Under review for possible publication in

Two-Timescale Design for Simultaneous Transmitting and Reflecting RIS-Assisted Massive MIMO Systems with Imperfect CSI

Journal:	*IEEE Transactions on Communications*
Manuscript ID	TCOM-TPS-23-1233
Manuscript Type:	Transactions Paper Submissions
Date Submitted by the Author:	17-Sep-2023
Complete List of Authors:	Dai, Jianxin; Nanjing University of Posts and Telecommunications, School of Science; Southeast University, National Mobile Communications Research Laboratory Zhang, Shilong; Nanjing University of Posts and Telecommunications, Pan, Cunhua; Southeast University, Zhi, Kangda; Queen Mary University of London, Wang, Xianbin; The University of Western Ontario, Electrical and Computer Engineering Wang, Cheng-Xiang; Southeast University - Sipailou Campus, School of Information Science and Engineering; Purple Mountain Laboratories, Pervasive Communication Research Center
Keyword:	Broadband communication, Communication system performance

Two-Timescale Design for Simultaneous Transmitting and Reflecting RIS-Assisted Massive MIMO Systems with Imperfect CSI

Jianxin Dai, Shilong Zhang, Kangda Zhi, Cunhua Pan, Hong Ren, Xianbin Wang, *Fellow, IEEE*, Cheng-Xiang Wang, *Fellow, IEEE*

***Abstract*—This paper investigates the performance of simultaneous transmitting and reflecting reconfigurable intelligent surface (STAR-RIS) assisted massive multiple-input multiple-output (MIMO) systems with Rician fading channels and channel estimation errors. We adopt the two-timescale scheme to design the systems, namely, applying the instantaneous channel state information (CSI) to design the base station (BS) beamforming and leveraging the statistical CSI to design the phase shifts of the STAR-RIS. Specifically, we estimate the overall channels based on the linear minimum mean-squared error (LMMSE) estimator and derive the closed-form expression of the average achievable rate. Based on the derived rate, we analyze the power scaling laws in which the transmit power is respectively reduced inversely proportional to the number of BS antennas and STAR-RIS elements. Besides, we draw insights from the comparison between STAR-RIS and conventional RIS under the same condition, and optimize the phase shifts of the STAR-RIS to maximize the sum rate using accelerated gradient ascent-based algorithm. Finally, numerical results are provided to validate our theoretical insights. We show that STAR-RIS outperforms conventional RIS and draw insight behind this phenomenon.**

***Index Terms*—Simultaneously transmitting and reflecting reconfigurable intelligent surface (STAR-RIS), massive MIMO, two-timescale, imperfect CSI.**

I. INTRODUCTION

During the past thirty years, the focus of multiple-input multiple-output (MIMO) systems has also evolved from the beginning of point-to-point MIMO in the beginning to the current massive MIMO, and MIMO technology combined with active relay has been recently applied to the fifth-generation (5G) communication [1], [2]. However, the conventional active relay requires complicated feeding networks and high power consumption, which greatly restricts the capability of massive MIMO systems [3], [4]. Reconfigurable intelligent surface (RIS) was proposed to tackle this dilemma and identified as a pivotal technology for the sixth-generation (6G) communication systems [5], [6]. It is made up of many passive and low-cost reflecting elements which can flexibly adjust their phase, amplitude, and frequency of incident electromagnetic (EM) waves benefitting from its programable EM property [7], [8]. After constructive phase shift adjustment and signal combination, the signal impinging on RIS elements will be strengthened without an active amplifier [9]. Besides, RIS also has the advantages of compatibility with existing hardware, various installation environments, and enhancing spectral efficiency [10], [11]. Hitherto, many contributions have been devoted to RIS [12]–[16]. An overall description of the physical structure, application scenarios, related challenges, and development prospects of RIS are given in [12] and [13]. The authors of [14] derived the expression of outage probability, average bit error rate (BER), and average capacity, and validated the effectiveness of integrating RIS into unmanned aerial vehicle (UAV) communication systems. RIS-aided non-orthogonal multiple access (NOMA) systems were studied in [15] in which the authors theoretically analyzed the closed-form expression of the outage probability and the ergodic rate and verified the superiority of the proposed system. In [16], the authors compared the active RIS-aided system with the passive RIS-aided system based on the assumption that these two systems have the same overall power budget and concluded that the active RIS is superior to the passive RIS when having a small power budget.

However, existing reflection-only RIS suffers from the geographic topology constraint, where the transmitter and the receiver should be located on the same side of the RIS [17]. This means that the conventional RIS can only serve users in half-space with 180° coverage, greatly influencing the performance of users behind RIS. We follow a novel concept of simultaneous transmitting and reflecting RIS (STAR-RIS) to overcome this disadvantage. STAR-RIS can reflect and transmit the received signal to the sections in front of and behind it simultaneously, thus achieving 360° full-space coverage [18]. Furthermore, STAR-RIS can flexibly control the power ratio between reflection and transmission with the assistance of the transmitting and reflecting coefficients (TRCs), which

This work was supported by the National Mobile Communications Research Laboratory, Southeast University.(No.2023D03).

Jianxin Dai is with the School of Science, Nanjing University of Posts and Telecommunications, Nanjing 210023, China, and also with the National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China (e-mail: daijx@njupt.edu.cn). Shilong Zhang is with the College of Telecommunications and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210096, China (e-mail:1221014214@njupt.edu.cn). Kangda Zhi is with the School of Electronic Engineering and Computer Science, Queen Mary University of London, London E1 4NS, U.K. (e-mail: k.zhi@qmul.ac.uk). Cunhua Pan and Hong Ren are with with the National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China (e-mail:cpan@seu.edu.cn; hren@seu.edu.cn). Xianbin Wang is with the Department of Electrical and Computer Engineering, Western University, Canada (email: xianbin.wang@uwo.ca). Cheng-Xiang Wang is with the National Mobile Communications Research Laboratory, School of Information Science and Engineering, Southeast University, Nanjing, China, and Purple Mountain Laboratories, Nanjing, China (e-mail:chxwang@seu.edu.cn).

(Corresponding author: Cunhua Pan).

enlarges the degree of freedom (DoF) of the systems. In [19], the authors comprehensively evaluated STAR-RIS in multiple aspects, compared STAR-RIS with conventional passive and active RISs, and introduced three working protocols of STAR-RIS, including time switching (TS), mode switching (MS), and energy splitting (ES). For these protocols, the authors of [20] focused on minimizing the power consumption based on TS and ES protocols, and concluded that the ES protocol is more suitable for the broadcast transmission. Similarly, [21] optimized the weighted sum rate in STAR-RIS-aided MIMO systems, and likewise proved the superiority of STAR-RIS over conventional RIS and the advantage of ES protocol in broadcast transmission.

Besides, the majority of existing works on RIS are focused on designing the phase shifts of reflecting elements using instantaneous channel state information (CSI), which will significantly increase the channel estimation overhead and power consumption [22]. Due to the frequent changes in the user's locations and the passive feature of RIS, it is difficult to update this information frequently and timely [23]. Two-timescale (TTS) transmission scheme is an effective solution for this problem [24]–[26], by designing the beamforming at BS based on both instantaneous CSI and statistical CSI for the phase shifts design of the RIS. In [24], the authors considered RIS-aided multi-user millimeter Wave (mmWave) systems. Besides, they formulated and optimized the TTS-based beamforming problem based on the deep unrolling method and verified that this design scheme could approach an instantaneous CSI-only scheme. RIS-aided massive MIMO systems based on the TTS scheme were studied in [25], and the authors comprehensively analyzed the power scaling laws of the system. [26] provided and validated the solution for channel estimation of RIS-aided wireless communication systems under the TTS scheme. Considering that both the transmission and reflection phase shifts need to be optimized for STAR-RIS, the complexity and the overhead could be higher than conventional RIS. Therefore, TTS scheme would be a very suitable enabling technique for STAR-RIS.

In addition to the issues mentioned above, the general Rician fading channel model and imperfect CSI should also be considered for STAR-RIS-aided systems. Due to the fact that the BS and RIS are often installed at higher locations with few barriers and RIS generally serves users in a close range, there is a high possibility of existing line-of-sight links between the RIS, the BS, and the users. Meanwhile, it is impractical to assume the perfectly estimation of the channels because of the passive property of the RIS and the limited training period, which results in the estimation errors [27]. Hence, it is essential to consider the Rician fading channel model and imperfect CSI. To the best of our knowledge, none of the existing research studied the performance of STAR-RIS-aided multi-user massive MIMO systems with Rician fading channels in the presence of imperfect CSI which motivates our work.

Our main contributions are summarized as follows:

- The two-timescale scheme is adopted to design the beamforming at the BS and STAR-RIS, respectively based on instantaneous CSI and statistical CSI for less computational complexity and less channel estimation overhead.
- Considering imperfect CSI, the linear minimum mean-squared error (LMMSE) estimator is utilized to estimate the overall transmission and reflection channels. At the same time, the low-complexity maximum ratio combining (MRC) receiver is adopted for the BS.
- We derive the closed-form achievable rate expression, which holds for the arbitrary number of BS antennas and RIS elements under Rician fading and channel estimation errors. The comparison between STAR-RIS and conventional RIS are conducted and the power scaling laws in different cases are analyzed to draw useful insights.
- The accelerated gradient ascent-based method is applied to maximize the sum rate by optimizing the phase shifts of the STAR-RIS.
- Finally, numerical results validate our analytical conclusions and reveal the STAR-RIS's superiority over conventional RIS. Meanwhile, it demonstrates that integrating STAR-RIS into the massive MIMO systems is feasible and effective.

The rest of this work is organized as follows. Section II describes the model of uplink STAR-RIS-aided massive MIMO systems. Section III derives the channel estimation based on LMMSE. In Section IV, we provide the closed-form expressions for the uplink achievable rate and analyze its characteristics. Section V introduces the accelerated gradient ascent-based method to maximize the sum user rate. Section VI provides extensive numerical results to validate our analytical conclusions and characterize the advantages brought by STAR-RIS compared with conventional RIS. Finally, Section VII provide the summary of this work.

Notations: The vectors and the matrices are respectively expressed in lowercase blodface and uppercase blodface letters. $\mathbf{X}^T$, $\mathbf{X}^H$, $\mathbf{X}^{-1}$ and $\mathbf{X}^*$ denote the transpose, conjugate transpose, inverse and conjugate operators, respectively. $[\mathbf{X}]_{m,n}$ denotes the (m,n)-th entry of the matrix. $[\mathbf{a}]_m$ denotes the m-th entry of the vector. $|\mathbf{a}|$ denotes the modulus of the complex number and $\|\mathbf{a}\|$ denotes l_2-norm of the vector. $\mathrm{Re}\{\cdot\}$ and $\mathrm{Cov}\{\cdot\}$ represent the real part and covariance, respectively. the expectation and trace are denoted by $\mathbb{E}\{\cdot\}$ and $\mathrm{Tr}\{\cdot\}$ respectively. $\mathbf{I}_N$ denotes the $N\times N$ identity matrix. $\mathbb{C}^{M\times N}$ represents the $M\times N$ complex-valued matrix. Besides, $x\sim\mathcal{CN}\left(\mu,\sigma^2\right)$ denotes that random variable x follows the complex Gaussian distribution with mean μ and variance σ^2.

II. SYSTEM MODEL

As shown in Fig. 1, we consider the uplink transmission of STAR-RIS-aided massive MIMO systems. The BS is equipped with M antennas, the STAR-RIS is composed of N passive reflecting elements, and the users are assumed to be equipped with a single antenna. We refer to the areas in front of and behind the STAR-RIS as the reflection section (section R) and transmission section (section T), respectively. Then, it is assumed that there are K_1 and K_2 users in section R and T, respectively.

The direct links from the users to the BS are given by $\mathbf{D}_R=\left[\mathbf{d}_1^r,\mathbf{d}_2^r,...,\mathbf{d}_{K_1}^r\right]\in\mathbb{C}^{M\times K_1}$ and $\mathbf{D}_T=\left[\mathbf{d}_1^t,\mathbf{d}_2^t,...,\mathbf{d}_{K_2}^t\right]\in$

$\mathbb{C}^{M\times K_2}$, the channels between the users and the STAR-RIS are denoted by $\mathbf{H}_R=\left[\mathbf{h}_1^r,\mathbf{h}_2^r,...,\mathbf{h}_{K_1}^r\right]\in\mathbb{C}^{N\times K_1}$ and $\mathbf{H}_T=\left[\mathbf{h}_1^t,\mathbf{h}_2^t,...,\mathbf{h}_{K_2}^t\right]\in\mathbb{C}^{N\times K_2}$, and $\mathbf{H}\in\mathbb{C}^{M\times N}$ represents the channel between the BS and the STAR-RIS.

Let $\mathbf{Q}=[\mathbf{Q}_R,\mathbf{Q}_T]$ where $\mathbf{Q}_R=\mathbf{D}_R+\mathbf{H}\mathbf{\Phi}_R\mathbf{H}_R$ and $\mathbf{Q}_T=\mathbf{D}_T+\mathbf{H}\mathbf{\Phi}_T\mathbf{H}_T$ represent the overall user-RIS-BS channels in sections R and T, respectively. Consequently, the signal vector received by the BS is given by

$$\mathbf{y}=\sqrt{p_{k_1}}\mathbf{Q}_R\mathbf{x}_1+\sqrt{p_{k_2}}\mathbf{Q}_T\mathbf{x}_2+\mathbf{n}, \tag{1}$$

where p_{k_1} and p_{k_2} are the average transmit power of users in R and T sections, respectively, $\mathbf{x}_1=[x_1^r,x_2^r,...,x_{K_1}^r]^T$, $\mathbf{x}_2=[x_1^t,x_2^t,...,x_{K_2}^t]^T$ represent users' transmit symbols, $\mathbf{n}\sim\mathcal{CN}\left(0,\sigma^2\mathbf{I}\right)$ is the additional white Gaussian noise (AWGN).

The configuration matrices for the STAR-RIS can be expressed as $\mathbf{\Phi}_R=\sqrt{A_R}\,\mathrm{diag}\left\{e^{j\theta_1^R},e^{j\theta_2^R},...,e^{j\theta_N^R}\right\}$ and $\mathbf{\Phi}_T=\sqrt{A_T}\,\mathrm{diag}\left\{e^{j\theta_1^T},e^{j\theta_2^T},...,e^{j\theta_N^T}\right\}$, where $\theta_n^T,\theta_n^R\in[0,2\pi)$ and $A_R,A_T\in[0,1]$ are the phase shifts and amplitudes of the STAR-RIS, respectively. The reflection and transmission coefficients should obey $A_R+A_T=1$ due to the law of energy conservation.

By employing the low-complexity maximal-ratio-combining (MRC) technique at the BS, the transmitted symbols from K_1+K_2 users can be obtained as

$$\mathbf{r}=\mathbf{Q}^H\mathbf{y}=\sqrt{p_{k_1}}\mathbf{Q}^H\mathbf{Q}_R\mathbf{x}_1+\sqrt{p_{k_2}}\mathbf{Q}^H\mathbf{Q}_T\mathbf{x}_2+\mathbf{Q}^H\mathbf{n}. \tag{2}$$

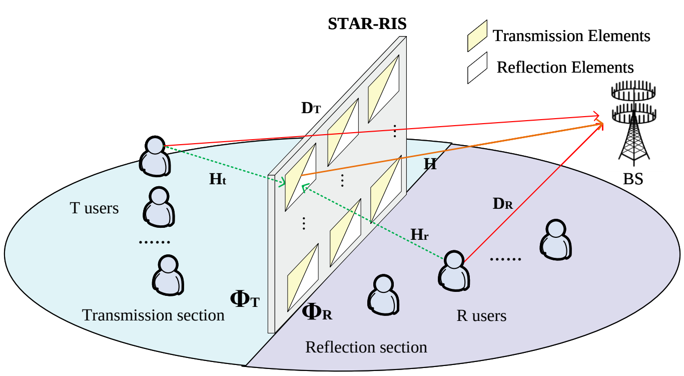

Fig. 1: STAR-RIS-assisted massive MIMO communication system.

For the convenience of subsequent derivation, we denote $\chi,\overline{\chi}\in\{k_1,k_2\}$, $1\le k_1\le K_1$, $1\le k_2\le K_2$. When $\chi=k_1$, then $\chi_{\max}=K_1$, $\overline{\chi}=k_2$, $\overline{\chi}_{\max}=K_2$, and when $\chi=k_2$, then $\chi_{\max}=K_2$, $\overline{\chi}=k_1$, $\overline{\chi}_{\max}=K_1$. In addition, when $\chi=k_1$, A_χ and $\mathbf{\Phi}_\chi$ respectively denote A_R and $\mathbf{\Phi}_R$. Similarly, when $\chi=k_2$, A_χ and $\mathbf{\Phi}_\chi$ respectively denote A_T and $\mathbf{\Phi}_T$.

A. Channel Model

In consideration of a number of blockages near the ground and long distances between users and BS, the line-of-sight (LoS) channel components barely exist between them. Hence, we use the Rayleigh fading model to characterize the direct links between the BS and users as

$$\mathbf{D}_\chi=\sqrt{\gamma_\chi}\widetilde{\mathbf{d}}_\chi, \tag{3}$$

where γ_χ denotes the large-scale path-loss and $\widetilde{\mathbf{d}}_\chi$ is an independently and identically distributed (i.i.d) complex Gaussian random variable with zero mean and unit variance.

Given that STAR-RIS is generally installed at a high altitude and close to the users, it is highly likely that there are line-of-sight (LoS) channel components between them. Hence, we utilize the Rician fading to model the RIS-related channels as

$$\mathbf{h}_\chi=\sqrt{\frac{\alpha_\chi}{\varepsilon_\chi+1}}\left(\sqrt{\varepsilon_\chi}\overline{\mathbf{h}}_\chi+\widetilde{\mathbf{h}}_\chi\right), \tag{4}$$

$$\mathbf{H}=\sqrt{\frac{\beta}{\delta+1}}\left(\sqrt{\delta}\overline{\mathbf{H}}+\widetilde{\mathbf{H}}\right), \tag{5}$$

where α_χ and β are large-scale path-loss factors, ε_χ and δ are Rician factors. $\widetilde{\mathbf{h}}_\chi$ and $\widetilde{\mathbf{H}}$ denote non-line-of-sight (NLoS) channel components and their entries are i.i.d complex Gaussian random variables following $\mathcal{CN}(0,1)$. $\overline{\mathbf{h}}_\chi$ and $\overline{\mathbf{H}}$ stand for LoS channel components. Based on the uniform squared planar array (USPA) model, they can be formulated as

$$\overline{\mathbf{h}}_\chi=\mathbf{a}_N\left(\psi_{\chi r}^a,\psi_{\chi r}^e\right), \tag{6}$$

$$\overline{\mathbf{H}}=\mathbf{a}_M\left(\phi_r^a,\phi_r^e\right)\mathbf{a}_N^H\left(\varphi_t^a,\varphi_t^e\right), \tag{7}$$

with

$$\mathbf{a}_X\left(\upsilon^a,\upsilon^e\right)=\Big[1,...,e^{j2\pi\frac{d}{\lambda}(x\sin\upsilon^e\sin\upsilon^a+y\cos\upsilon^e)},...,\\ e^{j2\pi\frac{d}{\lambda}\left(\left(\sqrt{X}-1\right)\sin\upsilon^e\sin\upsilon^a+\left(\sqrt{X}-1\right)\cos\upsilon^e\right)}\Big]^T, \tag{8}$$

where d is the element spacing, λ is the wavelength, $\left(\psi_{\chi r}^a,\psi_{\chi r}^e\right)$, $\left(\phi_r^a,\phi_r^e\right)$ are respectively the azimuth and elevation angles of arrival (AoA) from user χ to the STAR-RIS. $\left(\varphi_t^a,\varphi_t^e\right)$ are respectively the azimuth and elevation angles of departure (AoD) from the STAR-RIS to the BS. As a result, the overall channel can be detailed as (9) on the top of next page, where $c_\chi=\frac{\beta\alpha_\chi}{(\delta+1)(\varepsilon_\chi+1)}$.

III. Channel Estimation

The BS will estimate the instantaneous CSI matrix Q according to received pilot signals. Assume that the channel coherence interval consists of τ_c samples. Let $\tau\ge K_1+K_2$ denote the length of samples used for pilot signals. During the training part of each channel coherence interval, users will simultaneously send mutually orthogonal pilot sequences to the BS. The transmit pilot power is τp_χ and the $\chi_{\max}\times\tau$ orthogonal pilot sequences are given by $\mathbf{S}_\chi=[s_1;...;s_{\chi_{\max}}]$ with $\mathbf{S}_\chi\mathbf{S}_\chi^H=I_{\chi_{\max}}$. The pilot signals received at the BS can be written as

$$\mathbf{Y}_\chi=\sqrt{\tau p_\chi}\mathbf{Q}_\chi\mathbf{S}_\chi+\mathbf{N}, \tag{10}$$

where N denotes the $M\times\tau$ noise matrix whose entries are i.i.d. complex Gaussian random variables with zero mean and σ^2 variance. Correlating $\mathbf{Y}_\chi$ with $\frac{\mathbf{s}_\chi^H}{\sqrt{\tau p_\chi}}$, the observation vector for user χ can be given by

$$\mathbf{y}_{p_\chi}^\chi=\frac{1}{\sqrt{\tau p_\chi}}\mathbf{Y}_\chi\mathbf{s}_\chi^H=\mathbf{q}_\chi+\frac{1}{\sqrt{\tau p_\chi}}\mathbf{N}\mathbf{s}_\chi^H. \tag{11}$$

$$\mathbf{q}_\chi = \mathbf{g}_\chi + \mathbf{d}_\chi = \mathbf{H}\boldsymbol{\Phi}_\chi \mathbf{h}_\chi + \mathbf{d}_\chi = \underbrace{\underbrace{\sqrt{c_\chi\delta\varepsilon_\chi}\,\overline{\mathbf{H}}\boldsymbol{\Phi}_\chi\overline{\mathbf{h}}_\chi}_{\mathbf{g}_\chi^1} + \underbrace{\sqrt{c_\chi\delta}\,\overline{\mathbf{H}}\boldsymbol{\Phi}_\chi\widetilde{\mathbf{h}}_\chi}_{\mathbf{g}_\chi^2} + \underbrace{\sqrt{c_\chi\varepsilon_\chi}\widetilde{\mathbf{H}}\boldsymbol{\Phi}_\chi\overline{\mathbf{h}}_\chi}_{\mathbf{g}_\chi^3} + \underbrace{\sqrt{c_\chi}\widetilde{\mathbf{H}}\boldsymbol{\Phi}\widetilde{\mathbf{h}}_\chi}_{\mathbf{g}_\chi^4}}_{\mathbf{g}_\chi} + \sqrt{\gamma_\chi}\widetilde{\mathbf{d}}_\chi, \quad (9)$$

Lemma 1 *The mean and covariance matrices used in the LMMSE estimator are given by*

$$\mathbb{E}\{\mathbf{q}_\chi\} = \mathbb{E}\{\mathbf{y}_{p_\chi}^\chi\} = \sqrt{c_\chi\delta\varepsilon_\chi}\overline{\mathbf{H}}\boldsymbol{\Phi}_\chi\overline{\mathbf{h}}_\chi, \quad (12)$$

$$\begin{aligned}\mathrm{Cov}\{\mathbf{q}_\chi, \mathbf{y}_{p_\chi}^\chi\} &= \mathrm{Cov}\{\mathbf{y}_{p_\chi}^\chi, \mathbf{q}_\chi\} = \mathrm{Cov}\{\mathbf{q}_\chi, \mathbf{y}_{p_\chi}^\chi\} \\ &= a_{\chi_1}\mathbf{a}_M{\mathbf{a}_M}^H + a_{\chi_2}\mathbf{I}_M,\end{aligned} \quad (13)$$

$$\begin{aligned}\mathrm{Cov}\{\mathbf{y}_{p_\chi}^\chi, \mathbf{y}_{p_\chi}^\chi\} &= \mathrm{Cov}\{\mathbf{q}_\chi, \mathbf{q}_\chi\} + \frac{\sigma^2}{\tau p_\chi}\mathbf{I}_M \\ &= a_{\chi_1}\mathbf{a}_M\mathbf{a}_M^H + \left(a_{\chi_2} + \frac{\sigma^2}{\tau p_\chi}\right)\mathbf{I}_M,\end{aligned} \quad (14)$$

where $a_{\chi_1} = Nc_\chi\delta A_\chi$, $a_{\chi_2} = Nc_\chi A_\chi(\varepsilon_\chi + 1) + \gamma_\chi$ *are two auxiliary variables.*

Proof: Please refer to Appendix A. ■

Theorem 1 *Using the observation vector* $\mathbf{y}_\chi$*, the LMMSE estimate of the channel vector* $\mathbf{q}_\chi, \chi \in \{k_1, k_2\}$*, is given by (15) on the top of next page, where*

$$\mathbf{A}_\chi = \mathbf{A}_\chi^H = a_{\chi_3}\mathbf{a}_M\mathbf{a}_M^H + a_{\chi_4}\mathbf{I}_M, \quad (16)$$

$$\mathbf{B}_\chi = (\mathbf{I}_M - \mathbf{A}_\chi)\sqrt{c_\chi\delta\varepsilon_\chi}\,\overline{\mathbf{H}}\boldsymbol{\Phi}_\chi\overline{\mathbf{h}}_\chi, \quad (17)$$

$$a_{\chi_3} = \frac{a_{\chi_1}\frac{\sigma^2}{\tau p_\chi}}{\left(a_{\chi_2} + \frac{\sigma^2}{\tau p_\chi}\right)\left(a_{\chi_2} + \frac{\sigma^2}{\tau p_\chi} + Ma_{\chi_1}\right)}, \quad (18)$$

$$a_{\chi_4} = \frac{a_{\chi_2}}{a_{\chi_2} + \frac{\sigma^2}{\tau p_\chi}}. \quad (19)$$

Proof: Please refer to Appendix B. ■

Lemma 2 *Some useful properties are presented as follows*

$$\mathrm{Tr}\{\mathbf{A}_\chi\} = M(a_{\chi_3} + a_{\chi_4}) \triangleq Me_{\chi_1}, \quad (20)$$

$$\begin{aligned}\mathbf{A}_\chi\overline{\mathbf{H}} &= a_{\chi_3}\mathbf{a}_M\mathbf{a}_M^H\mathbf{a}_M\mathbf{a}_N^H + a_{\chi_4}\mathbf{I}_M\mathbf{a}_M\mathbf{a}_N^H \\ &= (Ma_{\chi_3} + a_{\chi_4})\overline{\mathbf{H}} \triangleq e_{\chi_2}\overline{\mathbf{H}},\end{aligned} \quad (21)$$

$$\begin{aligned}\mathbf{A}_\chi\mathbf{A}_\chi &= a_{\chi_3}\mathbf{a}_M\mathbf{a}_M^H\left(a_{\chi_3}\mathbf{a}_M\mathbf{a}_M^H + a_{\chi_4}\mathbf{I}_M\right) \\ &\quad + a_{\chi_4}\mathbf{I}_M\left(a_{\chi_3}\mathbf{a}_M\mathbf{a}_M^H + a_{\chi_4}\mathbf{I}_M\right) \\ &= Ma_{\chi_3}^2\mathbf{a}_M\mathbf{a}_M^H + 2a_{\chi_3}a_{\chi_4}\mathbf{a}_M\mathbf{a}_M^H + a_{\chi_4}^2\mathbf{I}_M,\end{aligned} \quad (22)$$

$$\mathrm{Tr}\{\mathbf{A}_\chi\mathbf{A}_\chi\} = M\left(Ma_{\chi_3}^2 + 2a_{\chi_3}a_{\chi_4} + a_{\chi_4}^2\right) \triangleq Me_{\chi_3}, \quad (23)$$

where

$$e_{\chi_1} \triangleq a_{\chi_3} + a_{\chi_4}, \quad (24)$$

$$e_{\chi_2} \triangleq Ma_{\chi_3} + a_{\chi_4}, \quad (25)$$

$$e_{\chi_3} \triangleq Ma_{\chi_3}^2 + 2a_{\chi_3}a_{\chi_4} + a_{\chi_4}^2. \quad (26)$$

IV. ANALYSIS OF UPLINK ERGODIC RATE

In this section, the closed-form expression is derived for the achievable rate and the benefits of the STAR-RIS are analyzed in some special cases. Based on the use-and-then-forget bound, the received signal for user χ can be rewritten as (27) at the top of the next page. Then, the achievable rate is obtained in the following theorem.

Theorem 2 *The achievable rate of user* χ *is lower bounded by* $R_\chi = \frac{\tau_c - \tau}{\tau_c}\log_2(1 + \mathrm{SINR}_\chi)$, $\forall\chi$, *with*

$$\mathrm{SINR}_\chi = \frac{p_\chi E_\chi^{\mathrm{signal}}}{p_\chi E_\chi^{\mathrm{leakage}} + p_\chi\sum_{i=1, i\neq\chi}^{\chi_{\max}} I_{\chi i} + p_{\overline{\chi}}\sum_{\overline{\chi}=1}^{\overline{\chi}_{\max}} I_{\chi\overline{\chi}} + \sigma^2 E_\chi^{\mathrm{noise}}}, \quad (28)$$

where

$$E_\chi^{\mathrm{signal}} \triangleq \left|\mathbb{E}\left\{\hat{\mathbf{q}}_\chi^H\mathbf{q}_\chi\right\}\right|^2, \quad (29)$$

$$E_\chi^{\mathrm{leakage}} \triangleq \mathbb{E}\left\{\left|\hat{\mathbf{q}}_\chi^H\mathbf{q}_\chi\right|^2\right\} - \left|\mathbb{E}\left\{\hat{\mathbf{q}}_\chi^H\mathbf{q}_\chi\right\}\right|^2, \quad (30)$$

$$I_{\chi i} \triangleq \mathbb{E}\left\{\left|\hat{\mathbf{q}}_\chi^H\mathbf{q}_i\right|^2\right\}, \quad (31)$$

$$I_{\chi\overline{\chi}} \triangleq \mathbb{E}\left\{\left|\hat{\mathbf{q}}_\chi^H\mathbf{q}_{\overline{\chi}}\right|^2\right\}, \quad (32)$$

$$E_\chi^{\mathrm{noise}} \triangleq \mathbb{E}\left\{\left\|\hat{\mathbf{q}}_\chi\right\|^2\right\}, \quad (33)$$

are respectively given as (34-37) at the top of the next page.

Proof: Please refer to Appendix C. ■

Corollary 1 *In conventional RIS-aided systems, by setting* $A_T = 0$ *and* $A_R = 1$ *in Theorem 2, we can obtain the achievable rate as* $R_{k_3} = \frac{\tau_c - \tau}{\tau_c}\log_2(1 + \mathrm{SINR}_{k_3})$, $\forall k_3$, $1 \leq k_3 \leq K_3 = K_1 + K_2$, *with*

$$\mathrm{SINR}_{k_3} = \frac{p_{k_3}E_{k_3}^{\mathrm{signal}}}{p_{k_3}E_{k_3}^{\mathrm{leakage}} + p_{k_3}\sum_{i=1, i\neq k_3}^{K_3} I_{k_3 i} + \sigma^2 E_{k_3}^{\mathrm{noise}}}, \quad (38)$$

where $E_{k_3}^{\mathrm{signal}} \triangleq \left|\mathbb{E}\left\{\hat{\mathbf{q}}_{k_3}^H\mathbf{q}_{k_3}\right\}\right|^2$, $I_{k_3 i} \triangleq \mathbb{E}\left\{\left|\hat{\mathbf{q}}_{k_3}^H\mathbf{q}_i\right|^2\right\}$, and $E_{k_3}^{\mathrm{noise}} \triangleq \mathbb{E}\left\{\left\|\hat{\mathbf{q}}_{k_3}^H\right\|^2\right\}$. It is readily to check that this result is consistent with the result in [25], which is a special case of our work.

It is evident that the closed-form expression in Theorem 2 does not rely on the NLOS components, including $\widetilde{\mathbf{d}}_\chi$, $\widetilde{\mathbf{h}}_\chi$, $\widetilde{\mathbf{H}}$. It is only affected by statistical CSI, such as the large-scale path-loss factors, the Rician factors, and the AoA and AoD, which changes slowly over a period of time. Hence, following the two-timescale framework, the design frequency of the RIS can be significantly reduced thanks to the slowly varying CSI, diminishing the computational complexity and power consumption.

When introducing the splitting coefficients, the power of signal, interference, and leakage are all scaled down accordingly. However, the direct link, namely users-STAR-RIS channels, is independent of these coefficients. It will lead to the impact of the direct link being relatively amplified, which results in better performance of STAR-RIS.

Besides, it can be seen that coefficients e_{χ_1}, e_{χ_2}, e_{χ_3} completely illustrate the effect of channel estimation errors

$$
\begin{aligned}
\hat{\mathbf{q}}_\chi &= \mathbf{A}_\chi \mathbf{y}_\chi + \mathbf{B}_\chi \\
&= \underbrace{\underbrace{\sqrt{c_\chi \delta \varepsilon_\chi}\, \overline{\mathbf{H}} \boldsymbol{\Phi}_\chi \overline{\mathbf{h}}_\chi}_{\hat{\mathbf{g}}_\chi^1} + \underbrace{\left(M a_{\chi_3} + a_{\chi_4}\right) \sqrt{c_\chi \delta}\, \overline{\mathbf{H}} \boldsymbol{\Phi}_\chi \tilde{\mathbf{h}}_\chi}_{\hat{\mathbf{g}}_\chi^2} + \underbrace{\sqrt{c_\chi \varepsilon_\chi} \mathbf{A}_\chi \tilde{\mathbf{H}} \boldsymbol{\Phi}_\chi \overline{\mathbf{h}}_\chi}_{\hat{\mathbf{g}}_\chi^3} + \underbrace{\sqrt{c_\chi} \mathbf{A}_\chi \tilde{\mathbf{H}} \boldsymbol{\Phi}_\chi \tilde{\mathbf{h}}_\chi}_{\hat{\mathbf{g}}_\chi^4}}_{\hat{\mathbf{g}}_\chi} + \sqrt{\gamma_\chi} \mathbf{A}_\chi \tilde{\mathbf{d}}_\chi + \frac{1}{\sqrt{\tau p_\chi}} \mathbf{A}_\chi \mathbf{N} \mathbf{s}_\chi^H,
\end{aligned} \tag{15}
$$

$$
r_\chi = \underbrace{\sqrt{p_\chi}\, \mathbb{E}\left\{\hat{\mathbf{q}}_\chi^H \mathbf{q}_\chi\right\} x_\chi}_{\text{Desired signal}} + \underbrace{\sqrt{p_\chi}\left(\hat{\mathbf{q}}_\chi^H \mathbf{q}_\chi - \mathbb{E}\left\{\hat{\mathbf{q}}_\chi^H \mathbf{q}_\chi\right\}\right) x_\chi}_{\text{Signal leakage}} + \underbrace{\sqrt{p_\chi} \sum_{i=1, i \neq \chi}^{\chi_{\max}} \mathbf{q}_\chi^H \mathbf{q}_i x_i}_{\text{Multi-user interference}} + \underbrace{\sqrt{p_{\overline{\chi}}} \sum_{\overline{\chi}=1}^{\overline{\chi}_{\max}} \mathbf{q}_\chi^H \mathbf{q}_{\overline{\chi}} + \mathbf{q}_\chi^H \mathbf{n}}_{\text{Inter-group interference}} + \underbrace{\hat{\mathbf{q}}_\chi^H \mathbf{n}}_{\text{Noise}}. \tag{27}
$$

$$
E_\chi^{\text{signal}} = \left\{E_\chi^{\text{noise}}\right\}^2, \qquad E_\chi^{\text{noise}} = M\left\{\left|f_\chi(\boldsymbol{\Phi}_\chi)\right|^2 c_\chi \delta \varepsilon_\chi + A_\chi N c_\chi \delta e_{\chi_2} + \left(A_\chi N c_\chi \left(\varepsilon_\chi + 1\right) + \gamma_\chi\right) e_{\chi_1}\right\}, \tag{34}
$$

$$
\begin{aligned}
E_\chi^{\text{leakage}} =& M\left|f_\chi(\boldsymbol{\Phi}_\chi)\right|^2 c_\chi \delta \varepsilon_\chi \left[c_\chi A_\chi \left\{N\left(M\delta + \varepsilon_\chi + 1\right)\left(e_{\chi_2}^2 + 1\right) + 2\left(M e_{\chi_1} + e_{\chi_2}\right)\left(e_{\chi_2} + 1\right)\right\} + \left\{\gamma_\chi + \left(\gamma_\chi + \frac{\sigma^2}{\tau p_\chi}\right) e_{\chi_2}^2\right\}\right. \\
& \left. + M\gamma_\chi \left(\gamma_\chi + \frac{\sigma^2}{\tau p_\chi}\right) e_{\chi_3}\right] + M^2 N^2 A_\chi^2 c_\chi^2 \delta^2 e_{\chi_2}^2 + M N^2 A_\chi^2 c_\chi^2 \left\{2\delta\left(\varepsilon_\chi + 1\right) e_{\chi_2}^2 + \left(\varepsilon_\chi + 1\right)^2 e_{\chi_3}\right\} + M^2 N A_\chi^2 c_\chi^2 \times \\
& \left\{\left(2\varepsilon_\chi + 1\right) e_{\chi_1}^2 + 2\delta e_{\chi_1} e_{\chi_2}\right\} + M N c_\chi \left\{A_\chi^2 c_\chi \left(2\delta e_{\chi_2}^2 + \left(2\varepsilon_\chi + 1\right) e_{\chi_3}\right) + \left(2\gamma_\chi + \frac{\sigma^2}{\tau p_\chi}\right)\left(\delta e_{\chi_2}^2 + \left(\varepsilon_\chi + 1\right) e_{\chi_3}\right)\right\},
\end{aligned} \tag{35}
$$

$$
\begin{aligned}
I_{\chi i} =& M^2 \left|f_\chi(\boldsymbol{\Phi}_\chi)\right|^2 \left|f_i(\boldsymbol{\Phi}_i)\right|^2 c_\chi c_i \delta^2 \varepsilon_\chi \varepsilon_i + M\left|f_\chi(\boldsymbol{\Phi}_\chi)\right|^2 c_\chi \delta \varepsilon_\chi \left\{A_\chi c_i \left(MN\delta + N\varepsilon_i + N + 2M e_{\chi_1}\right) + \gamma_i\right\} \\
& + M\left|f_i(\boldsymbol{\Phi}_i)\right|^2 c_i \delta \varepsilon_i \left\{A_\chi c_\chi e_{\chi_2}\left(MN\delta e_{\chi_2} + N\varepsilon_\chi e_{\chi_2} + N e_{\chi_2} + 2M e_{\chi_1}\right) + \left(\gamma_\chi + \frac{\sigma^2}{\tau p_\chi}\right) e_{\chi_2}^2\right\} + M^2 N^2 A_\chi^2 c_\chi c_i \delta^2 e_{\chi_2}^2 \\
& + M N^2 A_\chi^2 c_\chi c_i \left\{\delta\left(\varepsilon_\chi + \varepsilon_i + 2\right) e_{\chi_2}^2 + \left(\varepsilon_\chi + 1\right)\left(\varepsilon_i + 1\right) e_{\chi_3}\right\} + M^2 N A_\chi^2 c_\chi c_i e_{\chi_1} \left\{\left(\varepsilon_\chi + \varepsilon_i + 1\right) e_{\chi_1} + 2\delta e_{\chi_2}\right\} \\
& + M^2 c_\chi c_i \varepsilon_\chi \varepsilon_i e_{\chi_1} \left(A_\chi^2 \left|\overline{\mathbf{h}}_\chi^H \overline{\mathbf{h}}_i\right|^2 e_{\chi_1} + 2A_\chi \delta\, \mathrm{Re}\left\{f_\chi^H(\boldsymbol{\Phi}_\chi) f_i(\boldsymbol{\Phi}_i) \overline{\mathbf{h}}_i^H \overline{\mathbf{h}}_\chi\right\}\right) + M\gamma_i \left(\gamma_\chi + \frac{\sigma^2}{\tau p_\chi}\right) e_{\chi_3} \\
& + M N A_\chi \left\{\left(\gamma_\chi + \frac{\sigma^2}{\tau p_\chi}\right) c_i \left(\delta e_{\chi_2}^2 + \left(\varepsilon_i + 1\right) e_{\chi_3}\right) + \gamma_i c_\chi \left(\delta e_{\chi_2}^2 + \left(\varepsilon_\chi + 1\right) e_{\chi_3}\right)\right\},
\end{aligned} \tag{36}
$$

$$
\begin{aligned}
I_{\chi\overline{\chi}} =& M^2 \left|f_\chi(\boldsymbol{\Phi}_\chi)\right|^2 \left|f_{\overline{\chi}}(\boldsymbol{\Phi}_{\overline{\chi}})\right|^2 c_\chi c_{\overline{\chi}} \delta^2 \varepsilon_\chi \varepsilon_{\overline{\chi}} + M\left|f_\chi(\boldsymbol{\Phi}_\chi)\right|^2 c_\chi \delta \varepsilon_\chi \left\{A_{\overline{\chi}} c_{\overline{\chi}} \left(MN\delta + N\varepsilon_{\overline{\chi}} + N + 2M e_{\chi_1}\right) + \gamma_{\overline{\chi}}\right\} \\
& + M\left|f_{\overline{\chi}}(\boldsymbol{\Phi}_{\overline{\chi}})\right|^2 c_{\overline{\chi}} \delta \varepsilon_{\overline{\chi}} \left\{A_\chi c_\chi e_{\chi_2}\left(MN\delta e_{\chi_2} + N\varepsilon_\chi e_{\chi_2} + N e_{\chi_2} + 2M e_{\chi_1}\right) + \left(\gamma_\chi + \frac{\sigma^2}{\tau p_\chi}\right) e_{\chi_2}^2\right\} \\
& + M N^2 A_\chi A_{\overline{\chi}} c_\chi c_{\overline{\chi}} \left\{\delta\left(\varepsilon_\chi + \varepsilon_{\overline{\chi}} + 2\right) e_{\chi_2}^2 + \left(\varepsilon_\chi + 1\right)\left(\varepsilon_{\overline{\chi}} + 1\right) e_{\chi_3}\right\} + M^2 N A_\chi A_{\overline{\chi}} c_\chi c_{\overline{\chi}} e_{\chi_1} \left\{\left(\varepsilon_\chi + \varepsilon_{\overline{\chi}} + 1\right) e_{\chi_1} + 2\delta e_{\chi_2}\right\} \\
& + M^2 c_\chi c_{\overline{\chi}} \varepsilon_\chi \varepsilon_{\overline{\chi}} e_{\chi_1} \left(\left|\overline{\mathbf{h}}_\chi^H \boldsymbol{\Phi}_\chi^H \boldsymbol{\Phi}_{\overline{\chi}} \overline{\mathbf{h}}_{\overline{\chi}}\right|^2 e_{\chi_1} + 2\sqrt{A_\chi}\sqrt{A_{\overline{\chi}}}\,\delta\, \mathrm{Re}\left\{f_\chi^H(\boldsymbol{\Phi}_\chi) f_{\overline{\chi}}(\boldsymbol{\Phi}_{\overline{\chi}}) \overline{\mathbf{h}}_{\overline{\chi}}^H \overline{\mathbf{h}}_\chi\right\}\right) + M\gamma_{\overline{\chi}} \left(\gamma_\chi + \frac{\sigma^2}{\tau p_\chi}\right) e_{\chi_3} \\
& + M N \left\{A_\chi \left(\gamma_\chi + \frac{\sigma^2}{\tau p_\chi}\right) c_{\overline{\chi}} \left(\delta e_{\chi_2}^2 + \left(\varepsilon_{\overline{\chi}} + 1\right) e_{\chi_3}\right) + A_{\overline{\chi}} \gamma_{\overline{\chi}} c_\chi \left(\delta e_{\chi_2}^2 + \left(\varepsilon_\chi + 1\right) e_{\chi_3}\right)\right\} + M^2 N^2 A_\chi A_{\overline{\chi}} c_\chi c_{\overline{\chi}} \delta^2 e_{\chi_2}^2,
\end{aligned} \tag{37}
$$

on this system. Meanwhile, Theorem 2 enable us to reveal the difference between the considered STAR-RIS and the conventional reflection-only RIS.

On one hand, it can be seen that STAR-RIS introduces two extra power splitting coefficients A_R and A_T which can be exploited to adjust the power distribution. On the other hand, STAR-RIS bring a new inter-T/R section interference which is highly coupled by two different phase shift matrices. For this extra interference term, we will leverage some special cases to find out whether it will limit the potential of STAR-RIS systems, compared with the reflection-only RIS.

For a more straightforward comparison between STAR-RIS and conventional RIS and a better analysis of the influence of this inter-group interference term on the achievable rate, we consider a special case where users in the STAR-RIS-aided system and the conventional RIS-aided system are both placed closely and have the same transmit power at the same time, namely $c_\chi = c_{k_3} = c$, $\gamma_\chi = \gamma_{k_3} = \gamma$, $p_\chi = p_{k_3} = p$, $\forall\chi$. Additionally, we set $A_R = A_T = 1/2$ to divide the energy equally between the reflection region and the transmission region.

Corollary 2 *Consider a special case that the cascaded channels are pure NLoS, namely* $\delta = \varepsilon_\chi = \varepsilon_{k_3} = 0$, $\forall\chi$, *and STAR-RIS and RIS are deployed in the environment with rich scatters and* $\gamma > 0$, $\alpha > 0$, *In this case, the rate of user* χ *and* k_3 *are respectively lower bounded by* $R_\chi^{(\mathrm{FNL})} = \frac{\tau_c - \tau}{\tau_c}\log_2\left(1 + \mathrm{SINR}_\chi^{(\mathrm{FNL})}\right)$, $R_{k_3}^{(\mathrm{FNL})} = \frac{\tau_c - \tau}{\tau_c}\log_2\left(1 + \mathrm{SINR}_{k_3}^{(\mathrm{FNL})}\right)$, *and* SINR *are respectively given as*

$$R_{\mathrm{STAR}}^{(\mathrm{NL})}=\sum_{\chi=1}^{\chi_{\max}} R_{\chi}{}^{(\mathrm{NL})}+\sum_{\overline{\chi}=1}^{\overline{\chi}_{\max}} R_{\overline{\chi}}^{(\mathrm{NL})} \tag{39}$$
$$\triangleq \sum_{\chi=1}^{\chi_{\max}} \log_2\left(1+\mathrm{SINR}_{\chi}^{(\mathrm{NL})}\right)+\sum_{\overline{\chi}=1}^{\overline{\chi}_{\max}} \log_2\left(1+\mathrm{SINR}_{\overline{\chi}}^{(\mathrm{NL})}\right)$$
$$=\sum_{\chi=1}^{K_3} \log_2\left(1+\mathrm{SINR}_{\chi}^{(\mathrm{NL})}\right),$$
$$R_{\mathrm{Con}}^{(\mathrm{NL})}=\sum_{k_3=1}^{K_3} R_{k_3}^{(\mathrm{NL})} \triangleq \sum_{k_3=1}^{K_3} \log_2\left(1+\mathrm{SINR}_{k_3}^{(\mathrm{NL})}\right), \tag{40}$$

and these SINR are given as

$$\mathrm{SINR}_{\chi}^{(\mathrm{NL})} \approx \tag{41}$$
$$\frac{pM\left(Nc+2\gamma\right)^2 e_{\chi 1}}{p_\chi E_\chi^{\text{leakage}}+p_\chi \sum\limits_{i=1, i\neq\chi}^{\chi_{\max}} I_{\chi i}+p_{\overline{\chi}} \sum\limits_{\overline{\chi}=1}^{\overline{\chi}_{\max}} I_{\chi\overline{\chi}}+2\sigma^2\left(Nc+2\gamma\right)},$$
$$\mathrm{SINR}_{k_3}^{(\mathrm{NL})} \approx \tag{42}$$
$$\frac{pM\left(Nc+\gamma\right)^2 e_1^{\mathrm{Con}}}{p_{k_3} E_{k_3}^{\text{leakage}}+p_{k_3} \sum\limits_{i=1, i\neq k_3}^{K_3} I_{k_3 i}+\sigma^2\left(Nc+\gamma\right)},$$

where

$$p_\chi E_\chi^{\text{leakage}}+p_\chi \sum_{i=1, i\neq\chi}^{\chi_{\max}} I_{\chi i}+p_{\overline{\chi}} \sum_{\overline{\chi}=1}^{\overline{\chi}_{\max}} I_{\chi\overline{\chi}} \tag{43}$$
$$=pK_3\left[c^2 e_{\chi_1}\left(N^2+MN\right)+4\gamma e_{\chi_1}\left(\gamma+\frac{\sigma^2}{\tau p}\right)+Nce_{\chi_1}\left(4\gamma+\frac{2\sigma^2}{\tau p}\right)\right]+pNce_{\chi_1}\left(c+4\gamma+\frac{2\sigma^2}{\tau p}\right),$$
$$p_{k_3} E_{k_3}^{\text{leakage}}+p_{k_3} \sum_{i=1, i\neq k_3}^{K_3} I_{k_3 i} \tag{44}$$
$$=pK_3\left[c^2 e_1^{\mathrm{Con}}\left(N^2+MN\right)+\gamma e_1^{\mathrm{Con}}\left(\gamma+\frac{\sigma^2}{\tau p}\right)+Nce_1^{\mathrm{Con}}\left(2\gamma+\frac{\sigma^2}{\tau p}\right)\right]+pNc^2 e_1^{\mathrm{Con}},$$
$$e_{\chi_1}=\frac{N\beta\alpha_\chi+2\gamma_\chi}{N\beta\alpha_\chi+2\gamma_\chi+\frac{2\sigma^2}{\tau p}},\ e_1^{\mathrm{Con}}=\frac{N\beta\alpha_{k_3}+\gamma_{k_3}}{N\beta\alpha_{k_3}+\gamma_{k_3}+\frac{\sigma^2}{\tau p}}. \tag{45}$$

Corollary 2 indicates that the achievable rate is not affected by phase shifts when the cascaded channels are pure NLoS. We observe that the power terms in (41) and (42) are both on the order of $\mathcal{O}\left(MN^2\right)$, and the overall interference terms are both on the order of $\mathcal{O}\left(MN\right)$ or $\mathcal{O}\left(N^2\right)$, when $N, M \to \infty$. Hence, the rate will keep growing without bounds when $N, M \to \infty$.

To show the difference intuitively, we consider two special cases with a large number of BS antennas and RIS elements, i.e., $M \to \infty$ and $N \to \infty$.

Corollary 3 *When the number of BS antennas is large, namely $M \to \infty$, the SINR in (41) and (42) can be written as*

$$\mathrm{SINR}_{\chi}^{(\mathrm{NL})} \approx \frac{\left(Nc+2\gamma\right)^2}{K_3 c^2 N} > \mathrm{SINR}_{k_3}^{(\mathrm{NL})} \approx \frac{\left(Nc+\gamma\right)^2}{K_3 c^2 N}. \tag{46}$$

From Corollary 3, we can observe that the STAR-RIS-aided systems outperform conventional RIS-aided systems as $M \to \infty$. The result exhibits that STAR-RIS is appealing for massive MIMO communication than conventional RIS, which is consistent with our analysis in Corollary 1 that the impact of direct links is relatively amplified due to its independence from cascaded channels.

Corollary 4 *When the number of RIS elements is large, namely $N \to \infty$, the SINR in (41) and (42) can be written as*

$$\mathrm{SINR}_{\chi}^{(\mathrm{NL})}=\mathrm{SINR}_{k_3}^{(\mathrm{NL})} \approx \frac{M}{K_3}. \tag{47}$$

The above result shows that when $N \to \infty$, the performance of STAR-RIS is equal to that of conventional RIS, assuming that the same number of total served users. However, since the conventional reflection-only RIS can only serve users in half of the space, the STAR-RIS is expected to serve more users and therefore provide higher sum user rate.

Corollary 5 *Assume that the transmit power p is scaled as $p = E_u/M$. As $M \to \infty$, the rate of user χ is lower bounded by (48) at the top of the next page, with*

$$E_\chi^{\text{leakage}}=c_\chi^2\delta^2\left(N\left|f_\chi(\mathbf{\Phi}_\chi)\right|^2\varepsilon_\chi A_\chi\left(e_{\chi_2}^2+1\right)+N^2 A_\chi^2 e_{\chi_2}^2\right)+\frac{\sigma^2}{\tau E_u}c_\chi\delta e_{\chi_2}^2\left(\varepsilon_\chi\left|f_\chi(\mathbf{\Phi}_\chi)\right|^2+N\right), \tag{49}$$
$$I_{\chi i}=\left|f_\chi(\mathbf{\Phi}_\chi)\right|^2\left|f_i(\mathbf{\Phi}_i)\right|^2 c_\chi c_i\delta^2\varepsilon_\chi\varepsilon_i+A_\chi N\left|f_\chi(\mathbf{\Phi}_\chi)\right|^2 c_\chi c_i\delta^2\varepsilon_\chi+N^2A_\chi^2 c_\chi c_i\delta^2 e_{\chi_2}^2+NA_\chi\frac{\sigma^2}{\tau E_u}c_i\delta e_{\chi_2}^2+\left|f_i(\mathbf{\Phi}_\chi)\right|^2 c_i\delta\varepsilon_i e_{\chi_2}^2\left(A_\chi N c_\chi\delta+\frac{\sigma^2}{\tau E_u}\right), \tag{50}$$
$$I_{\chi\overline{\chi}}=\left|f_\chi(\mathbf{\Phi}_\chi)\right|^2\left|f_{\overline{\chi}}(\mathbf{\Phi}_{\overline{\chi}})\right|^2 c_\chi c_{\overline{\chi}}\delta^2\varepsilon_\chi\varepsilon_{\overline{\chi}}+A_{\overline{\chi}}N\left|f_\chi(\mathbf{\Phi}_\chi)\right|^2 c_\chi c_{\overline{\chi}}\delta^2\varepsilon_\chi+N^2A_\chi A_{\overline{\chi}}c_\chi c_{\overline{\chi}}\delta^2 e_{\chi_2}^2+NA_\chi\frac{\sigma^2}{\tau E_u}c_{\overline{\chi}}\delta e_{\chi_2}^2+\left|f_{\overline{\chi}}(\mathbf{\Phi}_{\overline{\chi}})\right|^2 c_{\overline{\chi}}\delta\varepsilon_{\overline{\chi}} e_{\chi_2}^2\left(A_\chi N c_\chi\delta+\frac{\sigma^2}{\tau E_u}\right), \tag{51}$$
$$e_{\chi_2}=\frac{Nc_\chi\delta}{\frac{\sigma^2}{\tau E_u}+Nc_\chi\delta}. \tag{52}$$

Proof: By replacing p_χ with E_u/M in (28) when $M \to \infty$, we can select the significant terms which are on the order of $\mathcal{O}\left(M\right)$. After some simplification, we complete this proof. ■

Corollary 5 shows that similar to conventional RIS, the rate in a STAR-RIS-aided system remains nonzero, even if the transmit power is scaled with E_u/M. Hence, the power scaling laws still apply to the STAR-RIS-aided system. Different from conventional RIS, the rate in (48) not only depends on the $\mathbf{\Phi}_\chi$ where the user χ is located but also depends on the $\mathbf{\Phi}_{\overline{\chi}}$. It can be seen that almost every term in (48) is related to c_χ and δ. Accordingly, the rate in (48) will degrade to zero when $c_\chi = 0$ or $\delta = 0$. $\delta = 0$ and $c_\chi = 0$ correspond to the scenarios that the RIS-BS channels are full NLoS and the communication

$$\underline{R}_\chi \to \frac{\tau_c - \tau}{\tau_c} \log_2 \left(1 + \frac{E_u c_\chi^2 \delta^2 \left(|f_\chi(\mathbf{\Phi}_\chi)|^2 \varepsilon_\chi + A_\chi N e_{\chi_2} \right)^2}{E_u E_\chi^{\text{leakage}} + E_u \sum_{i=1, i\neq\chi}^{\chi_{\max}} I_{\chi i} + \sigma^2 c_\chi \delta \left(|f_\chi(\mathbf{\Phi}_\chi)|^2 \varepsilon_\chi + N e_{\chi_2} \right) + E_u \sum_{\overline{\chi}=1}^{\overline{\chi}_{\max}} I_{\chi\overline{\chi}}} \right), \tag{48}$$

system is of non-RIS, respectively. The power scaling laws in non-RIS and conventional RIS-aided systems with fully NLOS RIS-BS channels was studied in [25]. To reveal the new feature of the power scaling laws in the considered STAR-RIS system with $\delta = 0$, we first derive its closed-form expression in the next corollary.

Corollary 6 *If the STAR-RIS-BS channel is Rayleigh distributed ($\delta = 0$), the rate of user χ is lower bounded by $R_\chi^{(\text{NL})} = \frac{\tau_c - \tau}{\tau_c} \log_2 \left(1 + \text{SINR}^{(\text{NL})}\right)$, with*

$$\text{SINR}^{(\text{NL})} = \frac{p_\chi E_\chi^{\text{signal}}}{p_\chi E_\chi^{\text{leakage}} + p_\chi \sum_{i=1,i\neq\chi}^{\chi_{\max}} I_{\chi i} + \sigma^2 E_\chi^{\text{noise}} + p_{\overline{\chi}} \sum_{\overline{\chi}=1}^{\overline{\chi}_{\max}} I_{\chi\overline{\chi}}}, \tag{53}$$

where

$$E_\chi^{\text{signal}} = M \left(A_\chi N c_\chi (\varepsilon_\chi + 1) + \gamma_\chi\right)^2 e_{\chi_1}, \tag{54}$$

$$E_\chi^{\text{noise}} = A_\chi N c_\chi (\varepsilon_\chi + 1) + \gamma_\chi, \tag{55}$$

$$\begin{aligned} E_\chi^{\text{leakage}} &= N^2 A_\chi^2 c_\chi^2 (\varepsilon_\chi + 1)^2 e_{\chi_1} + MNA_\chi^2 c_\chi^2 (2\varepsilon_\chi + 1) e_{\chi_1} \\ &+ N c_\chi \left\{ A_\chi^2 c_\chi (2\varepsilon_\chi + 1) + \left(2\gamma_\chi + \frac{\sigma^2}{\tau p_\chi}\right)(\varepsilon_\chi + 1) \right\} e_{\chi_1} \\ &+ \gamma_\chi \left(\gamma_\chi + \frac{\sigma^2}{\tau p_\chi}\right) e_{\chi_1}, \end{aligned} \tag{56}$$

$$\begin{aligned} I_{\chi i} &= N^2 A_\chi^2 c_\chi c_i (\varepsilon_\chi + 1)(\varepsilon_i + 1) e_{\chi_1} + MNA_\chi^2 c_\chi c_i (\varepsilon_\chi + \varepsilon_i \\ &+1) e_{\chi_1} + M A_\chi^2 c_\chi c_i \varepsilon_\chi \varepsilon_i \left| \overline{\mathbf{h}}_\chi^H \overline{\mathbf{h}}_i \right|^2 e_{\chi_1} + \left(\gamma_\chi + \frac{\sigma^2}{\tau p_\chi}\right) \gamma_i e_{\chi_1} \\ &+ N A_\chi \left\{ \left(\gamma_\chi + \frac{\sigma^2}{\tau p_\chi}\right) c_i (\varepsilon_i + 1) + \gamma_i c_\chi (\varepsilon_\chi + 1) \right\} e_{\chi_1}, \end{aligned} \tag{57}$$

$$\begin{aligned} I_{\chi\overline{\chi}} &= N^2 A_\chi A_{\overline{\chi}} c_\chi c_{\overline{\chi}} \left\{ (\varepsilon_\chi + 1)(\varepsilon_{\overline{\chi}} + 1) e_{\chi_1} \right\} + MNA_\chi A_{\overline{\chi}} \\ & c_\chi c_{\overline{\chi}} (\varepsilon_\chi + \varepsilon_{\overline{\chi}} + 1) e_{\chi_1} + \gamma_{\overline{\chi}} \left(\gamma_\chi + \frac{\sigma^2}{\tau p_\chi}\right) e_{\chi_1} \\ &+ M c_\chi c_{\overline{\chi}} \varepsilon_\chi \varepsilon_{\overline{\chi}} \left| \overline{\mathbf{h}}_\chi^H \mathbf{\Phi}_\chi^H \mathbf{\Phi}_{\overline{\chi}} \overline{\mathbf{h}}_{\overline{\chi}} \right|^2 e_{\chi_1} \\ &+ N \left\{ A_\chi \left(\gamma_\chi + \frac{\sigma^2}{\tau p_\chi}\right) c_{\overline{\chi}} (\varepsilon_{\overline{\chi}} + 1) + A_{\overline{\chi}} \gamma_{\overline{\chi}} c_\chi (\varepsilon_\chi + 1) \right\} e_{\chi_1}, \end{aligned} \tag{58}$$

$$e_{\chi_1} = \frac{N\beta\alpha_\chi A_\chi + \gamma_\chi}{N\beta\alpha_\chi A_\chi + \gamma_\chi + \frac{\sigma^2}{\tau p_\chi}}. \tag{59}$$

It is interesting to find that unlike the conventional RIS-aided systems, the rate in Corollary 6 still depends on the phase shifts. Only the case that $\mathbf{\Phi}_\chi$ is equal to $\mathbf{\Phi}_{\overline{\chi}}$ makes the rate independent of the phase shifts. Therefore, the STAR-RIS can still play its role in mitigating the interference under rich-scattering channel conditions. It can be seen that if the transmit power p_χ and $p_{\overline{\chi}}$ is scaled as E_u/M as in Corollary 6, e_1 tends to $\mathcal{O}(1/M)$ when $M \to \infty$ and therefore the rate degrades to zero in this scenario. Hence, we will investigate the power scaling law under this rich-scattering channel condition where the transmit power p_χ and $p_{\overline{\chi}}$ are scaled as $E_u/\sqrt{M}$ and E_u/N, respectively.

Corollary 7 *If the STAR-RIS-BS channel is Rayleigh distributed ($\delta = 0$), and the power is scaled as $p_\chi = p_{\overline{\chi}} = E_u/\sqrt{M}$ with $M \to \infty$, the rate of user χ tends to $R_\chi^{(\text{NL}_1)} = \frac{\tau_c - \tau}{\tau_c} \log_2 \left(1 + \text{SINR}^{(\text{NL}_1)}\right)$, with*

$$\text{SINR}^{(\text{NL}_1)} = \frac{E_u E_\chi^{\text{signal}}}{E_u E_\chi^{\text{leakage}} + E_u \sum_{i=1,i\neq\chi}^{\chi_{\max}} I_{\chi i} + \sigma^4 + E_u \sum_{\overline{\chi}=1}^{\overline{\chi}_{\max}} I_{\chi\overline{\chi}}}, \tag{60}$$

where

$$E_\chi^{\text{signal}} = \tau E_u \left(A_\chi N c_\chi (\varepsilon_\chi + 1) + \gamma_\chi\right)^2, \tag{61}$$

$$E_\chi^{\text{leakage}} = \tau E_u N A_\chi^2 c_\chi^2 (2\varepsilon_\chi + 1), \tag{62}$$

$$I_{\chi i} = \tau E_u A_\chi^2 c_\chi c_i \left\{ N (\varepsilon_\chi + \varepsilon_i + 1) + \varepsilon_\chi \varepsilon_i \left| \overline{\mathbf{h}}_\chi^H \overline{\mathbf{h}}_i \right|^2 \right\}, \tag{63}$$

$$\begin{aligned} I_{\chi\overline{\chi}} &= \tau E_u c_\chi c_{\overline{\chi}} \left\{ A_\chi A_{\overline{\chi}} N (\varepsilon_\chi + \varepsilon_{\overline{\chi}} + 1) + \varepsilon_\chi \varepsilon_{\overline{\chi}} \right. \\ & \left. \left| \overline{\mathbf{h}}_\chi^H \mathbf{\Phi}_\chi^H \mathbf{\Phi}_{\overline{\chi}} \overline{\mathbf{h}}_{\overline{\chi}} \right|^2 \right\}, \end{aligned} \tag{64}$$

Proof: It follows by substituting $p_\chi = p_{\overline{\chi}} = E_u/\sqrt{M}$ into Corollary 7 and neglecting the terms that tend to zero as $M \to \infty$. ■

It can be seen that nearly every term in (60) is proportional to N, except the signal term which is proportional to N^2. It indicates that the performance of STAR-RIS-aided systems will greatly increase as RIS elements increases, which is a promising feature considering the low cost of the RIS hardware. In the following, we will reveal the power scaling law when the transmit power is scaled as E_u/N.

Corollary 8 *If the STAR-RIS-BS channel is Rayleigh distributed ($\delta = 0$) and the power is scaled as $p_\chi = p_{\overline{\chi}} = E_u/N$ with $N \to \infty$, the rate of user χ is lower bounded by $R_\chi^{(\text{NL}_2)} = \frac{\tau_c - \tau}{\tau_c} \log_2 \left(1 + \text{SINR}^{(\text{NL}_2)}\right)$, with*

$$\text{SINR}^{(\text{NL}_2)} = \frac{E_u M A_\chi \beta \alpha_\chi}{\sum_{i=1}^{\chi_{\max}} E_u A_\chi \beta \alpha_i + \sigma^2 \left(1 + \frac{\sigma^2}{\tau A_\chi E_u \beta \alpha_\chi}\right) + \sum_{\overline{\chi}=1}^{\overline{\chi}_{\max}} E_u A_{\overline{\chi}} \beta \alpha_{\overline{\chi}}}. \tag{65}$$

Similar to Corollary 7, Corollary 8 proves that the rate in STAR-RIS-aided systems will maintain nonzero when transmit power is reduced as E_u/N and $\delta = 0$. Secondly, the power scaling law is not influenced by ε_χ, indicating that the power scaling laws are robust to the scattering environment (Los/NLoS) of the users-STAR-RIS channels. Finally, it can be seen that (65) is in proportion to M since the numerator of (65) is on the order of $\mathcal{O}(1/M)$ while its denominator is a constant. It indicates that the performance of STAR-RIS-aided systems will greatly increase as the number of BS antennas increases, which is feasible for the massive MIMO structure.

V. PHASE SHIFTS DESIGN

In this section, we will design the phase shifts of STAR-RIS based on the long-term statistical CSI. For a comprehensive investigation, we formulate the sum user rate-oriented optimization problem. Specifically, this optimization problem is formulated as

$$\max_{\mathbf{\Phi_R},\mathbf{\Phi_T}} \sum_{\chi=1}^{\chi_{\max}} R_\chi + \sum_{\overline{\chi}=1}^{\overline{\chi}_{\max}} R_{\overline{\chi}} \triangleq R_{\text{sum}}, \tag{66a}$$

$$\text{s.t.} \quad \theta_n^R, \theta_n^T \in [0,2\pi), \forall n. \tag{66b}$$

To deal with this problem, we propose an accelerated gradient ascent-based algorithm to solve the optimization problem. To begin with, we assume that $\chi = k_1$, $\overline{\chi} = k_2$, and introduce several auxiliary vectors such as $\boldsymbol{\theta}_\chi = \boldsymbol{\theta}_R = [\theta_1^R, \theta_2^R, \ldots, \theta_N^R]^T$, $\boldsymbol{\theta}_{\overline{\chi}} = \boldsymbol{\theta}_T = [\theta_1^T, \theta_2^T, \ldots, \theta_N^T]^T$, and $\boldsymbol{z}_\chi = \boldsymbol{z}_R = [e^{j\theta_1^R}, e^{j\theta_2^R}, \ldots, e^{j\theta_N^R}]^T$, $\boldsymbol{z}_{\overline{\chi}} = \boldsymbol{z}_T = [e^{j\theta_1^T}, e^{j\theta_2^T}, \ldots, e^{j\theta_N^T}]^T$, so that $\boldsymbol{z}_R = e^{j\boldsymbol{\theta}_R}$, $\boldsymbol{z}_T = e^{j\boldsymbol{\theta}_T}$ and $\mathbf{\Phi}_\chi = \mathbf{\Phi}_R = A_R \text{diag}(\boldsymbol{z}_R)$, $\mathbf{\Phi}_{\overline{\chi}} = \mathbf{\Phi}_T = A_T \text{diag}(\boldsymbol{z}_T)$. It is worth noting that (66b) does not need to be considered due to the periodicity of the objective functions with respect to $\boldsymbol{\theta}_R$, $\boldsymbol{\theta}_T$. Since the partial derivatives of R_{sum} with respect to $\boldsymbol{\theta}_R$ and $\boldsymbol{\theta}_T$ are similar, we only expand the derivations of $\frac{\partial R_{\text{sum}}}{\partial \boldsymbol{\theta}_{\overline{\chi}}}$ as follows:

$$\frac{\partial R_{\text{sum}}}{\partial \boldsymbol{\theta}_\chi} = \frac{\frac{\tau_c - \tau}{\tau_c}}{\ln 2} \left(\sum_{\chi=1}^{\chi_{\max}} \frac{\frac{\partial \text{SINR}_\chi}{\partial \boldsymbol{\theta}_\chi}}{1+\text{SINR}_\chi} + \sum_{\overline{\chi}=1}^{\overline{\chi}_{\max}} \frac{\frac{\partial \text{SINR}_{\overline{\chi}}}{\partial \boldsymbol{\theta}_\chi}}{1+\text{SINR}_{\overline{\chi}}} \right). \tag{67}$$

and $\frac{\partial \text{SINR}_{\overline{\chi}}}{\partial \boldsymbol{\theta}_\chi}$, $\frac{\partial \text{SINR}_{\overline{\chi}}}{\partial \boldsymbol{\theta}_{\overline{\chi}}}$ are respectively at top of the next page, where $\frac{\partial E_\chi^{\text{signal}}}{\partial \boldsymbol{\theta}_\chi}$, $\frac{\partial E_\chi^{\text{leak}}}{\partial \boldsymbol{\theta}_\chi}$, $\frac{\partial I_{\chi i}}{\partial \boldsymbol{\theta}_\chi}$, $\frac{\partial I_{\chi\overline{\chi}}}{\partial \boldsymbol{\theta}_\chi}$, $\frac{\partial E_\chi^{\text{noise}}}{\partial \boldsymbol{\theta}_\chi}$ and $\frac{\partial I_{\overline{\chi}\chi}}{\partial \boldsymbol{\theta}_\chi}$ can respectively be deduced as

$$\frac{\partial |f_\chi(\mathbf{\Phi}_\chi)|^2}{\partial \boldsymbol{\theta}_\chi} = 2\,\text{Im}\left\{ \mathbf{\Phi}_\chi^H \left(\mathbf{a}_N \mathbf{a}_N^H \odot \left(\overline{\mathbf{h}}_\chi \overline{\mathbf{h}}_\chi^H \right)^T \right) \boldsymbol{z}_\chi \right\}, \tag{70}$$

$$\frac{\partial |f_{\overline{\chi}}(\mathbf{\Phi}_{\overline{\chi}})|^2}{\partial \boldsymbol{\theta}_\chi} = 2\,\text{Im}\left\{ \mathbf{\Phi}_{\overline{\chi}}^H \left(\mathbf{a}_N \mathbf{a}_N^H \odot \left(\overline{\mathbf{h}}_{\overline{\chi}} \overline{\mathbf{h}}_{\overline{\chi}}^H \right)^T \right) \boldsymbol{z}_{\overline{\chi}} \right\}, \tag{71}$$

$$\frac{\partial |f_i(\mathbf{\Phi}_i)|^2}{\partial \boldsymbol{\theta}_\chi} = 2\,\text{Im}\left\{ \mathbf{\Phi}_\chi^H \left(\mathbf{a}_N \mathbf{a}_N^H \odot \left(\overline{\mathbf{h}}_i \overline{\mathbf{h}}_i^H \right)^T \right) \boldsymbol{z}_\chi \right\}, \tag{72}$$

$$\frac{\partial f_\chi^H(\mathbf{\Phi}_\chi) f_i(\mathbf{\Phi}_i)|}{\partial \boldsymbol{\theta}_\chi} = f_d \left(\mathbf{a}_N \mathbf{a}_N^H, \overline{\mathbf{h}}_i \overline{\mathbf{h}}_\chi^H \right), \tag{73}$$

Algorithm 1 Accelerated Gradient Ascent Algorithm

Initialize $\boldsymbol{\theta}_0$ randomly, $n=0$, $y_0 = 1$, $\boldsymbol{x}_{-1} = \boldsymbol{\theta}_0$;
while 1 **do**
  Calculate the gradient vector $\boldsymbol{f}_\chi'(\boldsymbol{\theta}_n) = \frac{\partial f_\chi(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_i}$;
  Obtain the step size κ_n based on the backtracking line search;
  $\boldsymbol{x}_n = \boldsymbol{\theta}_n + \varrho_n \boldsymbol{f}_\chi'(\boldsymbol{\theta}_n)$;
  $y_{n+1} = (1+\sqrt{4y_n^2+1})/2$;
  $\boldsymbol{\theta}_{n+1} = \boldsymbol{x}_i + (y_n - 1)(\boldsymbol{x}_n - \boldsymbol{x}_{n-1})/y_{n+1}$;
  if $f_\chi(\boldsymbol{\theta}_{n+1}) - f_\chi(\boldsymbol{\theta}_n) < 10^{-4}$ **then**
    $\boldsymbol{\theta}^* = \boldsymbol{\theta}_{n+1}$, break;
  end if
  $n = n+1$;
end while

$$\frac{\partial f_i^H(\mathbf{\Phi}_i) f_\chi(\mathbf{\Phi}_\chi)|}{\partial \boldsymbol{\theta}_\chi} = f_d \left(\mathbf{a}_N \mathbf{a}_N^H, \overline{\mathbf{h}}_\chi \overline{\mathbf{h}}_i^H \right), \tag{74}$$

$$\frac{\partial E_\chi^{\text{noise}}}{\partial \boldsymbol{\theta}_\chi} = 2Mc_\chi \delta \varepsilon_\chi \frac{\partial |f_\chi(\mathbf{\Phi}_\chi)|^2}{\partial \boldsymbol{\theta}_\chi}, \tag{75}$$

$$\frac{\partial E_\chi^{\text{signal}}}{\partial \boldsymbol{\theta}_\chi} = 2E_\chi^{\text{noise}} \frac{\partial E_\chi^{\text{noise}}}{\partial \boldsymbol{\theta}_\chi}, \tag{76}$$

$$\begin{aligned}
\frac{\partial I_{\chi i}}{\partial \boldsymbol{\theta}_\chi} &= M^2 c_\chi c_i \delta^2 \varepsilon_\chi \varepsilon_i \left(\frac{|f_\chi(\mathbf{\Phi}_\chi)|^2}{\partial \boldsymbol{\theta}_\chi} |f_i(\mathbf{\Phi}_i)|^2 + |f_\chi(\mathbf{\Phi}_\chi)|^2 \right. \\
&\left. \frac{|f_i(\mathbf{\Phi}_i)|^2}{\partial \boldsymbol{\theta}_\chi} \right) + Mc_\chi \delta \varepsilon_\chi \left\{ A_\chi c_i \left(MN\delta + N\varepsilon_i + N + 2Me_{\chi_1} \right) \right. \\
&\left. + \gamma_i \right\} \frac{|f_\chi(\mathbf{\Phi}_\chi)|^2}{\partial \boldsymbol{\theta}_\chi} + Mc_i \delta \varepsilon_i \left\{ A_\chi c_\chi e_{\chi_2} \left(MN\delta e_{\chi_2} + N\varepsilon_\chi e_{\chi_2} \right. \right. \\
&\left. \left. + Ne_{\chi_2} + 2Me_{\chi_1} \right) + \left(\gamma_\chi + \frac{\sigma^2}{\tau p_\chi} \right) e_{\chi_2}^2 \right\} \frac{|f_i(\mathbf{\Phi}_i)|^2}{\partial \boldsymbol{\theta}_\chi} \\
&+ A_\chi M^2 c_\chi c_i \varepsilon_\chi \varepsilon_i e_{\chi_1} \delta \overline{\mathbf{h}}_i^H \overline{\mathbf{h}}_\chi \frac{\partial f_\chi^H(\mathbf{\Phi}_\chi) f_i(\mathbf{\Phi}_i)|}{\partial \boldsymbol{\theta}_\chi} \\
&+ A_\chi M^2 c_\chi c_i \varepsilon_\chi \varepsilon_i e_{\chi_1} \delta \overline{\mathbf{h}}_\chi^H \overline{\mathbf{h}}_i \frac{\partial f_i^H(\mathbf{\Phi}_i) f_\chi(\mathbf{\Phi}_\chi)|}{\partial \boldsymbol{\theta}_\chi}.
\end{aligned} \tag{77}$$

The remaining terms $\frac{\partial I_{\chi\overline{\chi}}}{\partial \boldsymbol{\theta}_\chi}$ and $\frac{\partial I_{\overline{\chi}\chi}}{\partial \boldsymbol{\theta}_\chi}$ can be derived similarly to (77). In the following, we show the complete algorithm steps in Algorithm 1. It is worth noting that we adopt Nesterovs accelerated gradient method here to speed up the convergence of the gradient algorithm.

VI. NUMERICAL RESULTS

In this section, we provide numerical results to verify our analysis and characterize the impact of deploying STAR-RIS into massive MIMO systems with imperfect CSI. At the same time, we also simulate the performance of conventional RIS-aided systems under the same scenario. Unless stated otherwise, the parameters are set to $M = N = 64$, $\sigma^2 = -104$ dBm, number of users in the reflection region and transmission region of $K_1 = K_2 = 4$, accordingly $K_3 = K_1 + K_2 = 8$, the Rician factors of $\varepsilon_\chi = \varepsilon_{k_3} = 10$, $\delta_\chi = \delta_{k_3} = 1$, transmit power of $p_\chi = p_{k_3} = 0$ dBm, $\forall \chi, k_3$. Eight users are located on a circle with a radius of $d_{\text{UI}} = 20$m, which is centered at the STAR-RIS. STAR-RIS-BS distance

$$\frac{\partial\,\mathrm{SINR}_\chi}{\partial\boldsymbol{\theta}_\chi}=\frac{p_\chi\frac{\partial E_\chi^{\mathrm{signal}}}{\partial\boldsymbol{\theta}_\chi}}{p_\chi E_\chi^{\mathrm{leak}}+p_\chi\sum\limits_{\substack{i=1,\\ i\neq\chi}}^{\chi_{\max}}I_{\chi i}+p_{\overline{\chi}}\sum\limits_{\overline{\chi}=1}^{\overline{\chi}_{\max}}I_{\chi\overline{\chi}}+\sigma^2E_\chi^{\mathrm{noise}}}-\frac{p_\chi E_\chi^{\mathrm{signal}}\left(p_\chi\frac{\partial E_\chi^{\mathrm{leak}}}{\partial\boldsymbol{\theta}_\chi}+p_\chi\sum\limits_{\substack{i=1,\\ i\neq\chi}}^{K}\frac{\partial I_{\chi i}}{\partial\boldsymbol{\theta}_\chi}+p_{\overline{\chi}}\sum\limits_{\overline{\chi}=1}^{\overline{\chi}_{\max}}\frac{\partial I_{\chi\overline{\chi}}}{\partial\boldsymbol{\theta}_\chi}+\sigma^2\frac{\partial E_\chi^{\mathrm{noise}}}{\partial\boldsymbol{\theta}_\chi}\right)}{\left(p_\chi E_\chi^{\mathrm{leak}}+p_\chi\sum\limits_{\substack{i=1,\\ i\neq\chi}}^{\chi_{\max}}I_{\chi i}+p_{\overline{\chi}}\sum\limits_{\overline{\chi}=1}^{\overline{\chi}_{\max}}I_{\chi\overline{\chi}}+\sigma^2E_\chi^{\mathrm{noise}}\right)^2},\tag{68}$$

$$\frac{\partial\,\mathrm{SINR}_\chi}{\partial\boldsymbol{\theta}_{\overline{\chi}}}=\frac{-p_{\overline{\chi}}p_\chi E_{\overline{\chi}}^{\mathrm{signal}}\sum\limits_{\chi=1}^{\chi_{\max}}\frac{\partial I_{\overline{\chi}\chi}}{\partial\boldsymbol{\theta}_{\overline{\chi}}}}{\left(p_{\overline{\chi}}E_{\overline{\chi}}^{\mathrm{leak}}+p_{\overline{\chi}}\sum\limits_{\substack{i=1,\\ i\neq\overline{\chi}}}^{\overline{\chi}_{\max}}I_{\overline{\chi}i}+p_\chi\sum\limits_{\chi=1}^{\chi_{\max}}I_{\overline{\chi}\chi}+\sigma^2E_{\overline{\chi}}^{\mathrm{noise}}\right)^2}.\tag{69}$$

is $d_{\mathrm{IB}}=1000$m, and the distance between users and BS is set to $\left(d_\chi^{\mathrm{UB}}\right)^2=\left(d_{\mathrm{IB}}-d_{\mathrm{UI}}\sin(\frac{\pi}{9}\chi)\right)^2+\left(d_{\mathrm{UI}}\cos(\frac{\pi}{9}\chi)\right)^2$. The AoA and AoD are all generated randomly from $[0,2\pi]$. The large-scale pathloss are set to $\alpha_\chi=\alpha_{k_3}=10^{-3}d_{\mathrm{UI}}^{-2}$, $\beta=10^{-3}d_{\mathrm{IB}}^{-2.1}$ and $\gamma_\chi=\gamma_{k_3}=10^{-3}\left(d_\chi^{\mathrm{UB}}\right)^{-4}$, $\forall\chi,k_3$. In order to divide the energy equally between the reflection and transmission regions, we set $A_R=A_T=\frac{1}{2}$.

Fig. 2 illustrates the relation between the number of BS antennas and the rates. Firstly, we can infer from this figure that the system with STAR-RIS has better performance than conventional RIS. Besides, we also add a sample with $N=16$ to show that increasing N significantly increases the performance of systems as M approaches infinity. If conventional RIS intends to achieve the same performance as STAR-RIS, it needs to increase the number of antennas which requires larger-sized array, higher power consumption, and higher hardware cost. As shown in Fig. 2, we can find that 50 antennas with 64 STAR-RIS elements can be equivalent to 120 antennas with 16 STAR-RIS elements. It means that STAR-RIS-aided massive MIMO systems can reduce the cost of BS antennas and power consumption by increasing the number of RIS elements, which is promising to be applied to future communication systems. These two points fit perfectly with Corollary 3.

Fig. 3 shows that the relation between the user rate and the number of RIS elements respectively in Rician channel and NLOS cascaded channels. It can be seen that with the increase of N, the sum rate of RIS and STAR-RIS increase rapidly, and it will converge to a constant. Besides, it is consistent with Corollary 4 that the performance of RIS and STAR-RIS will be equal when the number of RIS elements tend to be infinite and the cascaded channels are fully NLOS. It can be seen from this figure that the performance of STAR-RIS and RIS in Rician channel are also similar.

In Fig. 4 and Fig. 5, the power scaling laws are studied in Rician channel and NLOS BS-RIS channel respectively. They validate the analytical results in Corollary 5, Corollary 7 and Corollary 8. The results show that the rate will maintain a constant value even if the transmit power is scaled as $100/M$, $100/\sqrt{M}$, or $100/N$, which indicate that the power scaling laws are still applicable to STAR-RIS. It can be seen in Fig. 5 that even if the rate tends to be constant when RIS elements N tend to be infinite, increasing M can greatly improve the rate. It is consistent with out analysis in Corollary 8. At the same time, these results validate the feasibility of applying the gradient ascent method.

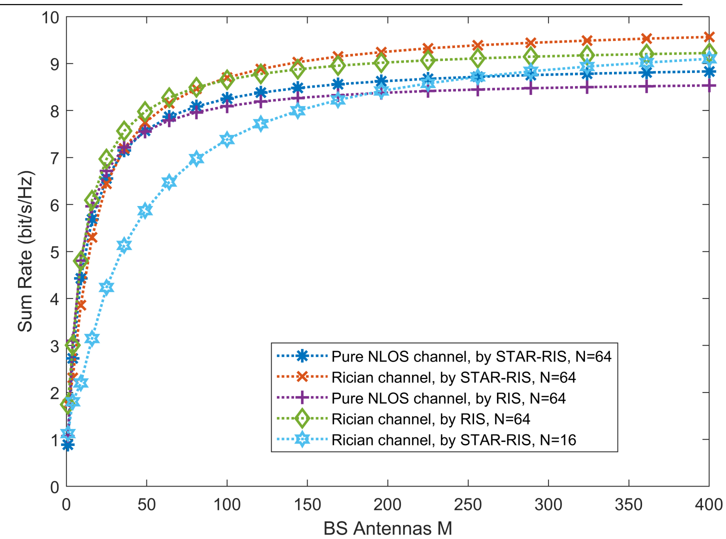

Fig. 2: Rate versus the number of BS antennas M.

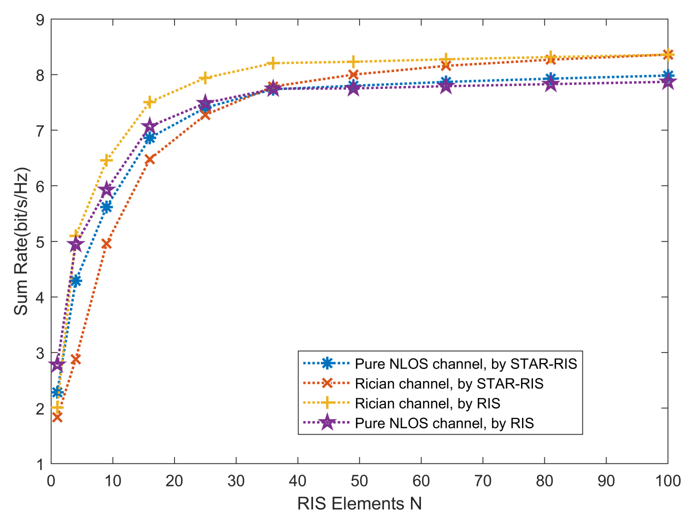

Fig. 3: Rate versus the number of RIS elements N.

VII. CONCLUSION

This paper has investigated the closed-form achievable rates of STAR-RIS-aided massive MIMO systems over Rician fading channels with imperfect CSI. Based on the derived

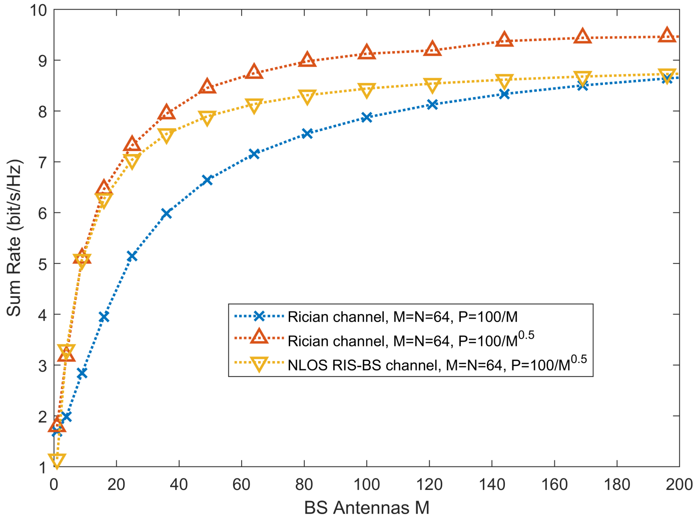

Fig. 4: Rate versus the number of BS antennas M, with scaled transmit power.

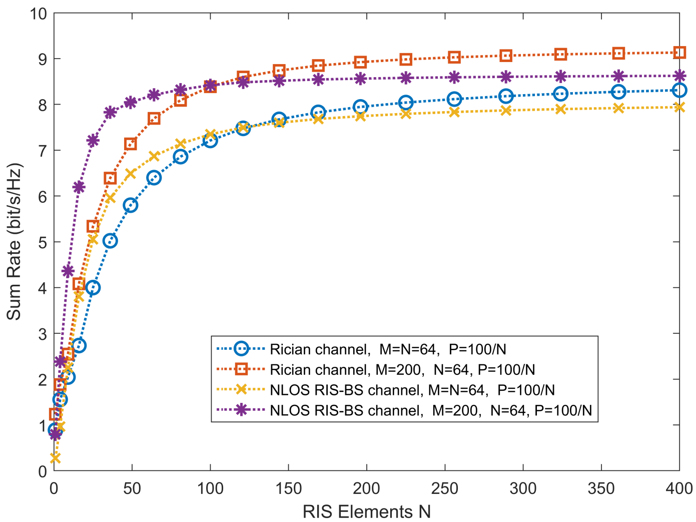

Fig. 5: Rate versus the number of RIS elements N, with scaled transmit power.

expressions, we analyzed the power scaling laws in a general case, and used some simple cases to shed light on the condition when systems with STAR-RIS outperform those with reflection-only RIS. Besides, gradient ascent-based method is adopted to optimize sum rate and minimum user rate maximization problem. Finally, extensive simulation results validated our derivations and showed the various advantages of deploying STAR-RIS.

APPENDIX A

Since $\widetilde{\mathbf{H}}$, $\widetilde{\mathbf{h}}_\chi$, $\widetilde{\mathbf{d}}_\chi$ and N are independent of each other, we have

$$\mathbb{E}\left\{\mathbf{q}_\chi\right\} = \sqrt{c_\chi \delta \varepsilon_\chi}\overline{\mathbf{H}}\mathbf{\Phi}_\chi \overline{\mathbf{h}}_\chi, \tag{78}$$

$$\mathbb{E}\left\{\mathbf{y}_{p_\chi}^\chi\right\} = \mathbb{E}\left\{\mathbf{q}_\chi\right\} + \frac{1}{\sqrt{\tau p_\chi}}\mathbb{E}\left\{\mathbf{N}\mathbf{s}_\chi^H\right\} = \mathbb{E}\left\{\mathbf{q}_\chi\right\}. \tag{79}$$

The covariance matrix between the unknown channel $\mathbf{q}_\chi$ and the observation vector $\mathbf{y}_{p_\chi}^\chi$ can be written as

$$\begin{aligned}
&\mathrm{Cov}\left\{\mathbf{q}_\chi, \mathbf{y}_{p_\chi}^\chi\right\} = \mathbb{E}\left\{\left(\mathbf{q}_\chi - \mathbb{E}\left\{\mathbf{q}_\chi\right\}\right)\left(\mathbf{y}_{p_\chi}^\chi - \mathbb{E}\left\{\mathbf{y}_{p_\chi}^\chi\right\}\right)^H\right\} \\
&= \mathbb{E}\left\{\left(\mathbf{q}_\chi - \mathbb{E}\left\{\mathbf{q}_\chi\right\}\right)\left(\mathbf{q}_\chi + \frac{1}{\sqrt{\tau p_\chi}}\mathbf{N}\mathbf{s}_\chi - \mathbb{E}\left\{\mathbf{q}_\chi\right\}\right)^H\right\} \\
&= \mathbb{E}\left\{\left(\mathbf{q}_\chi - \mathbb{E}\left\{\mathbf{q}_\chi\right\}\right)\left(\mathbf{q}_\chi - \mathbb{E}\left\{\mathbf{q}_\chi\right\}\right)^H\right\} \\
&= \mathrm{Cov}\left\{\mathbf{q}_\chi, \mathbf{q}_\chi\right\},
\end{aligned} \tag{80}$$

$$\begin{aligned}
&\mathrm{Cov}\left\{\mathbf{y}_p^\chi, \mathbf{q}_\chi\right\} = \left(\mathrm{Cov}\left\{\mathbf{q}_\chi, \mathbf{y}_p^\chi\right\}\right)^H \\
&= \left(\mathrm{Cov}\left\{\mathbf{q}_\chi, \mathbf{q}_\chi\right\}\right)^H = \mathrm{Cov}\left\{\mathbf{q}_\chi, \mathbf{q}_\chi\right\}, \\
&= \mathbb{E}\left\{c_\chi \delta \overline{\mathbf{H}}\mathbf{\Phi}_\chi \tilde{\mathbf{h}}_\chi \tilde{\mathbf{h}}_\chi^H \mathbf{\Phi}_\chi^H \overline{\mathbf{H}}^H + c_k \varepsilon_k \tilde{\mathbf{H}}\mathbf{\Phi}_\chi \overline{\mathbf{h}}_\chi \overline{\mathbf{h}}_\chi^H \mathbf{\Phi}_\chi^H \tilde{\mathbf{H}}^H\right. \\
&\left. + c_\chi \tilde{\mathbf{H}}\mathbf{\Phi}_\chi \tilde{\mathbf{h}}_\chi \tilde{\mathbf{h}}_\chi^H \mathbf{\Phi}_\chi^H \tilde{\mathbf{H}}^H + \gamma_\chi \tilde{\mathbf{d}}_\chi \tilde{\mathbf{d}}_\chi^H\right\} \\
&= N c_\chi \delta A_\chi \mathbf{a}_M \mathbf{a}_M^H + \left(N c_\chi A_\chi \left(\varepsilon_\chi + 1\right) + \gamma_k\right)\mathbf{I}_M,
\end{aligned} \tag{81}$$

$$\begin{aligned}
&\mathrm{Cov}\left\{\mathbf{y}_{p_\chi}^\chi, \mathbf{y}_{p_\chi}^\chi\right\} = \mathbb{E}\left\{\left(\mathbf{q}_\chi - \mathbb{E}\left\{\mathbf{q}_\chi\right\}\right)\left(\mathbf{q}_\chi - \mathbb{E}\left\{\mathbf{q}_\chi\right\}\right)^H\right\} \\
&+ \frac{1}{\tau p_\chi}\mathbb{E}\left\{\mathbf{N}\mathbf{s}_\chi \mathbf{s}_\chi^H \mathbf{N}^H\right\} = \mathrm{Cov}\left\{\mathbf{q}_\chi, \mathbf{q}_\chi\right\} + \frac{\sigma^2}{\tau p_\chi}\mathbf{I}_M.
\end{aligned} \tag{82}$$

Finally, we introduce the auxiliary variables of $a_{\chi_1} = N c_\chi A_\chi \delta$ and $a_{\chi_2} = N c_\chi A_\chi \left(\varepsilon_\chi + 1\right) + \gamma_\chi$.

APPENDIX B

The LMMSE estimate of the channel $\mathbf{q}_\chi$ based on the observation vector $\mathbf{y}_{p_\chi}^\chi$ can be written as

$$\begin{aligned}
\hat{\mathbf{q}}_\chi = &\mathbb{E}\left\{\mathbf{q}_\chi\right\} + \mathrm{Cov}\left\{\mathbf{q}_\chi, \mathbf{y}_{p_\chi}^\chi\right\} \mathrm{Cov}^{-1}\left\{\mathbf{y}_{p_\chi}^\chi, \mathbf{y}_{p_\chi}^\chi\right\} \times \\
&\left(\mathbf{y}_{p_\chi}^\chi - \mathbb{E}\left\{\mathbf{y}_{p_\chi}^\chi\right\}\right),
\end{aligned} \tag{83}$$

where the mean and covariance matrices have been obtained in Appendix A.

The remaining term in (83) is derived as

$$\begin{aligned}
&\mathrm{Cov}\left\{\mathbf{q}_\chi, \mathbf{y}_{p_\chi}^\chi\right\} \mathrm{Cov}^{-1}\left\{\mathbf{y}_{p_\chi}^\chi, \mathbf{y}_{p_\chi}^\chi\right\} \\
&= \left(a_{\chi_1}\mathbf{a}_M \mathbf{a}_M^H + \left(a_{\chi_2} + \frac{\sigma^2}{\tau p_\chi}\right)\mathbf{I}_M\right)^{-1}\left(a_{\chi_1}\mathbf{a}_M \mathbf{a}_M^H + a_{\chi_2}\mathbf{I}_M\right) \\
&= \left\{\left(a_{\chi_2} + \frac{\sigma^2}{\tau p_\chi}\right)^{-1}\mathbf{I}_M - \frac{a_{\chi_1}\left(a_{\chi_2} + \frac{\sigma^2}{\tau p_\chi}\right)^{-2}}{1 + M a_{\chi_1}\left(a_{\chi_2} + \frac{\sigma^2}{\tau p_\chi}\right)^{-1}}\mathbf{a}_M \mathbf{a}_M^H\right\} \\
&\times \left(a_{\chi_1}\mathbf{a}_M \mathbf{a}_M^H + a_{\chi_2}\mathbf{I}_M\right) \\
&= \frac{a_{\chi_1}\frac{\sigma^2}{\tau p_\chi}\mathbf{a}_M \mathbf{a}_M^H}{\left(a_{\chi_2} + \frac{\sigma^2}{\tau p_\chi}\right)\left\{\left(a_{\chi_2} + \frac{\sigma^2}{\tau p_\chi}\right) + M a_{\chi_1}\right\}} + \frac{a_{\chi_2}}{a_{\chi_2} + \frac{\sigma^2}{\tau p_\chi}}\mathbf{I}_M \\
&\triangleq a_{\chi_3}\mathbf{a}_M \mathbf{a}_M^H + a_{\chi_4}\mathbf{I}_M \triangleq \mathbf{A}_\chi = \mathbf{A}_\chi^H.
\end{aligned} \tag{84}$$

Hence, the LMMSE channel estimate in (83) is calculated as

$$\begin{aligned}
\hat{\mathbf{q}}_\chi &= \sqrt{c_\chi \delta \varepsilon_\chi}\,\overline{\mathbf{H}}\mathbf{\Phi}_\chi \overline{\mathbf{h}}_\chi + \mathbf{A}_\chi\left(\mathbf{y}_{p_\chi}^\chi - \sqrt{c_\chi \delta \varepsilon_\chi}\,\overline{\mathbf{H}}\mathbf{\Phi}_\chi \overline{\mathbf{h}}_\chi\right) \\
&= \mathbf{A}_\chi \mathbf{y}_{p_\chi}^\chi + \left(\mathbf{I}_M - \mathbf{A}_\chi\right)\sqrt{c_\chi \delta \varepsilon_\chi}\,\overline{\mathbf{H}}\mathbf{\Phi}_\chi \overline{\mathbf{h}}_\chi \\
&\triangleq \mathbf{A}_\chi \mathbf{y}_{p_\chi}^\chi + \mathbf{B}_\chi
\end{aligned} \tag{85}$$

$$=\mathbf{q}_\chi^1+\sum\nolimits_{\omega=2}^4\mathbf{A}_\chi\mathbf{q}_\chi^\omega+\sqrt{\gamma_\chi}\mathbf{A}_\chi\tilde{\mathbf{d}}_\chi+\frac{1}{\sqrt{\tau p_\chi}}\mathbf{A}_\chi\mathbf{N}\mathbf{s}_\chi^H.$$

Finally, by exploiting the property $\mathbf{A}_\chi\overline{\mathbf{H}}=\left(a_{\chi_3}\mathbf{a}_M\mathbf{a}_M^H+a_{\chi_4}\mathbf{I}_M\right)\mathbf{a}_M\mathbf{a}_N^H=\left(Ma_{\chi_3}+a_{\chi_4}\right)\overline{\mathbf{H}}$, the proof is completed.

Appendix C

A. Signal Term and Noise Term

Since $\mathbf{e}_\chi$ is independent of the channel estimation, we obtain $\mathbb{E}\left\{\hat{\mathbf{q}}_\chi^H\mathbf{e}_\chi\right\}=0$. Therefore, we have

$$\mathbb{E}\left\{\hat{\mathbf{q}}_\chi^H\mathbf{q}_\chi\right\}=\mathbb{E}\left\{\hat{\mathbf{q}}_\chi^H\hat{\mathbf{q}}_\chi\right\}+\mathbb{E}\left\{\hat{\mathbf{q}}_\chi^H\mathbf{e}_\chi\right\}=\mathbb{E}\left\{\|\hat{\mathbf{q}}_\chi\|^2\right\}.\tag{86}$$

Hence, the signal power can be transformed to

$$E_\chi^{\text{signal}}=\left|\mathbb{E}\left\{\hat{\mathbf{q}}_\chi^H\mathbf{q}_\chi\right\}\right|^2=\left(\mathbb{E}\left\{\|\hat{\mathbf{q}}_\chi\|^2\right\}\right)^2=\left(E_\chi^{\text{noise}}\right)^2.\tag{87}$$

We can derive the term $\mathbb{E}\left\{\hat{\mathbf{q}}_\chi^H\mathbf{q}_\chi\right\}$ by selecting the non-zero terms in the expansion as

$$\begin{aligned}&E_\chi^{\text{noise}}=\mathbb{E}\left\{\|\hat{\mathbf{q}}_\chi\|^2\right\}=\mathbb{E}\left\{\hat{\mathbf{q}}_\chi^H\mathbf{q}_\chi\right\}\\&=\sum_{\omega=1}^4\mathbb{E}\left\{\left(\hat{\mathbf{g}}_\chi^\omega\right)^H\mathbf{g}_\chi^\omega\right\}+\gamma_\chi\mathbb{E}\left\{\tilde{\mathbf{d}}_\chi^H\mathbf{A}_\chi^H\tilde{\mathbf{d}}_\chi\right\}\\&=c_\chi\delta\varepsilon_\chi M\left|f_\chi(\mathbf{\Phi}_\chi)\right|^2+A_\chi c_\chi\delta MNe_{\chi_2}+A_\chi c_\chi\varepsilon_\chi MNe_{\chi_1}\\&+A_\chi c_\chi MNe_{\chi_1}+\gamma_\chi Me_{\chi_1}\\&=M\left\{\left|f_\chi(\mathbf{\Phi}_\chi)\right|^2c_\chi\delta\varepsilon_\chi+A_\chi Nc_\chi\delta e_{\chi_2}+\left(A_\chi Nc_\chi\left(\varepsilon_\chi+1\right)\right.\right.\\&\left.\left.+\ \gamma_\chi\right)e_{\chi_1}\right\}.\end{aligned}\tag{88}$$

We conclude this subsection by providing some useful results similar to (88). To be specific, we aim to derive $\mathbb{E}\left\{\mathbf{q}_\chi^H\mathbf{q}_\chi\right\}$, $\mathbb{E}\left\{\mathbf{g}_\chi^H\mathbf{g}_\chi\right\}$, $\mathbb{E}\left\{\mathbf{g}_\chi^H\mathbf{A}_\chi\mathbf{A}_\chi^H\mathbf{g}_\chi\right\}$, $\mathbb{E}\left\{\mathbf{g}_i^H\mathbf{A}_\chi\mathbf{A}_\chi^H\mathbf{g}_i\right\}$, $\mathbb{E}\left\{\mathbf{g}_{\overline{\chi}}^H\mathbf{A}_\chi\mathbf{A}_\chi^H\mathbf{g}_{\overline{\chi}}\right\}$, and $\mathbb{E}\left\{\hat{\mathbf{g}}_\chi^H\mathbf{g}_\chi\right\}$.

Firstly, we have

$$\begin{aligned}&\mathbb{E}\left\{\mathbf{q}_\chi^H\mathbf{q}_\chi\right\}=\sum_{\omega=1}^4\mathbb{E}\left\{\left(\mathbf{q}_\chi^\omega\right)^H\mathbf{q}_\chi^\omega\right\}+\gamma_\chi\mathbb{E}\left\{\tilde{\mathbf{d}}_\chi^H\tilde{\mathbf{d}}_\chi\right\}\\&=c_\chi\delta\varepsilon_\chi\overline{\mathbf{h}}_\chi^H\mathbf{\Phi}_\chi^H\overline{\mathbf{H}}^H\overline{\mathbf{H}}\mathbf{\Phi}_\chi\overline{\mathbf{h}}_\chi+c_\chi\delta\mathbb{E}\left\{\tilde{\mathbf{h}}_\chi^H\mathbf{\Phi}_\chi^H\overline{\mathbf{H}}^H\overline{\mathbf{H}}\mathbf{\Phi}_\chi\tilde{\mathbf{h}}_\chi\right\}\\&+c_\chi\varepsilon_\chi\overline{\mathbf{h}}_\chi^H\mathbf{\Phi}_\chi^H\mathbb{E}\left\{\tilde{\mathbf{H}}^H\tilde{\mathbf{H}}\right\}\mathbf{\Phi}_\chi\overline{\mathbf{h}}_\chi+c_\chi\mathbb{E}\left\{\tilde{\mathbf{h}}_\chi^H\mathbf{\Phi}_\chi^H\mathbb{E}\left\{\tilde{\mathbf{H}}^H\tilde{\mathbf{H}}\right\}\right.\\&\left.\quad\mathbf{\Phi}_\chi\tilde{\mathbf{h}}_\chi\right\}+\gamma_\chi\mathbb{E}\left\{\tilde{\mathbf{d}}_\chi^H\tilde{\mathbf{d}}_\chi\right\}\qquad(89)\\&=M\left\{\left|f_\chi(\mathbf{\Phi}_\chi)\right|^2c_\chi\delta\varepsilon_\chi+A_\chi Nc_\chi\left(\delta+\varepsilon_\chi+1\right)+\gamma_\chi\right\}.\end{aligned}$$

Secondly, by using the expression of $\hat{\mathbf{g}}_\chi$, we have

$$\begin{aligned}&\mathbb{E}\left\{\hat{\mathbf{g}}_\chi^H\hat{\mathbf{g}}_\chi\right\}=\sum_{\omega=1}^4\mathbb{E}\left\{\left\|\hat{\mathbf{g}}_\chi^\omega\right\|^2\right\}\qquad(90)\\&=M\left\{\left|f_\chi(\mathbf{\Phi}_\chi)\right|^2c_\chi\delta\varepsilon_\chi+A_\chi Nc_\chi\delta e_{\chi_2}^2+A_\chi Nc_\chi\left(\varepsilon_\chi+1\right)e_{\chi_3}\right\}\end{aligned}$$

Thirdly, using $\mathbf{A}_\chi^H=\mathbf{A}_\chi$ and $\mathbf{A}_\chi\overline{\mathbf{H}}=e_{\chi_2}\overline{\mathbf{H}}$, we have

$$\begin{aligned}&\mathbb{E}\left\{\mathbf{g}_\chi^H\mathbf{A}_\chi\mathbf{A}_\chi^H\mathbf{g}_\chi\right\}=\mathbb{E}\left\{\sum_{\omega=1}^4\sum_{\psi=1}^4\left(\mathbf{A}_\chi\mathbf{g}_\chi^\omega\right)^H\left(\mathbf{A}_\chi\mathbf{g}_\chi^\psi\right)\right\}\\&=\left\|\sqrt{c_\chi\delta\varepsilon_\chi}\mathbf{A}_\chi\overline{\mathbf{H}}\mathbf{\Phi}_\chi\overline{\mathbf{h}}_\chi\right\|^2+\sum_{\omega=2}^4\mathbb{E}\left\{\left\|\hat{\mathbf{g}}_\chi^\omega\right\|^2\right\}\\&=M\left\{\left|f_\chi(\mathbf{\Phi}_\chi)\right|^2c_\chi\delta\varepsilon_\chi e_{\chi_2}^2+A_\chi Nc_\chi\delta e_{\chi_2}^2\right.\\&\left.+A_\chi Nc_\chi\left(\varepsilon_\chi+1\right)e_{\chi_3}\right\}.\end{aligned}\tag{91}$$

Besides, for $i\neq\chi$ and $\overline{\chi}$, we have

$$\begin{aligned}\mathbb{E}\left\{\mathbf{g}_i^H\mathbf{A}_\chi\mathbf{A}_\chi^H\mathbf{g}_i\right\}=&M\left\{\left|f_i(\mathbf{\Phi}_\chi)\right|^2c_i\delta\varepsilon_ie_{\chi_2}^2+A_\chi Nc_i\delta e_{\chi_2}^2\right.\\&\left.+A_\chi Nc_i\left(\varepsilon_i+1\right)e_{\chi_3}\right\},\end{aligned}\tag{92}$$

$$\begin{aligned}\mathbb{E}\left\{\mathbf{g}_{\overline{\chi}}^H\mathbf{A}_\chi\mathbf{A}_\chi^H\mathbf{g}_{\overline{\chi}}\right\}=&M\left\{\left|f_{\overline{\chi}}(\mathbf{\Phi}_{\overline{\chi}})\right|^2c_{\overline{\chi}}\delta\varepsilon_{\overline{\chi}}e_{\chi_2}^2+A_{\overline{\chi}}Nc_{\overline{\chi}}\delta e_{\chi_2}^2\right.\\&\left.+A_{\overline{\chi}}Nc_{\overline{\chi}}\left(\varepsilon_{\overline{\chi}}+1\right)e_{\chi_3}\right\}.\end{aligned}\tag{93}$$

Finally, by substituting $\gamma_\chi=0$ into (88), we arrive at

$$\begin{aligned}\mathbb{E}\left\{\hat{\mathbf{g}}_\chi^H\mathbf{g}_\chi\right\}=&M\left\{\left|f_\chi(\mathbf{\Phi}_\chi)\right|^2c_\chi\delta\varepsilon_\chi+A_\chi Nc_\chi\delta e_{\chi_2}+A_\chi N\times\right.\\&\left.c_\chi\left(\varepsilon_\chi+1\right)e_{\chi_1}\right\}.\end{aligned}\tag{94}$$

B. Interference Term

The interference term can be expanded as

$$\begin{aligned}&I_{\chi i}=\mathbb{E}\left\{\left|\hat{\mathbf{q}}_\chi^H\mathbf{q}_i\right|^2\right\}\\&=\mathbb{E}\left\{\left|\left(\hat{\mathbf{g}}_\chi+\mathbf{A}_\chi\mathbf{d}_\chi+\frac{1}{\sqrt{\tau p_\chi}}\mathbf{A}_\chi\mathbf{N}\mathbf{s}_\chi\right)^H\left(\mathbf{g}_i+\mathbf{d}_i\right)\right|^2\right\}\\&=\mathbb{E}\left\{\left|\hat{\mathbf{g}}_\chi^H\mathbf{g}_i\right|^2\right\}+\mathbb{E}\left\{\left|\hat{\mathbf{g}}_\chi^H\mathbf{d}_i\right|^2\right\}+\mathbb{E}\left\{\left|\mathbf{d}_\chi^H\mathbf{A}_\chi^H\mathbf{g}_i\right|^2\right\}\\&+\mathbb{E}\left\{\left|\mathbf{d}_\chi^H\mathbf{A}_\chi^H\mathbf{d}_i\right|^2\right\}+\frac{1}{\tau p_\chi}\mathbb{E}\left\{\left|\mathbf{s}_\chi^H\mathbf{N}^H\mathbf{A}_\chi^H\mathbf{g}_i\right|^2\right\}\\&+\frac{1}{\tau p_\chi}\mathbb{E}\left\{\left|\mathbf{s}_\chi^H\mathbf{N}^H\mathbf{A}_\chi^H\mathbf{d}_i\right|^2\right\}.\end{aligned}\tag{95}$$

The six expectations in (95) will be calculated one by one, and the derivation of $\mathbb{E}\left\{\left|\hat{\mathbf{g}}_\chi^H\mathbf{g}_i\right|^2\right\}$ will be scheduled last. The last five terms can be respectively tackled as

$$\mathbb{E}\left\{\left|\hat{\mathbf{g}}_\chi^H\mathbf{d}_i\right|^2\right\}=\mathbb{E}\left\{\hat{\mathbf{g}}_\chi^H\mathbf{d}_i\mathbf{d}_i^H\hat{\mathbf{g}}_\chi\right\}=\gamma_i\mathbb{E}\left\{\hat{\mathbf{g}}_\chi^H\hat{\mathbf{g}}_\chi\right\},\tag{96}$$

$$\begin{aligned}\mathbb{E}\left\{\left|\mathbf{d}_\chi^H\mathbf{A}_\chi^H\mathbf{g}_i\right|^2\right\}&=\mathbb{E}\left\{\mathbf{g}_i^H\mathbf{A}_\chi\mathbb{E}\left\{\mathbf{d}_\chi\mathbf{d}_\chi^H\right\}\mathbf{A}_\chi^H\mathbf{g}_i\right\}\\&=\gamma_\chi\mathbb{E}\left\{\mathbf{g}_i^H\mathbf{A}_\chi\mathbf{A}_\chi^H\mathbf{g}_i\right\},\end{aligned}\tag{97}$$

$$\mathbb{E}\left\{\left|\mathbf{d}_\chi^H\mathbf{A}_\chi^H\mathbf{d}_i\right|^2\right\}=\mathbb{E}\left\{\mathbf{d}_\chi^H\mathbf{A}_\chi^H\mathbb{E}\left\{\mathbf{d}_i\mathbf{d}_i^H\right\}\mathbf{A}_\chi\mathbf{d}_\chi\right\}\tag{98}$$

$$=\gamma_i\mathbb{E}\left\{\mathbf{d}_\chi^H\mathbf{A}_\chi^H\mathbf{A}_\chi\mathbf{d}_\chi\right\}=\gamma_\chi\gamma_i\,\mathrm{Tr}\left\{\mathbf{A}_\chi^H\mathbf{A}_\chi\right\}=\gamma_\chi\gamma_iMe_{\chi_3},$$

$$\frac{1}{\tau p_\chi}\mathbb{E}\left\{\left|\mathbf{s}_\chi^H\mathbf{N}^H\mathbf{A}_\chi^H\mathbf{g}_i\right|^2\right\}=\frac{\sigma^2}{\tau p_\chi}\mathbb{E}\left\{\mathbf{g}_i^H\mathbf{A}_\chi\mathbf{A}_\chi^H\mathbf{g}_i\right\},\tag{99}$$

$$\frac{1}{\tau p_\chi}\mathbb{E}\left\{\left|\mathbf{s}_\chi^H\mathbf{N}^H\mathbf{A}_\chi^H\mathbf{d}_i\right|^2\right\}=\frac{\sigma^2}{\tau p_\chi}\gamma_iMe_{\chi_3},\tag{100}$$

where $\mathbb{E}\left\{\hat{\mathbf{g}}_\chi^H\hat{\mathbf{g}}_\chi\right\}$ and $\mathbb{E}\left\{\mathbf{g}_i^H\mathbf{A}_\chi\mathbf{A}_\chi^H\mathbf{g}_i\right\}$ are given in (90) and (92).

Finally, we derive the first term $\mathbb{E}\left\{\left|\hat{\mathbf{g}}_\chi^H \mathbf{g}_i\right|^2\right\}$, which can be expanded as

$$\mathbb{E}\left\{\left|\hat{\mathbf{g}}_\chi^H \mathbf{g}_i\right|^2\right\} = \mathbb{E}\left\{\left|\sum_{\omega=1}^{4}\sum_{\psi=1}^{4}\left(\hat{\mathbf{g}}_\chi^\omega\right)^H \mathbf{g}_i^\psi\right|^2\right\} \tag{101}$$
$$= \sum_{\omega=1}^{4}\sum_{\psi=1}^{4}\mathbb{E}\left\{\left|\left(\hat{\mathbf{g}}_\chi^\omega\right)^H \mathbf{g}_i^\psi\right|^2\right\}$$
$$+ \sum_{\substack{\omega1,\psi1,\omega2,\psi2,\\(\omega1,\psi1)\neq(\omega2,\psi2)}}^{4}\mathbb{E}\left\{\left(\left(\hat{\mathbf{g}}_\chi^{\omega1}\right)^H \mathbf{g}_i^{\psi1}\right)\left(\left(\hat{\mathbf{g}}_\chi^{\omega2}\right)^H \mathbf{g}_i^{\psi2}\right)^H\right\}.$$

To begin with, we consider the terms with $\omega = 1$. We have

$$\mathbb{E}\left\{\left|\left(\hat{\mathbf{g}}_\chi^1\right)^H \mathbf{g}_i^1\right|^2\right\} \tag{102}$$
$$= \mathbb{E}\left\{\left|\sqrt{c_\chi\delta\varepsilon_\chi}\sqrt{c_i\delta\varepsilon_i}\overline{\mathbf{h}}_\chi^H\mathbf{\Phi}_\chi^H\overline{\mathbf{H}}^H\overline{\mathbf{H}}\mathbf{\Phi}_i\overline{\mathbf{h}}_i\right|^2\right\}$$
$$= c_\chi c_i \delta^2 \varepsilon_\chi\varepsilon_i M^2 \left|f_\chi(\mathbf{\Phi}_\chi)\right|^2\left|f_i(\mathbf{\Phi}_i)\right|^2.$$

Similarly, we have

$$\mathbb{E}\left\{\left|\left(\hat{\mathbf{g}}_\chi^1\right)^H \mathbf{g}_i^2\right|^2\right\} = A_\chi c_\chi c_i \delta^2\varepsilon_\chi M^2 N\left|f_\chi(\mathbf{\Phi}_\chi)\right|^2, \tag{103}$$
$$\mathbb{E}\left\{\left|\left(\hat{\mathbf{g}}_\chi^1\right)^H \mathbf{g}_i^3\right|^2\right\} = A_\chi c_\chi c_i \delta\varepsilon_\chi\varepsilon_i MN\left|f_\chi(\mathbf{\Phi}_\chi)\right|^2, \tag{104}$$
$$\mathbb{E}\left\{\left|\left(\hat{\mathbf{g}}_\chi^1\right)^H \mathbf{g}_i^4\right|^2\right\} = A_\chi c_\chi c_i \delta\varepsilon_\chi MN\left|f_\chi(\mathbf{\Phi}_\chi)\right|^2. \tag{105}$$

When $\omega = 2, 3, 4$, $\sum_{\psi=1}^{4}\mathbb{E}\left\{\left|\left(\hat{\mathbf{g}}_\chi^\omega\right)^H \mathbf{g}_i^\psi\right|^2\right\}$ can be respectively expanded as

$$\sum_{\psi=1}^{4}\mathbb{E}\left\{\left|\left(\hat{\mathbf{g}}_\chi^2\right)^H \mathbf{g}_i^\psi\right|^2\right\} \tag{106}$$
$$= A_\chi e_{\chi_2}^2 c_\chi c_i\delta^2\varepsilon_i M^2 N\left|f_i(\mathbf{\Phi}_i)\right|^2 + A_\chi^2 e_{\chi_2}^2 c_\chi c_i \delta^2 M^2 N^2$$
$$+ A_\chi^2 e_{\chi_2}^2 c_k c_i\delta\varepsilon_i MN^2 + A_\chi^2 e_{\chi_2}^2 c_\chi c_i \delta MN^2,$$

$$\sum_{\psi=1}^{4}\mathbb{E}\left\{\left|\left(\hat{\mathbf{g}}_\chi^3\right)^H \mathbf{g}_i^\psi\right|^2\right\} \tag{107}$$
$$= A_\chi e_{\chi_2}^2 c_\chi c_i\delta\varepsilon_\chi\varepsilon_i MN\left|f_i(\mathbf{\Phi}_i)\right|^2 + A_\chi^2 e_{\chi_2}^2 c_\chi c_i\delta\varepsilon_\chi MN^2$$
$$+ A_\chi^2 c_\chi c_i \varepsilon_\chi\varepsilon_i\left(e_{\chi_1}^2 M^2\left|\overline{\mathbf{h}}_\chi^H\overline{\mathbf{h}}_i\right|^2 + e_{\chi_3}MN^2\right)$$
$$+ A_\chi^2 c_\chi c_i\varepsilon_\chi\left(e_{\chi_1}^2 M^2 + e_{\chi_3}MN\right)N,$$
$$\sum_{\psi=1}^{4}\mathbb{E}\left\{\left|\left(\hat{\mathbf{g}}_\chi^4\right)^H \mathbf{g}_i^\psi\right|^2\right\} = A_\chi e_{\chi_2}^2 c_\chi c_i\delta\varepsilon_i MN\left|f_i(\mathbf{\Phi})\right|^2$$
$$+ A_\chi^2 e_{\chi_2}^2 c_\chi c_i \delta MN^2 + A_\chi^2 c_\chi c_i\varepsilon_i\left(e_{\chi_1}^2 M^2 N + e_{\chi_3}MN^2\right)$$
$$+ A_\chi^2 c_\chi c_i\left(e_{\chi_1}^2 M^2 N + e_{\chi_3}MN^2\right). \tag{108}$$

Then, we focus on the remaining cross-terms in (101). They can be combined as

$$\sum_{\substack{\omega1,\psi1,\omega2,\psi2,\\(\omega1,\psi1)\neq(\omega2,\psi2)}}^{4}\mathbb{E}\left\{\left(\left(\hat{\mathbf{g}}_\chi^{\omega1}\right)^H \mathbf{g}_i^{\psi1}\right)\left(\left(\hat{\mathbf{g}}_\chi^{\omega2}\right)^H \mathbf{g}_i^{\psi2}\right)^H\right\}$$
$$= \sum_{\substack{\omega1,\psi1,\\\omega2,\psi2}} 2\,\mathrm{Re}\left\{\mathbb{E}\left\{\left(\left(\hat{\mathbf{g}}_\chi^{\omega1}\right)^H \mathbf{g}_i^{\psi1}\right)\left(\left(\hat{\mathbf{g}}_\chi^{\omega2}\right)^H \mathbf{g}_i^{\psi2}\right)^H\right\}\right\}, \tag{109}$$

and $(\omega1, \psi1, \omega2, \psi2)$ includes four cases, i.e., $(1, 1, 3, 3)$, $(1, 2, 3, 4)$, $(2, 1, 4, 3)$, $(2, 2, 4, 4)$. Then, we calculate these 4 terms in (109) one by one. The first one is given by

$$2\,\mathrm{Re}\left\{\mathbb{E}\left\{\left(\left(\hat{\mathbf{g}}_\chi^1\right)^H \mathbf{g}_i^1\right)\left(\left(\hat{\mathbf{g}}_\chi^3\right)^H \mathbf{g}_i^3\right)^H\right\}\right\}$$
$$= 2A_\chi c_\chi c_i\delta\varepsilon_\chi\varepsilon_i e_{\chi_1} M\,\mathrm{Re}\left\{\overline{\mathbf{h}}_\chi^H\mathbf{\Phi}_\chi^H\overline{\mathbf{H}}^H\overline{\mathbf{H}}\mathbf{\Phi}_i\overline{\mathbf{h}}_i\overline{\mathbf{h}}_i^H\overline{\mathbf{h}}_\chi\right\}$$
$$= 2A_\chi c_\chi c_i\delta\varepsilon_\chi\varepsilon_i e_{\chi_1} M\,\mathrm{Re}\left\{\overline{\mathbf{h}}_\chi^H\mathbf{\Phi}_\chi^H\mathbf{a}_N\mathbf{a}_M^H\mathbf{a}_M\mathbf{a}_N^H\mathbf{\Phi}_i\overline{\mathbf{h}}_i\overline{\mathbf{h}}_i^H\overline{\mathbf{h}}_\chi\right\}$$
$$= 2A_\chi c_\chi c_i\delta\varepsilon_\chi\varepsilon_i e_{\chi_1} M^2\,\mathrm{Re}\left\{f_\chi^H(\mathbf{\Phi}_\chi)f_i(\mathbf{\Phi}_i)\overline{\mathbf{h}}_i^H\overline{\mathbf{h}}_\chi\right\}. \tag{110}$$

Then, the remaining cross-term is

$$\sum_{\substack{(1,2,3,4)\\(2,1,4,3)(2,2,4,4)}} 2\,\mathrm{Re}\left\{\mathbb{E}\left\{\left(\left(\hat{\mathbf{g}}_\chi^{\omega1}\right)^H \mathbf{g}_i^{\psi1}\right)\left(\left(\hat{\mathbf{g}}_\chi^{\omega2}\right)^H \mathbf{g}_i^{\psi2}\right)^H\right\}\right\}$$
$$= 2A_\chi c_\chi c_i\delta\varepsilon_k e_{\chi_1}M^2\left|f_\chi(\mathbf{\Phi}_\chi)\right|^2 + 2A_\chi^2 c_\chi c_i\delta e_{\chi_1}e_{\chi_2}M^2 N$$
$$+ 2A_\chi c_\chi c_i\delta\varepsilon_i e_{\chi_1}e_{\chi_2}M^2\left|f_i(\mathbf{\Phi}_i)\right|^2. \tag{111}$$

C. Signal Leakage

In this subsection, we derive the signal leakage term as

$$E_\chi^{(\mathrm{leakage})} = \mathbb{E}\left\{\left|\hat{\mathbf{q}}_\chi^H\mathbf{q}_\chi\right|^2\right\} - \left|\mathbb{E}\left\{\hat{\mathbf{q}}_\chi^H\mathbf{q}_\chi\right\}\right|^2, \tag{112}$$

where $\mathbb{E}\left\{\hat{\mathbf{q}}_\chi^H\mathbf{q}_\chi\right\}$ has been derived in (88). The remaining term can be decoupled as

$$\mathbb{E}\left\{\left|\hat{\mathbf{q}}_\chi^H\mathbf{q}_\chi\right|^2\right\} \tag{113}$$
$$= \mathbb{E}\left\{\left|\hat{\mathbf{g}}_\chi^H\mathbf{g}_\chi\right|^2\right\} + \mathbb{E}\left\{\left|\hat{\mathbf{g}}_\chi^H\mathbf{d}_\chi\right|^2\right\} + \mathbb{E}\left\{\left|\mathbf{d}_\chi^H\mathbf{A}_\chi^H\mathbf{g}_\chi\right|^2\right\}$$
$$+ \mathbb{E}\left\{\left|\mathbf{d}_\chi^H\mathbf{A}_\chi^H\mathbf{d}_\chi\right|^2\right\} + 2\,\mathrm{Re}\left\{\mathbb{E}\left\{\hat{\mathbf{g}}_\chi^H\mathbf{g}_\chi\left(\mathbf{d}_\chi^H\mathbf{A}_\chi^H\mathbf{d}_\chi\right)^H\right\}\right\}$$
$$+ \mathbb{E}\left\{\left|\frac{1}{\sqrt{\tau p_\chi}}\mathbf{s}_\chi^H\mathbf{N}^H\mathbf{A}_\chi^H\mathbf{d}_\chi\right|^2\right\} + \mathbb{E}\left\{\left|\frac{1}{\sqrt{\tau p_\chi}}\mathbf{s}_\chi^H\mathbf{N}^H\mathbf{A}_\chi^H\mathbf{g}_\chi\right|^2\right\},$$
$$= \mathbb{E}\left\{\left|\hat{\mathbf{g}}_\chi^H\mathbf{g}_\chi\right|^2\right\} + \gamma_\chi\mathbb{E}\left\{\hat{\mathbf{g}}_\chi^H\hat{\mathbf{g}}_\chi\right\} + \gamma_\chi\mathbb{E}\left\{\mathbf{g}_\chi^H\mathbf{A}_\chi\mathbf{A}_\chi^H\mathbf{g}_\chi\right\}$$
$$+ \gamma_\chi^2 M e_{\chi_3} + \gamma_\chi^2 M^2 e_{\chi_1}^2 + 2\gamma_\chi M e_{\chi_1}\mathbb{E}\left\{\hat{\mathbf{g}}_\chi^H\mathbf{g}_\chi\right\}$$
$$+ \frac{\sigma^2}{\tau p_\chi}\mathbb{E}\left\{\mathbf{g}_\chi^H\mathbf{A}_\chi\mathbf{A}_\chi^H\mathbf{g}_\chi\right\} + \frac{\sigma^2}{\tau p_\chi}\gamma_\chi M e_{\chi_3},$$

where $\mathbb{E}\left\{\hat{\mathbf{g}}_\chi^H\mathbf{g}_\chi\right\}$, $\mathbb{E}\left\{\hat{\mathbf{g}}_\chi^H\hat{\mathbf{g}}_\chi\right\}$, $\mathbb{E}\left\{\mathbf{g}_\chi^H\mathbf{A}_\chi\mathbf{A}_\chi^H\mathbf{g}_\chi\right\}$ are respectively given in (94), (90), (91).

Finally, the first term $\mathbb{E}\left\{\left|\hat{\mathbf{g}}_\chi^H\mathbf{g}_\chi\right|^2\right\}$ in (113) can be expanded as

$$\mathbb{E}\left\{\left|\hat{\mathbf{g}}_\chi^H\mathbf{g}_\chi\right|^2\right\} = \sum_{\omega=1}^{4}\sum_{\psi=1}^{4}\mathbb{E}\left\{\left|\left(\hat{\mathbf{g}}_\chi^\omega\right)^H\mathbf{g}_\chi^\psi\right|^2\right\} \tag{114}$$

$$+ \sum_{\substack{\omega 1,\psi 1,\omega 2,\psi 2\\(\omega 1,\psi 1)\neq(\omega 2,\psi 2)}}^{4} \mathbb{E}\left\{\left(\left(\hat{\mathbf{g}}_\chi^{\omega 1}\right)^H \mathbf{g}_\chi^{\psi 1}\right)\left(\left(\hat{\mathbf{g}}_\chi^{\omega 2}\right)^H \mathbf{g}_\chi^{\psi 2}\right)^H\right\}.$$

The terms $\sum_{\omega=1}^{4}\sum_{\psi=1}^{4}\mathbb{E}\left\{\left|\left(\hat{\mathbf{g}}_\chi^{\omega}\right)^H \mathbf{g}_\chi^{\psi}\right|^2\right\}$ can be derived similarly to (102). The remaining terms have 20 non-zero cross-terms. These terms can be combined into 10 terms, corresponding to $(\omega 1,\psi 1,\omega 2,\psi 2) = (1,1,2,2)$, $(1,1,3,3)$, $(1,1,4,4)$, $(1,2,3,4)$, $(1,3,2,4)$, $(2,1,4,3)$, $(2,2,3,3)$, $(2,2,4,4)$, $(3,1,4,2)$, $(3,3,4,4)$.

D. Inter-group interference

Similarly, we can expand this term as

$$\begin{aligned}
I_{\chi\overline{\chi}} &= \mathbb{E}\left\{\left|\hat{\mathbf{q}}_\chi^H \mathbf{q}_{\overline{\chi}}\right|^2\right\} = \mathbb{E}\left\{\left|\hat{\mathbf{g}}_\chi^H \mathbf{g}_{\overline{\chi}}\right|^2\right\} + \mathbb{E}\left\{\left|\hat{\mathbf{g}}_\chi^H \mathbf{d}_{\overline{\chi}}\right|^2\right\} \\
&+ \mathbb{E}\left\{\left|\mathbf{d}_\chi^H \mathbf{A}_\chi^H \mathbf{g}_{\overline{\chi}}\right|^2\right\} + \mathbb{E}\left\{\left|\mathbf{d}_\chi^H \mathbf{A}_\chi^H \mathbf{d}_{\overline{\chi}}\right|^2\right\} \qquad (115)\\
&+ \frac{1}{\tau p_\chi}\mathbb{E}\left\{\left|\mathbf{s}_\chi^H \mathbf{N}^H \mathbf{A}_\chi^H \mathbf{g}_{\overline{\chi}}\right|^2\right\} + \frac{1}{\tau p_\chi}\mathbb{E}\left\{\left|\mathbf{s}_\chi^H \mathbf{N}^H \mathbf{A}_\chi^H \mathbf{d}_{\overline{\chi}}\right|^2\right\} \\
&= \mathbb{E}\left\{\left|\hat{\mathbf{g}}_\chi^H \mathbf{g}_{\overline{\chi}}\right|^2\right\} + \gamma_{\overline{\chi}}\mathbb{E}\left\{\hat{\mathbf{g}}_\chi^H \hat{\mathbf{g}}_\chi\right\} + \gamma_\chi \mathbb{E}\left\{\mathbf{g}_{\overline{\chi}}^H \mathbf{A}_\chi \mathbf{A}_\chi^H \mathbf{g}_{\overline{\chi}}\right\} \\
&+ \gamma_\chi\gamma_{\overline{\chi}} M e_{\chi_3} + \frac{\sigma^2}{\tau p_\chi}\mathbb{E}\left\{\mathbf{g}_{\overline{\chi}}^H \mathbf{A}_\chi \mathbf{A}_\chi^H \mathbf{g}_{\overline{\chi}}\right\} + \frac{\sigma^2}{\tau p_\chi}\gamma_{\overline{\chi}} M e_{\chi_3},
\end{aligned}$$

where $\mathbb{E}\left\{\hat{\mathbf{g}}_\chi^H \hat{\mathbf{g}}_\chi\right\}$, $\mathbb{E}\left\{\mathbf{g}_{\overline{\chi}}^H \mathbf{A}_\chi \mathbf{A}_\chi^H \mathbf{g}_{\overline{\chi}}\right\}$ are respectively given in (90) and (92). The first term can be derived in the same way as (101), but it is worth noting when it comes to $\omega = 3, \psi = 3$, we have

$$\begin{aligned}
\mathbb{E}\left\{\left|\left(\hat{\mathbf{g}}_\chi^3\right)^H \mathbf{g}_\chi^3\right|^2\right\} &= c_\chi c_{\overline{\chi}} \varepsilon_\chi \varepsilon_{\overline{\chi}} e_{\chi_1}^2 M^2 \left|\overline{\mathbf{h}}_\chi^H \mathbf{\Phi}_\chi^H \mathbf{\Phi}_{\overline{\chi}} \overline{\mathbf{h}}_{\overline{\chi}}\right|^2 \\
&+ A_\chi A_{\overline{\chi}} c_\chi c_{\overline{\chi}} \varepsilon_\chi \varepsilon_{\overline{\chi}} e_{\chi_3} M N^2. \qquad (116)
\end{aligned}$$

References

[1] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive MIMO: Benefits and challenges," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 742–758, 2014.

[2] W. Tang, M. Z. Chen, X. Chen, J. Y. Dai, Y. Han, M. Di Renzo, Y. Zeng, S. Jin, Q. Cheng, and T. J. Cui, "Wireless communications with reconfigurable intelligent surface: Path loss modeling and experimental measurement," *IEEE Trans. Wireless Commun.*, vol. 20, no. 1, pp. 421–439, 2021.

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

[4] E. Björnson, E. G. Larsson, and T. L. Marzetta, "Massive MIMO: ten myths and one critical question," *IEEE Commun. Mag.*, vol. 54, no. 2, pp. 114–123, 2016.

[5] C.-X. Wang, X. You, X. Gao, X. Zhu, Z. Li, C. Zhang, H. Wang, Y. Huang, Y. Chen, H. Haas, J. S. Thompson, E. G. Larsson, M. D. Renzo, W. Tong, P. Zhu, X. Shen, H. V. Poor, and L. Hanzo, "On the road to 6G: Visions, requirements, key technologies, and testbeds," *IEEE Commun. Surveys Tuts.*, vol. 25, no. 2, pp. 905–974, 2023.

[6] J. Huang, C.-X. Wang, Y. Sun, R. Feng, J. Huang, B. Guo, Z. Zhong, and T. J. Cui, "Reconfigurable intelligent surfaces: Channel characterization and modeling," *Proceedings of the IEEE*, vol. 110, no. 9, pp. 1290–1311, 2022.

[7] C. Huang, R. Mo, and C. Yuen, "Reconfigurable intelligent surface assisted multiuser MISO systems exploiting deep reinforcement learning," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 8, pp. 1839–1850, 2020.

[8] W. Tang, M. Z. Chen, X. Chen, J. Y. Dai, Y. Han, M. Di Renzo, Y. Zeng, S. Jin, Q. Cheng, and T. J. Cui, "Wireless communications with reconfigurable intelligent surface: Path loss modeling and experimental measurement," *IEEE Trans. Wireless Commun.*, vol. 20, no. 1, pp. 421–439, 2021.

[9] C. Huang, A. Zappone, G. C. Alexandropoulos, M. Debbah, and C. Yuen, "Reconfigurable intelligent surfaces for energy efficiency in wireless communication," *IEEE Trans. Wireless Commun.*, vol. 18, no. 8, pp. 4157–4170, 2019.

[10] S. Zhou, W. Xu, K. Wang, M. Di Renzo, and M.-S. Alouini, "Spectral and energy efficiency of IRS-assisted MISO communication with hardware impairments," *IEEE Wireless Commun. Lett.*, vol. 9, no. 9, pp. 1366–1369, 2020.

[11] Y. Liu, X. Liu, X. Mu, T. Hou, J. Xu, M. Di Renzo, and N. Al-Dhahir, "Reconfigurable intelligent surfaces: Principles and opportunities," *IEEE Commun. Surveys Tuts.*, vol. 23, no. 3, pp. 1546–1577, 2021.

[12] M. Di Renzo, A. Zappone, M. Debbah, M.-S. Alouini, C. Yuen, J. de Rosny, and S. Tretyakov, "Smart radio environments empowered by reconfigurable intelligent surfaces: How it works, state of research, and the road ahead," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2450–2525, 2020.

[13] C. Pan, H. Ren, K. Wang, J. F. Kolb, M. Elkashlan, M. Chen, M. Di Renzo, Y. Hao, J. Wang, A. L. Swindlehurst, X. You, and L. Hanzo, "Reconfigurable intelligent surfaces for 6G systems: Principles, applications, and research directions," *IEEE Commun. Mag.*, vol. 59, no. 6, pp. 14–20, 2021.

[14] L. Yang, F. Meng, J. Zhang, M. O. Hasna, and M. D. Renzo, "On the performance of RIS-assisted dual-hop UAV communication systems," *IEEE Trans. Veh. Technol.*, vol. 69, no. 9, pp. 10 385–10 390, 2020.

[15] T. Hou, Y. Liu, Z. Song, X. Sun, Y. Chen, and L. Hanzo, "Reconfigurable intelligent surface aided NOMA networks," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2575–2588, 2020.

[16] K. Zhi, C. Pan, H. Ren, K. K. Chai, and M. Elkashlan, "Active RIS versus passive RIS: Which is superior with the same power budget?" *IEEE Communications Letters*, vol. 26, no. 5, pp. 1150–1154, 2022.

[17] W. Khalid, Z. Kaleem, R. Ullah, T. V. Chien, S. Noh, and H. Yu, "Simultaneous transmitting and reflecting-reconfigurable intelligent surface in 6G: Design guidelines and future perspectives," *IEEE Netw.*, pp. 1–9, 2022.

[18] J. Xu, Y. Liu, X. Mu, and O. A. Dobre, "STAR-RISs: Simultaneous transmitting and reflecting reconfigurable intelligent surfaces," *IEEE Commun. Lett.*, vol. 25, no. 9, pp. 3134–3138, 2021.

[19] M. Ahmed, A. Wahid, S. S. Laique, W. U. Khan, A. Ihsan, F. Xu, S. Chatzinotas, and Z. Han, "A survey on STAR-RIS: Use cases, recent advances, and future research challenges," *IEEE Internet of Things Journal*, vol. 10, no. 16, pp. 14 689–14 711, 2023.

[20] X. Mu, Y. Liu, L. Guo, J. Lin, and R. Schober, "Simultaneously transmitting and reflecting (STAR) RIS aided wireless communications," *IEEE Trans. Wireless Commun.*, vol. 21, no. 5, pp. 3083–3098, 2022.

[21] H. Niu, Z. Chu, F. Zhou, P. Xiao, and N. Al-Dhahir, "Weighted sum rate optimization for STAR-RIS-assisted MIMO system," *IEEE Trans. Veh. Technol.*, vol. 71, no. 2, pp. 2122–2127, 2022.

[22] A. Zappone, M. Di Renzo, F. Shams, X. Qian, and M. Debbah, "Overhead-aware design of reconfigurable intelligent surfaces in smart radio environments," *IEEE Trans. Wireless Commun.*, vol. 20, no. 1, pp. 126–141, 2021.

[23] X. Gan, C. Zhong, C. Huang, and Z. Zhang, "RIS-assisted multi-user MISO communications exploiting statistical CSI," *IEEE Trans. Commun.*, vol. 69, no. 10, pp. 6781–6792, 2021.

[24] P. Wang, J. Fang, Z. Wu, and H. Li, "Two-timescale beamforming for IRS-assisted millimeter wave systems: A deep unrolling-based stochastic optimization approach," in *2022 IEEE 12th Sensor Array and Multichannel Signal Processing Workshop (SAM)*, 2022, pp. 191–195.

[25] K. Zhi, C. Pan, H. Ren, K. Wang, M. Elkashlan, M. D. Renzo, R. Schober, H. V. Poor, J. Wang, and L. Hanzo, "Two-timescale design for reconfigurable intelligent surface-aided massive MIMO systems with imperfect CSI," *IEEE Trans. Inf. Theory*, vol. 69, no. 5, pp. 3001–3033, 2023.

[26] C. Hu, L. Dai, S. Han, and X. Wang, "Two-timescale channel estimation for reconfigurable intelligent surface aided wireless communications," *IEEE Trans. Commun.*, vol. 69, no. 11, pp. 7736–7747, 2021.

[27] Z. Peng, Z. Chen, C. Pan, G. Zhou, and H. Ren, "Robust transmission design for RIS-aided communications with both transceiver hardware impairments and imperfect CSI," *IEEE Wireless Commun. Lett.*, vol. 11, no. 3, pp. 528–532, 2022.