

Intelligent Reflecting Surface Aided Multi-User Communication: Capacity Region and Deployment Strategy

Shuowen Zhang, *Member, IEEE* and Rui Zhang, *Fellow, IEEE*

Abstract—Intelligent reflecting surface (IRS) is a new promising technology that is able to reconfigure the wireless propagation channel via smart and passive signal reflection. In this paper, we investigate the *capacity region* of a two-user communication network with one access point (AP) aided by M IRS elements for enhancing the user-AP channels, where the IRS incurs negligible delay, thus the user-AP channels via the IRS follow the classic discrete memoryless channel model. In particular, we consider two practical IRS deployment strategies that lead to different effective channels between the users and AP, namely, the *distributed deployment* where the M elements form two IRSs, each deployed in the vicinity of one user, versus the *centralized deployment* where all the M elements are deployed in the vicinity of the AP. First, we consider the uplink multiple-access channel (MAC) and derive the capacity/achievable rate regions for both deployment strategies under different multiple access schemes. It is shown that the centralized deployment generally outperforms the distributed deployment under symmetric channel setups in terms of achievable user rates. Next, we extend the results to the downlink broadcast channel (BC) by leveraging the celebrated uplink-downlink (or MAC-BC) duality framework, and show that the superior rate performance of centralized over distributed deployment also holds. Numerical results are presented that validate our analysis, and reveal new and useful insights for optimal IRS deployment in wireless networks.

Index Terms—Intelligent reflecting surface (IRS), capacity region, IRS deployment, multiple-access channel (MAC), broadcast channel (BC).

I. INTRODUCTION

Driven by the recent advancement in metamaterial technology, *intelligent reflecting surface (IRS)* has become a new solution to meet the increasingly high communication demands in beyond the fifth-generation (5G) wireless networks [2]–[4]. Specifically, IRS is a reconfigurable metasurface consisting of a large number of passive elements, each of which is able to introduce an independent and controllable phase shift to the impinging electromagnetic wave, thereby collaboratively altering the propagation channels between the wireless transceivers. By properly designing the IRS reflection coefficients (i.e., phase shifts), IRS has been shown effective in adaptively reconfiguring the wireless environment for enhancing desired signal strength and mitigating interference,

thus improving the achievable rate and/or reliability of various wireless communication systems. Moreover, IRSs are spectral-efficient and energy conservative since they perform passive signal reflection in the full-duplex mode without the need of any transmit radio frequency (RF) chains, which makes them suitable to be densely deployed in wireless networks.

To fully reap the benefits of IRS, IRS needs to be efficiently integrated into existing wireless communication systems, which brings new challenges. First, it is of paramount importance to optimally design the IRS reflection coefficients such that the wireless channels are properly altered in favor of communication performance, which has attracted a great deal of research interests recently (see, e.g., [5]–[10] for single-user systems and [11]–[24] for multi-user systems). Particularly, from an information theoretical viewpoint, it is crucial to characterize the fundamental *capacity limits* of IRS-aided communication systems so as to unveil the maximum performance gains brought by IRS, which, however, has only been pursued recently in [5], [6], [25] under the single-user setup. While for the more complex IRS-aided multi-user systems of which the performance limit is characterized by the *capacity region* that constitutes all the achievable rate-tuples of the users in the system, there has been very limited work in the literature, to the authors' best knowledge. This thus motivates the current work to characterize the capacity region of two fundamental multi-user communication systems aided by IRS, namely, the *multiple-access channel (MAC)* in the uplink and the *broadcast channel (BC)* in the downlink, respectively.

Besides capacity characterization through IRS reflection optimization, another important problem not well understood for IRS-aided multi-user systems is *IRS deployment*. The existing works on IRS have mostly assumed *given* IRS deployment (or IRS locations) without exploiting its design flexibility. However, with a given number of IRS elements, there are assorted approaches to deploy them in the network (e.g., by forming them as multiple IRSs, and placing them at different locations), which can lead to different IRS channels in general and thus impact the system capacity/achievable rates significantly. For instance, for the basic point-to-point communication system with one access point (AP) serving one user, if all elements form one single IRS, the IRS should be placed in close vicinity of the AP or the user so as to minimize the overall path loss of the AP-IRS-user channel, which increases with the product of the AP-IRS and IRS-user distances [2]. In contrast, by splitting the elements into two IRSs and deploying them near

This paper was presented in part at the IEEE International Workshop on Signal Processing Advances for Wireless Communications (SPAWC), 2020 [1].

S. Zhang is with the Department of Electronic and Information Engineering, The Hong Kong Polytechnic University (e-mail: shuowen.zhang@polyu.edu.hk). She was with the Department of Electrical and Computer Engineering, National University of Singapore. R. Zhang is with the Department of Electrical and Computer Engineering, National University of Singapore (e-mail: elezhang@nus.edu.sg). The work of R. Zhang is supported by National University of Singapore under research grant R-261-518-005-720 and R-263-000-E86-112.

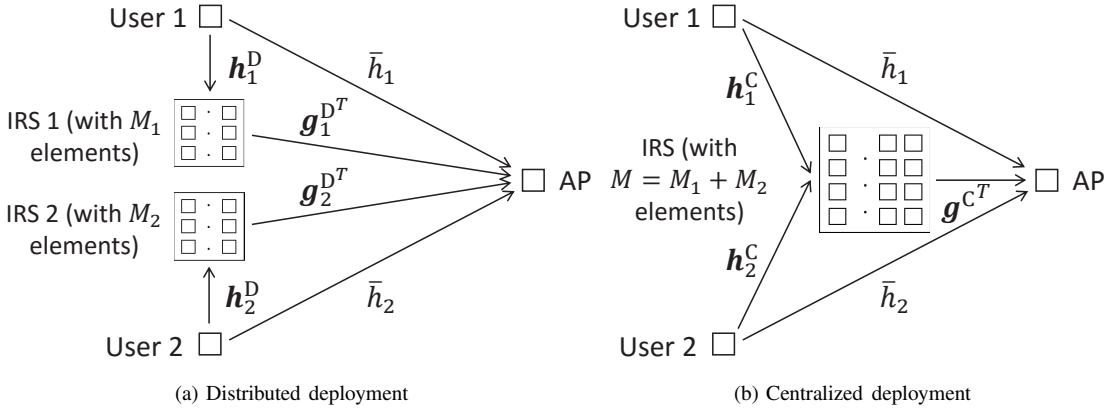


Fig. 1. An IRS-aided two-user communication system with different IRS deployment strategies (in the uplink MAC case).

the AP and the user, respectively, the received signal-to-noise ratio (SNR) can be significantly improved as compared to the optimal single-IRS deployment by exploiting the cooperative beamforming design of double IRSs, provided that the total number of IRS elements is sufficiently large [26]. However, for the more general system with multiple users or user clusters that are located far apart from one another, the optimal IRS deployment design has not been investigated yet. Generally speaking, motivated by the above single-user setup, there are two strategies to deploy a given number of IRS elements in the multi-user case: *distributed IRS deployment* where the available elements form multiple distributed IRSs each deployed near one user (or user cluster), as illustrated in Fig. 1 (a), and *centralized IRS deployment* where all elements form one IRS deployed near the AP, as illustrated in Fig. 1 (b). Note that these two strategies will lead to different user-AP effective channels in general and hence different user achievable rates. Specifically, with distributed deployment, each user can only enjoy the passive beamforming gain brought by its nearby IRS (since its signals reflected by other far-apart IRSs are too weak due to much higher path loss), which is thus smaller than the passive beamforming gain under the centralized deployment with a larger-size IRS where all the elements can be used for enhancing the channel of any user. However, the IRS passive beamforming gain under the centralized deployment needs to be shared by all the simultaneously served users in general, which may result in a reduced gain for each user. To the best of our knowledge, it is yet unclear which of the above two IRS deployment strategies achieves larger capacity region or achievable rate regions with practical orthogonal multiple access (OMA) schemes such as time-division multiple access (TDMA) and frequency-division multiple access (FDMA) in multi-user systems such as MAC and BC.

To tackle the above capacity characterization and IRS deployment problems, we study in this paper a two-user communication system aided by M IRS elements, as shown in Fig. 1. Specifically, this paper considers the most practical scenario of employing IRS, where the phase shifts at the IRS elements are tuned to collaboratively alter the effective user-AP channels only. We aim to derive the capacity region as well as achievable rate regions with TDMA/FDMA under the proposed two IRS deployment strategies, for both the uplink

MAC and downlink BC. To pursue the fundamental performance limit, we assume the availability of perfect channel state information (CSI) for all the channels shown in Fig. 1, which can be acquired via various existing channel training and estimation methods proposed in e.g., [7], [27]–[32]. Our main contributions are summarized as follows.

- First, we investigate the uplink MAC. For the distributed IRS deployment, we provide *closed-form* characterizations of its capacity region as well as the achievable rate regions with TDMA and FDMA. For the centralized IRS deployment, we develop a *rate-profile* based method to characterize the capacity region by solving a series of sum-rate maximization problems via joint IRS reflection and transmit power optimization, each corresponding to a different rate-ratio between the two users. An efficient alternating optimization (AO) algorithm is proposed to find a high-quality approximate solution to each problem, where we iteratively obtain the *optimal* solution to each IRS reflection coefficient with the others being fixed. Based on this, a capacity region inner bound (or achievable rate region) can be obtained with polynomial complexity. We further propose a capacity region outer bound via the semi-definite relaxation (SDR) technique. In addition, we characterize the achievable rate region with TDMA in closed-form and that with FDMA via a similar rate-profile method.
- Moreover, we analytically prove that the capacity region as well as achievable rate regions with TDMA and FDMA for the case of centralized deployment *contain* the corresponding regions for the case of distributed deployment under a practical symmetric channel setup. Furthermore, despite the lack of frequency-selective signal reflection at the IRS, the achievable rate region with FDMA still contains that with TDMA for both deployment strategies.
- Next, we extend the results for the uplink MAC to the downlink BC by leveraging the *uplink-downlink (or MAC-BC) duality* framework. We propose computationally efficient methods to characterize the capacity region and achievable rate regions with TDMA/FDMA based on those for the dual MAC, and prove that the performance advantage of centralized over distributed IRS deployment is also valid for the downlink BC.
- Finally, numerical results validate our analysis as well as

tightness of the proposed bounds. By comparing the various capacity/rate regions, we also draw useful insights into optimal deployment of IRSs in practical systems. Particularly, it is shown that the capacity gain of centralized over distributed IRS deployment is most prominent when the rates of the two users are *asymmetric*; moreover, centralized IRS deployment is more effective in alleviating the “*near-far*” problem in multi-user communications, as a result of users with drastically different distances from the AP.

The rest of this paper is organized as follows. Section II presents the system models under the two IRS deployment strategies. For the uplink MAC, Section III and Section IV characterize the capacity region and achievable rate regions with TDMA/FDMA under distributed and centralized IRS deployment, respectively; Section V compares these regions under the two IRS deployment strategies. Section VI extends the above results to the downlink BC. Numerical examples and their pertinent discussions are presented in Section VII. Finally, Section VIII concludes this paper.

Notations: Vectors and matrices are denoted by boldface lower-case letters and boldface upper-case letters, respectively. $|x|$, x^* , $\arg\{x\}$, and $\Re\{x\}$ denote the absolute value, conjugate, angle, and real part of a complex number x , respectively. For a complex vector z , $\|z\|_p$ and z_k denote the l_p -norm and the k th element, respectively, and $\text{diag}\{z\}$ denotes a square diagonal matrix with the elements of z on its main diagonal. $\mathbb{C}^{M \times N}$ denotes the space of $M \times N$ complex matrices. $\mathbf{0}$ denotes an all-zero matrix with appropriate dimension. For an $M \times N$ matrix \mathbf{A} , $\text{rank}(\mathbf{A})$ and $[\mathbf{A}]_{i,j}$ denote the rank and (i,j) -th element of \mathbf{A} , respectively. For a square matrix \mathbf{S} , $\mathbf{S} \succeq \mathbf{0}$ means that \mathbf{S} is positive semi-definite. The distribution of a circularly symmetric complex Gaussian (CSCG) random variable with mean 0 and variance σ^2 is denoted by $\mathcal{CN}(0, \sigma^2)$; and \sim stands for “distributed as”. $\mathbb{E}[\cdot]$ denotes the statistical expectation. $\mathcal{O}(\cdot)$ denotes the standard big-O notation. $\text{Conv}(\cdot)$ denotes the convex hull operation. \cup denotes the union operation.

II. SYSTEM MODEL

We consider a communication network where one single-antenna AP serves two single-antenna users that are sufficiently far apart from each other, as illustrated in Fig. 1. In the uplink, each user aims to send an independent message to the AP; while in the downlink, the AP sends independent messages to the users. Moreover, M passive IRS elements are deployed to enhance the user communication rates. Specifically, this paper considers the most practical scenario of employing IRS, where each IRS element induces an independent phase shift to its incident signal for collaboratively altering the effective channels between the users and the AP.¹ Under this strategy, we define the *IRS-aided MAC* and the *IRS-aided BC* as the uplink (from the users to the AP) and downlink (from the AP to the users) transmissions, respectively, which are similar to the conventional discrete memoryless MAC and

¹It is worth noting that IRS can perform other functions besides altering the wireless channel. For example, if user messages are known at the IRS, IRS can forward messages to the users by varying its reflection over time [25]. However, this requires message sharing between the AP and IRS, which is not considered in this paper.

BC [33], but with the user effective channels controllable by the IRS.

For the IRS-aided MAC and BC, we propose two different deployment strategies for the M elements. Specifically, for the *distributed deployment*, the M elements form two IRSs (see Fig. 1 (a)), where IRS k , $k \in \{1, 2\}$, consists of M_k elements, with $M_k \geq 1$, and is placed in the vicinity of user k , subject to $\sum_{k=1}^2 M_k = M$. In contrast, for the *centralized deployment*, all the M elements form one single IRS located in the vicinity of the AP (see Fig. 1 (b)). For the purpose of exposition, we will first focus our study on the IRS-aided MAC in the uplink shown in Fig. 1, and then extend our study to the IRS-aided BC in the downlink (see Section VI). Specifically, for the IRS-aided MAC, we denote the baseband equivalent direct channel from user k to the AP as $\bar{h}_k \in \mathbb{C}$, $k = 1, 2$. In the following, we describe the system models for the two IRS deployment cases, respectively.

A. Distributed IRS Deployment

For distributed IRS deployment, we denote $\mathbf{h}_k^D \in \mathbb{C}^{M_k \times 1}$ as the channel vector from user k to its serving (nearby) IRS, and $\mathbf{g}_k^{D^T} \in \mathbb{C}^{1 \times M_k}$ as the channel vector from its serving IRS to the AP. Denote $\Phi_k^D = \text{diag}\{\phi_{k1}^D, \dots, \phi_{kM_k}^D\} \in \mathbb{C}^{M_k \times M_k}$ as the IRS reflection matrix for IRS k , with $|\phi_{km}^D| = 1$, $\forall m \in \mathcal{M}_k$, where $\mathcal{M}_k = \{1, \dots, M_k\}$. Since the locations of the two users are sufficiently far apart, we assume that the signal transmitted by one user and reflected by the other user’s serving IRS is negligible at the AP due to the severe path loss. Hence, the effective channel from user k to the AP by combining the direct link and the reflected link by its serving IRS is given by

$$\tilde{h}_k^D(\Phi_k^D) = \bar{h}_k + \mathbf{g}_k^{D^T} \Phi_k^D \mathbf{h}_k^D, \quad k = 1, 2. \quad (1)$$

Let $s_k \in \mathbb{C}$ denote the desired information symbol for user k with zero mean and unit variance. Note that s_k ’s are independent over k . The transmitted signal by user k is modeled as $x_k = \sqrt{p_k} s_k$, which satisfies $\mathbb{E}[|x_k|^2] = p_k \leq P_k$, with p_k denoting the transmit power of user k and P_k denoting its maximum value. The received signal at the AP is thus modeled as

$$y = \tilde{h}_1^D(\Phi_1^D)x_1 + \tilde{h}_2^D(\Phi_2^D)x_2 + z, \quad (2)$$

where $z \sim \mathcal{CN}(0, \sigma^2)$ denotes the CSCG noise at the AP receiver with average power σ^2 . For each user k , let R_k^D denote its achievable rate in bits per second per Hertz (bps/Hz) under the distributed IRS deployment.

B. Centralized IRS Deployment

For centralized IRS deployment, we denote $\mathbf{h}_k^C \in \mathbb{C}^{M \times 1}$ as the channel vector from user k to the (single) IRS, and $\mathbf{g}^{C^T} \in \mathbb{C}^{1 \times M}$ as the channel vector from the IRS to the AP. Denote $\Phi^C = \text{diag}\{\phi_1^C, \dots, \phi_M^C\} \in \mathbb{C}^{M \times M}$ as the IRS reflection matrix, with $|\phi_m^C| = 1$, $\forall m \in \mathcal{M}$, where $\mathcal{M} = \{1, \dots, M\}$. Thus, the effective channel from user k to the AP is given by

$$\tilde{h}_k^C(\Phi^C) = \bar{h}_k + \mathbf{g}^{C^T} \Phi^C \mathbf{h}_k^C, \quad k = 1, 2. \quad (3)$$

Note that different from the distributed deployment where the effective channel between user k and the AP is only dependent on the M_k reflection coefficients of its own serving IRS in Φ_k^D , the effective channels for both users under the centralized

deployment depend on all the M reflection coefficients in Φ^C . Under the same transmitted signal and receiver noise model as in the distributed deployment case, the received signal at the AP is modeled similarly as (2) by replacing each $\tilde{h}_k^D(\Phi^D)$ with $\tilde{h}_k^C(\Phi^C)$. For user k , let R_k^C denote its achievable rate in bps/Hz under the centralized IRS deployment.

In the following, we characterize the *capacity region* of the IRS-aided two-user MAC under each of the two deployment strategies, which constitutes all the achievable rate-pairs (R_1^D, R_2^D) 's or (R_1^C, R_2^C) 's. We also derive their *achievable rate regions* under practical OMA schemes including TDMA and FDMA, where the two users communicate with the AP in orthogonal time slots or frequency bands, respectively. We then compare these capacity (rate) regions and draw useful insights into the optimal IRS deployment strategy.

III. DISTRIBUTED IRS DEPLOYMENT

In this section, we characterize the capacity region as well as the achievable rate regions with TDMA/FDMA under the distributed IRS deployment.

A. Capacity Region

First, we derive the capacity region to unveil the fundamental limit. Note that with given IRS reflection coefficients $\{\Phi_k^D\}$, the channels from the two users to the AP are determined as $\{\tilde{h}_k^D(\Phi_k^D)\}$ given in (1), and the capacity region of the two-user MAC is well-known as the pentagon region consisting of all rate-pairs that satisfy the following constraints [33]:

$$R_1^D \leq \log_2(1 + P_1 |\tilde{h}_1^D(\Phi_1^D)|^2 / \sigma^2) \triangleq r_1^D(\Phi_1^D), \quad (4)$$

$$R_2^D \leq \log_2(1 + P_2 |\tilde{h}_2^D(\Phi_2^D)|^2 / \sigma^2) \triangleq r_2^D(\Phi_2^D), \quad (5)$$

$$R_1^D + R_2^D \leq \log_2(1 + (P_1 |\tilde{h}_1^D(\Phi_1^D)|^2 + P_2 |\tilde{h}_2^D(\Phi_2^D)|^2) / \sigma^2) \triangleq r_{12}^D(\{\Phi_k^D\}), \quad (6)$$

which is denoted as $\mathcal{C}^D(\{\Phi_k^D\})$. Note that by flexibly designing the IRS reflection coefficients $\{\Phi_k^D\}$, any rate-pair within the union set of $\mathcal{C}^D(\{\Phi_k^D\})$'s over all feasible $\{\Phi_k^D\}$'s can be achieved. By further considering *time sharing* among different $\{\Phi_k^D\}$'s, the capacity region of IRS-aided MAC for distributed IRS deployment is defined as the convex hull of such a union set [33]:

$$\mathcal{C}^D \triangleq \text{Conv}\left(\bigcup_{\{\Phi_k^D\} \in \mathcal{F}^D} \mathcal{C}^D(\{\Phi_k^D\})\right), \quad (7)$$

where $\mathcal{F}^D \triangleq \{\{\Phi_k^D\} : |\phi_{km}^D| = 1, \forall k, m\}$ denotes the feasible set of $\{\Phi_k^D\}$. It is worth noting that for any given $\{\Phi_k^D\}$, the optimal input distribution for achieving $\mathcal{C}^D(\{\Phi_k^D\})$ is the CSCG distribution with $s_k \sim \mathcal{CN}(0, 1)$, $\forall k$ [33]. Therefore, it follows from (7) that the capacity-achieving optimal input distribution for the IRS-aided MAC under distributed deployment is still $s_k \sim \mathcal{CN}(0, 1)$, $\forall k$, which will be considered throughout this section.

In the following, we characterize \mathcal{C}^D in closed-form by exploiting the peculiar effective channel structure under the distributed deployment. Specifically, according to the triangle inequality, for any $\{\Phi_k^D\} \in \mathcal{F}^D$, the effective channel gain for each user k is upper-bounded by

$$\begin{aligned} |\tilde{h}_k^D(\Phi_k^D)| &= |\tilde{h}_k + \sum_{m=1}^{M_k} g_{km}^D \phi_{km}^D h_{km}^D| \\ &\leq |\tilde{h}_k| + \sum_{m=1}^{M_k} |g_{km}^D| |h_{km}^D| \triangleq \tilde{h}_{k,U}^D, \quad k = 1, 2, \quad (8) \end{aligned}$$

where the inequality holds with equality if and only if $\{\Phi_k^D\}$ is designed as follows:

$$\phi_{km}^D = e^{j(\arg\{\tilde{h}_k\} - \arg\{g_{km}^D h_{km}^D\})}, \quad k = 1, 2, \quad m \in \mathcal{M}_k. \quad (9)$$

Based on this result, we obtain the following theorem.

Theorem 1: The capacity region of the IRS-aided two-user MAC in (1) under the distributed deployment is given by $\mathcal{C}^D = \{(R_1^D, R_2^D) : R_1^D \leq r_1^{D*}, R_2^D \leq r_2^{D*}, R_1^D + R_2^D \leq r_{12}^{D*}\}$, (10)

where $r_1^{D*} \triangleq \log_2(1 + P_1 \tilde{h}_{1,U}^{D2} / \sigma^2)$, $r_2^{D*} \triangleq \log_2(1 + P_2 \tilde{h}_{2,U}^{D2} / \sigma^2)$, and $r_{12}^{D*} \triangleq \log_2(1 + (P_1 \tilde{h}_{1,U}^{D2} + P_2 \tilde{h}_{2,U}^{D2}) / \sigma^2)$.

Proof: Theorem 1 can be proved by noting that \mathcal{C}^D given in (10) is an achievable rate region with $\{\Phi_k^D\}$ given in (9), and also provides a convex-shape outer bound for all achievable $\mathcal{C}^D(\{\Phi_k^D\})$'s given in (4)–(6) (thus, the convex-hull operation in (7) is not needed with $\{\Phi_k^D\}$ given in (9)). ■

Note that to achieve the above capacity region \mathcal{C}^D , successive interference cancellation (SIC) or joint decoding needs to be performed at the AP in general [33]. For example, with SIC, the AP needs to first decode the message of one user by treating the signal of the other user as noise, then cancel the decoded signal and decode the other user's message [33]. Next, we derive the achievable rate regions with TDMA and FDMA, where the SIC/joint decoding operation is not needed since the signals of the two users are already separated in the time and frequency domains, respectively. Note that these achievable rate regions generally serve as inner bounds of the capacity region.

B. Achievable Rate Region with TDMA

With TDMA, the two users transmit in two orthogonal time slots, where we let $\rho_T \in [0, 1]$ denote the fraction of time that user 1 sends its message. In this case, under distributed IRS deployment, the achievable rate of each user k only depends on the reflection matrix of IRS k at its assigned time slot, which is denoted as $\Phi_k^D[k]$. For any given $\{\Phi_k^D[k]\}$, the achievable rate region is defined as $\mathcal{R}_T^D(\{\Phi_k^D[k]\}) = \bigcup_{\rho_T \in [0, 1]} \{(R_1^D, R_2^D) : R_1^D \leq \rho_T \log_2(1 + P_1 |\tilde{h}_1^D(\Phi_1^D[1])|^2 / \sigma^2), R_2^D \leq (1 - \rho_T) \log_2(1 + P_2 |\tilde{h}_2^D(\Phi_2^D[2])|^2 / \sigma^2)\}$. Note that both $|\tilde{h}_1^D(\Phi_1^D[1])|$ and $|\tilde{h}_2^D(\Phi_2^D[2])|$ can be maximized as $\tilde{h}_{1,U}^D$ and $\tilde{h}_{2,U}^D$ in (8) by setting $\Phi_1^D[1]$ and $\Phi_2^D[2]$ as Φ_1^D and Φ_2^D given in (9), respectively; moreover, $\mathcal{R}_T^D(\{\Phi_k^D[k]\})$ with the aforementioned $\{\Phi_k^D[k]\}$ can be easily shown to be a convex region. Therefore, it can be proved similarly as Theorem 1 that the achievable rate region with TDMA under distributed deployment is

$$\mathcal{R}_T^D = \bigcup_{\rho_T \in [0, 1]} \left\{ (R_1^D, R_2^D) : R_1^D \leq \rho_T \log_2 \left(1 + \frac{P_1 \tilde{h}_{1,U}^{D2}}{\sigma^2} \right), R_2^D \leq (1 - \rho_T) \log_2 \left(1 + \frac{P_2 \tilde{h}_{2,U}^{D2}}{\sigma^2} \right) \right\}. \quad (11)$$

C. Achievable Rate Region with FDMA

With FDMA, the two users transmit simultaneously over two orthogonal frequency bands, where we let $\rho_F \in [0, 1]$ denote the fraction of bandwidth assigned to user 1.² With any given $\{\Phi_k^D\}$, the achievable rate region is

²It is worth noting that the IRS reflection coefficients impact the channels at different frequency bands identically without frequency selectivity, due to the lack of RF chains and baseband processing at the IRS.

given by $\mathcal{R}_F^D(\{\Phi_k^D\}) = \bigcup_{\rho_F \in [0,1]} \{(R_1^D, R_2^D) : R_1^D \leq \rho_F \log_2(1 + P_1 |\tilde{h}_1^D(\Phi_1^D)|^2 / (\rho_F \sigma^2)), R_2^D \leq (1 - \rho_F) \log_2(1 + P_2 |\tilde{h}_2^D(\Phi_2^D)|^2 / ((1 - \rho_F) \sigma^2))\}$ [33]. Note that $|\tilde{h}_1^D(\Phi_1^D)|$ and $|\tilde{h}_2^D(\Phi_2^D)|$ can be simultaneously maximized as $\tilde{h}_{1,U}^D$ and $\tilde{h}_{2,U}^D$ in (8) with $\{\Phi_k^D\}$ given in (9), and the corresponding $\mathcal{R}_F^D(\{\Phi_k^D\})$ is a convex region [33]. Thus, similar to the TDMA case, the achievable rate region with FDMA under distributed deployment is

$$\mathcal{R}_F^D = \bigcup_{\rho_F \in [0,1]} \left\{ (R_1^D, R_2^D) : R_1^D \leq \rho_F \log_2 \left(1 + \frac{P_1 \tilde{h}_{1,U}^{D^2}}{\rho_F \sigma^2} \right), R_2^D \leq (1 - \rho_F) \log_2 \left(1 + \frac{P_2 \tilde{h}_{2,U}^{D^2}}{(1 - \rho_F) \sigma^2} \right) \right\}. \quad (12)$$

To summarize, the Pareto boundaries of all the capacity region and TDMA/FDMA achievable rate regions under distributed IRS deployment are achieved by setting the reflection coefficients at each IRS based on its nearby user's channel as given in (9). Moreover, by comparing (12) with (11), it can be easily shown that the achievable rate region of FDMA contains that of TDMA since in both cases the optimal reflection coefficients are identical and thus yield the same user-AP effective channels, while the signal energy transmitted by the two users in TDMA is generally less than their counterparts in FDMA [33]. In addition, $\mathcal{R}_F^D \subseteq \mathcal{C}^D$ holds since $\mathcal{R}_F^D(\{\Phi_k^D\}) \subseteq \mathcal{C}^D(\{\Phi_k^D\})$ holds for any $\{\Phi_k^D\}$ [33], and the optimal $\{\Phi_k^D\}$'s that achieve the Pareto boundaries of \mathcal{C}^D and \mathcal{R}_F^D are the same, as shown in Section III-A and Section III-C. Therefore, we have $\mathcal{R}_T^D \subseteq \mathcal{R}_F^D \subseteq \mathcal{C}^D$ in the case of distributed IRS deployment.

IV. CENTRALIZED IRS DEPLOYMENT

In this section, we characterize the capacity region and achievable rate regions with TDMA/FDMA under the centralized IRS deployment. Similar to the distributed IRS case, the capacity region with given reflection coefficients Φ^C of the centralized IRS is the pentagon region consisting of all rate-pairs that satisfy the following constraints:

$$R_1^C \leq \log_2(1 + P_1 |\tilde{h}_1^C(\Phi^C)|^2 / \sigma^2) \triangleq r_1^C(\Phi^C), \quad (13)$$

$$R_2^C \leq \log_2(1 + P_2 |\tilde{h}_2^C(\Phi^C)|^2 / \sigma^2) \triangleq r_2^C(\Phi^C), \quad (14)$$

$$R_1^C + R_2^C \leq \log_2(1 + (P_1 |\tilde{h}_1^C(\Phi^C)|^2 + P_2 |\tilde{h}_2^C(\Phi^C)|^2) / \sigma^2) \triangleq r_{12}^C(\Phi^C), \quad (15)$$

which is denoted as $\mathcal{C}^C(\Phi^C)$. By tuning the IRS reflection coefficients Φ^C and performing time sharing among different Φ^C 's, the capacity region is defined as

$$\mathcal{C}^C \triangleq \text{Conv} \left(\bigcup_{\Phi^C \in \mathcal{F}^C} \mathcal{C}^C(\Phi^C) \right), \quad (16)$$

where $\mathcal{F}^C \triangleq \{\Phi^C : |\phi_m^C| = 1, \forall m\}$ denotes the feasible set of Φ^C . Similar to the distributed case, for any given Φ^C , the optimal input distribution that achieves the capacity region $\mathcal{C}^C(\Phi^C)$ is $s_k \sim \mathcal{CN}(0, 1)$, $\forall k$. It thus follows from (16) that it is also the capacity-achieving input distribution for the IRS-aided MAC under centralized deployment, which will be assumed throughout this section.

Compared to the distributed IRS case, the capacity region in (16) is more challenging to characterize. This is because the effective channels of the two users, $\tilde{h}_1^C(\Phi^C)$ and $\tilde{h}_2^C(\Phi^C)$, are coupled through all the M reflection coefficients in Φ^C . Thus,

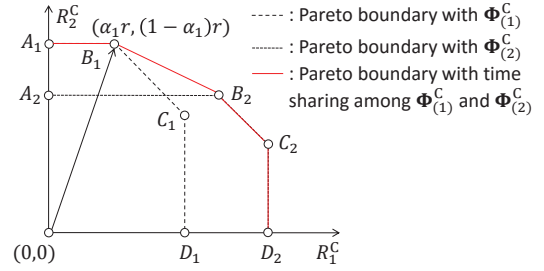


Fig. 2. Illustration of the capacity region under centralized IRS deployment. SIC-achievable rate-pairs with given IRS reflection coefficients $\Phi_{(n)}^C$ are shown on the line segments A_n-B_n and C_n-D_n , $n = 1, 2$.

different portions of the Pareto boundary of the capacity region \mathcal{C}^C are generally achieved by different Φ^C to strike a balance between $\tilde{h}_1^C(\Phi^C)$ and $\tilde{h}_2^C(\Phi^C)$, as illustrated in Fig. 2. Finding such capacity-achieving sets of reflection coefficients is more challenging as compared to the distributed IRS case where the entire Pareto boundary of the capacity region is achieved by a single set of $\{\Phi_k^D\}$ given in (9), since the effective channel of each user is maximized by the reflection coefficients of its own serving IRS.³ Although \mathcal{C}^C can be characterized via the *exhaustive search* method by first obtaining $\mathcal{C}^C(\Phi^C)$'s for all feasible $\Phi^C \in \mathcal{F}^C$ and then taking the convex hull of their union set, the required complexity is at least $\mathcal{O}(L_0^M)$ if the $[0, 2\pi)$ phase range for each ϕ_m^C in Φ^C is approximated by L_0 uniformly sampled points, which is *exponential* over M and thus prohibitive for practically large M . To avoid such high complexity for characterizing the capacity region \mathcal{C}^C , in the following subsections, we provide an alternative method to characterize \mathcal{C}^C , and develop efficient algorithms to find both inner and outer bounds of \mathcal{C}^C , whose tightness will be evaluated via numerical results in Section VII.

A. Rate-Profile based Capacity Region Characterization

To start with, note that for each Φ^C , all the achievable rate-pairs on the Pareto boundary of its corresponding $\mathcal{C}^C(\Phi^C)$ except those requiring time sharing/rate splitting of the two users can be attained via SIC at the AP [33]. Motivated by this result, we propose to first characterize the Pareto boundary of the union set of the above SIC-achievable rate-pairs for all feasible $\Phi^C \in \mathcal{F}^C$, and then perform time sharing among the obtained rate-pairs on the Pareto boundary to further enlarge the achievable rate region, as illustrated in Fig. 2. For the first task, we propose to adopt the *rate-profile* approach in [34].⁴ Specifically, let π denote the decoding order indicator, with $\pi = [1, 2]^T \triangleq \pi^I$ representing that user 1 is decoded before user 2, and $\pi = [2, 1]^T \triangleq \pi^{II}$ otherwise. Let $\alpha_{\pi_1} \in [0, 1]$ denote the rate ratio between the firstly decoded user and the users' sum-rate. We further

³It is worth noting that the user channels under distributed deployment can be expressed as a special case of those under centralized deployment, thus the capacity/rate region characterization methods for the centralized case are also readily applicable to the distributed case.

⁴It is worth noting that another approach to characterize the Pareto boundary is by solving a series of *weighted sum-rate maximization (WSRmax)* problems [34]. However, this approach is generally not guaranteed to obtain the complete Pareto boundary \mathcal{C}^C since the region formed by the SIC-achievable rate-pairs without time sharing is non-convex in general [34]; moreover, such WSRmax problems are also challenging to solve since the rates of the two users are coupled in the objective function in a complicated manner.

denote $\alpha_1 \in [0, 1]$ as the rate ratio between user 1 and the users' sum-rate, $\alpha_2 = 1 - \alpha_1 \in [0, 1]$ as that between user 2 and the users' sum-rate, and $\boldsymbol{\alpha} = [\alpha_1, 1 - \alpha_1]^T$ as the rate-profile vector, with $\boldsymbol{\alpha} = [\alpha_{\pi_1}, 1 - \alpha_{\pi_1}]^T$ if $\boldsymbol{\pi} = \boldsymbol{\pi}^I$ and $\boldsymbol{\alpha} = [1 - \alpha_{\pi_1}, \alpha_{\pi_1}]^T$ if $\boldsymbol{\pi} = \boldsymbol{\pi}^{II}$. Based on the above, we formulate the following problem to maximize the sum-rate of the two users with given $\boldsymbol{\alpha}$ and $\boldsymbol{\pi}$ by jointly optimizing the IRS reflection coefficients and user transmit powers:

$$(P1) \quad \max_{r, p_1, p_2, \Phi^C} r \quad (17)$$

$$\text{s.t.} \quad \log_2 \left(1 + \frac{p_{\pi_1} |\tilde{h}_{\pi_1}^C(\Phi^C)|^2}{p_{\pi_2} |\tilde{h}_{\pi_2}^C(\Phi^C)|^2 + \sigma^2} \right) \geq \alpha_{\pi_1} r \quad (18)$$

$$\log_2 \left(1 + \frac{p_{\pi_2} |\tilde{h}_{\pi_2}^C(\Phi^C)|^2}{\sigma^2} \right) \geq (1 - \alpha_{\pi_1}) r \quad (19)$$

$$p_k \leq P_k, \quad \forall k \in \{1, 2\} \quad (20)$$

$$\Phi^C = \text{diag}\{\phi_1^C, \dots, \phi_M^C\} \quad (21)$$

$$|\phi_m^C| = 1, \quad \forall m \in \mathcal{M}. \quad (22)$$

For each rate-profile vector $\boldsymbol{\alpha}$, let $r_1^*(\boldsymbol{\alpha})$ and $r_{II}^*(\boldsymbol{\alpha})$ denote the optimal value to (P1) with $\boldsymbol{\pi} = \boldsymbol{\pi}^I$ and $\boldsymbol{\pi} = \boldsymbol{\pi}^{II}$, respectively. Note that $r_1^*(\boldsymbol{\alpha}) \geq r_{II}^*(\boldsymbol{\alpha})$ represents that decoding order $\boldsymbol{\pi}^I$ is optimal for the given $\boldsymbol{\alpha}$, and $r_1^*(\boldsymbol{\alpha}) < r_{II}^*(\boldsymbol{\alpha})$ otherwise. Therefore, the Pareto-optimal rate-pair (R_1^C, R_2^C) along the rate-profile vector $\boldsymbol{\alpha}$ denoted by $(R_1^{C*}(\boldsymbol{\alpha}), R_2^{C*}(\boldsymbol{\alpha}))$ is given by $(\alpha_1, 1 - \alpha_1) \max(r_1^*(\boldsymbol{\alpha}), r_{II}^*(\boldsymbol{\alpha}))$. Moreover, we prove below that although only the SIC-achievable rate-pairs are considered in (P1), its optimal solutions $(R_1^{C*}(\boldsymbol{\alpha}), R_2^{C*}(\boldsymbol{\alpha}))$'s are able to fully characterize the capacity region \mathcal{C}^C .

Proposition 1: The capacity region of the IRS-aided two-user MAC in (3) under centralized IRS deployment is

$$\mathcal{C}^C = \text{Conv}((0, 0) \cup_{\boldsymbol{\alpha}: \alpha_1 \in [0, 1]} (R_1^{C*}(\boldsymbol{\alpha}), R_2^{C*}(\boldsymbol{\alpha}))). \quad (23)$$

Proof: Please refer to Appendix A. ■

Proposition 1 indicates that the optimal solutions to (P1) with different $\boldsymbol{\alpha}$'s provide an alternative characterization of \mathcal{C}^C . However, (P1) is a non-convex optimization problem due to the uni-modular constraints on ϕ_m^C 's in (22) as well as the complicated coupling among p_1 , p_2 , and Φ^C in (18)–(19). Therefore, finding the optimal solution to (P1) is generally a difficult task. In the following, we propose a high-quality suboptimal solution to (P1), based on which an inner bound of the capacity region can be obtained.

B. Capacity Region Inner Bound

In this subsection, we derive an inner bound of the capacity region \mathcal{C}^C (or an achievable rate region) based on the rate-profile method presented above. Specifically, we propose an AO based algorithm to find a high-quality suboptimal solution to the sum-rate maximization problem (P1) efficiently.

First, we exploit the structure of (P1) to transform it into a more tractable form.

Proposition 2: (P1) is equivalent to the following problem:

$$(P2) \quad \max_{r, \Phi^C: (21), (22)} r \quad (24)$$

$$\text{s.t.} \quad \log_2 \left(1 + \frac{P_{\pi_1} |\tilde{h}_{\pi_1}^C(\Phi^C)|^2}{2^{(1 - \alpha_{\pi_1})r} \sigma^2} \right) \geq \alpha_{\pi_1} r \quad (25)$$

$$\log_2 \left(1 + \frac{P_{\pi_2} |\tilde{h}_{\pi_2}^C(\Phi^C)|^2}{\sigma^2} \right) \geq (1 - \alpha_{\pi_1}) r \quad (26)$$

Proof: Proposition 2 can be proved by noting that the inequality in (19) can be replaced with equality without loss of optimality. We omit the details here for brevity. ■

Note that for the case of $\alpha_{\pi_1} = 1$, the optimal Φ^C to (P2) can be readily derived as $\phi_m^C = e^{j(\arg\{\tilde{h}_{\pi_1}\} - \arg\{g_m^C h_{\pi_1 m}^C\})}$, $\forall m$, similar to (9). Thus, we focus on solving (P2) with $\alpha_{\pi_1} \in [0, 1)$ in the next. To further simplify (P2), we define an auxiliary variable $\beta \triangleq 2^{(1 - \alpha_{\pi_1})r}$, which is an increasing function of r for any $\alpha_{\pi_1} \in [0, 1)$. (P2) is then equivalently rewritten as

$$(P3) \quad \max_{\beta, \Phi^C: (21), (22)} \beta \quad (27)$$

$$\text{s.t.} \quad |\tilde{h}_{\pi_1}^C(\Phi^C)|^2 \geq \frac{(\beta^{\frac{1}{1 - \alpha_{\pi_1}}} - \beta)\sigma^2}{P_{\pi_1}} \quad (28)$$

$$|\tilde{h}_{\pi_2}^C(\Phi^C)|^2 \geq \frac{(\beta - 1)\sigma^2}{P_{\pi_2}}. \quad (29)$$

(P3) is still non-convex due to the uni-modular constraints on ϕ_m^C 's as well as the quadratic terms on the left-hand sides (LHSs) of (28) and (29), for which the optimal solution is difficult to obtain. In the following, we adopt an AO approach for finding a high-quality suboptimal solution to (P3). Specifically, note that each quadratic term $|\tilde{h}_k^C(\Phi^C)|^2$ can be expressed as the following *affine* form over each ϕ_m^C with the other reflection coefficients $\{\phi_i^C, i \neq m\}_{i=1}^M$ being fixed:

$$|\tilde{h}_k^C(\Phi^C)|^2 = 2\Re\{f_{2,km}\phi_m^C\} + f_{1,km}, \quad k = 1, 2, \quad (30)$$

where $f_{1,km} \triangleq |\tilde{h}_k + \sum_{i \neq m} g_i^C \phi_i^C h_{ki}^C|^2 + |g_m^C h_{km}^C|^2$ and $f_{2,km} \triangleq g_m^C h_{km}^C (\tilde{h}_k^* + \sum_{i \neq m} g_i^{C*} \phi_i^{C*} h_{ki}^{C*})$, and the equality in (30) holds due to $|\phi_m^C| = 1$. Hence, with given $\{\phi_i^C, i \neq m\}_{i=1}^M$, (P3) is reduced to the following problem:

$$(P3-m) \quad \max_{\beta} \beta \quad (31)$$

$$\phi_m^C: |\phi_m^C| = 1$$

$$\text{s.t.} \quad 2\Re\{f_{2,\pi_1 m}\phi_m^C\} \geq \frac{(\beta^{\frac{1}{1 - \alpha_{\pi_1}}} - \beta)\sigma^2}{P_{\pi_1}} - f_{1,\pi_1 m} \quad (32)$$

$$2\Re\{f_{2,\pi_2 m}\phi_m^C\} \geq \frac{(\beta - 1)\sigma^2}{P_{\pi_2}} - f_{1,\pi_2 m}. \quad (33)$$

Note that the only non-convexity in (P3-m) lies in the uni-modular constraint on ϕ_m^C , thus motivating us to apply the *convex relaxation* technique on this constraint. Specifically, we relax (P3-m) by replacing the constraint $|\phi_m^C| = 1$ with a new convex constraint $|\phi_m^C| \leq 1$, and denote the relaxed problem as (P3-m-R). We then have the following proposition.

Proposition 3: There exists an optimal solution of ϕ_m^C to (P3-m-R) that satisfies $|\phi_m^C| = 1$.

Proof: Please refer to Appendix B. ■

Proposition 3 indicates that the convex relaxation from (P3-m) to (P3-m-R) is *tight*, and the optimal solution to (P3-

m-R) that satisfies $|\phi_m^C| = 1$ is also optimal for (P3-m). Thanks to the above transformations, (P3-m-R) is a convex optimization problem, which can be efficiently solved via the interior-point method with complexity $\mathcal{O}(1)$. If the obtained optimal solution does not satisfy $|\phi_m^C| = 1$, another optimal solution with $|\phi_m^C| = 1$ can be constructed via proper scaling and rotation according to Appendix B, with complexity $\mathcal{O}(1)$. Therefore, by iteratively optimizing (β, ϕ_m^C) with all the other variables $\{\phi_i^C, i \neq m\}_{i=1}^M$ being fixed at each time via solving (P3-m), we can obtain a feasible solution to (P3) as well as (P1), which is in general suboptimal. An initial point for the above algorithm can be found by randomly generating $Q > 1$ realizations of Φ^C with the phase of each ϕ_m^C following uniform distribution in $[0, 2\pi)$, and selecting the realization with the largest sum-rate under the given rate-profile vector. Note that since (P3-m) is solved *optimally* in every iteration, the objective value of (P3), β , is non-decreasing over the iterations, which guarantees the *monotonic convergence* of this algorithm since the sum-rate r and hence β is bounded above due to the finite transmit powers. For each α , let $\tilde{r}_I(\alpha)$ and $\tilde{r}_{II}(\alpha)$ denote the obtained solutions to (P1) with $\pi = \pi^I$ and $\pi = \pi^{II}$, respectively. Between their corresponding rate-pairs, we further select the one with larger sum-rate as

$$(\tilde{R}_1^C(\alpha), \tilde{R}_2^C(\alpha)) = (\alpha_1, 1 - \alpha_1) \max(\tilde{r}_I(\alpha), \tilde{r}_{II}(\alpha)). \quad (34)$$

By performing time sharing among the obtained $(\tilde{R}_1^C(\alpha), \tilde{R}_2^C(\alpha))$'s, an inner bound of the capacity region (or an achievable rate region) for the centralized IRS case is obtained as

$$\mathcal{C}_1^C = \text{Conv}\left((0, 0) \bigcup_{\alpha: \alpha_1 \in [0, 1]} (\tilde{R}_1^C(\alpha), \tilde{R}_2^C(\alpha))\right) \subseteq \mathcal{C}^C. \quad (35)$$

Note that the complexity for obtaining the above proposed solution to (P1) with both decoding orders can be shown to be $\mathcal{O}(2M(Q + I))$, where I denotes the number of outer iterations (each requires solving (P3-m) for M times from $m = 1$ to $m = M$); moreover, the complexity for taking the convex hull of $L + 1$ points in (35) is $\mathcal{O}(L \log L)$. Therefore, by approximating the $[0, 1]$ range of the rate ratio α_1 with L uniformly sampled points, the overall complexity for obtaining \mathcal{C}_1^C is $\mathcal{O}(2M(Q + I)L + L \log L)$, which is *polynomial* over M and thus much lower than that of the exhaustive search (i.e., $\mathcal{O}(L_0^M)$).

C. Capacity Region Outer Bound

Next, we provide an outer bound of the capacity region \mathcal{C}^C . Specifically, it follows from (13)–(15) that an outer bound of \mathcal{C}^C can be constructed by finding an upper bound for each of $r_1^C(\Phi^C)$, $r_2^C(\Phi^C)$, and $r_{12}^C(\Phi^C)$ separately, for which the details are given as follows.

First, similar to (8), it can be shown that for each user k , the effective channel gain $|\tilde{h}_k^C(\Phi^C)|$ is upper-bounded by

$$|\tilde{h}_k^C(\Phi^C)| \leq |\tilde{h}_k| + \sum_{m=1}^M |g_m^C| |h_{km}^C| \triangleq \tilde{h}_{k,U}^C, \quad (36)$$

where the inequality holds with equality if and only if all the IRS reflection coefficients are designed to maximize user k 's effective channel gain, i.e.,

$$\phi_m^C = e^{j(\arg\{h_k\} - \arg\{g_m^C h_{km}^C\})}, \quad m \in \mathcal{M}. \quad (37)$$

Thus, based on (13)–(14), each $r_k^C(\Phi^C)$ is upper-bounded as $r_k^C(\Phi^C) \leq \log_2(1 + P_k \tilde{h}_{k,U}^C / \sigma^2) \triangleq r_{k,U}^C$, $k = 1, 2$. (38)

Next, we derive an upper bound for $r_{12}^C(\Phi^C)$, which is a challenging task since Φ^C can change both $\tilde{h}_1^C(\Phi^C)$ and

$\tilde{h}_2^C(\Phi^C)$ in $r_{12}^C(\Phi^C)$. To achieve this goal, we formulate the following optimization problem:

$$(P4) \quad \max_{\Phi^C: |\phi_m^C|=1, \forall m \in \mathcal{M}} P_1 |\tilde{h}_1^C(\Phi^C)|^2 + P_2 |\tilde{h}_2^C(\Phi^C)|^2. \quad (39)$$

Let s_0^* denote the optimal value of (P4). Note that for any $s_0 \geq s_0^*$, $\log_2(1 + s_0/\sigma^2)$ is an upper bound for $r_{12}^C(\Phi^C)$. However, (P4) is a non-convex optimization problem due to the uni-modular constraints on ϕ_m^C 's, thus s_0^* is generally difficult to obtain. In the following, we find an upper bound for s_0^* instead. First, we transform (P4) into a more tractable form. Define $\mathbf{q}_k^H \triangleq \mathbf{g}^{C^T} \text{diag}\{\mathbf{h}_k^C\}$, $\mathbf{v} \triangleq P_1 \tilde{h}_1 \mathbf{q}_1 + P_2 \tilde{h}_2 \mathbf{q}_2$, and $\phi^C \triangleq [\phi_1^C, \dots, \phi_M^C]^T$. Consequently, the objective function of (P4) can be rewritten as $P_1 |\tilde{h}_1^C(\Phi^C)|^2 + P_2 |\tilde{h}_2^C(\Phi^C)|^2 = P_1 |\tilde{h}_1|^2 + P_2 |\tilde{h}_2|^2 + \mathbf{v}^H \phi^C + \phi^{C^H} \mathbf{v} + \phi^{C^H} (P_1 \mathbf{q}_1 \mathbf{q}_1^H + P_2 \mathbf{q}_2 \mathbf{q}_2^H) \phi^C$, which is a quadratic function of ϕ^C . Thus, we can apply the SDR technique for finding an upper bound for the optimal value of (P4). By introducing auxiliary variables $\mathbf{w} = [\phi^{C^T}, t]^T$ and $\mathbf{W} = \mathbf{w} \mathbf{w}^H$, (P4) can be shown to be equivalent to the following problem with an additional constraint of $\text{rank}(\mathbf{W}) = 1$:

$$(P4\text{-SDR}) \quad \max_{\mathbf{W}} P_1 |\tilde{h}_1|^2 + P_2 |\tilde{h}_2|^2 + \text{tr}\{\mathbf{W} \mathbf{Q}\} \quad (40)$$

$$\text{s.t.} \quad \mathbf{W} \succeq \mathbf{0} \quad (41)$$

$$[\mathbf{W}]_{m,m} = 1, m = 1, \dots, M + 1, \quad (42)$$

where $\mathbf{Q} \triangleq [P_1 \mathbf{q}_1 \mathbf{q}_1^H + P_2 \mathbf{q}_2 \mathbf{q}_2^H, \mathbf{v}; \mathbf{v}^H, 0]$. (P4-SDR) is a semi-definite program (SDP) which can be efficiently solved via the interior-point method with complexity $\mathcal{O}(M^{6.5})$ [35]. Denote s^* as the optimal value of (P4-SDR). Note that $s^* \geq s_0^*$ holds due to the relaxation of the rank-one constraint. Thus, we have $r_{12}^C(\Phi^C) \leq \log_2(1 + s^*/\sigma^2) \triangleq r_{12,U}^C$, which yields an outer bound of \mathcal{C}^C given by

$$\mathcal{C}_0^C = \{(R_1^C, R_2^C) : R_1^C \leq r_{1,U}^C, R_2^C \leq r_{2,U}^C, R_1^C + R_2^C \leq r_{12,U}^C\} \supseteq \mathcal{C}^C. \quad (43)$$

Besides the above bounds on the capacity region, we characterize the achievable rate regions with TDMA and FDMA for centralized IRS deployment as follows.

D. Achievable Rate Region with TDMA

With TDMA, the centralized IRS should apply two different sets of reflection coefficients over the two time slots, each tailored for one of the two users without loss of generality. Let $\Phi^C[k]$ denote the reflection coefficients at the time slot allocated to the k th user. For any given $\{\Phi^C[k]\}$, the TDMA achievable rate region is given by $\mathcal{R}_T^C(\{\Phi^C[k]\}) = \bigcup_{\rho_T \in [0, 1]} \{(R_1^C, R_2^C) : R_1^C \leq \rho_T \log_2(1 + P_1 |\tilde{h}_1^C(\Phi^C[1])|^2 / \sigma^2), R_2^C \leq (1 - \rho_T) \log_2(1 + P_2 |\tilde{h}_2^C(\Phi^C[2])|^2 / \sigma^2)\}$. Note that the channel gain of the k th user at its assigned time slot, $|\tilde{h}_k^C(\Phi^C[k])|$, can be maximized as $\tilde{h}_{k,U}^C$ by setting $\Phi^C[k]$ as Φ^C given in (37). Hence, similar to the distributed IRS case, the TDMA achievable rate region with centralized IRS is given by

$$\mathcal{R}_T^C = \bigcup_{\rho_T \in [0, 1]} \left\{ (R_1^C, R_2^C) : R_1^C \leq \rho_T \log_2 \left(1 + \frac{P_1 \tilde{h}_{1,U}^C{}^2}{\sigma^2} \right), R_2^C \leq (1 - \rho_T) \log_2 \left(1 + \frac{P_2 \tilde{h}_{2,U}^C{}^2}{\sigma^2} \right) \right\}. \quad (44)$$

E. Achievable Rate Region with FDMA

With FDMA, the achievable rate region for any given Φ^C is given by $\mathcal{R}_F^C(\Phi^C) = \bigcup_{\rho_F \in [0,1]} \{(R_1^C, R_2^C) : R_1^C \leq \rho_F \log_2(1 + P_1 |\tilde{h}_1^C(\Phi^C)|^2 / (\rho_F \sigma^2)), R_2^C \leq (1 - \rho_F) \log_2(1 + P_2 |\tilde{h}_2^C(\Phi^C)|^2 / ((1 - \rho_F) \sigma^2))\}$. After time sharing among different Φ^C 's, the overall achievable rate region is given by $\mathcal{R}_F^C = \text{Conv}(\bigcup_{\Phi^C \in \mathcal{F}^C} \mathcal{R}_F^C(\Phi^C))$. Similar to the characterization of \mathcal{C}^C in Section IV-A, \mathcal{R}_F^C can be characterized via the rate-profile based method by solving the following problem for $\alpha_1 \in [0, 1]$:

$$(P5) \max_{r, \rho_F} r \quad (45)$$

$$\Phi^C: (21), (22)$$

$$\text{s.t. } \rho_F \log_2 \left(1 + \frac{P_1 |\tilde{h}_1^C(\Phi^C)|^2}{\rho_F \sigma^2} \right) \geq \alpha_1 r \quad (46)$$

$$(1 - \rho_F) \log_2 \left(1 + \frac{P_2 |\tilde{h}_2^C(\Phi^C)|^2}{(1 - \rho_F) \sigma^2} \right) \geq (1 - \alpha_1) r \quad (47)$$

$$0 \leq \rho_F \leq 1. \quad (48)$$

Note that (P5) is a non-convex optimization problem due to the complicated coupling between ρ_F and Φ^C . Nevertheless, a suboptimal solution to (P5) can be found via AO in a similar manner as that for (P1). For brevity, the details are given in Appendix C. Based on this, an inner bound of \mathcal{R}_F^C , denoted as $\mathcal{R}_{F,1}^C$, can be similarly obtained as \mathcal{C}_1^C . On the other hand, we have $\mathcal{R}_F^C(\Phi^C) \subseteq \mathcal{C}^C(\Phi^C)$ for any Φ^C [33], and consequently $\mathcal{R}_F^C \subseteq \mathcal{C}^C$. Therefore, the outer bound for the capacity region \mathcal{C}^C , \mathcal{C}_O^C , is also an outer bound of \mathcal{R}_F^C , i.e., $\mathcal{R}_F^C \subseteq \mathcal{R}_{F,O}^C \triangleq \mathcal{C}_O^C$.

Finally, we show that $\mathcal{R}_T^C \subseteq \mathcal{R}_F^C$. Let $\Phi_{(k)}^C$ denote the IRS reflection coefficients that achieve the maximum effective channel gain for the k th user, $\tilde{h}_{k,U}^C$, which are given in (37). It can be easily shown that time sharing of the FDMA achievable rate regions with $\Phi_{(1)}^C$ and $\Phi_{(2)}^C$ suffices to contain the TDMA achievable rate region due to the curved Pareto boundary of $\mathcal{R}_F^C(\Phi^C)$. Therefore, we have $\mathcal{R}_T^C \subseteq \text{Conv}(\mathcal{R}_F^C(\Phi_{(1)}^C) \cup \mathcal{R}_F^C(\Phi_{(2)}^C)) \subseteq \mathcal{R}_F^C$, namely, the FDMA achievable rate region contains the TDMA achievable rate region. It is worth noting that although with any given Φ^C , the FDMA achievable rate region may not contain that of TDMA, time sharing over different Φ^C 's enables FDMA to outperform TDMA in terms of achievable rate region.

V. PERFORMANCE COMPARISON: DISTRIBUTED VS. CENTRALIZED IRS DEPLOYMENT

In this section, we compare the capacity regions and TDMA/FDMA achievable rate regions under the two IRS deployment strategies. For simplicity, we assume that the direct user-AP channels are negligible as compared to the IRS-reflected channels and thus $\bar{h}_1 = \bar{h}_2 = 0$, which is practically valid for IRSs with large M (and thus M_1 and M_2).⁵ Moreover, for fair comparison, we consider the following *twin* channels (defined in Assumption 1 below) between the two deployment cases, where the two distributed user-IRS channels constitute the centralized IRS-AP channel, and each user-IRS

⁵The general case with non-zero \bar{h}_1 and \bar{h}_2 is more difficult to analyze, which is thus considered for the numerical example in Section VII.

channel in the centralized case contains the corresponding IRS-AP channel in the distributed case.

Assumption 1 (Twin Channels): For the channel coefficients illustrated in Fig. 1, we assume $\mathbf{g}^C = [\mathbf{h}_1^{D^T}, \mathbf{h}_2^{D^T}]^T$, $h_{1m}^C = g_{1m}^D$, $\forall m \in \mathcal{M}_1$, and $h_{2(m+M_1)}^C = g_{2m}^D$, $\forall m \in \mathcal{M}_2$.

The above twin channels hold in practice if the user-IRS channels in the distributed case have the same statistical distribution and link distance as the IRS-AP channel in the centralized case, and the IRS-AP channels in the distributed case have the same statistical distribution and link distance as the corresponding user-IRS channels in the centralized case (see Fig. 1).⁶

A. Capacity Region Comparison

First, we have the following proposition for the capacity region comparison.

Proposition 4: Under $\bar{h}_1 = \bar{h}_2 = 0$ and Assumption 1, the capacity region of the centralized IRS deployment contains that of the distributed IRS deployment, i.e., $\mathcal{C}^D \subseteq \mathcal{C}^C$.

Proof: We construct $\tilde{\Phi}^C$ for the centralized IRS such that the reflection coefficients of its two sub-surfaces, $\{\tilde{\phi}_m^C\}_{m=1}^{M_1}$ and $\{\tilde{\phi}_m^C\}_{m=M_1+1}^M$, correspond to the capacity-achieving reflection coefficients at IRS 1 and 2 for the distributed deployment shown in (9), respectively, but each being rotated by a common phase $\theta_1 \in [0, 2\pi)$ or $\theta_2 \in [0, 2\pi)$, i.e.,

$$\tilde{\phi}_m^C = \begin{cases} e^{j(\arg\{\tilde{h}_1\} - \arg\{g_{1m}^D h_{1m}^D\} + \theta_1)}, & m \in \mathcal{M}_1 \\ e^{j(\arg\{\tilde{h}_2\} - \arg\{g_{2(m-M_1)}^D h_{2(m-M_1)}^D\} + \theta_2)}, & m \in \mathcal{M} \setminus \mathcal{M}_1. \end{cases} \quad (49)$$

Then, we have $|\tilde{h}_k^C(\tilde{\Phi}^C)| = |\tilde{h}_{k,U}^D + \tilde{f}_k e^{j(\theta_2 - \theta_1)}|$, with $\tilde{f}_1 = \sum_{m=M_1+1}^M g_m^C h_{1m}^C e^{-j \arg\{g_{2(m-M_1)}^D h_{2(m-M_1)}^D\}}$ and $\tilde{f}_2 = \sum_{m=1}^{M_1} g_m^C h_{2m}^C e^{j \arg\{g_{1m}^D h_{1m}^D\}}$ (recall $\tilde{h}_{k,U}^D$'s are the capacity-achieving effective channel gains for the distributed case). Then, we prove that we can always design θ_1 and θ_2 such that $|\tilde{h}_k^C(\tilde{\Phi}^C)| \geq \tilde{h}_{k,U}^D$ holds for any $k \in \{1, 2\}$. To this end, we present Lemma 1 as below.

Lemma 1: For any complex numbers $\{a_k, b_k\}_{k=1,2}^2$, denote $c_k = \arg\{b_k\} - \arg\{a_k\} \in [0, 2\pi)$, $k = 1, 2$. Then, $|a_k + b_k e^{j\theta}| \geq |a_k|$, $k = 1, 2$ holds with $\theta = \frac{\pi}{2} - \min(c_1, c_2)$ if $|c_1 - c_2| \geq \pi$, and $\theta = \frac{\pi}{2} - \max(c_1, c_2)$ otherwise.

Proof: Please refer to Appendix D. ■

By substituting a_k and b_k with $\tilde{h}_{k,U}^D$ and \tilde{f}_k , respectively, and setting $\theta_2 - \theta_1$ as θ in Lemma 1, we have $|\tilde{h}_k^C(\tilde{\Phi}^C)| \geq \tilde{h}_{k,U}^D$ for any $k \in \{1, 2\}$, and consequently $\mathcal{C}^D \subseteq \mathcal{C}^C(\tilde{\Phi}^C) \subseteq \mathcal{C}^C$. This thus completes the proof of Proposition 4. ■

Proposition 4 indicates that by judiciously designing the IRS reflection, the larger passive beamforming gain at the centralized IRS can benefit the two users at the same time, thus yielding a larger capacity region than the case with two distributed IRSs each serving one user only.

B. Achievable Rate Region Comparison with TDMA and FDMA

For TDMA, note from (11) and (44) that the expression for \mathcal{R}_T^C is the same as \mathcal{R}_T^D by replacing $\tilde{h}_{k,U}^D$ with $\tilde{h}_{k,U}^C$. Under the

⁶For scenarios where the two deployment strategies may lead to different channel statistical distributions, their capacity comparison is more complicated and difficult to be performed analytically, which will be further discussed in Section VIII to motivate future work.

TABLE I
SUMMARY OF MAIN RESULTS FOR IRS-AIDED TWO-USER MAC

	Distributed	Centralized	Comparison (Twin Channels)
Capacity Region	\mathcal{C}^D (closed-form)	\mathcal{C}_I^C (rate-profile); \mathcal{C}_O^C (SDR) $\mathcal{C}_I^C \subseteq \mathcal{C}^C \subseteq \mathcal{C}_O^C$	$\mathcal{C}^D \subseteq \mathcal{C}^C$ (with $\bar{h}_k = 0, \forall k$)
Achievable Rate Region with TDMA	\mathcal{R}_T^D (closed-form)	\mathcal{R}_T^C (closed-form)	$\mathcal{R}_T^D \subseteq \mathcal{R}_T^C$
Achievable Rate Region with FDMA	\mathcal{R}_F^D (closed-form) $\mathcal{R}_T^D \subseteq \mathcal{R}_F^D \subseteq \mathcal{C}^D$	$\mathcal{R}_{F,I}^C$ (rate-profile); $\mathcal{R}_{F,O}^C$ $\mathcal{R}_{F,I}^C \subseteq \mathcal{R}_F^C \subseteq \mathcal{R}_{F,O}^C$ $\mathcal{R}_T^C \subseteq \mathcal{R}_F^C \subseteq \mathcal{C}^C$	$\mathcal{R}_F^D \subseteq \mathcal{R}_F^C$ (with $\bar{h}_k = 0, \forall k$)

twin channel condition specified in Assumption 1, it can be shown from (8) and (36) that the maximum effective channel gain for each user under centralized deployment (i.e., $\tilde{h}_{k,U}^C$) is larger than that under distributed deployment (i.e., $\tilde{h}_{k,U}^D$) due to the larger-size IRS available for passive beamforming, thus we have $\mathcal{R}_T^D \subseteq \mathcal{R}_T^C$, i.e., the TDMA achievable rate region under centralized deployment contains that under distributed deployment.

For FDMA, recall from the proof of Proposition 4 that we can always construct $\tilde{\Phi}^C$ such that $|\tilde{h}_k^C(\tilde{\Phi}^C)| \geq \tilde{h}_{k,U}^D$ holds for both $k = 1$ and $k = 2$ under the twin channel condition in Assumption 1 and negligible direct channels $\bar{h}_1 = \bar{h}_2 = 0$. Therefore, it follows that $\mathcal{R}_F^D \subseteq \mathcal{R}_F^C(\tilde{\Phi}^C) \subseteq \mathcal{R}_F^C$ holds, i.e., the FDMA achievable rate region under centralized deployment contains that under distributed deployment under Assumption 1 and $\bar{h}_1 = \bar{h}_2 = 0$.

The above results indicate that the superior rate performance of centralized IRS deployment over distributed IRS deployment still holds for practical OMA schemes (i.e., TDMA/FDMA) under the assumed channel conditions. For ease of reference, we summarize in Table I our main results on the capacity/rate region characterization and comparison for the two-user MAC.

VI. EXTENSION TO IRS-AIDED TWO-USER BC

In this section, we extend our capacity/rate region characterization of the uplink IRS-aided MAC to the downlink IRS-aided BC, by leveraging the celebrated *uplink-downlink duality* (or MAC-BC duality) framework [33]. For ease of exposition, we consider a dual channel setup where all the downlink channels equal to their uplink counterparts, thus the effective channels from the AP to user k under distributed and centralized deployment are the corresponding effective user-AP channels in the MAC case, i.e., $\tilde{h}_k^D(\Phi_k^D)$ and $\tilde{h}_k^C(\Phi_k^C)$ given in (1) and (3), respectively.

A. Distributed IRS Deployment

Under distributed IRS deployment, the received signal at the k th user is given by

$$y_k = \tilde{h}_k^D(\Phi_k^D)x + z_k, \quad k = 1, 2, \quad (50)$$

where $x = \sqrt{p_1}s_1 + \sqrt{p_2}s_2$ denotes the transmitted signal at the AP, with p_k and s_k denoting the transmit power and information symbol for user k , respectively; $z_k \sim \mathcal{CN}(0, \sigma^2)$ denotes the receiver noise at user k . We consider a transmit power constraint P at the AP, thus we have $p_1 + p_2 \leq P$. For any given IRS reflection coefficients $\{\Phi_k^D\}$, it follows from

the uplink-downlink duality that the capacity region of the two-user BC equals to the union set of its dual MAC capacity regions with transmit power constraint pairs (P_1, P_2) 's that satisfy $P_1 + P_2 = P$ [33], which is given by

$$\mathcal{C}_{BC}^D(\{\Phi_k^D\}) = \bigcup_{(P_1, P_2): P_1 + P_2 = P} \mathcal{C}^D(\{\Phi_k^D\}). \quad (51)$$

By considering time sharing among different $\{\Phi_k^D\}$'s, the overall capacity region is given by

$$\begin{aligned} \mathcal{C}_{BC}^D &= \text{Conv}\left(\bigcup_{\{\Phi_k^D\} \in \mathcal{F}^D} \mathcal{C}_{BC}^D(\{\Phi_k^D\})\right) \\ &= \text{Conv}\left(\bigcup_{\{\Phi_k^D\} \in \mathcal{F}^D} \bigcup_{(P_1, P_2): P_1 + P_2 = P} \mathcal{C}^D(\{\Phi_k^D\})\right) \\ &= \text{Conv}\left(\bigcup_{(P_1, P_2): P_1 + P_2 = P} \text{Conv}\left(\bigcup_{\{\Phi_k^D\} \in \mathcal{F}^D} \mathcal{C}^D(\{\Phi_k^D\})\right)\right) \\ &= \text{Conv}\left(\bigcup_{(P_1, P_2): P_1 + P_2 = P} \mathcal{C}^D\right). \end{aligned} \quad (52)$$

Note that the capacity-achieving optimal input distribution for any given $\{\Phi_k^D\}$ is the CSCG distribution with $s_k \sim \mathcal{CN}(0, 1)$, $\forall k$, which is thus also optimal for the IRS-aided BC under distributed deployment according to (52) and will be assumed in the sequel. For the purpose of exposition, we define $\alpha_P \triangleq \frac{P_1}{P} \in [0, 1]$. Recall from Section III that the capacity region for the dual MAC, \mathcal{C}^D , is derived in closed-form in Theorem 1. Hence, \mathcal{C}_{BC}^D can be characterized based on (52) by obtaining \mathcal{C}^D for every (P_1, P_2) with $P_1 + P_2 = P$ via one-dimensional search over α_P , and then taking the convex hull of their union set. By approximating the $[0, 1]$ range of the power ratio α_P with L_P uniformly sampled points, the overall complexity for obtaining \mathcal{C}_{BC}^D is $\mathcal{O}(L_P \log(L_P))$, which is dominated by the convex hull operation.

Similarly, the achievable rate regions with TDMA and FDMA can be characterized based on their MAC counterparts \mathcal{R}_T^D and \mathcal{R}_F^D (which are available in closed-form) as $\mathcal{R}_{BC,T}^D = \text{Conv}\left(\bigcup_{(P_1, P_2): P_1 + P_2 = P} \mathcal{R}_T^D\right)$ and $\mathcal{R}_{BC,F}^D = \text{Conv}\left(\bigcup_{(P_1, P_2): P_1 + P_2 = P} \mathcal{R}_F^D\right)$, respectively. Note that since $\mathcal{R}_T^D \subseteq \mathcal{R}_F^D \subseteq \mathcal{C}^D$ holds for MAC, we have $\mathcal{R}_{BC,T}^D \subseteq \mathcal{R}_{BC,F}^D \subseteq \mathcal{C}_{BC}^D$ for the dual BC.

B. Centralized IRS Deployment

Under centralized IRS deployment, the received signal at each k th user is given by

$$y_k = \tilde{h}_k^C(\Phi_k^C)x + z_k, \quad k = 1, 2. \quad (53)$$

By leveraging the uplink-downlink duality, the capacity region of the two-user BC with any given IRS reflection coefficients Φ^C is given by $\mathcal{C}_{BC}^C(\Phi^C) = \bigcup_{(P_1, P_2): P_1 + P_2 = P} \mathcal{C}^C(\Phi^C)$. Similar to the distributed IRS case, the BC capacity region can be obtained by considering time sharing among different

Φ^C 's and thus expressed in terms of the dual MAC capacity region \mathcal{C}^C as follows:

$$\begin{aligned} \mathcal{C}_{\text{BC}}^C &= \text{Conv} \left(\bigcup_{\Phi^C \in \mathcal{F}^C} \mathcal{C}_{\text{BC}}^C(\Phi^C) \right) \\ &= \text{Conv} \left(\bigcup_{(P_1, P_2): P_1 + P_2 = P} \mathcal{C}^C \right). \end{aligned} \quad (54)$$

Notice from (54) that the optimal input distribution for IRS-aided BC under the centralized deployment is still $s_k \sim \mathcal{CN}(0, 1)$, $\forall k$ due to its optimality for any given Φ^C , which is thus assumed in the sequel. Since no closed-form expression is available for \mathcal{C}^C , we find inner and outer bounds for $\mathcal{C}_{\text{BC}}^C$ based on those for \mathcal{C}^C in the following.

1) *Capacity Region Inner Bound*: Recall from Proposition 1 that \mathcal{C}^C can be characterized via the rate-profile method by solving a series of sum-rate maximization problems in (P1) with different rate-profile vectors α . Motivated by this and based on (54), we take a similar approach to characterize $\mathcal{C}_{\text{BC}}^C$. Specifically, we formulate the following sum-rate maximization problem with given rate-profile vector α and decoding order π by replacing the individual transmit power constraints in (P1) for the MAC case with a sum transmit power constraint:

$$(P6) \quad \max_{r, p_1, p_2, \Phi^C: (18), (19), (21), (22)} r \quad (55)$$

$$\text{s.t.} \quad p_1 + p_2 \leq P. \quad (56)$$

For each α , let $r_{\text{BC},\text{I}}^*(\alpha)$ and $r_{\text{BC},\text{II}}^*(\alpha)$ denote the optimal solutions to (P6) with $\pi = \pi^{\text{I}}$ and $\pi = \pi^{\text{II}}$, respectively. Similar to the MAC case elaborated in Section IV-A, the directly achievable Pareto-optimal rate-pair (without the need of time sharing/rate splitting) for the two-user BC along the rate-profile vector $\alpha = [\alpha_1, 1 - \alpha_1]^T$ is given by

$$\begin{aligned} &(R_{\text{BC},\text{I}}^{C^*}(\alpha), R_{\text{BC},\text{II}}^{C^*}(\alpha)) \\ &= (\alpha_1, 1 - \alpha_1) \max(r_{\text{BC},\text{I}}^*(\alpha), r_{\text{BC},\text{II}}^*(\alpha)). \end{aligned} \quad (57)$$

The following proposition then follows directly from the above and (54).

Proposition 5: The capacity region of the IRS-aided two-user BC with centralized IRS deployment is

$$\mathcal{C}_{\text{BC}}^C = \text{Conv} \left((0, 0) \bigcup_{\alpha: \alpha_1 \in [0, 1]} (R_{\text{BC},\text{I}}^{C^*}(\alpha), R_{\text{BC},\text{II}}^{C^*}(\alpha)) \right). \quad (58)$$

Next, we proceed to solve (P6). Similar to (P1), (P6) is a non-convex optimization problem whose optimal solution is difficult to obtain. Thus, we adopt the AO technique to find a suboptimal solution for it. Specifically, we define an auxiliary variable $\beta \triangleq 2^{(1-\alpha_{\pi_1})r}$. With any given Φ^C , (P6) is equivalent to the following optimization problem over (β, p_1, p_2) :

$$(P6\text{-P}) \quad \max_{\beta, p_1, p_2: p_1 + p_2 \leq P} \beta \quad (59)$$

$$\text{s.t.} \quad \frac{p_{\pi_1} |\tilde{h}_{\pi_1}^C(\Phi^C)|^2}{\sigma^2} \geq \beta^{\frac{1}{1-\alpha_{\pi_1}}} - \beta \quad (60)$$

$$\frac{p_{\pi_2} |\tilde{h}_{\pi_2}^C(\Phi^C)|^2}{\sigma^2} \geq \beta - 1. \quad (61)$$

(P6-P) is a convex optimization problem that can be solved efficiently with complexity $\mathcal{O}(1)$. On the other hand, with any given p_1, p_2 and $\{\phi_i^C, i \neq m\}_{i=1}^M$, (P6) is equivalent to (P3-m) by replacing each P_{π_k} with p_{π_k} , $k = 1, 2$, which is denoted as (P6-m). The optimal solution to (P6-m) can be obtained similarly as that of (P3-m) with complexity $\mathcal{O}(1)$. Therefore, similar to the proposed algorithm for (P1), by iteratively optimizing (β, p_1, p_2) or (β, ϕ_m^C) for one element $m \in \mathcal{M}$ with

all the other optimization variables being fixed at each time, a feasible solution to (P6) can be obtained, which is generally suboptimal. Note that *monotonic convergence* is guaranteed for the proposed algorithm since the *optimal* solution is found for every sub-problem, and the maximum sum-rate is bounded above. For each α , let $\tilde{r}_{\text{BC},\text{I}}(\alpha)$ and $\tilde{r}_{\text{BC},\text{II}}(\alpha)$ denote the obtained solutions to (P6) with $\pi = \pi^{\text{I}}$ and $\pi = \pi^{\text{II}}$, respectively. We further select the one with larger sum-rate between their corresponding rate-pairs as $(\tilde{R}_{\text{BC},\text{I}}^C(\alpha), \tilde{R}_{\text{BC},\text{II}}^C(\alpha))$ in a similar manner as the selection of $(R_{\text{BC},\text{I}}^{C^*}(\alpha), R_{\text{BC},\text{II}}^{C^*}(\alpha))$ in (57) by replacing $r_{\text{BC},\text{I}}^*(\alpha)$ and $r_{\text{BC},\text{II}}^*(\alpha)$ with $\tilde{r}_{\text{BC},\text{I}}(\alpha)$ and $\tilde{r}_{\text{BC},\text{II}}(\alpha)$, respectively. By performing time sharing among the obtained $(R_{\text{BC},\text{I}}^{C^*}(\alpha), R_{\text{BC},\text{II}}^{C^*}(\alpha))$'s, an inner bound of the two-user BC capacity region (or an achievable rate region) is obtained as

$$\begin{aligned} \mathcal{C}_{\text{BC},\text{I}}^C &= \text{Conv} \left((0, 0) \bigcup_{\alpha: \alpha_1 \in [0, 1]} (\tilde{R}_{\text{BC},\text{I}}^C(\alpha), \tilde{R}_{\text{BC},\text{II}}^C(\alpha)) \right) \\ &\subseteq \mathcal{C}_{\text{BC}}^C. \end{aligned} \quad (62)$$

The complexity for obtaining $\mathcal{C}_{\text{BC},\text{I}}^C$ can be shown to be $\mathcal{O}(2M(Q_{\text{BC}} + I_{\text{BC}})L + L \log L)$, with Q_{BC} denoting the number of random realizations of Φ^C in the initialization, I_{BC} denoting the number of outer iterations (each requiring solving (P6-P) once and (P6-m) for M times), and L denoting the number of points for approximating the rate ratio range $[0, 1]$.

2) *Capacity Region Outer Bound*: On the other hand, recall from Section IV-C that an outer bound for \mathcal{C}^C , \mathcal{C}_O^C , can be obtained with complexity $\mathcal{O}(M^{6.5})$. Based on this and (54), we have $\mathcal{C}_{\text{BC}}^C \subseteq \text{Conv} \left(\bigcup_{(P_1, P_2): P_1 + P_2 = P} \mathcal{C}_O^C \right) \triangleq \mathcal{C}_{\text{BC},\text{O}}^C$, thus an outer bound for $\mathcal{C}_{\text{BC}}^C$, $\mathcal{C}_{\text{BC},\text{O}}^C$, can be characterized by obtaining \mathcal{C}_O^C for all (P_1, P_2) 's that satisfy $P_1 + P_2 = P$ and taking the convex hull of their union set. The required complexity is $\mathcal{O}(L_P M^{6.5} + L_P \log L_P)$ by approximating the $[0, 1]$ range of $\alpha_P = \frac{P_1}{P}$ with L_P uniformly sampled points.

3) *Achievable Rate Regions with TDMA and FDMA*: Similar to the distributed IRS case, the achievable rate region with TDMA for the case of centralized IRS can be characterized as $\mathcal{R}_{\text{BC},\text{T}}^C = \text{Conv} \left(\bigcup_{(P_1, P_2): P_1 + P_2 = P} \mathcal{R}_{\text{T}}^C \right)$, by leveraging the closed-form expression of \mathcal{R}_{T}^C derived in Section IV-D. For FDMA, we characterize inner and outer bounds of the achievable rate region (denoted as $\mathcal{R}_{\text{BC},\text{F}}^C$) similarly as those for the MAC case. Specifically, an inner bound of $\mathcal{R}_{\text{BC},\text{F}}^C$ (denoted by $\mathcal{R}_{\text{BC},\text{F},\text{I}}^C$) can be found via the rate-profile method by solving a series of sum-rate maximization problems with different rate-profile vectors, where each problem is an extended version of (P5) for the MAC case by including the power allocations among the two users, p_1 and p_2 , as optimization variables under the constraint $p_1 + p_2 \leq P$, and replacing P_1, P_2 with p_1, p_2 . A suboptimal solution can be found for each problem via a similar AO algorithm as that for (P5) by iteratively optimizing (r, p_1, p_2) , (r, ρ_{F}) , or (r, ϕ_m^C) at each time with all the other variables being fixed. On the other hand, the capacity region outer bound $\mathcal{C}_{\text{BC},\text{O}}^C$ serves as an outer bound for $\mathcal{R}_{\text{BC},\text{F}}^C$ since $\mathcal{R}_{\text{BC},\text{F}}^C \subseteq \mathcal{C}_{\text{BC}}^C$ [33], i.e., $\mathcal{R}_{\text{BC},\text{F}}^C \subseteq \mathcal{R}_{\text{BC},\text{F},\text{O}}^C \triangleq \mathcal{C}_{\text{BC},\text{O}}^C$.

C. Performance Comparison for IRS-Aided Two-User BC

Finally, we extend the capacity/rate region comparison of the two deployment strategies in Section V for the MAC to

TABLE II
SUMMARY OF MAIN RESULTS FOR IRS-AIDED TWO-USER BC

	Distributed	Centralized	Comparison (Twin Channels)
Capacity Region	\mathcal{C}_{BC}^D (duality)	$\mathcal{C}_{BC,I}^C$ (rate-profile); $\mathcal{C}_{BC,O}^C$ (duality) $\mathcal{C}_{BC,I}^C \subseteq \mathcal{C}_{BC}^C \subseteq \mathcal{C}_{BC,O}^C$	$\mathcal{C}_{BC}^D \subseteq \mathcal{C}_{BC}^C$ (with $\bar{h}_k = 0, \forall k$)
Achievable Rate Region with TDMA	$\mathcal{R}_{BC,T}^D$ (duality)	$\mathcal{R}_{BC,T}^C$ (duality)	$\mathcal{R}_{BC,T}^D \subseteq \mathcal{R}_{BC,T}^C$
Achievable Rate Region with FDMA	$\mathcal{R}_{BC,F}^D$ (duality) $\mathcal{R}_{BC,T}^D \subseteq \mathcal{R}_{BC,F}^D \subseteq \mathcal{C}_{BC}^D$	$\mathcal{R}_{BC,F,I}^C$ (rate-profile); $\mathcal{R}_{BC,F,O}^C$ $\mathcal{R}_{BC,F,I}^C \subseteq \mathcal{R}_{BC,F}^C \subseteq \mathcal{R}_{BC,F,O}^C$ $\mathcal{R}_{BC,T}^C \subseteq \mathcal{R}_{BC,F}^C \subseteq \mathcal{C}_{BC}^C$	$\mathcal{R}_{BC,F}^D \subseteq \mathcal{R}_{BC,F}^C$ (with $\bar{h}_k = 0, \forall k$)

its dual BC.

Proposition 6: Under $\bar{h}_1 = \bar{h}_2 = 0$ and the twin channel condition in Assumption 1, the capacity region of the two-user BC under centralized IRS deployment contains that under the distributed IRS deployment, i.e., $\mathcal{C}_{BC}^D \subseteq \mathcal{C}_{BC}^C$.

Proof: Recall from Proposition 4 that $\mathcal{C}^D \subseteq \mathcal{C}^C$ holds for the two-user MAC under the channel assumptions. Hence, it follows from (52) and (54) that $\mathcal{C}_{BC}^D = \text{Conv}\left(\bigcup_{(P_1, P_2): P_1 + P_2 = P} \mathcal{C}^D\right) \subseteq \text{Conv}\left(\bigcup_{(P_1, P_2): P_1 + P_2 = P} \mathcal{C}^C\right) = \mathcal{C}_{BC}^C$. This thus completes the proof of Proposition 6. ■

Similarly, under the twin channels, by leveraging the MAC-BC duality and the results in Section V, we have $\mathcal{R}_{BC,T}^D \subseteq \mathcal{R}_{BC,T}^C$ for any \bar{h}_k 's and $\mathcal{R}_{BC,F}^D \subseteq \mathcal{R}_{BC,F}^C$ for $\bar{h}_k = 0, \forall k$. In Table II, we summarize our main results for the two-user BC.

VII. NUMERICAL EXAMPLES

In this section, we provide numerical examples to validate our analytical results. We set $M = 30$, $M_1 = M_2 = 15$ unless specified otherwise. Under a three-dimensional coordinate system, the AP is located at $(0, 0, 10)$ in meter (m), and the two users are located at $(\bar{d}_1, 0, 1)$ m and $(-\bar{d}_2, 0, 1)$ m, respectively, with \bar{d}_k denoting the horizontal AP-user distance for user k . The IRS under centralized deployment is located at $(0, 0, 9)$ m, and the two IRSs under distributed deployment are located at $(\bar{d}_1, 0, 2)$ m and $(-\bar{d}_2, 0, 2)$ m, respectively. Thus, the IRS-AP distance under centralized deployment equals to the IRS-user distances under distributed deployment, and the IRS-user distances under centralized deployment equal to the corresponding IRS-AP distances under distributed deployment. The above setup is consistent to the twin channel condition in Assumption 1. We consider the independent and identically distributed (i.i.d.) Rayleigh fading channel model,⁷ where the entries in $\{\bar{h}_k\}$, $\{h_k^C\}$ and \mathbf{g}^C are generated as independent CSCG random variables with zero mean and variance equal to the path loss of the corresponding link modeled as $\gamma = \gamma_0(1/d)^{\bar{\alpha}}$, where $\gamma_0 = -30$ dB (corresponding to a carrier frequency of 755 MHz), d denotes the link distance in m, and $\bar{\alpha}$ denotes the path loss exponent. Note that our analytical results apply to arbitrary channels and due to the space limitation, we omit the numerical results for other channel models such as Rician fading. We set $\bar{\alpha} = 3.5$ for the

⁷We assume that the elements in both distributed and centralized IRSs follow a uniform linear array (ULA) configuration with half-wavelength spacing, which usually leads to i.i.d. Rayleigh fading in an isotropic scattering environment [36].

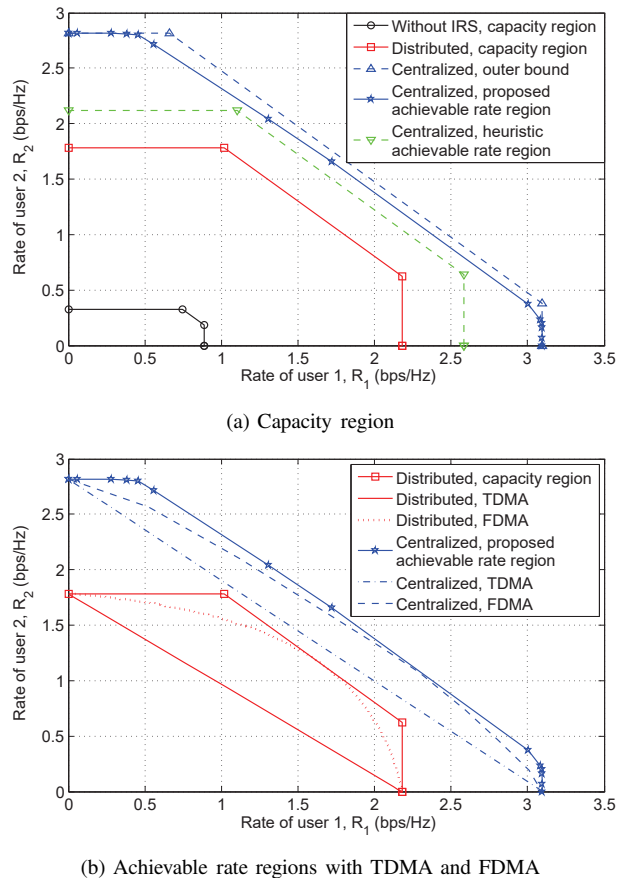
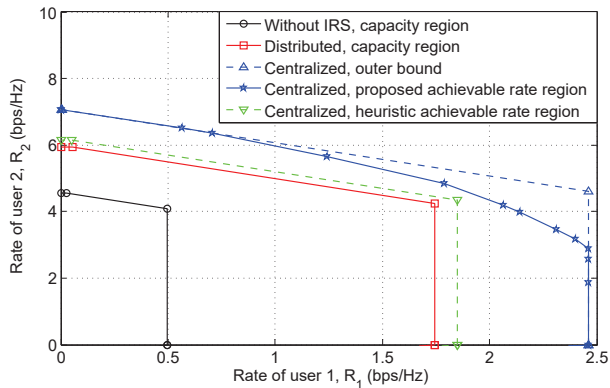


Fig. 3. Capacity/rate region comparison for IRS-aided two-user MAC under homogeneous user distance setup ($\bar{d}_1 = \bar{d}_2 = 500$ m).

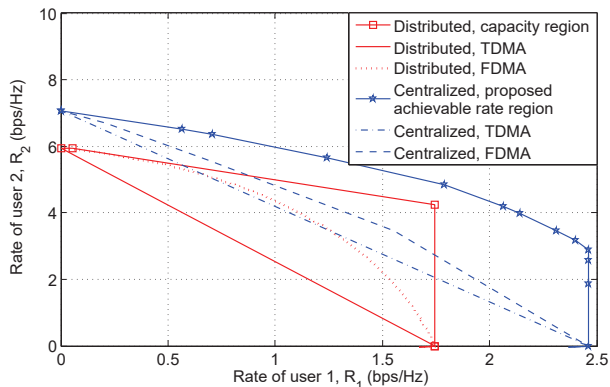
direct AP-user channels in $\{\bar{h}_k\}$ and $\bar{\alpha} = 3$ for the reflected channels in $\{h_k^C\}$ and \mathbf{g}^C . We also generate $\{h_k^D\}$ and $\{g_k^D\}$ similarly according to Assumption 1. The numbers of equally spaced points for approximating the rate ratio α_1 and the power ratio α_P are set as $L = 100$ and $L_P = 100$, respectively.

A. IRS-Aided Two-User MAC

First, we focus on the two-user MAC investigated in Sections II–V. We set $\frac{P_1}{\sigma^2} = \frac{P_2}{\sigma^2} = 120$ dB and $Q = 200$. First, we consider a *homogeneous* user distance setup where both users have a horizontal distance of $\bar{d}_1 = \bar{d}_2 = 500$ m from the AP, and randomly generate their channels based on i.i.d. Rayleigh fading. In Fig. 3 (a), we show the capacity region for the traditional MAC without IRS and that with two distributed IRSs, as well as the outer and inner capacity region bounds with a centralized IRS. It is observed that the capacity region



(a) Capacity region



(b) Achievable rate regions with TDMA and FDMA

Fig. 4. Capacity/rate region comparison for IRS-aided two-user MAC under heterogeneous user distance setup ($\bar{d}_1=500$ m, $\bar{d}_2=200$ m).

inner bound for centralized deployment contains the capacity region with distributed deployment, while the latter also contains the capacity region without IRS. This thus validates the effectiveness of deploying IRS in enlarging the capacity region as well as the advantage of centralized IRS deployment over distributed IRS deployment (even with the user-AP direct channels) under our assumed twin channel conditions. It is also interesting to observe that the capacity gain of centralized deployment over distributed deployment is more pronounced when the rates of the two users are *asymmetric*, since the larger passive beamforming gain provided by the centralized IRS is more useful for the user with larger rate requirement. In addition, we show the achievable rate region under centralized deployment by a heuristic scheme with $\tilde{\Phi}^C$ given in (49) by setting $\theta_1 = \theta_2 = 0$ (i.e., without the additional phase rotations designed for the two sub-surfaces to further align their reflected signals, as given in the proof of Proposition 4). The resultant achievable rate region is observed to be significantly smaller than our proposed one, which validates the efficacy of our proposed rate-profile based optimization. Furthermore, we show in Fig. 3 (b) the achievable rate regions with TDMA and FDMA under the two deployment strategies. It is observed that for both deployment cases, FDMA outperforms TDMA, which is consistent with our analytical results; moreover, the achievable rate region of centralized deployment contains that of distributed deployment for both TDMA and FDMA cases, even when the direct channels are present. The above results

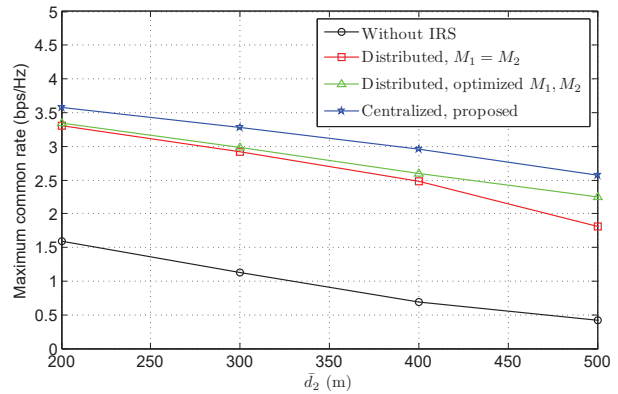
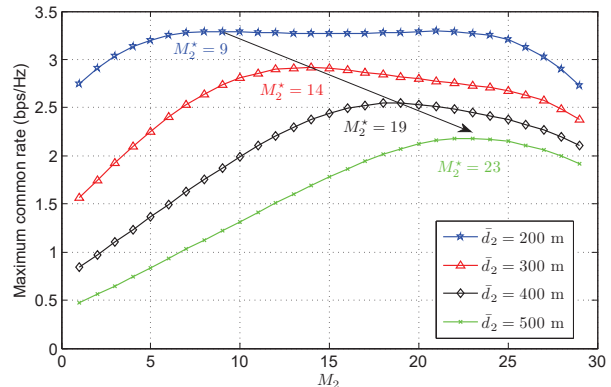
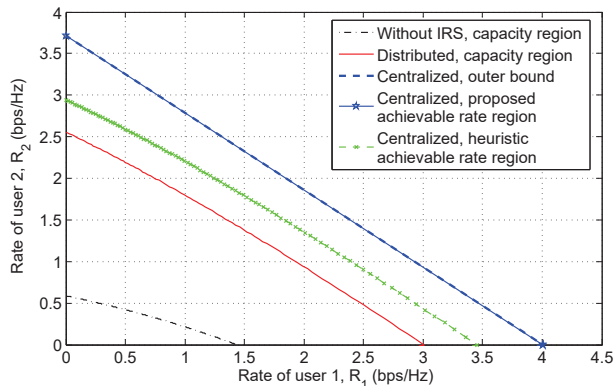
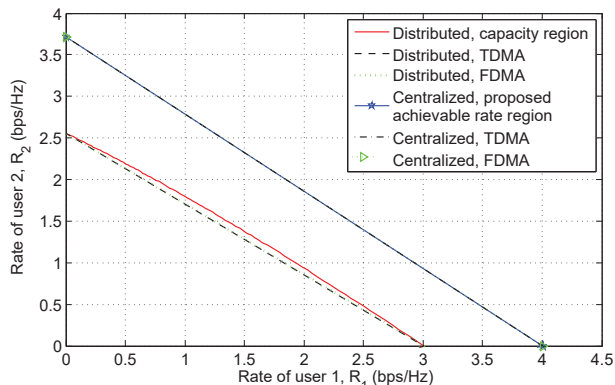
(a) Maximum common rate vs. \bar{d}_2 (b) Maximum common rate vs. M_2 under distributed deployment

Fig. 5. Maximum common rate comparison for IRS-aided two-user MAC. indicate that centralized deployment outperforms distributed deployment in both the capacity-achieving non-orthogonal multiple access (NOMA) and practical OMA schemes.

Next, we consider a *heterogeneous* user distance setup with $\bar{d}_1 = 500$ m and $\bar{d}_2 = 200$ m. In Fig. 4 (a) and Fig. 4 (b), we show the corresponding capacity region and achievable rate regions with TDMA/FDMA for different IRS deployment strategies, respectively. It is observed that the comparison results among different regions are similar as those under the homogeneous user distance setup shown in Fig. 3, as expected. However, it is observed that under the heterogeneous user distance setup, the performance gain of centralized deployment over distributed deployment is more pronounced for the user farther away from the AP, for both the NOMA and OMA schemes. This suggests that centralized deployment is more effective to alleviate the “near-far” problem and yield more fair achievable rates for the users in the network. To show this benefit more clearly, we fix $\bar{d}_1 = 200$ m and show in Fig. 5 (a) the *maximum common rate* achievable for the two users on the capacity regions of different deployment strategies versus (vs.) the horizontal AP-user distance for user 2, \bar{d}_2 , where the results are averaged over 100 independent fading channel realizations. It is observed that as \bar{d}_2 increases from 200 m (i.e., the “near-far” problem becomes more severe), the rate gain of the centralized deployment with our proposed design over distributed deployment becomes more prominent. Moreover, it is worth noting that the performance of distributed deployment can be further enhanced by optimizing the element



(a) Capacity region



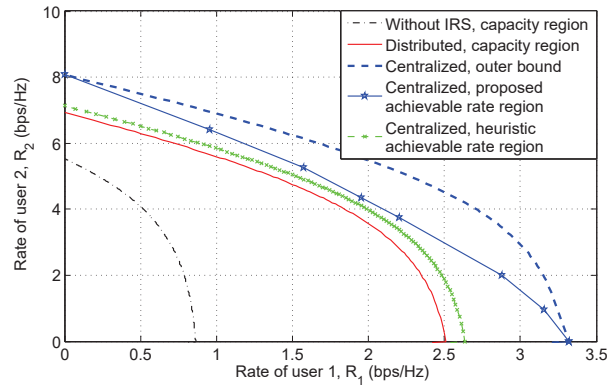
(b) Achievable rate regions with TDMA and FDMA

Fig. 6. Capacity/rate region comparison for IRS-aided two-user BC under homogeneous user distance setup ($\bar{d}_1 = \bar{d}_2 = 500$ m).

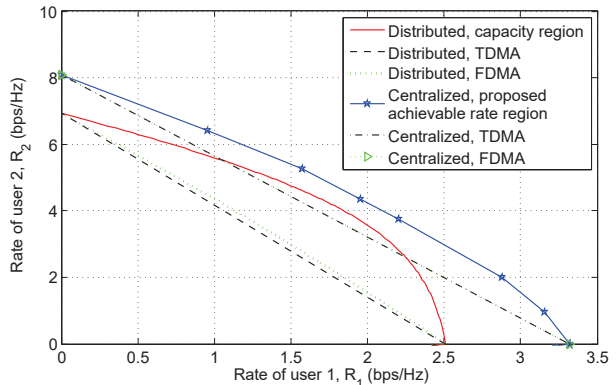
allocations among the two IRSs, i.e., M_1 and M_2 , which can be seen from Fig. 5 (b) on the maximum common rate versus M_2 . It can be observed from Fig. 5 (b) that as \bar{d}_2 increases, the optimal number of elements for IRS 2, M_2 , generally increases, i.e., more elements should be allocated to IRS 2 that serves the farther-away user 2. For example, when the two users have the same distance to the AP and thus similar path loss, the common rate is dominated by the interference from the second-decoded user on the first-decoded user, thus the two users should be allocated with slightly different numbers of elements to enhance the SIC performance (e.g., $M_2^* = 9$ or 21 when $\bar{d}_2 = 200$ m); however, when user 2 moves away from the AP, it should be decoded secondly and the common rate is dominated by its own signal power, thus more elements should be placed in its vicinity (e.g., $M_2^* = 23$ when $\bar{d}_2 = 500$ m). We also show the maximum common rate with the optimized element allocations under the distributed deployment in Fig. 5 (a), which outperforms that with equal element allocations but is also outperformed by the centralized deployment, for which there is no issue of elements allocation.

B. IRS-Aided Two-User BC

Finally, we consider the two-user BC under the setup of $\frac{P}{\sigma^2} = 123$ dB and $Q_{BC} = 200$, where the channels are dual to the two-user MAC considered above. In Fig. 6 and Fig. 7, we show the capacity and rate regions of various schemes under a homogeneous user distance setup with $\bar{d}_1 = \bar{d}_2 = 500$ m and a heterogeneous user distance setup with $\bar{d}_1 = 500$



(a) Capacity region



(b) Achievable rate regions with TDMA and FDMA

Fig. 7. Capacity/rate region comparison for IRS-aided two-user BC under heterogeneous user distance setup ($\bar{d}_1 = 500$ m, $\bar{d}_2 = 200$ m).

m, $\bar{d}_2 = 200$ m, respectively. It is observed that for both setups, our proposed achievable rate region under centralized deployment contains the capacity region under distributed deployment, while the latter also contains the capacity region without the IRS; moreover, the relationships between different capacity/rate regions are also consistent with our analytical results in Section VI. For comparison, we show a heuristic achievable rate region for the centralized deployment case by taking the convex hull of the union set of the heuristic achievable rate regions for its dual MAC shown in Section VII-A with different user transmit powers, which is observed to be substantially smaller than our proposed one based on rate-profile. In addition, the rate gain of centralized deployment is more pronounced for the farther away user from the AP under the heterogeneous distance setup, which is consistent with our results for the MAC case as discussed above. Furthermore, by comparing Fig. 6 (b) and Fig. 7 (b), it is observed that for both deployment strategies, the achievable rate regions with TDMA and FDMA approach the capacity region quite well under the homogeneous distance setup, while they are strictly suboptimal under the heterogeneous distance setup. This is because when the average channel gains of the two users are similar, the overall BC capacity region with optimized IRS reflection coefficients is approximately a triangle, which can be achieved by both TDMA and FDMA with the same IRS reflection coefficients through time sharing among the two users.

TABLE III
SUMMARY OF $\Delta\lambda$

		$\Delta\lambda$	
$\Re\{f_{2,\pi_1 m}\phi_m^C\} \geq 0, \Re\{f_{2,\pi_2 m}\phi_m^C\} \geq 0$		0	
$\Re\{f_{2,\pi_1 m}\phi_m^C\} < 0, \Re\{f_{2,\pi_2 m}\phi_m^C\} < 0$		π	
$\Re\{f_{2,\pi_i m}\phi_m^C\} \geq 0, \Re\{f_{2,\pi_j m}\phi_m^C\} \leq 0,$ $i, j \in \{1, 2\}, i \neq j$	$\eta_i + \lambda \in [0, \frac{\pi}{2}], \eta_j + \lambda \in (\frac{\pi}{2}, \pi]$	$-\arccos\{b\}$	
	$\eta_i + \lambda \in [0, \frac{\pi}{2}], \eta_j + \lambda \in [\pi, \frac{3\pi}{2})$	$\eta_j - \pi < \eta_i$	$-\eta_i - \lambda - \arccos\{b \cos(\eta_i + \lambda)\}$
		$\eta_j - \pi \geq \eta_i$	$-\eta_i - \lambda + \arccos\{b \cos(\eta_i + \lambda)\}$
	$\eta_i + \lambda \in [\frac{3\pi}{2}, 2\pi), \eta_j + \lambda \in (\pi, \frac{3\pi}{2})$		$\arccos\{b\}$
	$\eta_i + \lambda \in [\frac{3\pi}{2}, 2\pi), \eta_j + \lambda \in (\frac{\pi}{2}, \pi]$	$\eta_j + \pi \geq \eta_i$	$-\eta_i - \lambda + \arccos\{b \cos(\eta_i + \lambda)\}$
$\eta_j + \pi < \eta_i$		$-\eta_i - \lambda - \arccos\{b \cos(\eta_i + \lambda)\}$	

VIII. CONCLUDING REMARKS

This paper studied the capacity region of an IRS-aided two-user communication system, under two practical IRS deployment strategies. For the uplink IRS-aided MAC, we characterized the capacity region and achievable rate regions with TDMA/FDMA for both deployment strategies, and proved that the regions under centralized IRS deployment contain the corresponding ones under distributed IRS deployment, assuming a practical ‘‘twin channel’’ setup. The results were also extended to the downlink IRS-aided BC by leveraging the MAC-BC duality, where the performance gain of centralized over distributed IRS deployment was proved to be also valid. Numerical results validated our analysis and revealed that the superiority of centralized over distributed IRS deployment is more prominent when the two users have asymmetric rate requirements and/or channel conditions.

It is worth noting that the established rate-profile based framework for capacity region characterization can be readily extended to the more general IRS-aided K -user MAC and BC with arbitrary K , by considering all the possible decoding orders among the users. Besides, there are other appealing directions worth pursuing in future works. First, besides MAC and BC studied in this paper, it is interesting to extend our proposed framework for characterizing the capacity region of IRS-aided channels to other channel models such as the interference channel (IC), multicast channel, etc. Moreover, it is worthwhile to compare the performance of distributed versus centralized IRS deployment under more general network models such as the multi-cell multi-user network [17], where the complicated interplay between the inter-cell and intra-cell interference needs to be judiciously considered in the IRS deployment design. Finally, it is worth mentioning that the IRS deployment problem is addressed in this paper from an information-theoretic viewpoint, while from an implementation perspective, other practical factors may also need to be considered, such as the backhaul cost for information exchange (e.g., CSI), site/space constraint, availability of line-of-sight (LoS) channels, etc. Furthermore, the results in this paper are based on the assumption of identical channel distribution for both distributed and centralized IRS cases, while in practice, the channel distribution may vary at different IRS locations due to distinct terrain features and as a result, the performance comparison between the two IRS deployment strategies is more involved.

APPENDIX

A. Proof of Proposition 1

Let \mathcal{C}^{C^*} denote the right-hand side (RHS) of (23). First, the achievability of \mathcal{C}^{C^*} (i.e., $\mathcal{C}^{C^*} \subseteq \mathcal{C}^C$) is evident from the problem formulation of (P1). Next, we prove the converse, i.e., \mathcal{C}^{C^*} is an outer bound of \mathcal{C}^C . Specifically, for any given α and Φ^C , let $r_1^*(\alpha, \Phi^C)$ and $r_2^*(\alpha, \Phi^C)$ denote the optimal value of (P1) with $\pi = \pi^I$ and $\pi = \pi^{II}$, respectively, and $(R_1^{C^*}(\alpha, \Phi^C), R_2^{C^*}(\alpha, \Phi^C))$ denote the Pareto-optimal rate-pair along α defined similarly as $(R_1^C(\alpha), R_2^C(\alpha))$. We then have $\mathcal{C}^C(\Phi^C) = \text{Conv}((0, 0) \cup_{\alpha: \alpha_1 \in [0, 1]} (R_1^{C^*}(\alpha, \Phi^C), R_2^{C^*}(\alpha, \Phi^C)))$ and

$$\begin{aligned} \mathcal{C}^C &= \text{Conv}\left(\bigcup_{\Phi^C \in \mathcal{F}^C} \mathcal{C}^C(\Phi^C)\right) \\ &= \text{Conv}((0, 0) \bigcup_{\Phi^C \in \mathcal{F}^C} \bigcup_{\alpha: \alpha_1 \in [0, 1]} (R_1^{C^*}(\alpha, \Phi^C), R_2^{C^*}(\alpha, \Phi^C))) \\ &\subseteq \mathcal{C}^{C^*}, \end{aligned} \quad (63)$$

since $R_1^{C^*}(\alpha) \geq R_1^C(\alpha, \Phi^C)$ and $R_2^{C^*}(\alpha) \geq R_2^C(\alpha, \Phi^C)$ hold for any Φ^C . This completes the proof of the converse part. Consequently, we have $\mathcal{C}^C = \mathcal{C}^{C^*}$ and Proposition 1 is thus proved.

B. Proof of Proposition 3

We prove Proposition 3 by showing that for any feasible solution ϕ_m^C to (P3-m) with $|\phi_m^C| < 1$, we can always construct a new solution $\bar{\phi}_m^C$ with $|\bar{\phi}_m^C| = 1$ that yields a no smaller objective value of (P3-m). Specifically, define $f_{2,\pi_1 m} = a_1 e^{j\eta_1}$ and $f_{2,\pi_2 m} = a_2 e^{j\eta_2}$ with $a_1, a_2 \geq 0$, $\eta_1, \eta_2 \in [0, 2\pi)$; $\phi_m^C = b e^{j\lambda}$ with $0 \leq b < 1$, $\lambda \in [0, 2\pi)$; and $\bar{\phi}_m^C = e^{j(\lambda + \Delta\lambda)}$ with $|\bar{\phi}_m^C| = 1$, $\Delta\lambda \in [0, 2\pi)$. We show that we can always find a $\Delta\lambda$ such that $\Re\{f_{2,\pi_1 m}\bar{\phi}_m^C\} \geq \Re\{f_{2,\pi_1 m}\phi_m^C\}$ and $\Re\{f_{2,\pi_2 m}\bar{\phi}_m^C\} \geq \Re\{f_{2,\pi_2 m}\phi_m^C\}$ hold, and consequently, the objective value of (P3-m) with $\bar{\phi}_m^C$ is no smaller than that of ϕ_m^C since $\beta^{\frac{1}{1-\alpha\pi_1}} - \beta$ and $\beta - 1$ are non-decreasing and increasing functions of β , respectively. Due to the space limitation, we summarize the choice of such $\Delta\lambda$ for different cases in Table III, for which the detailed derivations are omitted for brevity.

C. Proposed Solution to (P5)

Note that with given Φ^C , (P5) can be shown to be a convex optimization problem over (r, ρ_F) . On the other hand, with

given ρ_F and $\{\phi_i, i \neq m\}_{i=1}^M$, (P5) is equivalent to

(P5-m)

$$\max_{r, \phi_m^C, |\phi_m^C|=1} r \quad (64)$$

$$\text{s. t. } 2\Re\{f_{2,1m}\phi_m^C\} + f_{1,1m} \geq \frac{(2^{\frac{\alpha_1 r}{\rho_F}} - 1)\rho_F \sigma^2}{P_1} \quad (65)$$

$$2\Re\{f_{2,2m}\phi_m^C\} + f_{1,2m} \geq \frac{(2^{\frac{(1-\alpha_1)r}{1-\rho_F}} - 1)(1-\rho_F)\sigma^2}{P_2}. \quad (66)$$

where $f_{1,km}$ and $f_{2,km}$ are defined in Section IV-B. Similar to (P3-m), (P5-m) can be shown to be equivalent to its relaxed version with $|\phi_m^C| = 1$ replaced by $|\phi_m^C| \leq 1$, which is a convex optimization problem and can be solved efficiently via the interior-point method. Hence, by iteratively optimizing (r, ρ_F) or (r, ϕ_m^C) for one $m \in \mathcal{M}$ with all the other variables being fixed at each time, we can obtain a feasible solution to (P5) denoted by $(\tilde{r}(\alpha), \tilde{\rho}_F(\alpha), \tilde{\Phi}^C(\alpha))$. Based on this, an inner bound of \mathcal{R}_F^C can be obtained as $\mathcal{R}_{F,I}^C = \text{Conv}((0, 0) \cup_{\alpha: \alpha_1 \in [0,1]} (\alpha_1, 1 - \alpha_1)\tilde{r}(\alpha)) \subseteq \mathcal{R}_F^C$.

D. Proof of Lemma 1

For ease of exposition, we assume that $c_2 \geq c_1$ without loss of generality. For the case of $|c_2 - c_1| \geq \pi$ and $\theta = \pi/2 - \min(c_1, c_2)$, we have $\cos(c_1 + \theta) = \cos(\pi/2) = 0$, and $c_2 + \theta = \pi/2 + c_2 - c_1 \in [3\pi/2, 5\pi/2]$, thus $\cos(c_2 + \theta) \geq 0$. On the other hand, for the case of $|c_2 - c_1| < \pi$ and $\theta = \pi/2 - \max(c_1, c_2)$, we have $\cos(c_2 + \theta) = \cos(\pi/2) = 0$, and $c_1 + \theta = \pi/2 + c_1 - c_2 \in (-\pi/2, \pi/2]$, thus $\cos(c_1 + \theta) \geq 0$. Therefore, we have $\cos(c_1 + \theta) \geq 0$ and $\cos(c_2 + \theta) \geq 0$ for both cases, and thus $|a_k + b_k e^{j\theta}| = (|a_k|^2 + |b_k|^2 + 2|a_k||b_k|\cos(c_k + \theta))^{1/2} \geq |a_k|$ holds for both $k = 1$ and $k = 2$. This thus completes the proof of Lemma 1.

REFERENCES

- [1] S. Zhang and R. Zhang, "Intelligent reflecting surface aided multiple access: Capacity region and deployment strategy," in *Proc. IEEE Int. Wkshps. Signal Process. Adv. Wireless Commun. (SPAWC)*, May 2020.
- [2] Q. Wu, S. Zhang, B. Zheng, C. You, and R. Zhang, "Intelligent reflecting surface aided wireless communications: A tutorial," *IEEE Trans. Commun.*, Early Access.
- [3] Q. Wu and R. Zhang, "Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network," *IEEE Commun. Mag.*, vol. 58, no. 1, pp. 106–112, Jan. 2020.
- [4] E. Basar, M. D. Renzo, J. Rosny, M. Debbah, M.-S. Alouini, and R. Zhang, "Wireless communications through reconfigurable intelligent surfaces," *IEEE Access*, vol. 7, pp. 116 753–116 773, 2019.
- [5] S. Zhang and R. Zhang, "On the capacity of intelligent reflecting surface aided MIMO communication," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2020.
- [6] S. Zhang and R. Zhang, "Capacity characterization for intelligent reflecting surface aided MIMO communication," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 8, pp. 1823–1838, Aug. 2020.
- [7] Y. Yang, B. Zheng, S. Zhang, and R. Zhang, "Intelligent reflecting surface meets OFDM: Protocol design and rate maximization," *IEEE Trans. Commun.*, vol. 68, no. 7, pp. 4522–4535, Jul. 2020.
- [8] J. Ye, S. Guo, and M.-S. Alouini, "Joint reflecting and precoding designs for SER minimization in reconfigurable intelligent surfaces assisted MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 19, no. 8, pp. 5561–5574, Aug. 2020.
- [9] E. Björnson, Ö. Özdogan, and E. G. Larsson, "Intelligent reflecting surface vs. decode-and-forward: How large surfaces are needed to beat relaying?" *IEEE Wireless Commun. Lett.*, vol. 9, no. 2, pp. 244–248, Feb. 2020.

- [10] M. Cui, G. Zhang, and R. Zhang, "Secure wireless communication via intelligent reflecting surface," *IEEE Wireless Commun. Lett.*, vol. 8, no. 5, pp. 1410–1414, Oct. 2019.
- [11] C. Huang, A. Zappone, G. C. Alexandropoulos, M. Debbah, and C. Yuen, "Reconfigurable intelligent surfaces for energy efficiency in wireless communication," *IEEE Trans. Wireless Commun.*, vol. 18, no. 8, pp. 4157–4170, Aug. 2019.
- [12] Q. Wu and R. Zhang, "Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming," *IEEE Trans. Wireless Commun.*, vol. 18, no. 11, pp. 5394–5409, Nov. 2019.
- [13] Q. Wu and R. Zhang, "Beamforming optimization for wireless network aided by intelligent reflecting surface with discrete phase shifts," *IEEE Trans. Commun.*, vol. 68, no. 3, pp. 1838–1851, Mar. 2020.
- [14] H. Guo, Y.-C. Liang, J. Chen, and E. G. Larsson, "Weighted sum-rate maximization for reconfigurable intelligent surface aided wireless networks," *IEEE Trans. Wireless Commun.*, vol. 19, no. 5, pp. 3064–3076, May 2020.
- [15] Y. Yang, S. Zhang, and R. Zhang, "IRS-enhanced OFDMA: Joint resource allocation and passive beamforming optimization," *IEEE Wireless Commun. Lett.*, vol. 9, no. 6, pp. 760–764, Jun. 2020.
- [16] Q. Tao, S. Zhang, C. Zhong, and R. Zhang, "Intelligent reflecting surface aided multicasting with random passive beamforming," *IEEE Wireless Commun. Lett.*, vol. 10, no. 1, pp. 92–96, Jan. 2021.
- [17] C. Pan, H. Ren, K. Wang, W. Xu, M. ElKashlan, A. Nallanathan, and L. Hanzo, "Multicell MIMO communications relying on intelligent reflecting surfaces," *IEEE Trans. Wireless Commun.*, vol. 19, no. 8, pp. 5218–5233, Aug. 2020.
- [18] 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 J. Sel. Areas Commun.*, vol. 38, no. 8, pp. 1719–1734, Aug. 2020.
- [19] X. Yu, D. Xu, Y. Sun, D. W. K. Ng, and R. Schober, "Robust and secure wireless communications via intelligent reflecting surfaces," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2637–2652, Nov. 2020.
- [20] T. Hou, Y. Liu, Z. Song, X. Sun, Y. Chen, and L. Hanzo, "Reconfigurable intelligent surface aided NOMA networks," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2575–2588, Nov. 2020.
- [21] X. Liu, Y. Liu, Y. Chen, and H. V. Poor, "RIS enhanced massive non-orthogonal multiple access networks: Deployment and passive beamforming design," [Online]. Available: <https://arxiv.org/abs/2001.10363>.
- [22] D. Xu, X. Yu, Y. Sun, D. W. K. Ng, and R. Schober, "Resource allocation for IRS-assisted full-duplex cognitive radio systems," *IEEE Trans. Commun.*, vol. 68, no. 12, pp. 7376–7394, Dec. 2020.
- [23] X. Mu, Y. Liu, L. Guo, J. Lin, and R. Schober, "Joint deployment and multiple access design for intelligent reflecting surface assisted networks," [Online]. Available: <https://arxiv.org/abs/2005.11544>.
- [24] G. Zhou, C. Pan, H. Ren, K. Wang, and A. Nallanathan, "Intelligent reflecting surface aided multigroup multicast MISO communication systems," *IEEE Trans. Signal Process.*, vol. 68, pp. 3236–3251, Apr. 2020.
- [25] R. Karasik, O. Simeone, M. Renzo, and S. Shamai, "Beyond max-SNR: Joint encoding for reconfigurable intelligent surfaces," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2020.
- [26] Y. Han, S. Zhang, L. Duan, and R. Zhang, "Cooperative double-IRS aided communication: Beamforming design and power scaling," *IEEE Wireless Commun. Lett.*, vol. 9, no. 8, pp. 1206–1210, Aug. 2020.
- [27] B. Zheng and R. Zhang, "Intelligent reflecting surface-enhanced OFDM: Channel estimation and reflection optimization," *IEEE Wireless Commun. Lett.*, vol. 9, no. 4, pp. 518–522, Apr. 2020.
- [28] Z. Wang, L. Liu, and S. Cui, "Channel estimation for intelligent reflecting surface assisted multiuser communications: Framework, algorithms, and analysis," *IEEE Trans. Wireless Commun.*, vol. 19, no. 10, pp. 6607–6620, Oct. 2020.
- [29] D. Mishra and H. Johansson, "Channel estimation and low-complexity beamforming design for passive intelligent surface-assisted MISO wireless energy transfer," in *Proc. IEEE Int. Conf. Acoustics Speech Signal Process. (ICASSP)*, May 2019.
- [30] Z.-Q. He and X. Yuan, "Cascaded channel estimation for large intelligent metasurface assisted massive MIMO," *IEEE Wireless Commun. Lett.*, vol. 9, no. 2, pp. 210–214, Feb. 2020.
- [31] C. You, B. Zheng, and R. Zhang, "Channel estimation and passive beamforming for intelligent reflecting surface: Discrete phase shift and progressive refinement," *IEEE J. Sel. Areas Commun.*, Early Access.
- [32] B. Zheng, C. You, and R. Zhang, "Intelligent reflecting surface assisted multi-user OFDMA: Channel estimation and training design," [Online]. Available: <https://arxiv.org/abs/2003.00648>.

- [33] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Wiley, 2006.
- [34] R. Zhang and S. Cui, "Cooperative interference management with MISO beamforming," *IEEE Trans. Signal Process.*, vol. 58, no. 10, pp. 5450–5458, Oct. 2010.
- [35] Z.-Q. Luo, W.-K. Ma, A. M.-C. So, Y. Ye, and S. Zhang, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 20–34, May 2010.
- [36] E. Björnson and L. Sanguinetti, "Rayleigh fading modeling and channel hardening for reconfigurable intelligent surfaces," *IEEE Wireless Commun. Lett.*, Early Access.