Capacity Enhancement for Irregular Reconfigurable Intelligent Surface-Aided Wireless Communications

Ruochen Su*, Linglong Dai*, Jingbo Tan*, Mo Hao†, and Richard MacKenzie‡
*Beijing National Research Center for Information Science and Technology (BNRist),
Department of Electronic Engineering, Tsinghua University, Beijing 100084, China
†Tsinghua SEM Advanced ICT LAB, Tsinghua University, Beijing 100084, China
‡BT Technology, Adastral Park, Ipswich IP5 3RE, U.K.
Email: src18@mails.tsinghua.edu.cn

GLOBECOM 2020 - 2020 IEEE Global Communications Conference | 978-1-7281-8298-8/20/$31.00 ©2020 IEEE | DOI: 10.1109/GLOBECOM42002.2020.9348273

***Abstract*—Reconfigurable intelligent surface (RIS) is an emerging technology to improve the spectral efficiency of wireless communication systems. However, the increase of RIS elements results in the non-negligible overhead of channel estimation and channel feedback, as well as the high complexity of beam design. Therefore, how to improve the system capacity with a limited number of RIS elements becomes a challenge. In this paper, we propose a brand new irregular RIS structure to enhance the capacity of RIS-aided wireless communications. The key idea is to irregularly configure a given number of RIS elements on an enlarged surface, which provides extra degrees of freedom and spatial diversity compared with the classical regular RIS. For the proposed irregular RIS-aided communication, we then formulate the joint topology and beamforming design problem to maximize the system capacity. Accordingly, we propose a joint optimization framework with low complexity to alternately optimize the RIS topology and the corresponding beamforming design. Finally, simulation results demonstrate that the proposed irregular RIS with a limited number of RIS elements can significantly enhance the system capacity compared with the traditional regular RIS.**

I. Introduction

Conventional wireless environment between the base station (BS) and the user ends (UEs) is generally regarded as uncontrollable. However, the emerging reconfigurable intelligent surface (RIS), i.e., a two-dimensional electromagnetic metasurface, is able to control the propagation of incident signals via the interaction between electromagnetic waves and materials coated on the surface [1]. Thus the development of RIS makes it possible to manipulate the wireless environment by tuning the reflection coefficients of a large number of RIS elements individually, which effectively improves the system capacity and reduces the communication outage [2], [3]. In view of the potential to improve the spectral efficiency, RIS is becoming a promising technology in beyond 5G and 6G wireless communications [4].

In RIS-aided wireless communications, beamforming design is essential to realize the potential benefits of RIS, which aims at jointly optimizing the precoding at the BS and the reflection coefficients at the RIS. Prior works on beamforming optimization have only considered the regular RIS structure. Specifically, the optimal beamforming design for multiple RISs was derived in [5], which aimed to maximize the received signal power of a single-antenna user. The analysis of the energy efficiency was focused on by joint transmit power allocation and beamforming optimization for multi-user cases in [6]. Several joint beamforming algorithms were also proposed for different optimization objectives, for instance, to maximize the weighted sum-rate (WSR) or to minimize the transmit power [7], [8]. Furthermore, [1] has revealed that the received signal power of a regular RIS-aided wireless communication system is quadratic proportional to the number of RIS elements. In consideration of the performance bound based on regular arrays, the system capacity can only be improved by increasing the number of RIS elements. However, the overhead of channel estimation and channel feedback, as well as the complexity of beam design will be prohibitively high with a large number of RIS elements [1], [9]. In view of this, the number of RIS elements in practical systems cannot be too large, which limits the system capacity. Therefore, it is vital to improve the capacity of RIS-aided communication systems with a limited number of RIS elements.

In this paper, inspired by the advantages of signal enhancement and interference suppression by irregular configuration for phased arrays [10], we propose an irregular RIS structure to enhance the received signal and thus improve the capacity of RIS-aided communications. To the best of our knowledge, the irregular RIS has not been investigated in the literature, which is studied in this paper for the first time. Specifically, we propose a brand new topologic structure of RIS based on irregular arrays, where a given number of RIS elements are irregularly arranged on an enlarged surface. The selection of feasible locations for RIS elements leads to additional degrees of freedom (DoFs) and spatial diversity of the RIS, which thus enhances the system capacity with a limited number of RIS elements. Furthermore, we analyze the model of proposed irregular RIS-aided communications, and formulate the joint topology and beamforming design problem to improve the WSR. Accordingly, we propose a joint optimization framework to alternately optimize the RIS topology and the beamforming design. Simulation results[1] show that the proposed irregular RIS can significantly improve the capacity of RIS-aided communication systems.

[1]Simulation codes are provided to reproduce the results presented in this paper: http://oa.ee.tsinghua.edu.cn/dailinglong/publications/publications.html.

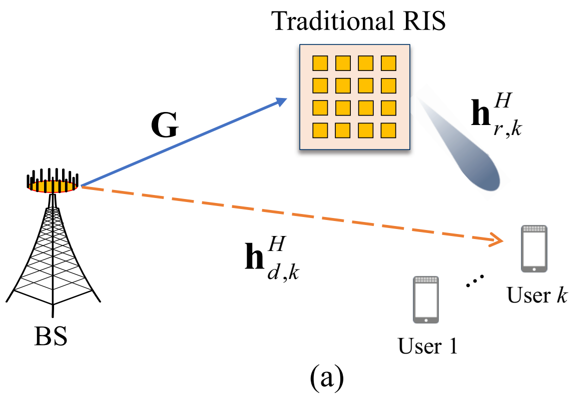

Fig. 1. The RIS-aided communication system: (a) Traditional regular RIS; (b) Proposed irregular RIS.

The rest of the paper is organized as follows. Section II introduces the system model and the problem formulation for proposed irregular RIS-aided communications. In Section III, we propose a joint optimization framework to solve the formulated optimization problem. The simulation results are shown in Section IV, and the conclusions are finally drawn in Section V.

Notations: Vectors and matrices are denoted by lower-case and upper-case boldface letters, respectively. $\mathbf{A}^H$, $\mathbf{A}^T$, and $\mathbf{A}^{-1}$ denote the conjugate transpose, transpose, and inverse of matrix $\mathbf{A}$, respectively. $\|\mathbf{a}\|_1$ and $\|\mathbf{a}\|_2$ denote the ℓ_1-norm and Euclidean norm of vector $\mathbf{a}$, respectively. $\mathrm{diag}(\mathbf{a})$ denotes a diagonal matrix whose diagonal elements consist of corresponding entries in vector $\mathbf{a}$. $\mathbf{I}_K$ denotes the identity matrix of size $K \times K$. Finally, $\mathbf{1}_{K\times L}$ denotes the all-one matrix of size $K \times L$.

II. System Model and Problem Formulation of The Proposed Irregular RIS-Aided Communications

In this section, we first introduce the concept of the proposed irregular RIS and analyze the system model. Then, the WSR maximization problem of irregular RIS-aided communications is illuminated.

A. System model

We consider an RIS-aided multi-user multiple-input multiple-output (MIMO) downlink wireless communication system. The traditional regular RIS-aided communication system is shown in Fig. 1 (a), where a BS equipped with M antennas and a regular RIS comprising N elements simultaneously serve K single-antenna users. Different from the regular RIS whose elements are arranged on a regular surface with constant interelement spacing, we propose a brand new concept of irregular RIS. To simplify discussion while verifying the effectiveness of the proposed concept, the grid constraint with fixed grid spacing is considered, where N reflecting elements are sparsely distributed over N_s grid points of an enlarged surface, as shown in Fig. 1 (b). Without loss of generality, the spacing between adjacent grid points is assumed half of the signal wavelength [10]. Note that the proposed irregular RIS can be realized in many ways. For example, this can be equivalently implemented by selecting a subset of RIS elements from a large set, which is similar to the antenna selection technology for phased arrays adopted in IEEE 802.11n [11].

On account of the severe propagation loss [12], we only consider the signal reflected by the irregular RIS for the first time and ignore multiple reflections. Let $\mathbf{Z} = \mathrm{diag}(\mathbf{z})$ denote the selection matrix representing the topology of RIS, where $\mathbf{z} = [z_1, z_2, \cdots, z_{N_s}]^T$. We have $z_i \in \{1, 0\}$ with "1" standing for selection and "0" on behalf of the opposite, corresponding to whether an RIS element is deployed at the ith grid point or not, respectively. Thus the received signal $\mathbf{y} \in \mathbb{C}^{K\times 1}$ for all K users can be expressed as

$$\mathbf{y} = \left(\mathbf{H}_r^H \mathbf{Z}\boldsymbol{\Theta}\mathbf{G} + \mathbf{H}_d^H\right)\mathbf{x} + \mathbf{n}, \tag{1}$$

where $\mathbf{x} \in \mathbb{C}^{M\times 1}$ represents the transmitted signal at the BS, each component of $\mathbf{n} \in \mathbb{C}^{K\times 1}$ donates the additive white Gaussian noise (AWGN) with zero mean and variance σ^2. $\boldsymbol{\Theta} = \mathrm{diag}([\beta_1 e^{j\theta_1}, \beta_2 e^{j\theta_2}, \cdots, \beta_{N_s} e^{j\theta_{N_s}}])$ represents the reflection coefficients of N_s grid points of the irregular RIS. $\mathbf{G} \in \mathbb{C}^{N_s\times M}$ denotes the BS-RIS channel. Besides, we have $\mathbf{H}_d^H = [\mathbf{h}_{d,1}, \mathbf{h}_{d,2}, \cdots, \mathbf{h}_{d,K}]^H \in \mathbb{C}^{K\times M}$, $\mathbf{H}_r^H = [\mathbf{h}_{r,1}, \mathbf{h}_{r,2}, \cdots, \mathbf{h}_{r,K}]^H \in \mathbb{C}^{K\times N_s}$, where $\mathbf{h}_{d,k}^H$ and $\mathbf{h}_{r,k}^H$ represent the channels from the BS to user k, and the RIS to user k, respectively.

In this paper, we consider the fully digital precoding at the BS with $\mathbf{x} = \sum_{k=1}^K \mathbf{w}_k s_k$, where $\mathbf{w}_k \in \mathbb{C}^{M\times 1}$ denotes the precoding vector for user k, and $\mathbf{s} = [s_1, s_1, \cdots, s_K]^H \in \mathbb{C}^{K\times 1}$ denotes the transmitted symbol vector satisfying $E[\mathbf{s}\mathbf{s}^H] = \mathbf{I}_K$. Considering practical hardware implementation, constant reflection amplitude constraints and finite discrete phase shifts constraints at the RIS are assumed [13]. To this end, let $\beta_n = 1$, θ_n take discrete values from the set $\mathcal{F} = \{0, \frac{2\pi}{2^b}, \cdots, \frac{2\pi}{2^b}(2^b - 1)\}$ for $\forall n = 1, 2, \cdots, N_s$, where b is the number of quantized bits of finite discrete phase shifts.

Thus the signal-to-interference-plus-noise ratio (SINR) of user k is

$$\gamma_k = \frac{\left|\left(\mathbf{h}_{r,k}^H \mathbf{Z}\boldsymbol{\Theta}\mathbf{G} + \mathbf{h}_{d,k}^H\right)\mathbf{w}_k\right|^2}{\sum_{i\neq k}^{K}\left|\left(\mathbf{h}_{r,k}^H \mathbf{Z}\boldsymbol{\Theta}\mathbf{G} + \mathbf{h}_{d,k}^H\right)\mathbf{w}_i\right|^2 + \sigma^2}. \tag{2}$$

B. Problem formulation

Let $\mathbf{W} = [\mathbf{w}_1, \mathbf{w}_2, \cdots, \mathbf{w}_K]$ denote the digital precoding matrix at the BS, P_T denote the maximum transmit power at the BS, ω_k denote the weight of user k. The WSR maximization problem subject to the transmit power constraint in (3b), discrete phase shifts constraint in (3c), and sparsity constraints in (3d) as well as (3e) can be formulated as

$$\begin{aligned}
\mathcal{P}_1: \quad & \max_{\mathbf{Z},\mathbf{W},\boldsymbol{\Theta}} \; R = \sum_{k=1}^{K} \omega_k \log_2(1+\gamma_k) && \text{(3a)}\\
\text{s.t.} \quad & C_1: \sum_{k=1}^{K} \|\mathbf{w}_k\|_2^2 \le P_\mathrm{T}, && \text{(3b)}\\
& C_2: \theta_n \in \mathcal{F}, \forall n = 1, 2, \cdots, N_\mathrm{s}, && \text{(3c)}\\
& C_3: z_i(z_i - 1) = 0, \forall i = 1, 2, \cdots, N_\mathrm{s}, && \text{(3d)}\\
& C_4: \mathbf{1}^T\mathbf{z} = N. && \text{(3e)}
\end{aligned}$$

C_3 and C_4 restrict the sparse deployment of the irregular RIS, where N diagonal elements of the topology matrix $\mathbf{Z}$ are assigned the value of 1, while the rest of which are assigned 0. Note that the objective function in (3a) and the phase shifts constraint C_2 are non-convex. Besides, the sparsity constraints C_3 and C_4 are also non-convex, thus making it more difficult to be solved than the traditional optimization problem in regular RIS-based systems. One possible solution to tackle this problem is to decouple the RIS deployment and the corresponding beamforming optimization. Specifically, for a given topology $\mathbf{Z}_0$, $\mathcal{P}_1$ is reduced to

$$\begin{aligned}
\mathcal{P}_2: \quad & \max_{\mathbf{W},\boldsymbol{\Theta}} \; R = \sum_{k=1}^{K} \omega_k \log_2(1+\gamma_k) && \text{(4a)}\\
\text{s.t.} \quad & C_1: \sum_{k=1}^{K} \|\mathbf{w}_k\|_2^2 \le P_\mathrm{T}, && \text{(4b)}\\
& C_2: \theta_n \in \mathcal{F}, \forall n = 1, 2, \cdots, N_s, && \text{(4c)}\\
& C_3: \mathbf{Z} = \mathbf{Z}_0. && \text{(4d)}
\end{aligned}$$

Note that the complete channel state information (CSI) relevant to the number of grid points of RIS is required to solve $\mathcal{P}_1$, which results in the high overhead and complexity. Fortunately, considering that the BS-RIS channel is quasi-static [9], the topology of RIS only needs to be changed adaptively in a large timescale, corresponding to solving $\mathcal{P}_1$. While the beamforming design for a specific RIS topology, i.e., the solution to $\mathcal{P}_2$, can be frequently optimized in a small timescale to track the change of users. Since the required CSI for a fixed RIS topology is relevant to the number of RIS elements, the resultant overhead and complexity are much lower than that of a large-scale system with excessive RIS elements.

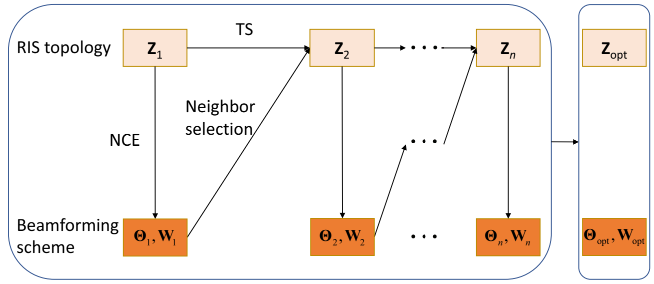

Fig. 2. The proposed joint optimization framework.

Moreover, the proposed model is equivalent to the system model of regular RIS-aided wireless communications by simply assuming that $N_s = N$ and $\mathbf{Z} = \mathbf{I}_N$. Thus the problem in $\mathcal{P}_1$ can be seen as a general formulation based on various topologies of RIS. Therefore, the proposed solution in Section III can serve as a general solution to the sum-rate optimization problem in classical RIS-aided communications as well.

III. The Proposed Joint Optimization Framework

In this section, we propose a joint optimization framework with low complexity to solve $\mathcal{P}_1$. Specifically, we propose a sub-optimal algorithm to carry out the sparse deployment of RIS elements. Then, given a specific RIS topology, an alternating optimization algorithm is proposed to jointly optimize the precoding at the BS and the phase shifts at the RIS.

A. Overview of the proposed joint optimization framework

As shown in Fig. 2, we first decouple the decision variables in $\mathcal{P}_1$. Then, we alternately optimize the RIS topology and the beamforming design. Instead of traversing all possible topologic structures of RIS, we search for the sub-optimal deployment of RIS iteratively by a tabu search (TS) method. In each iteration, the problem $\mathcal{P}_1$ is reduced to $\mathcal{P}_2$, where the RIS topology $\mathbf{Z}$ is fixed. Specifically, we generate the neighbors of $\mathbf{Z}$ and obtain corresponding beamforming design by a neighbor extraction-based cross-entropy (NCE) method with zero-forcing (ZF) precoding at the BS. By comparing the WSR, the optimal neighbor is selected for the next iteration. After maximum iterations or termination conditions are reached, we obtain the sub-optimal solution to $\mathcal{P}_1$, denoted by $\mathbf{Z}_{\mathrm{opt}}$, $\boldsymbol{\Theta}_{\mathrm{opt}}$, and $\mathbf{W}_{\mathrm{opt}}$. The details of sparse deployment and beamforming optimization of the irregular RIS are provided in Subsection III-B and III-C, respectively.

B. TS-based sparse deployment of RIS

Inspired by the Turbo-TS beamforming design in [14], we propose a TS-based algorithm to implement the sparse deployment of RIS elements.

Firstly we initialize the RIS topology $\mathbf{Z}_1$ by randomly selecting N ones and $N_s - N$ zeros. For a given $\mathbf{Z}$, the WSR R can be figured out based on ZF precoding at the

BS and discrete phase shifts adjustment at the RIS, which will be further discussed in the next subsection. In the ith iteration, we focus on the neighbors of $\mathbf{Z}_i$ generated by a redefined move criterion, where we randomly swap p ones for p zeros among the diagonal elements of $\mathbf{Z}_i$. Here p is defined as the neighbor distance, which will be dynamically adjusted on the basis of the value of i. We can choose a large p at the beginning of iterations for a wide searching range. As the iteration progresses, we decrease the neighbor distance for fine tuning. Next, the obtained neighbors should be checked in the tabu list, and those who are tabu solutions will be discarded. In this way we obtain Q neighbors and separately calculate the WSR for all the candidates $\{\mathbf{Z}_i^q\}_{q=1}^Q$. Then we select the candidate with the maximum WSR, which is saved as the new topology $\mathbf{Z}_{i+1}$ for the next iteration. The global optimum and the tabu list are updated accordingly. Specifically, the feasible solution $\mathbf{Z}_i$ is added to the tabu list to avoid cyclic search, and the earliest tabu solution in the current list is removed. After the iteration threshold $I_{\rm T}$ is reached, we can obtain the sub-optimal topology of RIS.

C. NCE-based beamforming optimization

The TS-based sparse deployment algorithm requires joint beamforming design at the BS and the RIS for a specific RIS topology, corresponding to solving $\mathcal{P}_2$. In [15], cross-entropy (CE) algorithm is developed to tackle combinatorial optimization problems with low complexity. Inspired by this idea, we further propose a NCE algorithm to solve $\mathcal{P}_2$, which utilizes neighbor extraction of the current optimal solution during the iterative process.

Let $\mathbf{P} = [\mathbf{p}_1, \mathbf{p}_2, \cdots, \mathbf{p}_{N_s}]$ denote the probability matrix of $\mathbf{\Theta}$ for a given $\mathbf{Z}$, where $\mathbf{p}_n \in \mathbb{R}^{2^b \times 1}$ is the probability vector of θ_n satisfying $\|\mathbf{p}_n\|_1 = 1$. Each component of $\mathbf{p}_n$ denotes the probability of taking different values in $\mathcal{F}$. Firstly we initialize $\mathbf{P}^{(1)} = \frac{1}{2^b} \times \mathbf{1}_{2^b \times N_s}$, which represents that the value of θ_n is selected from the elements in $\mathcal{F}$ with the same probability. In the ith iteration, we randomly generate C candidates $\{\mathbf{\Theta}^c\}_{c=1}^C$ based on the probability distribution function (PDF) $\Xi(\mathbf{\Theta}; \mathbf{P}^{(i)})$. Let $\delta(t)$ be 1 when $t = 0$, otherwise 0. Let $\mathcal{F}(k)$ denote the kth element of $\mathcal{F}$. Thus the PDF can be expressed as

$$\Xi(\mathbf{\Theta}; \mathbf{P}^{(i)}) = \prod_{n=1}^{N_s} \left(\prod_{k=1}^{2^b} (p_{n,k}^{(i)})^{\delta(\theta_n - \mathcal{F}(k))} \right). \tag{5}$$

Then we are able to calculate the effective channel $\mathbf{H}_{\rm eq}^c = \mathbf{H}_r^H \mathbf{Z} \mathbf{\Theta}^c \mathbf{G} + \mathbf{H}_d^H$ for all candidates $\{\mathbf{\Theta}^c\}_{c=1}^C$. The precoding matrices $\{\mathbf{W}^c\}_{c=1}^C$ at the BS are acquired by the ZF precoder expressed as [8]

$$\mathbf{W} = \mathbf{H}_{\rm eq}^H (\mathbf{H}_{\rm eq} \mathbf{H}_{\rm eq}^H)^{-1} \mathbf{P}_{\rm B}^{\frac{1}{2}}, \tag{6}$$

where $\mathbf{P}_{\rm B}$ represents the power allocation at the BS. The WSR $\{R(\mathbf{\Theta}^c)\}_{c=1}^C$ are acquired according to (3a). Next, we sort $\{R(\mathbf{\Theta}^c)\}_{c=1}^C$ in a descending order as $R(\mathbf{\Theta}^{[1]}) \geq R(\mathbf{\Theta}^{[2]}) \geq \cdots \geq R(\mathbf{\Theta}^{[C]})$ and select $C_{\rm pr}$ optimal candidates as primary elites. Then we propose a neighbor extraction method to expand the searching range through changing one of the diagonal elements of the current optimal solution $\mathbf{\Theta}^{[1]}$. Only N diagonal elements are effective, whose values can be selected from $\mathcal{F}$. By this means we obtain $N(2^b - 1)$ extra candidates and calculate the corresponding precoders at the BS as well as the WSR. The candidate whose WSR is larger than $R(\mathbf{\Theta}^{[1]})$ is selected as a supplementary elite. The total number of elites is updated as $C_{\rm elite} = C_{\rm pr} + C_{\rm sup}$, where $C_{\rm sup}$ is the number of supplementary elites. After that, we reorganize the elites as $\{\mathbf{\Theta}^{(c)}\}_{c=1}^{C_{\rm elite}}$, and the probability transfer criterion is expressed as

$$\mathbf{P}^{(i+1)} = \arg\max_{\mathbf{P}^{(i)}} \frac{1}{C_{\rm elite}} \sum_{c=1}^{C_{\rm elite}} \eta_c \ln \Xi\left(\mathbf{\Theta}^{(c)}; \mathbf{P}^{(i)}\right), \tag{7}$$

where $\eta_c = \frac{R(\mathbf{\Theta}^{(c)}) C_{\rm elite}}{\sum_{c=1}^{C_{\rm elite}} R(\mathbf{\Theta}^{(c)})}$ is defined as the weight of the cth elite, which represents the ratio of the WSR of elite c to the average WSR of all elites. Note that the larger WSR corresponds to the larger weight. The updated probability matrix $\mathbf{P}^{(i+1)}$ is employed in the next loop until the iteration threshold $I_{\rm N}$ is reached. Finally, the sub-optimal beamforming design can be obtained.

IV. Simulations

In this section, we provide simulation results to evaluate the performance of the proposed irregular RIS-based systems by employing the joint optimization framework.

In our simulation scenario, K single-antenna users are served by a BS equipped with M antennas and an irregular RIS equipped with N elements which are distributed over a rectangular surface with N_s grid points. We set the distance between the BS and UEs, the RIS and UEs as $d_{\rm BU} = 50$ m, $d_{\rm RU} = 2$ m, respectively. The distance between the BS and RIS is set as $d_{\rm BR} = \sqrt{d_{\rm BU}^2 + d_{\rm RU}^2}$. This corresponds to a hot spot scenario in practice where RIS plays a vital role [8]. The parameters of the proposed iteration algorithms are set as $Q = 15$, $I_{\rm T} = 40$, $I_{\rm N} = 15$, $C = 200$, and $C_{\rm pr} = 40$. The neighbor distance p is set as 3 at the beginning and is reduced to 2 as the iteration progresses. Due to the excessive possible cases of RIS topologies, the size of the tabu list can be simply set as 1 to expand the searching range, while with a very low probability of falling into a loop. Other parameters are set as $b = 1$, $\sigma^2 = -80$ dBm, and $\omega_k = 1, \forall k = 1, 2, \cdots, K$.

The uncorrelated Rayleigh fading channel model is adopted to account for the small-scale fading. The large-scale fading, i.e. the distance-dependent pass loss, is considered as well. Specifically, the path loss of the BS-RIS-UE channel can be expressed as [12], [16]

$$f_r(d_{\rm BR}, d_{\rm RU}) = C_{\rm r} d_{\rm BR}^{-\alpha_{\rm BR}} d_{\rm RU}^{-\alpha_{\rm RU}}, \tag{8}$$

where $d_{\rm BR}$ and $d_{\rm RU}$ are the distance between the BS and RIS, the RIS and UEs, respectively. $C_{\rm r}$ denotes the effect of channel fading and antenna gain. $\alpha_{\rm BR}$ and $\alpha_{\rm RU}$ denote the

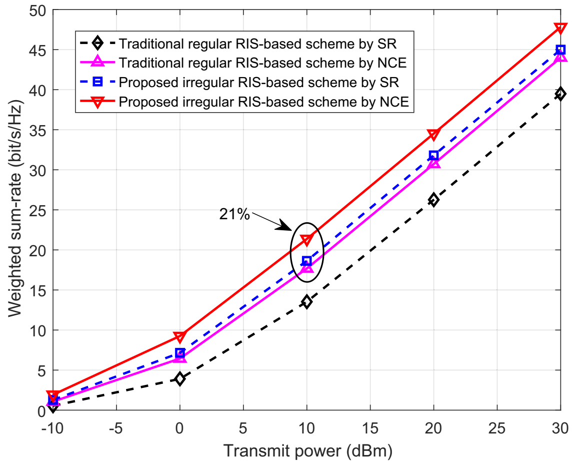

Fig. 3. WSR versus the transmit power. $M = 4$, $N = 32$, $N_s = 64$, $K = 4$.

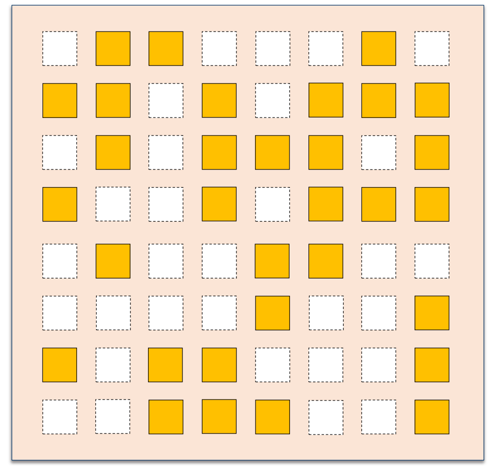

Fig. 4. The optimized RIS topology. $N = 32$, $N_s = 64$.

path loss exponents. Similarly, the path loss of the BS-UE channel is expressed as $f_d(d_{\text{BU}}) = C_{\text{d}} d_{\text{BU}}^{-\alpha_{\text{BU}}}$. The parameters are set as $C_{\text{r}} = -60$ dB, $C_{\text{d}} = -30$ dB, $\alpha_{\text{BR}} = 2$, $\alpha_{\text{RU}} = 2$, and $\alpha_{\text{BU}} = 3.5$ [7].

To validate the superiority of the proposed irregular RIS structure and joint optimization framework, we assume the CSI is perfectly known. We consider a scenario with $M = 4$, $N = 32$, $N_s = 64$, and $K = 4$. Since the complexity of the optimal solution based on the exhaustive search method is unacceptable for such large system size, we provide the sub-optimal solution by the proposed joint optimization framework. To the best of our knowledge, the irregular RIS-based scheme has not been investigated in the literature. Thus the beamforming optimization for traditional regular RIS-aided communications is considered as the benchmark. Besides, the successive refinement (SR) algorithm proposed in [8] is adopted to optimize the phase shifts of RIS for comparison with the NCE-based beamforming algorithm proposed in this paper.

The WSR versus the transmit power is shown in Fig. 3. By comparing the results of the traditional regular scheme and the irregular scheme, one can observe that the proposed irregular RIS-based scheme with a limited number of elements substantially enhances the system capacity. For example, provided the NCE algorithm is adopted, the WSR of the proposed irregular scheme increases by 21% compared with that of the traditional regular scheme at the transmit power of 10 dBm. Note that the capacity enhancement does not depend on a specific beamforming optimization algorithm, and the simpler SR algorithm also works. This is for the reason that the selection of flexible locations for RIS elements leads to additional DoFs at the cost of space sacrificing, which enables us to select a N-element subset with the optimal channel conditions. Therefore, the sparse configuration of the irregular RIS employing N elements can achieve the full spatial diversity benefits of N_s grid points, which enhances the received signal and thus improves the system capacity via the well-designed topology and beamforming design. For instance, we provide an example of the optimized RIS topology in Fig. 4, where the colored squares represent the selected locations for RIS elements. In addition, the proposed NCE algorithm outperforms the SR algorithm for both cases of the traditional regular scheme and the irregular scheme, which confirms the effectiveness of our proposed optimization framework.

To further show the effect of different sparse ratios of RIS, we consider an irregular RIS with a fixed number of elements, where $M = 4$, $N = 20$, $K = 4$. While the size of the irregular RIS is variable, which is represented by N_s, i.e., the number of grid points of the surface. The simulation results are provided based on the proposed joint optimization framework with the transmit power set as $P_{\text{T}} = 10$ dBm. The traditional regular RIS-based scheme with $M = 4$, $N = 20$, and $K = 4$ serves as a benchmark, whose surface size is fixed. The WSR versus the size of the irregular RIS is shown in Fig. 5. It is observed that altering the sparse ratio of the irregular RIS via enlarging the surface size effectively improves the WSR performance, where the number of RIS elements is assumed to be a constant. Moreover, we consider several traditional regular RIS-based scenarios where more antennas are deployed at the BS, namely, $M = 6$, $M = 7$, and $M = 8$. The number of RIS elements and the number of users remain the same. It is revealed that enlarging the size of the irregular RIS to $N_s = 40$, $N_s = 60$, and $N_s = 80$ outperforms the traditional regular schemes with $M = 6$, $M = 7$, and $M = 8$, respectively. Note that the irregular scheme when the sparse ratio is 25%, i.e., $N = 20$ and $N_s = 80$, saves half of the number of antennas at the BS. Thus the sparse configuration of the irregular RIS can provide a feasible solution to improve the system capacity without increasing antennas and RF chains at the BS, whose hardware cost and power comsumption are usually high. Nevertheless, with the increasing surface size of the irregular RIS, the growth of performance slows down, while the complexity of RIS increases. Thus we should make

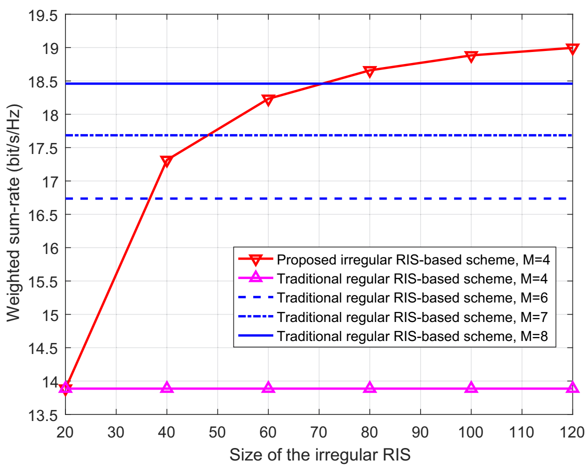

Fig. 5. WSR versus the size of the irregular RIS. $P_{\mathrm{T}} = 10$ dBm, $N = 20$, $K = 4$.

a tradeoff between the cost and the performance by carefully designing the sparse ratio of RIS in practice.

V. Conclusions

Considering the high overhead and complexity caused by the large number of RIS elements in traditional regular RIS-aided wireless communications, we investigate the design paradigm of irregular RIS in this paper for the first time. Firstly, we propose an irregular RIS structure with a given number of elements distributed over an enlarged surface. Then, for the proposed irregular RIS-aided communication, we formulate the WSR maximization problem to optimize the system capacity. Finally, a joint optimization framework is proposed to alternately solve the optimization problem. Simulation results validate that with a limited number of RIS elements, the proposed irregular RIS significantly enhances the system capacity compared to the traditional regular RIS. In addition, we can obtain better performance by further optimizing the sparse ratio of the irregular RIS.

References

[1] E. Basar, M. Di Renzo, J. De Rosny, M. Debbah, M.-S. Alouini, and R. Zhang, "Wireless communications through reconfigurable intelligent surfaces," *IEEE Access*, vol. 7, pp. 116 753–116 773, Aug. 2019.

[2] M. D. Renzo, K. Ntontin, J. Song, F. H. Danufane, X. Qian, F. Lazarakis, J. D. Rosny, D. Phan-Huy, O. Simeone, R. Zhang, M. Debbah, G. Lerosey, M. Fink, S. Tretyakov, and S. Shamai, "Reconfigurable intelligent surfaces vs. relaying: Differences, similarities, and performance comparison," *IEEE Open J. Commun. Soc.*, vol. 1, pp. 798–807, Jun. 2020.

[3] M. Jung, W. Saad, Y. Jang, G. Kong, and S. Choi, "Reliability analysis of large intelligent surfaces (LISs): Rate distribution and outage probability," *arXiv preprint arXiv:1903.11456v2*, Aug. 2019.

[4] J. Zhao and Y. Liu, "A survey of intelligent reflecting surfaces (IRSs): Towards 6G wireless communication networks," *arXiv preprint arXiv:1907.04789v3*, Nov. 2019.

[5] P. Wang, J. Fang, X. Yuan, Z. Chen, H. Duan, and H. Li, "Intelligent reflecting surface-assisted millimeter wave communications: Joint active and passive precoding design," *arXiv preprint arXiv:1908.10734*, Aug. 2019.

[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] H. Guo, Y.-C. Liang, J. Chen, and E. G. Larsson, "Weighted sum-rate optimization for intelligent reflecting surface enhanced wireless networks," *arXiv preprint arXiv:1905.07920v2*, May 2019.

[8] Q. Wu and R. Zhang, "Beamforming optimization for wireless network aided by intelligent reflecting surface with discrete phase shifts," *IEEE Trans. Commun.*, vol. 68, no. 3, pp. 1838–1851, Mar. 2020.

[9] C. Hu and L. Dai, "Two-timescale channel estimation for reconfigurable intelligent surface aided wireless communications," *arXiv preprint arXiv:1912.07990*, Dec. 2019.

[10] X. Wang, M. Amin, and X. Cao, "Analysis and design of optimum sparse array configurations for adaptive beamforming," *IEEE Trans. Signal Process.*, vol. 66, no. 2, pp. 340–351, Jan. 2018.

[11] N. B. Mehta, S. Kashyap, and A. F. Molisch, "Antenna selection in LTE: From motivation to specification," *IEEE Commun. Mag.*, vol. 50, no. 10, pp. 144–150, Oct. 2012.

[12] Ö. Özdogan, E. Björnson, and E. G. Larsson, "Intelligent reflecting surfaces: Physics, propagation, and pathloss modeling," *IEEE Wireless Commun. Lett.*, vol. 9, no. 5, pp. 581–585, May 2020.

[13] L. Dai, B. Wang, M. Wang, X. Yang, J. Tan, S. Bi, S. Xu, F. Yang, Z. Chen, M. D. Renzo, C. Chae, and L. Hanzo, "Reconfigurable intelligent surface-based wireless communications: Antenna design, prototyping, and experimental results," *IEEE Access*, vol. 8, pp. 45 913–45 923, Mar. 2020.

[14] X. Gao, L. Dai, C. Yuen, and Z. Wang, "Turbo-like beamforming based on tabu search algorithm for millimeter-wave massive MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 65, no. 7, pp. 5731–5737, Jul. 2016.

[15] R. Y. Rubinstein and D. P. Kroese, "The cross-entropy method: A unified approach to combinatorial optimization, Monte-Carlo simulation and machine learning," *Springer Science and Business Media*, 2013.

[16] Z. Zhang and L. Dai, "A joint precoding framework for wideband reconfigurable intelligent surface-aided cell-free network," *arXiv preprint arXiv:2002.03744v2*, Feb. 2020.