NRIS, we develop a downlink NRIS-enabled THz MIMO system model and propose a gradient descent based method to optimize the phase shifts of NRIS. Compared with the conventional MIMO system without NRIS [24], [25], the maximum rate optimization problem of the NRIS-enabled THz MIMO system involves multiple matrix variables, and thus is more sophisticated. The main contributions of this paper can be summarized as follows.

- To begin with, we design a practical graphene-based NRIS hardware structure, where its phase response can be controlled up to 306.82 degrees, and the reflecting amplitude efficiency is more than 50% at 1.6 THz. Furthermore, the design theory and working principle of the NRIS are also provided. Then, the electric properties of the graphene are introduced by revealing the relationship between conductivity and applied voltage, which is the foundation of forming an electrically-controlled NRIS.

- Based on the hardware features of NRIS, we employ singular value decomposition (SVD) based method to obtain the optimally active beamforming at transceiver and propose an adaptive gradient descent (A-GD) algorithm to achieve the passive beamforming design for NRIS. Specifically, compared with conventional gradient descent (C-GD) algorithm with the fixed step size, the proposed A-GD algorithm turns out to be more efficient by dynamically updating the step size during the iterative process, and thus is able to realize a better achievable rate performance.

- Finally, simulation results demonstrate that our designed NRIS hardware structure with the practical phase response of 306.82 degrees achieves basically identical performance in contrast with the ideal phase response of 360 degrees. More importantly, with the assistance of NRIS, the achievable rate performance of our proposed A-GD algorithm greatly outweighs C-GD algorithm and the conventional THz systems without NRIS.

The rest of this paper is organized as follows. In Section II, the electric properties and hardware structure of graphene-based NRIS is presented. Section III describes the NRIS based software design for THz communications. Finally, the simulation results and conclusion are presented in Section IV and Section V, respectively.

II. GRAPHENE BASED HARDWARE DESIGN OF NRIS

In this section, an efficient graphene-based NRIS is designed that realizes a wide range of amplitude response and phase response. The key component of NRIS is the reflecting elements with sub-wavelength thickness, and each reflecting element needs to be designed as a tunable resonant structure. In addition, by arranging these reflecting elements in a specific way, the whole NRIS is capable of achieving diverse functions, such as phase control, anomalous reflection and planar focusing. Finally, the detailed working principle and device performance can be presented in the following.

A. Electric Properties of Graphene At THz Band

In order to achieve the beam controllability, tunable components are embedded into the NRIS elements [26]. In general, semiconductor devices are extensively utilized at microwave or millimeter wave frequency band (e.g., varactor diode [27], switching diode [28]). Since the physical size of each reflecting element at THz band is extremely small, the diodes and transistors cannot be integrated into such a nanoscale structure. In this case, graphene is a kind of appropriate material to facilitate the NRIS with ultra-small size and tunable property.

Graphene is a two-dimensional material consisting of a single layer of carbon atoms. The conductivity of graphene can be altered through the applied voltage bias in a relatively wide range. Therefore, graphene provides various resonant states for each NRIS element. According to [29], the conductivity of graphene at THz band can be written as

$$\sigma = \frac{2e^2}{\pi \hbar^2} k_B T \cdot \ln \left(2 \cos \left(\frac{E_F}{2k_BT}\right)\right) \frac{i}{\omega + i\tau^{-1}},$$

where $e$ is the elementary charge; $\hbar$ is the reduced Planck constant; $k_B$ is the Boltzmann constant; $T$ is the temperature; $E_F$ is the Fermi level; $\tau$ is the relaxation time; and $\omega$ is the angular frequency, respectively. It can be concluded that at a certain frequency point, the conductivity is only determined by Fermi level [30]. Then we can get the following expression as

$$|E_F| = \hbar \nu_F \sqrt{\pi n_d},$$

where $\nu_F$ is the Fermi velocity and $n_d$ is the carrier density which can be expressed as

$$n_d = \sqrt{n_0^2 + \alpha_c |V_{CNP} - V_g|^2},$$

where $n_0$ is the residual carrier density and $\alpha_c$ is capacitivity related to the electrode. Besides, $V_{CNP}$ is the compensating voltage, and $V_g$ is the applied voltage [31]. In summary, the conductivity of graphene can be continuously changed by the applied voltages, which is the foundation of forming an electrically controlled NRIS.
Fig. 1. Hardware structure of a nanoscale reflecting element that consists of graphene, quartz (substrate) and gold (ground plane) with \( a = 66 \) µm, \( g = 70 \) µm, \( t = 38 \) µm and \( t_m = 1 \) µm.

B. Hardware Design of Graphene Based NRIS

The EM responses of the reflecting elements play an important role in the hardware structure of the NRIS. Fig. 1 shows a typical hardware design of a NRIS element, which can be divided into three parts from top to the bottom: the graphene layer, the substrate and the metallic ground plane. The resonance model of this NRIS architecture can be described as a Fabry-Perot cavity, where EM waves reflect back and forth between the top and the bottom surfaces. In addition, the resonance responses are caused by constructive or destructive interference of the multiple reflections [32]. In terms of such a reflecting element structure as shown in Fig. 1, the phase response based on [33] can be expressed as

\[
\varphi = m\pi - a k_0 \text{Re}(n_{\text{eff}}),
\]

where \( m \) is an integer; \( a \) is the width of graphene patch; \( k_0 \) is the wave number of free space; and \( n_{\text{eff}} \) is the effective refraction index of the resonant structure, which is related to the effective permittivity \( \varepsilon_{\text{eff}} \) of the graphene. In light of [34], the parameter \( \varepsilon_{\text{eff}} \) can be written as

\[
\varepsilon_{\text{eff}} = 1 + \frac{i \sigma_c}{\omega \varepsilon_0 t_g},
\]

where \( \sigma_c \) is the conductivity and \( t_g \) denotes the thickness of graphene. Combining (4) and (5), the phase response can be altered by the conductivity of graphene as well as the applied voltages.

The reflecting elements are simulated by leveraging the frequency domain solver in the simulation environment of CST Microwave Studio 2016. It is worth noting that a single reflecting element is unable to work since the miniature size causes the strong scattering. As a result, the boundary condition of the NRIS is set as ‘unit cell’ to mimic the repeated arrangement of the NRIS elements. Fig. 2 illustrates the reflecting coefficients from 1.4 THz to 1.8 THz with various Fermi levels. By combining Fig. 2(a) and Fig. 2(b), our designed NRIS performs relatively stable broadband characteristics. However, a narrowband working mode of the NRIS elements is selected in this paper where the center frequency is located at 1.6 THz. Fig. 2(c) verifies that the amplitude efficiency of our designed reflecting element at 1.6 THz is more than 50% and the phase response reaches 306.82 degrees along with the chemical potential ranging from 0 eV to 2 eV. Moreover, the discrete phase shifts at 1.6 THz are also well-distributed with diverse Fermi levels, which lays the foundation for the bit quantization operation of phase shifts.

Once the hardware structure of a single reflecting element is designed, the whole NRIS is able to be accomplished by...
arranging massive reflecting elements closely in an array structure, as shown in Fig. 3. Then, various beam steering functions can be realized by controlling the phase response of all the NRIS elements. But in practice, the appropriate number of the NRIS elements needs to be carefully selected, which can make a better tradeoff between performance and deployment cost. According to Fig. 2(c), any expected phase shift within the phase response range can be obtained via applying the voltages continuously. However, the continuous phase control for each NRIS element results in practical problems [35], such as hardware complexity, power consumption, size limitation and the accuracy of phase control. To this end, the discrete phase shifts are considered as the hardware structure of NRIS. We define the discrete phase set of each NRIS element as \( \mathcal{F} = \{0, \varphi_{\text{max}}/2^b, \ldots, 2^b - 1 \varphi_{\text{max}}/2^b\} \), where \( \varphi_{\text{max}} \) is the maximum phase response and \( b \) is the bit quantization number. Then, we define the reflecting amplitude set as \( \mathcal{A} = \{\mu_1, \ldots, \mu_{2^b}\} \), where \( |\mathcal{F}| = 2^b \). From Fig. 2(c) we can note that when the distribution of the discrete phase shifts is determined, \( \mathcal{A} \) can be acquired accordingly. In other words, there is a fixed mapping relationship between phase shift and reflecting amplitude for a specific hardware structure of NRIS. Thus, the reflecting amplitude for each reflecting element can be further defined as \( \bar{\mu} = (\sum_{i=1}^{2^b} \mu_i) / |\mathcal{F}| \), where the averaged amplitude \( \bar{\mu} \in [0.5, 1] \) is determined by parameter \( b \) with the given NRIS structure. Specifically, considering these practical constraints, the maximum phase and amplitude response of NRIS are set as 306.82 degrees and 0.8, respectively.

III. SYSTEM MODEL AND BEAMFORMING DESIGN

In this section, we mainly introduce the NRIS-aided THz system model and present a joint passive and active beamforming design framework.

A. System Model

Considering a downlink THz MIMO system with hybrid beamforming architecture as shown in Fig. 4, a BS employs \( N_{\text{BS}} \) antennas to serve a MS equipped with \( N_{\text{MS}} \) antennas. Since the LoS path between the BS and the MS is usually blocked by the obstacles, we suppose that the NRIS is installed to assist THz communication links, which consists of \( N_{\text{RIS}} \) passive reflecting elements. Due to the serious scattering loss of THz wave, we only consider the one-bounce reflected NLoS paths in the channel model [36]. In addition, a central controller that connects the BS and the NRIS is also deployed to realize the phase adjustment of reflecting elements. Also, we assume that there are \( M_{\text{BS}} \) RF chains at BS side, and \( M_{\text{MS}} \) RF chains at MS side. Considering the heavy power consumption of RF circuits, the hybrid beamforming architecture is employed at the BS and MS, and the number of array antennas is larger than the number of RF chains, i.e., \( N_{\text{BS}} > M_{\text{BS}}, N_{\text{MS}} > M_{\text{MS}} \).

When the BS sends \( \mathbf{s} \) data streams \( \mathbf{s} \in \mathbb{C}^{N_{\text{BS}} \times 1} \) that satisfies \( \mathbb{E}[\mathbf{s}^H \mathbf{s}] = \frac{1}{N_{\text{BS}}} \mathbf{I}_{N_{\text{BS}}} \), the MS employs \( M_{\text{MS}} \) RF chains to receive the processed signals. Thus, the received signal \( \mathbf{y}_r \in \mathbb{C}^{N_{\text{MS}} \times 1} \) can be expressed as

\[
\mathbf{y}_r = \sqrt{\rho} \mathbf{W}_{\text{BB}}^H \mathbf{W}_{\text{RF}}^H \mathbf{H}_2 \Phi \mathbf{H}_1 \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}} \mathbf{s} + \mathbf{W}_{\text{BB}}^H \mathbf{W}_{\text{RF}}^H \mathbf{n},
\]

(6)

where \( \rho \) is transmit power; \( \mathbf{H}_1 \in \mathbb{C}^{N_{\text{BS}} \times N_{\text{RIS}}} \) denotes BS-NRIS channel; \( \mathbf{H}_2 \in \mathbb{C}^{N_{\text{MS}} \times N_{\text{RIS}}} \) is NRIS-MS channel; \( \mathbf{F}_{\text{RF}} \in \mathbb{C}^{N_{\text{BS}} \times M_{\text{BS}}} \) (\( \mathbf{F}_{\text{BB}} \in \mathbb{C}^{N_{\text{MS}} \times M_{\text{MS}}} \)) denotes the analog (digital) precoding matrix; \( \mathbf{W}_{\text{RF}} \in \mathbb{C}^{N_{\text{MS}} \times M_{\text{BS}}} \) and \( \mathbf{W}_{\text{BB}} \in \mathbb{C}^{N_{\text{MS}} \times N_{\text{MS}}} \) are analog and digital combining matrices; \( \mathbf{n} \in \mathbb{C}^{N_{\text{MS}} \times 1} \) represents the additive white Gaussian noise (AWGN) vector following the distribution of \( \mathcal{CN}(0, \sigma^2 \mathbf{I}_{N_{\text{MS}}}) \); and \( \Phi = \text{diag}(\mu_1 e^{j\varphi_1}, \mu_2 e^{j\varphi_2}, \ldots, \mu_{N_{\text{RIS}}} e^{j\varphi_{N_{\text{RIS}}}}) \) is a diagonal matrix with the dimension of \( N_{\text{RIS}} \times N_{\text{RIS}} \), respectively. Each entry \( \{\mu_1 e^{j\varphi_1}\} \) of \( \Phi \) indicates the reflecting coefficient of a NRIS element, and is composed of the reflecting amplitude \( \mu \) and the phase shift \( \{\varphi_n\} \in [0, \pi] \). Both \( \mu \) and \( \varphi_n \) are closely related with the hardware structure of NRIS that is concretely introduced in Section II.

B. Channel Model

The NRIS-enabled THz MIMO channel model contains \( \mathbf{H}_1 \), \( \Phi \) and \( \mathbf{H}_2 \), and the cascaded channel is denoted as \( \mathbf{H}_c = \mathbf{H}_2 \Phi \mathbf{H}_1 \). We assume that both \( \mathbf{H}_2 \) and \( \mathbf{H}_2 \) consist of a LoS path and several non-line-of-sight (NLoS) paths, as we take the sparse nature of the THz channel into consideration. On the basic of geometric channel model [37], the BS-NRIS channel \( \mathbf{H}_1 \) can be written as

\[
\mathbf{H}_1 = \sqrt{N_{\text{BS}}N_{\text{RIS}}} \alpha_0 \mathbf{a}_R (\theta_{R,0}^1, \theta_{R,0}^2) \mathbf{a}_B^H (\theta_{B,0}^1, \theta_{B,0}^2) + \sqrt{N_{\text{BS}}N_{\text{RIS}}} \sum_{l=1}^{L} \alpha_1 \mathbf{a}_R (\theta_{R,l}^1, \theta_{R,l}^2) \mathbf{a}_B^H (\theta_{B,l}^1, \theta_{B,l}^2),
\]

(7)

where \( L \) is the number of NLoS paths; \( \theta_{R}^1, \theta_{R}^2 \) denotes the angle of arrival (AoA) of \( \mathbf{H}_1 \) in the azimuth (elevation) domain; and \( \theta_{B}^1, \theta_{B}^2 \) denotes the angles of departure (AoD) of \( \mathbf{H}_1 \) in the azimuth (elevation) domain, respectively. Considering the large number of THz array antennas, the uniform planar array (UPA) structure is adopted as the array geometry. The normalized array response for the \( N_{\nu} \)-element UPA on \( x-y \)-plane can be expressed as

\[
\mathbf{a}_B = \frac{1}{\sqrt{N_{\nu}}} \left[ e^{j \frac{2\pi}{\lambda} \left( a \sin(\theta_B^e) \cos(\theta_B^h) + q \cos(\theta_B^e) \right)}, \ldots \right]^T,
\]

(8)
where \( p \in [0, N_x - 1], q \in [0, N_y - 1] \) and \( N_x N_y = N_{BS} \), respectively. The spacing of THz array antenna is \( d = \lambda/2 \) where \( \lambda \) is the incident wavelength. Similar to (8), \( \alpha_{1k}(\theta_{1k}^R, \theta_{2k}^R) \) also employs the UPA structure, but the spacing of the adjacent NRIS elements is the side length of each reflecting element. In addition, \( \alpha_0 \) is the LoS path gain of \( H_1 \). As discussed in [38], \( \alpha_0 \) consists of spreading loss and molecular absorbing loss, which can be expressed as

\[
\alpha_0 = \frac{c}{4\pi f} r_0 \left( e^{-\frac{1}{2}c\lambda(f) r_0} e^{-j2\pi f \tau_{\text{los}}}, \right)
\]

where \( c \) is the speed of light; \( r_0 \) is the straight distance between the BS and the MS; \( \tau_{\text{los}} = r_0/c \) is the time-of-arrival of the LoS path; and \( \kappa(f) \) is the molecular absorbing coefficient. In addition, the channel gain \( \alpha_l \) for the \( l \)th reflected path based on [38] and [39] can be written as

\[
\alpha_l = \frac{c\xi(f)}{4\pi f} (r_1 + r_2) e^{-\frac{1}{2}c\lambda(f)(r_1 + r_2) - j2\pi f \tau_{\text{refl}}},
\]

where \( \xi(f) \) is the reflection coefficient of the reflecting materials (e.g., concrete, plastic, glass); \( r_1 \) is the distance between the transmitter and the reflecting material; \( r_2 \) is the distance between the receiver and the reflecting material; and \( \tau_{\text{refl}} = \tau_{\text{los}} + (r_1 + r_2 - r)/c \) is the time-of-arrival of the reflected path, respectively. Besides, the channel characteristics of \( H_2 \) are identical to \( H_1 \), so we can generate \( H_2 \) in the same way. We assume that the channel estimation problem has been extensively studied in [11], [40], and these effective channel estimation methods can be well leveraged in this paper.

### C. SVD-Based Active Beamforming Design

In light of the system model and hardware architecture of NRIS, the achievable rate of the NRIS-aided THz MIMO system can be expressed as

\[
R = \log_2 \left| I_{N_s} + \frac{\rho}{\delta^2 N_s} (W^H W)^{-1} W^H \right| \times (H_2 \Phi H_1)^F (H_2 \Phi H_1)^F W^H,
\]

where we define \( F = F_{RF} F_{BB} \) and \( W = W_{RF} W_{BB} \). Then problem (11) can be reformulated as

\[
(\Phi_{\text{opt}}, W_{\text{opt}}, F_{\text{opt}}) = \arg \max_{\Phi, W, F} R
\]

\[
\text{s.t. } \varphi_{\text{max}} = 306.82^\circ,
\]

\[
\mu = 0.8, \forall n = 1, \ldots, N_{RIS},
\]

\[
\varphi_n \in \mathcal{F}, \forall n = 1, \ldots, N_{RIS},
\]

\[
||F||^2_F = N_s,
\]

where (12b) is from our designed discrete phase response and (12c) stems from our designed reflecting amplitude. Due to the non-convex and discrete constraints, it is challenging to directly solve the problem (12). Fortunately, one available way of settling such an optimization problem is to first design hybrid beamforming \( F \) and \( W \), and then optimize the phase shift matrix \( \Phi \), respectively.

Given a fixed \( \Phi \), the singular value decomposition (SVD) of the cascaded channel \( H_c \) can be written as

\[
H_c = U \Lambda V^H
\]

\[
= \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Lambda_1 & 0 \\ 0 & \Lambda_2 \end{bmatrix} \begin{bmatrix} V_1 & V_2 \end{bmatrix}^H,
\]

where \( U \) is a \( N_{BS} \times G \) unitary matrix; \( \Lambda \) is a \( G \times G \) dimensional matrix; \( V \) is a \( N_{BS} \times G \) unitary matrix; \( U_1 \) is a \( N_{BS} \times N_s \) submatrix of \( U = [U_1, U_2] \); and \( V_1 \) is a \( N_{BS} \times N_s \) submatrix of \( V = [V_1, V_2] \), in which \( G \triangleq \text{rank}(H_a) \). Especially, \( \Lambda_1 \) and \( \Lambda_2 \) are \( N_s \times N_s \) and \( (G - N_s) \times (G - N_s) \) diagonal matrices with the singular values arranged in a decreasing order. According to [25], the optimal precoding matrix and combining matrix can be expressed as

\[
F_{\text{opt}} = V_1, \quad W_{\text{opt}} = U_1,
\]

where the transmit power satisfies \( ||F_{\text{opt}}||^2_F = N_s \).

### D. Proposed A-GD for Passive Beamforming Design

Given the active beamforming matrices \( F_{\text{opt}} \) and \( W_{\text{opt}} \), the following target is to optimize the phase shift matrix \( \Phi \). In
addition, the achievable rate $R$ in (11) can be rewritten as
\[
R = \log_2 \left| \mathbf{I}_{N_s} + \frac{\rho}{\delta^2 N_s} \mathbf{A}_1 \mathbf{A}_1^H \right|. \quad (15)
\]

Subsequently, (15) can be further simplified as
\[
R \leq N_s \log_2 \left( 1 + \frac{\rho}{\delta^2 N_s} \right) \left[ \text{tr} (\mathbf{A}_1^2) \right] \quad (a)
\]
\[
\leq N_s \log_2 \left( 1 + \frac{\rho}{\delta^2 N_s} \right) \left[ \text{tr} (\mathbf{H}_s \mathbf{H}_s^H) \right], \quad (b)
\]
where (a) comes from Jensen’s inequality and (b) takes the mark of equality when $N_s = G$. Therefore, the optimization problem (12) can be formulated as
\[
\Phi_{\text{opt}} = \arg \max_{\Phi} \text{tr} (\mathbf{H}_s \mathbf{H}_s^H) \quad (17a)
\]
\[
s.t. \quad \varphi_{\text{max}} = 306.82^\circ, \quad (17b)
\]
\[
\varphi_e = 0.8, \quad \forall n = 1, \ldots, N_{\text{RIS}}, \quad (17c)
\]
\[
\varphi_{n} \in \mathcal{F}, \quad \forall n = 1, \ldots, N_{\text{RIS}}, \quad (17d)
\]

Nevertheless, it is worth noting that problem (17) is still a constrained optimization problem, since $\Phi$ possesses discrete phase shifts and constant-magnitude entries. Let us define $\varphi = [\varphi_1, \varphi_2, \ldots, \varphi_{N_{\text{RIS}}}]$, and then consider $\Phi$ as a function of $\theta$ where $\theta = \Phi \mathbf{1}_{N_{\text{RIS}}} = [\mu \varphi_{N_{\text{RIS}}}, \mu \varphi_{N_{\text{RIS}}+1}, \ldots, \mu \varphi_{N_{\text{RIS}}} - 1]$. Given this, we temporarily consider the continuous phase, and thus problem (17) can be reformulated as
\[
\arg \min_{\theta} f(\theta) = -\theta^H \mathbf{D} \theta \quad (18a)
\]
\[
s.t. \quad \varphi_{\text{max}} = 306.82^\circ, \quad (18b)
\]
\[
\varphi_n = 0.8, \quad \forall n = 1, \ldots, N_{\text{RIS}}, \quad (18c)
\]
where $f(\theta)$ is a compound function, $\mathbf{D} = \mathbf{D}^H \mathbf{D}$ and $\mathbf{D} = [\cdots, (\mathbf{H}_s \otimes \mathbf{H}_s)_{\mathcal{F}}, (\cdots, N_{\text{RIS}}+n+1, \cdots), \cdots], \forall n = 1, 2, \ldots, N_{\text{RIS}}$.

Note that problem (18) still suffers from constant-magnitude constraint due to the variable $\theta$. To obtain an unconstrained problem, $f(\theta)$ can be equivalently rewritten as
\[
f(\varphi) = -\sum_{p=1}^{N_{\text{RIS}}} \sum_{q=1}^{N_{\text{RIS}}} \theta_p^H \mathbf{D}_{p,q} \theta_q = -\mu^2 \sum_{p=1}^{N_{\text{RIS}}} \sum_{q=1}^{N_{\text{RIS}}} e^{-j\varphi_p} \mathbf{D}_{p,q} e^{j\varphi_q}. \quad (19)
\]

Since $\mathbf{D} \in \mathcal{C}^{N_{\text{RIS}} \times N_{\text{RIS}}}$ is a positive-definite Hermitian matrix, we can get $\mathbf{D} = \mathbf{D}^H$. Given continuous phase shifts $\{\varphi_n\}_{N_{\text{RIS}}}^1$, the $n$th element of the gradient vector $\nabla \varphi f(\varphi)$ can be calculated as
\[
\frac{\partial f(\varphi)}{\partial \varphi_n} = \mu^2 j e^{-j\varphi_n} \sum_{q=1}^{N_{\text{RIS}}} \mathbf{D}_{n,q} e^{j\varphi_q}. \quad (20)
\]

After calculating all the $\partial f(\varphi)/\partial \varphi_n, \forall n = 1, \ldots, N_{\text{RIS}}$, $\nabla \varphi f(\varphi)$ can be expressed as
\[
\nabla \varphi f(\varphi) = \left[ \frac{\partial f(\varphi)}{\partial \varphi_1}, \frac{\partial f(\varphi)}{\partial \varphi_2}, \ldots, \frac{\partial f(\varphi)}{\partial \varphi_{N_{\text{RIS}}}} \right]^T. \quad (21)
\]

On basis of the gradient direction $\nabla \varphi f(\varphi)$, the objective function $f(\varphi)$ is able to descend by replacing $\varphi$ with $\varphi - \lambda \text{diag}(\nabla \varphi f(\varphi))$, where $\lambda$ is the iterative step size. During the $i$th iteration, the updated $\varphi^{(i+1)}$ and the renewed $f(\varphi^{(i+1)})$ can be respectively written as
\[
\varphi^{(i+1)} = \varphi^{(i)} - \lambda \nabla \varphi f(\varphi^{(i)}) \quad (22a)
\]
\[
f(\varphi^{(i+1)}) = f(\varphi^{(i)} - \lambda \nabla \varphi f(\varphi^{(i)})) \quad (22b)
\]

Since the C-GD algorithm obtains the fixed step size $\lambda$ by simulation experiment [6], it suffers from high complexity and low efficiency. To this end, a novel A-GD algorithm is developed to determine the adaptive step size $\lambda^i$. Based on above definition, (22) can be formulated as
\[
f(\varphi^{(i)}) = -\sum_{p=1}^{N_{\text{RIS}}} \sum_{q=1}^{N_{\text{RIS}}} (\theta_p^{(i)})^H \mathbf{D}_{p,q} \theta_q^{(i)} = -\mu^2 \sum_{p=1}^{N_{\text{RIS}}} \sum_{q=1}^{N_{\text{RIS}}} \mathbf{D}_{p,q} e^{j\varphi_p^{(i)}} e^{j\varphi_q^{(i)}} \Delta \varphi_p^{(i)} \Delta \varphi_q^{(i)}, \quad (23)
\]

where $\Delta \varphi_p^{(i)} = -\lambda^i \partial f(\varphi^{(i)})/\partial \varphi_p^{(i)}$ for $n = 1, 2, \ldots, N_{\text{RIS}}$. Note that $f(\varphi^{(i+1)})$ is only determined by $\lambda^i$ during $(i+1)$th iteration. Hence, the optimization problem with regard to $\lambda^i$ can be expressed as
\[
\lambda^i \approx \arg \min_{\lambda^i} -\mu^2 \sum_{p=1}^{N_{\text{RIS}}} \sum_{q=1}^{N_{\text{RIS}}} \mathbf{D}_{p,q} e^{j\varphi_p^{(i)}} e^{j\varphi_q^{(i)}} \times e^{-j \lambda^i \varphi_p^{(i)}} e^{-j \lambda^i \varphi_q^{(i)}} \quad (24)
\]

To smartly control $\lambda^i$ during iterative process, we replace $e^{j \lambda^i \varphi_p^{(i)}}$ by the second-order Taylor expansion formulation. Then, (24) can be approximated as
\[
\lambda^i \approx \arg \min_{\lambda^i} -\mu^2 \sum_{p=1}^{N_{\text{RIS}}} \sum_{q=1}^{N_{\text{RIS}}} \mathbf{D}_{p,q} e^{j\varphi_p^{(i)}} e^{j\varphi_q^{(i)}} e^{-j \lambda^i \varphi_p^{(i)}} e^{-j \lambda^i \varphi_q^{(i)}} \times \left( 1 + j \lambda^i \Gamma_{p,q} + \frac{(j \lambda^i \Gamma_{p,q})^2}{2} \right) = \arg \min_{\lambda^i} C_0 + C_1 \lambda^i + C_2 \lambda^i, \quad (25)
\]

where $\Gamma_{p,q} = \partial f(\varphi^{(i)})/\partial \varphi_p^{(i)} - \partial f(\varphi^{(i)})/\partial \varphi_q^{(i)}$.

It is worth noting that there are two different cases for the quadratic function in (25), including the positive value of term $C_2$ and the negative value of term $C_2$. For these two cases, $\lambda^i$
Algorithm 1: Proposed A-GD Algorithm.

Require: $H_1$, $H_2$, $\rho$, $N_s$, $\delta$, $F$, $b$, $\phi_{\max}$, $I$.
1: Initialize $\mathbf{D}$, $\mathbf{W}^{\text{opt}}$, $\mathbf{F}^{\text{opt}}$, $\Phi = \mu I_{N_{\text{RIS}} \times N_{\text{RIS}}}$, $\phi^0 = 0_{N_{\text{RIS}} \times 1}$, $I_{\max} = 0$, $i = 0$.
2: while $i \leq I_{\max}$ do
3: Calculate the gradient vector $\nabla \phi f(\phi^i)$ in (21),
4: Calculate the step size $\lambda^i$ according to (26),
5: Update $\phi^{i+1} = \phi^i - \lambda^i \nabla \phi f(\phi^i)$,
6: if $f(\phi^{i+1}) > f_{\max}$ do
7: $f_{\max} = f(\phi^{i+1})$, $\phi^{\text{opt}} = \phi^{i+1}$,
8: end if
9: $i = i + 1$;
10: end while
11: Map each entry of $\phi^{\text{opt}}$ into discrete phase from $F$.
Ensure: $\Phi^{\text{opt}}$, $R$

The main steps of our proposed A-GD algorithm are illustrated in Algorithm 1.

E. Complexity Analysis

The main computational complexity of Algorithm 1 comes from updating $f(\phi^{i+1})$ in (22) and $\lambda^i$ in (26). Specifically, the complexity of calculating (22) for each iteration can be expressed as $\mathcal{O}(4N_{\text{RIS}}^2 + 9N_{\text{RIS}}^2)$. For the sake of updating step size $\lambda^i$, the complexity of executing (26) for each iteration is $\mathcal{O}(2N_{\text{RIS}}^3 + 8N_{\text{RIS}}^2)$. As a consequence, the total complexity of Algorithm 1 can be given by

$$\mathcal{O}\left(I_{\max}\left(6N_{\text{RIS}}^3 + 17N_{\text{RIS}}^2\right)\right),$$

where $I_{\max}$ represents the maximum iterations of our proposed A-GD Algorithm.

IV. SIMULATION RESULTS

In this section, simulation results are provided to examine the effectiveness of joint hardware and algorithm design for NRIS-enabled THz system, including A-GD algorithm, C-GD algorithm, random phase algorithm and the conventional THz system without NRIS. In addition, the distances of BS-NRIS, NRIS-MS and BS-MS are set as $r_0 = 10$ m, $r_0 = 20$ m, and $r_0 = 25$ m respectively, which can commendably meet the indoor communication scenarios. The LoS path of BS-MS link is blocked by the obstacle, and thus the NLoS paths are assisted by NRIS. Considering the sparse nature of THz channel, we assume that $H_2$ contains $L_1 = 3$ propagation paths, e.g., one LoS path and two NLoS paths. More specifically, the complex gain of LoS path is generated based on (9) and the complex gains of NLoS paths are computed by (11). The molecular absorbing coefficient and the reflection coefficient of ceramic tile are set as $\kappa(f) = 0.2$ and $\xi(f) = 10^{-6}$ [38]. Similarly, the parameter settings of $H_2$ are consistent with $H_1$. We further set $N_{\text{RIS}} = 512$, $N_{\text{MS}} = 256$, $N_{\text{BS}} = 6$, $M_{\text{BS}} = 32$, and $M_{\text{MS}} = 4$. Here we define the signal-to-noise ratio (SNR) as $\text{SNR} = \rho/\sigma^2$, and all simulation results are averaged over 1000 random channel realizations.

A. Hardware Influence for THz System

Fig. 5 investigates the achievable rate with the increasing number of the maximum phase response $\phi_{\max}$, which aims to validate the effectiveness of the proposed hardware architecture mentioned in Section II. From Fig. 5 we can note that the achievable rates of the considered NRIS-enabled algorithms improve firstly and then converge to a fixed value. Remarkably, the achievable rate of our proposed A-GD algorithm with $\phi_{\max} = 306.82^\circ$ have already converged, and possess the same performance as the ideal case with $\phi_{\text{ideal}} = 360^\circ$. Compared with low phase response $\phi_{\max} = 60^\circ$, the considered A-GD algorithm and the random phase scheme with $\phi_{\max} = 306.82^\circ$ are able to realize the achievable rate enhancement of 3.59 bps/Hz and 2.68 bps/Hz, respectively. Hence, the numerical results reveal that our proposed graphene-based NRIS is an efficient hardware structure.

Fig. 6 provides the achievable rate performance versus different bit quantization values $b$. Fig. 6 indicates that the achievable rates of the random phase scheme and the conventional THz system without NRIS are insensitive to $b$. Instead, the achievable rates of A-GD and AO algorithms are greatly affected by $b$. In the case of $b = 1$, our developed A-GD algorithm suffers from the obvious performance degradation due to the limited quantization precision of NRIS. In contrast with $b = 2$, our proposed A-GD algorithm with $b = 1$ endures about 0.93 bps/Hz performance penalty, respectively. Intriguingly, the achievable rates of our proposed algorithms with $b = 2$ almost yields the similar performance compared with $b \geq 3$, indicating that $b = 2$ is sufficient to quantize the discrete phase for NRIS in practice.

B. Performance Analysis for NRIS-Aided THz System

Fig. 7 depicts the achievable rate versus diverse SNR in NRIS-aided THz MIMO system. As shown in Fig. 7, the achievable rate of the NRIS-aided THz system greatly outstrips the conventional
THz system without NRIS that hardly meet the future communication requirements. Numerically, the performance gap between the random phase scheme and the THz system without NRIS is about 4.24 bps/Hz under the condition of SNR = 10 dB. Meanwhile, Fig. 7 also indicates that the achievable rate of our proposed A-GD algorithm is around 8.4 bps/Hz higher than the random phase scheme, respectively. More importantly, by dynamically adjusting the step size, our proposed A-GD algorithm is superior to the C-GD algorithm with fixed step size, which demonstrates that our proposed optimization algorithm can be employed to further enhance the achievable rate performance for the NRIS-aided THz MIMO systems.

Fig. 8 discusses the achievable rate comparisons of the considered schemes versus the number of NRIS elements. From Fig. 8, we can note that the achievable rate of the THz system without NRIS case has the worst performance due to the lack of NRIS, and remains unchanged along with the diverse values of $N_{RIS}$. In terms of these NRIS-aided algorithms, the achievable rate of our developed A-GD algorithm outstrips the C-GD algorithm and the random phase scheme. Moreover, when the number of reflecting elements increases, the performance gaps between A-GD algorithm and the random phase scheme become much larger. Specifically, our developed A-GD algorithm realizes around 9.21 bps/Hz performance improvement compared with the random phase scheme with $N_{RIS} = 192$. Hence, NRIS can provide obvious performance gain for THz communications, and our proposed software design is capable of further improving the achievable rate. Last but not least, it should be pointed out that the achievable rates of NRIS-aided algorithms will converge to finite values even if the number of reflecting elements goes to infinity, which is caused by the power budget limitation existing in practical communications.

V. CONCLUSION

This paper jointly considered the joint hardware and software design for the NRIS-empowered THz MIMO communications. Considering the working principle of the NRIS, we primarily probed into the characteristics of the practical graphene-based NRIS. In light of these practical hardware constraints, the A-GD algorithm was developed to settle the passive beamforming design at NRIS and obtained much better performance than the C-GD algorithm. In the near future, our research work will concentrate on the practical measurements for the NRIS-empowered THz MIMO communication system.

REFERENCES

[1] X. You et al., “Towards 6G wireless communication networks: Vision, enabling technologies, and new paradigm shifts,” Sci. China Inf. Sci., vol. 64, no. 1, pp. 1–74, Jan. 2021.
[2] Z. Chen et al., “A survey on terahertz communications,” China Commun., vol. 16, no. 2, pp. 1–35, Feb. 2019.
[3] I. F. Akyildiz, J. M. Jornet, and C. Han, “Terahertz band: next frontier for wireless communications,” Phys. Commun. (Elsevier), vol. 12, no. 4, pp. 16–32, 2014.
[4] I. F. Akyildiz, C. Han, and S. Nie, “Combating the distance problem in the millimeter wave and terahertz frequency bands,” IEEE Commun. Mag., vol. 56, no. 6, pp. 102–108, Jun. 2018.
[5] C. Han and Y. Chen, “Propagation modeling for wireless communications in the terahertz band,” IEEE Commun. Mag., vol. 56, no. 6, pp. 96–101, Jun. 2018.
[6] C. Huang, A. Zappone, G. C. Alexandropoulos, M. Debbah, and C. Yuen, “Reconfigurable intelligent surfaces for energy efficiency in wireless communication,” *IEEE Trans. Wireless Commun.*, vol. 18, no. 8, pp. 4157–4170, Aug. 2019.

[7] C. Liaskos, S. Nie, A. Tsiofillidou, A. Pitsillides, S. Ioannidis, and I. Akyildiz, “A new wireless communication paradigm through Software-Controlled metasurfaces,” *IEEE Commun. Mag.*, vol. 56, no. 9, pp. 162–169, Sep. 2018.

[8] L. Dai et al., “Reconfigurable intelligent surface-based wireless communications: Antenna design, prototyping, and experimental results,” *IEEE Access*, vol. 8, pp. 45913–45923, 2020.

[9] C. Huang et al., “Holographic MIMO surfaces for 6G wireless networks: Opportunities, challenges, and trends,” *IEEE Wireless Commun.*, vol. 27, no. 5, pp. 118–125, Oct. 2020.

[10] F. Gao, T. Cui, and A. Nallanathan, “On channel estimation and optimal training design for amplify and forward relay networks,” *IEEE Trans. Wireless Commun.*, vol. 7, no. 5, pp. 1907–1916, May 2008.

[11] X. Ma et al., “Joint channel estimation and data rate maximization for intelligent reflecting surface assisted terahertz MIMO communication systems,” *IEEE Access*, vol. 8, pp. 99565–99581, 2020.

[12] Q. Wu and R. Zhang, “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.

[13] W. Mei and R. Zhang, “Multi-beam multi-hop routing for intelligent reflecting surfaces aided massive MIMO,” *IEEE Trans. Wireless Commun.*, vol. 21, no. 3, pp. 1897–1912, Mar. 2022.

[14] W. Mei and R. Zhang, “Cooperative beam routing for Multi-IRS aided communication,” *IEEE Wireless Commun. Lett.*, vol. 10, no. 2, pp. 426–430, Feb. 2021.

[15] C. Huang et al., “Multi-hop RIS-Empowered terahertz communications: A DRL-Based hybrid beamforming design,” *IEEE J. Sel. Areas Commun.*, vol. 39, no. 6, pp. 1663–1677, Jun. 2021.

[16] C. Huang, G. C. Alexandropoulos, A. Zappone, M. Debbah, and C. Yuen, “Energy efficient multi-user MISO communication using low resolution large intelligent surfaces,” in *Proc. IEEE GLOBECOM Workshops*, 2018, pp. 1–6.

[17] Q. Wu and R. Zhang, “Intelligent reflecting surface enhanced wireless network: Joint active and passive beamforming design,” in *Proc. IEEE GLOBECOM*, 2018, pp. 1–6.

[18] X. Ma et al., “Intelligent reflecting surface enhanced indoor terahertz communication systems,” *Nano Commun. Netw.*, vol. 24, May 2020, Art. no. 100284.

[19] C. Pan et al., “Multicell MIMO communications relying on intelligent reflecting surfaces,” *IEEE Trans. Wireless Commun.*, vol. 19, no. 8, pp. 5218–5233, Aug. 2020.

[20] H. Shen, W. Xu, S. Gong, Z. He, and C. Zhao, “Secrecy rate maximization for intelligent reflecting surface assisted multi-antenna communications,” *IEEE Commun. Lett.*, vol. 23, no. 9, pp. 1488–1492, Sep. 2019.

[21] L. Liang et al., “Anomalous terahertz reflection and scattering by flexible and conformal coding metamaterials,” *Adv. Opt. Mater.*, vol. 3, no. 10, pp. 1311–1315, 2015.

[22] S. Liu et al., “Frequency dependent dual functional coding metasurfaces at terahertz frequencies,” *Adv. Opt. Mater.*, vol. 4, no. 12, pp. 1965–1973, 2016.

[23] Z. Chen, X. Ma, C. Han, and Q. Wen, “Towards intelligent reflecting surface empowered 6G terahertz communications: A survey,” *China Commun.*, vol. 18, no. 5, pp. 93–119, May 2021.

[24] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, “Capacity limits of MIMO channels,” *IEEE J. Sel. Areas Commun.*, vol. 21, no. 5, pp. 684–702, Jun. 2003.

[25] Z. Wang, L. Liu, and S. Cui, “A reconfigurable polarization converter based on active metasurface,” in *Proc. Cross Strait Quad-Regional Radio Sci. Wireless Technol. Conf.*, 2018, pp. 1–2.

[26] L. Li et al., “Electromagnetic reprogrammable codingmetasurface holograms,” *Nat. Commun.*, vol. 8, no. 1, Aug. 2017, Art. no. 197.

[27] H. Wong and D. Akinwande, *Carbon Nanotube and Graphene Device Physics*, Cambridge, U.K.: Cambridge Univ. Press, 2010.

[28] D. K. Efetov and P. Kim, “Controlling electron-phonon interactions in graphene at ultrahigh carrier densities,” *Phys. Rev. Lett.*, vol. 105, no. 25, Dec. 2010, Art. no. 256805.

[29] S. Lee et al., “Switching terahertz waves with gate-controlled active graphene metamaterials,” *Nat. Mater.*, vol. 11, pp. 936–941, Sep. 2012.

[30] X. Yang et al., “Dielectric meta-reflectarray for broadband linear polarization conversion and optical vortex generation,” *Nano Lett.*, vol. 14, no. 3, pp. 1394–1399, Feb. 2014.

[31] A. Pors and S. I. Bozhevolnyi, “Plasmonic metasurfaces for efficient phase control in reflection,” *Opt. Exp.*, vol. 21, no. 22, pp. 27438–27451, Nov. 2013.

[32] Z. Li et al., “Graphene plasmonic metasurfaces to steer infrared light,” *Sci. Rep.*, vol. 5, Jul. 2015, Art. no. 12423.

[33] B. Di, H. Zhang, L. Song, Y. Li, Z. Han, and H. V. Poor, “Hybrid beamforming for reconfigurable intelligent surface based multi-user communications: Achievable rates with limited discrete phase shifts,” *IEEE J. Sel. Areas Commun.*, vol. 38, no. 8, pp. 1809–1822, Aug. 2020.

[34] X. Ma et al., “Channel estimation for intelligent reflecting surface enabled terahertz MIMO systems,” in *Proc. IEEE Int. Conf. Commun. Workshops*, 2020, pp. 1–6.

[35] A. A. M. Saleh and R. Valenzuela, “A statistical model for indoor multipath propagation,” *IEEE J. Sel. Areas Commun.*, vol. 5, no. 2, pp. 128–137, Feb. 1987.

[36] C. Han, A. O. Bicen, and I. F. Akyildiz, “Multi-ray channel modeling and wideband characterization for wireless communications in the terahertz band,” *IEEE Trans. Wireless Commun.*, vol. 14, no. 5, pp. 2402–2412, May 2015.

[37] W. Tang et al., “Wireless communications with reconfigurable intelligent surface: Path loss modeling and experimental measurement,” *IEEE Trans. Wireless Commun.*, vol. 20, no. 1, pp. 421–439, Jan. 2021.

[38] Z. Wang, L. Liu, and S. Cui, “Channel estimation for intelligent reflecting surface assisted multiuser communications: Framework, algorithms, and analysis,” *IEEE Trans. Wireless Commun.*, vol. 19, no. 10, pp. 6607–6620, Oct. 2020.