

A Framework for Transmission Design for Active RIS-Aided Communication with Partial CSI

Gui Zhou, Cunhua Pan, Hong Ren, Dongfang Xu, Zaichen Zhang, Jiangzhou Wang,
IEEE Fellow, and Robert Schober, *IEEE Fellow*

Abstract

Active reconfigurable intelligent surfaces (RISs) have recently been proposed to compensate for the severe multiplicative fading effect of conventional passive RIS-aided systems. Each reflecting element of active RISs is assisted by an amplifier such that the incident signal can be reflected and amplified instead of only being reflected as in passive RIS-aided systems. This work addresses the practical challenge that, on the one hand, in active RIS-aided systems the perfect individual CSI of the RIS-aided channels cannot be acquired due to the lack of signal processing power at the active RISs, but, on the other hand, this CSI is required to calculate the expected system data rate and RIS transmit power needed for transceiver design. To address this issue, we first derive closed-form expressions for the average achievable rate and the average RIS transmit power based on partial CSI of the RIS-aided channels. Then, we formulate an average achievable rate maximization problem for jointly optimizing the active beamforming at both the base station (BS) and the RIS. This problem is then tackled using the majorization–minimization (MM) algorithm framework, and, for each iteration, semi-closed-form solutions for the BS and RIS beamforming are derived based on the Karush-Kuhn-Tucker (KKT) conditions. To ensure the quality of service (QoS) of each user, we further formulate a rate outage constrained beamforming problem, which is solved using the Bernstein-Type inequality (BTI) and semidefinite relaxation (SDR) techniques. Numerical results show that the proposed algorithms can efficiently overcome the challenges imposed by imperfect CSI in active RIS-aided wireless systems.

Index Terms

Reconfigurable intelligent surface (RIS), intelligent reflecting surface (IRS), active RIS, beamforming, partial channel state information (CSI).

(Corresponding author: Cunhua Pan) Part of this work has been published in WCSP Conference [1].

G. Zhou, D. Xu, and R. Schober are with the Institute for Digital Communications, Friedrich-Alexander-University Erlangen-Nürnberg (FAU), 91054 Erlangen, Germany (email: gui.zhou, dongfang.xu, robert.schober@fau.de). C. Pan, H. Ren, and Z. Zhang are with the National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China. (cpan, hren, zczhang@seu.edu.cn). J. Wang is with the School of Engineering, University of Kent, Canterbury CT2 7NT, U.K. (e-mail: j.z.wang@kent.ac.uk).

I. INTRODUCTION

Reconfigurable intelligent surfaces (RISs) have attracted extensive research attention from both academia and industry thanks to their appealing features of low cost, low power consumption, programmability, and easy deployment [2], [3]. In fact, they are envisioned to be one of the key candidate technologies of sixth generation (6G) mobile communication systems [4], [5]. The existing literature has mainly focused on the investigation of passive RISs, where each reflecting element can only reflect the incident signals. However, passive RISs have an inherent disadvantage: the signals reflected by the RISs suffer from multiplicative fading, which causes the received signal to be extremely weak. Multiplicative fading implies that the equivalent pathloss of the transmitter-RIS-receiver link is the product of the transmitter-RIS link pathloss and the RIS-receiver link pathloss, which is typically thousands of times higher than that of the unobstructed direct BS-receiver link [6]. Most of the existing works on passive RISs bypass this issue by assuming a much larger pathloss exponent for the direct link than for the reflected links [7], [8].

To overcome the multiplicative fading effect, the authors of [9] and [10] recently proposed a new active RIS architecture. Unlike passive RISs, active RISs are additionally equipped with active reflective amplifiers. Therefore, active RISs can not only adjust the phase of the reflected signal, but also amplify the reflected signal. The authors of [10] showed that in an application scenario with direct links, passive RISs can only obtain a 3% data rate gain, while active RISs can obtain a 108% gain. In addition, they also presented a hardware platform for active RISs.

Different from traditional active antenna arrays, active RISs do not require radio frequency (RF) chains and digital signal processing circuits, such that active RISs can be relatively thin, which facilitates deployment. Since active RISs comprise amplifiers, their hardware power consumption is increased compared to passive RISs. However, the authors of [11] recently compared the performances of passive RISs and active RISs for the same total power consumption (including hardware power consumption), and showed that active RISs outperform passive RISs when the number of reflecting elements is small and the system power budget is sufficiently large. Therefore, active RISs can mitigate the multiplicative fading effect while retaining the benefits of passive RISs.

Due to the above advantages, active RISs have attracted significant research interest recently. The authors of [12] compared the performances of passive RISs and active RISs, and optimized the RIS location and number of reflecting elements. The authors of [13] investigated the resource allocation design for active RIS-aided multiuser systems. Furthermore, active RISs have been considered for wireless powered communications to enhance throughput and energy efficiency [14].

It is widely known that, for passive RIS-aided systems, only the cascaded channel state information (CSI) of the transmitter-RIS-receiver link is needed for transceiver design. However, due to the introduction of amplifiers, for active RIS-aided systems, the RIS transmit power and the thermal noise amplified by the RISs need to be taken into account for transceiver design, which requires the individual CSI of the transmitter-RIS link and the RIS-receiver link. However, active RISs cannot estimate the two individual channels as they are not equipped with an RF chain. To the best of the authors' knowledge, all existing works on active RISs assume the availability of perfect CSI of the transmitter-RIS and RIS-receiver links, respectively, which is challenging to obtain in practice. Therefore, it is imperative to study the system design for the case, where only partial CSI of the individual active RIS-aided channels is available.

Against this background, in this work, average achievable rate maximization and average power consumption minimization are addressed, respectively, if only partial CSI of the individual RIS-aided channels is available. To this end, we assume that the RIS-aided channels are Rician distributed. Although the perfect CSI of the individual RIS-aided channels is not available, knowledge of the deterministic line-of-sight (LoS) components and the statistics of the Gaussian distribution of the non-LoS (NLoS) components can be acquired. In particular, the angle and distance information of the LoS links can be determined via localization techniques [15]. Based on this partial CSI, we derive analytical expressions for the average achievable rate and the average RIS transmit power. Then, the average achievable rate is maximized by jointly optimizing the active beamforming at the BS and RIS under an RIS average transmit power constraint. Since system designs based on average achievable rate maximization cannot guarantee the QoS of each user, we further study designs based on a rate outage constrained power minimization problem.

The main contributions of this work can be summarized as follows:

- To the best of the authors' knowledge, this is the first work on active RIS-aided systems that investigates the practical issue of partial CSI knowledge. Based on the distributions of the individual RIS-aided channels, we propose a joint active beamforming design at the BS and the RIS for maximization of the average achievable rate for partial CSI. In addition, we also study the robust active beamforming design to minimize the average total power consumption under rate outage probability constraints.
- Closed-form expressions for the average achievable rate and the average RIS transmit power in the presence of partial CSI are derived. Furthermore, the average achievable rate maximization problem is efficiently solved in an iterative manner exploiting the majorization–minimization (MM) concept.

Specifically, a surrogate quadratic function for active beamforming is constructed to minorize the original non-concave objective function. Then, alternating optimization (AO) is employed to decouple the BS and RIS beamforming vectors. For each subproblem, a semi-closed-form solution is obtained based on the Karush–Kuhn–Tucker (KKT) conditions.

- To guarantee a predefined outage probability, we develop an outage constrained beamforming design that minimizes the average transmit power subject to constraints on the RIS amplification gain and the rate outage probability, respectively. The Bernstein-type inequality (BTI) is applied to safely approximate the outage probability constraints such that the non-convexity of the constraints is mitigated. Then, the beamforming vectors at both the BS and the RIS are updated by using semidefinite relaxation (SDR) in an iterative manner.
- Our simulation results demonstrate that active RISs can effectively overcome the negative impact of the multiplicative fading effect and perform much better than the conventional passive RISs. Furthermore, since the RIS amplifier circuits consume power, there exists an optimal number of RIS reflecting elements.

The rest of this paper is organized as follows. In Section II, we introduce the considered system model. The average achievable rate maximization problem and the average power minimization problem are respectively revealed in Sections III and IV. Finally, Sections V and VI report numerical results and conclusions, respectively.

Notations: The following mathematical notations and symbols are used throughout this paper. Vectors and matrices are denoted by boldface lowercase letters and boldface uppercase letters, respectively. \mathbf{X}^* , \mathbf{X}^T , \mathbf{X}^H , and $\|\mathbf{X}\|_F$ denote the conjugate, transpose, Hermitian (conjugate transpose), and Frobenius norm of matrix \mathbf{X} , respectively. $\text{vec}(\mathbf{X})$ denotes the vectorization of matrix \mathbf{X} . $\|\mathbf{x}\|_2$ denotes the L2-norm of vector \mathbf{x} . Operations $\text{Tr}\{\cdot\}$, $\text{Re}\{\cdot\}$, $|\cdot|$, $\lambda(\cdot)$, and $\angle(\cdot)$ denote the trace, real part, modulus, eigenvalue, and angle of a complex number, respectively. $\text{Diag}(\mathbf{x})$ is a diagonal matrix with the entries of \mathbf{x} on its main diagonal. Furthermore, $\text{diag}(\mathbf{X})$ is a vector whose entries are the main diagonal elements of matrix \mathbf{X} . $[\mathbf{x}]_m$ denotes the m -th element of vector \mathbf{x} . $[\mathbf{X}]_{m:n,p:q}$ is a matrix consisting of the m -th to the n -th rows and the p -th to the q -th columns of matrix \mathbf{X} . The Kronecker product, Hadamard product, and Khatri-Rao product between two matrices \mathbf{X} and \mathbf{Y} are respectively denoted by $\mathbf{X} \otimes \mathbf{Y}$, $\mathbf{X} \odot \mathbf{Y}$, and $\mathbf{X} \diamond \mathbf{Y}$. $\mathbf{X} \succeq \mathbf{Y}$ means that $\mathbf{X} - \mathbf{Y}$ is positive semidefinite. \mathbb{C} denotes the complex field, \mathbb{R} denotes the real field, and $j \triangleq \sqrt{-1}$ is the imaginary unit. $\mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ represents the distribution of a circularly symmetric complex Gaussian random vector with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$.

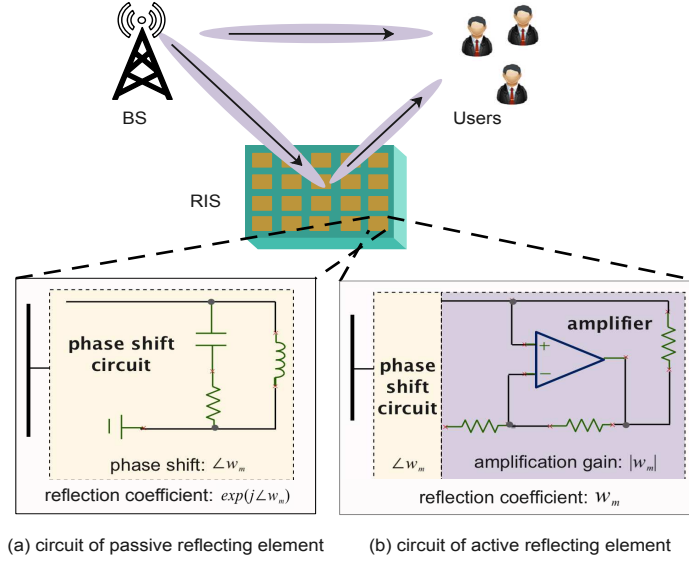


Fig. 1: Active and passive RIS-aided communication system, respectively.

II. SYSTEM MODEL

A. Signal Transmission Model

As shown in Fig. 1, we consider an RIS-aided downlink multiple-input single-output (MISO) system, where an N -antenna BS communicates with K single-antenna users. The RIS is assumed to be equipped with M reflecting elements, and its reflection coefficient matrix is given by $\mathbf{\Lambda}_{\mathbf{w}} = \text{Diag}(w_1, \dots, w_M) \in \mathbb{C}^{M \times M}$. Here, $\angle w_m$ and $|w_m|$ denote the phase shift and the reflection gain of the m -th RIS element, respectively. For passive RISs, each RIS element comprises an impedance adjustable circuit to vary the phase shift [16]. Thus, passive RISs are capable of reflecting the incident signal¹ without consuming direct-current (DC) power, which leads to a reflection gain of $|w_m|^2 = 1$ and negligible thermal noise. However, the multiplicative fading effect results in a weak received signal power for the passive RIS reflection link. To address this issue, the authors of [9] and [10] proposed a new active RIS architecture, where each active RIS element includes an active reflection-type amplifier to also amplify the incident signals. Therefore, the reflection gain is given by $1 \leq |w_m|^2 \leq a_{\max}$, where a_{\max} is the maximum amplification gain.

The BS transmits K data symbols collected in vector $\mathbf{s} = [s_1, \dots, s_K]^T \in \mathbb{C}^{K \times 1}$ to the users by applying precoder matrix $\mathbf{F} = [\mathbf{f}_1, \dots, \mathbf{f}_K] \in \mathbb{C}^{N \times K}$. By assuming independent complex Gaussian signals with $\mathbb{E}[\mathbf{s}\mathbf{s}^H] = \mathbf{I}_N$, the BS transmit power is given by $\mathbb{E}\{\|\mathbf{F}\mathbf{s}\|_2^2\} = \|\mathbf{F}\|_F^2 \leq P_N$, where P_N is

¹Here, we assume an ideal reflective material without reflection loss. If reflection loss is considered, then $|w_m|^2 < 1$.

the BS transmit power budget. Denote by $\mathbf{H}_{\text{dr}} \in \mathbb{C}^{M \times N}$ the channel from the BS to the RIS, and by $\mathbf{h}_k \in \mathbb{C}^{N \times 1}$ and $\mathbf{h}_{\text{r},k} \in \mathbb{C}^{M \times 1}$ the channels from user k to the BS and to the RIS, respectively. Then, the signal received by user k is given by

$$\begin{aligned} y_k &= \mathbf{h}_k^H \mathbf{F} \mathbf{s} + \mathbf{h}_{\text{r},k}^H \mathbf{\Lambda}_{\mathbf{w}} (\mathbf{H}_{\text{dr}} \mathbf{F} \mathbf{s} + \mathbf{z}) + n_k \\ &= (\mathbf{h}_k^H + \mathbf{h}_{\text{r},k}^H \mathbf{\Lambda}_{\mathbf{w}} \mathbf{H}_{\text{dr}}) \mathbf{F} \mathbf{s} + \mathbf{h}_{\text{r},k}^H \mathbf{\Lambda}_{\mathbf{w}} \mathbf{z} + n_k, \end{aligned} \quad (1)$$

where n_k and \mathbf{z} are the zero-mean additive white Gaussian noise (AWGN) at the user and the RIS, respectively, which follow distributions $n_k \sim \mathcal{CN}(0, \sigma_k^2)$ and $\mathbf{z} \sim \mathcal{CN}(\mathbf{0}, \sigma_z^2 \mathbf{I}_M)$ with noise powers σ_k^2 and σ_z^2 , respectively. Notice that the thermal noise \mathbf{z} , which can be ignored in passive RIS-aided systems, has to be considered in active RIS-aided systems because of the amplification. The transmit power of the active RIS is given by

$$\mathbb{E}\{||\mathbf{\Lambda}_{\mathbf{w}}(\mathbf{H}_{\text{dr}} \mathbf{F} \mathbf{s} + \mathbf{z})||_2^2\} = ||\mathbf{\Lambda}_{\mathbf{w}} \mathbf{H}_{\text{dr}} \mathbf{F}||_F^2 + ||\mathbf{w}||_2^2 \sigma_z^2, \quad (2)$$

where $\mathbf{w} = [w_1, \dots, w_M]^H$. Furthermore, the achievable rate of user k is given by

$$R_k(\mathbf{F}, \mathbf{w}) = \log_2 \left(1 + \frac{|(\mathbf{h}_k^H + \mathbf{h}_{\text{r},k}^H \mathbf{\Lambda}_{\mathbf{w}} \mathbf{H}_{\text{dr}}) \mathbf{f}_k|^2}{\sum_{i=1, i \neq k}^K |(\mathbf{h}_i^H + \mathbf{h}_{\text{r},i}^H \mathbf{\Lambda}_{\mathbf{w}} \mathbf{H}_{\text{dr}}) \mathbf{f}_i|^2 + \sigma_k^2 + ||\mathbf{h}_{\text{r},k}^H \mathbf{\Lambda}_{\mathbf{w}}||_2^2 \sigma_z^2} \right). \quad (3)$$

B. Channel Model

In practice, not all of the channels connecting the BS and the users can be individually estimated. Specifically, the direct BS-user channels $\{\mathbf{h}_k\}_{k=1}^K$ can be estimated by turning off the RIS [17]. Thus, it is reasonable to assume that perfect CSI of the direct BS-user channels is available at the BS. However, for the passive/active RIS-aided channels, we can only estimate the cascaded BS-RIS-user channel \mathbf{G}_k , which is the product of the BS-RIS channel \mathbf{H}_{dr} and the RIS-user channel $\mathbf{h}_{\text{r},k}$, i.e., $\mathbf{G}_k = \text{diag}(\mathbf{h}_{\text{r},k}^H) \mathbf{H}_{\text{dr}}$. As a result, we cannot estimate \mathbf{H}_{dr} and $\mathbf{h}_{\text{r},k}$ individually due to the lack of signal processing capability at the passive/active RIS [18]. There is an extensive literature on cascaded CSI estimation in RIS-aided communication systems [17], [19]–[21]. Thus, in this work, we also assume that $\{\mathbf{G}_k\}_{k=1}^K$ is perfectly known at the BS. In passive RIS-aided communication systems, knowledge of the CSI of the cascaded channels is typically sufficient for transceiver and RIS reflection phase shift design [8], [22]–[24]. However, in active RIS-aided systems, the instantaneous RIS transmit power in (2) and the instantaneous achievable rate in (3) depend on the individual instantaneous CSI of \mathbf{H}_{dr} and $\mathbf{h}_{\text{r},k}$, respectively, and this CSI is impossible to obtain. To address this issue, in this work, we focus on the investigation of the average achievable rate and the average RIS transmit power based on statistical CSI of \mathbf{H}_{dr} and $\mathbf{h}_{\text{r},k}$.

In particular, \mathbf{H}_{dr} and $\mathbf{h}_{r,k}$ are modelled as correlated Rician fading channels as follows

$$\mathbf{H}_{\text{dr}} = \sqrt{\beta_0/(\delta_0 + 1)}(\sqrt{\delta_0}\bar{\mathbf{H}}_{\text{dr}} + \tilde{\mathbf{H}}_{\text{dr}}), \quad (4)$$

$$\mathbf{h}_{r,k} = \sqrt{\beta_k/(\delta_k + 1)}(\sqrt{\delta_k}\bar{\mathbf{h}}_{r,k} + \tilde{\mathbf{h}}_{r,k}), \forall k, \quad (5)$$

where $\{\delta_k\}_{k=0}^K$ are the Rician factors, and $\{\beta_k\}_{k=0}^K$ are the distance-dependent large-scale pathloss coefficients. The LoS components $\bar{\mathbf{H}}_{\text{dr}}$ and $\{\bar{\mathbf{h}}_{r,k}\}_{k=1}^K$ are determined by the angles-of-arrival (AoAs) and the angles-of-departure (AoDs) [25]. The physical positions of the BS and the RIS are generally fixed and known in advance, and the users' locations can be determined by GPS positioning [26] or pilot-based positioning algorithms [27]. Thus, the communication distance and LoS angle information can be assumed to be known by the BS. The NLoS components are distributed as $\tilde{\mathbf{H}}_{\text{dr}} \sim \mathcal{CN}(\mathbf{0}, \Sigma_{\text{R}} \otimes \Sigma_{\text{B}})$ and $\tilde{\mathbf{h}}_{r,k} \sim \mathcal{CN}(\mathbf{0}, \Sigma_{r,k})$, where $\Sigma_{\text{B}} \succeq \mathbf{0}$ is the spatial correlation matrix with unit diagonal elements at the BS for channel $\tilde{\mathbf{H}}_{\text{dr}}$, and $\Sigma_{\text{R}} \succeq \mathbf{0}$ and $\Sigma_{r,k} \succeq \mathbf{0}$ are the spatial correlation matrices with unit diagonal elements at the RIS for channels $\tilde{\mathbf{H}}_{\text{dr}}$ and $\tilde{\mathbf{h}}_{r,k}$, respectively. The spatial covariance matrices can be estimated with the method proposed in [28] or the model proposed in [29]. Thus, in the following, we model $\tilde{\mathbf{H}}_{\text{dr}}$ and $\tilde{\mathbf{h}}_{r,k}$ as $\tilde{\mathbf{H}}_{\text{dr}} = \Sigma_{\text{R}}^{1/2} \mathbf{E} \Sigma_{\text{B}}^{1/2}$ with $\text{vec}(\mathbf{E}) \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_M \otimes \mathbf{I}_N)$ and $\tilde{\mathbf{h}}_{r,k} = \Sigma_{r,k}^{1/2} \mathbf{e}_{r,k}$ with $\mathbf{e}_{r,k} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_M)$, respectively.

Since the communication distances between the BS and the users are generally long and the electromagnetic environment is complex, we assume channels $\{\mathbf{h}_k\}_{k=1}^K$ to be Rayleigh distributed.

C. Average Achievable Rate and Average RIS Transmit Power

Since perfect instantaneous CSI of \mathbf{H}_{dr} and $\{\mathbf{h}_{r,k}\}_{k=1}^K$ is not available, in this work, we consider the average achievable rate and average RIS transmit power, denoted as $\mathbb{E}_{\mathbf{h}_{r,k}|\mathbf{G}_k} \{R_k(\mathbf{F}, \mathbf{w})\}$ for all k and $\mathbb{E}_{\mathbf{H}_{\text{dr}}|\mathbf{G}_k} \{||\Lambda_{\mathbf{w}}\mathbf{H}_{\text{dr}}\mathbf{F}||_F^2 + ||\mathbf{w}||_F^2 \sigma_z^2\}$, respectively. Here, the average rate and average RIS transmit power are short-term (instantaneous) measures that capture the expected performance over the distributions of $\{\mathbf{h}_{r,k}\}_{k=1}^K$ and \mathbf{H}_{dr} for given $\{\mathbf{G}_k\}_{k=1}^K$.

First, we derive an analytical expression for the average achievable rate. With the definitions $\tilde{\mathbf{w}} = [\mathbf{w}^H, 1]^H$ and $\mathbf{H}_k = [\mathbf{G}_k^H, \mathbf{h}_k]^H$ such that $\tilde{\mathbf{w}}^H \mathbf{H}_k = \mathbf{h}_k^H + \mathbf{h}_{r,k}^H \Lambda_{\mathbf{w}}^H \mathbf{H}_{\text{dr}}$, (3) can be reformulated as follows

$$R_k(\mathbf{F}, \mathbf{w}) = \log_2 \left(1 + \frac{|\tilde{\mathbf{w}}^H \mathbf{H}_k \mathbf{f}_k|^2}{\sum_{i=1, i \neq k}^K |\tilde{\mathbf{w}}^H \mathbf{H}_k \mathbf{f}_i|^2 + \sigma_z^2 ||\mathbf{h}_{r,k}^H \Lambda_{\mathbf{w}}^H||_2^2 + \sigma_k^2} \right). \quad (6)$$

Since the function $f(x) = \log_2(1 + \frac{1}{x})$ is convex in x , by using Jensen's inequality, we can obtain a lower bound for $\mathbb{E}_{\mathbf{h}_{r,k}|\mathbf{G}_k} \{R_k(\mathbf{F}, \mathbf{w})\}$ as follows

$$\mathbb{E}_{\mathbf{h}_{r,k}|\mathbf{G}_k} \{R_k(\mathbf{F}, \mathbf{w})\}$$

$$\begin{aligned}
&\geq \bar{R}_k(\mathbf{F}, \mathbf{w}) \\
&= \log_2 \left(1 + \frac{|\tilde{\mathbf{w}}^H \mathbf{H}_k \mathbf{f}_k|^2}{\sum_{i=1, i \neq k}^K |\tilde{\mathbf{w}}^H \mathbf{H}_k \mathbf{f}_i|^2 + \sigma_z^2 \mathbb{E}_{\mathbf{h}_{r,k} | \mathbf{G}_k} \{ \|\mathbf{h}_{r,k}^H \Lambda_{\mathbf{w}}^H\|_2^2 \} + \sigma_k^2} \right) \\
&= \log_2 \left(1 + \frac{|\tilde{\mathbf{w}}^H \mathbf{H}_k \mathbf{f}_k|^2}{\sum_{i=1, i \neq k}^K |\tilde{\mathbf{w}}^H \mathbf{H}_k \mathbf{f}_i|^2 + \sigma_z^2 \text{Tr} \{ \Lambda_{\mathbf{w}} \mathbb{E}_{\mathbf{h}_{r,k} | \mathbf{G}_k} \{ \mathbf{h}_{r,k} \mathbf{h}_{r,k}^H \} \Lambda_{\mathbf{w}}^H \} + \sigma_k^2} \right) \\
&\stackrel{(a)}{=} \log_2 \left(1 + \frac{|\tilde{\mathbf{w}}^H \mathbf{H}_k \mathbf{f}_k|^2}{\sum_{i=1, i \neq k}^K |\tilde{\mathbf{w}}^H \mathbf{H}_k \mathbf{f}_i|^2 + \sigma_z^2 \text{Tr} \left\{ \Lambda_{\mathbf{w}} \left(\frac{\beta_k \delta_k}{\delta_k + 1} \bar{\mathbf{h}}_{r,k} \bar{\mathbf{h}}_{r,k}^H + \frac{\beta_k}{\delta_k + 1} \Sigma_{r,k} \right) \Lambda_{\mathbf{w}}^H \right\} + \sigma_k^2} \right) \\
&\stackrel{(b)}{=} \log_2 \left(1 + \frac{|\tilde{\mathbf{w}}^H \mathbf{H}_k \mathbf{f}_k|^2}{\sum_{i=1, i \neq k}^K |\tilde{\mathbf{w}}^H \mathbf{H}_k \mathbf{f}_i|^2 + \sigma_z^2 \left(\frac{\beta_k \delta_k}{\delta_k + 1} \|\text{Diag}(\bar{\mathbf{h}}_{r,k}) \mathbf{w}\|_2^2 + \frac{\beta_k}{\delta_k + 1} \|\mathbf{w}\|_2^2 \right) + \sigma_k^2} \right) \\
&= \log_2 \left(1 + \frac{|\tilde{\mathbf{w}}^H \mathbf{H}_k \mathbf{f}_k|^2}{\sum_{i=1, i \neq k}^K |\tilde{\mathbf{w}}^H \mathbf{H}_k \mathbf{f}_i|^2 + \mathbf{w}^H \Psi_k \mathbf{w} + \sigma_k^2} \right), \tag{7}
\end{aligned}$$

where $\Psi_k = \frac{\beta_k \sigma_z^2}{\delta_k + 1} (\delta_k \text{Diag}(\bar{\mathbf{h}}_{r,k} \odot \bar{\mathbf{h}}_{r,k}^*) + \mathbf{I})$ and \mathbf{I} denotes the identity matrix. Equality (a) in (7) is obtained due to $\mathbf{h}_{r,k} \sim \mathcal{CN} \left(\sqrt{\beta_k \delta_k / (\delta_k + 1)} \bar{\mathbf{h}}_{r,k}, \frac{\beta_k}{\delta_k + 1} \Sigma_{r,k} \right)$ and $\mathbb{E}_{\mathbf{h}_{r,k} | \mathbf{G}_k} \{ \mathbf{h}_{r,k} \mathbf{h}_{r,k}^H \} = \frac{\beta_k \delta_k}{\delta_k + 1} \bar{\mathbf{h}}_{r,k} \bar{\mathbf{h}}_{r,k}^H + \frac{\beta_k}{\delta_k + 1} \Sigma_{r,k}$. Equality (b) in (7) is due to $\text{Tr} \{ \Lambda_{\mathbf{w}} \bar{\mathbf{h}}_{r,k} \bar{\mathbf{h}}_{r,k}^H \Lambda_{\mathbf{w}}^H \} = \|\text{Diag}(\bar{\mathbf{h}}_{r,k}) \mathbf{w}\|_2^2$ and $\frac{\beta_k}{\delta_k + 1} \text{Tr} \{ \Lambda_{\mathbf{w}} \Sigma_{r,k} \Lambda_{\mathbf{w}}^H \} = \frac{\beta_k}{\delta_k + 1} \text{Tr} \{ \Lambda_{\mathbf{w}} \Lambda_{\mathbf{w}}^H \} = \frac{\beta_k}{\delta_k + 1} \|\mathbf{w}\|_2^2$, as $\Sigma_{r,k}$ has unit diagonal elements.

Next, to derive an analytical expression for the average RIS transmit power, we provide a useful lemma, as follows.

Lemma 1 Let $\mathbf{H} \in \mathbb{C}^{M \times N} = \bar{\mathbf{H}} + \Sigma_r^{1/2} \mathbf{H}^w \Sigma_t^{1/2}$ represent a random matrix following distribution $\mathbf{H} \sim \mathcal{CN}(\bar{\mathbf{H}}, \Sigma_r \otimes \Sigma_t)$ with mean $\bar{\mathbf{H}}$ and covariance $\Sigma_r \otimes \Sigma_t$, where \mathbf{H}^w is a complex Gaussian random matrix with independent and identically distributed (i.i.d.) entries of zero mean and unit variance. Given matrix $\mathbf{X} \in \mathbb{C}^{N \times N}$, we have

$$\mathbb{E}_{\mathbf{H}} \{ \mathbf{H} \mathbf{X} \mathbf{H}^H \} = \bar{\mathbf{H}} \mathbf{X} \bar{\mathbf{H}}^H + \text{Tr} \{ \mathbf{X} \Sigma_t \} \Sigma_r.$$

Proof: Please refer to the proof of Lemma 2 in [30]. ■

By using $\mathbf{H}_{\text{dr}} \sim \mathcal{CN} \left(\sqrt{\frac{\beta_0 \delta_0}{\delta_0 + 1}} \bar{\mathbf{H}}_{\text{dr}}, \frac{\beta_0}{\delta_0 + 1} (\Sigma_R \otimes \Sigma_B) \right)$ and Lemma 1, the average RIS transmit power $P(\mathbf{F}, \mathbf{w})$ can be obtained as

$$\begin{aligned}
P(\mathbf{F}, \mathbf{w}) &= \mathbb{E}_{\mathbf{H}_{\text{dr}} | \mathbf{G}_k} \{ \|\Lambda_{\mathbf{w}} \mathbf{H}_{\text{dr}} \mathbf{F}\|_F^2 + \|\mathbf{w}\|_2^2 \sigma_z^2 \} \\
&= \text{Tr} \{ \Lambda_{\mathbf{w}} \mathbb{E}_{\mathbf{H}_{\text{dr}}} \{ \mathbf{H}_{\text{dr}} \mathbf{F} \mathbf{F}^H \mathbf{H}_{\text{dr}}^H \} \Lambda_{\mathbf{w}}^H \} + \|\mathbf{w}\|_2^2 \sigma_z^2 \\
&= \text{Tr} \{ \Lambda_{\text{RIS}}^a \mathbf{Q} \} + \|\mathbf{w}\|_2^2 \sigma_z^2, \tag{8}
\end{aligned}$$

where $\mathbf{Q} = \left(\frac{\beta_0 \delta_0}{\delta_0 + 1} \bar{\mathbf{H}}_{\text{dr}} \mathbf{F} \mathbf{F}^H \bar{\mathbf{H}}_{\text{dr}}^H + \frac{\beta_0}{\delta_0 + 1} \text{Tr} \{ \mathbf{F} \mathbf{F}^H \Sigma_B \} \Sigma_R \right)$ and $\Lambda_{\text{RIS}}^a = \Lambda_{\mathbf{w}}^H \Lambda_{\mathbf{w}}$.

III. AVERAGE ACHIEVABLE RATE MAXIMIZATION

In this section, we maximize the average achievable rate under a constraint on the average RIS transmit power. To this end, a concave lower bound of the non-concave objective function is constructed, and a KKT-based AO algorithm is developed.

A. Problem Formulation

The proposed problem can be formulated as follows

$$\max_{\mathbf{F}, \mathbf{w}} \sum_{k=1}^K \mathbb{E}_{\mathbf{h}_{r,k} | \mathbf{G}_k} \{R_k(\mathbf{F}, \mathbf{w})\} \quad (9a)$$

$$\text{s.t. } \|\mathbf{F}\|_F^2 \leq P_N, \quad (9b)$$

$$\mathbb{E}_{\mathbf{H}_{dr} | \mathbf{G}_k} \{ \|\mathbf{\Lambda}_w \mathbf{H}_{dr} \mathbf{F}\|_F^2 + \|\mathbf{w}\|_F^2 \sigma_z^2 \} \leq P_M, \quad (9c)$$

$$1 \leq |w_m|^2 \leq a_{max}, \forall m, \quad (9d)$$

where P_M is the maximum average RIS transmit power.

Since $\mathbb{E}_{\mathbf{h}_{r,k} | \mathbf{G}_k} \{R_k(\mathbf{F}, \mathbf{w})\}$ and $\mathbb{E}_{\mathbf{H}_{dr} | \mathbf{G}_k} \{ \|\mathbf{\Lambda}_w \mathbf{H}_{dr} \mathbf{F}\|_F^2 + \|\mathbf{w}\|_F^2 \sigma_z^2 \}$ are not analytically tractable, Problem (9) cannot be solved directly. Thus, based on (8) and (7), Problem (9) is lower bounded as follows

$$\max_{\mathbf{F}, \mathbf{w}} \sum_{k=1}^K \bar{R}_k(\mathbf{F}, \mathbf{w}) \quad (10a)$$

$$\text{s.t. } \|\mathbf{F}\|_F^2 \leq P_N, \quad (10b)$$

$$P(\mathbf{F}, \mathbf{w}) \leq P_M, \quad (10c)$$

$$1 \leq |w_m|^2 \leq a_{max}, \forall m. \quad (10d)$$

Problem (10) is still difficult to solve due to the non-concave objective function in (10a), the non-convex amplification gain constraints in (10d), and the coupling of variables \mathbf{F} and \mathbf{w} .

B. Problem Reformulation

In the following, we propose an AO algorithm to solve Problem (10) based on the MM algorithm (see, e.g., [31], [32] for tutorial introductions to MM algorithms). Specifically, the key idea of MM algorithms is to construct an easy-to-solve surrogate problem by deriving a minorizer of the original non-convex objective function, which is then used for optimization. Specifically, assuming that $f(\mathbf{x})$

is the original objective function which needs to be maximized over a convex set \mathcal{S}_x , its minorizers (denoted by $\tilde{f}(\mathbf{x}|\mathbf{x}^n)$) at a given point \mathbf{x}^n should satisfy the following conditions [32]:

- (A1) : $\tilde{f}(\mathbf{x}^n|\mathbf{x}^n) = f(\mathbf{x}^n), \forall \mathbf{x}^n \in \mathcal{S}_x$;
- (A2) : $\tilde{f}(\mathbf{x}|\mathbf{x}^n) \leq f(\mathbf{x}), \forall \mathbf{x}, \mathbf{x}^n \in \mathcal{S}_x$;
- (A3) : $\tilde{f}'(\mathbf{x}|\mathbf{x}^n; \mathbf{d})|_{\mathbf{x}=\mathbf{x}^n} = f'(\mathbf{x}^n; \mathbf{d}), \forall \mathbf{d} \text{ with } \mathbf{x}^n + \mathbf{d} \in \mathcal{S}_x$;
- (A4) : $\tilde{f}(\mathbf{x}|\mathbf{x}^n)$ is continuous in \mathbf{x} and \mathbf{x}^n ,

where $f'(\mathbf{x}^n; \mathbf{d})$, defined as the direction derivative of $f(\mathbf{x}^n)$ in direction \mathbf{d} , is given by

$$f'(\mathbf{x}^n; \mathbf{d}) = \lim_{\kappa \rightarrow 0} \frac{f(\mathbf{x}^n + \kappa \mathbf{d}) - f(\mathbf{x}^n)}{\kappa}.$$

Based on the MM framework, we derive a quadratic lower bound of $\bar{R}_k(\mathbf{F}, \mathbf{w})$ shown in the following lemma, the proof of which is similar to the proof in [22, Appendix A].

Lemma 2 *For a fixed point $\{\mathbf{F}^n, \mathbf{w}^n\}$, $\bar{R}_k(\mathbf{F}, \mathbf{w})$ is minorized by the concave surrogate function $\tilde{R}_k(\mathbf{F}, \mathbf{w}|\mathbf{F}^n, \mathbf{w}^n)$ given by*

$$\tilde{R}_k(\mathbf{F}, \mathbf{w}|\mathbf{F}^n, \mathbf{w}^n) = \text{const}_k + 2\text{Re} \{a_k \tilde{\mathbf{w}}^H \mathbf{H}_k \mathbf{f}_k\} - b_k(\|\tilde{\mathbf{w}}^H \mathbf{H}_k \mathbf{F}\|_2^2 + \mathbf{w}^H \Psi_k \mathbf{w}), \quad (11)$$

where

$$a_k = \frac{t_k^{n,*}}{r_k^n - |t_k^n|^2}, \quad b_k = \frac{|t_k^n|^2}{r_k^n(r_k^n - |t_k^n|^2)}, \quad \text{const}_k = R_k(\mathbf{F}^n, \mathbf{w}^n) - b_k(\sigma_k^2 + r_k^n),$$

$$t_k^n = (\tilde{\mathbf{w}}^n)^H \mathbf{H}_k \mathbf{f}_k^n, \quad r_k^n = \sum_{i=1}^K |(\tilde{\mathbf{w}}^n)^H \mathbf{H}_k \mathbf{f}_i^n|^2 + (\tilde{\mathbf{w}}^n)^H \Psi_k \mathbf{w}^n + \sigma_k^2.$$

Function (11) is biconcave in \mathbf{F} and \mathbf{w} , which motivates us to update \mathbf{F} and \mathbf{w} in an iterative manner. In particular, in the proposed AO algorithm, we first update \mathbf{F} based on the concave function $\tilde{R}_k(\mathbf{F}|\mathbf{F}^n) = \tilde{R}_k(\mathbf{F}, \mathbf{w}|\mathbf{F}^n, \mathbf{w}^n)$ for a given \mathbf{w} , and then we update \mathbf{w} based on the concave function $\tilde{R}_k(\mathbf{w}|\mathbf{w}^n) = \tilde{R}_k(\mathbf{F}, \mathbf{w}|\mathbf{F}^n, \mathbf{w}^n)$ for a given \mathbf{F} .

C. Optimization of Precoding Matrix \mathbf{F}

By using Lemma 2, a lower bound of the objective function in (10a) with respect to \mathbf{F} , denoted by $\tilde{R}_{\text{sum}}(\mathbf{F})$, is obtained as

$$\begin{aligned} \tilde{R}_{\text{sum}}(\mathbf{F}) &= \sum_{k=1}^K \tilde{R}_k(\mathbf{F}, \mathbf{w}|\mathbf{F}^n, \mathbf{w}^n) \\ &= \text{const}_F + 2\text{Re} \{ \text{Tr} \{ \mathbf{C}_F^H \mathbf{F} \} \} - \text{Tr} \{ \mathbf{F}^H \mathbf{A}_F \mathbf{F} \}, \end{aligned} \quad (12)$$

where $\text{const}_F = \sum_{k=1}^K \text{const}_k - \mathbf{w}^H (\sum_{k=1}^K b_k \mathbf{\Psi}_k) \mathbf{w}$, $\mathbf{C}_F = \sum_{k=1}^K a_k^* \mathbf{H}_k^H \tilde{\mathbf{w}} \mathbf{t}_k^H$, $\mathbf{A}_F = \sum_{k=1}^K b_k \mathbf{H}_k^H \tilde{\mathbf{w}} \tilde{\mathbf{w}}^H \mathbf{H}_k$, and $\mathbf{t}_k \in \mathbb{R}^{K \times 1}$ is a selection vector in which the k -th element is equal to one and all the other elements are equal to zero.

After some manipulations, the average RIS transmit power in (8) can be rewritten as a quadratic function of \mathbf{F} as follows

$$\begin{aligned} P(\mathbf{F}) &= \frac{\beta_0 \delta_0}{\delta_0 + 1} \text{Tr} \{ \mathbf{F}^H \bar{\mathbf{H}}_{\text{dr}}^H \mathbf{\Lambda}_{\text{RIS}}^a \bar{\mathbf{H}}_{\text{dr}} \mathbf{F} \} + \frac{\beta_0}{\delta_0 + 1} \text{Tr} \{ \mathbf{\Lambda}_{\text{RIS}}^a \mathbf{\Sigma}_{\text{R}} \} \text{Tr} \{ \mathbf{F}^H \mathbf{\Sigma}_{\text{B}} \mathbf{F} \} + \|\mathbf{w}\|_2^2 \sigma_z^2 \\ &= \frac{\beta_0 \delta_0}{\delta_0 + 1} \text{Tr} \{ \mathbf{F}^H \bar{\mathbf{H}}_{\text{dr}}^H \mathbf{\Lambda}_{\text{RIS}}^a \bar{\mathbf{H}}_{\text{dr}} \mathbf{F} \} + \frac{\beta_0}{\delta_0 + 1} \text{Tr} \{ \mathbf{\Lambda}_{\text{RIS}}^a \} \text{Tr} \{ \mathbf{F}^H \mathbf{\Sigma}_{\text{B}} \mathbf{F} \} + \|\mathbf{w}\|_2^2 \sigma_z^2 \\ &= \text{Tr} \{ \mathbf{F}^H \mathbf{D}_F \mathbf{F} \} + \|\mathbf{w}\|_2^2 \sigma_z^2, \end{aligned} \quad (13)$$

where we have $\text{Tr} \{ \mathbf{\Lambda}_{\text{RIS}}^a \mathbf{\Sigma}_{\text{R}} \} = \|\mathbf{w}\|_2^2$ and $\mathbf{D}_F = \frac{\beta_0 \delta_0}{\delta_0 + 1} \bar{\mathbf{H}}_{\text{dr}}^H \mathbf{\Lambda}_{\text{RIS}}^a \bar{\mathbf{H}}_{\text{dr}} + \frac{\beta_0}{\delta_0 + 1} \|\mathbf{w}\|_2^2 \mathbf{\Sigma}_{\text{B}}$.

Combining (12) with (13), and ignoring irrelevant constant terms, the surrogate subproblem of (10) with respect to \mathbf{F} for a given \mathbf{w} is formulated as follows

$$\max_{\mathbf{F}} 2\text{Re} \{ \text{Tr} \{ \mathbf{C}_F^H \mathbf{F} \} \} - \text{Tr} \{ \mathbf{F}^H \mathbf{A}_F \mathbf{F} \} \quad (14a)$$

$$\text{s.t. } \|\mathbf{F}\|_F^2 \leq P_N, \quad (14b)$$

$$\text{Tr} \{ \mathbf{F}^H \mathbf{D}_F \mathbf{F} \} + \|\mathbf{w}\|_2^2 \sigma_z^2 \leq P_M. \quad (14c)$$

Problem (14) is a standard second-order cone programming (SOCP) problem and can be solved with CVX. However, the computational complexity of SOCP-based algorithms is high. In the following, we solve Problem (14) by exploiting the standard dual decomposition method. In particular, the Lagrange function of Problem (14) is given by

$$\begin{aligned} \mathcal{L}(\mathbf{F}, \gamma_F, \mu_F) \\ = \text{Tr} \{ \mathbf{F}^H \mathbf{A}_F \mathbf{F} \} - 2\text{Re} \{ \text{Tr} \{ \mathbf{C}_F^H \mathbf{F} \} \} + \gamma_F (\|\mathbf{F}\|_F^2 - P_N) + \mu_F (\text{Tr} \{ \mathbf{F}^H \mathbf{D}_F \mathbf{F} \} + \|\mathbf{w}\|_2^2 \sigma_z^2 - P_M), \end{aligned}$$

where Lagrange multipliers $\gamma_F \geq 0$ and $\mu_F \geq 0$ are associated with constraints (14b) and (14c), respectively. The dual function $d(\gamma_F, \mu_F)$ is given by

$$d(\gamma_F, \mu_F) = \max_{\mathbf{F}} \mathcal{L}(\mathbf{F}, \gamma_F, \mu_F). \quad (15)$$

Thus, the dual problem of Problem (14) can be formulated as follows

$$\min_{\gamma_F \geq 0, \mu_F \geq 0} d(\gamma_F, \mu_F). \quad (16)$$

Firstly, by exploiting the first-order KKT necessary and sufficient condition of the problem in (15), i.e., $\frac{\partial \mathcal{L}(\mathbf{F}, \gamma_F, \mu_F)}{\partial \mathbf{F}^*} = (\mathbf{A}_F + \gamma_F^{\text{opt}} \mathbf{I} + \mu_F^{\text{opt}} \mathbf{D}_F) \mathbf{F}_{\text{opt}} - \mathbf{C}_F = \mathbf{0}$, we obtain the optimal solution of \mathbf{F} for fixed dual variables $(\gamma_F^{[v]}, \mu_F^{[v]})$ in iteration v as follows

$$\mathbf{F}(\gamma_F^{[v]}, \mu_F^{[v]}) = \left(\mathbf{A}_F + \gamma_F^{[v]} \mathbf{I} + \mu_F^{[v]} \mathbf{D}_F \right)^{-1} \mathbf{C}_F. \quad (17)$$

Then, the dual problem (16) can be solved by the gradient projection algorithm, i.e., the dual variables are updated as follows

$$\gamma_F^{[v+1]} = \left[\gamma_F^{[v]} + \varsigma^{[v]} \nabla_{\gamma_F} d(\gamma_F^{[v]}, \mu_F^{[v]}) \right]^+, \quad (18a)$$

$$\mu_F^{[v+1]} = \left[\mu_F^{[v]} + \varsigma^{[v]} \nabla_{\mu_F} d(\gamma_F^{[v]}, \mu_F^{[v]}) \right]^+, \quad (18b)$$

where

$$\nabla_{\gamma_F} d(\gamma_F^{[v]}, \mu_F^{[v]}) = \|\mathbf{F}(\gamma_F^{[v]}, \mu_F^{[v]})\|_F^2 - P_N \quad (19a)$$

$$\nabla_{\mu_F} d(\gamma_F^{[v]}, \mu_F^{[v]}) = \text{Tr} \left\{ (\mathbf{F}(\gamma_F^{[v]}, \mu_F^{[v]}))^H \mathbf{D}_F \mathbf{F}(\gamma_F^{[v]}, \mu_F^{[v]}) \right\} + \|\mathbf{w}\|_2^2 \sigma_z^2 - P_M. \quad (19b)$$

The initialization points can be set as $\gamma_F^{[0]} = 0$ and $\mu_F^{[0]} = 0$. Then, (17) and (19) are updated in an alternating manner until $\|(\gamma_F^{[v+1]}, \mu_F^{[v+1]}) - (\gamma_F^{[v]}, \mu_F^{[v]})\| \rightarrow 0$. Note that although the algorithm proposed for solving Problem (14) requires iterations, the computational complexity is comparatively low due to closed-form expressions employed in each iteration, see Section III-F.

D. Optimization of Reflection Vector \mathbf{w}

In order to facilitate the subsequent derivations, we convert the surrogate objective function $\sum_{k=1}^K \tilde{R}_k(\mathbf{F}, \mathbf{w} | \mathbf{F}^n, \mathbf{w}^n)$ and average RIS transmit power in (8) into quadratic functions of \mathbf{w} as follows

$$\begin{aligned} \tilde{R}_{\text{sum}}(\mathbf{w}) &= \sum_{k=1}^K \tilde{R}_k(\mathbf{F}, \mathbf{w} | \mathbf{F}^n, \mathbf{w}^n) \\ &= \text{const}_w + 2\text{Re} \{ \mathbf{w}^H \mathbf{c}_w \} - \mathbf{w}^H \mathbf{A}_w \mathbf{w}, \end{aligned} \quad (20)$$

with $\text{const}_w = \sum_{k=1}^K \text{const}_k + 2\text{Re} \{ \sum_{k=1}^K a_k \mathbf{h}_k^H \mathbf{f}_k \} - \sum_{k=1}^K b_k \mathbf{h}_k^H \mathbf{F} \mathbf{F}^H \mathbf{h}_k$, $\mathbf{c}_w = \sum_{k=1}^K a_k \mathbf{G}_k \mathbf{f}_k - \sum_{k=1}^K b_k \mathbf{G}_k \mathbf{F} \mathbf{F}^H \mathbf{h}_k$, and $\mathbf{A}_w = \sum_{k=1}^K b_k \mathbf{G}_k \mathbf{F} \mathbf{F}^H \mathbf{G}_k^H + (\sum_{k=1}^K b_k \Psi_k)$, and

$$P(\mathbf{w}) = \mathbf{w}^H \mathbf{D}_w \mathbf{w}, \quad (21)$$

with $\mathbf{D}_w = \frac{\beta_0 \delta_0}{\delta_0 + 1} \text{Diag}(\text{diag}(\bar{\mathbf{H}} \mathbf{F} \mathbf{F}^H \bar{\mathbf{H}}^H)) + \frac{\beta_0}{\delta_0 + 1} \text{Tr} \{ \mathbf{F}^H \Sigma_B \mathbf{F} \} \mathbf{I} + \sigma_z^2 \mathbf{I}$.

Exploiting (20) and (21), we formulate a surrogate subproblem for (10) with respect to \mathbf{w} for a given \mathbf{F} as follows

$$\max_{\mathbf{w}} 2\text{Re} \{ \mathbf{w}^H \mathbf{c}_w \} - \mathbf{w}^H \mathbf{A}_w \mathbf{w} \quad (22a)$$

$$\text{s.t. } \mathbf{w}^H \mathbf{D}_w \mathbf{w} \leq P_M, \quad (22b)$$

$$1 \leq |w_m|^2 \leq a_{max}, \forall m. \quad (22c)$$

Problem (22) can be transformed into an SOCP by relaxing the non-convex constraint $1 \leq |w_m|^2$ in (22c) via a linear approximate constraint $1 \leq 2\text{Re}\{(w_m^n)^* w_m\} - |w_m^n|^2$ by using the first-order Taylor approximation at fixed point w_m^n . The resulting approximate SOCP problem is given by

$$\max_{\mathbf{w}} 2\text{Re}\{\mathbf{w}^H \mathbf{c}_w\} - \mathbf{w}^H \mathbf{A}_w \mathbf{w} \quad (23a)$$

$$\text{s.t. } \mathbf{w}^H \mathbf{D}_w \mathbf{w} \leq P_M, \quad (23b)$$

$$1 \leq 2\text{Re}\{(w_m^n)^* w_m\} - |w_m^n|^2, \forall m, \quad (23c)$$

$$|w_m|^2 \leq a_{max}, \forall m. \quad (23d)$$

To find a low-complexity solution for (23), we adopt the alternating direction method of multipliers (ADMM) [33]. In particular, we introduce auxiliary variable $\mathbf{u} = [u_1, \dots, u_M]^T$ such that $\mathbf{u} = \mathbf{w}$ and $1 \leq |u_m|^2 \leq a_{max}, \forall m$. The augmented Lagrangian of the optimization problem is given by

$$\mathcal{L}_\xi(\mathbf{w}, \mathbf{u}, \boldsymbol{\eta}) = \mathbf{w}^H \mathbf{A}_w \mathbf{w} - 2\text{Re}\{\mathbf{w}^H \mathbf{c}_w\} + \zeta \|\mathbf{w} - \mathbf{u} + \boldsymbol{\eta}\|_2^2,$$

where $\zeta > 0$ is a penalty parameter. The benefit of including the penalty term is to make the dual function differentiable. The ADMM method comprises the following steps ²:

$$\mathbf{w}^{[i+1]} = \arg \min_{\mathbf{w} \in \{\mathbf{w}^H \mathbf{D}_w \mathbf{w} \leq P_M\}} \mathcal{L}_\xi(\mathbf{w}, \mathbf{u}^{[i]}, \boldsymbol{\eta}^{[i]}), \quad (24)$$

$$\mathbf{u}^{[i+1]} = \arg \min_{\mathbf{u} \in \{1 \leq |u_m|^2 \leq a_{max}, \forall m\}} \mathcal{L}_\xi(\mathbf{w}^{[i+1]}, \mathbf{u}, \boldsymbol{\eta}^{[i]}), \quad (25)$$

$$\boldsymbol{\eta}^{[i+1]} = \boldsymbol{\eta}^{[i]} + \mathbf{w}^{[i+1]} - \mathbf{u}^{[i+1]}. \quad (26)$$

• Updating \mathbf{w} : $\mathbf{w}^{[i+1]}$ can be obtained using the KKT conditions. We form the Lagrangian function of Problem (24) with Lagrange multiplier γ_w as $\mathcal{L}(\mathbf{w}, \gamma_w) = \mathbf{w}^H \mathbf{A}_w \mathbf{w} - 2\text{Re}\{\mathbf{w}^H \mathbf{c}_w\} + \zeta \|\mathbf{w} - \mathbf{u}^{[i]} + \boldsymbol{\eta}^{[i]}\|_2^2 + \gamma_w (\mathbf{w}^H \mathbf{D}_w \mathbf{w} - P_M)$, and obtain the first-order KKT necessary condition for the optimal $\mathbf{w}^{[i+1]}$ as $\frac{\partial \mathcal{L}(\mathbf{w}^{[i+1]}, \gamma_w^{\text{opt}})}{\partial \mathbf{w}^*} = (\mathbf{A}_w + \gamma_w^{\text{opt}} \mathbf{D}_w + \zeta \mathbf{I}) \mathbf{w}^{[i+1]} - \mathbf{c}_w + \zeta(\boldsymbol{\eta}^{[i]} - \mathbf{u}^{[i]}) = \mathbf{0}$. Then, $\mathbf{w}^{[i+1]}$ is given by

$$\mathbf{w}^{[i+1]}(\gamma_w^{\text{opt}}) = (\mathbf{A}_w + \gamma_w^{\text{opt}} \mathbf{D}_w + \zeta \mathbf{I})^{-1} (\mathbf{c}_w - \zeta(\boldsymbol{\eta}^{[i]} - \mathbf{u}^{[i]})). \quad (27)$$

Function $g_w(\gamma_w) = (\mathbf{w}^{[i+1]}(\gamma_w))^H \mathbf{D}_w \mathbf{w}^{[i+1]}(\gamma_w)$ is a monotonically decreasing function of γ_w . If $g_w(0) \leq P_M$, then $\mathbf{w}^{[i+1]} = (\mathbf{A}_w + \zeta \mathbf{I})^{-1} (\mathbf{c}_w - \zeta(\boldsymbol{\eta}^{[i]} - \mathbf{u}^{[i]}))$. Otherwise, $g_w(0) > P_M$. Based

²Please note that w_m^n in Problem (23) is the updated value in each iteration of the MM algorithm, while $\mathbf{w}^{[i+1]}$ in Problem (24) is the updated value in each iteration of the ADMM method.

on the complementary condition of $\gamma_w(g_w(\gamma_w) - P_M) = 0$, we need to find a positive γ_w such that $g_w(\gamma_w) - P_M = 0$. Defining $\gamma_{w,1} = \sqrt{\frac{(\mathbf{c}_w - \zeta(\boldsymbol{\eta}^{[i]} - \mathbf{u}^{[i]}))^H \mathbf{D}_w^{-1} (\mathbf{c}_w - \zeta(\boldsymbol{\eta}^{[i]} - \mathbf{u}^{[i]})}{P_M}}$, we have $g_w(\gamma_{w,1}) < (\gamma_{w,1})^{-2} (\mathbf{c}_w - \zeta(\boldsymbol{\eta}^{[i]} - \mathbf{u}^{[i]}))^H \mathbf{D}_w^{-1} (\mathbf{c}_w - \zeta(\boldsymbol{\eta}^{[i]} - \mathbf{u}^{[i]})) = P_M$. Then, a unique $\gamma_w^{\text{opt}} \in (0, \gamma_{w,1})$ must exist such that $g_w(\gamma_w^{\text{opt}}) = P_M$, and thus γ_w^{opt} can be found by using a one-dimensional search.

- Updating \mathbf{u} : The optimization problem in (25) is equivalent to

$$\min_{\mathbf{u}} \|\mathbf{w} - \mathbf{u} + \boldsymbol{\eta}\|_2^2 \quad \text{s.t. } 1 \leq |u_m|^2 \leq a_{max}, \forall m. \quad (28)$$

Its solution is given by $\mathbf{u}^{[i+1]} = [\|\mathbf{w}^{[i+1]} + \boldsymbol{\eta}^{[i]}\|_1^{\sqrt{a_{max}}} \exp(j\angle(\mathbf{w}^{[i+1]} + \boldsymbol{\eta}^{[i]}))]$, where operator $|\cdot|$ returns the elementwise absolute value and operator $[\mathbf{x}]_{\underline{u}}^{\overline{u}}$ maps \mathbf{x} elementwise onto the interval $[\underline{u}, \overline{u}]$.

E. Algorithm Development

Under the MM framework, the solution \mathbf{F} of Problem (14) and the solution \mathbf{w} of Problem (23) in each AO iteration can be obtained with low complexity using the proposed KKT-based and ADMM methods, respectively. The convergence speed of the MM algorithm will be affected by the tightness of the lower bound of the original objective function given in Lemma 2. Thus, an acceleration method, called SQUAREM [34], is adopted to accelerate the MM-based algorithm, as is summarized in Algorithm 1. $F_F(\mathbf{F}^n)$ in Step 9 and $F_w(\mathbf{w}^n)$ in Step 20 represent the objective function values of Problem (14) and Problem (23) in the n -th iteration, respectively.

$\mathcal{P}_F(\cdot)$ in Step 8 and $\mathcal{P}_w(\cdot)$ in Step 19 are projection operations onto the nonlinear constraint sets of \mathbf{F} and \mathbf{w} , respectively, which ensure the feasibility of the updated solutions. The projection operation is defined as $\mathcal{P}(\mathbf{x}) = \arg \min_{\mathbf{z} \in \mathcal{S}} \|\mathbf{z} - \mathbf{x}\|_2^2$, where \mathcal{S} is the constraint set of \mathbf{z} [35, Equ. (4.4.13)]. Therefore, for the power constraint set of \mathbf{F} , $\mathcal{P}_F(\cdot)$ is obtained as follows

$$\mathcal{P}_F(\mathbf{X}) = \arg \min_{\mathbf{F}} \|\mathbf{F} - \mathbf{X}\|_F \quad \text{s.t. (14b), (14c)}. \quad (29)$$

and for the power and amplification gain constraints of \mathbf{w} , $\mathcal{P}_w(\cdot)$ is obtained as follows

$$\mathcal{P}_w(\mathbf{x}) = \arg \min_{\mathbf{w}} \|\mathbf{w} - \mathbf{x}\|_2 \quad \text{s.t. (22b), (22c)}. \quad (30)$$

Steps 9 to 12 and Steps 20 to 23 are used to maintain the monotonicity of the objective function values.

F. Complexity Analysis

Algorithm 1 requires solving Problem (14) and Problem (23). In the following complexity analysis, we neglect terms with low-order complexity. To solve Problem (14), we first need to calculate \mathbf{A}_F and

Algorithm 1 Low-complexity MM algorithm

Initialize: Initialize \mathbf{F}^0 and \mathbf{w}^0 . Set $n = 1$

- 1: **repeat**
 - 2: Set $\mathbf{w} = \mathbf{w}^{n-1}$
 - 3: Obtain \mathbf{F}_1 from Problem (14) based on \mathbf{F}^{n-1}
 - 4: Obtain \mathbf{F}_2 from Problem (14) based on \mathbf{F}_1
 - 5: $\mathbf{R}_1 = \mathbf{F}_1 - \mathbf{F}^{n-1}$
 - 6: $\mathbf{R}_2 = \mathbf{F}_2 - \mathbf{F}_1 - \mathbf{R}_1$
 - 7: $\omega_F = -\frac{\|\mathbf{R}_1\|_F}{\|\mathbf{R}_2\|_F}$
 - 8: $\mathbf{F}^n = -\mathcal{P}_F(\mathbf{F}^{n-1} - 2\omega_F \mathbf{R}_1 + \omega_F^2 \mathbf{R}_2)$
 - 9: **while** $F_F(\mathbf{F}^n) < F_F(\mathbf{F}^{n-1})$ **do**
 - 10: $\omega_F = (\omega_F - 1)/2$
 - 11: $\mathbf{F}^n = -\mathcal{P}_F(\mathbf{F}^{n-1} - 2\omega_F \mathbf{R}_1 + \omega_F^2 \mathbf{R}_2)$
 - 12: **end while**
 - 13: Set $\mathbf{F} = \mathbf{F}^n$
 - 14: Obtain \mathbf{w}_1 from Problem (23) based on \mathbf{w}^{n-1}
 - 15: Obtain \mathbf{w}_2 from Problem (23) based on \mathbf{w}^1
 - 16: $\mathbf{r}_1 = \mathbf{w}_1 - \mathbf{w}^{n-1}$
 - 17: $\mathbf{r}_2 = \mathbf{w}_2 - \mathbf{w}_1 - \mathbf{r}_1$
 - 18: $\omega_w = -\frac{\|\mathbf{r}_1\|_2}{\|\mathbf{r}_2\|_2}$
 - 19: $\mathbf{w}^n = -\mathcal{P}_w(\mathbf{w}^{n-1} - 2\omega_w \mathbf{r}_1 + \omega_w^2 \mathbf{r}_2)$
 - 20: **while** $F_w(\mathbf{w}^n) < F_w(\mathbf{w}^{n-1})$ **do**
 - 21: $\omega_w = (\omega_w - 1)/2$
 - 22: $\mathbf{w}^n = -\mathcal{P}_w(\mathbf{w}^{n-1} - 2\omega_w \mathbf{r}_1 + \omega_w^2 \mathbf{r}_2)$
 - 23: **end while**
 - 24: $n = n + 1$
 - 25: **until** $|F_F(\mathbf{F}^n) - F_F(\mathbf{F}^{n-1})| \rightarrow 0$ and $|F_w(\mathbf{w}^n) - F_w(\mathbf{w}^{n-1})| \rightarrow 0$
-

\mathbf{D}_F , which have computational complexity orders of $\mathcal{O}(M^2NK + MNK)$ and $\mathcal{O}(M^2NK)$, respectively. The calculation of \mathbf{C}_F is similar to that of \mathbf{A}_F . Then, the inverse operation in (17) has complexity $\mathcal{O}(M^3)$. Therefore, the approximate computational complexity of solving Problem (14) is $\mathcal{O}(M^3 + M^2NK + MNK)$, where constant coefficients are ignored. The computational complexity of the ADMM algorithm used for solving Problem (23) is mainly determined by the calculation of \mathbf{A}_w , \mathbf{D}_w , and the inverse operation in (27), which have complexities of $\mathcal{O}(MNK^2 + M^2NK)$, $\mathcal{O}(MNK)$, and $\mathcal{O}(M^3)$, respectively. Also, computing \mathbf{c}_w involves similar steps as computing \mathbf{A}_w . Neglecting the constant coefficients, the approximate complexity of the ADMM algorithm is given by $\mathcal{O}(M^3 + MNK^2 + M^2NK + MNK)$. Thus, the approximate complexity of Algorithm 1 per iteration is $\mathcal{O}(M^3 + MNK^2 + M^2NK + MNK)$.

G. Convergence Analysis

Next, we analyze the convergence of the proposed algorithm. The monotonic convergence of the MM algorithm has been proved in [32] and [36]. In the following, we prove the monotonic convergence of Algorithm 1. Let $f(\mathbf{F}, \mathbf{w}) = \sum_{k=1}^K \bar{R}_k(\mathbf{F}, \mathbf{w})$ denote the objective value of Problem (10) and $\tilde{f}(\mathbf{F}, \mathbf{w}) = \sum_{k=1}^K \tilde{R}_k(\mathbf{F}, \mathbf{w})$ represent its minorizer. In the n^{th} iteration, given \mathbf{w}^n , we have

$$f(\mathbf{F}^n, \mathbf{w}^n) = \tilde{f}(\mathbf{F}^n, \mathbf{F}^n) \leq \tilde{f}(\mathbf{F}^{n+1}, \mathbf{F}^n) \leq f(\mathbf{F}^{n+1}, \mathbf{w}^n),$$

where the first equality follows from condition (A1), the first inequality is due to the optimal solution of Problem (14), and the second inequality follows from condition (A2). Subsequently, given \mathbf{F}^{n+1} , it is straightforward to show that

$$f(\mathbf{F}^{n+1}, \mathbf{w}^n) = \tilde{f}(\mathbf{w}^n, \mathbf{w}^n) \leq \tilde{f}(\mathbf{w}^{n+1}, \mathbf{w}^n) \leq f(\mathbf{F}^{n+1}, \mathbf{w}^{n+1}).$$

Therefore, the sequence of objective values $\{f(\mathbf{F}^{n+1}, \mathbf{w}^{n+1})\}$ generated by the AO algorithm is monotonically non-decreasing. Since \mathbf{F} belongs to a convex set, every limit point of \mathbf{F}^n is a d-stationary point of Problem (10) [22]. Furthermore, since \mathbf{w} belongs to a non-convex set, every limit point of \mathbf{w}^n is a B-stationary point of Problem (10) [22].

IV. OUTAGE CONSTRAINED AVERAGE POWER MINIMIZATION

In the previous section, we have investigated the average rate maximization problem for the practical case where only partial CSI of the RIS-aided channels is available. However, this problem formulation cannot guarantee the QoS of the individual users and outages may occur in an uncontrolled manner. Thus, in order to ensure the QoS of the individual users, in this section, we jointly optimize the

beamforming matrices at both the BS and the RIS to guarantee that the probability that the instantaneous achievable rate of each user exceeds a target rate is larger than a predefined value, while minimizing the total transmit power consumption comprising the BS transmit power and the average RIS transmit power. To obtain a tractable problem formulation, the outage probability constraint is approximated by the BTI, and then an SDR-based AO algorithm is proposed to optimize the beamforming matrices.

A. Problem Formulation

The proposed optimization problem is formulated as follows

$$\min_{\mathbf{F}, \mathbf{w}} \|\mathbf{F}\|_F^2 + \mathbb{E}_{\mathbf{H}_{\text{dr}} | \mathbf{G}_k} \{ \|\mathbf{\Lambda}_{\mathbf{w}} \mathbf{H}_{\text{dr}} \mathbf{F}\|_F^2 \} + \|\mathbf{\Lambda}_{\mathbf{w}}\|_F^2 \sigma_z^2 \quad (31a)$$

$$\text{s.t. } \Pr\{R_k(\mathbf{F}, \mathbf{w}) \geq r_k\} \geq 1 - \rho_k, \forall k, \quad (31b)$$

$$1 \leq |w_m|^2 \leq a_{\max}, \forall m, \quad (31c)$$

where (31b) ensures that the probability that each user can successfully decode its message for a data rate of r_k is no less than $1 - \rho_k$, where $\rho_1, \dots, \rho_K \in (0, 1]$ are the corresponding maximum outage probabilities.

B. Problem Reformulation

Problem (31) is computationally intractable since outage probability constraint (31b) does not have an analytical expression. Therefore, we safely approximate (31b) by some easy-to-handle constraints by exploiting the following lemma.

Lemma 3 (*Bernstein-Type Inequality: Lemma 1 in [37]*) Assume $f(\mathbf{x}) = \mathbf{x}^H \mathbf{U} \mathbf{x} + 2\text{Re}\{\mathbf{u}^H \mathbf{x}\} + u$, where $\mathbf{U} \in \mathbb{H}^{n \times n}$, $\mathbf{u} \in \mathbb{C}^{n \times 1}$, $u \in \mathbb{R}$, and $\mathbf{x} \in \mathbb{C}^{n \times 1} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$. Then, for any $\rho \in [0, 1]$, the following approximation holds:

$$\Pr\{\mathbf{x}^H \mathbf{U} \mathbf{x} + 2\text{Re}\{\mathbf{u}^H \mathbf{x}\} + u \leq 0\} \geq 1 - \rho \quad (32a)$$

$$\Rightarrow \text{Tr}\{\mathbf{U}\} + \sqrt{2 \ln(1/\rho)} x - \ln(\rho) \lambda_{\max}^+(\mathbf{U}) + u \leq 0 \quad (32b)$$

$$\Rightarrow \begin{cases} \text{Tr}\{\mathbf{U}\} + \sqrt{2 \ln(1/\rho)} x - \ln(\rho) y + u \leq 0 \\ \sqrt{\|\mathbf{U}\|_F^2 + 2\|\mathbf{u}\|_2^2} \leq x \\ y\mathbf{I} - \mathbf{U} \succeq \mathbf{0}, y \geq 0, \end{cases} \quad (32c)$$

where $\lambda_{\max}^+(\mathbf{U}) = \max(\lambda_{\max}(\mathbf{U}), 0)$ and $\lambda_{\max}(\mathbf{U})$ denotes the maximum eigenvalue of \mathbf{U} . x and y are slack variables.

Please refer to [37] for a proof of Lemma 3.

To utilize Lemma 3, the outage probability in (31b) is firstly reformulated into the form of (32a):

$$\Pr \{R_k(\mathbf{F}, \mathbf{w}) \geq r_k\} = \Pr \{ \tilde{\mathbf{w}}^H \mathbf{H}_k \mathbf{\Phi}_k \mathbf{H}_k^H \tilde{\mathbf{w}} - \|\mathbf{h}_{r,k}^H \mathbf{\Lambda}_w\|_2^2 \sigma_z^2 - \sigma_k^2 \geq 0 \}, \quad (33)$$

where

$$\mathbf{\Phi}_k = \mathbf{f}_k \mathbf{f}_k^H / (2^{r_k} - 1) - \mathbf{F}_{-k} \mathbf{F}_{-k}^H, \quad (34a)$$

$$\mathbf{F}_{-k} = [\mathbf{f}_1, \dots, \mathbf{f}_k, \mathbf{f}_{k+1}, \dots, \mathbf{f}_K]. \quad (34b)$$

Substituting (5) and $\tilde{\mathbf{h}}_{r,k} = \Sigma_{r,k}^{1/2} \mathbf{e}_{r,k}$ into (33), we have

$$\Pr \{R_k(\mathbf{F}, \mathbf{w}) \geq r_k\} = \Pr \{ \mathbf{e}_{r,k}^H \mathbf{U}_k \mathbf{e}_{r,k} + 2\text{Re}\{\mathbf{u}_k^H \mathbf{e}_{r,k}\} + u_k \leq 0 \},$$

where

$$\mathbf{U}_k = \frac{\beta_k}{\delta_k + 1} \Sigma_{r,k}^{1/2} \mathbf{\Lambda}_{\text{RIS}}^a \Sigma_{r,k}^{1/2}, \quad (35a)$$

$$\mathbf{u}_k = \frac{\beta_k \sqrt{\delta_k}}{\delta_k + 1} \Sigma_{r,k}^{1/2} \mathbf{\Lambda}_{\text{RIS}}^a \bar{\mathbf{h}}_{r,k}, \quad (35b)$$

$$u_k = \frac{\beta_k \delta_k}{\delta_k + 1} \bar{\mathbf{h}}_{r,k}^H \mathbf{\Lambda}_{\text{RIS}}^a \bar{\mathbf{h}}_{r,k} - \frac{1}{\sigma_z^2} (\tilde{\mathbf{w}}^H \mathbf{H}_k \mathbf{\Phi}_k \mathbf{H}_k^H \tilde{\mathbf{w}} - \sigma_k^2). \quad (35c)$$

Furthermore, the following theorem is provided to facilitate the subsequent derivations.

Theorem 1 Given matrices $\mathbf{A} \in \mathbb{C}^{N \times N}$, $\mathbf{b} \in \mathbb{C}^N$, and $\mathbf{C} \in \mathbb{C}^{N \times N}$, we have

$$\text{Tr} \{ \mathbf{A} \text{Diag}(\mathbf{b}) \mathbf{C} \text{Diag}(\mathbf{b}) \} = \mathbf{b}^T (\mathbf{A}^T \odot \mathbf{C}) \mathbf{b}. \quad (36)$$

Proof: Please refer to Appendix A. ■

Then, we establish the following identities:

$$\text{Tr} \{ \mathbf{U}_k \} = \frac{\beta_k}{\delta_k + 1} \text{Tr} \{ \mathbf{\Lambda}_{\text{RIS}}^a \}, \quad (37a)$$

$$\begin{aligned} \|\mathbf{U}_k\|_F^2 &= \left(\frac{\beta_k}{\delta_k + 1} \right)^2 \text{Tr} \{ \Sigma_{r,k} \mathbf{\Lambda}_{\text{RIS}}^a \Sigma_{r,k} \mathbf{\Lambda}_{\text{RIS}}^a \} \\ &= \left(\frac{\beta_k}{\delta_k + 1} \right)^2 \mathbf{p}_{\text{RIS}}^T (\Sigma_{r,k}^T \odot \Sigma_{r,k}) \mathbf{p}_{\text{RIS}}, \end{aligned} \quad (37b)$$

$$\begin{aligned} \|\mathbf{u}_k\|_2^2 &= \left(\frac{\beta_k \sqrt{\delta_k}}{\delta_k + 1} \right)^2 \text{Tr} \{ \bar{\mathbf{h}}_{r,k} \bar{\mathbf{h}}_{r,k}^H \mathbf{\Lambda}_{\text{RIS}}^a \Sigma_{r,k} \mathbf{\Lambda}_{\text{RIS}}^a \} \\ &= \left(\frac{\beta_k \sqrt{\delta_k}}{\delta_k + 1} \right)^2 \mathbf{p}_{\text{RIS}}^T ((\bar{\mathbf{h}}_{r,k}^* \bar{\mathbf{h}}_{r,k}^T) \odot \Sigma_{r,k}) \mathbf{p}_{\text{RIS}}, \end{aligned} \quad (37c)$$

$$\lambda(\mathbf{U}_k) = \lambda \left(\frac{\beta_k}{\delta_k + 1} \Sigma_{r,k}^{1/2} \mathbf{\Lambda}_{\text{RIS}}^a \Sigma_{r,k}^{1/2} \right)$$

$$= \frac{\beta_k}{\delta_k + 1} \lambda(\Sigma_{r,k} \Lambda_{\text{RIS}}^a), \quad (37d)$$

where $\mathbf{p}_{\text{RIS}} = \text{diag}(\Lambda_{\text{RIS}}^a)$. Equations (37b) and (37c) are obtained based on Theorem 1.

Therefore, applying Lemma 3, constraint (31b) can be approximated as follows

$$\left\{ \begin{array}{l} \frac{\beta_k}{\delta_k + 1} \text{Tr} \{ \Lambda_{\text{RIS}}^a \} + \sqrt{2 \ln(1/\rho_k)} x_k - \ln(\rho_k) y_k + \frac{\beta_k \delta_k}{\delta_k + 1} \bar{\mathbf{h}}_{r,k}^H \Lambda_{\text{RIS}}^a \bar{\mathbf{h}}_{r,k} \\ - \frac{1}{\sigma_z^2} (\tilde{\mathbf{w}}^H \mathbf{H}_k \Phi_k \mathbf{H}_k^H \tilde{\mathbf{w}} - \sigma_k^2) \leq 0, \forall k, \\ \frac{\beta_k}{\delta_k + 1} \| \mathbf{C}_k^{1/2} \mathbf{p}_{\text{RIS}} \| \leq x_k, \forall k, \\ y_k \mathbf{I} - \frac{\beta_k}{\delta_k + 1} \Sigma_{r,k} \Lambda_{\text{RIS}}^a \succeq \mathbf{0}, \forall k, \\ y_k \geq 0, \forall k, \end{array} \right. \quad (38)$$

where $\mathbf{C}_k = (\Sigma_{r,k}^T + 2\delta_k (\bar{\mathbf{h}}_{r,k}^* \bar{\mathbf{h}}_{r,k}^T)) \odot \Sigma_{r,k}$, and $\mathbf{x} = [x_1, \dots, x_K]^T$ and $\mathbf{y} = [y_1, \dots, y_K]^T$ are auxiliary variables.

Using $P(\mathbf{F}, \mathbf{w})$, defined in (8), and the analytical constraints in (38), Problem (31) can be equivalently transformed into

$$\min_{\mathbf{F}, \mathbf{w}, \mathbf{x}, \mathbf{y}} \|\mathbf{F}\|_F^2 + P(\mathbf{F}, \mathbf{w}) \quad (39a)$$

$$\text{s.t. (38),}$$

$$1 \leq |w_m|^2 \leq a_{\max}, \forall m. \quad (39b)$$

To overcome the coupling of variables \mathbf{F} and \mathbf{w} , we employ AO to solve Problem (39). The resulting non-convex subproblems for \mathbf{F} and \mathbf{w} are separately relaxed by using SDR [38] and then solved with CVX in an iterative manner.

C. Optimization of Precoding Matrix \mathbf{F}

Given \mathbf{w} , we define new variables $\Gamma_k = \mathbf{f}_k \mathbf{f}_k^H$ constrained by $\Gamma_k \succeq \mathbf{0}$ and $\text{rank}(\Gamma_k) = 1, \forall k$. Correspondingly, Φ_k in (34a) can be rewritten as $\Phi_k = \Gamma_k / (2^{r_k} - 1) - \sum_{i=1, i \neq k}^K \Gamma_i$, and the objective function in (39a) can be re-expressed as

$$\|\mathbf{F}\|_F^2 + P(\mathbf{F}) = \text{Tr} \left\{ (\mathbf{I} + \mathbf{D}_F) \sum_{k=1}^K \Gamma_k \right\} + \sigma_z^2 \text{Tr} \{ \Lambda_{\text{RIS}}^a \}. \quad (40)$$

Since constraint $\text{rank}(\Gamma_k) = 1$ is non-convex and hard to be tackled directly. We adopt the SDR technique, that is, we first obtain an intermediate solution by dropping the rank-one constraint, and then construct a rank-one optimal solution from the intermediate solution. Specifically, by removing the rank-one constraints, the relaxed subproblem for $\Gamma = [\Gamma_1, \dots, \Gamma_K]$ of Problem (39) is given by

$$\min_{\Gamma, \mathbf{x}, \mathbf{y}} \text{Tr} \left\{ (\mathbf{I} + \mathbf{D}_F) \sum_{k=1}^K \Gamma_k \right\} \quad (41a)$$

s.t. (38),

$$\mathbf{\Gamma}_k \succeq \mathbf{0}, \forall k, \quad (41b)$$

where we have omitted all irrelative constant terms that do not depend on $\mathbf{\Gamma}$. Problem (41) is a standard SDP and can be solved using CVX. The following theorem further reveals the tightness of SDR for Problem (41), the proof of which can be found in [39, Appendix C].

Theorem 2 *Assuming that the relaxed Problem (41) is feasible, there always exists a feasible solution $\{\mathbf{\Gamma}_k^*\}_{k=1}^K$ satisfying $\text{rank}(\mathbf{\Gamma}_k^*) = 1, \forall k$.*

Based on Theorem 2, the optimal BS beamforming vectors $\{\mathbf{f}_k^*\}_{k=1}^K$ can be obtained from $\{\mathbf{\Gamma}_k^*\}_{k=1}^K$ via eigenvalue decomposition.

D. Optimization of Reflection Vector \mathbf{w}

Next, we consider the subproblem of solving \mathbf{w} for a given \mathbf{F} . We introduce auxiliary variable $\tilde{\mathbf{W}} = \tilde{\mathbf{w}}\tilde{\mathbf{w}}^H$ with constraints $\tilde{\mathbf{W}} \succeq \mathbf{0}$ and $\text{rank}(\tilde{\mathbf{W}}) = 1$, thus $\mathbf{\Lambda}_{\text{RIS}}^a$ defined in (8) and \mathbf{p}_{RIS} defined in (37) can be expressed as $\mathbf{\Lambda}_{\text{RIS}}^a = \text{Diag}(\text{diag}([\tilde{\mathbf{W}}]_{1:M,1:M}))$ and $\mathbf{p}_{\text{RIS}} = \text{diag}([\tilde{\mathbf{W}}]_{1:M,1:M})$, respectively. Correspondingly, the objective function in (21) and the constraints in (38) are equivalent to

$$P(\tilde{\mathbf{W}}) = \mathbf{w}^H \mathbf{D}_w \mathbf{w} = \text{Tr} \left\{ \mathbf{D}_w \text{Diag}(\text{diag}([\tilde{\mathbf{W}}]_{1:M,1:M})) \right\}, \quad (42)$$

and

$$\left\{ \begin{array}{l} \frac{\beta_k}{\delta_k+1} \text{Tr} \left\{ \text{Diag}(\text{diag}([\tilde{\mathbf{W}}]_{1:M,1:M})) \right\} + \sqrt{2 \ln(1/\rho_k)} x_k - \ln(\rho_k) y_k \\ - \frac{1}{\sigma_z^2} \left(\text{Tr} \left\{ \mathbf{H}_k \mathbf{\Phi}_k \mathbf{H}_k^H \tilde{\mathbf{W}} \right\} - \sigma_z^2 \right) \leq 0, \forall k, \\ \frac{\beta_k}{\delta_k+1} \|\mathbf{C}_k^{1/2} \text{diag}([\tilde{\mathbf{W}}]_{1:M,1:M})\| \leq x_k, \forall k, \\ y_k \mathbf{I} - \frac{\beta_k}{\delta_k+1} \mathbf{\Sigma}_{r,k} \text{Diag}(\text{diag}([\tilde{\mathbf{W}}]_{1:M,1:M})) \succeq \mathbf{0}, \forall k, \\ y_k \geq 0, \forall k. \end{array} \right. \quad (43)$$

Adopting again the SDR technique and removing non-convex constraint $\text{rank}(\tilde{\mathbf{W}}) = 1$, we obtain the rank-relaxed subproblem for $\tilde{\mathbf{W}}$ of Problem (39) by ignoring irrelevant constants as follows

$$\min_{\tilde{\mathbf{W}}, \mathbf{x}, \mathbf{y}} P(\tilde{\mathbf{W}}) \quad (44a)$$

$$\text{s.t. (43),}$$

$$1 \leq [\text{diag}(\tilde{\mathbf{W}})]_m \leq a_{\max}, 1 \leq m \leq M, \quad (44b)$$

$$[\text{diag}(\tilde{\mathbf{W}})]_{M+1} = 1, \quad (44c)$$

$$\tilde{\mathbf{W}} \succeq \mathbf{0}, \quad (44d)$$

which is a standard SDP and can be solved using CVX. Since the diagonal elements of $\tilde{\mathbf{W}}$ are independently constrained in (44b) and (44c), the optimal solution, $\tilde{\mathbf{W}}^*$, of (44) may not be rank-one. Therefore, only a suboptimal $\tilde{\mathbf{w}}^*$ can be constructed from $\tilde{\mathbf{W}}^*$ by using the Gaussian decomposition technique. Specifically, we consider the eigenvalue decomposition of $\tilde{\mathbf{W}}^*$, $\tilde{\mathbf{W}}^* = \mathbf{E}\mathbf{\Upsilon}\mathbf{E}^H$, where the columns of \mathbf{E} are the eigenvectors of $\tilde{\mathbf{W}}^*$, and diagonal matrix $\mathbf{\Upsilon}$ contains the corresponding eigenvalues. Then, we compute 1000 candidate vectors, $\{\mathbf{v}_i = \mathbf{E}\mathbf{\Upsilon}^{1/2}\mathbf{e}_i/[\mathbf{E}\mathbf{\Upsilon}^{1/2}\mathbf{e}_i]_{M+1}\}_{i=1}^{1000}$ with $\mathbf{e}_i \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{M+1})$, such that each \mathbf{v}_i satisfies the QoS constraints. Then, the \mathbf{v}_i that minimizes the total power consumption is selected as the optimal $\tilde{\mathbf{w}}^*$. In order to ensure convergence of the proposed AO algorithm, in each iteration, we need to find a $\tilde{\mathbf{w}}^*$ that decreases the objective function value compared with the previous iteration, see Section IV-F. This can always be achieved empirically by generating a sufficient number of trial vectors for Gaussian randomization.

E. Computational Complexity

As CVX employs the interior point method, the computational complexity of solving Problems (41) and (44) is given by [40]

$$\mathcal{O}\left(\underbrace{\left(\sum_{j=1}^J c_j + 2I\right)^{1/2} n(n^2 + n \sum_{j=1}^J c_j^2 + \sum_{j=1}^J c_j^3 + n \sum_{i=1}^I v_i^2)}_{\text{due to LMI}}\right),$$

$\underbrace{\hspace{10em}}_{\text{due to SOC}}$

where n is the number of variables, J is the number of linear matrix inequalities (LMIs) of size c_j , and I is the number of second-order cone (SOC) constraints of size v_i . For Problem (41), the number of variables is $n_1 = NK$, (38) only contains linear constraints, and (41b) contains K LMIs of size N . Therefore, the approximate complexity of Problem (41) is $o_F = \mathcal{O}([KN]^{1/2}n_1[n_1^2 + n_1KN^2 + KN^3])$. For Problem (44), there are $n_2 = M$ variables, K LMIs of size M and K SOC of size M in (43), and one LMI of size $M+1$ in (44d). The remaining constraints are linear. Therefore, the approximate complexity of Problem (44) is $o_e = \mathcal{O}([KM + 2K]^{1/2}n_2[n_2^2 + 2n_2KM^2 + KM^3])$ by neglecting terms with low-order complexity. Finally, the approximate complexity per iteration is $o_F + o_e$.

F. Convergence Analysis

Finally, we analyze the convergence behavior obtained by alternately solving Problems (41) and (44) for solving Problem (39). Let $g(\mathbf{F}, \mathbf{w})$ denote the objective value of Problem (39). Given \mathbf{w}^n in the n^{th} iteration, we have

$$g(\mathbf{F}^n, \mathbf{w}^n) \geq g(\mathbf{F}^{n+1}, \mathbf{w}^n),$$

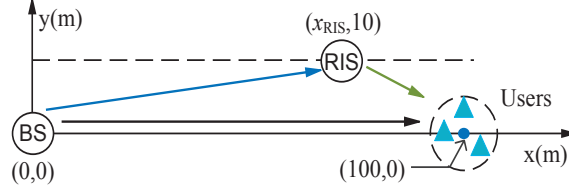


Fig. 2: The simulation system setup.

as we can find the global optimal solution for Problem (41) based on Theorem 2. Then, given \mathbf{F}^{n+1} , we can always find a \mathbf{w}^{n+1} for Problem (44) satisfying

$$g(\mathbf{F}^{n+1}, \mathbf{w}^n) \geq g(\mathbf{F}^{n+1}, \mathbf{w}^{n+1})$$

by using the Gaussian decomposition technique. Therefore, the sequence of objective values $\{g(\mathbf{F}^{n+1}, \mathbf{w}^{n+1})\}$ generated in an alternating manner is monotonically non-increasing. Thus, the obtained solutions are stationary points of Problem (39).

V. NUMERICAL RESULTS

In this section, we provide numerical results to evaluate the performance of an active RIS-aided system, where the BS and an active or passive RIS are located at (0 m, 0 m) and (x_{RIS} m, 10 m), as shown in Fig. 2. K users are randomly and uniformly distributed in a circle with a radius of 5 m and centered at (100 m, 0 m). The large-scale pathloss coefficients are modelled as $\beta = -\text{PL}_0 - 10\alpha \log_{10}(d)$ dB, where d is the link distance in meters and α is the pathloss exponent which is set to 3.5 and 2 for the BS-user and the RIS-aided links, respectively. $\text{PL}_0 = 40$ dB denotes the pathloss at a distance of 1 meter, i.e., we assume a carrier frequency of 3.5 GHz [41]. Unless specified otherwise, the BS and the RIS are equipped with $N = 8$ antennas and $M = 32$ reflecting elements, respectively, the maximum amplification gain of the RIS is assumed to be $a_{\text{max}} = 40$ dB, the Rician factors are $\delta_0 = \dots = \delta_K = \delta = 10$, and the noise power at the RIS and the users are set to $\sigma_z^2 = \sigma_1^2 = \dots = \sigma_K^2 = -80$ dBm.

Compared with passive RISs, the power consumed by active RISs also includes the transmit power and the circuit power for amplification. Thus, the maximum total RIS power consumption is given by $P_{\text{RIS}} = P_M + P_{\text{cir}}$. Here, the circuit power $P_{\text{cir}} = M(P_c + P_{\text{DC}})$ comprises the power consumed by the phase shifters and the control circuits of the RIS elements, P_c , and the DC biasing power, P_{DC} , used to drive the amplifiers of the active RIS elements. For consistency, the circuit power of each RF chain, P_{RF} , is also accounted for in the maximum BS power consumption, denoted by $P_{\text{BS}} = P_N + NP_{\text{RF}}$. According to [9], we set $P_{\text{DC}} = -5$ dBm, $P_c = -10$ dBm, and $P_{\text{RF}} = 23$ dBm.

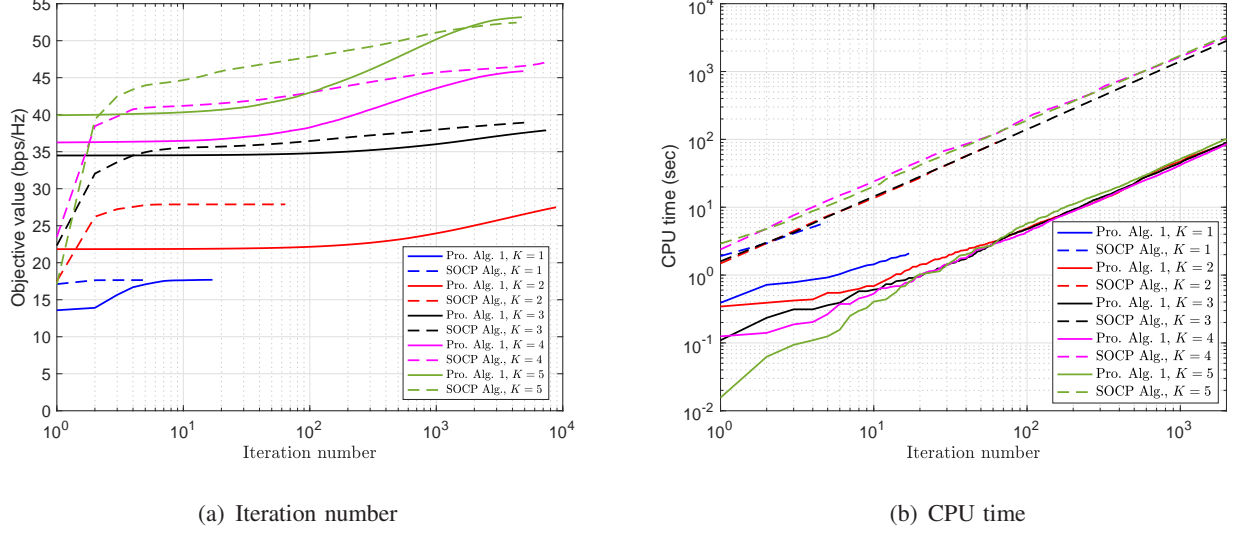


Fig. 3: The convergence behaviour of different algorithms for 1 randomly generated channel realization, when $N = 8$, $M = 32$, and $a_{max} = 40$ dB.

A. Maximum Achievable Rate

In this subsection, we evaluate the ergodic achievable rate of the active RIS-aided system as discussed in Section III for 500 independent realizations of $\{\mathbf{h}_k, \mathbf{G}_k\}_{k=1}^K$, except for the convergence analysis in Fig. 3. We denote the ergodic achievable rate obtained by averaging the lower bound on the average achievable rate in (7)³ and the corresponding instantaneous achievable rate in (6) over all channel realizations as “Act. RIS-LB” and “Act. RIS”, respectively. For comparison, we also consider an upper bound for “Act. RIS”, denoted as “Act. RIS-perfect”, for which we assume the availability of perfect CSI, and a corresponding lower bound, denoted as “Act. RIS-non-robust”, for which we ignore the NLoS components of the RIS-aided channels in (6) for beamformer design. In addition, systems with passive RIS and without RIS are also considered as performance benchmarks, and are denoted as “Pas. RIS” and “No RIS”, respectively. We determine the total power consumption of the active RIS-aided system as $P_{BS} + P_{RIS}$, that of the passive RIS-aided system as $P_{BS} + MP_c$, and that of the system without RIS as P_{BS} . For a fair comparison, the maximum total power consumption is set to the same value for all considered schemes.

Fig. 3 illustrates the convergence and complexity of the proposed Algorithm 1, denoted as “Pro. Alg. 1”, where the RIS is located at (80 m, 10 m). An SOCP-based algorithm is considered as a

³The ergodic achievable rate and the average achievable rate should not be confused: the latter is given by $\mathbb{E}_{\mathbf{h}_{r,k}|\mathbf{G}_k} \{R_k(\mathbf{F}, \mathbf{w})\}$ in (7), while the former is given by $\mathbb{E}_{\mathbf{h}_k, \mathbf{G}_k} \{R_k(\mathbf{F}, \mathbf{w})\} = \mathbb{E}_{\mathbf{h}_k, \mathbf{G}_k} \{\mathbb{E}_{\mathbf{h}_{r,k}|\mathbf{G}_k} \{R_k(\mathbf{F}, \mathbf{w})\}\}$.

benchmark algorithm, denoted as “SOCP Alg.”. For the “SOCP Alg.”, auxiliary variables are introduced to transfer the non-concave rate expression in the objective function to the constraints, and then SCA is used to handle the non-convex constraints. The resulting SOCP problem with multiple constraints can be directly solved using CVX. An SOCP based algorithms can handle optimization problems with multiple and complex constraints, but for a large number of variables, their complexity becomes high. As can be seen in Fig. 3(a), “SOCP Alg.” converges faster than “Pro. Alg. 1” if the number of users is small ($K = 1, 2$), while it loses its advantage for large numbers of users ($K = 3, 4, 5$). Fig. 3(b) shows that the CPU time consumed by “Pro. Alg. 1” is significantly smaller than that of “SOCP Alg.”, especially when the number of users is large. This is because the number of variables in multi-user systems is high, which causes a high computational complexity per iteration in “SOCP Alg.”, while the complexity per iteration of Algorithm 1, benefiting from semi-closed-form solutions, is low and not sensitive to the number of variables.

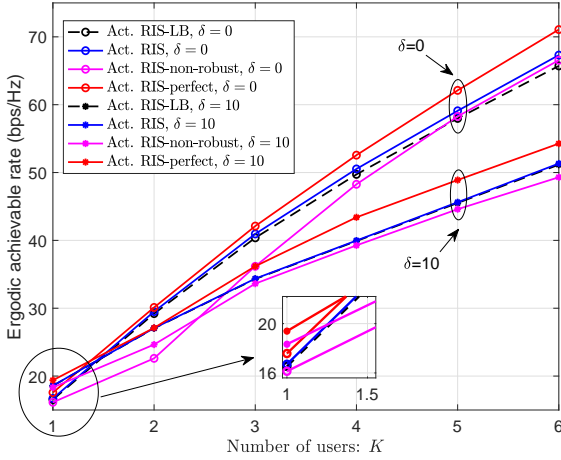


Fig. 4: Achievable rate versus number of users, when $N = 8$, $M = 32$, and $a_{max} = 40$ dB.

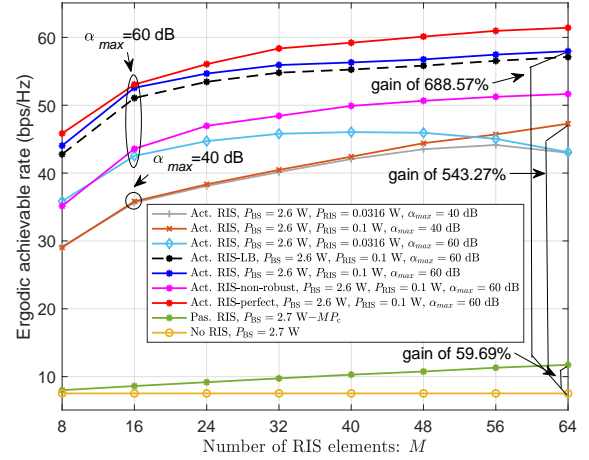


Fig. 5: Achievable rate versus the number of RIS elements, when $N = 8$ and $K = 4$.

Fig. 4 investigates the ergodic achievable rate for RIS-aided systems as a function of the number of users. Here, P_{BS} and P_{RIS} are set to 2.6 W and 0.1 W, respectively. *First*, as can be observed, the ergodic achievable rate of “Act. RIS-LB” is only slightly lower than that of “Act. RIS”, when the RIS-aided channels follow a Rayleigh distribution ($\delta = 0$). When the Rician factor increases to 10, which means a reduction of the uncertain NLoS components, the ergodic achievable rate of “Act. RIS-LB” is almost equal to that of “Act. RIS”. This confirms the tightness of the proposed lower bound expression in (7). *Furthermore*, the ergodic achievable rates for “Act. RIS” and “Act. RIS-perfect” are almost the same in the single-user case, and the gap between them increases with the number of

users. This is because the negative impact of partial CSI becomes more significant as the number of users increases. *Finally*, the proposed “Act. RIS” always outperforms “Act. RIS-non-robust”, which reveals that a robust design is needed and that the proposed problem formulation and the corresponding algorithm can efficiently mitigate the performance loss caused by partial CSI.

The ergodic achievable rate as a function of the number of RIS elements is shown in Fig. 5 for $K = 4$ users and with a passive or active RIS fixed at (80 m, 10 m). It is observed that the improvement in ergodic achievable rate provided by the active RIS is affected by the amplification gain α_{max} and the RIS power consumption P_{RIS} . *First*, for $a_{max} = 40$ dB, increasing P_{RIS} from 0.0316 W to 1 W yields little performance improvement, which means that each reflecting element is operating with the maximum amplification gain in this scenario, i.e., $a_{max} = 40$ dB limits the RIS power consumption even for $P_{\text{RIS}} = 0.0316$ W. For $a_{max} = 60$ dB, increasing P_{RIS} to 1 W yields a further improvement in the ergodic achievable rate. *Next*, for $P_{\text{RIS}} = 0.0316$ W, more RIS reflecting elements may actually reduce the ergodic achievable rate. This is because the RIS circuit power consumption increases with the number of RIS elements, and as a result, the available RIS transmit power decreases and leads to a performance loss for $P_{\text{RIS}} = 0.0316$ W. Although the available RIS transmit power is also reduced for $P_{\text{RIS}} = 1$ W, the remaining transmit power is sufficient to support the additional RIS elements to achieve a performance improvement due to the resulting increased beamforming gain. *Furthermore*, for $P_{\text{RIS}} = 0.1$ W, the performance gap between “Act. RIS” and “Act. RIS-perfect” increases with the number of RIS reflecting elements, which reveals a higher performance loss for RIS-aided channels with more coefficients due to the higher impact of the imperfection caused by partial CSI. Nevertheless, compared to “Act. RIS-non-robust”, the proposed “Act. RIS” can still efficiently mitigate the uncertainty of the partial CSI. *Finally*, for $M = 64$ and a total power consumption of 2.7 W, compared with the No RIS scenario, the passive RIS yields a maximum performance gain of 59.96%, while the active RIS with $a_{max} = 40$ dB and $a_{max} = 60$ dB achieves performance gains of 543.27% and 688.57%, respectively.

B. Minimum Average Power

The minimum average power consumption investigated in Section IV is evaluated in this subsection. Each point in the following figures is obtained by averaging over 500 independent channel realizations. The maximum outage probabilities and target rates of all users are respectively assumed to be identical, i.e., $\rho_1 = \dots = \rho_K = \rho = 0.05$ and $r_1 = \dots = r_K = r$. For a fair comparison, the total achievable power consumption including the total transmit power and the circuit power consumption, i.e., $\|\mathbf{F}\|_F^2 + P(\mathbf{F}, \mathbf{w}) + M(P_c + P_{\text{DC}}) + NP_{\text{RF}}$, is adopted as performance metric. For the benchmark

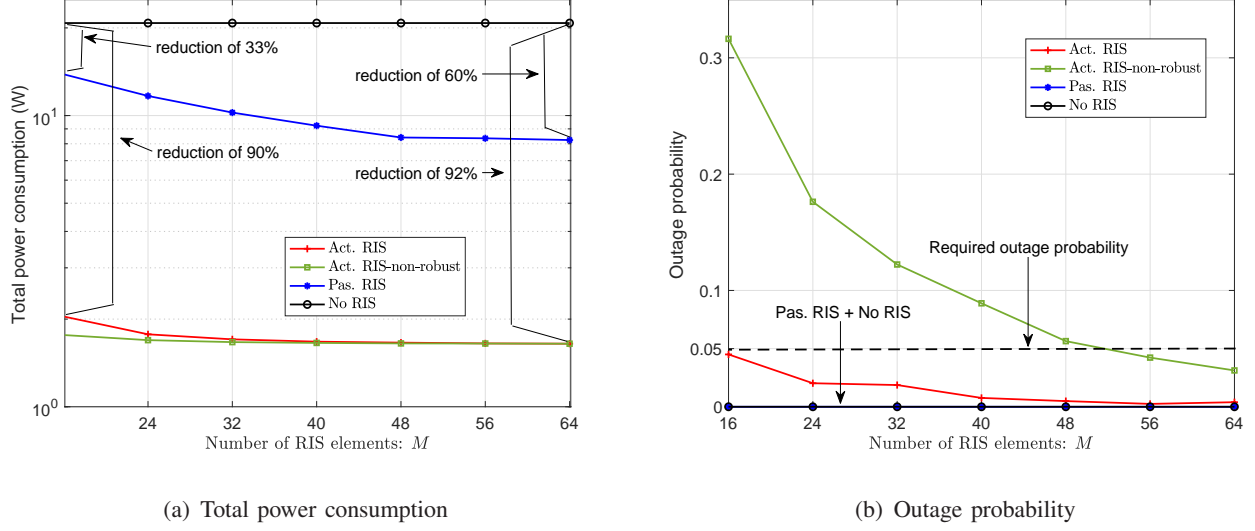


Fig. 6: Total power consumption and outage probability versus the number of RIS elements, when $N = 8$, $K = 4$, $\rho = 0.95$, and $r = 5$ bps/Hz.

“Act. RIS-non-robust”, the beamformer in Problem (39) is obtained by ignoring the NLoS components of the RIS-aided channels.

Fig. 6 shows the minimum total power consumption and the outage probability versus the number of RIS elements for a multi-user system ($K = 4$), wherein the target rate of each user is $r = 5$ bps/Hz. The other parameters are set to the same values as for Fig. 5. *First*, as can be observed in Fig. 6(a), an RIS equipped with only 16 active elements can reduce the total power consumption by 90% compared to the “No RIS” scenario, while an RIS with 16 passive elements can reduce the total power consumption by only 33%. Increasing the number of active reflecting elements further to $M = 64$ can reduce the total power consumption by 92%, compared to the case without RIS. *Second*, the “Act. RIS-non-robust” scheme consumes the least power as the NLoS components of RIS-user links are ignored for beamformer design. However, this comes at the expense of a high outage probability, cf. Fig. 6(b).

To further demonstrate the effectiveness of the proposed “Act. RIS”, Fig. 6(b) compares the outage probabilities of “Act. RIS” and “Act. RIS-non-robust”. In particular, the outage probability for each channel realization $\{\mathbf{h}_k, \mathbf{G}_k\}_{k=1}^K$ is calculated as follows: For a given channel realization $\{\mathbf{h}_k, \mathbf{G}_k\}_{k=1}^K$, 1000 conditional channel realizations $\{(\mathbf{H}_{\text{dr}}, \{\mathbf{h}_{r,k}\}_{k=1}^K)^{(z)}\}_{z=1}^{1000}$ are drawn from their distributions. Then, the outage probability is defined as the ratio of the number of outage conditional channel realizations to the total number of conditional channel realizations, where an outage is declared when the target

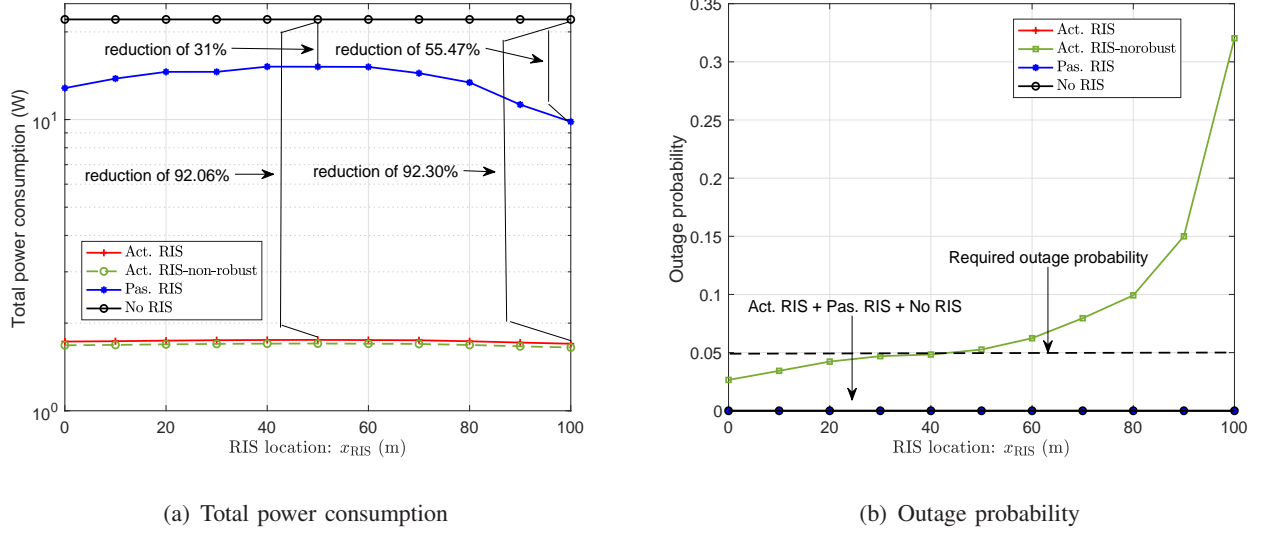


Fig. 7: Average power consumption versus RIS location, when $N = 8$, $M = 8$, $K = 1$, $\rho = 0.95$, and $r = 10$ bps/Hz.

rate of at least one user cannot be satisfied. Fig. 6(b) reveals that system outages occur frequently with “Act. RIS-non-robust”, especially for low values of M ($M \leq 48$) due to low spatial diversity gain. However, the proposed “Act. RIS” scheme can effectively control the system outage probability to a very low level and meets the required outage probability, i.e., $\rho = 0.05$, for $M \geq 16$. This illustrates the ability of the proposed scheme to mitigate the uncertainties of partial CSI.

To investigate the impact of the location of the RIS on the total power consumption and the outage probability, Fig. 7 considers only a single user located at (100 m, 0 m), cf. Fig. 2. *First*, compared with the “No RIS” scenario, the “Pas. RIS” system can reduce the total power consumption by 31% ~ 55.47%, while the total power consumption reduction in the “Act. RIS” system is considerably higher and reaches 92%, as shown in Fig. 7(a). *Second*, the “Pas. RIS” system has the worst performance if it is placed in the middle of the BS-user link due to the severe multiplicative fading. On the other hand, “Act. RIS” can significantly mitigate the impact of multiplicative fading in the middle of the BS-user link and yields a better performance. *Finally*, the total power consumption of the proposed “Act. RIS” scheme and “Act. RIS-non-robust” are almost the same, cf. Fig. 7(a), but, “Act. RIS” yields zero outage probability, cf. Fig. 7(b), which again underscores the benefit of the proposed scheme in mitigating the uncertainties introduced by CSI.

VI. CONCLUSIONS

In this work, we have addressed the practical problem that perfect individual CSI knowledge of the RIS-aided channels in active RIS systems is not available. Considering this limitation, we have derived

analytical expressions for the average achievable rate and the average RIS transmit power taking into account the partial CSI knowledge of the individual RIS-aided channels. To address the uncertainty caused by the partial CSI, we formulated joint BS and RIS beamforming optimization problems to respectively maximize the average sum achievable rate and minimize the average total transmit power subject to rate outage probability constraints. For the average sum achievable rate maximization problem, a computationally efficient AO algorithm exploiting closed-form expressions in every iteration has been proposed under the MM framework. Furthermore, to facilitate the beamforming design for average transmit power minimization, we adopted the BTI to bound the rate outage probability constraints. Subsequently, an AO algorithm with guaranteed convergence was developed by exploiting SDR. Our simulation results confirmed that the proposed design for average achievable rate maximization closely approaches the performance obtained for perfect CSI. Moreover, our results revealed that compared to the non-robust scheme ignoring the unknown NLoS components, the proposed rate outage constrained design guarantees the QoS of each user.

APPENDIX A

THE PROOF OF THEOREM 1

To prove Property 1, we exploit $\text{vec}(\mathbf{A}\text{Diag}(\mathbf{b})\mathbf{C}) = (\mathbf{C}^T \diamond \mathbf{A})\mathbf{b}$ [35, Equ. (1.11.21)]. Then, we have

$$\begin{aligned} \text{Tr} \{ \mathbf{A}\text{Diag}(\mathbf{b})\mathbf{C}\text{Diag}(\mathbf{b}) \} &= (\text{vec}(\mathbf{C}^T \text{Diag}(\mathbf{b})\mathbf{A}^T))^T \text{vec}(\text{Diag}(\mathbf{b})) \\ &= ((\mathbf{A} \diamond \mathbf{C}^T)\mathbf{b})^T \text{vec}(\text{Diag}(\mathbf{b})) \end{aligned} \quad (45)$$

$$\begin{aligned} &= \mathbf{b}^T (\mathbf{A} \diamond \mathbf{C}^T)^T \text{vec}(\text{Diag}(\mathbf{b})) \\ &= \mathbf{b}^T (\mathbf{A}^T \odot \mathbf{C})\mathbf{b}, \end{aligned} \quad (46)$$

where (45) is due to property $\text{vec}(\mathbf{A}\text{Diag}(\mathbf{b})\mathbf{C}) = (\mathbf{C}^T \diamond \mathbf{A})\mathbf{b}$, and (46) is due to

$$\begin{aligned} (\mathbf{A} \diamond \mathbf{C}^T)^T \text{vec}(\text{Diag}(\mathbf{b})) &= \begin{bmatrix} (\mathbf{a}_1 \otimes \mathbf{c}_1)^T \\ (\mathbf{a}_2 \otimes \mathbf{c}_2)^T \\ \vdots \\ (\mathbf{a}_N \otimes \mathbf{c}_N)^T \end{bmatrix} \text{vec}(\text{Diag}(\mathbf{b})) \\ &= \begin{bmatrix} (\mathbf{a}_1 \odot \mathbf{c}_1)^T \\ (\mathbf{a}_2 \odot \mathbf{c}_2)^T \\ \vdots \\ (\mathbf{a}_N \odot \mathbf{c}_N)^T \end{bmatrix} \mathbf{b} \\ &= (\mathbf{A}^T \odot \mathbf{C})\mathbf{b}, \end{aligned}$$

where $\mathbf{A} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_N \end{bmatrix}$ and $\mathbf{C}^T = \begin{bmatrix} \mathbf{c}_1 & \mathbf{c}_2 & \cdots & \mathbf{c}_N \end{bmatrix}$.

Thus, Theorem 1 is proved.

REFERENCES

- [1] G. Zhou, C. Pan, and H. Ren, "Active reconfigurable intelligent surface aided communication with partial CSI," *The 14th International Conference on Wireless Communications and Signal Processing (WCSP 2022)*, *accepted*.
- [2] 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.
- [3] C. Pan, H. Ren, K. Wang *et al.*, "Reconfigurable intelligent surfaces for 6G systems: Principles, applications, and research directions," *IEEE Commun. Mag.*, vol. 59, no. 6, pp. 14–20, Jun. 2021.
- [4] X. You, C.-X. Wang, J. Huang *et al.*, "Towards 6G wireless communication networks: Vision, enabling technologies, and new paradigm shifts," *Sci. China Inf. Sci.*, vol. 64, no. 1, pp. 1–74, 2021.
- [5] W. Saad, M. Bennis, and M. Chen, "A vision of 6G wireless systems: Applications, trends, technologies, and open research problems," *IEEE Network*, vol. 34, no. 3, pp. 134–142, May/Jun. 2020.
- [6] M. Najafi, V. Jamali, R. Schober, and H. V. Poor, "Physics-based modeling and scalable optimization of large intelligent reflecting surfaces," *IEEE Trans. Commun.*, vol. 69, no. 4, pp. 2673–2691, Apr. 2021.
- [7] 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, Jun. 2020.
- [8] H. Shen, W. Xu, S. Gong, Z. He, and C. Zhao, "Secrecy rate maximization for intelligent reflecting surface assisted multi-antenna communications," *IEEE Commun. Lett.*, vol. 23, no. 9, pp. 1488–1492, Jun. 2019.
- [9] R. Long, Y. Liang, Y. Pei, and E. G. Larsson, "Active reconfigurable intelligent surface-aided wireless communications," *IEEE Trans. Wireless Commun.*, vol. 20, no. 8, pp. 4962–4975, Aug. 2021.
- [10] Z. Zhang, L. Dai, X. Chen, C. Liu, F. Yang, R. Schober, and H. V. Poor, "Active RIS vs. passive RIS: Which will prevail in 6G?" 2021. [Online]. Available: <https://arxiv.org/abs/2103.15154>
- [11] K. Zhi, C. Pan, H. Ren, K. K. Chai, and M. ElKashlan, "Active RIS versus passive RIS: Which is superior with the same power budget?" *IEEE Commun. Lett.*, vol. 26, no. 5, pp. 1150–1154, May 2022.
- [12] C. You and R. Zhang, "Wireless communication aided by intelligent reflecting surface: Active or passive?" *IEEE Wireless Commun. Lett.*, vol. 10, no. 12, pp. 2659–2663, Dec. 2021.
- [13] D. Xu, X. Yu, D. W. Kwan Ng, and R. Schober, "Resource allocation for active IRS-assisted multiuser communication systems," in *Proc. 55th Asilomar Conference on Signals, Systems, and Computers*, Nov. 2021, pp. 113–119.
- [14] P. Zeng, D. Qiao, Q. Wu, and Y. Wu, "Active IRS aided WPCNs: A new paradigm towards higher efficiency and wider coverage," 2021. [Online]. Available: <https://arxiv.org/abs/2111.11600v1>
- [15] C. Pan, G. Zhou, K. Zhi *et al.*, "An overview of signal processing techniques for RIS/IRS-aided wireless systems," 2021. [Online]. Available: <https://arxiv.org/abs/2112.05989>
- [16] S. Shen, B. Clerckx, and R. Murch, "Modeling and architecture design of reconfigurable intelligent surfaces using scattering parameter network analysis," *IEEE Trans. Wireless Commun.*, vol. 21, no. 2, pp. 1229–1243, Feb. 2022.
- [17] 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.
- [18] A. L. Swindlehurst, G. Zhou, R. Liu, C. Pan, and M. Li, "Channel estimation with reconfigurable intelligent surfaces-a general framework," *Proc. IEEE*, vol. 110, no. 9, pp. 1312–1338, Sept. 2022.

- [19] Y. Wei, M.-M. Zhao, M.-J. Zhao, and Y. Cai, "Channel estimation for IRS-aided multiuser communications with reduced error propagation," *IEEE Trans. Wireless Commun.*, vol. 21, no. 4, pp. 2725–2741, Apr. 2022.
- [20] X. Wei, D. Shen, and L. Dai, "Channel estimation for RIS assisted wireless communications: Part II - an improved solution based on double-structured sparsity," *IEEE Commun. Lett.*, vol. 25, no. 5, pp. 1403–1407, May 2021.
- [21] G. Zhou, C. Pan, H. Ren, P. Popovski, and A. L. Swindlehurst, "Channel estimation for RIS-aided multiuser millimeter-wave systems," *IEEE Trans. Signal Process.*, vol. 70, pp. 1478–1492, Mar. 2022.
- [22] 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.
- [23] X. Yu, D. Xu, and R. Schober, "Enabling secure wireless communications via intelligent reflecting surfaces," in *Proc. IEEE Global Communications Conference (GLOBECOM)*, Dec. 2019, pp. 1–6.
- [24] 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, Jun 2020.
- [25] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2005.
- [26] Z. Abu-Shaban, X. Zhou, T. Abhayapala, G. Seco-Granados, and H. Wymeersch, "Error bounds for uplink and downlink 3D localization in 5G millimeter wave systems," *IEEE Trans. Wireless Commun.*, vol. 17, no. 8, pp. 4939–4954, Aug. 2018.
- [27] A. Elzanaty, A. Guerra, F. Guidi, and M.-S. Alouini, "Reconfigurable intelligent surfaces for localization: Position and orientation error bounds," *IEEE Transactions on Signal Processing*, vol. 69, pp. 5386–5402, Aug. 2021.
- [28] Y.-C. Liang and F. Chin, "Downlink channel covariance matrix (DCCM) estimation and its applications in wireless DS-CDMA systems," *IEEE J. Sel. Areas Commun.*, vol. 19, no. 2, pp. 222–232, Feb. 2001.
- [29] E. Björnson and L. Sanguinetti, "Rayleigh fading modeling and channel hardening for reconfigurable intelligent surfaces," *IEEE Wireless Commun. Lett.*, vol. 10, no. 4, pp. 830–834, Apr. 2021.
- [30] R. Zhang, S.-H. Leung, Z. Luo, and H. Wang, "Precoding design for correlated MIMO-AF relay networks with statistical channel state information," *IEEE Trans. Signal Process.*, vol. 66, no. 22, pp. 5902–5916, Nov. 2018.
- [31] Hunter, D. R., and K. Lange, "A tutorial on MM algorithms," *The American Statistician*, vol. 58, no. 1, pp. 30–37, 2004.
- [32] Y. Sun, P. Babu, and D. P. Palomar, "Majorization-minimization algorithms in signal processing, communications, and machine learning," *IEEE Trans. Signal Process.*, vol. 65, no. 3, pp. 794–816, Feb. 2017.
- [33] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, *Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers*, Apr. 2011, vol. 3, no. 1.
- [34] R. Varadhan and C. Roland, "Simple and globally convergent methods for accelerating the convergence of any EM algorithm," *Scand. J. Statist.*, vol. 35, no. 2, pp. 335–353, 2008.
- [35] X.-D. Zhang, *Matrix Analysis and Applications*. Cambridge University Press., 2017.
- [36] M. W. Jacobson and J. A. Fessler, "An expanded theoretical treatment of iteration-dependent majorize-minimize algorithms," *IEEE Trans. Image Process.*, vol. 16, no. 10, pp. 2411–2422, Oct. 2007.
- [37] K. Wang, A. M. So, T. Chang, W. Ma, and C. Chi, "Outage constrained robust transmit optimization for multiuser MISO downlinks: Tractable approximations by conic optimization," *IEEE Trans. Signal Process.*, vol. 62, no. 21, pp. 5690–5705, Nov. 2014.
- [38] Z. Luo, W. Ma, A. M. 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.
- [39] G. Zhou, C. Pan, H. Ren, K. Wang, and A. Nallanathan, "A framework of robust transmission design for IRS-aided MISO communications with imperfect cascaded channels," *IEEE Trans. Signal Process.*, vol. 68, pp. 5092–5106, Aug. 2020.
- [40] A. Ben-Tal and A. Nemirovski, (*Lectures on modern convex optimization: Analysis, algorithms, and engineering applications*). Philadelphia, PA, USA: SIAM. MPSSIAM Ser. Optim., 2001.
- [41] 3GPP, "Technical specification group radio access network; study on 3D channel model for LTE (release 12)," *TR 36.873 V12.7.0*, Dec. 2017.