Achievable Rate of a STAR-RIS Assisted Massive MIMO System Under Spatially-Correlated Channels

Anastasios Papazafeiropoulos, Le-Nam Tran, Zaid Abdullah, Pandelis Kourtessis, Symeon Chatzinotas

Abstract—Reconfigurable intelligent surfaces (RIS)-assisted massive multiple-input multiple-output (mMIMO) is a promising technology for applications in next-generation networks. However, reflecting-only RIS provides limited coverage compared to a simultaneously transmitting and reflecting RIS (STAR-RIS). Hence, in this paper, we focus on the downlink achievable rate and its optimization of a STAR-RIS-assisted mMIMO system. Contrary to previous works on STAR-RIS, we consider mMIMO, correlated fading, and multiple user equipments (UEs) at both sides of the RIS. In particular, we introduce an estimation approach of the aggregated channel with the main benefit of reduced overhead links instead of estimating the individual channels. Next, leveraging channel hardening in mMIMO and the use-and-forget bounding technique, we obtain an achievable rate in closed-form that only depends on statistical channel state information (CSI). To optimize the amplitudes and phase shifts of the STAR-RIS, we employ a projected gradient ascent method (PGAM) that simultaneously adjusts the amplitudes and phase shifts for both energy splitting (ES) and mode switching (MS) STAR-RIS operation protocols. By considering large-scale fading, the proposed optimization can be performed every several coherence intervals, which can significantly reduce overhead. Considering that STAR-RIS has twice the number of controllable parameters compared to conventional reflecting-only RIS, this accomplishment offers substantial practical benefits. Simulations are carried out to verify the analytical results, reveal the interplay of the achievable rate with fundamental parameters, and show the superiority of STAR-RIS regarding its achievable rate compared to its reflecting-only counterpart.

Index Terms—Simultaneously transmitting and reflecting RIS, correlated Rayleigh fading, imperfect CSI, achievable rate, 6G networks.

I. INTRODUCTION

Reconfigurable intelligent surfaces (RIS) have emerged as a promising technology to meet the requirements of sixth-generation (6G) networks such as a 1000-fold capacity increase together with increased connectivity among billions of devices [1]–[3]. A RIS consists of a metamaterial layer of low-cost controllable elements. Among its significant benefits is that its control signals can be dynamically adjusted to steer the impinging waves in specific directions and shape the propagation environment while providing uninterrupted service not only with low hardware cost, but also with low power consumption due to the absence of any power amplifiers.

Most of the existing works on RIS have assumed that both the transmitter and the receiver are found on the same side of the surface, i.e., only reflection takes place [1], [2], [4]–[7]. However, practical applications might include user equipments (UEs) on both sides of the RIS, which contain the spaces in front and behind the surface. Recently, advancements in programmable metamaterials have enabled the technology of simultaneously transmitting and reflecting RIS (STAR-RIS).[1] Hence, STAR-RIS has been proposed as a technology to satisfy this demand, i.e., it provides full space coverage by changing the amplitudes and phases of the impinging waves [8]–[12]. For instance, in [8], the authors provided a general hardware model and two-channel models corresponding to the near-field region and the far-field region of STAR-RIS with only two UEs. Also, they showed that the coverage and diversity gain are greater than reflecting-only/conventional RIS-assisted systems. Furthermore, in [9], three operating protocols for adjusting the transmission and reflection coefficients of the transmitted and reflected signals were suggested, namely, energy splitting (ES), mode switching (MS), and time switching (TS).

In particular, most existing works on RIS-aided systems have assumed perfect CSI, but this is a highly unrealistic assumption since practical systems have imperfect CSI. The accuracy of the channel state information (CSI) at the transmitter side is crucial to achieving a high beamforming gain of RIS, which includes channels between the transmitter and the UEs [13]. However, the acquisition of CSI is challenging because of the following reasons. First, RIS, in general, consists of passive elements to perform the desired reflecting operation, which makes any active transmission or reception infeasible, i.e., it cannot perform any sampling or processing of the pilots [1]. For this reason, an alternative method is the estimation of the aggregated transmitter-RIS-receiver channel by sending appropriate pilot symbols [14]. Second, RIS are generally large and consist of a large number of elements, and thus, induce high training overhead for channel estimation (CE), which results in spectral efficiency (SE) reduction [15].

Various CE schemes have been proposed to address this issue [16]–[20]. For example, in [17], an ON/OFF CE method was proposed, where the estimates of all RIS-assisted channels for a single-user MISO system are obtained one-by-one. Note that in the case of multi-user systems, this model was extended, assuming all RIS elements to be active during training, but the number of sub-phases is required to be at least equal to the number of RIS elements [20]. Although that method provides better CE as the number of sub-phases increases,

---

[1]We note that the word "transmitting" does not correspond to active transmission but implies coverage of the UEs at the other side of the RIS.

the achievable rate decreases because the data transmission phase takes a smaller fraction of the coherence time due to excessive training overhead. Also, that method computes the estimates of the channels of the individual RIS elements but the covariance of the channel vector from all RIS elements to a specific UE is unknown. Especially, in the case of STAR-RIS, CE becomes more challenging because UEs are located in both transmission and reflection regions, which requires different passive beamforming matrices (PBMs). In [11], a CE scheme was presented but did not account for multiple antennas at the BS, multiple UEs, and correlated fading.

In parallel, many early works on conventional RIS assumed independent Rayleigh fading such as [1], but recently, it was shown that RIS correlation should be considered because it is unavoidable in practical systems [4]. To this end, several works on conventional RIS have taken into account the impact of RIS correlation [5], [6], but only [21] has considered fading correlation on a STAR-RIS assisted system. Furthermore, except [9], [10], [12], all other works have assumed a single-antenna transmitter. Also, all previous studies on STAR-RIS have only considered a single UE on each side of the STAR-RIS. In this paper, we consider a more general case where multiple UEs are present on each side of the STAR-RIS.

Contributions: The observations above indicate the topic of this work, which concerns the study and design of a STAR-RIS assisted mMIMO system under the realistic conditions of imperfect CSI and correlated fading. These realistic assumptions and the consideration of multiple UEs at each side of the STAR-RIS make it extremely difficult for the derivations of the achievable rate and the resulting optimization of the amplitudes and phase shifts of the STAR-RIS. Our main contributions are summarized as follows:

- Aiming to characterize the potentials of STAR-RIS under realistic assumptions, we include the effect of spatially correlated fading at both the BS and the STAR-RIS.[2] In particular, we consider a massive multiple-input multiple-output (mMIMO) system with a BS having a large but finite number of antennas. Under this general setup, we derive the downlink achievable spectral efficiency (SE) of a STAR-RIS-assisted mMIMO system with imperfect CSI and correlated fading in closed form that depends only on large-scale statistics, which has not been known previously. Moreover, we achieve this by a unified analysis of the channel estimation and data transmission phases for UEs located in either the t or r regions, which distinguishes our work from previous research.
- Contrary to [8]–[12] which have assumed a single UE at each side of the surface, we consider multiple UEs

at each side of the RIS, which are served in the same time-frequency response.
- We apply the linear minimum mean square error (LMMSE) method to perform CE, and obtain closed-form expressions with lower overhead than other CE methods suggested for RIS-assisted systems. Specifically, we demonstrate that LMMSE can be applied without the need for a tailored design for STAR-RIS under conditions of statistical CSI. Note that previous works do not provide analytical expressions and/or do not take into account the spatial correlation at the RIS [8]–[12].
- Our analysis relies on statistical CSI, meaning that our closed-form expressions are dependent only on large-scale fading that changes at every several coherence intervals. Thus, the proposed optimization of the STAR-RIS can take place at every several coherence intervals, which saves significant overhead. On the contrary, previous studies, which are based on instantaneous CSI changing at each coherence interval, might not be feasible in practice due to inherent large overheads.[3]
- We formulate the problem of finding the amplitudes and phase shifts of the STAR-RIS to maximize the achievable sum SE. Our optimization framework considers multiple users at each side of the STAR-RIS in a unified manner. Despite its non-convexity, we derive an iterative efficient method based on the projected gradient ascent method in which both amplitudes and phase shifts of the STAR-RIS are updated simultaneously at each iteration. To the best of our knowledge, we are the first to optimize simultaneously the amplitudes and the phase shifts of the PBM in a STAR-RIS system. This is a significant contribution since other works optimize only the phase shifts or optimize both the amplitudes and the phase shifts in an alternating optimization manner. Moreover, as large-scale fading is considered, our optimization has very lower overhead in terms of complexity, training, and feedback compared to other works which rely on instantaneous CSI such as [9]. Notably, this property is important for STAR-RIS applications, which have twice the number of optimization variables compared to reflecting-only RIS. We also remark that the beamforming optimization based on statistical CSI for STAR-RIS has not been investigated previously.
- Simulations and analytical results are provided to shed light on the impact of various parameters and to show the superiority of STAR-RIS over conventional RIS. For example, we find that the system performance decreases as the RIS correlation increases.

Paper Outline: The remainder of this paper is organized as follows. Section II presents the system model of a STAR-RIS-assisted mMIMO system with correlated Rayleigh fading. Section III provides the CE. Section IV presents the downlink

---

[2]In the case of the active beamforming, being MRT in this work, it is designed based on the instantaneous channel, which depends on the correlation of the aggregated channel described by (6). In the case of the passive beamforming, it is designed based on statistical CSI in terms of path loss and correlation. Specifically, the sum-rate expression depends only on these large scale statistics, which vary every several coherence intervals. Hence, passive beamforming can be optimized at every several coherence intervals. Moreover, given that we rely on the statistical CSI approach, if no correlation is considered, the aggregated correlation will not depend on the phase shifts, which means that the sum rate cannot be optimized with respect to the phase shifts.

[3]In this work, we have followed the two-timescale transmission protocol approach as in [22], where a maximisation of the achievable sum rate of a RIS-assisted multi-user multi-input single-output (MU-MISO) system took place. According to this approach, the precoding is designed in terms of instantaneous CSI, while the RIS phase shifts is optimized by using statistical CSI. Notably, all works, which are based on statistical CSI, have relied on the two-timescale protocol. Examples are the study of the impact of hardware impairments on the sum rate and the minimum rate in [6] and [23], respectively.

data transmission with the derived downlink sum SE. Section V provides the simultaneous amplitudes and phase-shifts configuration concerning both the PBMs for the transmission and reflection regions. The numerical results are placed in Section VI, and Section VII concludes the paper.

Notation: Vectors and matrices are denoted by boldface lower and upper case symbols, respectively. The notations $(\cdot)^{\mathsf{T}}$, $(\cdot)^{\mathsf{H}}$, and $\mathrm{tr}(\cdot)$ describe the transpose, Hermitian transpose, and trace operators, respectively. Moreover, the notations $\arg(\cdot)$, $\mathbb{E}[\cdot]$, and $\mathrm{Var}(\cdot)$ express the argument function, the expectation, and variance operators, respectively. The notation $\mathrm{diag}(\mathbf{A})$ describes a vector with elements equal to the diagonal elements of $\mathbf{A}$, the notation $\mathrm{diag}(\mathbf{x})$ describes a diagonal matrix whose elements are $\mathbf{x}$, while $\mathbf{b} \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{\Sigma})$ describes a circularly symmetric complex Gaussian vector with zero mean and a covariance matrix $\boldsymbol{\Sigma}$.

II. SYSTEM MODEL

We consider a STAR-RIS-aided system, where a BS with an M-element uniform linear array (ULA) serves simultaneously K single-antenna UEs that are distributed on both sides of the STAR-RIS, as illustrated in Fig. 1. Specifically, $\mathcal{K}_t = \{1, \ldots, K_t\}$ UEs are located in the transmission region (t) and $\mathcal{K}_r = \{1, \ldots, K_r\}$ UEs are located in the reflection region (r), respectively, where $K_t + K_r = K$. Also, we denote by $\mathcal{W} = \{w_1, w_2, ..., w_K\}$ the set that defines the RIS operation mode for each of the K UEs. In particular, if the kth UE is located behind the STAR-RIS (i.e., $k \in \mathcal{K}_t$), then $w_k = t$, while $w_k = r$ when the kth UE is facing the STAR-RIS (i.e., $k \in \mathcal{K}_r$). Moreover, we assume direct links between the BS and UEs. The RIS consists of a uniform planar array (UPA) composed of N_{h} horizontally and N_{v} vertically passive elements, which belong to the set of $\mathcal{N} = \{1, \ldots, N\}$ elements, where $N = N_{\mathrm{h}} \times N_{\mathrm{v}}$ is the total number of RIS elements. The STAR-RIS is able to

Fig. 1: A mMIMO STAR-RIS assisted system with multiple UEs at transmission and reflection regions.

configure the transmitted (t) and reflected (r) signals by two independent coefficients. In particular, let $t_n = (\beta_n^t e^{j\phi_n^t})s_n$ and $r_n = (\beta_n^r e^{j\phi_n^r})s_n$ denote the transmitted and reflected signal

by the nth STAR-RIS element, respectively.[4] The amplitude and phase parameters $\beta_n^{w_k} \in [0, 1]$ and $\phi_n^{w_k} \in [0, 2\pi)$, where the kth UE can be in any of the two regions that corresponds also to the RIS mode, i.e. transmission (t) or reflection (r) [8], are independent. This model suggests that ϕ_n^t and ϕ_n^r can be chosen independently, but the choice of the amplitudes is based on the relationship expressed by the law of energy conservation as

$$(\beta_n^t)^2 + (\beta_n^r)^2 = 1, \forall n \in \mathcal{N}. \tag{1}$$

Henceforth, for the sake of exposition, we denote $\theta_i^{w_k} = e^{j\phi_i^{w_k}}$.

A. Operation Protocols

Our analysis is dedicated to the ES/MS protocols, which were presented in [9]. Herein, we outline them by providing their main points.

1) ES protocol: All RIS elements serve simultaneously all UEs in both t and r regions. Especially, the PBM for the kth UE is expressed as $\boldsymbol{\Phi}_{w_k}^{\mathrm{ES}} = \mathrm{diag}(\beta_1^{w_k}\theta_1^{w_k}, \ldots, \beta_N^{w_k}\theta_N^{w_k}) \in \mathbb{C}^{N \times N}$, where $\beta_n^{w_k} \geq 0$, $(\beta_n^t)^2 + (\beta_n^r)^2 = 1$, and $|\theta_n^{w_k}| = 1, \forall n \in \mathcal{N}$.

2) MS protocol: The RIS elements are partitioned into two groups of N_t and N_r elements that serve UEs in the t and r regions, respectively. In other words, $N_t + N_r = N$. The PBM for $k \in \mathcal{K}_t$ or $k \in \mathcal{K}_r$ is given by $\boldsymbol{\Phi}_{w_k}^{\mathrm{MS}} = \mathrm{diag}(\beta_1^{w_k}\theta_1^{w_k}, \ldots, \beta_N^{w_k}\theta_N^{w_k}) \in \mathbb{C}^{N \times N}$, where $\beta_n^{w_k} \in \{0, 1\}$, $(\beta_n^t)^2 + (\beta_n^r)^2 = 1$, and $|\theta_i^{w_k}| = 1, \forall n \in \mathcal{N}$. As can be seen, this protocol is a special case of the ES protocol, where the amplitude coefficients for transmission and reflection are restricted to binary values. As a result, the MS protocol is inferior of the ES counterpart since it cannot achieve the full-dimension transmission and reflection beamforming gain. Despite this performance degradation, it brings the advantage of lower computational complexity regarding the PBM design.
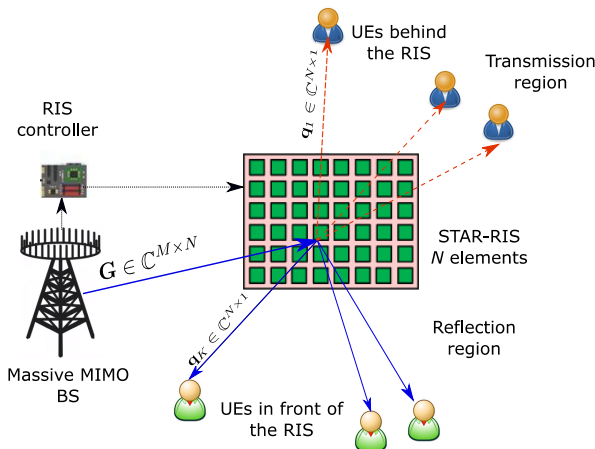
B. Channel Model

We assume narrowband quasi-static block fading channels with each block having a duration of τ_{c} channel uses. We adopt the standard time-division-duplex (TDD) protocol, which is preferable in mMIMO systems. Within TDD, we assume that each block includes τ channel uses for the uplink training phase and $\tau_{\mathrm{c}} - \tau$ channel uses for the downlink data transmission phase. Notably, contrary to other works, we aim to achieve a unified analysis regarding the channel estimation and data transmission phase that applies to a UE found in any of the t or r regions.

Let $\mathbf{G} = [\mathbf{g}_1 \ldots, \mathbf{g}_N] \in \mathbb{C}^{M \times N}$ be the channel between the BS and the STAR-RIS with $\mathbf{g}_i \in \mathbb{C}^{M \times 1}$ for $i \in \mathcal{N}$. Also, $\mathbf{q}_k \in \mathbb{C}^{N \times 1}$ denotes the channel between the STAR-RIS and UE k that can be found on either side. The direct link between the BS and UE k is denoted as $\mathbf{d}_k$. On this ground, we assume

---

[4]Note that here, we use $\beta_i^{w_k}$, instead of $\sqrt{\beta_i^{w_k}}$ as in [8], to denote the amplitude of the ith RIS element in mode w_k. The reason for this change will become clear when we present our proposed algorithm in Section V.

that all links are subject to correlated Rayleigh fading, which is normally the case in practice [4].[5] In particular, we have

$$\mathbf{G} = \sqrt{\tilde{\beta}_g} \mathbf{R}_{\mathrm{BS}}^{1/2} \mathbf{D} \mathbf{R}_{\mathrm{RIS}}^{1/2}, \tag{2}$$

$$\mathbf{q}_k = \sqrt{\tilde{\beta}_k} \mathbf{R}_{\mathrm{RIS}}^{1/2} \mathbf{c}_k, \tag{3}$$

$$\mathbf{d}_k = \sqrt{\bar{\beta}_k} \mathbf{R}_{\mathrm{BS}}^{1/2} \bar{\mathbf{c}}_k, \tag{4}$$

where $\mathbf{R}_{\mathrm{BS}} \in \mathbb{C}^{M \times M}$ and $\mathbf{R}_{\mathrm{RIS}} \in \mathbb{C}^{N \times N}$, assumed to be known by the network, express the deterministic Hermitian-symmetric positive semi-definite correlation matrices at the BS and the RIS respectively.[6] Regarding $\mathbf{R}_{\mathrm{BS}}$, it can be modeled e.g., as in [24], and $\mathbf{R}_{\mathrm{RIS}}$ is modeled as in [4]. Moreover, $\tilde{\beta}_g$, $\bar{\beta}_k$, and $\tilde{\beta}_k$ express the path-losses of the BS-RIS, BS-UE k, and RIS-UE k links in t or r region, respectively. Also, $\mathrm{vec}(\mathbf{D}) \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{MN})$, $\mathbf{c}_k \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_N)$, and $\bar{\mathbf{c}}_k \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_N)$ express the corresponding fast-fading components.

We note that the correlation matrices $\mathbf{R}_{\mathrm{RIS}}$ and $\mathbf{R}_{\mathrm{BS}}$ can be assumed to be known by the network since they can be obtained by existing estimation methods [25], [26]. Alternatively, we can practically calculate the covariance matrices for both $\mathbf{R}_{\mathrm{RIS}}$ and $\mathbf{R}_{\mathrm{BS}}$, despite the fact that $\mathbf{R}_{\mathrm{RIS}}$ is passive. Especially, the expressions for these covariance matrices depend on the distances between the RIS elements and the BS antennas, respectively, as well as the angles between them. The distances are known from the construction of the RIS and the BS, and the angles can be calculated when the locations are given. Hence, the covariance matrices can be considered to be known.

Given the PBM, the aggregated channel vector for UE k $\mathbf{h}_k = \mathbf{d}_k + \mathbf{G} \mathbf{\Phi}_{w_k} \mathbf{q}_k$ has a covariance matrix $\mathbf{R}_k = \mathbb{E}\{\mathbf{h}_k \mathbf{h}_k^{\mathsf{H}}\}$ given by

$$\mathbf{R}_k = \bar{\beta}_k \mathbf{R}_{\mathrm{BS}} + \hat{\beta}_k \, \mathrm{tr}\big(\mathbf{R}_{\mathrm{RIS}} \mathbf{\Phi}_{w_k} \mathbf{R}_{\mathrm{RIS}} \mathbf{\Phi}_{w_k}^{\mathsf{H}}\big) \mathbf{R}_{\mathrm{BS}}, \tag{5}$$

where we have used the independence between $\mathbf{G}$ and $\mathbf{q}_k$, $\hat{\beta}_k = \tilde{\beta}_g \tilde{\beta}_k$, $\mathbb{E}\{\mathbf{q}_k \mathbf{q}_k^{\mathsf{H}}\} = \tilde{\beta}_k \mathbf{R}_{\mathrm{RIS}}$, and $\mathbb{E}\{\mathbf{V}\mathbf{U}\mathbf{V}^{\mathsf{H}}\} = \mathrm{tr}(\mathbf{U})\mathbf{I}_M$ with $\mathbf{U}$ being a deterministic square matrix, and $\mathbf{V}$ being any matrix with independent and identically distributed (i.i.d.) entries of zero mean and unit variance. Notably, when $\mathbf{R}_{\mathrm{RIS}} = \mathbf{I}_N$, $\mathbf{R}_k$ does not depend on the phase shifts but only on the amplitudes, as also observed in [7].

Remark 1: As shown in (5), when independent Rayleigh fading is assumed, i.e., $\mathbf{R}_{\mathrm{RIS}} = \mathbf{I}_N$ and $\mathbf{R}_{\mathrm{BS}} = \mathbf{I}_M$, the covariance matrix of the aggregated channel becomes $\mathbf{R}_k = (\bar{\beta}_k + \hat{\beta}_k \sum_{i=1}^{N} (\beta_i^{w_k})^2) \mathbf{I}_M$, which is independent of the phase shifts. This reduces significantly the capability of the RIS in forming narrow beams and thus the performance is degraded accordingly. Therefore, it is not possible to optimize the achievable rate with respect to the phase shifts under

---

[5]The consideration of correlated Rician fading, which includes an LoS component, is the topic of future work.

[6]Many previous works have assumed that the channel between the BS and the RIS is deterministic expressing a line-of-sight (LoS) component [6], [20], while the analysis here is more general since we assume that all links are correlated Rayleigh fading distributed. In particular, $\mathbf{G}$, as expressed in (2) is based on the Kronecker channel model.

---

independent Rayleigh fading conditions.[7] However, in practice, correlated fading is unavoidable, which enables the optimization of the surface in terms of the phase shifts.

III. CHANNEL ESTIMATION

In practical systems, perfect CSI cannot be obtained. Especially, in mMIMO systems, the TDD protocol is adopted and channels are estimated by an uplink training phase with pilot symbols [27]. However, a RIS, being implemented by nearly passive elements without any RF chains, cannot process the estimated channels and obtain the received pilots by UEs. Also, it cannot transmit any pilot sequences to the BS for channel estimation.

In general, there are two approaches to channel estimation for RIS-aided communication systems, one focusing on the estimation of the individual channels such as [11], [13], [14], and the other obtaining the estimated aggregated channel such as [6], [20], [28]. The first benefit of the latter approach is that its implementation does not require any extra hardware and power cost. Also, the estimated aggregated BS-RIS-user channel is sufficient for the transmission beamforming design for the RIS-related links. It is easy to see that the BS-RIS channel has a large dimension in the considered system, which results in a prohibitively high pilot overhead if individual channels need to be estimated. This issue motivates us to employ the second approach in this paper, which has lower overhead and allows estimated channels to be expressed in closed form. We will now provide the details of the adopted channel estimation method.

We assume that all UEs either in t or r region send orthogonal pilot sequences. Specifically, we denote by $\mathbf{x}_k = [x_{k,1}, \ldots, x_{k,\tau}]^{\mathsf{H}} \in \mathbb{C}^{\tau \times 1}$ the pilot sequence of UE k that can be found in any of the two regions since the duration of the uplink training phase is τ channel uses. Note that $\mathbf{x}_k^{\mathsf{H}} \mathbf{x}_l = 0 \ \forall k \neq l$ and $\mathbf{x}_k^{\mathsf{H}} \mathbf{x}_k = \tau P$ joules with $P = |x_{k,i}|^2$, $\forall k, i$, i.e., it is assumed that all UEs use the same average transmit power during the training phase.

The received signal by the BS for the whole uplink training period is written as

$$\mathbf{Y}^{\mathrm{tr}} = \sum_{i=1}^{K} \mathbf{h}_i \mathbf{x}_i^{\mathsf{H}} + \mathbf{Z}^{\mathrm{tr}}, \tag{6}$$

where $\mathbf{Z}^{\mathrm{tr}} \in \mathbb{C}^{M \times \tau}$ is the received AWGN matrix having independent columns with each one distributed as $\mathcal{CN}(\mathbf{0}, \sigma^2 \mathbf{I}_M)$. Obviously, in (6), there is a contribution from UEs of both regions.

Multiplication of (6) with the transmit training sequence from UE k removes the interference by other UEs that can be found in the same or in the opposite region, and gives

$$\mathbf{r}_k = \mathbf{h}_k + \frac{\mathbf{z}_k}{\tau P}, \tag{7}$$

where $\mathbf{z}_k = \mathbf{Z}^{\mathrm{tr}} \mathbf{x}_k$.

---

[7]It is important to note that the recent works based on statistical CSI, such as [5]–[7], [22], [23], have shown a similar observation, i.e., in the case of no RIS correlation, the covariance matrix of the aggregated channel $\mathbf{R}_k$ does not depend on the phase shifts.

Lemma 1: The LMMSE estimate of the aggregated channel $\mathbf{h}_k$ between the BS and UE k is given by

$$\hat{\mathbf{h}}_k = \mathbf{R}_k \mathbf{Q}_k \mathbf{r}_k, \tag{8}$$

where $\mathbf{Q}_k = \left(\mathbf{R}_k + \frac{\sigma^2}{\tau P} \mathbf{I}_M \right)^{-1}$, and $\mathbf{r}_k$ is the noisy channel given by (7).

Proof: Please see Appendix A. ∎

The property of the orthogonality of LMMSE estimation gives the overall perfect channel in terms of the estimated channel $\hat{\mathbf{h}}_k$ and estimation channel error vectors $\tilde{\mathbf{h}}_k$ as

$$\mathbf{h}_k = \hat{\mathbf{h}}_k + \tilde{\mathbf{h}}_k. \tag{9}$$

Both $\hat{\mathbf{h}}_k$ and $\tilde{\mathbf{h}}_k$ have zero mean, and have variances (cf. (40))

$$\boldsymbol{\Psi}_k = \mathbf{R}_k \mathbf{Q}_k \mathbf{R}_k, \tag{10}$$

$$\tilde{\boldsymbol{\Psi}}_k = \mathbf{R}_k - \boldsymbol{\Psi}_k, \tag{11}$$

respectively. Given that $\mathbf{h}_k$ is not Gaussian, $\hat{\mathbf{h}}_k$ and $\tilde{\mathbf{h}}_k$ are not independent, but they are uncorrelated and each of them has zero mean [27].

It is clear from the above derivations that we indeed follow a conventional channel estimation method from standard mMIMO systems, where only the aggregated instantaneous BS-UE channel $\mathbf{h}_k$ is estimated. In this way, the minimum pilot sequence length is K, which is independent of the dimensions M and N. We note that if individual channels need to be estimated, the required complexity increases with M and N, which are large in the considered system. Thus, the presented method of estimating aggregated channels offers significant overhead reduction. It is important to note that this channel estimation method is sufficient for the design of the precoder at the BS in statistical CSI-based approaches.

We also remark that the method of estimating aggregated channels presented above has not been previously applied to STAR-RIS-aided systems based on the two-timescale method, which adds to the novelty of our paper. Specifically, we have demonstrated that the same expression for estimated channels can be used for users in either t or r regions, which has not been reported in previous papers studying STAR-RIS. Note also that the proposed two-timescale transmission approach has a channel estimation phase that does not depend on N, and thus, it is applicable to both ES and MS protocols. However, the ES protocol has twice the number of optimizable variables and thus requires higher complexity compared to the MS protocol. An advantage of the two-timescale approach is that the surface needs to be redesigned only when the statistical CSI changes. In contrast, instantaneous CSI-based schemes require beamforming calculations and information feedback in every channel coherence interval, leading to high computational complexity, power consumption, and feedback overhead. For such schemes, the ES protocol is not practically appealing, and thus, our proposed two-timescale approach is certainly more viable.

Remark 2: Our analysis presented above relies on large-scale statistics for a given PBM, which is obtained at every several coherence intervals. Thus, the optimization of the PBM that will be studied in the sequel is more practically appealing. Note

that our method provides the estimated aggregated channel vector in closed-form. Other methods in the RIS literature such as [18] do not result in analytical expressions, and do not capture the correlation effect since they obtain the estimated channel per RIS element [20]. Moreover, in the case of STAR-RIS, the only work on channel estimation is [11] but it does not consider practical effects such as correlation and multiple antennas at the BS.

IV. DOWNLINK DATA TRANSMISSION

The downlink data transmission from the BS to UE k in t or r region relies on TDD, which exploits channel reciprocity, i.e., the downlink channel equals the Hermitian transpose of the uplink channel. Hence, the received signal by UE k is expressed as

$$r_k = \mathbf{h}_k^{\mathsf{H}} \mathbf{s} + z_k, \tag{12}$$

where $\mathbf{s} = \sqrt{\lambda} \sum_{i=1}^{K} \sqrt{p_i} \mathbf{f}_i l_i$ expresses the transmit signal vector by the BS, p_i is the power allocated to UE i, and λ is a constant which is found such that $\mathbb{E}[\mathbf{s}^{\mathsf{H}}\mathbf{s}] = \rho$, where ρ is the total average power budget. Also, $z_k \sim \mathcal{CN}(0, \sigma^2)$ is the additive white complex Gaussian noise at UE k. Moreover, $\mathbf{f}_i \in \mathbb{C}^{M \times 1}$ is the linear precoding vector and l_i is the corresponding data symbol with $\mathbb{E}\{|l_i|^2\} = 1$. In this paper we adopt equal power allocation among all UEs as usually happens in the mMIMO literature, i.e. $p_i = \rho/K$ [24]. Thus, λ is found to ensure $\mathbb{E}[\mathbf{s}^{\mathsf{H}}\mathbf{s}] = \rho$, which gives $\lambda = \frac{K}{\mathbb{E}\{\mathrm{tr}\,\mathbf{F}\mathbf{F}^{\mathsf{H}}\}}$, where $\mathbf{F} = [\mathbf{f}_1, \ldots, \mathbf{f}_K] \in \mathbb{C}^{M \times K}$.

According to the technique in [29] and by exploiting that UEs do not have instantaneous CSI but are aware of only statistical CSI, the received signal by UE k can be written as

$$r_k = \frac{\sqrt{\lambda \rho}}{K} (\mathbb{E}\{\mathbf{h}_k^{\mathsf{H}} \mathbf{f}_k\} l_k + \mathbf{h}_k^{\mathsf{H}} \mathbf{f}_k l_k - \mathbb{E}\{\mathbf{h}_k^{\mathsf{H}} \mathbf{f}_k\} l_k$$
$$+ \sum_{i \neq k}^{K} \mathbf{h}_k^{\mathsf{H}} \mathbf{f}_i l_i Z + z_k. \tag{13}$$

Now, by using the use-and-then-forget bound [27], which relies on the common assumption of the worst-case uncorrelated additive noise for the inter-user interference, we obtain a lower bound on downlink average SE in bps/Hz. We remark that this lower bound is tight for mMIMO as the number of antennas is very large. Specifically, the achievable sum SE is given by

$$\mathrm{SE} = \frac{\tau_{\mathrm{c}} - \tau}{\tau_{\mathrm{c}}} \sum_{k=1}^{K} \log_2 (1 + \gamma_k), \tag{14}$$

where γ_k is the downlink signal-to-interference-plus-noise ratio (SINR), and the pre-log fraction corresponds to the percentage of samples per coherence block for downlink data transmission. Note that according to the use-and-forget bounding technique, the downlink SINR is given by

$$\gamma_k = \frac{S_k}{I_k}, \tag{15}$$

where

$$S_k = |\mathbb{E}\{\mathbf{h}_k^{\text{H}}\mathbf{f}_k\}|^2 \tag{16}$$

$$I_k = \mathbb{E}\big\{\big|\mathbf{h}_k^{\text{H}}\hat{\mathbf{h}}_k - \mathbb{E}\{\mathbf{h}_k^{\text{H}}\hat{\mathbf{h}}_k\}\big|^2\big\} + \sum_{i \neq k}^{K} |\mathbb{E}\{\mathbf{h}_k^{\text{H}}\mathbf{f}_i\}|^2 + \frac{K\sigma^2}{\rho\lambda}. \tag{17}$$

It is obvious that the final expressions for S_k and I_k depend on the choice of the precoder and the derivation of the SINR. In this regard, we note that maximum ratio transmission (MRT) and regularized zero-forcing (RZF) precoders are common options in the mMIMO literature. Herein, we select MRT for the sake of simplicity while RZF will be investigated in a future work. Since aggregated channels between the BS and users involve the indirect link through the STAR-RIS, it is challenging to evaluate (16) and (17) in closed-form. In the following proposition we present tight approximations of (16) and (17), which are then used to optimize the phase shifts.

Proposition 1: Let $\mathbf{f}_k = \hat{\mathbf{h}}_k$, i.e., MRT precoding being used, then a tight approximation of the downlink achievable SINR of UE k for a given PBM Φ_{w_k} in a STAR-RIS assisted mMIMO system, accounting for imperfect CSI, is given by

$$\gamma_k \approx \frac{S_k}{\tilde{I}_k}, \tag{18}$$

where

$$S_k = \text{tr}^2\left(\boldsymbol{\Psi}_k\right), \tag{19}$$

$$\tilde{I}_k = \sum_{i=1}^{K} \text{tr}(\mathbf{R}_k \boldsymbol{\Psi}_i) - \text{tr}\left(\boldsymbol{\Psi}_k^2\right) + \frac{K\sigma^2}{\rho} \sum_{i=1}^{K} \text{tr}(\boldsymbol{\Psi}_i). \tag{20}$$

Proof: Please see Appendix B. ∎

V. SIMULTANEOUS AMPLITUDES AND PHASE SHIFTS CONFIGURATION

It is critical to find the PBM to optimize a performance measure of STAR-RIS assisted systems. In this paper, assuming infinite-resolution phase shifters, we formulate the optimization problem for maximizing the sum SE with imperfect CSI and correlated fading. As mentioned in the preceding section, there are two operation protocols: ES protocol and MS protocol. In the following two subsections we deal with these two protocols.

A. Optimization of Amplitudes and Phase Shifts for ES protocol

For the ES protocol the formulated problem reads

$$\max_{\boldsymbol{\theta},\boldsymbol{\beta}} \quad f(\boldsymbol{\theta},\boldsymbol{\beta}) \triangleq \sum_{k=1}^{K} \log_2(1 + \frac{S_k}{\tilde{I}_k})$$
$$\text{s.t} \quad (\beta_n^t)^2 + (\beta_n^r)^2 = 1, \forall n \in \mathcal{N} \tag{P1}$$
$$\beta_n^t \geq 0, \beta_n^r \geq 0, \ \forall n \in \mathcal{N}$$
$$|\theta_n^t| = |\theta_n^r| = 1, \ \forall n \in \mathcal{N}$$

where $\boldsymbol{\theta} = [(\boldsymbol{\theta}^t)^{\text{T}}, (\boldsymbol{\theta}^r)^{\text{T}}]^{\text{T}}$ and $\boldsymbol{\beta} = [(\boldsymbol{\beta}^t)^{\text{T}}, (\boldsymbol{\beta}^r)^{\text{T}}]^{\text{T}}$. Note that to achieve a compact description we have vertically stacked $\boldsymbol{\theta}^t$ and $\boldsymbol{\theta}^r$ into a single vector $\boldsymbol{\theta}$, and $\boldsymbol{\beta}^t$ and $\boldsymbol{\beta}^r$ into a single vector $\boldsymbol{\beta}$, respectively. Also note that in the above problem formulation we have used the tight approximation of the SINR given in Proposition 1 to maximize the approximate sum SE,

denoted by $f(\boldsymbol{\theta},\boldsymbol{\beta})$. For ease of exposition, we define two sets: $\Theta = \{\boldsymbol{\theta} \mid |\theta_i^t| = |\theta_i^r| = 1, i = 1, 2, \ldots N\}$, and $\mathcal{B} = \{\boldsymbol{\beta} \mid (\beta_i^t)^2 + (\beta_i^r)^2 = 1, \beta_i^t \geq 0, \beta_i^r \geq 0, i = 1, 2, \ldots N\}$, which in fact together describe the feasible set of ($\mathcal{P}$1). Notably, the introduction of STAR-RIS imposes new challenges. In particular, the first constraint is not simple but includes the two types of passive beamforming, namely transmission and reflection beamforming, to be optimized, which are coupled with each other due to the energy conservation law.

The problem ($\mathcal{P}$1) is non-convex and is coupled among the optimization variables, which are the amplitudes and the phase shifts for transmission and reflection. For the development of an efficient algorithm to solve ($\mathcal{P}$1) we remark that the sets Θ and $\mathcal{B}$ are simple in the sense that their projection operators can be done in closed-form. This motivates us to apply the projected gradient ascent method (PGAM) [30, Ch. 2] to optimize $\boldsymbol{\theta}$ and $\boldsymbol{\beta}$, which is described next. However, in the case of independent Rayleigh fading, SE does not depend on $\boldsymbol{\theta}$, which means that optimization can take place only with respect to $\boldsymbol{\beta}$.

The proposed PGAM consists of the following iterations

$$\boldsymbol{\theta}^{n+1} = P_\Theta(\boldsymbol{\theta}^n + \mu_n \nabla_{\boldsymbol{\theta}} f(\boldsymbol{\theta}^n, \boldsymbol{\beta}^n)), \tag{21a}$$
$$\boldsymbol{\beta}^{n+1} = P_\mathcal{B}(\boldsymbol{\beta}^n + \mu_n \nabla_{\boldsymbol{\beta}} f(\boldsymbol{\theta}^n, \boldsymbol{\beta}^n)). \tag{21b}$$

In the above equations, the superscript denotes the iteration count. From the current iterate $(\boldsymbol{\theta}^n, \boldsymbol{\beta}^n)$ we move along the gradient direction to increase the objective. In (21), μ_n is the step size for both $\boldsymbol{\theta}$ and $\boldsymbol{\beta}$. Also, in (21), $P_\Theta(\cdot)$ and $P_\mathcal{B}(\cdot)$ are the projections onto Θ and $\mathcal{B}$, respectively.

The choice of the step size in (21a) and (21b) is important to make the proposed PGAM converge. The ideal step size should be inversely proportional to the Lipschitz constant of the corresponding gradient but this is difficult to find for the considered problem. For this reason, we apply the Armijo-Goldstein backtracking line search to find the step size at each iteration. To this end, we define a quadratic approximation of $f(\boldsymbol{\theta},\boldsymbol{\beta})$ as

$$Q_\mu(\boldsymbol{\theta},\boldsymbol{\beta};\mathbf{x},\mathbf{y}) = f(\boldsymbol{\theta},\boldsymbol{\beta}) + \langle \nabla_{\boldsymbol{\theta}} f(\boldsymbol{\theta},\boldsymbol{\beta}), \mathbf{x} - \boldsymbol{\theta} \rangle$$
$$- \frac{1}{\mu}\|\mathbf{x} - \boldsymbol{\theta}\|_2^2 + \langle \nabla_{\boldsymbol{\beta}} f(\boldsymbol{\theta},\boldsymbol{\beta}), \mathbf{y} - \boldsymbol{\beta} \rangle - \frac{1}{\mu}\|\mathbf{y} - \boldsymbol{\beta}\|_2^2. \tag{22}$$

Note that in this paper we define $\langle \mathbf{x}, \mathbf{y} \rangle = 2\,\text{Re}\{\mathbf{x}^{\text{H}}\mathbf{y}\}$ for complex-valued $\mathbf{x}$ and $\mathbf{y}$ and $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^{\text{T}}\mathbf{y}$ for non complex-valued $\mathbf{x}$ and $\mathbf{y}$. Now, let $L_n > 0$, and $\kappa \in (0, 1)$. Then the step size μ_n in (21) can be found as $\mu_n = L_n \kappa^{m_n}$, where m_n is the smallest nonnegative integer satisfying

$$f(\boldsymbol{\theta}^{n+1}, \boldsymbol{\beta}^{n+1}) \geq Q_{L_n\kappa^{m_n}}(\boldsymbol{\theta}^n, \boldsymbol{\beta}^n; \boldsymbol{\theta}^{n+1}, \boldsymbol{\beta}^{n+1}), \tag{23}$$

which can be done by an iterative procedure. In the proposed PGAM, we use the step size at iteration n as the initial step size at iteration $n + 1$. The proposed PGAM is summarized in Algorithm 1.

We present the complex-valued gradients in the following lemma.

Lemma 2: The complex gradients $\nabla_{\boldsymbol{\theta}} f(\boldsymbol{\theta},\boldsymbol{\beta})$ and $\nabla_{\boldsymbol{\beta}} f(\boldsymbol{\theta},\boldsymbol{\beta})$

Algorithm 1 Projected Gradient Ascent Method for the RIS Design

---

1: Input: $\boldsymbol{\theta}^0, \boldsymbol{\beta}^0, \mu_1 > 0, \kappa \in (0, 1)$
2: $n \leftarrow 1$
3: **repeat**
4:    **repeat**
5:       $\boldsymbol{\theta}^{n+1} = P_\Theta(\boldsymbol{\theta}^n + \mu_n \nabla_{\boldsymbol{\theta}} f(\boldsymbol{\theta}^n, \boldsymbol{\beta}^n))$
6:       $\boldsymbol{\beta}^{n+1} = P_B(\boldsymbol{\beta}^n + \mu_n \nabla_{\boldsymbol{\beta}} f(\boldsymbol{\theta}^n, \boldsymbol{\beta}^n))$
7:       **if** $f(\boldsymbol{\theta}^{n+1}, \boldsymbol{\beta}^{n+1}) \leq Q_{\mu_n}(\boldsymbol{\theta}^n, \boldsymbol{\beta}^n; \boldsymbol{\theta}^{n+1}, \boldsymbol{\beta}^{n+1})$
       then
8:          $\mu_n = \mu_n \kappa$
9:       **end if**
10:    **until** $f(\boldsymbol{\theta}^{n+1}, \boldsymbol{\beta}^{n+1}) > Q_{\mu_n}(\boldsymbol{\theta}^n, \boldsymbol{\beta}^n; \boldsymbol{\theta}^{n+1}, \boldsymbol{\beta}^{n+1})$
11:    $\mu_{n+1} \leftarrow \mu_n$
12:    $n \leftarrow n + 1$
13: **until**   convergence
14: Output: $\boldsymbol{\theta}^{n+1}, \boldsymbol{\beta}^{n+1}$

---

are given in closed-forms by

$$\nabla_{\boldsymbol{\theta}} f(\boldsymbol{\theta}, \boldsymbol{\beta}) = [\nabla_{\boldsymbol{\theta}^t} f(\boldsymbol{\theta}, \boldsymbol{\beta})^\top, \nabla_{\boldsymbol{\theta}^r} f(\boldsymbol{\theta}, \boldsymbol{\beta})^\top]^\top, \tag{24a}$$

$$\nabla_{\boldsymbol{\theta}^t} f(\boldsymbol{\theta}, \boldsymbol{\beta}) = \frac{\tau_c - \tau}{\tau_c \log 2} \sum_{k=1}^K \frac{\tilde{I}_k \nabla_{\boldsymbol{\theta}^t} S_k - S_k \nabla_{\boldsymbol{\theta}^t} \tilde{I}_k}{(1 + \gamma_k) \tilde{I}_k^2}, \tag{24b}$$

$$\nabla_{\boldsymbol{\theta}^r} f(\boldsymbol{\theta}, \boldsymbol{\beta}) = \frac{\tau_c - \tau}{\tau_c \log 2} \sum_{k=1}^K \frac{\tilde{I}_k \nabla_{\boldsymbol{\theta}^r} S_k - S_k \nabla_{\boldsymbol{\theta}^r} \tilde{I}_k}{(1 + \gamma_k) \tilde{I}_k^2}, \tag{24c}$$

where

$$\nabla_{\boldsymbol{\theta}^t} S_k = \begin{cases} \nu_k \mathrm{diag}(\mathbf{A}_t \mathrm{diag}(\boldsymbol{\beta}^t)) & w_k = t \\ 0 & w_k = r \end{cases} \tag{25a}$$

$$\nabla_{\boldsymbol{\theta}^r} S_k = \begin{cases} \nu_k \mathrm{diag}(\mathbf{A}_r \mathrm{diag}(\boldsymbol{\beta}^r)) & w_k = r \\ 0 & w_k = t \end{cases} \tag{25b}$$

$$\nabla_{\boldsymbol{\theta}^t} \tilde{I}_k = \mathrm{diag}(\tilde{\mathbf{A}}_{kt} \mathrm{diag}(\boldsymbol{\beta}^t)) \tag{25c}$$

$$\nabla_{\boldsymbol{\theta}^r} \tilde{I}_k = \mathrm{diag}(\tilde{\mathbf{A}}_{kr} \mathrm{diag}(\boldsymbol{\beta}^r)) \tag{25d}$$

with $\mathbf{A}_{w_k} = \mathbf{R}_{\mathrm{RIS}} \boldsymbol{\Phi}_{w_k} \mathbf{R}_{\mathrm{RIS}}$ for $w_k \in \{t, r\}$, $\nu_k = 2\hat{\beta}_k \mathrm{tr}(\boldsymbol{\Psi}_k) \mathrm{tr}((\mathbf{Q}_k \mathbf{R}_k + \mathbf{R}_k \mathbf{Q}_k - \mathbf{Q}_k \mathbf{R}_k \mathbf{Q}_k)\mathbf{R}_{\mathrm{BS}})$,

$$\tilde{\mathbf{A}}_{ku} = \begin{cases} (\bar{\nu}_k + \sum_{i \in \mathcal{K}_u}^K \tilde{\nu}_{ki}) \mathbf{A}_u & w_k = u \\ \sum_{i \in \mathcal{K}_u}^K \tilde{\nu}_{ki} \mathbf{A}_u & w_k \neq u, \end{cases} \tag{26}$$

$u \in \{t, r\}$, $\bar{\nu}_k = \hat{\beta}_k \mathrm{tr}(\check{\boldsymbol{\Psi}}_k \mathbf{R}_{\mathrm{BS}})$, $\tilde{\nu}_{ki} = \hat{\beta}_k \mathrm{tr}(\tilde{\mathbf{R}}_{ki} \mathbf{R}_{\mathrm{BS}})$, $\check{\boldsymbol{\Psi}}_k = \boldsymbol{\Psi} - 2(\mathbf{Q}_k \mathbf{R}_k \boldsymbol{\Psi}_k + \boldsymbol{\Psi}_k \mathbf{R}_k \mathbf{Q}_k - \mathbf{Q}_k \mathbf{R}_k \boldsymbol{\Psi}_k \mathbf{R}_k \mathbf{Q}_k)$, $\boldsymbol{\Psi} = \sum_{i=1}^K \boldsymbol{\Psi}_i$, $\tilde{\mathbf{R}}_{ki} = \mathbf{Q}_i \mathbf{R}_i \bar{\mathbf{R}}_k - \mathbf{Q}_i \mathbf{R}_i \bar{\mathbf{R}}_k \mathbf{R}_i \mathbf{Q}_i + \bar{\mathbf{R}}_k \mathbf{R}_i \mathbf{Q}_i$, and $\bar{\mathbf{R}}_k = \mathbf{R}_k + \frac{K\sigma^2}{\rho} \mathbf{I}_M$. Similarly, the gradient $\nabla_{\boldsymbol{\beta}} f(\boldsymbol{\theta}, \boldsymbol{\beta})$ is given by

$$\nabla_{\boldsymbol{\beta}} f(\boldsymbol{\theta}, \boldsymbol{\beta}) = [\nabla_{\boldsymbol{\beta}^t} f(\boldsymbol{\theta}, \boldsymbol{\beta})^\top, \nabla_{\boldsymbol{\beta}^r} f(\boldsymbol{\theta}, \boldsymbol{\beta})^\top]^\top, \tag{27a}$$

$$\nabla_{\boldsymbol{\beta}^t} f(\boldsymbol{\theta}, \boldsymbol{\beta}) = \frac{\tau_c - \tau}{\tau_c \log 2} \sum_{k=1}^K \frac{\tilde{I}_k \nabla_{\boldsymbol{\beta}^t} S_k - S_k \nabla_{\boldsymbol{\beta}^t} \tilde{I}_k}{(1 + \gamma_k) \tilde{I}_k^2}, \tag{27b}$$

$$\nabla_{\boldsymbol{\beta}^r} f(\boldsymbol{\theta}, \boldsymbol{\beta}) = \frac{\tau_c - \tau}{\tau_c \log 2} \sum_{k=1}^K \frac{\tilde{I}_k \nabla_{\boldsymbol{\beta}^r} S_k - S_k \nabla_{\boldsymbol{\beta}^r} \tilde{I}_k}{(1 + \gamma_k) \tilde{I}_k^2}, \tag{27c}$$

where

$$\nabla_{\boldsymbol{\beta}^t} S_k = \begin{cases} 2\nu_k \mathrm{Re}\{\mathrm{diag}(\mathbf{A}_k^{\mathtt{H}} \mathrm{diag}(\boldsymbol{\theta}^t))\} & w_k = t \\ 0 & w_k = r \end{cases} \tag{28a}$$

$$\nabla_{\boldsymbol{\beta}^r} S_k = \begin{cases} 2\nu_k \mathrm{Re}\{\mathrm{diag}(\mathbf{A}_k^{\mathtt{H}} \mathrm{diag}(\boldsymbol{\theta}^r))\} & w_k = r \\ 0 & w_k = t \end{cases} \tag{28b}$$

$$\nabla_{\boldsymbol{\beta}^t} \tilde{I}_k = 2 \mathrm{Re}\{\mathrm{diag}(\tilde{\mathbf{A}}_{kt}^{\mathtt{H}} \mathrm{diag}(\boldsymbol{\theta}^t))\} \tag{28c}$$

$$\nabla_{\boldsymbol{\beta}^r} \tilde{I}_k = 2 \mathrm{Re}\{\mathrm{diag}(\tilde{\mathbf{A}}_{kr}^{\mathtt{H}} \mathrm{diag}(\boldsymbol{\theta}^r))\}. \tag{28d}$$

Note that $\nabla_{\boldsymbol{\beta}} f(\boldsymbol{\theta}, \boldsymbol{\beta})$ is real-valued.

Proof: Please see Appendix C. ∎

Remark 3: As mentioned earlier we use $\beta_i^{w_k}$, instead of $\sqrt{\beta_i^{w_k}}$ as in [8], to denote the amplitude of the ith RIS element in mode w_k. The purpose of that maneuver is now clear. In fact, if $\sqrt{\beta_i^{w_k}}$ were used to represent the amplitude, then the gradient $\nabla_{\boldsymbol{\beta}} f(\boldsymbol{\theta}, \boldsymbol{\beta})$ would be similar to (27) but contain the term $\sqrt{\beta_i^{w_k}}$ in the denominator. This will make $\nabla_{\boldsymbol{\beta}} f(\boldsymbol{\theta}, \boldsymbol{\beta})$ ill-conditioned (i.e., extremely large), which in turn can cause numerical issues in the execution of Algorithm 1 in practice.

To conclude the description of Algorithm 1, we now provide the projection onto the sets Θ and $\mathcal{B}$. First, it is straightforward to check that, for a given $\boldsymbol{\theta} \in \mathbb{C}^{2N \times 1}$ $P_\Theta(\boldsymbol{\theta})$ is given by

$$P_\Theta(\boldsymbol{\theta}) = \boldsymbol{\theta}/|\boldsymbol{\theta}| = e^{j\angle \boldsymbol{\theta}}, \tag{29}$$

where the operations in the right-hand side of the above equation are performed entrywise.

The projection $P_\mathcal{B}(\boldsymbol{\beta})$ deserves special attention. Note that the constraint $(\beta_i^t)^2 + (\beta_i^r)^2 = 1, \beta_i^t \geq 0, \beta_i^r \geq 0$ indeed defines the first quadrant of the unit circle. Thus, the expression of the projection onto $\mathcal{B}$ is rather complicated. To make $P_\mathcal{B}(\boldsymbol{\beta})$ more efficient we allow β_i^t and β_i^r to take negative value during the iterative process. However, we remark that this step *does not* affect the optimality of the proposed solution since we can change the sign of both β_i^u and θ_i^u, $u \in t, r$ and still achieve the same objective. As a result, we can project β_i^t and β_i^r onto the entire unit circle, and thus we can write $P_\mathcal{B}(\boldsymbol{\beta})$ as

$$[P_\mathcal{B}(\boldsymbol{\beta})]_i = \frac{\beta_i}{\sqrt{\beta_i^2 + \beta_{i+N}^2}}, i = 1, 2, \ldots, N \tag{30a}$$

$$[P_\mathcal{B}(\boldsymbol{\beta})]_{i+N} = \frac{\beta_{i+N}}{\sqrt{\beta_i^2 + \beta_{i+N}^2}}, i = 1, 2, \ldots, N. \tag{30b}$$

Complexity Analysis of Algorithm 1: We now provide the complexity analysis for each iteration of Algorithm 1 in terms of the required number of complex multiplications using big-O notation, which is in particular relevant for large M and N as considered in this paper. We note Algorithm 1 only requires the first-order information, i.e., the objective and its gradient value. Let us analyze the complexity of computing the objective value. First, we need to compute $\mathbf{R}_k$ which can be written as $\mathbf{R}_k = \bar{\beta}_k \mathbf{R}_{\mathrm{BS}} + \hat{\beta}_k \mathrm{tr}(\mathbf{R}_{\mathrm{RIS}} \boldsymbol{\Phi}_{w_k} \mathbf{R}_{\mathrm{RIS}} \boldsymbol{\Phi}_{w_k}^{\mathtt{H}}) \mathbf{R}_{\mathrm{BS}} = \bar{\beta}_k \mathbf{R}_{\mathrm{BS}} + \hat{\beta}_k \mathrm{tr}(\mathbf{A}_{w_k} \boldsymbol{\Phi}_{w_k}^{\mathtt{H}}) \mathbf{R}_{\mathrm{BS}}$. Now it is obvious that we need to obtain the term $\mathrm{tr}(\mathbf{A}_{w_k} \boldsymbol{\Phi}_{w_k}^{\mathtt{H}})$. To this end, we note that since $\boldsymbol{\Phi}_{w_k}$ is diagonal, *only diagonal elements of $\mathbf{A}_{w_k}$, i.e., diag$(\mathbf{A}_{w_k})$, are required*. Next, computing $\mathbf{R}_{\mathrm{RIS}} \boldsymbol{\Phi}_{w_k}$ requires N^2 complex multiplications since $\boldsymbol{\Phi}_{w_k}$ is diagonal, and thus, diag$(\mathbf{A}_{w_k}) = $ diag$(\mathbf{R}_{\mathrm{RIS}} \boldsymbol{\Phi}_{w_k} \mathbf{R}_{\mathrm{RIS}})$ requires $O(N^2)$ complex multiplications.

As a result, the complexity to compute $\text{tr}\big(\mathbf{A}_{w_k}\boldsymbol{\Phi}_{w_k}^{\mathsf{H}}\big)$ is $O(N^2 + N)$, and thus, the complexity to obtain $\mathbf{R}_k$ is $O(N^2 + N + M^2)$ since $O(M^2)$ additional complexity multiplications are required to obtain $\text{tr}\big(\mathbf{A}_{w_k}\boldsymbol{\Phi}_{w_k}^{\mathsf{H}}\big)\mathbf{R}_{\text{BS}}$. Recall that $\boldsymbol{\Psi}_k = \mathbf{R}_k\mathbf{Q}_k\mathbf{R}_k = \mathbf{R}_k\big(\mathbf{R}_k + \frac{\sigma^2}{\tau P}\mathbf{I}_M\big)^{-1}\mathbf{R}_k$ and thus it would take $O(M^3)$ to compute it, which is due to the calculation of the involving matrix inversion. Here we present a more efficient way to compute $\boldsymbol{\Psi}_k$. Let $\mathbf{R}_{\text{BS}} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{U}^{\mathsf{H}}$, where $\boldsymbol{\Sigma}$ is diagonal and $\mathbf{U}$ is unitary, be the eigenvalue decomposition (EVD) of $\mathbf{R}_{\text{BS}}$ and $\alpha_k = \hat{\beta}_k\,\text{tr}\big(\mathbf{A}_{w_k}\boldsymbol{\Phi}_{w_k}^{\mathsf{H}}\big)$. We remark that *the EVD of $\boldsymbol{\Sigma}$ is only performed once* before Algorithm 1 is executed. Then, we can write

$$
\begin{aligned}
\mathbf{Q}_k &= \big(\mathbf{R}_k + \frac{\sigma^2}{\tau P}\mathbf{I}_M\big)^{-1} \\
&= \big(\alpha_k\mathbf{U}\boldsymbol{\Sigma}\mathbf{U}^{\mathsf{H}} + \frac{\sigma^2}{\tau P}\mathbf{I}_M\big)^{-1} \\
&= \mathbf{U}\big(\alpha_k\boldsymbol{\Sigma} + \frac{\sigma^2}{\tau P}\mathbf{I}_M\big)^{-1}\mathbf{U}^{\mathsf{H}},
\end{aligned} \tag{31}
$$

where we have used the fact that $\mathbf{R}_k = \alpha_k\mathbf{R}_{\text{BS}} = \alpha_k\mathbf{U}\boldsymbol{\Sigma}\mathbf{U}^{\mathsf{H}}$. Substituting (31) into (10), we immediately have

$$
\begin{aligned}
\boldsymbol{\Psi}_k &= \alpha_k^2\mathbf{U}\boldsymbol{\Sigma}\big(\alpha_k\boldsymbol{\Sigma} + \frac{\sigma^2}{\tau P}\mathbf{I}_M\big)^{-1}\boldsymbol{\Sigma}\mathbf{U}^{\mathsf{H}} \\
&= \mathbf{U}\boldsymbol{\Sigma}\big(\alpha_k^{-1}\boldsymbol{\Sigma} + \frac{\sigma^2}{\tau P\alpha_k^2}\mathbf{I}_M\big)^{-1}\boldsymbol{\Sigma}\mathbf{U}^{\mathsf{H}} \\
&= \mathbf{U}\boldsymbol{\Sigma}\bar{\boldsymbol{\Sigma}}_k\boldsymbol{\Sigma}\mathbf{U}^{\mathsf{H}},
\end{aligned} \tag{32}
$$

where $\bar{\boldsymbol{\Sigma}}_k = \big(\alpha_k^{-1}\boldsymbol{\Sigma} + \frac{\sigma^2}{\tau P\alpha_k^2}\mathbf{I}_M\big)^{-1}$. Note that $\bar{\boldsymbol{\Sigma}}_k$ is diagonal and takes $O(M)$ complex multiplications to compute. Now it is clear that $\text{tr}(\boldsymbol{\Psi}_k) = \text{tr}\big(\boldsymbol{\Sigma}\bar{\boldsymbol{\Sigma}}_k\boldsymbol{\Sigma}\big)$ requires $O(M)$ complex multiplications to compute, which is indeed the complexity to compute S_k in (19). To compute I_k we have

$$
\begin{aligned}
\boldsymbol{\Psi} &= \sum_{i=1}^{K}\boldsymbol{\Psi}_i \\
&= \mathbf{U}\boldsymbol{\Sigma}\Big(\sum_{i=1}^{K}\bar{\boldsymbol{\Sigma}}_i\Big)\boldsymbol{\Sigma}\mathbf{U}^{\mathsf{H}} \\
&= \mathbf{U}\boldsymbol{\Sigma}\bar{\boldsymbol{\Sigma}}\boldsymbol{\Sigma}\mathbf{U}^{\mathsf{H}},
\end{aligned} \tag{33}
$$

where $\bar{\boldsymbol{\Sigma}} = \sum_{i=1}^{K}\bar{\boldsymbol{\Sigma}}_i$. We note that obtaining $\bar{\boldsymbol{\Sigma}}$ once all $\bar{\boldsymbol{\Sigma}}_i$'s are known requires only KM complex additions, which is negligible. Thus, it follows that

$$
\begin{aligned}
\sum_{i=1}^{K}\text{tr}(\mathbf{R}_k\boldsymbol{\Psi}_i) &= \text{tr}(\mathbf{R}_k\boldsymbol{\Psi}) \\
&= \alpha_k\,\text{tr}\big(\mathbf{U}\boldsymbol{\Sigma}^2\bar{\boldsymbol{\Sigma}}\boldsymbol{\Sigma}\mathbf{U}^{\mathsf{H}}\big) \\
&= \alpha_k\,\text{tr}\big(\boldsymbol{\Sigma}^2\bar{\boldsymbol{\Sigma}}\boldsymbol{\Sigma}\big),
\end{aligned} \tag{34}
$$

and that

$$
\sum_{i=1}^{K}\text{tr}(\boldsymbol{\Psi}_i) = \text{tr}(\boldsymbol{\Psi}) = \text{tr}\big(\boldsymbol{\Sigma}\bar{\boldsymbol{\Sigma}}\boldsymbol{\Sigma}\big). \tag{35}
$$

Summarizing the above results, we can conclude that the complexity to compute $f(\boldsymbol{\theta},\boldsymbol{\beta})$ is $O(K(N^2 + M^2))$.

Next we present the complexity to compute $\nabla_{\boldsymbol{\theta}}f(\boldsymbol{\theta},\boldsymbol{\beta})$ and $\nabla_{\boldsymbol{\beta}}f(\boldsymbol{\theta},\boldsymbol{\beta})$. Recall that v_k in (25a) and (25b) is given by $v_k = 2\hat{\beta}_k\,\text{tr}(\boldsymbol{\Psi}_k)\,\text{tr}((\mathbf{Q}_k\mathbf{R}_k + \mathbf{R}_k\mathbf{Q}_k - \mathbf{Q}_k\mathbf{R}_k\mathbf{Q}_k)\mathbf{R}_{\text{BS}})$. Following the above analysis, we can write $\mathbf{Q}_k\mathbf{R}_k\mathbf{R}_{\text{BS}}$ as $\mathbf{Q}_k\mathbf{R}_k\mathbf{R}_{\text{BS}} = \mathbf{R}_k\mathbf{Q}_k = \alpha_k\mathbf{U}(\alpha_k\boldsymbol{\Sigma} + \frac{\sigma^2}{\tau P\alpha_k^2}\mathbf{I}_M)^{-1}\boldsymbol{\Sigma}^2\mathbf{U}^{\mathsf{H}}$

and $\mathbf{Q}_k\mathbf{R}_k\mathbf{Q}_k\mathbf{R}_{\text{BS}}$ as $\mathbf{Q}_k\mathbf{R}_k\mathbf{Q}_k)\mathbf{R}_{\text{BS}} = \alpha_k\mathbf{U}(\alpha_k\boldsymbol{\Sigma} + \frac{\sigma^2}{\tau P\alpha_k^2}\mathbf{I}_M)^{-2}\boldsymbol{\Sigma}^2\mathbf{U}^{\mathsf{H}}$. Thus, we can rewrite v_k equivalently as

$$
\begin{aligned}
v_k = {}& 2\hat{\beta}_k\alpha_k\,\text{tr}\,(\boldsymbol{\Psi}_k) \\
&\times \Big(2\,\text{tr}\big((\alpha_k\boldsymbol{\Sigma} + \frac{\sigma^2}{\tau P\alpha_k^2}\mathbf{I}_M)^{-1}\boldsymbol{\Sigma}^2\big) \\
&\quad - \text{tr}\big((\alpha_k\boldsymbol{\Sigma} + \frac{\sigma^2}{\tau P\alpha_k^2}\mathbf{I}_M)^{-2}\boldsymbol{\Sigma}^2\big)\Big),
\end{aligned} \tag{36}
$$

which requires $O(M)$ complex multiplications to obtain since $\text{tr}\,(\boldsymbol{\Psi}_k)$ is already computed and the involving matrices in the second term of the above equation are diagonal.

Next, to obtain $\nabla_{\boldsymbol{\theta}^t}S_k$ in (25a), we further need to calculate the diagonal elements of $\mathbf{A}_t\text{diag}(\boldsymbol{\beta}^t)$, which can be obtained by multiplying each diagonal element of $\mathbf{A}_t$ with the corresponding entry of $\boldsymbol{\beta}^t$, i.e., $\text{diag}\big(\mathbf{A}_t\text{diag}(\boldsymbol{\beta}^t)\big) = \text{diag}\big(\mathbf{A}_t\big)\odot\boldsymbol{\beta}^t$, where $\odot$ represents the entry-wise multiplication. We remark that the term $\text{diag}\big(\mathbf{A}_t\big)$ is already computed when calculating $\mathbf{R}_k$, and thus, the complexity to compute $\nabla_{\boldsymbol{\theta}^t}S_k$ is $O(N)$. Apparently, the same complexity is required to obtain $\nabla_{\boldsymbol{\theta}^r}S_k$ in (25b). The complexity of calculating $\nabla_{\boldsymbol{\theta}^u}I_k$, $u \in \{t,r\}$ follows similar lines. More specifically, both $\bar{\nu}_k$ and $\sum_{i\in\mathcal{K}_u}^{K}\tilde{\nu}_{ki}$ require $O(M^2)$ to compute. In summary, the complexity of computing the gradients for each iteration is $O(K(N^2 + M^2))$.

Convergence Analysis of Algorithm 1: The convergence of Algorithm 1 is guaranteed by following standard arguments for projected gradient methods. First, the gradients $\nabla_{\boldsymbol{\theta}}f(\boldsymbol{\theta},\boldsymbol{\beta})$ and $\nabla_{\boldsymbol{\beta}}f(\boldsymbol{\theta},\boldsymbol{\beta})$ are Lipschitz continuous[8] over the feasible set as they comprise basic functions as given above. Let $L_{\boldsymbol{\theta}}$ and $L_{\boldsymbol{\beta}}$ be the Lipschitz constant of $\nabla_{\boldsymbol{\theta}}f(\boldsymbol{\theta},\boldsymbol{\beta})$ and $\nabla_{\boldsymbol{\beta}}f(\boldsymbol{\theta},\boldsymbol{\beta})$, respectively. Then it holds that [30, Chapter 2]

$$
\begin{aligned}
f(\mathbf{x},\mathbf{y}) &\geq f(\boldsymbol{\theta},\boldsymbol{\beta}) + \langle\nabla_{\boldsymbol{\theta}}f(\boldsymbol{\theta},\boldsymbol{\beta}),\mathbf{x} - \boldsymbol{\theta}\rangle - \frac{1}{L_{\boldsymbol{\theta}}}\|\mathbf{x} - \boldsymbol{\theta}\|_2^2 \\
&\quad + \langle\nabla_{\boldsymbol{\beta}}f(\boldsymbol{\theta},\boldsymbol{\beta}),\mathbf{y} - \boldsymbol{\beta}\rangle - \frac{1}{L_{\boldsymbol{\beta}}}\|\mathbf{y} - \boldsymbol{\beta}\|_2^2 \\
&\geq f(\boldsymbol{\theta},\boldsymbol{\beta}) + \langle\nabla_{\boldsymbol{\theta}}f(\boldsymbol{\theta},\boldsymbol{\beta}),\mathbf{x} - \boldsymbol{\theta}\rangle - \frac{1}{L_{\max}}\|\mathbf{x} - \boldsymbol{\theta}\|_2^2 \\
&\quad + \langle\nabla_{\boldsymbol{\beta}}f(\boldsymbol{\theta},\boldsymbol{\beta}),\mathbf{y} - \boldsymbol{\beta}\rangle - \frac{1}{L_{\max}}\|\mathbf{y} - \boldsymbol{\beta}\|_2^2
\end{aligned}
$$

where $L_{\max} = \max(L_{\boldsymbol{\theta}}, L_{\boldsymbol{\beta}})$. Thus, the line search procedure of Algorithm 1 (i.e. the loop between Steps 4 – 10) terminates in finite iterations since the condition in Step 10 must be satisfied when $\mu_n < L_{\max}$. More specifically, given μ_{n-1}, the maximum number of steps in the line search procedure is $\left\lceil\frac{\log(L_{\max}\mu_{n-1})}{\log\kappa}\right\rceil$, where $\log()$ denotes the natural logarithm and $\lceil\cdot\rceil$ denotes the smallest integer that is larger than or equal to the argument. Also, due to the line search we automatically have an increasing sequence of objectives, i.e., $f(\boldsymbol{\theta}^{n+1},\boldsymbol{\beta}^{n+1}) \geq f(\boldsymbol{\theta}^n,\boldsymbol{\beta}^n)$. Since the feasible sets Θ and $\mathcal{B}$ are compact, $f(\boldsymbol{\theta}^n,\boldsymbol{\beta}^n)$ must converge. However, we remark that Algorithm 1 is only guaranteed to converge to a stationary point of $(\mathcal{P}1)$, which is not necessarily an optimal solution due

---

[8]A function $\mathbf{h}(\mathbf{x})$ is said to be Lipschitz continuous over the set D if there exists $L > 0$ such that $\|\mathbf{h}(\mathbf{x}) - \mathbf{h}(\mathbf{y})\| \leq L\|\mathbf{x} - \mathbf{y}\|_2$

to the nonconvexity of ($\mathcal{P}1$). We also note that $L_{\boldsymbol{\theta}}$ and $L_{\boldsymbol{\beta}}$ are not required to run Algorithm 1.

Remark 4: A question naturally arising is why we have optimized the amplitude, $[\boldsymbol{\beta}_u]_i$, and phase shift, $[\boldsymbol{\theta}_u]_i$, $u \in \{t, r\}$ separately, rather than optimizing them as a single complex, e.g., $[v_u]_i = [\boldsymbol{\beta}_u]_i e^{j[\boldsymbol{\theta}_u]_i}$. The latter would certainly make the presentation of the proposed method more elegant. However, interestingly enough, we find by extensive numerical experiments that both ways give the same performance in many cases. We note that this does not mean the amplitudes corresponding to the reflection or transmission mode of most of the STAR-IRS elements are close to 1. There is indeed a significant gap between the ES and MS mode as shown in the next section. However, in some cases, using two separate variables yields a better performance. This numerical observation has led to the current presentation of the proposed method where amplitudes and phase shifts are optimized separately.

B. Optimization of Amplitudes and Phase Shifts for MS protocol

In the case of the MS scheme, the values of amplitude are forced to be binary, i.e., $\boldsymbol{\beta}_n^t \in \{0,1\}$ and $\boldsymbol{\beta}_n^r \in \{0,1\}$. Thus, the optimization problem for the MS protocol is stated as

$$
\begin{aligned}
\max_{\boldsymbol{\theta}, \boldsymbol{\beta}} \quad & f(\boldsymbol{\theta}, \boldsymbol{\beta}) \\
\text{s.t} \quad & \beta_n^t + \beta_n^r = 1, \forall n \in \mathcal{N} \\
& \beta_n^t \in \{0,1\}, \beta_n^r \in \{0,1\}, \ \forall n \in \mathcal{N} \\
& |\theta_n^t| = |\theta_n^r| = 1, \ \forall n \in \mathcal{N}.
\end{aligned}
\quad (\mathcal{P}2)
$$

The binary constraints on $\boldsymbol{\beta}_n^t$ and $\boldsymbol{\beta}_n^r$ in ($\mathcal{P}2$) make it far more difficult to solve. In fact, ($\mathcal{P}2$) belongs to the class of binary nonconvex programming, which is generally NP-hard. For this type of problems, a pragmatic approach is to find a high-performing solution. To this end, we find that the simple solution obtained by rounding off the solution obtained by solving ($\mathcal{P}1$) to the nearest binary value can produce a reasonably good performance. This shall be numerically demonstrated in the next section. More advanced methods for solving ($\mathcal{P}2$) are thus left for future work.

VI. NUMERICAL RESULTS

In this section, we present numerical results for the sum SE in STAR-RIS-aided systems, using both analytical techniques and Monte Carlo simulations. Specifically, our analytical results for the sum SE are derived from equations (18)-(20), while for Monte Carlo simulations, we perform 1000 independent channel realizations to evaluate the expressions in equations (15)-(17). This is to verify the tightness of the approximation stated in Proposition 1 and the derivations in Appendix B. The results shown in Figs. 3 and 5 clearly demonstrate a close match between the analytical results and MC simulations, and thus, confirming that Proposition 1 indeed presents a very tight approximation of the SINR.

The simulation setup includes a STAR-RIS with a UPA of $N = 64$ elements assisting the communication between a uniform linear array (ULA) of $M = 64$ antennas at the BS that

serves $K = 4$ UEs. The $xy-$coordinates of the BS and RIS are given as $(x_B, y_B) = (0, 0)$ and $(x_R, y_R) = (50, 10)$, respectively, all in meter units. In addition, users in r region are located on a straight line between $(x_R - \frac{1}{2}d_0, y_R - \frac{1}{2}d_0)$ and $(x_R + \frac{1}{2}d_o, y_R - \frac{1}{2}d_0)$ with equal distances between each two adjacent users, and $d_0 = 20$ m in our simulations. Similarly, users in the t region are located between $(x_R - \frac{1}{2}d_0, y_R + \frac{1}{2}d_0)$ and $(x_R + \frac{1}{2}d_o, y_R + \frac{1}{2}d_0)$. The size of each RIS element is $d_H = d_V = \lambda/4$. Distance-based path-loss is considered in our work, such that the channel gain of a given link j is $\tilde{\beta}_j = A d_j^{-\alpha_j}$, where A is the area of each reflecting element at the RIS, and α_j is the path-loss exponent. Regarding $\tilde{\beta}_g$, we assume the same values as for $\tilde{\beta}_j$. Similar values are assumed for $\bar{\beta}_k$ but we also consider an additional penetration loss equal to 15 dB. The correlation matrices $\mathbf{R}_{BS}$ and $\mathbf{R}_{RIS}$ are computed according to [24] and [4], respectively. Also, $\sigma^2 = -174 + 10\log_{10} B_c$ in dBm, where $B_c = 200$ kHz is the bandwidth.

As a baseline scheme, we consider the RIS, which consists of transmitting-only or reflecting-only elements, each with N_t and N_r elements, such that $N_t + N_r = N$. Notably, this scheme resembles the MS protocol, where the first N_t elements operate in transmission mode and the N_r elements operate in reflection mode. Also, we have applied an ON/OFF scheme for channel estimation by following the idea in [17], the direct links are estimated with all sub-surfaces turned off and the cascaded links are estimated with one element turned on at transmission/reflection mode sequentially.
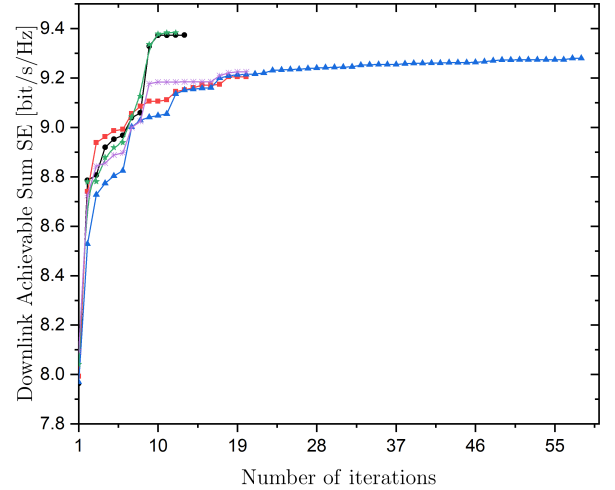

Fig. 2: Convergence of Algorithm 1 for an STAR-RIS assisted MIMO system with imperfect CSI ($M = 64$, $N = 64$, $\tilde{K} = 4$) for five different initial points.

In the first numerical experiment, we demonstrate the convergence of the proposed projected gradient algorithm. Specifically, we plot the achievable sum SE against the iteration count returned by Algorithm 1 from 5 different randomly generated initial points as shown in Fig. 2. More specifically, the initial points for Algorithm 1 are generated as follows. First, we set the amplitudes to $[\boldsymbol{\beta}_r^{(0)}]_n = [\boldsymbol{\beta}_t^{(0)}]_n = \sqrt{0.5}$, for all $n \in \mathcal{N}$, i.e, equal power splitting between transmission and reception
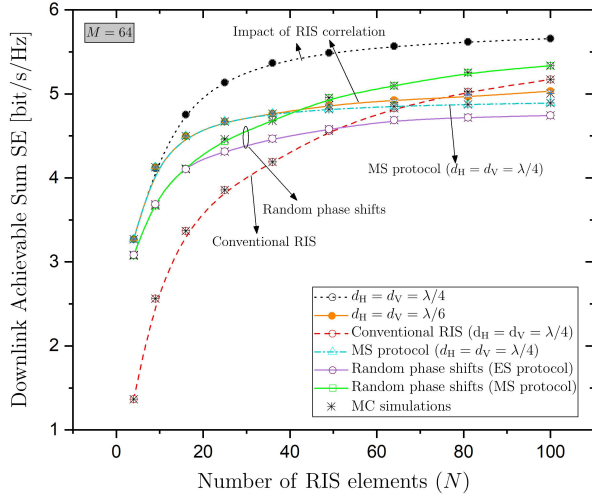
Fig. 3: Downlink achievable sum SE versus the number of RIS elements antennas N of a STAR-RIS assisted MIMO system with imperfect CSI ($N = 64$, $K = 4$) for varying conditions (Analytical results and MC simulations).

mode for all elements of the STAR-RIS. The initial values for the phase shifts are taken as $[\boldsymbol{\theta}_r^{(0)}]_n = e^{j\phi_n^r}$ and $[\boldsymbol{\theta}_t^{(0)}]_n = e^{j\phi_n^t}$, where ϕ_n^r and ϕ_n^t are independently drawn from the uniform distribution over $[0, 2\pi]$. We terminate Algorithm 1 when the increase of the objective between two last iterations is less than 10^{-5} or the number of iterations is larger than 200. Note that the considered problem in ($\mathcal{P}1$) is nonconvex, and thus, the proposed projected gradient algorithm can only guarantee a stationary solution that is not necessarily optimal. As a result, Algorithm 1 may converge to different points starting from different initial points, which is clearly seen in Fig. 2. Moreover, we can see that different initial points may lead to different convergence rates. Thus, to mitigate this performance sensitivity of Algorithm 1 on the initial points, we need to run it from different initial points and take the best convergent solutions. Through our extensive simulations, it is best to run Algorithm 1 from 5 randomly generated initial points to achieve a good trade-off between complexity and obtained sum SE.

Fig. 3 shows the achievable sum SE versus the number of STAR-RIS elements N while varying the effect of spatial correlation in terms of the size of each RIS element. First, as can be seen, the downlink sum SE increases with N as expected. Next, by focusing on the impact of spatial correlation at the STAR-RIS, we show that the performance decreases as the correlation increases. In particular, the sum SE decreases with increased correlation as the inter-element distance of the STAR-RIS decreases. Moreover, the MS protocol achieves a lower performance because it is a special case of the ES protocol. Especially, for a low number of RIS elements the curves coincide, while as N increases, an increasing gap appears. Furthermore, for the sake of comparison, we provide the performance of conventional RIS with reflection-only operation but this also appears lower performance since fewer degrees of freedom for just reflection can be exploited. We have also depicted the performance in the case of blocked direct signal.
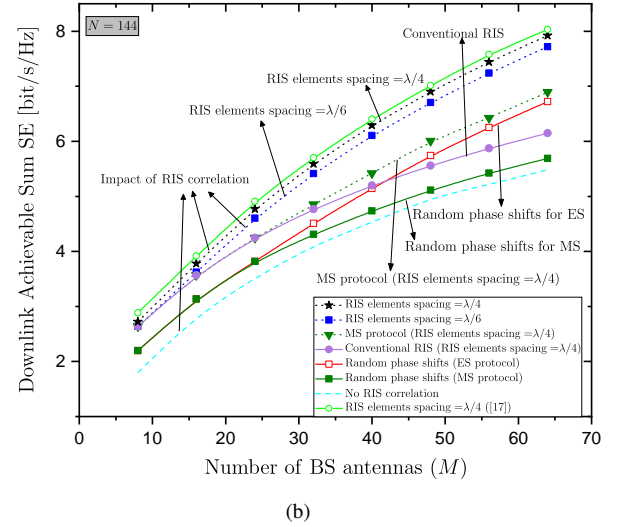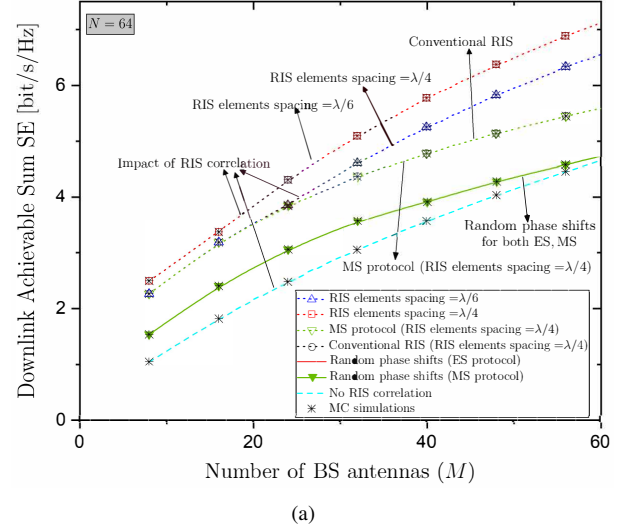

(a)

(b)

Fig. 4: Downlink achievable sum SE versus the number of BS antennas M of a STAR-RIS assisted MIMO system with imperfect CSI for: (a) $N = 64$, $K = 4$, (b) $N = 144$, $K = 4$ under varying conditions (Analytical results).

Obviously, the STAR-RIS contributes to the performance since the line corresponding the case with no direct signal is lower, which means that the performance is worse.

Figs. 4(a) and 4(b) illustrate the achievable sum SE versus the number of BS antennas M while shedding light on various effects. Obviously, the sum SE increases with M. In particular, regarding the RIS correlation, in Fig. 4(a), we observe that an increased correlation by reducing the distance among the RIS elements degrades the performance due to reduced diversity gains among the RIS elements. In the case of no RIS correlation, represented by the dashed cyan line, the performance is quite low due to the absence of capability for phase shift optimization as mentioned in Remark 1. Moreover, the ES protocol achieves better performance but with higher complexity compared to the MS protocol. The performance increases with more BS antennas. Also, in the case of random phase shifts, the sum SE is lower. Notably, for $N = 64$ elements, the two lines
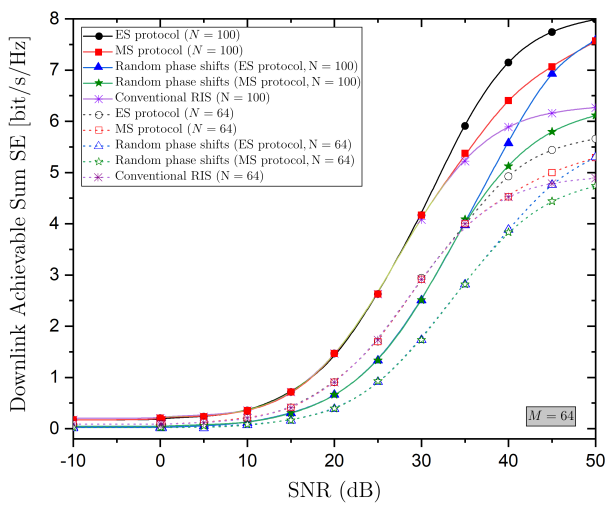
Fig. 5: Downlink achievable sum SE versus the SNR of a STAR-RIS assisted MIMO system with imperfect CSI ($M = 64$, $N = 64$, $K = 4$) for varying conditions (Analytical results).

we continue to focus on users in the reflection region. Thus, at high SNR, directing some power to users in the transmission region can improve the total SE. This leads to performance differences between the ES and MS protocols, as well as the conventional RIS.

VII. CONCLUSION

This paper presented a study of the achievable rate of STAR-RIS assisted mMIMO systems while accounting for imperfect CSI and correlated Rayleigh fading. Notably, we considered several UEs, each of which can lie on either side of the RIS, and we derived the achievable rate in closed-form. Also, we provided a low-complexity iterative optimization approach to maximizing the achievable rate, in which the amplitudes and the phase shifts of the RIS are updated simultaneously at each iteration. Furthermore, we provided useful insights into the impact of RIS correlation and showed that STAR-RIS is more beneficial compared to the traditional RIS which is reflecting only.

corresponding to the ES and MS protocols coincide but, according to Fig.4(b), which assumes $N = 144$ elements, a gap between the lines appears. The gap increases with increasing M. Similar to the previous figure, we have included the baseline scenario with reflecting-only capabilities having the half elements $N = 20$, and we witness the superiority of STAR-RIS. Moreover, a comparison of the two figures, reveals that an increase in RIS elements spacing has a greater impact on a lower number of RIS elements, i.e., $N = 64$. Furthermore, Fig. 4(b) shows that the scenario of no RIS correlation performs worse than the cases with random phases when M is large. Also, in this figure, it is shown that the achievable rate is higher than Fig 4(b). In addition, in Fig. 4(b), we have added a line corresponding to channel estimation based on the ON/OFF scheme in [17]. We observe that the achievable rate is higher in this case because in the case of statistical CSI we have loss of information. In other words, we observe a trade-off between a lower overhead of the proposed approach and a higher rate in the case of estimating individual channels.

Fig. 5 depicts the achievable sum SE versus the SNR under similar conditions, i.e., in the cases of $N = 100$ (solid lines) and $N = 64$ (dotted lines). As expected, when $N = 100$, the performance is better since a higher SE is achieved. In each case, for low SNR, the ES and MS protocols coincide, while for high SNR, an increasing gap is observed. In the case $N = 100$, conventional RIS and random MS protocols exhibit the same performance at low SNR, but a gap appears as the SNR increases. The behavior at low SNR is similar, however, the corresponding gaps are smaller. The reasons for these observations can be explained as follows. At low SNR, it is more beneficial to focus on users in the reflection region as they are closer to the BS. This is confirmed by the fact that, after running the proposed algorithm, $\beta_n^r \approx 1, \forall n \in \mathcal{N}$. As a result, the performances of the ES and MS protocols, as well as the conventional RIS, are nearly the same. However, as SNR increases, the increase in the sum SE becomes minimal if

APPENDIX A
PROOF OF LEMMA 1

The LMMSE estimator of $\mathbf{h}_k$, obtained by minimizing $\mathrm{tr}\big(\mathbb{E}\big\{(\hat{\mathbf{h}}_k - \mathbf{h}_k)(\hat{\mathbf{h}}_k - \mathbf{h}_k)^{\mathsf{H}}\big\}\big)$, is given by

$$\hat{\mathbf{h}}_k = \mathbb{E}\{\mathbf{r}_k \mathbf{h}_k^{\mathsf{H}}\}\left(\mathbb{E}\{\mathbf{r}_k \mathbf{r}_k^{\mathsf{H}}\}\right)^{-1}\mathbf{r}_k. \tag{37}$$

Given that the channel and the receiver noise are uncorrelated, we obtain

$$\mathbb{E}\{\mathbf{r}_k \mathbf{h}_k^{\mathsf{H}}\} = \mathbb{E}\{\mathbf{h}_k \mathbf{h}_k^{\mathsf{H}}\} = \mathbf{R}_k. \tag{38}$$

The second term in (37) is written as

$$\mathbb{E}\{\mathbf{r}_k \mathbf{r}_k^{\mathsf{H}}\} = \mathbf{R}_k + \frac{\sigma^2}{\tau P}\mathbf{I}_M. \tag{39}$$

The LMMSE estimate in (8) is obtained by inserting (38) and (39) into (37), which completes the proof. We further note that the covariance matrix of the estimated channel is

$$\mathbb{E}\left\{\hat{\mathbf{h}}_k \hat{\mathbf{h}}_k^{\mathsf{H}}\right\} = \mathbf{R}_k \mathbf{Q}_k \mathbf{R}_k. \tag{40}$$

APPENDIX B
PROOF OF PROPOSITION 1

Recalling the property $\mathbf{x}^{\mathsf{H}}\mathbf{y} = \mathrm{tr}(\mathbf{y}\mathbf{x}^{\mathsf{H}})$ for any vectors $\mathbf{x}$, $\mathbf{y}$, we can further rewrite S_k in (16) as

$$S_k = |\mathbb{E}\{\mathbf{h}_k^{\mathsf{H}}\hat{\mathbf{h}}_k\}|^2 = |\mathrm{tr}\left(\mathbb{E}\{\hat{\mathbf{h}}_k \mathbf{h}_k^{\mathsf{H}}\}\right)|^2 \tag{41}$$

$$= |\mathrm{tr}\left(\mathbb{E}\{\mathbf{R}_k \mathbf{Q}_k \mathbf{r}_k \mathbf{h}_k^{\mathsf{H}}\}\right)|^2 \tag{42}$$

$$= |\mathrm{tr}\left(\boldsymbol{\Psi}_k\right)|^2, \tag{43}$$

where, in (42), we have substituted (8). The last equation is obtained after computing the expectation between $\mathbf{r}_k$ and $\mathbf{h}_k$.

Next, the first term of I_k in (17) is written as

$$\mathbb{E}\big\{|\mathbf{h}_k^{\mathsf{H}}\hat{\mathbf{h}}_k - \mathbb{E}\{\mathbf{h}_k^{\mathsf{H}}\hat{\mathbf{h}}_k\}|^2\big\} = \mathbb{E}\big\{|\mathbf{h}_k^{\mathsf{H}}\hat{\mathbf{h}}_k|^2\big\} - |\mathbb{E}\{\mathbf{h}_k^{\mathsf{H}}\hat{\mathbf{h}}_k\}|^2 \tag{44}$$

$$= \mathbb{E}\big\{|\hat{\mathbf{h}}_k^{\mathsf{H}}\hat{\mathbf{h}}_k + \tilde{\mathbf{h}}_k^{\mathsf{H}}\hat{\mathbf{h}}_k|^2\big\} - |\mathbb{E}\{\hat{\mathbf{h}}_k^{\mathsf{H}}\hat{\mathbf{h}}_k\}|^2 \tag{45}$$

$$= \mathbb{E}\big\{|\tilde{\mathbf{h}}_k^{\mathsf{H}}\hat{\mathbf{h}}_k|^2\big\} + 2\,\mathrm{Re}\{\mathbb{E}\{\hat{\mathbf{h}}_k^{\mathsf{H}}\hat{\mathbf{h}}_k \hat{\mathbf{h}}_k^{\mathsf{H}}\tilde{\mathbf{h}}_k\}\} \tag{46}$$

where in (45), we have used (9) in the main text. In (46), to simplify the second term, we resort to the well-known channel hardening property in massive MIMO which intuitively states that channels behave as deterministic. The same property is also applied to the estimated channels, which means $\hat{\mathbf{h}}_k^{\mathsf{H}}\hat{\mathbf{h}}_k \approx \mathbb{E}\{\hat{\mathbf{h}}_k^{\mathsf{H}}\hat{\mathbf{h}}_k\}$ with high accuracy [27]. Using this property, we have

$$\mathbb{E}\{\hat{\mathbf{h}}_k^{\mathsf{H}}\hat{\mathbf{h}}_k\hat{\mathbf{h}}_k^{\mathsf{H}}\tilde{\mathbf{h}}_k\} \approx \mathbb{E}\{\hat{\mathbf{h}}_k^{\mathsf{H}}\hat{\mathbf{h}}_k\}\mathbb{E}\{\hat{\mathbf{h}}_k^{\mathsf{H}}\tilde{\mathbf{h}}_k\} = 0 \qquad (47)$$

and thus

$$\mathbb{E}\{|\mathbf{h}_k^{\mathsf{H}}\hat{\mathbf{h}}_k - \mathbb{E}\{\mathbf{h}_k^{\mathsf{H}}\hat{\mathbf{h}}_k\}|^2\} \approx \mathbb{E}\{|\tilde{\mathbf{h}}_k^{\mathsf{H}}\hat{\mathbf{h}}_k|^2\} \qquad (48)$$
$$= \operatorname{tr}(\mathbf{R}_k\mathbf{\Psi}_k) - \operatorname{tr}(\mathbf{\Psi}_k^2). \qquad (49)$$

We note that (49) holds because $\mathbb{E}\{|\tilde{\mathbf{h}}_k^{\mathsf{H}}\hat{\mathbf{h}}_k|^2\} = \operatorname{tr}(\mathbb{E}\{\tilde{\mathbf{h}}_k\tilde{\mathbf{h}}_k^{\mathsf{H}}\hat{\mathbf{h}}_k\hat{\mathbf{h}}_k^{\mathsf{H}}\}) \approx \operatorname{tr}(\mathbb{E}\{\tilde{\mathbf{h}}_k\tilde{\mathbf{h}}_k^{\mathsf{H}}\}\mathbb{E}\{\hat{\mathbf{h}}_k\hat{\mathbf{h}}_k^{\mathsf{H}}\}) = \operatorname{tr}((\mathbf{R}_k - \mathbf{\Psi}_k)\mathbf{\Psi}_k)$, where we have applied the approximations $\hat{\mathbf{h}}_k^{\mathsf{H}}\hat{\mathbf{h}}_k \approx \mathbb{E}\{\hat{\mathbf{h}}_k^{\mathsf{H}}\hat{\mathbf{h}}_k\}$ and $\tilde{\mathbf{h}}_k^{\mathsf{H}}\tilde{\mathbf{h}}_k \approx \mathbb{E}\{\tilde{\mathbf{h}}_k^{\mathsf{H}}\tilde{\mathbf{h}}_k\}$, which is due to the channel hardening property in massive MIMO as explained above.

For the second term of I_k in (17) it is easy to check that

$$\mathbb{E}\{|\mathbf{h}_k^{\mathsf{H}}\hat{\mathbf{h}}_i|^2\} = \mathbb{E}\{|(\hat{\mathbf{h}}_k^{\mathsf{H}} + \tilde{\mathbf{h}}_k)\hat{\mathbf{h}}_i|^2\} \qquad (50)$$
$$= \mathbb{E}\{|\hat{\mathbf{h}}_k^{\mathsf{H}}\hat{\mathbf{h}}_i|^2\} + \mathbb{E}\{|\tilde{\mathbf{h}}_k\hat{\mathbf{h}}_i|^2\} + 2\operatorname{Re}\{\mathbb{E}\{\hat{\mathbf{h}}_k^{\mathsf{H}}\hat{\mathbf{h}}_i\hat{\mathbf{h}}_i^{\mathsf{H}}\tilde{\mathbf{h}}_k\}\} \qquad (51)$$
$$= \operatorname{tr}(\mathbf{R}_k\mathbf{\Psi}_i). \qquad (52)$$

We note that (52) is true because the third term in (51) is zero as can be shown below. Specifically, this term can be written as

$$2\operatorname{Re}\{\mathbb{E}\{\hat{\mathbf{h}}_k^{\mathsf{H}}\hat{\mathbf{h}}_i\hat{\mathbf{h}}_i^{\mathsf{H}}\tilde{\mathbf{h}}_k\}\} = 2\operatorname{Re}\left\{\mathbb{E}\{\operatorname{tr}((\hat{\mathbf{h}}_i\hat{\mathbf{h}}_i^{\mathsf{H}})(\tilde{\mathbf{h}}_k\hat{\mathbf{h}}_k^{\mathsf{H}}))\}\right\} \qquad (53)$$
$$= 2\operatorname{Re}\left\{\operatorname{tr}(\mathbb{E}\{\hat{\mathbf{h}}_i\hat{\mathbf{h}}_i^{\mathsf{H}}\}\mathbb{E}\{\tilde{\mathbf{h}}_k\hat{\mathbf{h}}_k^{\mathsf{H}}\})\right\} \qquad (54)$$
$$= 0 \qquad (55)$$

where, in (54), we have accounted for the independence between $\hat{\mathbf{h}}_k$ and $\hat{\mathbf{h}}_i$, and, in (55), we have considered that $\tilde{\mathbf{h}}_k$ and $\hat{\mathbf{h}}_k$ are uncorrelated.

The normalization parameter is written as

$$\lambda = \frac{1}{\sum_{i=1}^{K}\mathbb{E}\{\mathbf{f}_i^{\mathsf{H}}\mathbf{f}_i\}} = \frac{1}{\sum_{i=1}^{K}\mathbb{E}\{\hat{\mathbf{h}}_i^{\mathsf{H}}\hat{\mathbf{h}}_i\}} = \frac{1}{\sum_{i=1}^{K}\operatorname{tr}(\mathbf{\Psi}_i)}. \qquad (56)$$

Combining (49), (52), and (56), we can approximate I_k as $\tilde{I}_k$ in (20), and thus complete the proof.

APPENDIX C
PROOF OF LEMMA 2

Let us first derive $\nabla_{\boldsymbol{\theta}^t}f(\boldsymbol{\theta},\boldsymbol{\beta})$ the complex gradient of the achievable sum SE with respect to $\boldsymbol{\theta}^{t*}$. From (14), it is easy to see that

$$\nabla_{\boldsymbol{\theta}^t}f(\boldsymbol{\theta},\boldsymbol{\beta}) = c\sum_{k=1}^{K}\frac{\tilde{I}_k\nabla_{\boldsymbol{\theta}^t}S_k - S_k\nabla_{\boldsymbol{\theta}^t}\tilde{I}_k}{(1+\gamma_k)\tilde{I}_k^2}, \qquad (57)$$

where $c = \frac{\tau_c - \tau}{\tau_c\log_2(e)}$. To compute $\nabla_{\boldsymbol{\theta}^t}S_k$ for a given user k, we immediately note that $\nabla_{\boldsymbol{\theta}^t}S_k = 0$ if $w_k = r$, i.e., if UE k is in the reflection region. This is obvious from (19), (10) and

(5). Thus, we only need to find $\nabla_{\boldsymbol{\theta}^t}S_k$ when $w_k = t$. In such a case, we can explicitly write $\mathbf{R}_k$

$$\mathbf{R}_k = \bar{\beta}_k\mathbf{R}_{\mathrm{BS}} + \hat{\beta}_k\operatorname{tr}(\mathbf{R}_{\mathrm{RIS}}\mathbf{\Phi}_t\mathbf{R}_{\mathrm{RIS}}\mathbf{\Phi}_t^{\mathsf{H}})\mathbf{R}_{\mathrm{BS}}$$
$$= \bar{\beta}_k\mathbf{R}_{\mathrm{BS}} + \hat{\beta}_k\operatorname{tr}(\mathbf{A}_t\mathbf{\Phi}_t^{\mathsf{H}})\mathbf{R}_{\mathrm{BS}}. \qquad (58)$$

where $\hat{\beta}_k = \tilde{\beta}_g\tilde{\beta}_k$ and $\mathbf{A}_t = \mathbf{R}_{\mathrm{RIS}}\mathbf{\Phi}_t\mathbf{R}_{\mathrm{RIS}}$. When $w_k = r$, we define $\mathbf{A}_r = \mathbf{R}_{\mathrm{RIS}}\mathbf{\Phi}_r\mathbf{R}_{\mathrm{RIS}}$.

To calculate $\nabla_{\boldsymbol{\theta}^t}S_k$, we follow steps detailed in [31, Chap. 3]. First, let us denote $d(\cdot)$ the complex differential of the function in the argument. Then, it holds that

$$d(S_k) = d(\operatorname{tr}(\mathbf{\Psi}_k)^2)$$
$$= 2\operatorname{tr}(\mathbf{\Psi}_k)d\operatorname{tr}(\mathbf{\Psi}_k)$$
$$= 2\operatorname{tr}(\mathbf{\Psi}_k)\operatorname{tr}(d\mathbf{\Psi}_k). \qquad (59)$$

Next, we apply [31, Eq. (3.35)], which gives

$$d(\mathbf{\Psi}_k) = d(\mathbf{R}_k\mathbf{Q}_k\mathbf{R}_k)$$
$$= d(\mathbf{R}_k)\mathbf{Q}_k\mathbf{R}_k + \mathbf{R}_kd(\mathbf{Q}_k)\mathbf{R}_k + \mathbf{R}_k\mathbf{Q}_kd(\mathbf{R}_k). \qquad (60)$$

The differentials $d(\mathbf{R}_k)$ and $d(\mathbf{Q}_k)$ are derived as follows. First, from (58) it is easy to check that

$$d(\mathbf{R}_k) = \hat{\beta}_k\mathbf{R}_{\mathrm{BS}}\operatorname{tr}(\mathbf{A}_t^{\mathsf{H}}d(\mathbf{\Phi}_t) + \mathbf{A}_td(\mathbf{\Phi}_t^{\mathsf{H}})), \qquad (61)$$

Since $\mathbf{\Phi}_t$ is diagonal, we can further write $d(\mathbf{R}_k)$ as

$$d(\mathbf{R}_k) = \hat{\beta}_k\mathbf{R}_{\mathrm{BS}}\big((\operatorname{diag}(\mathbf{A}_t^{\mathsf{H}}\operatorname{diag}(\boldsymbol{\beta}^t)))^{\mathsf{T}}d(\boldsymbol{\theta}^t)$$
$$+ (\operatorname{diag}(\mathbf{A}_t\operatorname{diag}(\boldsymbol{\beta}^t)))^{\mathsf{T}}d(\boldsymbol{\theta}^{t*})\big). \qquad (62)$$

Next, we use [31, eqn. (3.40)] to obtain

$$d(\mathbf{Q}_k) = d(\mathbf{R}_k + \frac{\sigma}{\tau P}\mathbf{I}_M)^{-1}$$
$$= -(\mathbf{R}_k + \frac{\sigma^2}{\tau P}\mathbf{I}_M)^{-1}d(\mathbf{R}_k + \frac{\sigma^2}{\tau P}\mathbf{I}_M)(\mathbf{R}_k + \frac{\sigma}{\tau P}\mathbf{I}_M)^{-1}$$
$$= -\mathbf{Q}_kd(\mathbf{R}_k)\mathbf{Q}_k. \qquad (63)$$

Combining (60) and (63) yields

$$d(\mathbf{\Psi}_k) = d(\mathbf{R}_k)\mathbf{Q}_k\mathbf{R}_k - \mathbf{R}_k\mathbf{Q}_kd(\mathbf{R}_k)\mathbf{Q}_k\mathbf{R}_k + \mathbf{R}_k\mathbf{Q}_kd\mathbf{R}_k. \qquad (64)$$

Thus, by inserting (64) and (62) into (59), we obtain

$$d(S_k) = 2\operatorname{tr}(\mathbf{\Psi}_k)\big(\operatorname{tr}(\mathbf{Q}_k\mathbf{R}_kd(\mathbf{R}_k)) + \operatorname{tr}(\mathbf{R}_k\mathbf{Q}_kd(\mathbf{R}_k))$$
$$- \operatorname{tr}(\mathbf{Q}_k\mathbf{R}_k^2\mathbf{Q}_kd(\mathbf{R}_k))\big) \qquad (65a)$$
$$= \nu_k\big(\operatorname{diag}((\mathbf{A}_t^{\mathsf{H}}\operatorname{diag}(\boldsymbol{\beta}^t)))^{\mathsf{T}}d\boldsymbol{\theta}^t + (\operatorname{diag}(\mathbf{A}_t\operatorname{diag}(\boldsymbol{\beta}^t)))^{\mathsf{T}}d\boldsymbol{\theta}^{t*}\big), \qquad (65b)$$

where

$$\nu_k = 2\hat{\beta}_k\operatorname{tr}(\mathbf{\Psi}_k)\operatorname{tr}((\mathbf{Q}_k\mathbf{R}_k + \mathbf{R}_k\mathbf{Q}_k - \mathbf{Q}_k\mathbf{R}_k^2\mathbf{Q}_k)\mathbf{R}_{\mathrm{BS}}). \qquad (66)$$

From (65), we can conclude that

$$\nabla_{\boldsymbol{\theta}^t}S_k = \frac{\partial}{\partial\boldsymbol{\theta}^{t*}}S_k$$
$$= \nu_k\operatorname{diag}(\mathbf{A}_t\operatorname{diag}(\boldsymbol{\beta}^t)) \qquad (67)$$

for $w_k = t$, which indeed proves (25a). Following the same procedure we can easily prove (25b). The details are skipped for the sake of brevity.

Now we turn our attention to $\nabla_{\boldsymbol{\theta}^t}\tilde{I}_k$. To this end, from (20), it is straightforward to check that

$$d(\tilde{I}_k) = \sum_{i=1}^{K}\operatorname{tr}(d(\mathbf{R}_k)\boldsymbol{\Psi}_i) + \sum_{i=1}^{K}\operatorname{tr}(\mathbf{R}_k d(\boldsymbol{\Psi}_i))$$
$$- 2\operatorname{tr}(\boldsymbol{\Psi}_k d(\boldsymbol{\Psi}_k)) + \frac{K\sigma^2}{\rho}\sum_{i=1}^{K}\operatorname{tr}(d(\boldsymbol{\Psi}_i)). \quad (68)$$

Note that $d(\boldsymbol{\Psi}_i) = 0$ if $w_i \neq t$ since $\boldsymbol{\Psi}_i$ is independent of $\boldsymbol{\theta}^t$ in this case. Thus, the above equation is reduced to

$$d(\tilde{I}_k) = \operatorname{tr}(\boldsymbol{\Psi} d(\mathbf{R}_k)) - 2\operatorname{tr}(\boldsymbol{\Psi}_k d(\boldsymbol{\Psi}_k)) + \sum_{i\in\mathcal{K}_t}\operatorname{tr}(\bar{\mathbf{R}}_k d(\boldsymbol{\Psi}_i)), \quad (69)$$

where $\boldsymbol{\Psi} = \sum_{i=1}^{K}\boldsymbol{\Psi}_i$ and $\bar{\mathbf{R}}_k = \mathbf{R}_k + \frac{K\sigma^2}{\rho}\mathbf{I}_M$. Using (64) into (69) gives

$$d(\tilde{I}_k) = \operatorname{tr}(\boldsymbol{\Psi} d\mathbf{R}_k)$$
$$+ \sum_{i\in\mathcal{K}_t}\operatorname{tr}(\bar{\mathbf{R}}_k(d(\mathbf{R}_i)\mathbf{Q}_i\mathbf{R}_i - \mathbf{R}_i\mathbf{Q}_i d(\mathbf{R}_i)\mathbf{Q}_i\mathbf{R}_i + \mathbf{R}_i\mathbf{Q}_i d(\mathbf{R}_i)))$$
$$- 2\operatorname{tr}(\boldsymbol{\Psi}_k(d(\mathbf{R}_k)\mathbf{Q}_k\mathbf{R}_k - \mathbf{R}_k\mathbf{Q}_k d(\mathbf{R}_k)\mathbf{Q}_k\mathbf{R}_k + \mathbf{R}_k\mathbf{Q}_k d(\mathbf{R}_k)))$$
$$= \operatorname{tr}(\check{\boldsymbol{\Psi}}_k d(\mathbf{R}_k)) + \sum_{i\in\mathcal{K}_t}\operatorname{tr}(\tilde{\mathbf{R}}_{ki}d(\mathbf{R}_i)), \quad (70)$$

where

$$\check{\boldsymbol{\Psi}}_k = \boldsymbol{\Psi} - 2(\mathbf{Q}_k\mathbf{R}_k\boldsymbol{\Psi}_k + \boldsymbol{\Psi}_k\mathbf{R}_k\mathbf{Q}_k - \mathbf{Q}_k\mathbf{R}_k\boldsymbol{\Psi}_k\mathbf{R}_k\mathbf{Q}_k) \quad (71)$$

and

$$\tilde{\mathbf{R}}_{ki} = \mathbf{Q}_i\mathbf{R}_i\bar{\mathbf{R}}_k - \mathbf{Q}_i\mathbf{R}_i\bar{\mathbf{R}}_k\mathbf{R}_i\mathbf{Q}_i + \bar{\mathbf{R}}_k\mathbf{R}_i\mathbf{Q}_i, i\in\mathcal{K}_t. \quad (72)$$

Again, we note that $d(\mathbf{R}_k) = 0$ if $w_k \neq t$. Thus, by using (62), we can write $\nabla_{\boldsymbol{\theta}^t}\tilde{I}_k$ as

$$\nabla_{\boldsymbol{\theta}^t}\tilde{I}_k = \frac{\partial}{\partial\boldsymbol{\theta}^{t*}}\tilde{I}_k = \operatorname{diag}(\tilde{\mathbf{A}}_{kt}\operatorname{diag}(\boldsymbol{\beta}^t)), \quad (73)$$

where $\bar{\nu}_k = \hat{\beta}_k\operatorname{tr}(\check{\boldsymbol{\Psi}}_k\mathbf{R}_{\mathrm{BS}})$, $\tilde{\nu}_{ki} = \hat{\beta}_k\operatorname{tr}(\tilde{\mathbf{R}}_{ki}\mathbf{R}_{\mathrm{BS}})$, and

$$\tilde{\mathbf{A}}_{kt} = \begin{cases} \bar{\nu}_k\mathbf{A}_t + \sum_{i\in\mathcal{K}_t}^{K}\tilde{\nu}_{ki}\mathbf{A}_t & w_k = t \\ \sum_{i\in\mathcal{K}_t}\tilde{\nu}_{ki}\mathbf{A}_t & w_k \neq t, \end{cases} \quad (74)$$

which is in fact the special case of (26) when $u = t$, meaning that (25c) has been proved. Following the same steps we can prove (25d), but again, we skip the details for the sake of brevity.

The expression for $\nabla_{\boldsymbol{\beta}^t}S_k$ is derived as follows. First we only need to consider $\nabla_{\boldsymbol{\beta}^t}S_k$ when $w_k = t$. For this case, from (61), we can write

$$d(\mathbf{R}_k) = \hat{\beta}_k\mathbf{R}_{\mathrm{BS}}\operatorname{tr}(\mathbf{A}_t^{\mathsf{H}}d(\boldsymbol{\Phi}_t) + \mathbf{A}_t d(\boldsymbol{\Phi}_t^{\mathsf{H}})) \quad (75a)$$
$$= \hat{\beta}_k\mathbf{R}_{\mathrm{BS}}(\operatorname{diag}(\mathbf{A}_t^{\mathsf{H}}\operatorname{diag}(\boldsymbol{\theta}^t))^{\mathsf{T}}d(\boldsymbol{\beta}^t)$$
$$+ \operatorname{diag}(\mathbf{A}_t\operatorname{diag}(\boldsymbol{\theta}^{t*}))^{\mathsf{T}}d(\boldsymbol{\beta}^t)) \quad (75b)$$
$$= 2\hat{\beta}_k\mathbf{R}_{\mathrm{BS}}\operatorname{Re}\{\operatorname{diag}(\mathbf{A}_t^{\mathsf{H}}\operatorname{diag}(\boldsymbol{\theta}^t)\}^{\mathsf{T}}d(\boldsymbol{\beta}^t). \quad (75c)$$

Now, using (75) in (65) yields

$$\nabla_{\boldsymbol{\beta}^t}S_k = 2\nu_k\operatorname{Re}\{\operatorname{diag}(\mathbf{A}_t^{\mathsf{H}}\operatorname{diag}(\boldsymbol{\theta}^t))\}. \quad (76)$$

Similarly, we can write $\nabla_{\boldsymbol{\beta}^r}S_k$ as

$$\nabla_{\boldsymbol{\beta}^r}S_k = 2\nu_k\operatorname{Re}\{\operatorname{diag}(\mathbf{A}_r^{\mathsf{H}}\operatorname{diag}(\boldsymbol{\theta}^r))\}. \quad (77)$$

For $\nabla_{\boldsymbol{\beta}^t}\tilde{I}_k$ and $\nabla_{\boldsymbol{\beta}^r}\tilde{I}_k$, we can follow the same steps above, which gives

$$\nabla_{\boldsymbol{\beta}^t}\tilde{I}_k = 2\operatorname{Re}\{\operatorname{diag}(\tilde{\mathbf{A}}_{kt}^{\mathsf{H}}\operatorname{diag}(\boldsymbol{\theta}^t))\}. \quad (78)$$
$$\nabla_{\boldsymbol{\beta}^r}\tilde{I}_k = 2\operatorname{Re}\{\operatorname{diag}(\tilde{\mathbf{A}}_{kr}^{\mathsf{H}}\operatorname{diag}(\boldsymbol{\theta}^r))\}. \quad (79)$$

REFERENCES

[1] 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, 2019.

[2] E. Basar *et al.*, "Wireless communications through reconfigurable intelligent surfaces," *IEEE Access*, vol. 7, pp. 116 753–116 773, 2019.

[3] S. Kisseleff *et al.*, "Reconfigurable intelligent surfaces for smart cities: Research challenges and opportunities," *IEEE Open J. Commun. Soc.*, vol. 1, pp. 1781–1797, 2020.

[4] 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, 2021.

[5] T. Van Chien *et al.*, "Outage probability analysis of IRS-assisted systems under spatially correlated channels," *IEEE Wireless Commun. Lett.*, vol. 10, no. 8, pp. 1815–1819, 2021.

[6] A. Papazafeiropoulos *et al.*, "Intelligent reflecting surface-assisted MU-MISO systems with imperfect hardware: Channel estimation and beamforming design," *IEEE Trans. Wireless Commun.*, vol. 21, no. 3, pp. 2077–2092, 2021.

[7] A. Papazafeiropoulos, "Ergodic capacity of IRS-assisted MIMO systems with correlation and practical phase-shift modeling," *IEEE Wireless Commun. Lett.*, vol. 11, no. 2, pp. 421–425, 2022.

[8] J. Xu *et al.*, "STAR-RISs: Simultaneous transmitting and reflecting reconfigurable intelligent surfaces," *IEEE Commun. Lett.*, vol. 25, no. 9, pp. 3134–3138, 2021.

[9] X. Mu *et al.*, "Simultaneously transmitting and reflecting (STAR) RIS aided wireless communications," *IEEE Trans. Wireless Commun.*, vol. 21, no. 5, pp. 3083–3098, 2021.

[10] H. Niu *et al.*, "Simultaneous transmission and reflection reconfigurable intelligent surface assisted secrecy MISO networks," *IEEE Commun. Lett.*, vol. 25, no. 11, pp. 3498–3502, 2021.

[11] C. Wu *et al.*, "Channel estimation for STAR-RIS-aided wireless communication," *IEEE Commun. Letters*, vol. 26, no. 3, pp. 652–656, 2021.

[12] H. Niu *et al.*, "Weighted sum rate optimization for STAR-RIS-assisted MIMO system," *IEEE Trans. Veh. Tech.*, vol. 71, no. 2, pp. 2122–2127, 2022.

[13] B. Zheng *et al.*, "A survey on channel estimation and practical passive beamforming design for intelligent reflecting surface aided wireless communications," *IEEE Commun. Sur. & Tut.*, vol. 24, no. 2, pp. 1035–1071, 2022.

[14] Y. Yang *et al.*, "Intelligent reflecting surface meets OFDM: Protocol design and rate maximization," *IEEE Trans. Commun.*, vol. 68, no. 7, pp. 4522–4535, 2020.

[15] 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, 2020.

[16] B. Zheng and R. Zhang, "Intelligent reflecting surface-enhanced OFDM: Channel estimation and reflection optimization," vol. 9, no. 4, pp. 518–522, 2019.

[17] D. Mishra and H. Johansson, "Channel estimation and low-complexity beamforming design for passive intelligent surface assisted MISO wireless energy transfer," in *International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2019, pp. 4659–4663.

[18] 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, 2019.

[19] A. M. Elbir *et al.*, "Deep channel learning for large intelligent surfaces aided mm-Wave massive MIMO systems," *IEEE Wireless Commun. Lett.*, vol. 9, no. 9, pp. 1447–1451, 2020.

[20] Q. Nadeem *et al.*, "Intelligent reflecting surface-assisted multi-user MISO Communication: Channel estimation and beamforming design," *IEEE Open J. Commun. Soc.*, vol. 1, pp. 661–680, 2020.

[21] T. Wang *et al.*, "Performance analysis of IOS-Assisted NOMA system with channel correlation and phase errors," *IEEE Trans. Veh. Tech.*, vol. 71, no. 11, pp. 11 861–11 875, 2022.

[22] M.-M. Zhao *et al.*, "Intelligent reflecting surface enhanced wireless networks: Two-timescale beamforming optimization," *IEEE Trans. Wireless Commun.*, vol. 20, no. 1, pp. 2–17, 2020.

[23] A. Papazafeiropoulos *et al.*, "Asymptotic analysis of Max-Min weighted SINR for IRS-assisted MISO systems with hardware impairments," *IEEE Wireless Commun. Lett.*, pp. 1–1, 2021.

[24] J. Hoydis, S. ten Brink, and M. Debbah, "Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?" *IEEE J. Select. Areas Commun.*, vol. 31, no. 2, pp. 160–171, 2013.

[25] D. Neumann, M. Joham, and W. Utschick, "Covariance matrix estimation in massive MIMO," *IEEE Signal Process. Lett.*, vol. 25, no. 6, pp. 863–867, 2018.

[26] K. Upadhya and S. A. Vorobyov, "Covariance matrix estimation for massive MIMO," *IEEE Signal Process. Lett.*, vol. 25, no. 4, pp. 546–550, 2018.

[27] E. Björnson *et al.*, "Massive MIMO networks: Spectral, energy, and hardware efficiency," *Foundations and Trends® in Signal Processing*, vol. 11, no. 3-4, pp. 154–655, 2017.

[28] N. V. Deshpande *et al.*, "Spatially-correlated irs-aided multiuser FD mMIMO systems: Analysis and optimization," *IEEE Trans. Commun.*, vol. 70, no. 6, pp. 3879–3896, 2022.

[29] M. Medard, "The effect upon channel capacity in wireless communications of perfect and imperfect knowledge of the channel," *IEEE Trans. Inf. Theory*, vol. 46, no. 3, pp. 933–946, May 2000.

[30] D. Bertsekas, *Nonlinear Programming*, 2nd ed., M. A. Scientific, Ed., 1999.

[31] A. Hjørungnes, *Complex-Valued Matrix Derivatives: With Applications in Signal Processing and Communications*. Cambridge University Press, 2011.