

STAR-RIS辅助URLLC-NOMA系统的联合波束成形设计

朱建月^{*①} 吴雨桐^② 陈晓^③ 谢亚琴^① 许尧^① 张治中^①

^①(南京信息工程大学电子与信息工程学院 南京 210044)

^②(东南大学网络空间安全学院 南京 211189)

^③(南京信息工程大学人工智能学院(未来技术学院) 南京 210044)

摘要: 针对超可靠低时延通信(URLLC)场景, 该文研究了融合透射与反射功能的智能超表面(STAR-RIS)辅助的非正交多址接入(NOMA)系统的传输设计。具体而言, 该文联合设计了基站端的波束成形向量、RIS端的透射相移矩阵和反射相移矩阵, 以在满足基站总功率约束的条件下实现能耗最小化。为解决所提出的非凸问题, 该文首先分析了有限块长传输下的用户速率函数特性, 并据此将优化问题进行等价转换。随后, 采用交替优化和半正定松弛(SDR)方法来解决联合波束设计问题。实验结果表明, 与正交多址接入和传统RIS方案相比, 所提出的方法在能耗性能上有显著提升。

关键词: 智能反射面; 非正交多址接入技术; 波束成形设计

中图分类号: TN929.5

文献标识码: A

文章编号: 1009-5896(2025)02-0409-09

DOI: 10.11999/JEIT240717

1 引言

随着用户和设备数量的爆炸式增长以及新兴应用的不断涌现, 第六代(the 6th Generation, 6G)网络对数据速率、时延和大规模连接的要求变得更加严苛^[1]。为了应对这些挑战, 现有的工作对多种前沿技术展开了深入的探讨与研究, 例如大规模MIMO技术、太赫兹通信技术和人工智能相关技术^[2]。由于具有低功耗、低成本和灵活部署的优势, 可重构智能超表面(Reconfigurable Intelligent Surfaces, RIS)已成为6G通信系统中的重要候选技术之一^[3,4]。RIS通过智能调控电磁波的传播路径, 可以显著提升信号覆盖范围和质量, 从而有效应对复杂多变的通信环境。这些技术的有机结合和创新应用, 不仅有望满足未来6G网络的高性能需求, 还将推动全球通信网络的发展与变革。

为了进一步提高通信系统的性能, RIS可以与其他关键技术相结合, 其中之一是非正交多址接入(Non-Orthogonal Multiple Access, NOMA)技术。NOMA的核心思想是通过用户信道条件的差异实现通信资源的复用, 从而能够充分利用RIS灵活调节用户信道的特性实现性能的提升^[5]。文献^[6]基于RIS辅助的NOMA系统, 通过设计RIS上的被动波束成形来服务于成对功率域NOMA用户以提高无线网络的频谱效率和能量效率。文献^[7]通过优化所有用户的最小信噪比来提升RIS辅助NOMA系统的通信速率和用户公平性。文献^[8]通过联合优化波束成形向量和RIS相移矩阵, 以最小化RIS辅助NOMA系统的能耗。

然而, 传统的RIS往往只通过反射信号来增强传输, 这实际上限制了RIS的辅助范围^[9]。为了解决这个问题, 文献^[10]提出了一种融合透射与反射功能的智能超表面(Simultaneous Transmitting And Reflecting Reconfigurable Intelligent Surface, STAR-RIS)。与仅支持反射的RIS不同, STAR-RIS能够在360°覆盖范围内为用户提供服务, 目前已有很多的工作致力于STAR-RIS辅助无线网络的研究^[11,12]。实际上, 相比于仅有反射功能的RIS, STAR-RIS能够更加灵活高效地调节用户的信道, 助力于NOMA网络性能的进一步提升。

当前, 很多的工作对STAR-RIS辅助的NOMA网络进行了深入的研究与探讨。例如, 文献^[13]在STAR-RIS辅助的NOMA网络中, 通过联合优化解码顺序、功率分配系数、主动波束成形以及被动波束成形, 以最大化系统的可达和速率。文献^[14]主

收稿日期: 2024-08-19; 改回日期: 2025-02-14; 网络出版: 2025-02-24

*通信作者: 朱建月 zhujy@nuist.edu.cn

基金项目: 国家自然科学基金(62101273), 江苏省自然科学基金(BK20220439, BK20220438), 江苏省高等学校基础科学(自然科学)研究面上项目(22KJB510005, 22KJB510033), 江苏省重点研发计划(BE2023088), 江苏省双创团队((2023)200008号)

Foundation Items: The National Natural Science Foundation of China (62101273), The Natural Science Foundation of Jiangsu Province of China (BK20220439, BK20220438), The Natural Science Foundation of the Jiangsu Higher Education Institutions of China (22KJB510005, 22KJB510033), Jiangsu Provincial Key Research and Development Program (BE2023088), Jiangsu Provincial Innovation and Entrepreneurship Team ((2023)200008)

要对STAR-RIS辅助双用户通信网络的覆盖范围进行了研究,并针对NOMA和正交多址接入(Orthogonal Multiple Access, OMA)分别提出了一个总覆盖范围最大化问题,验证了STAR-RIS以及NOMA的有效性。文献[15]研究了NOMA辅助STAR-RIS系统中的能效问题,通过联合优化基站的波束成形向量和STAR-RIS端的相移矩阵来最大化系统能效。文献[16]通过联合优化STAR-RIS的位置和相移阵,提出了两种NOMA方案以最大化用户加权和速率。文献[17]针对传统RIS无法透射信号的问题,利用STAR-RIS,提出了一种基于NOMA的信号增强算法,以提升系统的频谱效率。

然而,基于STAR-RIS辅助NOMA网络,大多数现有的研究考虑的是无限块长传输场景。实际上,超可靠低时延(Ultra-Reliable Low-Latency Communication, URLLC)场景是下一代移动通信的典型场景之一^[18],因此对有限块长(Finite Blocklength, FBL)传输的研究势在必行^[19]。FBL传输下的可达速率函数综合考虑了信干噪比(Signal to Interference plus Noise Ratio, SINR)、块错误率(Block Error Rate, BLER)和块长度,比传统的香农容量更为复杂^[20]。因此,复杂的速率函数给FBL传输场景下的优化问题带来了巨大挑战^[21]。在该文中,将通过联合优化基站端波束成形、RIS传输矩阵和RIS反射矩阵来最小化STAR-RIS辅助NOMA系统的传输功率。为了解决所提出的非凸问题,对波束成形和RIS矩阵进行了交替优化。该文的贡献总结如下:

(1) 该文首先分析了在有限块长传输下可达速率函数的性质,并将用户的服务质量约束进行了等价转换,简化了优化过程。

(2) 基于系统传输功率最小化的目标,对基站端主动波束成形向量和反射端被动波束成形向量进行联合优化,并提供了高效的联合波束设计算法。

(3) 通过实验结果验证了所提出联合波束成形设计方案的有效性,并对比了STAR-RIS在不同模式下的能耗性能。

2 系统模型与问题描述

2.1 系统模型

该文所研究的STAR-RIS辅助NOMA URLLC系统模型如图1所示。该系统中,具有 M 根天线的基站为2个单天线用户提供服务,并且传输通过具有 N 个元素的RIS进行增强。特别地,中心用户,即用户1,位于反射区域,STAR-RIS将以反射模式运行。同时,边缘用户,即用户2,位于透射区域,STAR-RIS将以透射模式运行。对于 $k = 1, 2$, s_k 表

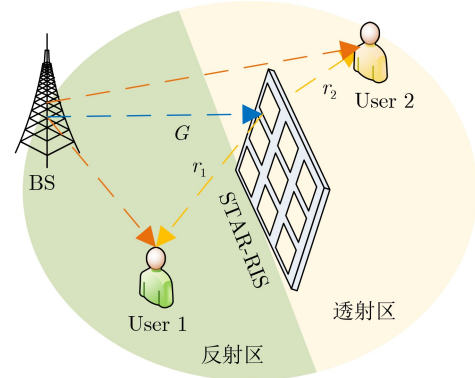


图1 STAR-RIS辅助NOMA URLLC系统模型

示发送给用户 k 的信号并满足 $E[s_k^2] = 1$, $\mathbf{w}_k \in \mathbb{C}^M$ 表示用户 k 的波束成形向量。基于NOMA的准则,在STAR-RIS辅助的通信系统中,基站发送叠加码,从而用户 k 的接收信号为

$$y_k = \mathbf{h}_k^H \sum_{j=1}^2 \mathbf{w}_j s_j + n_k, k = 1, 2 \quad (1)$$

其中, $n_k \sim \mathcal{CN}(0, \sigma^2)$ 为噪声项。在式(1)中, $\mathbf{h}_k$, $k = 1, 2$, 是用户 k 的信道, 具体由下式给出

$$\mathbf{h}_1^H = \mathbf{r}_1^H \boldsymbol{\Theta}_r \mathbf{G} + \mathbf{h}_{d1}^H \quad (2)$$

$$\mathbf{h}_2^H = \mathbf{r}_2^H \boldsymbol{\Theta}_t \mathbf{G} + \mathbf{h}_{d2}^H, \quad (3)$$

在式(2)和式(3)中, $\mathbf{h}_{dk}^H \in \mathbb{C}^{1 \times M}$ 表示基站和用户 k 之间的信道, $\mathbf{r}_k^H \in \mathbb{C}^{1 \times N}$ 表示RIS和用户 k 之间的信道, $\mathbf{G} \in \mathbb{C}^{N \times M}$ 表示基站和RIS之间的信道。同时,在STAR-RIS端,透射相移阵和反射相移阵分别定义为 $\boldsymbol{\Theta}_t = \sqrt{\beta_t} \text{diag}(e^{j\theta_1^t}, e^{j\theta_2^t}, \dots, e^{j\theta_N^t})$ 和 $\boldsymbol{\Theta}_r = \sqrt{\beta_r} \text{diag}(e^{j\theta_1^r}, e^{j\theta_2^r}, \dots, e^{j\theta_N^r})$, 其中 $\theta_n^r, \theta_n^t \in [0, 2\pi)$, $n = 1, 2, \dots, N$, $\beta_t, \beta_r \in [0, 1]$ 并满足 $\beta_t + \beta_r = 1$ ^[10]。

在STAR-RIS辅助的NOMA系统中,该文首先假设用户1和用户2分别为小区中心用户和小区边缘用户。从而基于NOMA准则,中心用户1可实行串行干扰消除技术,即用户1将首先对用户2的信号进行解码,并从中接收到的信号中减去该信号再解码得到自己的信号。相反,对于小区边缘用户2,其将直接解码其自身的信号,同时将用户1的信号作为干扰处理。因此,在NOMA传输方案下,用户1和用户2的信干噪比可表示为

$$\gamma_1 = \frac{|\mathbf{h}_1^H \mathbf{w}_1|^2}{\sigma^2} \quad (4)$$

$$\gamma_2 = \min \left\{ \frac{|\mathbf{h}_2^H \mathbf{w}_2|^2}{|\mathbf{h}_2^H \mathbf{w}_1|^2 + \sigma^2}, \frac{|\mathbf{h}_1^H \mathbf{w}_2|^2}{|\mathbf{h}_1^H \mathbf{w}_1|^2 + \sigma^2} \right\} \quad (5)$$

此外,需要注意到,传统的香农公式是基于假设块长趋近于无穷大且解码错误概率趋近于零的前

提条件而建立的。然而,在有限块长传输下,与传统的香农容量不同,用户的可达速率是用户信干噪比、译码错误概率以及块长的函数^[20]。鉴于该文所考察的URLLC场景中,大量的物联网设备需要使用短数据包进行通信,因此将会考虑有限块长传输,进而用户的速率将由式(6)给出

$$R_k = \ln(1 + \gamma_k) - \sqrt{\frac{V(\gamma_k)}{l_k}} Q^{-1}(\varepsilon), \quad k = 1, 2, \quad (6)$$

其中 l_k 表示用户 k 的块长, ε 表示解码错误概率, $Q^{-1}(\cdot)$ 表示高斯 Q 函数,即 $Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty \exp\left(-\frac{t^2}{2}\right) dt$,的逆函数, $V(\gamma)$ 表示信道色散函数,定义为 $V(\gamma) = 1 - \frac{1}{(1 + \gamma)^2}$ 。

2.2 问题描述

随着通信技术的不断进步,人们对于更快速、更可靠的连接需求不断增长。然而,这也伴随着更高的能耗挑战。高能耗不仅增加了运营成本,还对环境造成了不可忽视的影响。在如今注重可持续发展的时代,减少通信设备的能耗对于降低碳排放、保护环境至关重要。因此,该文旨在通过联合优化基站端的主动波束成形向量以及STAR-RIS端的被动波束成形向量,最大限度地降低STAR-RIS辅助MISO-NOMA系统的能耗。所考察的优化问题建立如下:

$$\begin{aligned} & \min_{\{\Theta_t, \Theta_r, \mathbf{w}_1, \mathbf{w}_2\}} \|\mathbf{w}_1\|^2 + \|\mathbf{w}_2\|^2 \\ & \text{s.t. C1: } R_1 \geq r_1^{\text{th}} \\ & \quad \text{C2: } R_2 \geq r_2^{\text{th}} \\ & \quad \text{C3: } \beta_t, \beta_r \in [0, 1], \beta_r + \beta_t = 1 \\ & \quad \text{C4: } 0 \leq \theta_n^r, \theta_n^t < 2\pi, n = 1, 2, \dots, N \end{aligned} \quad (7)$$

其中, r_1^{th} 和 r_2^{th} 分别为用户1和用户2的最小传输速率需求。在问题式(7)中,约束1和约束2为用户的QoS约束,约束3是基于STAR-RIS端的能量守恒定律,约束4为相移角约束。由于式(6)中定义的速率表达式相对于香农公式更为复杂,因此问题式(7)为一个非凸优化问题,这意味着难以直接获得其解决方案。在下文中,将首先对有限块长传输下速率函数的性质进行分析,然后基于这些分析结果,对基站端的主动波束成形向量和RIS端的被动波束成形向量进行迭代优化,即首先固定RIS端的相移阵来优化基站端的波束成形向量,随后再进一步优化RIS端的相移阵。

3 主被动波束成形联合优化

在本节中,考虑到式(7)中QoS约束的非凸性,

将首先分析有限块长传输下速率函数的性质,并基于分析的结果对QoS约束进行等价转换。然后进一步地使用迭代优化的方法对基站端波束成形向量以及RIS端的相移阵进行联合优化设计。

3.1 有限块长传输下速率函数分析

本小节将重点分析有限块长传输下的速率函数,即

$$R(\gamma) = \ln(1 + \gamma) - \Psi \sqrt{V(\gamma)} \quad (8)$$

其中 $\Psi = \frac{Q^{-1}(\varepsilon)}{\sqrt{l}}$ 。注意到, $Q^{-1}(\varepsilon)$ 是关于 ε 的减函数并且 $Q^{-1}(0) > 0$ 。与此同时,在URLLC系统中,理想的译码错误概率 ε 应小于 10^{-6} ^[22]。因此在不失一般性的前提下,可假设 $Q^{-1}(\varepsilon) > 0$,也就是 $\Psi = \frac{Q^{-1}(\varepsilon)}{\sqrt{l}} > 0$ 。

定理1 当 $\Psi > 0$ 时, $R(\gamma)_{\min} = R(\varphi_0) < 0$,其中 $\varphi_0 = \sqrt{\frac{1 + \sqrt{1 + 4\Psi^2}}{2}} - 1$ 。

证明: $R(\gamma)$ 的一阶导数可以表示为

$$R'(\gamma) = \frac{1}{1 + \gamma} \left(1 - \frac{\Psi}{(1 + \gamma)\sqrt{(1 + \gamma)^2 - 1}} \right) \quad (9)$$

令 $R'(\gamma) = 0$,其解为 $\varphi_0 = \sqrt{\frac{1 + \sqrt{1 + 4\Psi^2}}{2}} - 1$ 。

进一步地, $R(\gamma)$ 的2阶导数可以表示为

$$R''(\gamma) = -\frac{1}{(1 + \gamma)^2} \left[1 - \frac{\Psi \left(3(1 + \gamma)^2 - 2 \right)}{(1 + \gamma)\left((1 + \gamma)^2 - 1 \right)^{\frac{3}{2}}} \right] \quad (10)$$

从而,易得 $R''(\varphi_0) > 0$,即 φ_0 为 $R(\gamma)$ 的极小值点。此外,根据 $R(0) = 0$,可以得到 $R(\varphi_0) < 0$ 。

定理1 给出了短包传输下用户速率函数的性质分析。然而,由于速率表达式的复杂性,仍然无法给出 $R_k = r_k^{\text{th}}, k = 1, 2$ 的解,从而无法转换问题式(7)中的QoS约束。因此,本文将进一步给出如下的结论。

定理2 当 $\Psi > 0$ 时,对于 $\gamma > 0$, $R(\gamma) = a (a \geq 0)$ 的解可以表示为

$$\gamma^* = e^{a + \frac{\mathcal{W}(2\Psi, -2\Psi; -4\Psi^2 e^{-2a})}{2}} - 1 \quad (11)$$

其中, $\mathcal{W}(\iota_1, \iota_2; \mu) = \iota_1 - \sum_{m=1}^{\infty} \frac{1}{m \times m!} \left(\frac{\mu m e^{-\iota_1}}{\iota_2 - \iota_1} \right)^m \times B_{m-1} \left(\frac{-2}{m(\iota_2 - \iota_1)} \right)$, $B_m(z) = \sum_{k=0}^m \frac{(m+k)!}{k!(m-k)!} \left(\frac{z}{2} \right)^k$ 。

证明 $R(\gamma) = a$ 即 $\ln(1 + \gamma) - \Psi \sqrt{V(\gamma)} = a$ 可以重写为

$$\ln(e^{-a}(1 + \gamma)) - \Psi \sqrt{V(\gamma)} = 0, \quad (12)$$

令 $e^{-a} = \alpha$, $\ln(\alpha(1+\gamma)) = \beta$, 则有

$$(1+\gamma)^{-2} = \alpha^2 e^{-2\beta}, \quad (13)$$

此外, 由式(12)中, $\beta = \Psi\sqrt{V(\gamma)}$ 且 $\gamma \geq 0$, 因此可以发现 $0 < \beta \leq \Psi$ 。进一步来说, 还可以将式(13)重写为

$$\beta = \Psi\sqrt{1 - \alpha^2 e^{-2\beta}} \quad (14)$$

通过一些基本的数学运算后, 式(14)可以转换成

$$-\frac{1}{\alpha^2 \Psi^2}(\beta - \Psi)(\beta + \Psi) = e^{-2\beta} \quad (15)$$

这属于一类超越方程(广义Lambert W函数)^[22]。再令 $\eta = 2\beta$, $0 < \eta \leq 2\Psi$, 式(15)可以重新表述为

$$e^\eta(\eta - 2\Psi)(\eta + 2\Psi) = -4\alpha^2 \Psi^2 \quad (16)$$

在 $0 < \eta \leq 2\Psi$ 范围内, 利用拉格朗日反演定理^[23]可得到式(16)的解:

$$\eta^* = \mathcal{W}(2\Psi, -2\Psi; -4\alpha^2 \Psi^2) \quad (17)$$

其中, $\mathcal{W}(\iota_1, \iota_2; \mu)$ 和 $B_m(z)$ 分别被定义为

$$\begin{aligned} \mathcal{W}(\iota_1, \iota_2; \mu) &= \iota_1 - \sum_{m=1}^{\infty} \frac{1}{m * m!} \left(\frac{\mu m e^{-\iota_1}}{\iota_2 - \iota_1} \right)^m \\ &\quad \times B_{m-1} \left(\frac{-2}{m(\iota_2 - \iota_1)} \right) \end{aligned} \quad (18)$$

$$B_m(z) = \sum_{k=0}^m \frac{(m+k)!}{k!(m-k)!} \left(\frac{z}{2} \right)^k \quad (19)$$

因此, $R(\gamma) = a$ 的解可以表示为

$$\gamma^* = e^{\alpha+\beta^*} - 1 = e^{\alpha+\frac{\eta^*}{2}} - 1 = e^{a+\frac{\mathcal{W}(2\Psi, -2\Psi; -4\Psi^2 e^{-2a})}{2}} - 1 \quad (20)$$

推论 为了保证问题式(7)可行, 需满足 $\gamma_k^{\text{th}} \geq \varphi_2$, $k = 1, 2$, 其中 $\gamma_k^{\text{th}} = e^{r_k^{\text{th}} + \frac{\mathcal{W}(2\Psi_k, -2\Psi_k; -4\Psi_k^2 e^{-2r_k^{\text{th}}})}{2}}$, -1 , $\Psi_k = \frac{Q^{-1}(\varepsilon)}{\sqrt{l_k}}$ 。

证明 根据定理1和定理2可以得到 $R(\gamma) = 0$ 时的两个解分别为 $\varphi_1 = 0$ 和 $\varphi_2 = e^{\frac{\mathcal{W}(2\Psi, -2\Psi; -4\Psi^2)}{2}} - 1$, 且当 $\varphi_1 < \gamma < \varphi_2$ 时, $R(\gamma) < 0$ 。因此, 为了保证问题式(7)可行, 即 $R(\gamma) \geq 0$, 需满足 $\gamma_k^{\text{th}} \geq \varphi_2$, $k = 1, 2$, 其中 γ_k^{th} 为用户 k 的目标数据速率对应的信噪比。

因此基于定理1和定理2的分析, 用户的QoS约束, 即 $R_k \geq r_k^{\text{th}}$, $k = 1, 2$, 可等价转换为

$$\gamma_k \geq \gamma_k^{\text{th}}, k = 1, 2, \quad (21)$$

从而问题式(7)可以等价地转化为

$$\begin{aligned} &\min_{\{\Theta_t, \Theta_r, \mathbf{w}_1, \mathbf{w}_2\}} \|\mathbf{w}_1\|^2 + \|\mathbf{w}_2\|^2 \\ &\text{s.t. } \gamma_1 \geq \gamma_1^{\text{th}} \\ &\quad \gamma_{2 \rightarrow 1} \geq \gamma_2^{\text{th}} \\ &\quad \gamma_{2 \rightarrow 2} \geq \gamma_2^{\text{th}} \\ &\quad \beta_t, \beta_r \in [0, 1], \beta_r + \beta_t = 1 \\ &\quad 0 \leq \theta_n^r, \theta_n^t < 2\pi, n = 1, 2, \dots, N \end{aligned} \quad (22)$$

问题(22)的解决难点在于基站端主动波束成形向量与RIS端被动波束成形向量之间的耦合。该文将采用迭代优化的方法, 以高效地获得优化解。

3.2 基站端主动波束成形设计

本小节首先将在给定RIS端被动波束成形向量, 即相移阵 Θ_t 和 Θ_r 的前提下, 对基站端主动波束成形向量进行优化, 此时的优化问题可转换为

$$\begin{aligned} &\min_{\{\mathbf{w}_1, \mathbf{w}_2\}} \|\mathbf{w}_1\|^2 + \|\mathbf{w}_2\|^2 \\ &\text{s.t. } \gamma_1 \geq \gamma_1^{\text{th}} \\ &\quad \gamma_{2 \rightarrow 1} \geq \gamma_2^{\text{th}} \\ &\quad \gamma_{2 \rightarrow 2} \geq \gamma_2^{\text{th}} \end{aligned} \quad (23)$$

为了解决问题(23), 引入半正定矩阵 $\mathbf{Q}_k = \mathbf{w}_k \mathbf{w}_k^H$, $k = 1, 2$, 从而问题式(23)可等效转换为式(24)问题

$$\begin{aligned} &\min_{\{\mathbf{Q}_i \succeq 0\}} \text{Tr}(\mathbf{Q}_1 + \mathbf{Q}_2) \\ &\text{s.t. } \mathbf{h}_1^H \mathbf{Q}_1 \mathbf{h}_1 \geq \sigma^2 \gamma_1^{\text{th}} \\ &\quad \frac{\mathbf{h}_1^H \mathbf{Q}_2 \mathbf{h}_1}{\gamma_2^{\text{th}}} \geq \mathbf{h}_1^H \mathbf{Q}_1 \mathbf{h}_1 + \sigma^2 \\ &\quad \frac{\mathbf{h}_2^H \mathbf{Q}_2 \mathbf{h}_2}{\gamma_2^{\text{th}}} \geq \mathbf{h}_2^H \mathbf{Q}_1 \mathbf{h}_2 + \sigma^2 \\ &\quad \text{Rank}(\mathbf{Q}_k) = 1, k = 1, 2 \end{aligned} \quad (24)$$

此时问题式(24)的非凸性来自矩阵秩的约束。因此将采用半正定松弛(SemiDefinite Relaxation, SDR)的方法对问题(24)进行转化, 即

$$\begin{aligned} &\min_{\{\mathbf{Q}_i \succeq 0\}} \text{Tr}(\mathbf{Q}_1 + \mathbf{Q}_2) \\ &\text{s.t. } \mathbf{h}_1^H \mathbf{Q}_1 \mathbf{h}_1 \geq \sigma^2 \gamma_1^{\text{th}} \\ &\quad \frac{\mathbf{h}_1^H \mathbf{Q}_2 \mathbf{h}_1}{\gamma_2^{\text{th}}} \geq \mathbf{h}_1^H \mathbf{Q}_1 \mathbf{h}_1 + \sigma^2 \\ &\quad \frac{\mathbf{h}_2^H \mathbf{Q}_2 \mathbf{h}_2}{\gamma_2^{\text{th}}} \geq \mathbf{h}_2^H \mathbf{Q}_1 \mathbf{h}_2 + \sigma^2 \end{aligned} \quad (25)$$

注意到问题式(25)是一个标准的半正定规划(SemiDefinite Programming, SDP)问题, 可由标准的凸优化工具(比如CVX)解决。不过, 在下面的命题中, 将进一步证明秩约束松弛的紧密性, 即问题式(25)的最优解总能满足问题式(24)中的秩一约束。

命题1 问题式(25)的最优解满足

$$\text{Rank}(\mathbf{Q}_1) = \text{Rank}(\mathbf{Q}_2) = 1 \quad (26)$$

证明 根据文献[24]可知, 问题式(25)的最优解满足 $\text{Rank}^2(\mathbf{Q}_1) + \text{Rank}^2(\mathbf{Q}_2) \leq 3$, 此外, 为了满足其他3个约束, 我们有 $\mathbf{Q}_1, \mathbf{Q}_2 \neq \mathbf{0}$, 即 $\text{Rank}(\mathbf{Q}_1) \geq 1$, $\text{Rank}(\mathbf{Q}_2) \geq 1$, 由此可以推出式(26)。

从而, 根据命题1, 在给定RIS端相移阵的前提下, 最优的基站端波束成形向量可通过解决问题式(25)得到。

3.3 相移优化

在本小节中, 结合前面所给出的最优基站端波束成形解, 将进一步优化RIS端的被动波束成形向量, 即透射相移阵 Θ_t 和反射相移阵 Θ_r 。在给定 $\mathbf{w}_1$ 和 $\mathbf{w}_2$ 的情况下, 根据问题式(22), 被动波束成形向量的优化问题可以写为

$$\begin{aligned} & \text{Find } \{\Theta_t, \Theta_r\} \\ & \text{s.t. } \gamma_1 \geq \gamma_1^{\text{th}} \\ & \quad \gamma_{2 \rightarrow 1} \geq \gamma_2^{\text{th}} \\ & \quad \gamma_{2 \rightarrow 2} \geq \gamma_2^{\text{th}} \\ & \quad \beta_t, \beta_r \in [0, 1], \beta_r + \beta_t = 1 \\ & \quad 0 \leq \theta_n^r, \theta_n^t < 2\pi, n = 1, 2, \dots, N \end{aligned} \quad (27)$$

为了求解问题式(27), 定义 $\mathbf{v}_i^H = (\sqrt{\beta_i}e^{j\theta_1^i}, \sqrt{\beta_i}e^{j\theta_2^i}, \dots, \sqrt{\beta_i}e^{j\theta_N^i}, 1)$, $i \in \{r, t\}$ 。因此, STAR-RIS辅助下的用户信道可以转换为

$$\mathbf{h}_1^H = \mathbf{v}_r^H \mathbf{H}_1^H, \mathbf{h}_2^H = \mathbf{v}_t^H \mathbf{H}_2^H \quad (28)$$

其中, $\mathbf{H}_k = \begin{bmatrix} \text{diag}(\mathbf{r}_k^H) \mathbf{G} \\ \mathbf{h}_{di}^H \end{bmatrix}^H$, $k = 1, 2$ 。从而可进一步得到

$$|\mathbf{h}_1^H \mathbf{w}_k|^2 = \mathbf{h}_1^H \mathbf{w}_k \mathbf{w}_k^H \mathbf{h}_1 = \mathbf{v}_r^H \mathbf{H}_1^H \mathbf{w}_k \mathbf{w}_k^H \mathbf{H}_1 \mathbf{v}_r \quad (29)$$

$$|\mathbf{h}_2^H \mathbf{w}_k|^2 = \mathbf{h}_2^H \mathbf{w}_k \mathbf{w}_k^H \mathbf{h}_2 = \mathbf{v}_t^H \mathbf{H}_2^H \mathbf{w}_k \mathbf{w}_k^H \mathbf{H}_2 \mathbf{v}_t \quad (30)$$

引入 $\mathbf{V}_i = \mathbf{v}_i \mathbf{v}_i^H$, $i \in \{r, t\}$, 问题式(27)等价地转化为

$$\begin{aligned} & \text{Find } \{\mathbf{V}_r, \mathbf{V}_t\} \\ & \text{s.t. } \frac{\text{Tr}(\mathbf{V}_r \mathbf{H}_1^H \mathbf{w}_1 \mathbf{w}_1^H \mathbf{H}_1)}{\sigma^2} \geq \gamma_1^{\text{th}} \\ & \quad \frac{\text{Tr}(\mathbf{V}_r \mathbf{H}_1^H \mathbf{w}_2 \mathbf{w}_2^H \mathbf{H}_1)}{\text{Tr}(\mathbf{V}_r \mathbf{H}_1^H \mathbf{w}_1 \mathbf{w}_1^H \mathbf{H}_1) + \sigma^2} \geq \gamma_2^{\text{th}} \\ & \quad \frac{\text{Tr}(\mathbf{V}_t \mathbf{H}_2^H \mathbf{w}_2 \mathbf{w}_2^H \mathbf{H}_2)}{\text{Tr}(\mathbf{V}_t \mathbf{H}_2^H \mathbf{w}_1 \mathbf{w}_1^H \mathbf{H}_2) + \sigma^2} \geq \gamma_2^{\text{th}} \\ & \quad \beta_t, \beta_r \in [0, 1], \beta_r + \beta_t = 1 \\ & \quad \text{Diag}(\mathbf{V}_i) = \beta_i, \forall i \in \{r, t\} \\ & \quad \text{Rank}(\mathbf{V}_i) = 1, \forall i \in \{r, t\} \end{aligned} \quad (31)$$

其中, $\beta_i = [\beta_i, \beta_i, \dots, \beta_i, 1]$, $i \in \{r, t\}$ 。同样地, 通过使用SDR, 问题式(31)可以转化为式(32)形式

$$\begin{aligned} & \text{Find } \{\mathbf{V}_r, \mathbf{V}_t\} \\ & \text{s.t. } \text{Tr}(\mathbf{V}_r \mathbf{H}_1 \mathbf{w}_1 \mathbf{w}_1^H \mathbf{H}_1^H) \geq \gamma_1^{\text{th}} \sigma^2 \\ & \quad \text{Tr}(\mathbf{V}_r \mathbf{H}_1 \mathbf{w}_2 \mathbf{w}_2^H \mathbf{H}_1^H) \geq \gamma_2^{\text{th}} \text{Tr}(\mathbf{V}_r \mathbf{H}_1 \mathbf{w}_1 \mathbf{w}_1^H \mathbf{H}_1^H) \\ & \quad \quad + r_2^{\text{th}} \sigma^2 \\ & \quad \text{Tr}(\mathbf{V}_t \mathbf{H}_2 \mathbf{w}_2 \mathbf{w}_2^H \mathbf{H}_2^H) \geq \gamma_2^{\text{th}} \text{Tr}(\mathbf{V}_t \mathbf{H}_2 \mathbf{w}_1 \mathbf{w}_1^H \mathbf{H}_2^H) \\ & \quad \quad + r_2^{\text{th}} \sigma^2 \\ & \quad \beta_t, \beta_r \in [0, 1], \beta_r + \beta_t = 1 \\ & \quad \text{Diag}(\mathbf{V}_i) = \beta_i, \forall i \in \{r, t\} \end{aligned} \quad (32)$$

此时, 问题式(32)为标准的凸优化问题, 可通过标准的凸优化工具来求解。

3.4 联合优化算法

基于3.2节和3.3节两个部分所获得的基站端主动波束成形以及反射端被动波束成形设计的结果, 本节进一步给出了联合优化的过程, 并在算法1中给出。在算法1中, 首先通过求解问题式(25)得到基站端主动波束成形向量。然后, 通过求解问题式(32)进一步优化反射端被动波束成形向量。

4 仿真结果

在本节中, 对在STAR-RIS辅助URLLC-NOMA系统中所提出的功率最小化方法的性能进行了评估。在仿真中, BS位于小区的中心, 而用户则随机分散在半径为200 m的圆形区域内。当 $k = 1, 2$ 时, 用户 k 与RIS之间的信道为 $\mathbf{r}_k = d_{rk}^{-\alpha_{rk}} \mathbf{g}_{rk}$, 用户 k 与BS之间的信道为 $\mathbf{h}_{dk} = d_k^{-\alpha_k} \mathbf{g}_k$, 其中 d_{rk} 和 d_k 分别是用户 k 和RIS之间的距离以及用户 k 和BS之间的距离, $\mathbf{g}_{rk}$ 和 $\mathbf{g}_k$ 为瑞利分布系数, α_{rk} 和 α_k 为路径损耗指数。同样, RIS和BS之间的信道以 $\mathbf{G} = d_r^{-\alpha_r} \mathbf{G}_r$ 形式给出。由于反射链路或传输链路的“双衰落”效应[25], 路径损耗指数分别设为 $\alpha_{rk} = \alpha_r = 2.7$, $\alpha_k = 4$, $k = 1, 2$ 。噪声功率设置为-80 dBm。

图2评估了在用户不同的目标数据速率下天线

算法1 主被动波束成形联合优化设计算法

1. 初始化: 迭代次数 $\text{iter}=0$, $\Theta_t^{(0)}, \Theta_r^{(0)}$, 临界值 $\eta > 0$;
2. 重复步骤3-步骤4:
3. 给定 $\Theta_t^{(\text{iter})}, \Theta_r^{(\text{iter})}$, 通过求解问题式得到 $\mathbf{w}_i^*$, $i = 1, 2$, 令 $\mathbf{w}_i^{(\text{iter})} = \mathbf{w}_i^*$, $i = 1, 2$;
4. 给定 $\mathbf{w}_i^{(\text{iter})}$, $i = 1, 2$, 通过求解问题式得到 Θ_t^* 和 Θ_r^* , 令 $\Theta_t^{(\text{iter})} = \Theta_t^*$, $\Theta_r^{(\text{iter})} = \Theta_r^*$;
5. 直到系统的最小功率收敛到精度 η 或达到指定的迭代次数。

数量与发射功率的关系。由图2可见, 天线数量越多, 所需传输功率越少。这是因为随着天线数量的增加, 传输效率也会提升。同时, 当目标数据速率 r_k^{th} ($k=1,2$), 变高时, 所需传输功率明显变大。此外, 图2将本文所提出的方法与传统的STAR-RIS辅助OMA方案进行了对比, 可以观察到, 特别是在高目标数据速率要求的场景下, 本文提出的STAR-RIS辅助NOMA方案在能耗方面具有更好的性能。

图3讨论了在不同块长下天线数量与发射功率的关系。根据有限块长传输下的可达速率表达式, 我们可以得知, 随着块长的增加, 用户可达速率也增加, 这意味着用户将更容易满足其QoS需求。因此, 块长越大, 所需的传输功率越小。值得指出的是, 随着块长的增加, 所需传输功率降低的速率也会减小。特别是当块长度增加到一定程度时, 有限块长传输会逐渐接近于无限块长传输的情况。这是因为, 传输过程中的带宽利用率将趋于极限, 传输的效率也会逐渐接近于最优性能。此外, 图4进一步展示了不同解码错误概率 ε 下天线数量与发射功

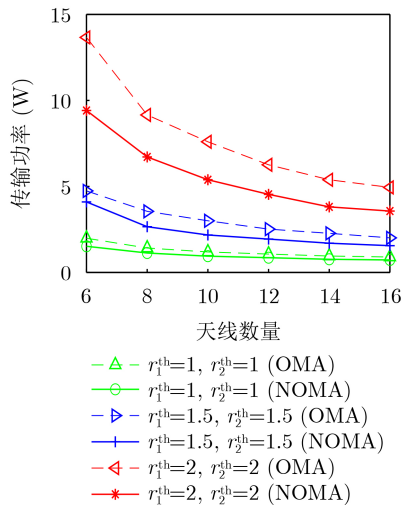


图2 在用户不同的目标数据速率下天线数量与传输功率的关系

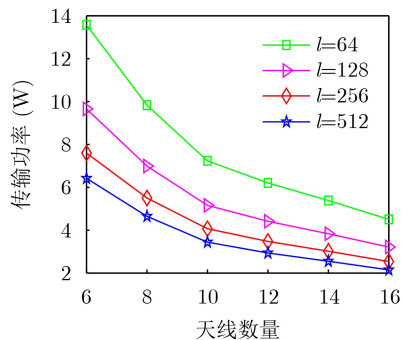


图3 在不同块长 l 下天线数量与发射功率的关系

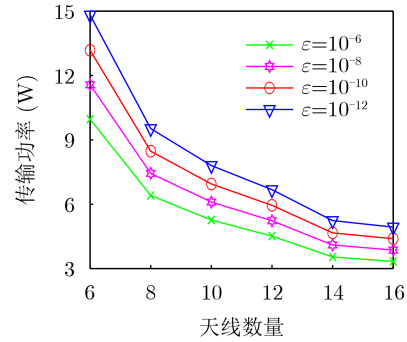


图4 在不同解码错误概率 ε 下天线数量与发射功率的关系

率的关系。正如预期的那样, 解码错误概率要求越高, 所需传输功率越大。

为了展示STAR-RIS的优势, 图5比较了 β_r 和 β_t 不同取值时天线数量与发射功率的关系。在图5中, $\beta_r=0.1, \beta_t=0.9$ 表示反射模式占主导的RIS, $\beta_r=0.1, \beta_t=0.9$ 表示传输模式占主导的RIS, 优化后的 β_r 和 β_t 表示STAR-RIS。很容易发现STAR-RIS优于单模RIS。另外, $\beta_r=0.1, \beta_t=0.9$ 时的最小传输功率高于 $\beta_r=0.1, \beta_t=0.9$ 时的最小传输功率。这是由于位于反射区的用户, 即用户1, 比位于发射区的用户, 即用户2, 更靠近BS。

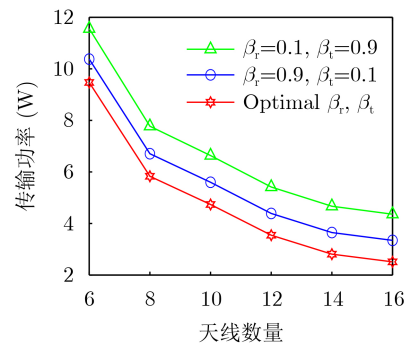


图5 β_r 和 β_t 不同取值时天线数量与发射功率的关系

5 结束语

本文深入研究了STAR-RIS在NOMA系统中的应用, 特别是针对URLLC场景下的有限块长传输问题。通过联合优化基站波束成形和RIS矩阵, 该文提出了一种高效的联合波束设计算法, 成功实现了传输功率的最小化。实验结果表明, 该方法显著提升了系统的能耗性能, 证明了STAR-RIS在多种模式下的有效性和优势。后续的工作, 将更进一步地探讨如何将本文的结果应用于多个用户的情况。

参考文献

- [1] DANG Shuping, AMIN O, SHIHADA B, et al. What should 6G be?[J]. *Nature Electronics*, 2020, 3(1): 20-29. doi: 10.1038/s41928-019-0355-6.

- [2] CHOWDHURY M Z, SHAHJALAL M, AHMED S, *et al.* 6G wireless communication systems: Applications, requirements, technologies, challenges, and research directions[J]. *IEEE Open Journal of the Communications Society*, 2020, 1: 957–975. doi: [10.1109/OJCOMS.2020.3010270](https://doi.org/10.1109/OJCOMS.2020.3010270).
- [3] CHENG Qiang, JIN Shi, CUI Tiejun. Reconfigurable intelligent surfaces for wireless communications[J]. *Frontiers of Information Technology Electronic Engineering*, 2023, 24(12): 1665–1668. doi: [10.1631/FITEE.2320000](https://doi.org/10.1631/FITEE.2320000).
- [4] ZHAO Yajun. Reconfigurable intelligent surfaces for 6G: Applications, challenges, and solutions[J]. *Frontiers of Information Technology & Electronic Engineering*, 2023, 24(12): 1669–1688. doi: [10.1631/FITEE.2200666](https://doi.org/10.1631/FITEE.2200666).
- [5] DING Zhiguo, LV Lu, FANG Fang, *et al.* A state-of-the-art survey on reconfigurable intelligent surface-assisted non-orthogonal multiple access networks[J]. *Proceedings of the IEEE*, 2022, 110(9): 1358–1379. doi: [10.1109/JPROC.2022.3174140](https://doi.org/10.1109/JPROC.2022.3174140).
- [6] HOU Tianwei, LIU Yuanwei, SONG Zhengyu, *et al.* Reconfigurable intelligent surface aided NOMA networks[J]. *IEEE Journal on Selected Areas in Communications*, 2020, 38(11): 2575–2588. doi: [10.1109/JSAC.2020.3007039](https://doi.org/10.1109/JSAC.2020.3007039).
- [7] YANG Gang, XU Xinyue, LIANG Yingchang, *et al.* Reconfigurable intelligent surface-assisted non-orthogonal multiple access[J]. *IEEE Transactions on Wireless Communications*, 2021, 20(5): 3137–3151. doi: [10.1109/TWC.2020.3047632](https://doi.org/10.1109/TWC.2020.3047632).
- [8] ZHU Jianyue, HUANG Yongming, WANG Jiaheng, *et al.* Power efficient IRS-assisted NOMA[J]. *IEEE Transactions on Communications*, 2021, 69(2): 900–913. doi: [10.1109/TCOMM.2020.3029617](https://doi.org/10.1109/TCOMM.2020.3029617).
- [9] WU Qingqing, ZHANG Shuowen, ZHENG Beixiong, *et al.* Intelligent reflecting surface-aided wireless communications: A tutorial[J]. *IEEE transactions on Communications*, 2021, 69(5): 3313–3351. doi: [10.1109/TCOMM.2021.3051897](https://doi.org/10.1109/TCOMM.2021.3051897).
- [10] LIU Yuanwei, MU Xidong, XU Jiaqi, *et al.* STAR: Simultaneous transmission and reflection for 360° coverage by intelligent surfaces[J]. *IEEE Wireless Communications*, 2021, 28(6): 102–109. doi: [10.1109/MWC.001.2100191](https://doi.org/10.1109/MWC.001.2100191).
- [11] MU Xidong, LIU Yuanwei, GUO Li, *et al.* Simultaneously transmitting and reflecting (STAR) RIS aided wireless communications[J]. *IEEE Transactions on Wireless Communications*, 2022, 21(5): 3083–3098. doi: [10.1109/TWC.2021.3118225](https://doi.org/10.1109/TWC.2021.3118225).
- [12] LIU Yuanwei, XU Jiaqi, WANG Zhaolin, *et al.* Simultaneously transmitting and reflecting (STAR) RISs for 6G: Fundamentals, recent advances, and future directions[J]. *Frontiers of Information Technology & Electronic Engineering*, 2023, 24(12): 1689–1707. doi: [10.1631/FITEE.2300490](https://doi.org/10.1631/FITEE.2300490).
- [13] ZUO Jiakuo, LIU Yuanwei, DING Zhiguo, *et al.* Joint design for simultaneously transmitting and reflecting (STAR) RIS assisted NOMA systems[J]. *IEEE Transactions on Wireless Communications*, 2023, 22(1): 611–626. doi: [10.1109/TWC.2022.3197079](https://doi.org/10.1109/TWC.2022.3197079).
- [14] WU Chenyu, LIU Yuanwei, MU Xidong, *et al.* Coverage characterization of STAR-RIS networks: NOMA and OMA[J]. *IEEE Communications Letters*, 2021, 25(9): 3036–3040. doi: [10.1109/LCOMM.2021.3091807](https://doi.org/10.1109/LCOMM.2021.3091807).
- [15] GUO Yi, FANG Fang, CAI Donghong, *et al.* Energy-efficient design for a NOMA assisted STAR-RIS network with deep reinforcement learning[J]. *IEEE Transactions on Vehicular Technology*, 2023, 72(4): 5424–5428. doi: [10.1109/TVT.2022.3224926](https://doi.org/10.1109/TVT.2022.3224926).
- [16] GAO Qiling, LIU Yuanwei, MU Xidong, *et al.* Joint location and beamforming design for STAR-RIS assisted NOMA systems[J]. *IEEE Transactions on Communications*, 2023, 71(4): 2532–2546. doi: [10.1109/TCOMM.2023.3247753](https://doi.org/10.1109/TCOMM.2023.3247753).
- [17] 侯天为, 李洁, 王俊, 等. 非正交多址接入的同步透射反射智能超表面信号增强算法[J]. *北京邮电大学学报*, 2024, 47(1): 7–12. doi: [10.13190/j.jbupt.2022-279](https://doi.org/10.13190/j.jbupt.2022-279).
HOU Tianwei, LI Jie, WANG Jun, *et al.* A signal enhancement algorithm for simultaneous-transmitting- and-reflecting reconfigurable intelligent surface based on non-orthogonal multiple access networks[J]. *Journal of Beijing University of Posts and Telecommunications*, 2024, 47(1): 7–12. doi: [10.13190/j.jbupt.2022-279](https://doi.org/10.13190/j.jbupt.2022-279).
- [18] YUE Chentao, MILOSLAVSKAYA V, SHIRVANIMOGHADDAM M, *et al.* Efficient decoders for short block length codes in 6G URLLC[J]. *IEEE Communications Magazine*, 2023, 61(4): 84–90. doi: [10.1109/MCOM.001.2200275](https://doi.org/10.1109/MCOM.001.2200275).
- [19] DURISI G, KOCH T, and POPOVSKI P. Toward massive, ultrareliable, and low-latency wireless communication with short packets[J]. *Proceedings of the IEEE*, 2016, 104(9): 1711–1726. doi: [10.1109/JPROC.2016.2537298](https://doi.org/10.1109/JPROC.2016.2537298).
- [20] POLYANSKIY Y, POOR H V, and VERDÚ S. Channel coding rate in the finite blocklength regime[J]. *IEEE Transactions on Information Theory*, 2010, 56(5): 2307–2359. doi: [10.1109/TIT.2010.2043769](https://doi.org/10.1109/TIT.2010.2043769).
- [21] HE Shiwen, AN Zhenyu, ZHU Jianyue, *et al.* Beamforming design for multiuser uRLLC with finite blocklength transmission[J]. *IEEE Transactions on Wireless Communications*, 2021, 20(12): 8096–8109. doi: [10.1109/TWC.2021.3090197](https://doi.org/10.1109/TWC.2021.3090197).
- [22] SUTTON G J, ZENG Jie, LIU Renping, *et al.* Enabling

- technologies for ultra-reliable and low latency communications: From PHY and MAC layer perspectives[J]. *IEEE Communications Surveys & Tutorials*, 2019, 21(3): 2488–2524. doi: [10.1109/COMST.2019.2897800](https://doi.org/10.1109/COMST.2019.2897800).
- [23] MAIGNAN A and SCOTT T C. Fleshing out the generalized Lambert W function[J]. *ACM Communications in Computer Algebra*, 2016, 50(2): 45–60. doi: [10.1145/2992274.2992275](https://doi.org/10.1145/2992274.2992275).
- [24] LUO Zhiqian, MA W K, SO A M C, *et al.* Semidefinite relaxation of quadratic optimization problems[J]. *IEEE Signal Processing Magazine*, 2010, 27(3): 20–34. doi: [10.1109/MSP.2010.936019](https://doi.org/10.1109/MSP.2010.936019).
- [25] GRIFFIN J D and DURGIN G D. Complete link budgets for backscatter-radio and RFID systems[J]. *IEEE Antennas and Propagation Magazine*, 2009, 51(2): 11–25. doi: [10.1109/MAP.2009.5162013](https://doi.org/10.1109/MAP.2009.5162013).
- 朱建月: 女, 讲师, 研究方向为多址接入技术、大规模MIMO无线传输技术、无线资源管理等.
- 吴雨桐: 女, 学士, 研究方向为多天线传输技术、多址接入技术等.
- 陈 晓: 女, 讲师, 研究方向为智能超表面、大规模MIMO系统、基于深度学习通信技术等.
- 谢亚琴: 女, 副教授, 研究方向为无线定位技术、卫星导航、通信网络规划与优化等.
- 许 尧: 男, 讲师, 博士, 研究方向为非正交多址接入、正交时频空间调制等.
- 张治中: 男, 教授, 博士, 研究方向为移动通信、通信网络与测试技术等.
- 责任编辑: 陈 倩

Joint Beamforming Design for STAR-RIS Assisted URLLC-NOMA System

ZHU Jianyue^① WU Yutong^② CHEN Xiao^③ XIE Yaqin^①
XU Yao^① ZHANG Zhizhong^①

^①(School of Electronic and Information Engineering, Nanjing University of Information Science and Technology, Nanjing 210044, China)

^②(School of Cyberspace Security, Southeast University, Nanjing 211189, China)

^③(School of Artificial Intelligence/School of Future Technology, Nanjing University of Information Science, Nanjing University of Information Science and Technology, Nanjing 210044, China)

Abstract:

Objective This paper addresses the energy efficiency challenge in Ultra-Reliable Low-Latency Communication (URLLC) systems, crucial for mission-critical applications such as industrial automation and remote surgery. The integration of Simultaneously Transmitting and Reflecting Reconfigurable Intelligent Surfaces (STAR-RIS) with Non-Orthogonal Multiple Access (NOMA) is proposed to improve spectral efficiency and coverage while meeting URLLC's stringent reliability and latency requirements. However, the joint optimization of base station beamforming, STAR-RIS transmission, and reflection matrices presents a non-trivial problem due to non-convexity and coupled variables. This work aims to minimize energy consumption under a total power constraint by jointly designing these parameters, advancing STAR-RIS-aided NOMA systems for URLLC.

Methods To address the non-convex optimization problem, the proposed methodology involves several key steps. First, the user rate function under finite blocklength transmission is analyzed, considering the specific requirements of URLLC. This analysis facilitates the reformulation of the original problem into an equivalent form more amenable to optimization. Specifically, the rate function is approximated using a Taylor series expansion, and the effect of finite blocklength on decoding error probability is incorporated into the optimization framework. Next, an alternating optimization framework is adopted to decouple the joint design problem into subproblems, each focused on optimizing either the base station beamforming, the STAR-RIS transmission matrix, or the reflection matrix. Semidefinite Relaxation (SDR) techniques are then applied to address the non-convexity of these subproblems, ensuring efficient and tractable solutions. The SDR method transforms the original non-convex constraints into convex ones by relaxing certain matrix rank constraints, which are subsequently recovered using randomization techniques. The proposed approach is validated through extensive simulations, comparing its performance with Orthogonal Multiple Access (OMA) and traditional RIS-

aided schemes. The simulation setup includes a multi-user scenario with varying channel conditions, blocklengths, and reliability requirements.

Results and Discussions The main contributions of this paper are summarized as follows: (1) Joint Optimization of Active and Passive Beamforming Vectors: To minimize system transmission power, the paper jointly optimizes the active beamforming vector at the base station and the passive beamforming vector at the reflective surface, presenting an efficient joint beamforming design algorithm (Algorithm 1). (2) Validation and Energy Efficiency Comparison: Experimental results confirm the effectiveness of the proposed joint beamforming design. A comparison of energy consumption performance for STAR-RIS under different modes is provided. Specifically, the proposed STAR-RIS-aided NOMA scheme demonstrates a significant reduction in power consumption compared to OMA and conventional RIS-aided systems (Fig. 2 and Fig. 5). The proposed joint beamforming and STAR-RIS optimization framework effectively addresses the trade-offs between energy consumption, reliability, and latency in URLLC systems.

Conclusions This paper presents a comprehensive framework for the transmission design of STAR-RIS-aided NOMA systems in URLLC scenarios. By jointly optimizing beamforming, transmission, and reflection matrices, the proposed method significantly enhances energy efficiency while meeting the stringent requirements of URLLC. The use of alternating optimization and SDR techniques effectively addresses the non-convexity of the problem, providing practical and scalable solutions. The results highlight the potential of STAR-RIS-aided NOMA systems to support next-generation wireless communication applications, laying the foundation for further research in this area. Future work will explore the integration of machine learning techniques to further enhance the performance and adaptability of the proposed framework. Additionally, the impact of hardware impairments and imperfect channel state information on system performance will be investigated to ensure robustness in real-world deployments.

Key words: Reconfigurable intelligent surface; Non-Orthogonal Multiple Access (NOMA); Beamforming design