

On the Capacity of Reconfigurable Intelligent Surface Assisted MIMO Symbiotic Communications

Jia Ye, *Student Member, IEEE*, Shuaishuai Guo, *Member, IEEE*, Shuping Dang, *Member, IEEE*, Basem Shihada, *Senior Member, IEEE* and Mohamed-Slim Alouini, *Fellow, IEEE*

Abstract—Reconfigurable intelligent surfaces (RISs) appear as one of the most promising paradigms for future wireless communications, because of their high adjustability for diverse communication demands and the additional information-carrying capability by reflecting patterns. This paper investigates the capacity of RIS-assisted multiple-input multiple-output (MIMO) symbiotic communications utilizing multiple reflecting patterns, where each reflecting pattern is non-uniformly activated to carry additional information. To enhance transmission performance, the reflecting patterns, reflecting activation probability, and the transmit covariance matrix are jointly designed. Since the exact expression of the system capacity is intractable, the lower and upper bounds on the capacity are derived and used for optimization in this paper. Based on the lower bound on the capacity, a gradient ascent algorithm is developed to find the optimal reflecting patterns, reflecting activation probability, and the transmit covariance matrix. By taking advantage of the concise-form upper bound on the capacity, closed-form solutions of the reflecting activation probability and transmit covariance matrix can be derived after optimizing the reflecting patterns. The superiority of the proposed design is investigated and verified by computer simulations. Some selected numerical results demonstrate that the proposed design can achieve a higher capacity than the benchmark adopting only one reflecting pattern.

Index Terms—Reconfigurable intelligent surface (RIS), multiple-input multiple-output (MIMO), symbiotic communications, capacity analysis.

I. INTRODUCTION

AS fifth generation (5G) wireless networks are gradually being deployed worldwide, the potential technologies applied in beyond 5G or even sixth generation (6G) networks attract researchers from both academia and industry [1]–[5]. The lifestyle of people becomes increasingly dependent on network connectivity, raising the expectations of transmission capacity, data rate, latency, coverage, and miscellaneous quality of service (QoS) requirements [1], [5]. To handle these

challenges, various promising technologies have been proposed, such as terahertz (THz) communications, various free-space optical networks, and integrated satellite-to-terrestrial systems. Among all prominent candidates, the reconfigurable intelligent surface (RIS) has recently emerged as one of the most promising paradigms to support the beyond 5G and 6G wireless communications [6].

In essence, the RIS is a planar meta-surface composed of a large number of passive and low-cost reflecting elements, such as positive-intrinsic-negative (PIN) diodes, each of which enables an independent phase shift and/or amplitude attenuation to the incident signal [7]. With the support of the RIS, the propagation environment between receiver and transmitter becomes adjustable in a flexible manner for QoS improvements corresponding to different communication scenarios and demands [8]. The signals reflected by a well designed RIS can be added at the receiver with other signals to increase the received signal strength, suppress symbol error, and mitigate co-channel interference [9].

Although the signal forwarding operation at the RIS sounds like a traditional relay, the most distinguishable difference is that the reflecting elements on the RIS are passive and without transmitting/receiving radio frequency (RF) chains [10]. As a result, an RIS operates in a much greener way with orders-of-magnitude lower energy cost/hardware compared to traditional active antenna arrays. Furthermore, due to the passivity, RISs can easily operate in a full-duplex mode and offers a much higher spectral efficiency (SE). As RISs are free of any antenna noise and self-interference, the hardware complexity of passive RISs is much lower than a conventional full-duplex relay, which needs additional techniques and architectures for self-interference cancellation [11]. Except for the aforementioned advantages, RISs are also favorable by researchers and engineers due to the high flexibility and superior compatibility for practical implementations [12]. The lightweight and conformal geometry of RISs enables the transparent integration into existing wireless communication networks without any change in the hardware and software architectures. Therefore, RISs can be easily installed on facades of buildings, objects, ceiling, and walls [13].

A. Prior Work

The promising advantages stated above envision the potential of RIS to significantly enhance wireless transmission in an energy- and cost-efficient way. However, the performance gain brought by RIS can only be realized under the precondition that the RIS reflection coefficients are well designed.

J. Ye, B. Shihada, and M.-S. Alouini are with Computer, Electrical and Mathematical Sciences and Engineering Division, King Abdullah University of Science and Technology (KAUST), Thuwal 23955-6900, Saudi Arabia (e-mail: {jia.ye, basem.shihada, slim.alouini}@kaust.edu.sa).

S. Dang was with Computer, Electrical and Mathematical Science and Engineering Division, King Abdullah University of Science and Technology (KAUST), Thuwal 23955-6900, Saudi Arabia, and is now with Department of Electrical and Electronic Engineering, University of Bristol, Bristol BS8 1UB, UK (e-mail: shuping.dang@bristol.ac.uk).

S. Guo was with Computer, Electrical and Mathematical Sciences and Engineering Division, King Abdullah University of Science and Technology (KAUST), Thuwal 23955-6900, Saudi Arabia. He is now with Shandong Provincial Key Laboratory of Wireless Communication Technologies and School of Control Science and Engineering, Shandong University, Jinan 250061, P.R. China (e-mail: shuaishuai_guo@sdu.edu.cn).

Recently, a great number of researchers from both industry and academia have been attracted by the peculiarities of RIS, who have provided a large and growing body of literature regarding system performance analysis and optimization from the perspectives of received signal power [14], [15], channel estimation [16]–[18], symbol error rate [19], deployment scheme [20], [21], physical layer security [22]–[24], and so on. Apart from these emerging research directions, studying the capacity of RIS assisted communication systems from an information theoretic viewpoint is crucial to unveil the upper bound on the performance gain brought by RISs. However, most existing works analyzing capacity of RIS assisted communication systems focus on single-input single-output (SISO) or multiple-input single-output (MISO) scenarios [25]–[27], while the capacity analysis of RIS assisted multiple-input multiple-output (MIMO) systems is still preliminary and awaits in-depth investigations.

There are a couple of works laying the foundation for the capacity analysis of RIS assisted MIMO systems. The authors in [28] demonstrated the impact of RIS on the channel capacity of an RIS assisted MIMO system, which can increase the channel power gains as a direct result of the improved channel matrix rank. The authors in [29], [30] made valuable contribution for maximizing the capacity of RIS assisted MIMO systems by jointly optimizing the transmit covariance matrix and reflection coefficients. The capacity of a practical RIS assisted MIMO system with transmitted symbols chosen from discrete signal constellation was investigated in [31], where the cut-off rate is adopted as the optimization objective to replace the actual discrete input system capacity that is not analytically tractable. The capacity maximization problem of an RIS assisted full-duplex MIMO two-way communication system was studied in [32], which has been solved efficiently in closed form. The investigation into the capacity of the RIS assisted MIMO system was also extended to higher frequency bands. The authors in [33] maximized the capacity of the RIS assisted MIMO system in an indoor millimeter-wave environment, ignoring the line-of-sight (LOS) path by only adjusting RIS reflection elements or jointly optimizing the RIS reflection elements and a transmit phase precoder. The authors in [34] converted the data rate maximization problem of RIS assisted THz MIMO systems to a discrete phase shift search problem and solved it. Moreover, RIS assisted multi-user MIMO systems and RIS assisted MIMO systems integrating other advanced communication techniques, e.g., simultaneous wireless information and power transfer (SWIPT) and cognitive radio (CR) have also been proposed and investigated [21], [35]–[38].

On the other hand, the RIS considered in the above-mentioned literature only functions as a passive reflector to reflect received signals. Actually, antennas have already been designed to realize passive information transfer through a single-RF chain [39], [40] or an ambient backscatter [41], [42], which motivate researchers to further explore the potential of RIS for acting as an active modulator and piggybacking its own information when helping the information transmission between the transmitter and the receiver at the same time. Karasik *et al* derived the capacity limit achieved by the RIS

assisted single-output multiple-input (SIMO) system in [43], where the information is encoded in the transmitted signal and the RIS reconfiguration. They recently implemented joint encoding of the single-RF chain transmitted signal and the RIS response in [44], [45], which has been shown to achieve information-theoretic optimality of the RIS-based modulation. This concept was also proposed and investigated by Basar in [46], [47], who has analyzed the error performance of RIS-based spatial modulation systems. Specifically, the effects of blind phases and modulation orders on the error performance of the RIS assisted system were studied. The proposed system adopts the reflecting patterns that steer beams to a single receive antenna and therefore cannot harvest the receive diversity gain. The authors in [48], [49] maximized the average signal-to-noise ratio (SNR) at the receiver under the assumption that the RIS can perform spatial modulation, and the reflecting elements are capable of carrying information by their on/off states. Recently, Guo *et al* in [50] laid a solid foundation in this field by proposing a reflecting modulation scheme for RIS assisted systems considering that the transmitter and RIS can either jointly or independently convey information. Given a signal candidate set and a reflecting pattern candidate set, a discrete optimization based joint signal mapping, shaping, and reflecting scheme was designed to minimize bit error rate, which has been proven to outperform existing modulation schemes. Furthermore, RIS-based modulation schemes have been introduced and summarized in [51], [52], highlighting the advantages of using RISs. These valuable findings revealed that an extra transmission opportunity brought by RIS can be provided, while the primary system performance is expected to be improved at the same time. Due to the mutualistic relationship between them, the concept of symbiotic communications is borrowed from biology, which not only supports the RIS information transmission but also enhances the primary transmission [53]–[56].

To the best of the authors' knowledge, most published works focus on analyzing the performance of MIMO systems equipped with passive beamforming RISs, where RISs can only reflect incident signals. Although RISs have already been integrated into an increasing number of complex communication systems for performance enhancement, neglecting the active role of RIS cannot fully exploit the potential of this novel paradigm. The literature regarding RIS based spatial modulation or information delivery mainly focuses on studying the error performance and bit error rate [53]. Although the authors in [43] studied the capacity limit of the RIS assisted system enabling information encoding in the RIS configuration, they only considered a special SIMO system with one signal data stream transmitted only. The parallel transmissions of multiple data streams make the analysis of the MIMO channel capacity problem a great challenge, since the channel gains corresponding to multiple spatial data streams should be balanced by properly designing reflection coefficients. In addition, the direct transmission link between the transmitter and the receiver is also ignored in [43], [54], which might not be practical and affects the designs of the covariance matrix at the transmitter and the reflecting patterns. Moreover, a finite alphabet for encoding is adopted in [43] to derive the capacity

limits, which cannot provide an upper bound with respect to a Gaussian distributed input. To the best of our knowledge, there is still no work jointly optimizing the reflecting patterns and the transmit covariance matrix at the transmitter for RIS assisted MIMO symbiotic communications through the capacity maximization criterion, where the information is encoded and carried by both reflecting patterns and transmitted signals.

B. Contributions

To bridge the gap of capacity analysis for RIS assisted MIMO systems, we maximize the capacity of RIS assisted MIMO symbiotic communication systems employing reflecting patterns at the RIS to carry additional information. In particular, the reflecting patterns are not uniformly activated to encode information but follow a discrete probability distribution for capacity maximization purposes. This setting is different from the uniform distribution of activation patterns in the previous literature. In this way, the reflecting patterns, reflecting set activation probability, as well as the transmit covariance matrix can be jointly designed to maximize the capacity. Since the exact expression of the capacity for MIMO systems is intractable, we develop algorithms based on the lower and upper bounds on capacity, respectively. For clarity, we list our contributions as follows:

- We analyze the capacity limits of an RIS assisted MIMO symbiotic communication system, where a multiple-antenna transmitter serves a multiple-antenna receiver with the help of an RIS. The transmitted information is signified by both reflection patterns at the RIS and transmitted data streams.
- We propose a new design of reflecting patterns at the RIS, relying on a discrete and non-uniform distribution for activation. In this way, the reflecting patterns with higher power gain will be activated with higher probabilities, leading to a higher capacity on average.
- We propose schemes to jointly design the reflecting patterns, reflecting set activation probability, and transmit covariance matrix at the transmitter based on the lower bound and upper bound on capacity. Based on the lower bound on capacity, a gradient ascent algorithm is proposed to find the optimal covariance matrix at the transmitter, reflecting patterns, and their related activation probabilities. Moreover, the reflecting patterns are also optimized by taking the criterion of maximizing the upper bound on capacity, and the closed-form solutions of reflecting activation probability and transmit covariance matrix are derived accordingly.
- We carry out a thorough investigation of the proposed design and algorithms by simulations. The superiority of the proposed design and algorithms are verified through the comparison to existing designs for optimizing capacity. In particular, we verify that the capacity of RIS assisted MIMO systems can be improved by considering the active role and information delivery functionality of RIS.

C. Organization

The remainder of the paper is organized as follows. Section II describes the system model and formulates the capacity of the considered system. Section III introduces the proposed reflecting and precoding designs based on the lower bound on capacity. In section IV, we jointly optimize the reflecting patterns, reflecting set activation probability, and transmit covariance matrix by maximizing the upper bound on capacity. Numerical comparisons and investigations are presented in section V, and conclusions of this paper are drawn in section VI.

D. Notations

In this paper, x denotes a scalar; \mathbf{x} represents a vector; \mathbf{X} stands for a matrix. $\|\mathbf{x}\|_2$, $\|\mathbf{x}\|_p$, and $\|\mathbf{x}\|_\infty$ represents l_2 norm, l_p norm and l_∞ norm of \mathbf{x} , respectively. $\|\mathbf{X}\|_F$ is the Frobenius norm of \mathbf{X} . $\text{Diag}(\mathbf{x})$ is a diagonal matrix whose diagonal entries are taken from vector \mathbf{x} , and inversely, $\text{diag}(\mathbf{X})$ is a vector whose elements are taken from the diagonal entries of matrix \mathbf{X} . $\det(\mathbf{X})$ represents the determinant of square matrix \mathbf{X} . $\text{rank}(\mathbf{X})$ denotes the rank of matrix \mathbf{X} . $\arg(x)$ denotes the angle of a complex number x . x_i denotes the i th entry of \mathbf{x} , and $x_{i,j}$ is the element in i th row and j th column of a matrix \mathbf{X} ; \mathbf{x}_i denotes the i th column of matrix \mathbf{X} . \otimes denotes Kronecker product. $(\cdot)^H$ is the conjugate transpose, and $(\cdot)^T$ denotes the transpose. $(\cdot)^+ = \max\{0, \cdot\}$ returns the positive argument or zero, otherwise. $\mathbf{X} \geq 0$ means that matrix \mathbf{X} is positive semidefinite. \mathbb{C} stands for the complex domain, and \mathbb{R} represents the real domain. $\mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ stands for the circularly symmetric complex Gaussian distribution with mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$. $\mathbf{1}_N$ represents an $N \times 1$ all-one vector and $\mathbf{I}_{N \times N}$ denotes an $N \times N$ identity matrix. $\mathbf{0}_N$ represents an $N \times 1$ all-zero vector, and $\mathbf{0}_{N \times M}$ stands for an $N \times M$ all-zero matrix. ∇ denotes the vector differential operator (gradient) of a function. $\mathbb{E}[\cdot]$ returns the expected value of the random variable/vector enclosed.

II. SYSTEM MODEL AND CAPACITY ANALYSIS

As depicted in Fig. 1, we consider an (N_t, N, N_r) RIS assisted MIMO system, where N_t , N , and N_r represent the numbers of transmit antennas, reflecting elements, and receive antennas, respectively. The transmitter equipped with N_t antennas leverages an RIS of N reflecting elements for the purpose of enhancing the data transmission to a receiver equipped with N_r antennas over quasi-static fading channels. The information is conveyed per channel use not only by the transmit signal vector, but also by the index of the reflecting pattern $\Phi_k \in \Psi$; Φ_k is the k th reflecting pattern at the RIS, which is a diagonal matrix defined by $(\Phi_k)_{n,n} \in \{0\} \cup \{e^{j\theta}, \theta \in \mathbb{R}\}$, $\forall n \in \{1, 2, \dots, N\}$; $\Psi = \{\Phi_1, \Phi_2, \dots, \Phi_K\}$ represents the reflection pattern candidate set with size $|\Psi| = K$. The channel between the transmitter and the receiver can be estimated through the training sequences fed back from the receiver by leveraging the channel reciprocity of the channel between two devices [57]. In our analysis, we assume that all involved channel state information (CSI) has been perfectly known at the transmitter and the receiver by adopting the

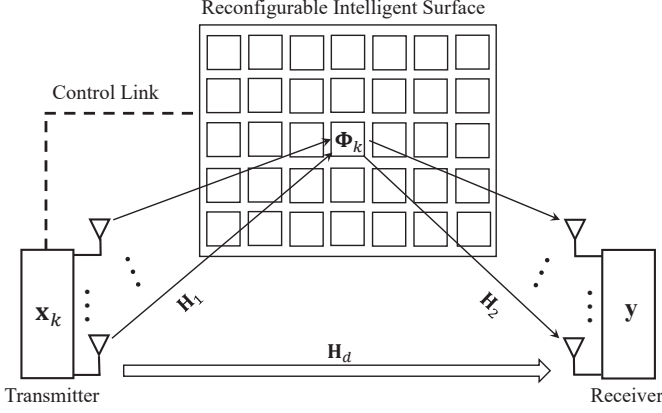


Fig. 1: A typical MIMO system assisted by an RIS with a control link connecting to the transmitter.

channel estimation techniques proposed in [16], [58], [59]. Thus, the received vector $\mathbf{y} \in N_r \times 1$ with activated Φ_k can be expressed as

$$\mathbf{y} = \sqrt{\gamma} (\mathbf{H}_2 \Phi_k \mathbf{H}_1 + \mathbf{H}_d) \mathbf{x}_k + \mathbf{n}, \quad (1)$$

where γ is the average receive SNR; $\mathbf{H}_2 \in \mathbb{C}^{N_r \times N}$ is the channel coefficient matrix characterizing the channels between the RIS and the receiver; $\mathbf{H}_1 \in \mathbb{C}^{N \times N_t}$ is the channel coefficient matrix characterizing the channels between the transmitter and the RIS; $\mathbf{H}_d \in \mathbb{C}^{N_r \times N_t}$ is the channel coefficient matrix characterizing the direct channels between the transmitter and the receiver¹; $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}_{N_r}, \mathbf{I}_{N_r \times N_r})$ is the additive white Gaussian noise (AWGN) vector with each entry abiding a zero-mean unit-variance complex Gaussian distribution; $\mathbf{x}_k \in \mathbb{C}^{N_t \times 1}$ denotes the transmitted signal vector when Φ_k is activated with zero mean and covariance matrix \mathbf{Q}_k , i.e., $\mathbf{x}_k \sim \mathcal{CN}(\mathbf{0}_{N_t}, \mathbf{Q}_k)$. It is assumed that all signal vectors are subject to a unit average power constraint $\mathbb{E}\{\text{Tr}(\mathbf{Q}_k)\} \leq 1$ for simplicity.

Since Φ_k is activated not only for assisting transmissions between the transmitter and the receiver, but also for carrying additional information, it becomes challenging to quantify the channel capacity in this case. To simplify the following discussion, the channel matrices, phase shift matrix, and the transmitted data stream are built in the new forms as follows [50]:

$$\hat{\mathbf{H}}_2 = \overbrace{[\mathbf{H}_2, \mathbf{H}_2, \dots, \mathbf{H}_2]}^K \in \mathbb{C}^{N_r \times KN}, \quad (2)$$

¹Here, we do not give any specific channel model, which means the following developed design can be applied to any channel model, including traditional Rayleigh, Rician, Nakagami-m, and even the recently proposed spatially correlated Rayleigh fading model for RIS-assisted systems [60].

$$\Theta = \begin{bmatrix} \Phi_1 & \mathbf{0}_{N \times N} & \cdots & \mathbf{0}_{N \times N} \\ \mathbf{0}_{N \times N} & \Phi_2 & \cdots & \mathbf{0}_{N \times N} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{N \times N} & \mathbf{0}_{N \times N} & \cdots & \Phi_K \end{bmatrix} \in \mathbb{C}^{KN \times KN}, \quad (3)$$

$$\hat{\mathbf{H}}_1 = \begin{bmatrix} \mathbf{H}_1 & \mathbf{0}_{N \times N_t} & \cdots & \mathbf{0}_{N \times N_t} \\ \mathbf{0}_{N \times N_t} & \mathbf{H}_1 & \cdots & \mathbf{0}_{N \times N_t} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{N \times N_t} & \mathbf{0}_{N \times N_t} & \cdots & \mathbf{H}_1 \end{bmatrix} \in \mathbb{C}^{KN \times KN_t}, \quad (4)$$

$$\hat{\mathbf{H}}_d = \overbrace{[\mathbf{H}_d, \mathbf{H}_d, \dots, \mathbf{H}_d]}^K \in \mathbb{C}^{N_r \times KN_t}, \quad (5)$$

and

$$\hat{\mathbf{x}} = \mathbf{g}_k \otimes \mathbf{x}_k \in \mathbb{C}^{KN_t \times 1}, \quad (6)$$

where $\mathbf{g}_k \in \mathbb{C}^{K \times 1}$ is the k th vector basis with all zeros except the k th entry being one. As a direct result of these expressions, the received signal vector can be rewritten as

$$\mathbf{y} = \sqrt{\gamma} (\hat{\mathbf{H}}_d + \hat{\mathbf{H}}_2 \Theta \hat{\mathbf{H}}_1) \hat{\mathbf{x}} + \mathbf{n}. \quad (7)$$

It can be proven in a straightforward manner that $\hat{\mathbf{x}}$ also satisfies the power constraint as follows:

$$\begin{aligned} \mathbb{E}\{\text{Tr}(\hat{\mathbf{x}}\hat{\mathbf{x}}^H)\} &= \mathbb{E}\left\{\text{Tr}\left(\left(\mathbf{g}_k \otimes \mathbf{x}_k\right)\left(\mathbf{g}_k \otimes \mathbf{x}_k\right)^H\right)\right\} \\ &= \mathbb{E}\left\{\text{Tr}\left(\left(\mathbf{g}_k \mathbf{g}_k^H\right) \otimes \left(\mathbf{x}_k \mathbf{x}_k^H\right)\right)\right\} \leq 1. \end{aligned} \quad (8)$$

Denoting $\mathbf{H} = \hat{\mathbf{H}}_d + \hat{\mathbf{H}}_2 \Theta \hat{\mathbf{H}}_1$, the channel capacity of the RIS assisted MIMO system detailed above can be derived by maximizing its mutual information [61], which is given by

$$\begin{aligned} I(\hat{\mathbf{x}}; \mathbf{y}) &= H(\mathbf{y}) - H(\mathbf{y} | \mathbf{H}, \hat{\mathbf{x}}) \\ &= \mathbb{E}[-\log_2 f(\mathbf{y})] - N_r \log_2(\pi e), \end{aligned} \quad (9)$$

where $H(\cdot)$ and $H(\cdot | \cdot)$ represent the entropy and conditional entropy of the enclosed; $f(\cdot)$ denotes the probability density function (PDF) of the enclosed. Accordingly, the channel capacity can be expressed as

$$\begin{aligned} C &= \max_{\substack{\mathbb{E}\{\text{Tr}(\hat{\mathbf{x}}\hat{\mathbf{x}}^H)\} \leq 1 \\ \hat{\mathbf{x}} = \mathbf{g}_k \otimes \mathbf{x}_k}} I(\hat{\mathbf{x}}; \mathbf{y}) \\ &= \max_{\substack{\mathbb{E}\{\text{Tr}(\mathbf{Q}_k)\} \leq 1 \\ \sum_{k=1}^K f_{\mathbf{g}_k}(\mathbf{g}_k) = 1}} I(\mathbf{x}_k, \mathbf{g}_k; \mathbf{y}), \end{aligned} \quad (10)$$

where $f_{\mathbf{g}_k}(\mathbf{g}_k)$ stands for the activation probability of the k th reflecting pattern. For denotation simplicity, we use a general discrete distribution $\alpha = [\alpha_1, \alpha_2, \dots, \alpha_K]$ to represent $\{f_{\mathbf{g}_k}(\mathbf{g}_k)\}$ and let

$$f_{\mathbf{g}_k}(\mathbf{g}_k) = \alpha_k, \quad (11)$$

subject to

$$\sum_{k=1}^K \alpha_k = 1. \quad (12)$$

With these denotations, the unit average power constraint can be written as

$$\sum_{k=1}^K \alpha_k \text{Tr}(\mathbf{Q}_k) \leq 1. \quad (13)$$

Defining $\mathcal{Q} = \{\mathbf{Q}_1, \mathbf{Q}_2, \dots, \mathbf{Q}_K\}$, we use $\mathcal{R}(\boldsymbol{\alpha}, \boldsymbol{\Psi}, \mathcal{Q})$ to represent the achievable capacity. Supposing Gaussian distributed signals as inputs, we can derive the mutual information between $\hat{\mathbf{x}}$ and \mathbf{y} as [62]

$$\mathcal{R}(\boldsymbol{\alpha}, \boldsymbol{\Psi}, \mathcal{Q}) = - \int_{\mathbb{C}^{N_r}} f(\mathbf{y}) \log_2 f(\mathbf{y}) d\mathbf{y} - N_r \log_2(\pi e), \quad (14)$$

where \mathbf{y} follows a Gaussian mixture distribution with mixture coefficients $\{\alpha_k\}$, which can be explicitly written as

$$f(\mathbf{y}) = \sum_{k=1}^K \alpha_k f(\mathbf{y} | k), \quad (15)$$

where $f(\mathbf{y} | k)$ is the conditional distribution of \mathbf{y} following a complex normal distribution with zero mean and covariance matrix \mathbf{D}_k , given by,

$$f(\mathbf{y} | k) = \frac{1}{\pi^{N_r} \det(\mathbf{D}_k)} \exp(-\mathbf{y}^H \mathbf{D}_k^{-1} \mathbf{y}), \quad (16)$$

with

$$\mathbf{D}_k = \mathbf{I}_{N_r \times N_r} + \gamma \mathbf{H}_k \mathbf{Q}_k \mathbf{H}_k^H, \quad (17)$$

where $\mathbf{H}_k = \mathbf{H}_2 \boldsymbol{\Phi}_k \mathbf{H}_1 + \mathbf{H}_d$. However, the capacity expressed in (14) cannot be solved in closed form, while it can only be evaluated via numerical integration methods, e.g., Monte Carlo sampling [63], which incurs prohibitive complexity for practical applications. Therefore for the mathematical tractability, we provide closed-form expressions of the lower and upper bounds on $\mathcal{R}(\boldsymbol{\alpha}, \boldsymbol{\Psi}, \mathcal{Q})$ in **Theorem 1** as follows, which are denoted as $\mathcal{R}^L(\boldsymbol{\alpha}, \boldsymbol{\Psi}, \mathcal{Q})$ and $\mathcal{R}^U(\boldsymbol{\alpha}, \boldsymbol{\Psi}, \mathcal{Q})$, respectively.

Theorem 1. Closed-form lower and upper bounds on the achievable SE of the proposed RIS assisted MIMO system are given by

$$\mathcal{R}^L(\boldsymbol{\alpha}, \boldsymbol{\Psi}, \mathcal{Q}) = - \sum_{k=1}^K \alpha_k \log_2 \left(\sum_{l=1}^K \alpha_l z_{k,l} \right) - N_r \log_2(e), \quad (18)$$

and

$$\mathcal{R}^U(\boldsymbol{\alpha}, \boldsymbol{\Psi}, \mathcal{Q}) = \sum_{k=1}^K \alpha_k (\log_2(\det \mathbf{D}_k) - \log_2(\alpha_k)), \quad (19)$$

where $z_{k,l} = (\det(\mathbf{D}_k + \mathbf{D}_l))^{-1}$.

Proof. See proof in Appendix A. \square

It should be noted that the lower bound can be negative in the low SNR regime. However, it will be always positive and tight in the low and high SNR regimes by adding a constant gap $N_r (\log_2 e - 1)$, which has been proven in [64]. Since the constant is independent of the three optimization variables, i.e., reflecting activation probability $\boldsymbol{\alpha}$, reflecting patterns $\boldsymbol{\Psi}$, and the transmit covariance matrix \mathcal{Q} , the optimization objective function can be both $\mathcal{R}^L(\boldsymbol{\alpha}, \boldsymbol{\Psi}, \mathcal{Q})$ given in (18) and $\mathcal{R}^L(\boldsymbol{\alpha}, \boldsymbol{\Psi}, \mathcal{Q}) + N_r (\log_2 e - 1)$. The upper bound is tight

in the high SNR regime, which has been proven in [65] and [66]. The tightness of both capacity bounds will also be numerically investigated and verified via simulations in the sequel. The closed-form lower bound $\mathcal{R}^L(\boldsymbol{\alpha}, \boldsymbol{\Psi}, \mathcal{Q})$ and upper bound $\mathcal{R}^U(\boldsymbol{\alpha}, \boldsymbol{\Psi}, \mathcal{Q})$ proposed by **Theorem 1** in this paper enable computationally efficient evaluation and optimization of RIS assisted MIMO systems by jointly designing the reflecting patterns, activation probability, and covariance matrix of input signals. We will present the details and show their usefulness in the following sections.

III. LOWER BOUND BASED OPTIMIZATION

In this section, relying on the derived lower bound, we find the optimal phase shift set $\boldsymbol{\Psi}$, activation distribution $\boldsymbol{\alpha}$, and transmit covariance matrix set \mathcal{Q} to maximize the capacity lower bound and obtain insights into the designs for the proposed RIS assisted MIMO system.

A. Problem Formulation

Based on the derived expression of lower bound $\mathcal{R}^L(\boldsymbol{\alpha}, \boldsymbol{\Psi}, \mathcal{Q})$, the optimization problem can be formulated as

$$\begin{aligned} (\mathbf{P1}) : \quad & \max_{\boldsymbol{\alpha}, \boldsymbol{\Psi}, \mathcal{Q}} \mathcal{R}^L(\boldsymbol{\alpha}, \boldsymbol{\Psi}, \mathcal{Q}) \\ \text{s.t.} : \quad & \mathbf{1}_K^T \boldsymbol{\alpha} = 1, \\ & \boldsymbol{\alpha} \geq \mathbf{0}_K, \\ & |\boldsymbol{\Phi}_k(i, i)| = 1, \\ & \sum_{k=1}^K \alpha_k \text{Tr}(\mathbf{Q}_k) \leq 1, \\ & \mathbf{Q}_k \succeq \mathbf{0}. \end{aligned} \quad (20)$$

Obviously, (P1) is a non-convex optimization problem since the uni-modular constraint on each reflection coefficient $\boldsymbol{\Phi}_k(i, i)$ is non-convex, and the objective function is shown to be non-concave over the reflection matrix $\boldsymbol{\Phi}_k$ in $\boldsymbol{\Psi}$. As a consequence of this non-convexity, it is difficult to solve (P1) because the activation probability $\boldsymbol{\alpha}$ and the transmit covariance matrix set \mathcal{Q} are coupled with $\boldsymbol{\Psi}$ in the objective function. To simplify the problem, we first re-express the non-convex constraint on the reflection matrix $\boldsymbol{\Phi}_k$ in the next step.

Since $\boldsymbol{\Phi}_k$ is a diagonal matrix, we define $\boldsymbol{\phi}_k = \text{diag}(\boldsymbol{\Phi}_k)$ and stack them into vector $\boldsymbol{\delta} = [\boldsymbol{\phi}_1^T, \dots, \boldsymbol{\phi}_K^T]^T \in \mathbb{C}^{NK \times 1}$. To handle the non-convex constraint $|\delta_i| = 1$, we re-express it to be the l_∞ constraint, which leads to the following new set

$$\begin{aligned} & \{\boldsymbol{\delta} \in \mathbb{C}^{NK \times 1} : |\delta_i| = 1, \forall i = 1, \dots, NK\} \\ & = \{\boldsymbol{\delta} \in \mathbb{C}^{NK \times 1} : \text{tr}(\boldsymbol{\delta} \boldsymbol{\delta}^H) = NK, \|\boldsymbol{\delta}\|_\infty \leq 1\}. \end{aligned} \quad (21)$$

It should be noted that the second set satisfies two conditions. One is $\|\boldsymbol{\delta}\|_\infty \leq 1$, which indicates $\max_{i=1, \dots, NK} |\delta_i| \leq 1$, while the other one $\text{tr}(\boldsymbol{\delta} \boldsymbol{\delta}^H) = NK$ means that $\sum_{i=1}^{NK} \delta_i^2 = NK$. It can be observed that there exists only a unique solution of δ_i , that is, $|\delta_i| = 1$, which exactly equals the first set.

Based on this transformation, the optimization problem can be reformulated as

$$\begin{aligned}
(\mathbf{P2}) : \quad & \max_{\alpha, \delta, \mathcal{Q}} \mathcal{R}^L(\alpha, \delta, \mathcal{Q}) \\
\text{s.t.} : \quad & \mathbf{1}_K^T \alpha = 1, \\
& \alpha \geq \mathbf{0}_K, \\
& \text{tr}(\delta \delta^H) = NK, \\
& \|\delta\|_\infty \leq 1, \\
& \sum_{k=1}^K \alpha_k \text{Tr}(\mathbf{Q}_k) \leq 1, \\
& \mathbf{Q}_k \succeq 0.
\end{aligned} \tag{22}$$

Since the ℓ_∞ constraint is non-differentiable, we can exploit the ℓ_p approximation with a gradually increased large p , leading to $\lim_{p \rightarrow \infty} \|\delta\|_p = \|\delta\|_\infty$ during the optimization process [64].

Moreover, we then consider the optimization of \mathcal{Q} , given δ and α . Given δ , the reflecting patterns and $\{\mathbf{H}_k\}$ are equivalently given. Provided that $\{\mathbf{H}_k\}$ and α are known, $(\mathbf{P2})$ is a convex optimization problem over \mathbf{Q}_k , and the optimal \mathbf{Q}_k is given by the eigenmode transmission. In details, denoting the maximum number of data streams that can be transmitted over \mathbf{H}_k as $N_s = \text{rank}(\mathbf{H}_k) \leq \min\{N_t, N_r\}$, and the singular value decomposition of \mathbf{H}_k as

$$\mathbf{H}_k = \mathbf{U}_k \Sigma_k \mathbf{V}_k^H, \tag{23}$$

we express the transmit covariance matrix as $\mathbf{Q}_k = \mathbf{V}_k \text{Diag}\{\lambda_k\} \mathbf{V}_k^H$, where \mathbf{V}_k is the right-singular matrix; $\lambda_k = [\lambda_{k,1}, \dots, \lambda_{k,N_s}]^T$ with entries $\lambda_{k,j}$ representing the optimal amount of power allocated to the j th data stream. By defining $\omega = [\lambda_1^T, \dots, \lambda_K^T]^T$ and $\beta = \alpha \otimes \mathbf{1}_{N_t}$, we can re-express the unit average power constraint to be $\beta^T \omega \leq 1$. It should be noted that because of the monotonicity of the achievable capacity with power, $\mathcal{R}^L(\alpha, \delta, \omega)$ is maximized when the power constraint is met with strict equality. Therefore, we transform the inequality power constraint to be the equality power constraint $\beta^T \omega = 1$. Accordingly, $(\mathbf{P2})$ can be transformed to be

$$\begin{aligned}
(\mathbf{P3}) : \quad & \max_{\alpha, \delta, \omega} \mathcal{R}^L(\alpha, \delta, \omega), \\
\text{s.t.} : \quad & \mathbf{1}_K^T \alpha = 1, \\
& \alpha \geq \mathbf{0}_K, \\
& \text{tr}(\delta \delta^H) = NK, \\
& \|\delta\|_\infty \leq 1, \\
& \beta^T \omega = 1, \\
& \omega > 0.
\end{aligned} \tag{24}$$

B. Transmitter and Reflector Designs

To solve $(\mathbf{P3})$, we resort to the barrier method and incorporate the non-negative constraint with the logarithmic barrier function $I(u)$ to approximate the penalty of violating the ℓ_p constraint, i.e., $I(u) = \begin{cases} \frac{1}{t} \ln(u), & u > 0 \\ -\infty, & u \leq 0, \end{cases}$ where t is used

to scale the barrier function's penalty [67]. Then, $(\mathbf{P3})$ can be rewritten as

$$\begin{aligned}
(\mathbf{P4}) : \quad & \max_{\alpha, \delta, \omega} f(\alpha, \delta, \omega) \\
\text{s.t.} : \quad & \mathbf{1}_K^T \alpha = 1, \\
& \text{tr}(\delta \delta^H) = NK, \\
& \beta^T \omega = 1,
\end{aligned} \tag{25}$$

where the objective function $f(\alpha, \delta, \omega)$ is given by

$$\begin{aligned}
f(\alpha, \delta, \omega) = & R^L(\alpha, \delta, \omega) + I(1 - \|\delta\|_p) \\
& + \sum_{k=1}^K I(\alpha_k) + \sum_{k=1}^K \sum_{i=1}^{N_s} I(\lambda_{ki}).
\end{aligned} \tag{26}$$

Accordingly, the gradient of $f(\alpha, \delta, \omega)$ with respect to α is derived to be

$$\nabla_{\alpha} f(\alpha, \delta, \omega) = [\nabla_{\alpha_1} f(\alpha, \delta, \omega), \dots, \nabla_{\alpha_K} f(\alpha, \delta, \omega)]^T, \tag{27}$$

where

$$\nabla_{\alpha_k} f(\alpha, \delta, \omega) = \nabla_{\alpha_k} R^L(\alpha, \delta, \omega) + \frac{1}{t \alpha_k}, \tag{28}$$

and

$$\begin{aligned}
\nabla_{\alpha_k} R^L(\alpha, \delta, \omega) = & -\log_2 \left(\sum_{i=1}^K \alpha_i z_{k,i} \right) \\
& - \frac{1}{\ln 2} \sum_{j=1}^K \frac{\alpha_j z_{j,k}}{\sum_{i=1}^K \alpha_i z_{j,i}}.
\end{aligned} \tag{29}$$

The gradient of $f(\alpha, \delta, \omega)$ with respect to δ can be expressed as

$$\begin{aligned}
\nabla_{\delta} f(\alpha, \delta, \omega) = & [\nabla_{\phi_1} R^L(\alpha, \delta, \omega), \dots, \nabla_{\phi_K} R^L(\alpha, \delta, \omega)]^T \\
& - \frac{\|\delta\|_p^{1-p} \mathbf{p}_{\delta}}{t(1 - \|\delta\|_p)},
\end{aligned} \tag{30}$$

where $\mathbf{p}_{\delta} \in \mathbb{C}^{KN \times 1}$ is given by $\mathbf{p}_{\delta} = [\phi_{1,1} \cdot |\phi_{1,1}|^{p-2}, \dots, \phi_{1,N} \cdot |\phi_{1,N}|^{p-2}, \dots, \phi_{K,N} \cdot |\phi_{K,N}|^{p-2}]^T$.

Therefore, $\nabla_{\phi_k} R^L(\boldsymbol{\alpha}, \boldsymbol{\delta}, \boldsymbol{\omega})$ can be determined as

$$\begin{aligned}
\nabla_{\phi_k} R^L(\boldsymbol{\alpha}, \boldsymbol{\delta}, \boldsymbol{\omega}) &= -\sum_{j=1}^K \alpha_j \nabla_{\phi_k} \log_2 \left(\sum_{i=1}^K \alpha_i z_{j,i} \right) \\
&= -\alpha_k \nabla_{\phi_k} \log_2 \left(\sum_{i=1}^K \alpha_i z_{k,i} \right) \\
&\quad - \sum_{j=1, j \neq k}^K \alpha_j \nabla_{\phi_k} \log_2 \left(\sum_{i=1}^K \alpha_i z_{j,i} \right) \\
&= -\frac{1}{\ln 2} \left[\alpha_k \frac{\sum_{i=1}^K \alpha_i \nabla_{\phi_k} z_{k,i}}{\sum_{l=1}^K \alpha_l z_{k,l}} + \sum_{j=1, j \neq k}^K \alpha_j \frac{\sum_{i=1}^K \alpha_i \nabla_{\phi_k} z_{j,i}}{\sum_{l=1}^K \alpha_l z_{j,l}} \right] \\
&= -\frac{1}{\ln 2} \left[\alpha_k \frac{\sum_{i=1}^K \alpha_i \nabla_{\phi_k} z_{k,i}}{\sum_{l=1}^K \alpha_l z_{k,l}} - \sum_{j=1, j \neq k}^K \alpha_j \frac{\alpha_k \nabla_{\phi_k} z_{j,k}}{\sum_{l=1}^K \alpha_l z_{j,l}} \right] \\
&= -\frac{1}{\ln 2} \left[\sum_{i=1}^K \frac{\alpha_k \alpha_i \nabla_{\phi_k} z_{k,i}}{\sum_{l=1}^K \alpha_l z_{k,l}} + \sum_{j=1, j \neq k}^K \frac{\alpha_j \alpha_k \nabla_{\phi_k} z_{j,k}}{\sum_{l=1}^K \alpha_l z_{j,l}} \right], \tag{31}
\end{aligned}$$

with

$$\nabla_{\phi_k} z_{k,i} = z_{k,i} \text{Tr} \left[(\mathbf{D}_k + \mathbf{D}_l)^{-1} \nabla_{\phi_k} \mathbf{D}_k \right]. \tag{32}$$

Since $\mathbf{D}_k = \mathbf{I}_{N_r \times N_r} + \gamma (\mathbf{H}_d + \mathbf{H}_2 \boldsymbol{\Phi}_k \mathbf{H}_1) \mathbf{Q}_k (\mathbf{H}_d + \mathbf{H}_2 \boldsymbol{\Phi}_k \mathbf{H}_1)^H$, \mathbf{D}_k can be further expanded to be

$$\begin{aligned}
\mathbf{D}_k &= \mathbf{I}_{N_r \times N_r} + \gamma \mathbf{H}_d \mathbf{Q}_k \mathbf{H}_d^H + \gamma \mathbf{H}_d \mathbf{Q}_k \mathbf{H}_1^H \boldsymbol{\Phi}_k^H \mathbf{H}_2^H \\
&\quad + \gamma \mathbf{H}_2 \boldsymbol{\Phi}_k \mathbf{H}_1 \mathbf{Q}_k \mathbf{H}_d^H + \gamma \mathbf{H}_2 \boldsymbol{\Phi}_k \mathbf{H}_1 \mathbf{Q}_k \mathbf{H}_1^H \boldsymbol{\Phi}_k^H \mathbf{H}_2^H. \tag{33}
\end{aligned}$$

Omitting the irrelevant terms in the above form, $\nabla_{\phi_k} \mathbf{D}_k$ can be recast as

$$\begin{aligned}
\nabla_{\phi_k} \mathbf{D}_k &= \gamma \nabla_{\phi_k} \mathbf{H}_2 \boldsymbol{\Phi}_k \mathbf{H}_1 \mathbf{Q}_k \mathbf{H}_d^H + \gamma \nabla_{\phi_k} \mathbf{H}_1 \mathbf{Q}_k \mathbf{H}_1^H \boldsymbol{\Phi}_k^H \mathbf{H}_2^H \\
&\quad + \gamma \nabla_{\phi_k} \mathbf{H}_2 \boldsymbol{\Phi}_k \mathbf{H}_1 \mathbf{Q}_k \mathbf{H}_1^H \boldsymbol{\Phi}_k^H \mathbf{H}_2^H, \tag{34}
\end{aligned}$$

where $\mathbf{H}_{1d} = \mathbf{H}_1 \mathbf{Q}_k \mathbf{H}_d^H$. We can easily obtain $\nabla_{\phi_k} \mathbf{H}_{1d}^H \boldsymbol{\Phi}_k^H \mathbf{H}_2^H = 0$ and the gradient of $\mathbf{H}_2 \boldsymbol{\Phi}_k \mathbf{H}_{1d}$ respect to the i th element of ϕ_k as

$$\nabla_{\phi_k, i} \mathbf{H}_2 \boldsymbol{\Phi}_k \mathbf{H}_{1d} = \mathbf{h}_{2,i} \mathbf{h}_{1d,i}^T, \tag{35}$$

where $\mathbf{h}_{2,i}$ is the i th column of \mathbf{H}_2 and $\mathbf{h}_{1d,i}$ is the i th column of \mathbf{H}_{1d}^T .

Similarly, the gradient of $\mathbf{H}_2 \boldsymbol{\Phi}_k \mathbf{H}_1 \mathbf{Q}_k \mathbf{H}_1^H \boldsymbol{\Phi}_k^H \mathbf{H}_2^H$ respect to ϕ_k, i can be calculated as

$$\nabla_{\phi_k, i} \mathbf{H}_2 \boldsymbol{\Phi}_k \mathbf{H}_1 \mathbf{Q}_k \mathbf{H}_1^H \boldsymbol{\Phi}_k^H \mathbf{H}_2^H = \mathbf{h}_{2,i} \mathbf{h}_{1,i}^T \mathbf{Q}_k \mathbf{H}_1^H \boldsymbol{\Phi}_k^H \mathbf{H}_2^H, \tag{36}$$

where $\mathbf{h}_{1,i}$ is the i th column of \mathbf{H}_1^T . Thus, $\nabla_{\phi_k} z_{k,i}$ can be finally obtained as

$$\begin{aligned}
\nabla_{\phi_k} z_{k,i} &= z_{k,i} \text{diag} \left[\mathbf{H}_{1d} (\mathbf{D}_k + \mathbf{D}_l)^{-1} \mathbf{H}_2 \right. \\
&\quad \left. + \mathbf{H}_1 \mathbf{Q}_k \mathbf{H}_1^H \boldsymbol{\Phi}_k^H \mathbf{H}_2^H (\mathbf{D}_k + \mathbf{D}_l)^{-1} \mathbf{H}_2 \right]. \tag{37}
\end{aligned}$$

The gradient of $f(\boldsymbol{\alpha}, \boldsymbol{\delta}, \boldsymbol{\omega})$ with respect to $\boldsymbol{\omega}$ can be expressed as

$$\nabla_{\boldsymbol{\omega}}(\boldsymbol{\alpha}, \boldsymbol{\delta}, \boldsymbol{\omega}) = [\nabla_{\lambda_1} f(\boldsymbol{\alpha}, \boldsymbol{\delta}, \boldsymbol{\omega}), \dots, \nabla_{\lambda_K} f(\boldsymbol{\alpha}, \boldsymbol{\delta}, \boldsymbol{\omega})]^T, \tag{38}$$

where

$$\nabla_{\lambda_k} f(\boldsymbol{\alpha}, \boldsymbol{\delta}, \boldsymbol{\omega}) = \nabla_{\lambda_k} R^L(\boldsymbol{\alpha}, \boldsymbol{\delta}, \boldsymbol{\omega}) + \frac{1}{t} \boldsymbol{\nu}_k, \tag{39}$$

$$\boldsymbol{\nu}_k = [\lambda_{k,1}^{-1}, \dots, \lambda_{k,N_s}^{-1}]^T, \tag{40}$$

and the derivation of $\nabla_{\lambda_k} R^L(\boldsymbol{\alpha}, \boldsymbol{\delta}, \boldsymbol{\omega})$ is similar to $\nabla_{\phi_k} R^L(\boldsymbol{\alpha}, \boldsymbol{\delta}, \boldsymbol{\omega})$ given in (31), which is given by

$$\begin{aligned}
\nabla_{\lambda_k} R^L(\boldsymbol{\alpha}, \boldsymbol{\delta}, \boldsymbol{\omega}) &= -\frac{1}{\ln 2} \left[\sum_{i=1}^K \frac{\alpha_k \alpha_i \nabla_{\lambda_k} z_{k,i}}{\sum_{l=1}^K \alpha_l z_{k,l}} \right. \\
&\quad \left. + \sum_{j=1, j \neq k}^K \frac{\alpha_j \alpha_k \nabla_{\lambda_k} z_{j,k}}{\sum_{l=1}^K \alpha_l z_{j,l}} \right], \tag{41}
\end{aligned}$$

and

$$\nabla_{\lambda_k} z_{k,i} = \gamma z_{k,i} \text{diag} \left[\mathbf{V}_k^H \mathbf{H}_k^H (\mathbf{D}_k + \mathbf{D}_i)^{-1} \mathbf{H}_k \mathbf{V}_k \right]. \tag{42}$$

To meet the equality constraints $\mathbf{1}_K^T \boldsymbol{\alpha} = 1$ and $\boldsymbol{\beta}^T \boldsymbol{\omega} = 1$, we introduce the following projections:

$$\Delta \boldsymbol{\alpha} = \mathbf{A}_1 \nabla_{\boldsymbol{\alpha}} f(\boldsymbol{\alpha}, \boldsymbol{\delta}, \boldsymbol{\omega}), \tag{43}$$

and

$$\Delta \boldsymbol{\omega} = (\mathbf{I}_{KN_s \times KN_s} - \mathbf{1}_{KN_s} \boldsymbol{\beta}^T) \nabla_{\boldsymbol{\omega}} f(\boldsymbol{\alpha}, \boldsymbol{\delta}, \boldsymbol{\omega}), \tag{44}$$

which ensure $\mathbf{1}_K^T \Delta \boldsymbol{\alpha} = 0$, $\mathbf{b}^T \Delta \boldsymbol{\alpha} = 0$, and $\boldsymbol{\beta}^T \Delta \boldsymbol{\omega} = 0$

with $\mathbf{b} = [\sum_{j=1}^{N_s} \lambda_{1,j}, \dots, \sum_{j=1}^{N_s} \lambda_{K,j}]^T$, $\mathbf{A}_1 = \mathbf{I}_{K \times K} - \mathbf{B}^T (\mathbf{B} \mathbf{B}^T)^{-1} \mathbf{B}$, and $\mathbf{B} = [\mathbf{1}_K \mathbf{b}]^T$ if $\mathbf{b} \neq \mathbf{1}_K$ or $\mathbf{B} = \mathbf{1}_K^T$ if $\mathbf{b} = \mathbf{1}_K$. Then, the iteration process for updating $\boldsymbol{\alpha}$ and $\boldsymbol{\omega}$ can be designed as

$$\boldsymbol{\alpha}^{(i+1)} \leftarrow \boldsymbol{\alpha}^{(i)} + \eta_1 \Delta \boldsymbol{\alpha}^{(i)}, \tag{45}$$

and

$$\boldsymbol{\omega}^{(i+1)} \leftarrow \boldsymbol{\omega}^{(i)} + \eta_3 \Delta \boldsymbol{\omega}^{(i)}, \tag{46}$$

where η_1 and η_3 are the step sizes. It should be noted that the searching steps may be small if some entries of $\boldsymbol{\alpha}$ and $\boldsymbol{\omega}$ approach zero in the iteration process, while leads to local optima of $\boldsymbol{\alpha}$ and $\boldsymbol{\omega}$. To avoid being trapped in local optima, we adopt a gradient modification scheme proposed in [65].

For non-linear equality constraint $\text{tr}(\boldsymbol{\delta} \boldsymbol{\delta}^H) = NK$, we project the search direction of $\boldsymbol{\delta}$ into its constrained tangent plane vertically by

$$\Delta \boldsymbol{\delta} = \nabla_{\boldsymbol{\delta}} f(\boldsymbol{\alpha}, \boldsymbol{\delta}, \boldsymbol{\omega}) - \frac{\langle \nabla_{\boldsymbol{\delta}} f(\boldsymbol{\alpha}, \boldsymbol{\delta}, \boldsymbol{\omega}), \boldsymbol{\delta} \rangle \boldsymbol{\delta}}{\|\boldsymbol{\delta}\|_2^2}, \tag{47}$$

which also ensures that $\boldsymbol{\delta}^H \Delta \boldsymbol{\delta} = 0$. After that, we renew $\boldsymbol{\delta}$ as

$$\boldsymbol{\delta}^{(i+1)} \leftarrow \cos(\eta_2) \boldsymbol{\delta}^{(i)} + \sin(\eta_2) \sqrt{NK} \frac{\Delta \boldsymbol{\delta}^{(i)}}{\|\Delta \boldsymbol{\delta}^{(i)}\|}, \tag{48}$$

where $\eta_2 \in [0, \frac{\pi}{2}]$ is the step size. This update step makes sure that the renewal solution satisfies equality constraint $\text{tr}(\boldsymbol{\delta}^{(i+1)}(\boldsymbol{\delta}^{(i+1)})^H) = NK$, which can be proved as

$$\begin{aligned} & \text{tr}(\boldsymbol{\delta}^{(i+1)}(\boldsymbol{\delta}^{(i+1)})^H) = (\boldsymbol{\delta}^{(i+1)})^H \boldsymbol{\delta}^{(i+1)} \\ &= \left(\cos(\eta_2) \boldsymbol{\delta}^{(i)} + \sin(\eta_2) \sqrt{NK} \frac{\Delta \boldsymbol{\delta}^{(i)}}{\|\Delta \boldsymbol{\delta}^{(i)}\|} \right)^H \\ & \times \left(\cos(\eta_2) \boldsymbol{\delta}^{(i)} + \sin(\eta_2) \sqrt{NK} \frac{\Delta \boldsymbol{\delta}^{(i)}}{\|\Delta \boldsymbol{\delta}^{(i)}\|} \right) \\ &= \cos(\eta_2)^2 (\boldsymbol{\delta}^{(i)})^H \boldsymbol{\delta}^{(i)} + \sin(\eta_2)^2 NK \\ &+ 2 \cos(\eta_2) \sin(\eta_2) (\boldsymbol{\delta}^{(i)})^H \frac{\Delta \boldsymbol{\delta}^{(i)}}{\|\Delta \boldsymbol{\delta}^{(i)}\|} \\ &= \cos(\eta_2)^2 NK + \sin(\eta_2)^2 NK \\ &= NK. \end{aligned} \quad (49)$$

Combining these obtained gradients and projections, we develop a gradient ascent algorithm to maximize the lower bound on channel capacity as stipulated in **Algorithm 1**. It should be noted that the amplitude of RIS elements in each iteration does not satisfy $|\phi_{k,n}| = 1$ due to the constraint relaxation given in (21) and the ℓ_p approximation. Therefore, an element-wise normalization for $\boldsymbol{\delta}^*$ by $\phi_{k,n}^* = \frac{\phi_{k,n}^*}{|\phi_{k,n}^*|}$ is added to the output by **Algorithm 1** to ensure the feasibility of the obtained solution. The normalization is also conditioned on that the objective function is not much sensitive to the normalization to achieve favorable performance.

C. Convergence and Complexity Analysis

The converge of our proposed algorithm can be observed from the updating process of $\boldsymbol{\alpha}$, $\boldsymbol{\delta}$, $\boldsymbol{\omega}$ in each iteration, that is,

$$f(\boldsymbol{\alpha}^{(i+1)}, \boldsymbol{\delta}^{(i)}, \boldsymbol{\omega}^{(i)}) - f(\boldsymbol{\alpha}^{(i)}, \boldsymbol{\delta}^{(i)}, \boldsymbol{\omega}^{(i)}) > 0, \quad (54)$$

$$f(\boldsymbol{\alpha}^{(i+1)}, \boldsymbol{\delta}^{(i+1)}, \boldsymbol{\omega}^{(i)}) - f(\boldsymbol{\alpha}^{(i+1)}, \boldsymbol{\delta}^{(i)}, \boldsymbol{\omega}^{(i)}) > 0, \quad (55)$$

and

$$f(\boldsymbol{\alpha}^{(i+1)}, \boldsymbol{\delta}^{(i+1)}, \boldsymbol{\omega}^{(i+1)}) - f(\boldsymbol{\alpha}^{(i+1)}, \boldsymbol{\delta}^{(i+1)}, \boldsymbol{\omega}^{(i)}) > 0. \quad (56)$$

The proofs of (54) and (56) have been provided in [65], while the proof of (55) is given in [19]. With these three inequalities, it can be ensured that the value of the objective function increases in each iteration and is upper bounded, which proves the convergence of our proposed algorithm.

The complexity of the proposed algorithm is dominated by the calculation of the gradients by (27), (30) and (38). In each iteration, the calculation of the gradient with respect to $\boldsymbol{\omega}$ needs K^3 calculations of the matrix determinant of size $N_r \times N_r$. The calculation of the gradient with respect to $\boldsymbol{\omega}$ and $\boldsymbol{\delta}$ requires K^3 calculations of matrix determinant, matrix inversion, and matrix multiplication of size $N_r \times N_r$. Thus, the overall complexity is given by $\mathcal{O}(K^3 N_r^3)$.

Algorithm 1 Gradient ascent algorithm maximizing the lower bound on channel capacity.

1: **BEGIN**

2: **Initialization:** Set a feasible initial solution $\boldsymbol{\delta}^{(0)}$, $\boldsymbol{\alpha}^{(0)}$ with $\alpha_k = \frac{1}{K}$, $\boldsymbol{\omega}^{(0)}$ with $\lambda_{k,j} = 1$, halting criterion $\varepsilon > 0$, and iteration index $i = 0$.

3: **Search and update $\boldsymbol{\alpha}$:** Compute the gradient $\nabla_{\boldsymbol{\alpha}} f(\boldsymbol{\alpha}^{(i)}, \boldsymbol{\delta}^{(i)}, \boldsymbol{\omega}^{(i)})$ and carry out the projection to obtain $\Delta \boldsymbol{\alpha}^{(i)}$. Use backtracking line search and update

$$\boldsymbol{\alpha}^{(i+1)} \leftarrow \boldsymbol{\alpha}^{(i)} + \eta_1^* \Delta \boldsymbol{\alpha}^{(i)}, \quad (50)$$

where η_1^* is computed by a line search process to satisfy the Wolfe conditions.

4: **Search and update $\boldsymbol{\delta}$:** Compute the gradient $\nabla_{\boldsymbol{\delta}} f(\boldsymbol{\alpha}^{(i+1)}, \boldsymbol{\delta}^{(i)}, \boldsymbol{\omega}^{(i)})$ and carry out the projection to obtain $\Delta \boldsymbol{\delta}^{(i)}$. Use backtracking line search and update

$$\boldsymbol{\delta}^{(i+1)} \leftarrow \cos(\eta_2^*) \boldsymbol{\delta}^{(i)} + \sin(\eta_2^*) \sqrt{NK} \frac{\Delta \boldsymbol{\delta}^{(i)}}{\|\Delta \boldsymbol{\delta}^{(i)}\|}, \quad (51)$$

where $\eta_2^* = \arg \max_{\eta_2 \in [0, \frac{\pi}{2}]} f(\boldsymbol{\alpha}^{(i+1)}, h(\boldsymbol{\delta}^{(i)}, \eta_2), \boldsymbol{\omega}^{(i)})$ and $h(\boldsymbol{\delta}^{(i)}, \eta_2) = \cos(\eta_2) \boldsymbol{\delta}^{(i)} + \sin(\eta_2) \sqrt{NK} \frac{\Delta \boldsymbol{\delta}^{(i)}}{\|\Delta \boldsymbol{\delta}^{(i)}\|}$.

5: **Search and update $\boldsymbol{\omega}$:** Compute $\boldsymbol{\beta}^{(i+1)} = \boldsymbol{\alpha}^{(i+1)} \otimes \mathbf{1}_{N_s}$ and the gradient $\nabla_{\boldsymbol{\omega}} f(\boldsymbol{\alpha}^{(i+1)}, \boldsymbol{\delta}^{(i+1)}, \boldsymbol{\omega}^{(i)})$ and carry out the projection to obtain $\Delta \boldsymbol{\omega}^{(i)}$. Use backtracking line search and update

$$\boldsymbol{\omega}^{(i+1)} \leftarrow \boldsymbol{\omega}^{(i)} + \eta_3^* \Delta \boldsymbol{\omega}^{(i)}, \quad (52)$$

and

$$i \leftarrow i + 1, \quad (53)$$

where η_3^* is computed by a line search process to satisfy the Wolfe conditions.

6: **Iteration:** Go to Step 6 if $\eta_1^* \|\Delta \boldsymbol{\alpha}^{(k)}\|_2 \leq \varepsilon \|\boldsymbol{\alpha}^{(k)}\|_2$, $\eta_2^* \|\Delta \boldsymbol{\delta}^{(k)}\|_2 \leq \varepsilon \|\boldsymbol{\delta}^{(k)}\|_2$ and $\eta_3^* \|\Delta \boldsymbol{\omega}^{(k)}\|_2 \leq \varepsilon \|\boldsymbol{\omega}^{(k)}\|_2$, or go to Step 2, otherwise.

7: **Output:** Normalize the optimized $\boldsymbol{\delta}^*$ by $\phi_{k,n}^* = \frac{\phi_{k,n}^*}{|\phi_{k,n}^*|}$, $k \in \{1, \dots, K\}$, $n \in \{1, 2, \dots, N\}$; use the normalized $\boldsymbol{\delta}^*$ and the optimized $\boldsymbol{\omega}^*$ to obtain $\boldsymbol{\Psi}$ and \mathcal{Q} ; output the optimized $\boldsymbol{\alpha}^*$ as the final outcome.

8: **END**

IV. UPPER BOUND BASED OPTIMIZATION

Except for the lower bound on channel capacity, we can also resort to maximizing the upper bound derived in (19) to obtain the optimized phase shift set, activation distribution $\boldsymbol{\alpha}$, and transmit covariance matrix set. Detailed optimization procedure and analysis based on the upper bound are presented in this section.

A. Problem Formulation

Observing from the expression of the upper bound on channel capacity given in (19), we cannot obtain the optimal phase shift set through traditional gradient optimization approaches. This is because the gradient of each ϕ_k is independent and will

converge to the same optimal value by an adequate number of iterations. In this case, the phase shift set is composed of K identical phase shift matrices, which will not provide additional information anymore. **To avoid this dilemma, we design the optimal phase shift set to maximize the equivalent channel capacity C^U , which maximizes the mutual information between re-shaped transmitted signal vector $\hat{\mathbf{x}}$ defined in (6) and the received signal vector \mathbf{y} :**

$$C^U = \max_{\mathbb{E}\{\text{Tr}(\hat{\mathbf{x}}\hat{\mathbf{x}}^H)\} \leq 1} \mathbf{I}(\hat{\mathbf{x}}; \mathbf{y}). \quad (57)$$

By reshaping the channel matrices, phase shift matrix and, the transmitted data stream derived in (2)-(5), C^U can be explicitly expressed as

$$C^U = \log_2 \det \left(\mathbf{I}_{N_r \times N_r} + \gamma (\hat{\mathbf{H}}_d + \hat{\mathbf{H}}_2 \Theta \hat{\mathbf{H}}_1) \mathbf{Q} (\hat{\mathbf{H}}_d + \hat{\mathbf{H}}_2 \Theta \hat{\mathbf{H}}_1)^H \right) \quad (58)$$

where \mathbf{Q} is the unit constrained transmit covariance matrix. We can see that the equivalent channel capacity C^U is related to the reshaped phase shift matrix Θ and transmit covariance matrix \mathbf{Q} . Therefore, both Θ and \mathbf{Q} should be jointly optimized when maximizing channel capacity. Since the channel capacity C^U is not related to distribution probability α and the transmit covariance matrix set \mathcal{Q} , the formed optimization problem relying on C^U can be written as

$$\begin{aligned} (\mathbf{P5}) : \quad & \max_{\Theta, \mathbf{Q}} C^U \\ \text{s.t.} : \quad & |\psi_{i,i}| = 1, i = 1, \dots, NK, \\ & \text{Tr}(\mathbf{Q}) \leq 1, \\ & \mathbf{Q} \succeq 0, \end{aligned} \quad (59)$$

where $\psi_{i,i}$ denotes the i th diagonal element of Θ . The main purpose of the proposed problem (P5) is to find the reshaped phase shift matrix Θ , where each reflecting coefficient is constrained by a uni-modular constraint. Although we do not need the optimized \mathbf{Q} in the following optimization process, \mathbf{Q} is still needed to be jointly optimized with Θ to achieve the maximum channel capacity C^U . By finding Θ , we can subsequently obtain each Φ_k . Since the optimization problem (P5) is independent of the activation probability α and the covariance matrix set \mathcal{Q} , the obtained optimal solution Θ^* from (P5) generates the optimality of the phase shift, which do not need to be jointly optimized with α and \mathcal{Q} .

With the optimized phase shift set, we can resort to the upper bound on channel capacity given in (19) to optimize the distribution probability α and the transmit covariance matrix set \mathcal{Q} . By transforming the optimization problem regarding the transmit covariance matrix set \mathcal{Q} to the problem related to the vector ω in the same way as applied for the lower bound based optimization, the upper bound based optimization problem can be formulated as

$$\begin{aligned} (\mathbf{P6}) : \quad & \max_{\alpha, \omega} \mathcal{R}^U(\alpha, \omega) \\ \text{s.t.} : \quad & \mathbf{1}_K^T \alpha = 1, \\ & \alpha \geq \mathbf{0}_K, \\ & \beta^T \omega = 1, \\ & \omega > 0. \end{aligned} \quad (60)$$

The constraints imposed on activation probability α and power allocation vector ω is same as that in the lower bound based optimization problem, i.e., (P3). Overall, the optimal reflecting patterns Ψ can be directly obtained from the solution to optimization problem (P5), while the activation probability α and the covariance matrix set \mathcal{Q} can be directly obtained from the solution to optimization problem (P6). In the following, we give the solutions to optimization problems (P5) and (P6), respectively.

B. Transmitter and Reflector Designs

Based on the formulated optimization problems above, (P5) should first be solved by alternatively optimizing the phase shift matrix Θ and transmit covariance matrix \mathbf{Q} . The covariance matrix \mathbf{Q} is then optimized by fixing the phase shift matrix Θ , and the phase shift matrix Θ is subsequently optimized given the covariance matrix \mathbf{Q} . The optimized phase shift matrix Θ is used to generate the phase shift set and solve (P6) to produce the optimal solutions of α and ω .

Given phase shift matrix Θ , the regenerated transmit covariance matrix \mathbf{Q} by (P5) is also given by the eigenmode transmission. Specifically, the re-shaped channel matrix can be truncated singular value decomposed as $\mathbf{H} = \hat{\mathbf{H}}_d + \hat{\mathbf{H}}_2 \Theta \hat{\mathbf{H}}_1 = \mathbf{U} \mathbf{\Lambda} \mathbf{V}^H$, where $\mathbf{U} \in \mathbb{C}^{N_r \times R}$ denotes the left-singular vectors of \mathbf{H} , $\mathbf{V} \in \mathbb{C}^{N_t \times R}$ represents the right-singular vectors of \mathbf{H} , and $R = \text{rank}(\mathbf{H})$ is the maximum number of data streams that can be transmitter over \mathbf{H} . As a result, the optimal \mathbf{Q} is given by

$$\mathbf{Q} = \mathbf{V} \text{Diag}(\mathbf{p}) \mathbf{V}^H, \quad (61)$$

where $\mathbf{p} \in \mathbb{R}^{R \times 1}$ represents the optimal power allocation vector following the water-filling strategy. Therefore, the i th element of \mathbf{p} , denoted as $p_i, \forall i \in \{1, \dots, R\}$, can be determined as

$$p_i = \left(\frac{1}{\xi \ln 2} - \frac{1}{\gamma [\mathbf{\Lambda}]_{i,i}^2} \right)^+, \quad (62)$$

with ξ satisfying the following constraint:

$$\sum_{i=1}^R \left(\frac{1}{\xi \ln 2} - \frac{1}{\gamma [\mathbf{\Lambda}]_{i,i}^2} \right)^+ = 1, \quad (63)$$

where $[\mathbf{\Lambda}]_{i,i}$ denote the i th singular value selected from \mathbf{H} . Having obtained covariance matrix \mathbf{Q} , the phase shift can be optimized through the following steps.

To facilitate the following analysis and description, we denote $\mathbf{H}' = \mathbf{H} \mathbf{U}_Q \Sigma_Q^{\frac{1}{2}}$ with $\mathbf{Q} = \mathbf{U}_Q \Sigma_Q \mathbf{U}_Q^H$ as the truncated singular decomposition of \mathbf{Q} ; we also denote $\hat{\mathbf{h}}_{2,i}$ as the i th column of $\hat{\mathbf{H}}_2$ and $\hat{\mathbf{h}}'_{1,i} = \Sigma_Q^{\frac{1}{2}} \mathbf{U}_Q^H \hat{\mathbf{h}}_{1,i}$ with $\hat{\mathbf{h}}_{1,i}^H$ as the i th row

of $\hat{\mathbf{H}}_1$. Consequently, we can derive the following relations among these matrices/vectors:

$$\begin{aligned} \hat{\mathbf{A}}_i &= \mathbf{I}_{N_r \times N_r} + \gamma \left(\mathbf{H}' + \sum_{j=1, j \neq i}^{KN} \psi_{j,j} \hat{\mathbf{h}}_{2,j} \left(\hat{\mathbf{h}}'_{1,j} \right)^H \right) \\ &\times \left(\mathbf{H}' + \sum_{j=1, j \neq i}^{KN} \psi_{j,j} \hat{\mathbf{h}}_{2,j} \left(\hat{\mathbf{h}}'_{1,j} \right)^H \right)^H \\ &+ \gamma \hat{\mathbf{h}}_{2,i} \left(\hat{\mathbf{h}}'_{1,i} \right)^H \hat{\mathbf{h}}'_{1,i} \hat{\mathbf{h}}_{2,i}, \end{aligned} \quad (64)$$

and

$$\hat{\mathbf{B}}_i = \gamma \hat{\mathbf{h}}_{2,i} \left(\hat{\mathbf{h}}'_{1,i} \right)^H \left(\mathbf{H}'^H + \sum_{j=1, j \neq i}^{KN} \hat{\mathbf{h}}'_{1,j} \hat{\mathbf{h}}_{2,j}^H \psi_{j,j}^H \right). \quad (65)$$

Therefore, the phase shift matrix Θ with given \mathbf{Q} can be optimized through the same algorithm applied in [29], and the solution is given by

$$\psi_{i,i} = \exp \{-j \arg\{\varphi_i\}\}, \quad (66)$$

where φ_i denotes the sole non-zero eigenvalue of $\hat{\mathbf{A}}_i^{-1} \hat{\mathbf{B}}_i$. Therefore, \mathbf{Q} and Θ can be alternatively optimized through (61) and (66) until a certain level of convergence has been reached. As stated before, we only adopt the optimized Θ to generate the optimized phase shift set Ψ and solve optimization problem (P6) related to α and ω . According to this optimization strategy, matrix \mathbf{Q} is only obtained to derive the optimized Θ and is not related to the following optimization process.

The solution to optimization problem (P6) has been proposed in [65], where the optimal α^* and λ_k^* are given by

$$\alpha_k^* = \frac{\det(\mathbf{D}_k^*)}{\sum_{k=1}^K \det(\mathbf{D}_k^*)}, \quad k = 1, \dots, K \quad (67)$$

and

$$\lambda_k^* = [\lambda_{k,1}^*, \dots, \lambda_{k,N_s}^*]^T, \quad (68)$$

where $\lambda_{k,j}^*$ is given by

$$\lambda_{k,j}^* = \left(\frac{1}{\xi_k \ln 2} - \frac{1}{\gamma \sigma_{k,j}^2} \right)^+, \quad j = 1, \dots, N_s, \quad (69)$$

and $\sigma_{k,j}$ is the j th element of $\Sigma_k = \text{diag}(\sigma_{k,1}, \dots, \sigma_{k,N_s})$, which is the diagonal matrix composed of non-zeros singular values of \mathbf{H}_k defined in (23); ξ_k is an optimization parameter satisfying

$$\sum_{j=1}^{N_s} \left(\frac{1}{\xi_k \ln 2} - \frac{1}{\gamma \sigma_{k,j}^2} \right)^+ = 1. \quad (70)$$

With the optimized reflecting patterns Ψ derived from optimization problem (P5), and the activation probability α and the covariance matrix set \mathcal{Q} derived from optimization problem (P6), the proposed optimization algorithm can be summarized in **Algorithm 2**.

Algorithm 2 Proposed algorithm maximizing the upper bound on channel capacity.

- 1: **BEGIN**
- 2: **Initialization:** Set a feasible initial solution $\Theta^{(0)}$, $\mathbf{Q}^{(0)}$, and iteration index $j = 0$.
- 3: **Update Θ and \mathbf{Q} :** Compute $\hat{\mathbf{A}}_i^{(j)}$ and $\hat{\mathbf{B}}_i^{(j)}$ according to (64) and (65) from $i = 1$ to $i = NK$ and obtain the optimal $\Theta^{(j+1)}$ and $\mathbf{Q}^{(j+1)}$ according to (66) and (61).
- 4: **Iteration:** Go to Step 4 if the convergence criterion is met, or let $j \leftarrow j + 1$ and go to Step 2, otherwise.
- 5: **Optimize probability distribution and transmit covariance matrix:** Use the optimized Θ^* to obtain Ψ^* and compute the optimal power allocation $\lambda_{k,j}^*$ according to (69) based on the optimized phase shift set. Use derived λ_k^* and optimized phase shift set to generate \mathbf{Q}_k^* as well as \mathbf{D}_k^* , and compute α^* according to (67).
- 6: **Output:** Output the optimized Ψ^* , \mathcal{Q}^* , and α^* .
- 7: **END**

C. Convergence and Complexity Analysis

As stated in **Algorithm 2**, the optimal phase shift set is obtained by iteratively updating \mathbf{Q} and solving optimization problem (P5). This result is then followed by the optimized probability distribution and transmit covariance matrix obtained by (67) and (68) through the solution to (P6). Since the optimal activation probability α^* and optimal transmit covariance matrix set \mathcal{Q}^* are closed-form solutions, the convergence of **Algorithm 2** is mainly dependent on the phase shift set optimization. Because the phase shift set optimization algorithm makes the objective function of (P5) non-decreasing and bounded over iterations, the monotonic nature of the phase shift optimization process is guaranteed. Also, any limit point of Θ and \mathbf{Q} in each iteration satisfies the Karush-Kuhn-Tucker (KKT) condition, because the objective function of (P5) is differentiable, and both Θ and \mathbf{Q} are not coupled in the constraints. Therefore, **Algorithm 2** is guaranteed to converge to at least a local optimum.

Based on the description given above, we can easily find that the complexity of **Algorithm 2** is dominated by the complexity incurred by the phase shift set optimization, the water-filling power allocation, and the computation of α . It can be shown that the complexity of optimizing Ψ is $\mathcal{O}(K^2 N_r^3)$, and the complexity of power allocation is $\mathcal{O}(KN_s)$. The optimization of α requires K calculations of matrix determinate of size $N_r \times N_r$, leading to the complexity of $\mathcal{O}(KN_r^3)$. Therefore, it can be summarized that the overall complexity of **Algorithm 2** is $\mathcal{O}(K^2 N_r^3)$ by neglecting the low-order terms. Here, it should be noted that no iteration is needed for optimizing α and \mathcal{Q} .

V. NUMERICAL RESULTS AND DISCUSSION

In this section, numerical results are presented to illustrate and discuss the performance of the proposed RIS assisted MIMO system with $N_t = 2$ and $N_r = 4$, in which information is carried by both of the reflecting patterns at the RIS and by the transmitted signals.

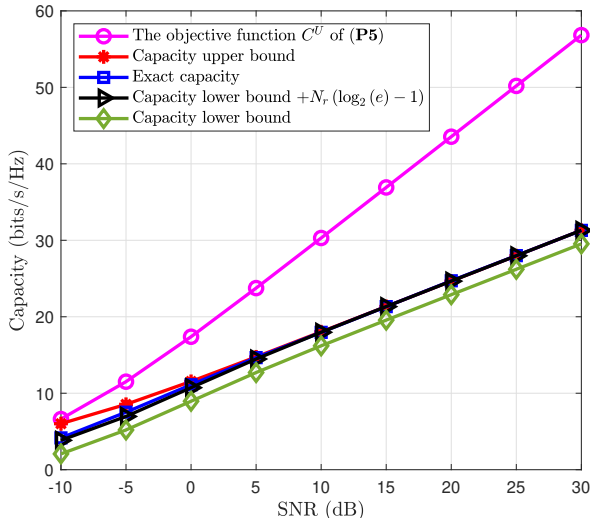


Fig. 2: Illustration of channel capacity and various capacity bounds, given $N = 20$, and $K = 6$.

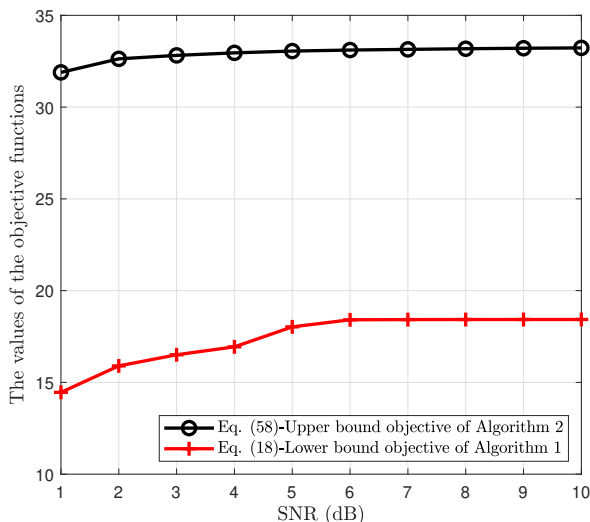


Fig. 3: Performance and convergence of **Algorithm 1** and **Algorithm 2**, given $N = 20$, and $K = 6$.

First, we validate the derived bounds on channel capacity from **Theorem 1** in Fig. 2 by plotting the exact capacity, capacity bounds of the proposed system with $N = 20$ and $K = 6$. As shown in Fig. 2, the lower bound adding a constant gap $N_r (\log_2(e) - 1)$ is tight over the whole SNR regime, while the upper bound is only tight in high SNR regime. However, the objective function C^U of the optimization problem **(P5)** is not tight but provides an easy-to-obtain solution of the reflecting patterns. Due to the tightness of the derived capacity bounds, it equivalently indicates that the exact capacity optimization can be achieved by maximizing the lower bound and upper bound on channel capacity. This verifies the effectiveness of the proposed algorithms.

To show the effectiveness as well as the convergence of the

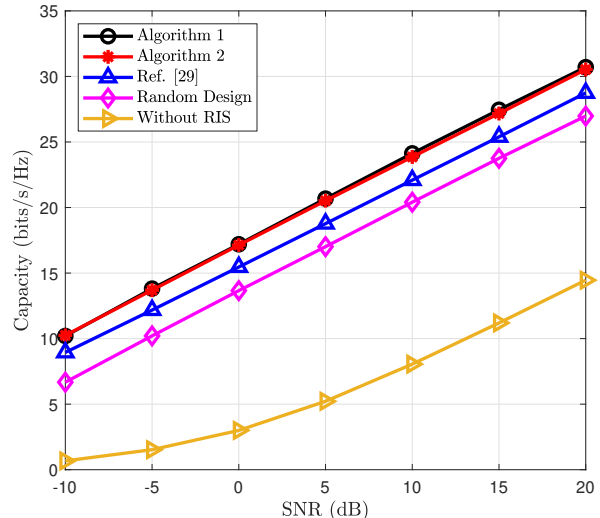


Fig. 4: Performance of **Algorithm 1** and **Algorithm 2**, given $N = 20$ and $K = 6$.

proposed algorithms, we evaluate the performance of $\alpha^{(i)}$, $\delta^{(i)}$, and $\omega^{(i)}$ of iteration i in **Algorithm 1** and $\Theta^{(j)}$ in **Algorithm 2**, given the SNR of 1 dB, $N = 20$, and $K = 6$. We calculate the values of the objective functions $R^L(\alpha^{(i)}, \delta^{(i)}, \omega^{(i)})$ given in **(P4)** and $C^U(\Theta^{(j)})$ defined in **(P5)**. The simulation results given in Fig. 3 demonstrate that both algorithms converge fast, where the iteration numbers for both algorithms are less than 10. It also can be observed that **Algorithm 2** converges faster than **Algorithm 1**.

In the following, the two proposed algorithms are compared and investigated under various system setups operating in standard Rayleigh channels. The algorithm proposed in [29] aiming to maximize the classic RIS assisted MIMO system is adopted in the simulations as the benchmark, which is denoted as 'Ref. [29]' in the following data plots. In these data plots, the exact capacity achieved with parameters updated through **Algorithm 1** based on the lower bound on channel capacity is denoted as 'Algorithm 1', while the one with parameters updated through **Algorithm 2** based on the upper bound on channel capacity is denoted as 'Algorithm 2'. Moreover, in the following figures, 'Random design' denotes the case considering the value of each reflecting element randomly chosen from $(0, 2\pi]$, while 'Without RIS' represents the conventional point-to-point MIMO system without the assistance of RIS.

In Fig. 4, we compare the capacity yielded by our proposed system design and algorithms to the benchmark proposed in [29] under the same channel realizations, given $N = 20$ and $K = 6$. We can see from the figure that our proposed designs achieve higher capacities than that produced by the benchmark. This verifies that the capacity of the RIS assisted MIMO system can be improved by incorporating the active information delivering functionality at the RIS. In particular, there are about 2-3 dB performance gains by employing 6 reflecting patterns at the RIS. Moreover, there is only a tiny difference between the capacities yielded by optimizing the lower and upper bounds. Hence, it can be confirmed that the

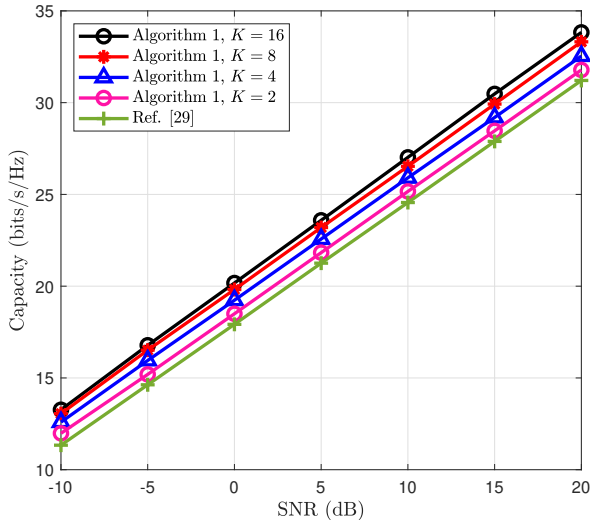


Fig. 5: Effects of the number of reflecting patterns K on the channel capacity, given $N = 32$.

lower bound based algorithm would be better because of its complexity advantage. Compared with K identical random phase shift patterns at the RIS with the optimal transmit covariance matrix and reflecting patterns activation probability, we can observe that an appropriate phase shift design for the RIS is crucial for achieving a higher capacity. More than 5 dB gains can be brought by the well designed reflecting patterns. Fig. 4 also proves that the system performance can be greatly enhanced by involving an RIS, which brings about 20 dB performance gains in comparison with the canonical MIMO system without the assistance of RIS.

The effect of the number of reflecting patterns on the capacity is shown in Fig. 5, given $N = 32$. As illustrated in this figure, the channel capacity can be greatly improved by increasing the number of reflecting patterns K , where 2^n reflecting patterns bring about n dB gains in our proposed system compared to the classic RIS assisted MIMO system. The numerical results also verify that the channel capacity of the proposed system can be improved by incorporating more reflecting patterns with optimized activation probabilities. However, more available reflecting patterns at the RIS also lead to higher computational complexity. Therefore, there exists a trade-off between capacity and computational complexity.

In Fig. 6, we show the influence of the number of reflecting elements N on the channel capacity, given $K = 6$. There are an approximate-6-dB performance gain with 20 more reflecting elements and an approximate-12-dB performance gain performance gains with 40 more reflecting elements if the transmit covariance matrix at the transmitter, reflecting patterns, and reflecting patterns activation probabilities have been jointly optimized in a proper way. In Fig. 7, we plot the achievable capacity versus N to obtain more insights from the relationship between the achievable capacity and the number of reflecting elements. To summarize, the channel capacity can be significantly increased by allowing more reflecting elements

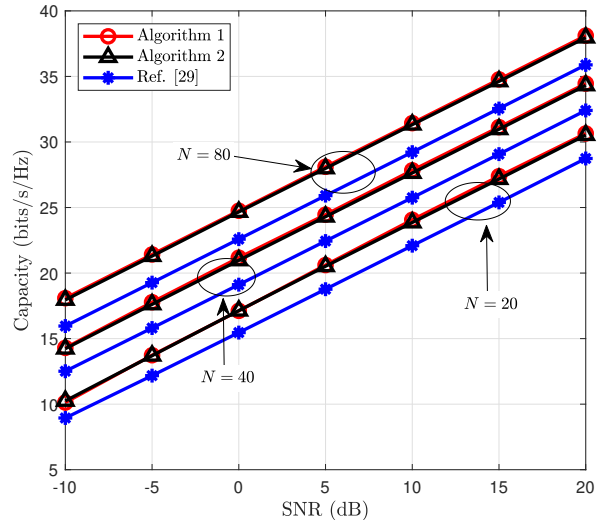


Fig. 6: Effects of the number of reflecting elements N on the channel capacity, given $K = 6$.

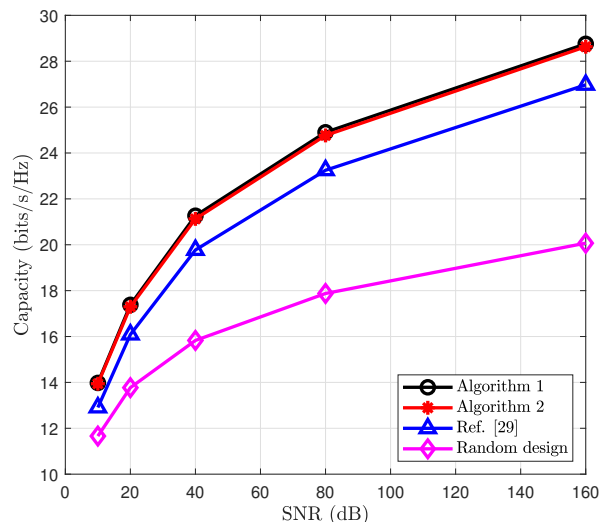


Fig. 7: Effects of the number of reflecting elements N on the channel capacity, given $K = 4$ and $\gamma = 1$ dB.

at the RIS, where 10×2^n reflecting elements brings about a $4n$ -dB gain to the proposed system.

Although the perfect CSI acquisition is assumed in the previous simulations for simplicity, the challenge of accurate channel estimation in practice should never be underestimated, especially when the number of elements/components in communication systems increases [68]. Considering this practical challenge, the capacity of the proposed system and designs subject to channel estimation errors is shown in Fig. 8. The channel estimation errors are modeled by $\mathbf{H}'_i = \mathbf{H}_i + \mathbf{H}_e$, where $i \in \{1, 2, d\}$, and \mathbf{H}_e is the error matrix whose entries follow an independent and identically distributed (i.i.d.) circularly symmetric complex Gaussian distribution $\mathcal{CN}(0, \sigma_e)$ [57], [69]. As expected, the channel capacity decreases with

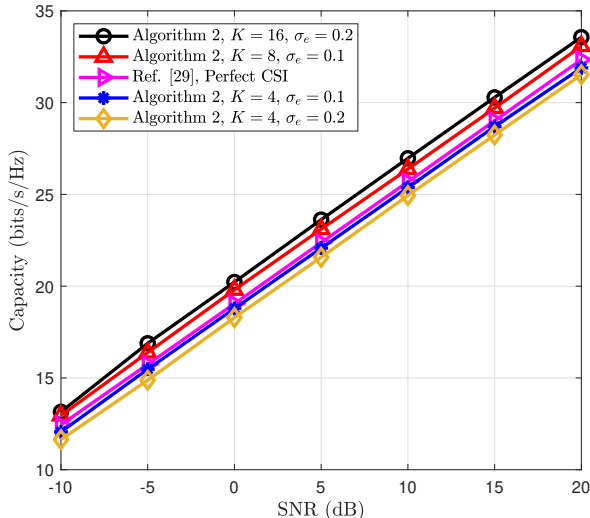


Fig. 8: Comparison of the capacities yielded by different optimization algorithms in the presence of imperfect CSI, given $N = 40$.

the increasing severity of channel estimation errors, and the capacity of the proposed system corresponding to $K = 4$ is even lower than that yielded by the benchmark. Fortunately, the channel capacity of the proposed system can still be raised by increasing the number of reflecting patterns K under the assumption of imperfect CSI. The proposed system with $K = 8$ and estimation error variance $\sigma_e = 0.1$ achieves a higher capacity than that of the classic RIS assisted MIMO system with perfect CSI. What is better, the proposed system with $K = 16$ and estimation error variance $\sigma_e = 0.2$ has even a 2 dB performance gain over the classic RIS assisted MIMO system. These results imply that the capacity degradation caused by imperfect CSI can be compensated by the active information delivery functionality of RIS.

It should be noted that although the proposed symbiotic system achieves a higher capacity than the conventional one without additional information carried by reflection patterns, the fronthaul link in the symbiotic systems also consumes more capacity compared to the conventional one. Therefore, there is a trade-off between the achieved capacity and the resource consumed by the fronthaul link. To better understand it, we take the transmission between the transmitter and RIS controller into consideration and assume a simplified transmission process within a fixed channel coherence time T_{coh} . The whole transmission process is composed of two parts. One is the fronthaul transmission phase between the transmitter and the controller, and another is the information transmission phase between the transmitter and receiver assisted by an RIS. Specifically, well-designed reflecting patterns will be sent from the transmitter to the controller once they are computed at the transmitter through the proposed algorithms. Assuming the phase of each reflecting element to be continuous during the computation process, they will be quantized as b phase resolution in number of bits [70], that is, $(\Phi_k)_{n,n} \in \mathcal{F} \triangleq$

$\left\{ \exp\left(\frac{j2\pi m}{2^b}\right) \right\}_{m=0}^{2^b-1}$, before transmission. In the conventional system, there is only one reflecting pattern with Nb bits to be transmitted, while there are K reflecting patterns with NKb bits to be transmitted in the symbiotic system. Moreover, since different activation probabilities are assigned to different reflecting matrices, the activation pattern is also needed to be transmitted to the controller before information transmission. Here, we assign b_{act} bits to each activation probability, which indicates that we need to transmit additional $(K-1)b_{act}$ bits from the transmitter to the controller. Denoting the channel between the transmitter and the controller as \mathbf{H}_{con} , the channel capacity between the transmitter and the controller can be calculated as

$$C_{con} = \log_2 \det \left(\mathbf{I}_{N_r^{con} \times N_r^{con}} + \gamma_{con} \mathbf{H}_{con} \mathbf{Q}_{con} \mathbf{H}_{con}^H \right), \quad (71)$$

where N_r^{con} is the receive antenna at the controller; \mathbf{Q}_{con} is the transmit signal covariance matrix for the control link, which can be optimally solved by the water-filling algorithm, and γ_{con} represents the average receive SNR for the fronthaul link. Therefore, part of the coherence time is used for the fronthaul transmission, while the remaining time is assigned to the information transmission between transmitter and receiver. Consequently, the total bits that can be transmitted within the channel coherence time can be calculated as

$$I = CB \left[T_{coh} - \frac{NKb + (K-1)b_{act}}{C_{con}B} \right], \quad (72)$$

where B is the bandwidth, and C is the achievable channel capacity of the proposed system. Similarly, the total bits can be transmitted through the conventional RIS-assisted MIMO system can be obtained as

$$I_{bench} = C_{bench}B \left[T_{coh} - \frac{Nb}{C_{con}B} \right], \quad (73)$$

where C_{bench} is the achievable channel capacity of the conventional RIS-assisted MIMO system with the RIS design proposed in [29]. Here, we neglect the propagation time between the transmitter the receiver since they are the same for both systems.

We can observe from (72) and (73) that the proposed symbiotic system yield less time for information transmission compared to the conventional system within the same channel coherence time due to the transmission of more than one reflecting patterns and the activation patterns. However, since the proposed system can achieve higher channel capacity than the conventional ones, the proposed system is possible to transmit more bits within the channel coherence time when the achievable capacity is high enough or the fronthaul transmission time only occupies a small part of the whole coherence time. Theoretically, we can derive the condition that the proposed system outperforms the conventional system:

$$T_{coh} > \frac{C [NKb + (K-1)b_{act}] - NbC_{bench}}{C_{con}B [C - C_{bench}]}. \quad (74)$$

To verify the above reasoning and derivation, we conduct several experiments and present the results in Figs. 9 and 10. We can observe that the conventional system outperforms the

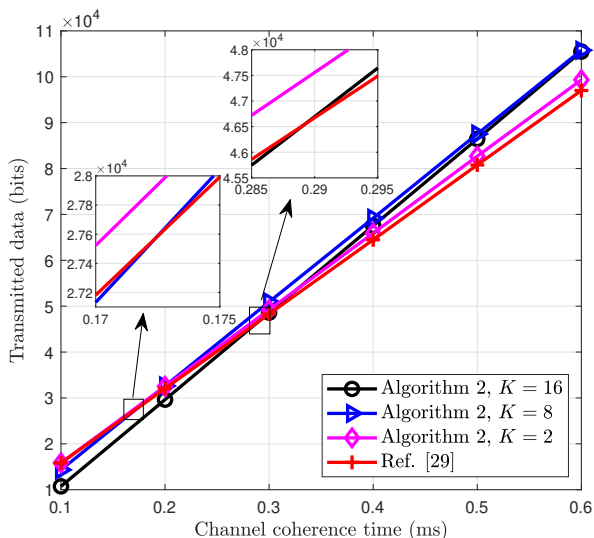


Fig. 9: Effects of the number of reflecting patterns on total transmitted data, given $N = 20$ and $\gamma = \gamma_{con} = 1$ dB, $N_r^{con} = 4$, $b_{act} = 8$ bits.

system. Meanwhile, through simulation results, it has also been found that the channel capacity can be raised by increasing the number of reflecting patterns. Furthermore, the robustness of the proposed system against imperfect CSI has also been verified through simulations. This outcome opens up the opportunity to effectively compensate the capacity degradation caused by channel estimation errors by increasing the number of reflecting patterns.

symbiotic system when the channel coherence time is short, but the symbiotic system performs better once T_{coh} surpasses a certain value, which is in agreement with the condition we derived before. Increasing the number of reflecting patterns yields a negative impact on the information transmission when the channel coherence time is small. Also, the total transmitted data in the symbiotic system increases firstly but then decreases as K increases. This is because the increased achievable capacity can compensate for the time used for additional fronthaul transmission due to the increase of K when K is small. However, as K increases, the fronthaul transmission requires most of the channel coherence time, and therefore it transmits fewer data due to the limited time, albeit with an improved achievable capacity. In summary, the symbiotic system would be a better choice when the channel coherence time exceeds a threshold (c.f. (74)), which is dominated by several system parameters, including channel condition, the number of reflecting patterns and reflecting elements.

VI. CONCLUSION AND DISCUSSION

In this paper, we proposed a novel RIS assisted MIMO symbiotic system and studied its channel capacity through several derived bounds. The proposed RIS assisted MIMO system is capable of not only assisting in delivering the information from the transmitter to the receiver but also activating different reflecting patterns for piggybacking additional information. Two optimization algorithms were proposed to optimize the reflecting patterns, reflecting patterns activation probability, and transmit covariance matrix based on the lower bound and upper bound on the exact channel capacity of the proposed system. Simulation results demonstrated that our proposed system equipped with the optimization designs can achieve a higher capacity than the classic RIS MIMO

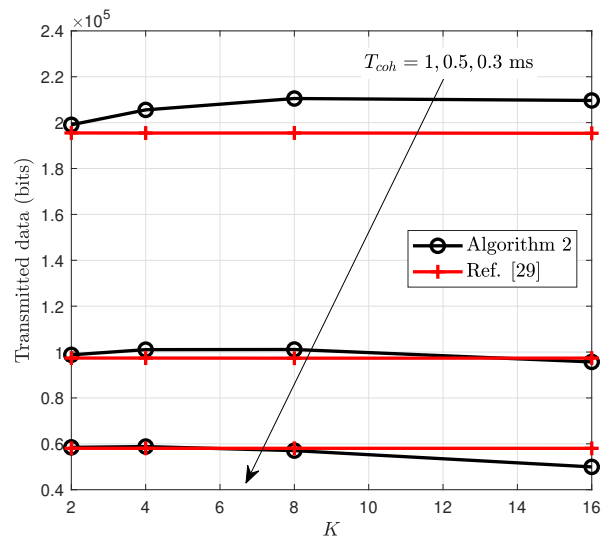


Fig. 10: Effects of the channel coherence time T_{coh} on total transmitted data, given $N = 40$ and $\gamma = \gamma_{con} = 1$ dB, $N_r^{con} = 4$, $b_{act} = 8$ bits.

This initial study on the classical point-to-point communication system also leaves us the following research directions:

- The control link in the considered symbiotic system costs more capacity than a conventional RIS assisted system with only a single fixed reflection pattern. Therefore, there is a trade-off between the achieved capacity and the resource consumed by the control link, which is worth investigating in depth.
- The analysis and design can be extended to multi-user application scenarios to study and minimize the impact of the mutual interference. In addition, the proposed communication paradigm integrating with other advanced communication techniques, e.g., non-orthogonal multiple access techniques, millimeter-wave or THz communications, and non-terrestrial communications, are expected to be investigated.
- Continuous controllable phase shifts of RIS elements are assumed in this work. As an ideal assumption, it is necessary to analyze the effect of hardware imperfections in a realistic phase controller that can only switch to a finite number of phase shifts.
- No specific channel model is assumed in this work. The channel assumption with certain distributions and properties considering spatial correlation, like Rayleigh, Rician, and Nakagami-m, can facilitate the analysis and design, and thereby worth further investigating.

- Our current analysis is based on the perfect CSI availability assumption, while only imperfect CSI or statistical CSI is available in reality. The impacts of imperfect and statistical CSI on the achievable capacity remain unknown and should be explored.

APPENDIX A
PROOF OF **THEOREM 1**

By substituting $f(\mathbf{y})$ given in (15) into (14), the capacity term can be re-expressed as follows:

$$\mathcal{R}(\boldsymbol{\alpha}, \boldsymbol{\Psi}, \mathcal{Q}) = -N_r \log_2(\pi e) - \sum_{k=1}^K \alpha_k \underbrace{\int_{\mathbb{C}^{N_r}} f(\mathbf{y}|k) \log_2 \left(\sum_{l=1}^K \alpha_l f(\mathbf{y}|l) \right) d\mathbf{y}}_I. \quad (75)$$

Since $\log_2(\cdot)$ is a concave function, and according to the Jensen's inequality, the integration operation of I is upper bounded by

$$I = \int_{\mathbb{C}^{N_r}} f(\mathbf{y}|k) \log_2 \left(\sum_{l=1}^K \alpha_l f(\mathbf{y}|l) \right) d\mathbf{y} \leq \log_2 \left(\sum_{l=1}^K \alpha_l \int_{\mathbb{C}^{N_r}} f(\mathbf{y}|k) f(\mathbf{y}|l) d\mathbf{y} \right). \quad (76)$$

Substituting the conditional distribution $f(\mathbf{y}|k)$ given in (16) into (76), we obtain the upper bound on I as

$$\begin{aligned} I &\leq \log_2 \left(\sum_{l=1}^K \alpha_l \int_{\mathbb{C}^{N_r}} \frac{1}{\pi^{N_r} \det(\mathbf{D}_k)} \exp(-\mathbf{y}^H \mathbf{D}_k^{-1} \mathbf{y}) \right. \\ &\quad \times \left. \frac{1}{\pi^{N_r} \det(\mathbf{D}_l)} \exp(-\mathbf{y}^H \mathbf{D}_l^{-1} \mathbf{y}) d\mathbf{y} \right) \\ &= \log_2 \left(\sum_{l=1}^K \frac{\alpha_l}{\pi^{N_r} \det(\mathbf{D}_k \mathbf{D}_l) \det(\mathbf{D}_k^{-1} + \mathbf{D}_l^{-1})} \right. \\ &\quad \times \left. \int_{\mathbb{C}^{N_r}} \frac{\det(\mathbf{D}_k^{-1} + \mathbf{D}_l^{-1})}{\pi^{N_r}} \exp(-\mathbf{y}^H (\mathbf{D}_k^{-1} + \mathbf{D}_l^{-1}) \mathbf{y}) d\mathbf{y} \right) \\ &= \log_2 \left(\sum_{l=1}^K \frac{\alpha_l}{\pi^{N_r} \det(\mathbf{D}_k \mathbf{D}_l) \det(\mathbf{D}_k + \mathbf{D}_l)} \right) \\ &= -N_r \log_2(\pi) + \log_2 \left(\sum_{l=1}^K \frac{\alpha_l}{\det(\mathbf{D}_k + \mathbf{D}_l)} \right). \end{aligned} \quad (77)$$

By applying the upper bound on I , we can obtain the lower bound on the channel capacity as

$$\begin{aligned} \mathcal{R}(\boldsymbol{\alpha}, \boldsymbol{\Psi}, \mathcal{Q}) &\geq -N_r \log_2(\pi e) + \sum_{k=1}^K \alpha_k N_r \log_2(\pi) \\ &\quad - \sum_{k=1}^K \alpha_k \log_2 \left(\sum_{l=1}^K \frac{\alpha_l}{\det(\mathbf{D}_k + \mathbf{D}_l)} \right) \\ &= -N_r \log_2(e) - \sum_{k=1}^K \alpha_k \log_2 \left(\sum_{l=1}^K \frac{\alpha_l}{\det(\mathbf{D}_k + \mathbf{D}_l)} \right), \end{aligned} \quad (78)$$

which is exactly the lower bound $\mathcal{R}^L(\boldsymbol{\alpha}, \boldsymbol{\Psi}, \mathcal{Q})$ given in (18).

To prove the upper bound on the channel capacity, we should look back on the integration operation of I in (75). Since $\log_2(\cdot)$ is a non-decreasing function and $\sum_{l=1}^K \alpha_l f(\mathbf{y}|l) \geq \alpha_k f(\mathbf{y}|k)$, it is obvious by the monotonicity of logarithmic functions that the following inequality is valid:

$$\log_2 \left(\sum_{l=1}^K \alpha_l f(\mathbf{y}|l) \right) \geq \log_2(\alpha_k f(\mathbf{y}|k)), \quad (79)$$

which leads to

$$\begin{aligned} I &= \int_{\mathbb{C}^{N_r}} f(\mathbf{y}|k) \log_2 \left(\sum_{l=1}^K \alpha_l f(\mathbf{y}|l) \right) d\mathbf{y} \\ &\geq \int_{\mathbb{C}^{N_r}} f(\mathbf{y}|k) \log_2(\alpha_k f(\mathbf{y}|k)) d\mathbf{y} \\ &= \log_2(\alpha_k) + \int_{\mathbb{C}^{N_r}} f(\mathbf{y}|k) \log_2(f(\mathbf{y}|k)) d\mathbf{y} \\ &= \log_2(\alpha_k) + \int_{\mathbb{C}^{N_r}} \frac{1}{\pi^{N_r} \det(\mathbf{D}_k)} \exp(-\mathbf{y}^H \mathbf{D}_k^{-1} \mathbf{y}) \\ &\quad \times \log_2 \left(\frac{1}{\pi^{N_r} \det(\mathbf{D}_k)} \exp(-\mathbf{y}^H \mathbf{D}_k^{-1} \mathbf{y}) \right) d\mathbf{y} \\ &= \log_2(\alpha_k) + \log_2 \left(\frac{1}{\pi^{N_r} \det(\mathbf{D}_k)} \right) - \frac{1}{\ln 2} \mathbb{E}(\mathbf{y}^H \mathbf{D}_k^{-1} \mathbf{y}), \end{aligned} \quad (80)$$

where the expectation $\mathbb{E}(\mathbf{y}^H \mathbf{D}_k^{-1} \mathbf{y})$ can be calculated as

$$\begin{aligned} \mathbb{E}(\mathbf{y}^H \mathbf{D}_k^{-1} \mathbf{y}) &= \mathbb{E}(\text{tr}(\mathbf{y}^H \mathbf{D}_k^{-1} \mathbf{y})) \\ &= \mathbb{E}(\text{tr}(\mathbf{y} \mathbf{y}^H \mathbf{D}_k^{-1})) \\ &= \text{tr}(\mathbb{E}(\mathbf{y} \mathbf{y}^H) \mathbf{D}_k^{-1}) \\ &= \text{tr}(\mathbb{E}(\mathbf{D}_k) \mathbf{D}_k^{-1}) \\ &= N_r, \end{aligned} \quad (81)$$

where the third line of the equation is obtained by exchanging the operator of expectation and the trace based on the linearity of the trace operator. Therefore, the lower bound on I can be further reduced to be

$$\begin{aligned} I &\geq \log_2(\alpha_k) + \log_2 \left(\frac{1}{\pi^{N_r} \det(\mathbf{D}_k)} \right) - \frac{N_r}{\ln 2} \\ &= \log_2(\alpha_k) - N_r \log_2(\pi) - \log_2(\det(\mathbf{D}_k)) - \frac{\ln e^{N_r}}{\ln 2} \\ &= \log_2(\alpha_k) - N_r \log_2(\pi e) - \log_2(\det(\mathbf{D}_k)). \end{aligned} \quad (82)$$

Therefore, the channel capacity can be upper bounded by

$$\begin{aligned} \mathcal{R}(\boldsymbol{\alpha}, \boldsymbol{\Psi}, \mathcal{Q}) &\leq -N_r \log_2(\pi e) - \sum_{k=1}^K \alpha_k \log_2(\alpha_k) \\ &\quad + \sum_{k=1}^K \alpha_k N_r \log_2(\pi e) + \sum_{k=1}^K \alpha_k \log_2(\det(\mathbf{D}_k)) \\ &= - \sum_{k=1}^K \alpha_k \log_2(\alpha_k) + \sum_{k=1}^K \alpha_k \log_2(\det(\mathbf{D}_k)). \end{aligned} \quad (83)$$

Since both terms summing from $k = 1$ to K involve the same factor α_k , it can be further simplified as

$$\begin{aligned} \mathcal{R}(\boldsymbol{\alpha}, \boldsymbol{\Psi}, \mathcal{Q}) &\leq \sum_{k=1}^K (\alpha_k \log_2(\det(\mathbf{D}_k)) - \alpha_k \log_2(\alpha_k)) \\ &= \sum_{k=1}^K \alpha_k (\log_2(\det(\mathbf{D}_k)) - \log_2(\alpha_k)), \end{aligned} \quad (84)$$

which is exactly the upper bound $\mathcal{R}^U(\boldsymbol{\alpha}, \boldsymbol{\Psi}, \mathcal{Q})$ given in (19). This completes the proof of **Theorem 1**.

REFERENCES

- [1] S. Dang, O. Amin, B. Shihada, and M.-S. Alouini, "What should 6G be?" *Nature Electronics*, vol. 3, no. 1, pp. 20–29, Jan. 2020.
- [2] P. Yang, Y. Xiao, M. Xiao, and S. Li, "6G wireless communications: Vision and potential techniques," *IEEE Network*, vol. 33, no. 4, pp. 70–75, Jul./Aug. 2019.
- [3] W. Saad, M. Bennis, and M. Chen, "A vision of 6G wireless systems: Applications, trends, technologies, and open research problems," *IEEE Network*, vol. 34, no. 3, pp. 134–142, May/Jun. 2020.
- [4] K. B. Letaief, W. Chen, Y. Shi, J. Zhang, and Y. A. Zhang, "The roadmap to 6G: AI empowered wireless networks," *IEEE Communications Magazine*, vol. 57, no. 8, pp. 84–90, Aug. 2019.
- [5] K. David and H. Berndt, "6G vision and requirements: Is there any need for beyond 5G?" *IEEE Vehicular Technology Magazine*, vol. 13, no. 3, pp. 72–80, Sept. 2018.
- [6] M. Di Renzo, M. Debbah, D.-T. Phan-Huy, A. Zappone, M.-S. Alouini, C. Yuen, V. Sciancalepore, G. C. Alexandropoulos, J. Hoydis, H. Gacanin *et al.*, "Smart radio environments empowered by reconfigurable AI meta-surfaces: An idea whose time has come," *EURASIP Journal on Wireless Communications and Networking*, vol. 2019, no. 1, pp. 1–20, May 2019.
- [7] C. Pradhan, A. Li, L. Song, J. Li, B. Vucetic, and Y. Li, "Reconfigurable intelligent surface (RIS)-enhanced two-way OFDM communications," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 12, pp. 16270–16275, Dec. 2020.
- [8] M. A. ElMossallamy, H. Zhang, L. Song, K. G. Seddik, Z. Han, and G. Y. Li, "Reconfigurable intelligent surfaces for wireless communications: Principles, challenges, and opportunities," *IEEE Transactions on Cognitive Communications and Networking*, vol. 6, no. 3, pp. 990–1002, Sept. 2020.
- [9] 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. 116753–116773, Sept. 2019.
- [10] M. Di Renzo, K. Ntontin, J. Song, F. H. Danufane, X. Qian, F. Lazarakis, J. De Rosny, D.-T. Phan-Huy, O. Simeone, R. Zhang *et al.*, "Reconfigurable intelligent surfaces vs. relaying: Differences, similarities, and performance comparison," *IEEE Open Journal of the Communications Society*, vol. 1, pp. 798–807, 2020.
- [11] S. Dang, G. Chen, and J. P. Coon, "Multicarrier relay selection for full-duplex relay-assisted OFDM D2D systems," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 8, pp. 7204–7218, Aug. 2018.
- [12] L. Subrt and P. Pechac, "Intelligent walls as autonomous parts of smart indoor environments," *IET communications*, vol. 6, no. 8, pp. 1004–1010, May 2012.
- [13] S. Kisseleff, W. A. Martins, H. Al-Hraishawi, S. Chatzinotas, and B. Ottersten, "Reconfigurable intelligent surfaces for smart cities: Research challenges and opportunities," *IEEE Open Journal of the Communications Society*, pp. 1–1, 2020.
- [14] A. Kammoun, A. Chaaban, M. Debbah, M.-S. Alouini *et al.*, "Asymptotic max-min SINR analysis of reconfigurable intelligent surface assisted MISO systems," *IEEE Transactions on Wireless Communications*, vol. 19, no. 12, pp. 7748–7764, Dec. 2020.
- [15] C. Huang, A. Zappone, G. C. Alexandropoulos, M. Debbah, and C. Yuen, "Reconfigurable intelligent surfaces for energy efficiency in wireless communication," *IEEE Transactions on Wireless Communications*, vol. 18, no. 8, pp. 4157–4170, Aug. 2019.
- [16] B. Zheng and R. Zhang, "Intelligent reflecting surface-enhanced OFDM: Channel estimation and reflection optimization," *IEEE Wireless Communications Letters*, vol. 9, no. 4, pp. 518–522, Apr. 2019.
- [17] Z.-Q. He and X. Yuan, "Cascaded channel estimation for large intelligent metasurface assisted massive MIMO," *IEEE Wireless Communications Letters*, vol. 9, no. 2, pp. 210–214, Feb. 2019.
- [18] C. You, B. Zheng, and R. Zhang, "Channel estimation and passive beamforming for intelligent reflecting surface: Discrete phase shift and progressive refinement," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 11, pp. 2604–2620, Nov. 2020.
- [19] J. Ye, S. Guo, and M. Alouini, "Joint reflecting and precoding designs for SER minimization in reconfigurable intelligent surfaces assisted MIMO systems," *IEEE Transactions on Wireless Communications*, vol. 19, no. 8, pp. 5561–5574, Aug. 2020.
- [20] H. Hashida, Y. Kawamoto, and N. Kato, "Intelligent reflecting surface placement optimization in air-ground communication networks toward 6G," *IEEE Wireless Communications*, vol. 27, no. 6, pp. 146–151, Dec. 2020.
- [21] L. Zhang, Y. Wang, W. Tao, Z. Jia, T. Song, and C. Pan, "Intelligent reflecting surface aided MIMO cognitive radio systems," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 10, pp. 11445–11457, Oct. 2020.
- [22] X. Yu, D. Xu, Y. Sun, D. W. K. Ng, and R. Schober, "Robust and secure wireless communications via intelligent reflecting surfaces," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 11, pp. 2637–2652, Nov. 2020.
- [23] H. Shen, W. Xu, S. Gong, Z. He, and C. Zhao, "Secrecy rate maximization for intelligent reflecting surface assisted multi-antenna communications," *IEEE Communications Letters*, vol. 23, no. 9, pp. 1488–1492, Sep. 2019.
- [24] M. Cui, G. Zhang, and R. Zhang, "Secure wireless communication via intelligent reflecting surface," *IEEE Wireless Communications Letters*, vol. 8, no. 5, pp. 1410–1414, Oct. 2019.
- [25] G. Zhou, C. Pan, H. Ren, K. Wang, and A. Nallanathan, "Intelligent reflecting surface aided multigroup multicast MISO communication systems," *IEEE Transactions on Signal Processing*, vol. 68, pp. 3236–3251, Apr. 2020.
- [26] Q. Tao, J. Wang, and C. Zhong, "Performance analysis of intelligent reflecting surface aided communication systems," *IEEE Communications Letters*, vol. 24, no. 11, pp. 2464–2468, Nov. 2020.
- [27] T. Van Chien, L. T. Tu, S. Chatzinotas, and B. Ottersten, "Coverage probability and ergodic capacity of intelligent reflecting surface-enhanced communication systems," *IEEE Communications Letters*, vol. 25, no. 1, pp. 69–73, Jan. 2020.
- [28] Ö. Özdogan, E. Björnson, and E. G. Larsson, "Using intelligent reflecting surfaces for rank improvement in MIMO communications," in *Proc. IEEE ICASSP*, May 2020, pp. 9160–9164.
- [29] S. Zhang and R. Zhang, "Capacity characterization for intelligent reflecting surface aided MIMO communication," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 8, pp. 1823–1838, Aug. 2020.
- [30] N. S. Perović, L.-N. Tran, M. Di Renzo, and M. F. Flanagan, "Achievable rate optimization for MIMO systems with reconfigurable intelligent surfaces," *IEEE Transactions on Wireless Communications*, vol. 20, no. 6, pp. 3865–3882, Jun. 2021.
- [31] N. S. Perović, L.-N. Tran, M. Di Renzo, and M. F. Flanagan, "Optimization of RIS-aided MIMO systems via the cutoff rate," *IEEE Wireless Communications Letters*, pp. 1–1, 2021.
- [32] Y. Zhang, C. Zhong, Z. Zhang, and W. Lu, "Sum rate optimization for two way communications with intelligent reflecting surface," *IEEE Communications Letters*, vol. 24, no. 5, pp. 1090–1094, May 2020.
- [33] N. S. Perović, M. D. Renzo, and M. F. Flanagan, "Channel capacity optimization using reconfigurable intelligent surfaces in indoor mmWave environments," in *Proc. IEEE ICC*, Jun. 2020, pp. 1–7.
- [34] X. Ma, Z. Chen, W. Chen, Z. Li, Y. Chi, C. Han, and S. Li, "Joint channel estimation and data rate maximization for intelligent reflecting surface assisted Terahertz MIMO communication systems," *IEEE Access*, vol. 8, pp. 99565–99581, Jun. 2020.
- [35] C. Pan, H. Ren, K. Wang, W. Xu, M. ElKashlan, A. Nallanathan, and L. Hanzo, "Multicell MIMO communications relying on intelligent reflecting surfaces," *IEEE Transactions on Wireless Communications*, vol. 19, no. 8, pp. 5218–5233, Aug. 2020.
- [36] S. Liu, L. Xiao, M. Zhao, X. Xu, and Y. Li, "Performance analysis of intelligent reflecting surface in multi-user MIMO systems," in *Journal of Physics: Conference Series*, vol. 1575, no. 1, 2020, p. 012078.
- [37] M. Hua, Q. Wu, D. W. K. Ng, J. Zhao, and L. Yang, "Intelligent reflecting surface-aided joint processing coordinated multipoint transmission," *IEEE Transactions on Communications*, vol. 69, no. 3, pp. 1650–1665, Mar. 2020.
- [38] C. Pan, H. Ren, K. Wang, M. ElKashlan, A. Nallanathan, J. Wang, and L. Hanzo, "Intelligent reflecting surface aided MIMO broadcasting for

- simultaneous wireless information and power transfer," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 8, pp. 1719–1734, Aug. 2020.
- [39] A. Kalis, A. G. Kanatas, and C. B. Papadias, "A novel approach to MIMO transmission using a single RF front end," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 6, pp. 972–980, Aug. 2008.
- [40] V. I. Barousis, A. G. Kanatas, and A. Kalis, "Beamspace-domain analysis of single-RF front-end MIMO systems," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 3, pp. 1195–1199, Mar. 2011.
- [41] G. Wang, F. Gao, R. Fan, and C. Tellambura, "Ambient backscatter communication systems: Detection and performance analysis," *IEEE Transactions on Communications*, vol. 64, no. 11, pp. 4836–4846, Nov. 2016.
- [42] V. Liu, A. Parks, V. Talla, S. Gollakota, D. Wetherall, and J. R. Smith, "Ambient backscatter: Wireless communication out of thin air," *ACM SIGCOMM Computer Communication Review*, vol. 43, no. 4, pp. 39–50, 2013.
- [43] R. Karasik, O. Simeone, M. Di Renzo, and S. S. Shitz, "Beyond max-SNR: Joint encoding for reconfigurable intelligent surfaces," in *IEEE ISIT*, Jun. 2020, pp. 2965–2970.
- [44] R. Karasik, O. Simeone, M. Di Renzo, and S. Shamai, "Single-RF multi-user communication through reconfigurable intelligent surfaces: An information-theoretic analysis," *arXiv preprint arXiv:2101.07556*, 2021.
- [45] —, "Adaptive coding and channel shaping through reconfigurable intelligent surfaces: An information-theoretic analysis," *arXiv preprint arXiv:2012.00407*, 2020.
- [46] E. Basar, "Transmission through large intelligent surfaces: A new frontier in wireless communications," in *IEEE EuCNC*, Jun. 2019, pp. 112–117.
- [47] —, "Reconfigurable intelligent surface-based index modulation: A new beyond MIMO paradigm for 6G," *IEEE Transactions on Communications*, vol. 68, no. 5, pp. 3187–3196, May 2020.
- [48] W. Yan, X. Yuan, and X. Kuai, "Passive beamforming and information transfer via large intelligent surface," *IEEE Wireless Communications Letters*, vol. 9, no. 4, pp. 533–537, Apr. 2019.
- [49] S. Lin, B. Zheng, G. C. Alexandropoulos, M. Wen, M. D. Renzo, and F. Chen, "Reconfigurable intelligent surfaces with reflection pattern modulation: Beamforming design and performance analysis," *IEEE Transactions on Wireless Communications*, vol. 20, no. 2, pp. 741–754, Feb. 2021.
- [50] S. Guo, S. Lv, H. Zhang, J. Ye, and P. Zhang, "Reflecting modulation," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 11, pp. 2548–2561, Nov. 2020.
- [51] Q. Li, M. Wen, and M. Di Renzo, "Single-RF MIMO: From spatial modulation to metasurface-based modulation," *arXiv preprint arXiv:2009.00789*, 2020.
- [52] M. Di Renzo, A. Zappone, M. Debbah, M.-S. Alouini, C. Yuen, J. de Rosny, and S. Tretjakov, "Smart radio environments empowered by reconfigurable intelligent surfaces: How it works, state of research, and the road ahead," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 11, pp. 2450–2525, Nov. 2020.
- [53] M. Hua, L. Yang, Q. Wu, C. Pan, C. Li, and A. L. Swindlehurst, "UAV-assisted intelligent reflecting surface symbiotic radio system," *IEEE Transactions on Wireless Communications*, pp. 1–1, 2021.
- [54] Q. Zhang, Y.-C. Liang, and H. V. Poor, "Large intelligent surface/antennas (LISA) assisted symbiotic radio for IoT communications," *arXiv preprint arXiv:2002.00340*, 2020.
- [55] —, "Symbiotic radio: A new application of large intelligent surface/antennas (LISA)," in *IEEE WCNC*. IEEE, May 2020, pp. 1–6.
- [56] J. Hu, Y.-C. Liang, and Y. Pei, "Reconfigurable intelligent surface enhanced multi-user MISO symbiotic radio system," *IEEE Transactions on Communications*, vol. 69, no. 4, pp. 2359–2371, Apr. 2020.
- [57] T. A. Le, T. Van Chien, and M. Di Renzo, "Robust probabilistic-constrained optimization for IRS-aided MISO communication systems," *IEEE wireless communications letters*, vol. 10, no. 1, pp. 1–5, Jan. 2020.
- [58] D. Mishra and H. Johansson, "Channel estimation and low-complexity beamforming design for passive intelligent surface assisted MISO wireless energy transfer," in *Proc. IEEE ICASSP*, May 2019, pp. 4659–4663.
- [59] Q.-U.-A. Nadeem, A. Kammoun, A. Chaaban, M. Debbah, and M.-S. Alouini, "Intelligent reflecting surface assisted multi-user MISO communication," *arXiv preprint arXiv:1906.02360*, 2019.
- [60] E. Björnson and L. Sanguinetti, "Rayleigh fading modeling and channel hardening for reconfigurable intelligent surfaces," *IEEE Wireless Communications Letters*, vol. 10, no. 4, pp. 830–834, Apr. 2020.
- [61] T. Cover and J. Thomas, *Elements of Information Theory*. Wiley, 2012.
- [62] A. A. Ibrahim, T. Kim, and D. J. Love, "On the achievable rate of generalized spatial modulation using multiplexing under a Gaussian mixture model," *IEEE Transactions on Communications*, vol. 64, no. 4, pp. 1588–1599, Apr. 2016.
- [63] J. R. Hershey and P. A. Olsen, "Approximating the Kullback Leibler divergence between gaussian mixture models," in *Proc. IEEE ICASSP*, vol. 4, Honolulu, HI, USA, Apr. 2007, pp. 317–320.
- [64] L. He, J. Wang, and J. Song, "Spatial modulation for more spatial multiplexing: RF-chain-limited generalized spatial modulation aided MM-wave MIMO with hybrid precoding," *IEEE Transactions on Communications*, vol. 66, no. 3, pp. 986–998, Mar. 2017.
- [65] S. Guo, H. Zhang, P. Zhang, P. Zhao, L. Wang, and M.-S. Alouini, "Generalized beamspace modulation using multiplexing: A breakthrough in mmWave MIMO," *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 9, pp. 2014–2028, Sept. 2019.
- [66] S. Guo, H. Zhang, and M. Alouini, "Asymptotic capacity for MIMO communications with insufficient radio frequency chains," *IEEE Transactions on Communications*, vol. 68, no. 7, pp. 4190–4201, Jul. 2020.
- [67] S. Boyd, S. P. Boyd, and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.
- [68] P. Liu and T. Jiang, "Channel estimation performance analysis of massive MIMO IoT systems with rician fading," *IEEE Internet of Things Journal*, vol. 8, no. 7, pp. 6114–6126, Apr. 2021.
- [69] S. Dang, J. P. Coon, and G. Chen, "Adaptive OFDM with index modulation for two-hop relay-assisted networks," *IEEE Transactions on Wireless Communications*, vol. 17, no. 3, pp. 1923–1936, Mar. 2018.
- [70] 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 *IEEE Globecom Workshops (GC Wkshps)*, Dec. 2018, pp. 1–6.