Sum Rate of Extremely Large STAR-RIS Aided Uplink NOMA

Hamza Ahmed Qureshi, Seunghyun Oh, Noureen Khan, Yun Hee Kim, *Senior Member, IEEE*, and Arumugam Nallanathan, *Fellow, IEEE*

***Abstract*—This letter investigates the effect of an extremely large (XL) simultaneously transmitting and reflecting intelligent reconfigurable surface (STAR-RIS) on the sum rate of the uplink nonorthogonal multiple access (NOMA). For the XL-STAR-RIS supporting NOMA in energy-splitting mode, we develop a sum rate maximization algorithm under the minimum rate constraints that accommodates XL-STAR-RIS in a computationally efficient way. We also analyze upper and lower bounds on the sum rate to identify governing factors of the channel models. The results show that the sum rate gain with respect to the number of STAR-RIS elements is reduced when the channels between the STAR-RIS and users become near-field, which entails the use of the correct channel model to estimate the exact value of an XL-STAR-RIS.**

***Index Terms*—Energy splitting, near-field channels, nonorthogonal multiple access, simultaneously transmitting and reflecting, reconfigurable intelligent surface**

I. Introduction

To fulfill the ambitious goals of the upcoming sixth generation (6G) wireless networks, not only existing technologies such as multiple input multiple output (MIMO) and nonorthogonal multiple access (NOMA) have kept enhancing but also new concepts such as reconfigurable intelligent surface (RIS) aided communications and integrated sensing and communication have emerged [1]–[3]. In particular, for the 6G networks operating at high frequency, RISs have been conceived as a key building block owing to their capability of creating full-duplex line-of-sight (LoS) channels and orchestrating channel environments toward higher spectral and energy efficiency at low complexity [1], [4]. Hence, various aspects of RIS aided communications have been studied such as system optimization with RIS reflection and deployment and new types of RISs such as simultaneously transmitting and reflecting (STAR)-RISs that provide full-space coverage instead of half-space coverage of the reflecting-only RISs [4]–[9].

Passive RISs and STAR-RISs suffer from a product path loss, which can be compensated by increasing the number of RIS elements and deploying RISs near to a transmitter or receiver [4], [8]. With an extremely large (XL) number of RIS elements, communication nodes closer to an RIS commence to go through a spherical-wave propagation called near-field (NF) channel rather than a planar-wave propagation called far-field (FF) channel, in particular at a higher frequency band. Nevertheless, extensive studies on RIS aided communications have been devoted mostly to FF channels [4]–[9] whilst XL-MIMO networks have started to exploit their NF beamfocusing capability in communication and sensing [3], [10], [11].

Recently, studies on XL-RIS aided NF communications have embarked [12]–[15] with challenges and opportunities observed in [16]. The sum rate of XL-RIS aided NF wideband uplink was investigated by employing a simple receive beamfomring and suboptimal RIS reflection matching to each of user channels using virtual subarrays [12]. NF multibeams were designed for an XL-RIS assisted downlink by optimizing the weights in a weighted sum of beams selected by users under the unit modulus constraints of the RIS [13]. In case of an XL-STAR-RIS, the downlink NF communications were considered with energy splitting (ES) and mode switching (MS) protocols of transmitting (T) and reflecting (R) coefficients for a metasurface-based RIS [14] and for a patch-array based RIS [15]. In [15], alternating optimization (AO) of downlink BF and T/R coefficients was developed to maximize the weighted sum rate.

We address a patch-array based XL-STAR-RIS [15] to the uplink NOMA [17] to identify the sum rate without and with the minimum quality-of-service (QoS) requirements of users in both FF and NF channels. For the STAR-RIS aided NOMA, a similar sum rate maximization problem was considered in the downlink [7] but only the transmit power minimization with the TS protocol and the sum rate of two users with the MS protocol were considered in the uplink [6], [9]. However, these studies and most literature surveyed in [8] are limited to FF channels with the algorithms inefficient to optimize XL-STAR-RIS coefficients exhibiting the NF effect. In this regard, the main contributions of this paper are summarized as follows.

- We optimize the QoS-constrained sum rate of an XL-STAR-RIS aided uplink NOMA with the ES protocol. To handle XL-STAR-RIS coefficients required for the NF channels, we develop a computationally efficient optimization algorithm replacing AO and semidefinite relaxation (SDR) procedures typically used for RIS and STAR-RIS BF [5], [7], [15].
- We analyze upper and lower bounds of the sum rate without QoS constraints for a quick performance estimation as well as to identify main factors governing the sum rate in NF and FF channels. Analysis shows that the correlations of the cascaded user channels, depending on the size and shape of the STAR-RIS planar array, affect the sum rate of the uplink NOMA.
- We first show the efficiency of the proposed algorithm and the gain of the ES over TS and MS protocols. We

H.A. Qureshi, S. Oh, N. Khan, and Y.H. Kim are with the Department of Electronics and Information Convergence Engineering, Kyung Hee University, Yongin 17104, Korea (e-mail: {hamza.ahmed, ohioandy99, noureen, yheekim}@khu.ac.kr) (*Corresponding author: Yun Hee Kim.*).

A. Nallanathan is with the School of Electronic Engineering and Computer Science, Queen Mary University of London, London (a.nallanathan@qmul.ac.uk).

This research was supported by the National Research Foundation of Korea (NRF) under Grant NRF-2021R1A2C1005869 and by the Institute of Information & Communications Technology Planning & Evaluation (IITP) under the Information Technology Research Center (ITRC) support program IITP-2023-2021-0-02046, with funding from the Ministry of Science and ICT (MSIT), Korea.

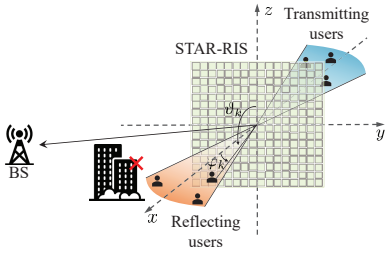


Fig. 1: System model for an XL-STAR-RIS assisted uplink.

then provide performance of the network optimized with perfect channel state information (CSI) in both NF and FF channels and with FF approximated CSI in NF channels to identify a performance gap incurred by the incorrect channel model and incorrect CSI.

Notation: The sets of $n \times m$ complex-valued (real-valued) matrices are denoted by $\mathbb{C}^{n\times m}$ ($\mathbb{R}^{n\times m}$). We denote by $[\mathbf{a}]_n$ and $\|\mathbf{a}\|_p$ the nth entry and ℓ^p-norm of vector $\mathbf{a}$, respectively. We use $\mathrm{diag}(\mathbf{a})$ for the diagonal matrix with $\mathbf{a}$ on the diagonal, $[\mathbf{A}]_{n,m}$ for the (n,m)th element of matrix $\mathbf{A}$, and $\otimes$ for the Kronecker product. In addition, $\mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ signifies complex Gaussian with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ and $\mathcal{U}(a,b)$ denotes uniform distribution over (a,b).

II. System and Signal Models

A. System and Channel Models

We consider an XL-STAR-RIS assisted uplink[1] depicted in Fig. 1, where a base station (BS) communicates with K users through an XL-STAR-RIS, referred to as STAR-RIS for brevity. The users in the transmitting (T) or reflecting (R) regimes of the STAR-RIS are grouped as $\mathcal{K}_t$ and $\mathcal{K}_r$, respectively, with $\mathcal{K}_t \cup \mathcal{K}_r = \mathcal{K} \triangleq \{1, 2, \cdots, K\}$. Due to the blockages between the BS and users, the single-antenna BS receives the signals from the single-antenna users via the STAR-RIS with N elements. We use $\boldsymbol{\theta}_t \in \mathbb{C}^{N\times 1}$ and $\boldsymbol{\theta}_r \in \mathbb{C}^{N\times 1}$ to denote the T and R coefficients of the STAR-RIS elements, respectively, where $\boldsymbol{\theta}_s = [\theta_{s1}, \theta_{s2}, \cdots, \theta_{sN}]^T$ for $s \in \{t, r\}$. The STAR-RIS operates in the ES protocol subject to $|\theta_{tn}|^2 + |\theta_{rn}|^2 = 1$ for $n \in \mathcal{N}$. The channel from user k to the STAR-RIS and the channel from the STAR-RIS to the BS are denoted by $\mathbf{v}_k \in \mathbb{C}^{N\times 1}$ for $k \in \mathcal{K}$ and $\mathbf{g} \in \mathbb{C}^{N\times 1}$.

The channels are described with the uniform planar array (UPA) STAR-RIS of size $N_y \times N_z$ and element spacing d, located in the yz-plane. The position of the STAR-IRS element is denoted by $\mathbf{u}^{\text{star}}_{n_y,n_z} = [0, (n_y - \frac{N_y}{2})d, (n_z - \frac{N_z}{2})d]$ for $n_y = 1, 2, \cdots, N_y$, $n_z = 1, 2, \cdots, N_z$ with $\mathbf{u}^{\text{star}}_{N_y/2,N_z/2}$ at the origin. The position of user k is expressed as

$$\mathbf{u}_k = [r_k \sin\vartheta_k \cos\varphi_k, r_k \sin\vartheta_k \sin\varphi_k, r_k \cos\vartheta_k] \tag{1}$$

with distance r_k, azimuth angle φ_k, and depression angle ϑ_k from the origin and the position $\mathbf{u}_0$ of the BS is described in a similar way with $(r_0, \varphi_0, \vartheta_0)$. We assume that the users are rather closer to the STAR-RIS than the BS so that an LoS channel is adopted for $\{\mathbf{v}_k\}_{k\in\mathcal{K}}$ whilst an LoS or non-LoS (NLoS) channel is adopted for $\mathbf{g}$.

The LoS channel is given by [13]

$$\mathbf{v}_k = \sqrt{N\beta_k}\mathbf{b}(\mathbf{u}_k),\ \mathbf{g} = \sqrt{N\beta_0}\mathbf{b}(\mathbf{u}_0), \tag{2}$$

where $\beta_m = (\frac{\lambda_c}{4\pi r_m})^2$ represents the free space path-loss with wavelength λ_c and

$$[\mathbf{b}(\mathbf{u}_m)]_{n_y+(n_z-1)N_y} = \tfrac{1}{\sqrt{N}} e^{-j\frac{2\pi}{\lambda_c}\|\mathbf{u}_m - \mathbf{u}^{\text{star}}_{n_y,n_z}\|_2} \tag{3}$$

for $m \in \mathcal{K}\cup\{0\}$. The array response exhibits the NF spherical wavefront if $r_m \ll \frac{2D^2}{\lambda_c}$ for STAR-RIS with dimension D and the FF planar wavefront, otherwise. The array response for the latter case is approximated as [13]

$$\mathbf{b}(\mathbf{u}_m) \approx e^{j\phi_{c,m}}\mathbf{a}_{\text{STAR}}(\varphi_m, \vartheta_m), \tag{4}$$

where $\phi_{c,m} = -\frac{2\pi}{\lambda_c}\|\mathbf{u}_m - \mathbf{u}^{\text{star}}_{0,0}\|_2$ and

$$\mathbf{a}_{\text{STAR}}(\varphi, \vartheta) = \mathbf{a}_{N_y}(\sin\varphi \sin\vartheta) \otimes \mathbf{a}_{N_z}(\cos\vartheta) \tag{5}$$

with $\mathbf{a}_Q(\nu) = \frac{1}{\sqrt{Q}}\left[1, e^{j\frac{2\pi d}{\lambda_c}\nu}, \cdots, e^{j\frac{2\pi d}{\lambda_c}\nu(Q-1)}\right]^T$. The NLoS $\mathbf{g}$ follows the geometric FF channel with L scatters as

$$\mathbf{g} = \sqrt{\frac{\beta_0^{\text{NLoS}} N}{L}} \sum_{l=1}^{L} \alpha_{0l}\mathbf{a}_{\text{STAR}}(\varphi_{0l}, \vartheta_{0l}), \tag{6}$$

where β_0^{NLoS} is the NLoS path-loss with α_{0l}, denotes the complex coefficient, and $(\varphi_{0l}, \vartheta_{0l})$ denote the azimuth and depression angles of the departure of the lth path.

B. Signal Model

For the STAR-RIS aided uplink NOMA with the ES protocol, all users transmit their symbols x_k at power p_k simultaneously. The received signal at the BS is given by

$$y = \sum_{k=1}^{K} \sqrt{p_k}\mathbf{g}_k^T \mathrm{diag}(\boldsymbol{\theta}_{s(k)})\mathbf{v}_k x_k + w, \tag{7}$$

where $s(k) \in \{t, r\}$ denotes T or R coefficients serving user k, $w \sim \mathcal{CN}(0, \sigma^2)$ is the background noise, and

$$\mathbf{h}_k = \mathrm{diag}(\mathbf{g})\mathbf{v}_k = \sqrt{N\beta_k}\mathrm{diag}(\mathbf{g})\mathbf{b}(\mathbf{u}_k) \tag{8}$$

is the cascaded channel from user k to the BS via the STAR-RIS. The symbols $\{x_k\}_{k\in\mathcal{K}}$ are decoded from (7) with successive interference cancellation (SIC), where the users are indexed by the SIC order. We adopt the SIC order [6] given by the maximally achievable channel power of each user, $\mathfrak{P}_k = \max_{\boldsymbol{\theta}_{s(k)}} |\mathbf{h}_k^T\boldsymbol{\theta}_{s(k)}|^2 = \|\mathbf{h}_k\|_1^2$, when the STAR-RIS serves only that user. The signal-to-interference-and-noise ratio (SINR) of user k with SIC becomes

$$\gamma_k = \frac{p_k|\mathbf{h}_k^T\boldsymbol{\theta}_{s(k)}|^2}{\sum_{l>k}^{K} p_l|\mathbf{h}_l^T\boldsymbol{\theta}_{s(l)}|^2 + \sigma^2}, \tag{9}$$

which leads to the achievable rate of user k as $R_k = \log_2(1+\gamma_k)$. The sum rate is then given by [5], [17]

$$R_{\text{sum}} = \sum_{k=1}^{K} \log_2\left(1 + \frac{p_k|\mathbf{h}_k^T\boldsymbol{\theta}_{s(k)}|^2}{\sum_{l>k}^{K} p_l|\mathbf{h}_l^T\boldsymbol{\theta}_{s(l)}|^2 + \sigma^2}\right) = \log_2\left(1 + \sum_{k=1}^{K}\frac{p_k|\mathbf{h}_k^T\boldsymbol{\theta}_{s(k)}|^2}{\sigma^2}\right). \tag{10}$$

[1]The channel models are applicable to the downlink but the signal model and consequent problem for the uplink are different from those for the downlink.

III. Sum Rate Maximization and Analysis

We aim to maximize the sum rate by optimizing STAR-RIS coefficients $\mathbf{\Theta} = [\boldsymbol{\theta}_t, \boldsymbol{\theta}_r]$ and power allocatioin (PA) $\mathbf{p} = [p_1, p_2, \cdots, p_K]^T$ as

$$\max_{\mathbf{\Theta}\in\mathbb{C}^{N\times 2}, \mathbf{p}\in\mathbb{R}^{K\times 1}} \quad R_{\text{sum}} \tag{11a}$$
$$\text{s.t.} \quad R_k \geq R_k^{\min}, \ \forall k, \tag{11b}$$
$$|\theta_{tn}|^2 + |\theta_{rn}|^2 = 1, \ \forall n, \tag{11c}$$
$$0 \leq p_k \leq P_k^{\max}, \ \forall k, \tag{11d}$$

where $R_k^{\min} \geq 0$ is a QoS requirement on the rate of user k. We first provide a computationally efficient optimization algorithm to handle a large number of T/R coefficients and then provide a sum rate analysis without QoS requirements.

A. Optimization Algorithm with QoS Constraints

We formulate an equivalent problem of (11) by removing the monotonically increasing log function as

$$\max_{\mathbf{\Theta}\in\mathbb{C}^{N\times 2}, \mathbf{p}\in\mathbb{R}^{K}} \quad \Gamma_{\text{sum}} \left(\triangleq \sum_{k=1}^{K} \frac{p_k \left|\mathbf{h}_k^T \boldsymbol{\theta}_{s(k)}\right|^2}{\sigma^2} \right) \tag{12a}$$
$$\text{s.t.} \quad \frac{p_k \left|\mathbf{h}_k^T \boldsymbol{\theta}_{s(k)}\right|^2}{\sum_{l>k}^{K} p_l \left|\mathbf{h}_l^T \boldsymbol{\theta}_{s(l)}\right|^2 + \sigma^2} \geq T_k^{\min}, \ \forall k, \tag{12b}$$
$$\text{(11c)}, \ \text{(11d)}, \tag{12c}$$

where $T_k^{\min} = 2^{R_k^{\min}} - 1$. To avoid the intolerable complexity of typical AO of PA and SDR-based (STAR-)RIS optimization computed at $\mathcal{O}(N^{4.5}) \sim \mathcal{O}(N^6)$ [5], [7], we propose to optimize $\mathbf{\Theta}$ and $\mathbf{p}$ simultaneously using new variables $\mathbf{v} = [v_1, v_2, \cdots, v_K]^T$ with

$$v_k \triangleq \frac{p_k}{\sigma^2} \left|\mathbf{h}_k^T \boldsymbol{\theta}_{s(k)}\right|^2 \leq \frac{P_{k,\max}}{\sigma^2} \left|\mathbf{h}_k^T \boldsymbol{\theta}_{s(k)}\right|^2. \tag{13}$$

We then transform problem (12) with (13) into

$$\max_{\mathbf{\Theta}\in\mathbb{C}^{N\times 2}, \mathbf{v}\in\mathbb{R}^{K\times 2}} \quad \sum_{k=1}^{K} v_k \tag{14a}$$
$$\text{s.t.} \quad -\frac{1}{T_k^{\min}} v_k + 1 + \sum\nolimits_{l>k}^{K} v_l \leq 0, \ \forall k, \tag{14b}$$
$$0 \leq v_k \leq \chi_k \zeta_k(\boldsymbol{\theta}_{s(k)}), \ \forall k, \tag{14c}$$
$$|\theta_{tn}|^2 + |\theta_{rn}|^2 \leq 1, \ \forall n, \tag{14d}$$

where

$$\chi_k = \frac{P_k^{\max}}{\sigma^2} \|\mathbf{h}_k\|_1^2 = \frac{P_k^{\max}}{\sigma^2} \beta_k \|\mathbf{g}\|_1^2 \tag{15}$$

represents the maximum SNR of user k from (8) and

$$\zeta_k(\boldsymbol{\theta}_{s(k)}) = \frac{|\mathbf{h}_k^T \boldsymbol{\theta}_{s(k)}|^2}{\|\mathbf{h}_k\|_1^2} \tag{16}$$

reflects the degree of channel matching of STAR-RIS coefficients $\boldsymbol{\theta}_{s(k)}$ to its serving channel $\mathbf{h}_k$. Although objective (14a) and constraints (14b) become linear, problem (14) is still nonconvex due to constraints (14c). To handle this nonconvexity, we solve (14) through successive convex approximation (SCA) by approximating $\zeta_k(\boldsymbol{\theta}_s)$ to its first-order Taylor series expansion around $\bar{\boldsymbol{\theta}}_s$ as

$$\hat{\zeta}_k(\boldsymbol{\theta}_s, \bar{\boldsymbol{\theta}}_s) = \zeta_k(\bar{\boldsymbol{\theta}}_s) + 2\Re\{\bar{\boldsymbol{\theta}}_s^H \frac{\mathbf{h}_k^* \mathbf{h}_k^T}{\|\mathbf{h}_k\|_1^2} (\boldsymbol{\theta}_s - \bar{\boldsymbol{\theta}}_s)\}. \tag{17}$$

We then solve the convex optimization problem

$$\max_{\mathbf{\Theta}\in\mathbb{C}^{N\times 2}, \mathbf{v}\in\mathbb{R}^{N\times 2}} \quad \sum\nolimits_{k=1}^{K} v_k \tag{18a}$$
$$\text{s.t.} \quad \text{(14b)}, \ \text{(14d)} \tag{18b}$$
$$0 \leq v_k \leq \chi_k \hat{\zeta}_k(\boldsymbol{\theta}_{s(k)}, \boldsymbol{\theta}_{s(k)}^{(q)}), \ \forall k. \tag{18c}$$

iteratively by updating $\hat{\zeta}_k(\boldsymbol{\theta}_{s(k)}, \boldsymbol{\theta}_{s(k)}^{(q)})$ with the solution $\mathbf{\Theta}^{(q)} = [\boldsymbol{\theta}_t^{(q)}, \boldsymbol{\theta}_r^{(q)}]$ obtained previously, starting from an initial point $\mathbf{\Theta}^{(0)}$ until its convergence within a tolerance error ϵ or its reaching to the maximum number I_{sca} of iterations. The SCA with (18) guarantees a convergence to a local minimum since it provides a feasible solution from $\hat{\zeta}_k(\boldsymbol{\theta}_{s(k)}, \boldsymbol{\theta}_{s(k)}^{(q)}) \leq \zeta_k(\boldsymbol{\theta}_s)$ and the optimal value $\Gamma^{(q+1)}$ of problem (18) satisfies $\Gamma^{(q+1)} \geq \Gamma^{(q)}$ since (18) is a second-order cone program producing a global maximum at complexity $\mathcal{O}(N^3)$.

B. Sum Rate Analysis without QoS Constraints

We analyze the optimal sum rate without minimum QoS constraints, which is obtained by maximizing Γ_{sum} in (12a) that is achieved with maximum PA $p_k = P_k^{\max}$. Thus,

$$\Gamma_{\text{sum}} = \sum_{k\in\mathcal{K}} \frac{P_k^{\max}}{\sigma^2} \left|\mathbf{h}_k^T \boldsymbol{\theta}_{s(k)}\right|^2 = \sum_{k\in\mathcal{K}} \chi_k \zeta_k(\boldsymbol{\theta}_{s(k)}), \tag{19}$$

where χ_k and $\zeta_k(\cdot)$ are defined in (15) and (16), respectively. Let us express $\boldsymbol{\theta}_s = \text{diag}(\mathbf{a}_s) e^{j\boldsymbol{\phi}_s}$ using amplitudes $\mathbf{a}_s = [a_{s1}, a_{s2}, \cdots, a_{sN}]^T$ and phases $\boldsymbol{\phi}_s = [\phi_{s1}, \phi_{s2}, \cdots, \phi_{sN}]^T$ for $s \in \{t, r\}$, where $a_{sn} \geq 0$, $a_{tn}^2 + a_{rn}^2 = 1$, and $\phi_{sn} \in [0, 2\pi)$.

For an upper bound, we assume coherent channel matching as $\phi_{s(k)} = -\angle \mathbf{h}_k$ which leads to

$$\zeta_k(\boldsymbol{\theta}_{s(k)}) = \frac{|\mathbf{h}_k^T \text{diag}(\mathbf{a}_{s(k)}) e^{-j\angle \mathbf{h}_k})|^2}{\|\mathbf{h}_k\|_1^2} = (\tilde{\mathbf{g}}^T \mathbf{a}_{s(k)})^2, \tag{20}$$

where $\tilde{\mathbf{g}}$ is given by $[\tilde{\mathbf{g}}]_n = \frac{|[\mathbf{h}_k]_n|}{\|\mathbf{h}_k\|_1} = \frac{|[\mathbf{g}]_n|}{\|\mathbf{g}\|_1}$ from (8). An upper bound is then obtained as

$$\Gamma_{\text{sum}}^{\text{ub}} = \chi_{\mathcal{K}_t} (\tilde{\mathbf{g}}^T \mathbf{a}_t)^2 + \chi_{\mathcal{K}_r} (\tilde{\mathbf{g}}^T \mathbf{a}_r)^2, \tag{21}$$

where $\chi_{\mathcal{K}_s} = \sum_{k\in\mathcal{K}_s} \chi_k$ for $s \in \{t, r\}$. The first term of $\Gamma_{\text{sum}}^{\text{ub}}$ is an increasing function of a_{tn} whilst the second term is a decreasing function of a_{tn} from $a_{rn} = \sqrt{1 - a_{tn}^2}$. To find a_{tn} maximizing $\Gamma_{\text{sum}}^{\text{ub}}$, we obtain a critical point as

$$\frac{\partial \Gamma_{\text{sum}}^{\text{ub}}}{\partial a_{tn}} = 2|[\tilde{\mathbf{g}}]_n| \left(\chi_{\mathcal{K}_t} \tilde{\mathbf{g}}^T \mathbf{a}_t - \chi_{\mathcal{K}_r} \frac{a_{tn}}{\sqrt{1 - a_{tn}^2}} \tilde{\mathbf{g}}^T \mathbf{a}_r \right) = 0, \tag{22}$$

or equivalently $\frac{a_{tn}}{\sqrt{1-a_{tn}^2}} = \frac{\chi_{\mathcal{K}_t} \tilde{\mathbf{g}}^T \mathbf{a}_t}{\chi_{\mathcal{K}_r} \tilde{\mathbf{g}}^T \mathbf{a}_r}$, $\forall n$, which implies that a_{tn} is identical for $n \in \mathcal{N}$ as $a_{tn} = a_t$. Hence, we have

$$\Gamma_{\text{sum}}^{\text{ub}} = \chi_{\mathcal{K}_t} a_t^2 + \chi_{\mathcal{K}_r} (1 - a_t^2), \tag{23}$$

which is maximized when $a_t = 1$ if $\chi_{\mathcal{K}_t} \geq \chi_{\mathcal{K}_r}$ and $a_t = 0$, otherwise, resulting in the maximum value

$$\Gamma_{\text{sum}}^{\text{ub}\star} = \max(\chi_{\mathcal{K}_t}, \chi_{\mathcal{K}_r}). \tag{24}$$

This upper bound is tight when the phase vectors, $\angle \mathbf{h}_k$, or equivalently $\angle \mathbf{v}_k$, are similar for the users in $\mathcal{K}_s$ for $s \in \{t, r\}$.

For different channel phase vectors, we obtain lower bounds by using suboptimal STAR-RIS BF matched to the channel of the maximum SNR user given by $\boldsymbol{\theta}_{s_o} = e^{-j\angle \mathbf{h}_{k_o}}$ and $\boldsymbol{\theta}_{s\neq s_o} =$

0, where $k_o = \arg\max_{k\in\mathcal{K}} \chi_k$ and $s_o = s(k_o) \in \{t, r\}$. A lower bound expression is given by

$$\Gamma_{\text{sum}}^{\text{lb}} = \sum_{k\in\mathcal{K}_{s_o}} \chi_k \zeta_k (e^{-j\angle \mathbf{h}_{k_o}}), \tag{25}$$

which leads to two lower bounds, $\Gamma_{\text{sum}}^{\text{lb1}} = \chi_{k_o}$ and

$$\Gamma_{\text{sum}}^{\text{lb2}} = \chi_{k_o} + (\chi_{\mathcal{K}_{s_o}} - \chi_{k_o})\zeta_{\min}, \tag{26}$$

with $\zeta_{\min} = \min_{k,k_o} \zeta_k(e^{-j\angle \mathbf{h}_{k_o}})$ for $k_o \in \mathcal{K}_{s_o}$. Notably, $\Gamma_{\text{sum}}^{\text{ub}\star}$ and $\Gamma_{\text{sum}}^{\text{lb1}}$ do not depend on the channel model, NF or FF, whilst $\Gamma_{\text{sum}}^{\text{lb2}}$ depends on the channel model with $\zeta_{\min}$ signifying the channel effect. To elaborate $\zeta_{\min}$, we define the RIS-perspective channel correlation $\rho_{i,j}$ of $\mathbf{h}_i$ and $\mathbf{h}_j$ as

$$\rho_{i,j} \triangleq \frac{\mathbf{h}_j^H \mathbf{h}_i}{\|\mathbf{h}_i\|_1 \|\mathbf{h}_j\|_1} = N\mathbf{b}^H(\mathbf{u}_j)\text{diag}(\tilde{\mathbf{g}})\mathbf{b}(\mathbf{u}_i). \tag{27}$$

Note that $|\rho_{k,k_o}|^2 = \zeta_k(e^{-j\angle \mathbf{h}_{k_o}})$ can be interpreted as a performance loss by sharing the STAR-RIS coefficients among multiple users with $\zeta_{\min} = \min_{k,k_o} |\rho_{k,k_o}|^2$. This channel correlation does not depend on the phases of the channel $\mathbf{g}$ common in $\{\mathbf{h}_k\}_{k\in\mathcal{K}}$ but depends on the array response $\mathbf{b}(\mathbf{u}_k)$ of the channel $\mathbf{v}_k$ between the STAR-RIS and user k. In addition, $\Gamma_{\text{sum}}^{\text{lb2}}$ with NF channels $\mathbf{v}_k$ would be lower than that with FF channels from the following observations. i) The channel correlation depends solely on $\mathbf{b}(\mathbf{u}_k)$ as $\rho_{i,j} = \mathbf{b}^H(\mathbf{u}_j)\mathbf{b}(\mathbf{u}_i)$ if the entries of $\mathbf{g}$ have the same amplitude, leading to $\text{diag}(\tilde{\mathbf{g}}) = \frac{1}{N}\mathbf{I}_N$. ii) For users (i, j) aligned at the same azimuth angle, $\rho_{i,j}^{\text{FF}} = 1$ for FF $\mathbf{v}_i$ and $\mathbf{v}_j$ with $\mathbf{b}(\mathbf{u}_i) = \mathbf{b}(\mathbf{u}_j)$ whilst $\rho_{i,j}^{\text{NF}} < 1$ for NF $\mathbf{v}_i$ and $\mathbf{v}_j$.

IV. Performance Evaluation

We evaluate the performance through simulation and analysis when the BS and STAR-RIS are located at $[40, 30, 0]$ and $[0, 0, 0]$ in meters (m), respectively. The user location $\mathbf{u}_k$ in (1) is determined by (r_k, φ_k) assuming $\vartheta_k = \frac{\pi}{2}$ for $k \in \mathcal{K}$. We set $f_c = \frac{3\times10^8}{\lambda_c} = 10$ GHz and $d = \frac{\lambda_c}{2} = 0.015$ m. NLoS $\mathbf{g}$ is generated with $L = 4$, $\beta_0^{\text{NLoS}} = 10^{-3} r_0^{-2.2}$, $\alpha_{0l} \sim \mathcal{CN}(0, 1)$, and $\varphi_{0l} \sim \mathcal{U}(-\frac{\pi}{2}, \frac{\pi}{2})$. We set the maximum transmit power to $P_k^{\max} = 23$ dBm, noise power to $\sigma^2 = -100$ dBm, maximum SCA iterations to $I_{sca} = 10$, and error tolerance to $\epsilon = 10^{-4}$.

We first explore the channel correlation $|\rho_{1,2}|$ of two users at the same azimuth angle $\varphi_1 = \varphi_2 = 0$ in Fig. 2 as distance r_1 varies while fixing $r_2 = 7$ m when $N = 512$. NF and FF denote the exact model (2) and its FF approximation (4) for $\mathbf{v}_k$ whilst $\mathbf{g}_{\text{LoS}}$ and $\mathbf{g}_{\text{NLoS}}$ indicate the LoS and NLoS models having identical and nonidentical amplitudes, respectively. The channel correlation $|\rho_{1,2}|$ of NF decreases as $|r_1 - 7|$ increases with a steeper slope for a larger N_y, implying a better azimuth beam resolution. It is also affected by the amplitude variation in $\mathbf{g}$ but only slightly. However, for FF, $|\rho_{1,2}| = 1$ irrespective of r_1, STAR-RIS shape N_y, and channel model for $\mathbf{g}$. We further illustrate $|\rho_{1,2}|$ in Fig. 3 as $(r_1, \varphi_1) = (r, \varphi)$ varies when $(r_2, \varphi_2) = (7, 0)$ for $\mathbf{g}_{\text{LoS}}$. The NF effect showing low channel correlation at different distances of the same angle is observed with $N_y = 512$ but not with $N_y = 64$.

We next evaluate the sum rate of users aligned in the same angle with $\varphi_k = 0$ for $k \in \mathcal{K}_r$ and $\varphi_k = \pi$ for $k \in \mathcal{K}_t$ using $N = N_y$ (unless stated otherwise) to maximize the NF effect.

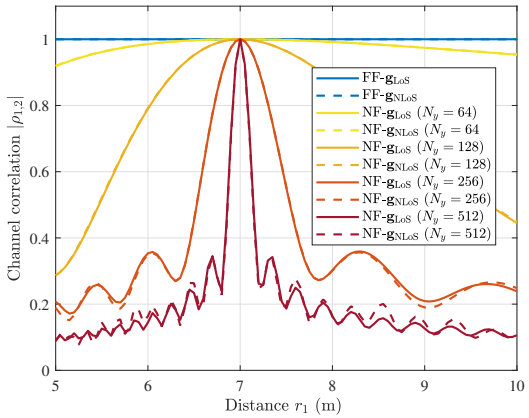
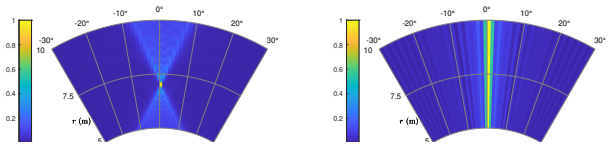


Fig. 2: Channel correlation of two users at the same azimuth angle but with different distances of r_1 and $r_2 = 7$ m.



(a) $(N_y, N_z) = (512, 1)$ (b) $(N_y, N_z) = (64, 8)$

Fig. 3: Channel correlation between user 1 at (r, φ) and user 2 at $(7, 0)$.

The average sum rate of $K = 8$ users is provided in Fig. 4 when $r_k \sim \mathcal{U}(5, 10)$ for T users and $r_k \sim \mathcal{U}(10, 15)$ for R users. The performance is shown as the number N of STAR-RIS elements increases when $R_k^{\min} = 0.5$ bps/Hz in Fig. 4(a) and as the rate requirement $R_k^{\min}$ increases when $N = 512$ for $N_y = 512$ and 64 in Fig. 4(b). ES (Prop) and ES (AO-SDR) denote the ES-NOMA using the proposed algorithm and conventional SDR-based AO in [5], respectively. TS (Opt) denotes the TS-NOMA in [6] which is optimized through a bisection search of the optimal TS factor with STAR-RIS coefficients and PA optimized as in the proposed method. MS (C-RIS) utilizes $N/2$ elements for T and R coefficients, respectively, as in [9], which is regarded as a conventional method with reflecting-only and transmitting-only RISs having $N/2$ elements for each as adopted in [15]. In Fig. 4(a), the performance of the proposed algorithm is indistinguishable from that of AO-SDR which is provided up to $N = 64$ due to its intolerable time cost for a large N; The computational time per channel realization averaged over 50 realizations is measured as 9.21 s, 32.2 s, and 62.2 s with ES (Prop) and 64.6 s, 944 s, and 46000 s with ES (AO-SDR), for $N = 64$, 128, and 256, respectively, with Intel(R) Core(TM) i9-12900K@3.20 GHz. We then use the proposed algorithm showing a similar performance with AO-SDR when $N = 32$ and 64 for a large N. The sum rate of all methods increases as N increases with a gain of the ES over TS and MS counterparts. The gain also increases as $R_k^{\min}$ increases, as observed in Fig. 4(b). The sum rate with $N_y = 512$ exhibiting stronger NF effect is worse than that with $N_y = 64$ although their gap is reduced as the minimum rate requirement increases. It implies that the NF effect is adverse to the sum rate of the NOMA but is favorable to the minimum rate requirement by reducing the interference.

We investigate the NF effect on the average sum rate of

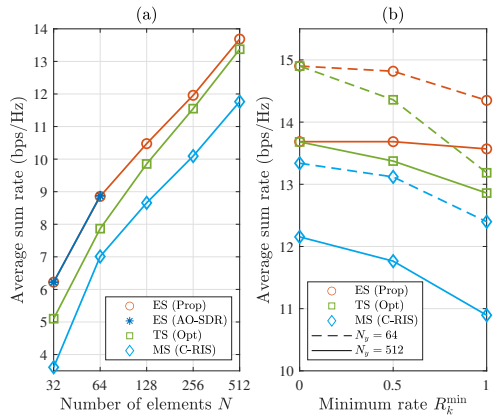


Fig. 4: Average sum rate of $K = 8$ users (a) as N increases with $R_k^{\min} = 0.5$ bps/Hz and (b) as $R_k^{\min}$ increases with $N = 512$.

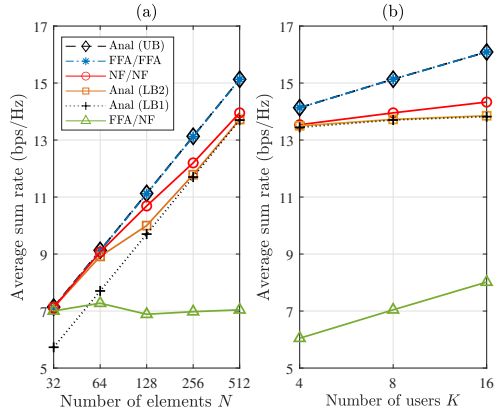


Fig. 5: Average sum rate with $R_k^{\min} = 0$ and $(N_y, N_z) = (N, 1)$ (a) as N increases for $K = 8$ and (b) as K increases for $N = 512$.

the ES-NOMA without minimum QoS constraints through analysis and simulation in Fig. 5 with $(N_y, N_z) = (N, 1)$ and $r_k \sim \mathcal{U}(5, 10)$ for all users. The sum rate is compared with respect to N when $K = 8$ in Fig. 5(a) and with respect to K when $N = 512$ in Fig. 5(b). Anal (UB), Anal (LB1), and Anal (LB2) indicate the analytical results with $\Gamma_{\text{sum}}^{\text{ub}}$, $\Gamma_{\text{sum}}^{\text{lb1}}$, and $\Gamma_{\text{sum}}^{\text{lb2}}$, respectively, whilst NF/NF and FFA/NF denote the simulation results of the proposed algorithm with perfect CSI and FF approximated (FFA) CSI, respectively. FFA/FFA denotes the wrong performance evaluated in FFA channels with their perfect CSI instead of true NF channels, which captures a performance discrepancy using a wrong channel and also validates the upper bound. The figures confirm that UB and LB1 provide upper and lower bounds irrelevant to the channel model but LB2 computed with the channel correlation provides a tighter lower bound for NF/NF. The sum rate increases as N or K increases, where the actual performance (NF/NF) tends to deviate from UB for larger N due to the NF effect. The performance of FFA/NF implies that the FFA CSI does not work in NF channels. Thus, the correct channel model and correct CSI should be employed to avoid a performance overestimation using FFA/FFA and a performance loss using FFA/NF.

V. Concluding Remarks

To identify the effect of the NF channel on the sum rate of a STAR-RIS aided uplink NOMA with single-antenna nodes, we have proposed a QoS-constrained sum rate maximization algorithm that can accommodate the XL-STAR-RIS and have analyzed the sum rate without QoS constraints. The analysis revealed that low channel correlations with the NF effect degrades the sum rate. The advantages of the ES-NOMA with the proposed algorithm were shown by comparing the performance and complexity with the benchmarks through simulation. In addition, the NF effect was shown to lower the sum rate slope of the STAR-RIS aided uplink NOMA with respect to the number of STAR-RIS elements, which entails the use of the exact channel model in performance investigation. In this regard, the NF effect on the performance of the downlink counterpart and extended networks with multiantenna BS and users would merit further investigation as future work.

References

[1] C. Pan *et al.*, "Reconfigurable intelligent surfaces for 6G systems: Principles, applications, and research directions," *IEEE Commun. Mag.*, vol. 59, no. 6, pp. 14–20, Jun. 2021.

[2] Y. Liu, W. Yi, Z. Ding, X. Liu, O. A. Dobre, and N. Al-Dhahir, "Developing NOMA to next generation multiple access: Future vision and research opportunities," *IEEE Wirel. Commun.*, vol. 29, no. 6, pp. 120–127, Dec. 2022.

[3] M. Cui, Z. Wu, Y. Lu, X. Wei, and L. Dai, "Near-field MIMO communications for 6G: Fundamentals, challenges, potentials, and future directions," *IEEE Commun. Mag.*, vol. 61, no. 1, pp. 40–46, Jan. 2023.

[4] Q. Wu, S. Zhang, B. Zheng, C. You, and R. Zhang, "Intelligent reflecting surface-aided wireless communications: A tutorial," *IEEE Trans. Commun.*, vol. 69, no. 5, pp. 3313–3351, Apr. 2021.

[5] M. Zeng, X. Li, G. Li, W. Hao, and O. A. Dobre, "Sum rate maximization for IRS-assisted uplink NOMA," *IEEE Commun. Lett.*, vol. 25, no. 1, pp. 234–238, Jan. 2021.

[6] H. Ma, H. Wang, H. Li, and Y. Feng, "Transmit power minimization for STAR-RIS-empowered uplink NOMA system," *IEEE Wirel. Commun. Lett.*, vol. 11, no. 11, pp. 2430–2434, Nov. 2022.

[7] J. Zuo, Y. Liu, Z. Ding, L. Song, and H. V. Poor, "Joint design for simultaneously transmitting and reflecting (STAR) RIS assisted NOMA systems," *IEEE Trans. Wirel. Commun.*, vol. 22, no. 1, pp. 611–626, Jan. 2023.

[8] M. Ahmed, et al. , "A survey on STAR-RIS: Use cases, recent advances, and future research challenges," *IEEE Internet Things J.*, vol. 10, no. 16, pp. 14 689–14 711, Aug. 2023.

[9] Q. Li, M. El-Hajjar, Y. Sun, I. Hemadeh, A. Shojaeifard, Y. Liu, and L. Hanzo, "Achievable rate analysis of the STAR-RIS-aided NOMA uplink in the face of imperfect CSI and hardware impairments," *IEEE Trans. Commun.*, vol. 71, no. 10, pp. 6100–6114, Jun. 2023.

[10] H. Zhang, N. Shlezinger, F. Guidi, D. Dardari, M. F. Imani, and Y. C. Eldar, "Beam focusing for near-field multiuser MIMO communications," *IEEE Trans. Wirel. Commun.*, vol. 21, no. 9, pp. 7476–7490, Sep. 2022.

[11] Y. Liu, Z. Wang, J. Xu, C. Ouyang, X. Mu, and R. Schober, "Near-field communications: A tutorial review," *IEEE Open J. Commun. Society*, vol. 4, pp. 1999–2049, Apr. 2023.

[12] Y. Cheng, C. Huang, W. Peng, M. Debbah, L. Hanzo, and C. Yuen, "Achievable rate optimization of the RIS-aided near-field wideband uplink," *IEEE Trans. Wirel. Commun.*, vol. 23, no. 3, pp. 2296–2311, Mar. 2024.

[13] D. Shen, L. Dai, X. Su, and S. Suo, "Multi-beam design for near-field extremely large-scale RIS-aided wireless communications," *IEEE Trans. Green Commun. Netw.*, vol. 7, no. 3, pp. 1542–1553, Sep. 2023.

[14] J. Xu, X. Mu, and Y. Liu, "Exploiting STAR-RISs in near-field communications," *IEEE Trans. Wirel. Commun.*, vol. 23, no. 3, pp. 2181–2196, Mar. 2024.

[15] H. Li, Y. Liu, X. Mu, Y. Chen, Z. Pan, and Y. C. Eldar, "Near-field beamforming for STAR-RIS networks," Jun. 2023, arXiv:2306.14587.

[16] X. Mu, J. Xu, Y. Liu, and L. Hanzo, "Reconfigurable intelligent surface-aided near-field communications for 6G: Opportunities and challenges," *IEEE Veh. Technol. Mag.*, vol. 19, no. 1, pp. 65–74, Mar. 2024.

[17] Z. Wei, L. Yang, D. W. K. Ng, J. Yuan, and L. Hanzo, "On the performance gain of NOMA over OMA in uplink communication systems," *IEEE Trans. Commun.*, vol. 68, no. 1, pp. 536–568, Jan. 2020.