SCIENCE CHINA
Information Sciences

• RESEARCH PAPER •

November 2023, Vol. 66 212301:1–212301:16
https://doi.org/10.1007/s11432-022-3723-5

RIS-enhanced multi-cell downlink transmission using statistical channel state information

Xiao LI[1*], Luoluo JIANG[1], Caihong LUO[1], Yu HAN[1],
Michail MATTHAIOU[2] & Shi JIN[1]

[1]*The National Communications Research Laboratory, Southeast University, Nanjing* 210096*, China;*
[2]*Centre for Wireless Innovation* (*CWI*)*, Queen's University Belfast, Belfast* BT3 9DT*, UK*

Received 23 October 2022/Revised 6 February 2023/Accepted 15 March 2023/Published online 25 September 2023

Abstract This paper considers a multi-cell downlink transmission system, where a reconfigurable intelligent surface (RIS) is deployed to support the connectivity of users in blind areas (i.e., at the cell edges). In order to reduce the instantaneous channel state information (CSI) acquisition overhead at each base station (BS), only statistical CSI is considered at both the BSs and the RIS. Particularly, each BS has only statistical CSI within its own cell. Under these assumptions, we first obtain the approximated ergodic spectral efficiency (SE) for each user. Then, we investigate the BSs beamforming and RIS phase shift joint optimization under different criteria. For the approximated ergodic weighted sum SE maximization criteria, a sub-optimal algorithm is proposed, based on complex circle manifold, to design the RIS phase shift while the statistical maximum ratio transmission beamforming is employed at each BS. To ensure fairness, we adopt the projected subgradient-based algorithm in order to optimize the approximated minimum ergodic user SE. Simulation results demonstrate that the approximated ergodic SE expression matches well with the exact expression, whilst the superiority of the weighted sum SE and fairness performance of the proposed algorithms are verified.

Keywords fairness transmission, RIS, statistical CSI, WSSE

Citation Li X, Jiang L L, Luo C H, et al. RIS-enhanced multi-cell downlink transmission using statistical channel state information. Sci China Inf Sci, 2023, 66(11): 212301, https://doi.org/10.1007/s11432-022-3723-5

1 Introduction

In the past decades, many potential technologies have been considered [1–6] to satisfy the increasingly high throughput demands of wireless communication systems, such as massive multiple-input multiple-output (MIMO), millimeter wave, and ultra-dense networks. However, the deployment of hundreds of antennas and densely distributed small cells brings up the issues of high cost, energy consumption, and complexity. In recent years, a low-cost and promising technology, known as reconfigurable intelligent surface (RIS), has been proposed to support the next generation wireless communication systems. It is composed of passive reconfigurable reflection elements [7–11], the number of which is usually high. Each element on it can be electronically controlled to adjust the phase of the incident signals independently. Moreover, large amounts of measurements in [12, 13] have shown that the power consumption of RIS is extremely low. Thus, an RIS provides us with an energy efficient way to reconfigure the transmission environment, so that the system performance can be enhanced cost-efficiently.

Thanks to the aforementioned advantages, RISs have been considered to be used in various scenarios, including power consumption saving [14], spectral efficiency (SE) enhancement [15–19], security improvement [20, 21], and to support unmanned aerial vehicle (UAV) communication [22, 23]. Specifically, for their application in single-cell downlink systems, the joint design of base station (BS) and RIS to reduce the BS power consumption was investigated in [14], while Refs. [15, 16] provided low latency deep reinforcement learning (DRL) based RIS phase shift design frameworks to optimize the SE. By appropriately

* Corresponding author (email: li_xiao@seu.edu.cn)

adjusting the phase shift, an RIS can help to suppress the signal at eavesdroppers [20, 21]. For multi-cell downlink systems, with the help of an optimization algorithm, i.e., semidefinite relaxation (SDR), majorization minimization (MM), successive convex approximation (SCA), and so on, RISs can be configured to improve the cell-edge user performance, such as the minimum users' rate [17], weighted sum rate (WSR) [18], and minimum weighted signal-to-interference-plus-noise ratio (SINR) [19]. Different from the centrally deployed BS in each cell, Ref. [24] analyzed the opportunities and challenges of wireless energy transfer with RIS-aided cell-free systems, in which access points (APs) are distributedly deployed so as to reduce the distance between users and APs.

In the above-mentioned studies, the design of RISs required instantaneous channel state information (CSI). However, for some scenarios, obtaining instantaneous CSI accurately might be very difficult, such as the scenario with high mobility. Moreover, an RIS usually has a large amount of reflecting elements on it. Hence, the overhead of acquiring instantaneous CSI might be prohibitively high. Another choice is to utilize statistical CSI [25–27], since it varies relatively slowly, and can be attained more accurately. For single-user downlink systems, Ref. [25] proposed a phase shift design method using statistical CSI, while instantaneous CSI was still required for BS beamforming. A similar two-timescale beamforming concept, that makes use of both the statistical and effective CSI, was applied to multi-user RIS-assisted systems in [26]. It was pointed out in [27] that this two-timescale beamforming concept is practical for high-speed train communication systems.

In this paper, a multi-cell downlink transmission system is investigated. To serve the users located in the blind coverage areas, an RIS is integrated into the system. We assume that only statistical CSI can be exploited, and each BS has no CSI of the users in other cells, so as to reduce the CSI acquisition overhead. Firstly, we obtain an approximating expression of the ergodic SE of each user. Then, we try to design the BSs beamforming and the RIS phase shift under two criteria, i.e., maximizing the approximated ergodic weighted sum SE (WSSE) and maximizing the approximated ergodic minimum user SE (referred to as max-min), respectively. Regarding the approximated ergodic WSSE maximization, we propose a sub-optimal algorithm, in which the RIS phase shift is obtained through optimization on a complex circle manifold (CCM). Moreover, the BSs beamforming vectors are obtained through statistical maximum ratio transmission (MRT). For the max-min criterion, which can ensure better fairness, we propose a projected subgradient (PSG) algorithm with statistical MRT. Thus, the BSs beamforming and RIS phase shift design are decoupled. In this way, no complex alternative optimization between these two elements is needed, which reduces the computational complexity. Simulations verify that the approximate expression of the ergodic SE matches well with the exact expression, whilst the superiority of the ergodic WSSE and fairness performance of these proposed algorithms are also demonstrated.

The remainder of the paper is organized as follows. We show the system model and formulate the problems in Section 2. An approximating expression of the ergodic SE is attained in Section 3. Then, the beamforming vector and transmit power are derived for each user, and an RIS elements phase shift matrix design algorithm is also proposed in Section 3 to maximize the approximated ergodic WSSE. After that, to maximize the minimal ergodic user SE so as to ensure better fairness, the PSG-based algorithm is proposed in Section 4. The simulation results are shown in Section 5 to validate the proposed algorithms, while our key conclusion is summarized in Section 6.

Notation. We use lower-case and upper-case bold-faced letters respectively to represent column vectors and matrices. The notations $\mathrm{Re}\{\cdot\}$ and $|\cdot|$ indicate taking the real part and modulus, $(\cdot)^{\mathrm{H}}$ and $\|\cdot\|$ represent the conjugate-transpose and Euclidean norm, $v \sim \mathcal{CN}(a, b)$ indicates a complex Gaussian distributed variable v with mean a and variance b, while $v \sim U(a, b)$ represents a uniformly distributed variable x between a and b. Moreover, $\mathbb{E}\{\cdot\}$ represents the expectation; $\mathrm{Conv}(\mathcal{C})$ denotes the convex hull of a set $\mathcal{C}$; $\mathrm{diag}\{x_1, \ldots, x_N\}$ is a diagonal matrix with the diagonal elements $x_1, \ldots, x_N$; $\mathbb{C}^{M\times N}$ represents the complex-valued matrix of dimension $M \times N$; $\otimes$ is the Kronecker product; and $\odot$ is the Hadamard product. Finally, $\nabla f(\boldsymbol{x})$ denotes the gradient vector of function $f(\boldsymbol{x})$.

2 System model and problem formulation

In this paper, a multi-cell downlink transmission system enhanced by an RIS is considered. As illustrated in Figure 1, there are K cells in the system. The RIS, which is a uniform planar array (UPA), is deployed at the cell-edge among them. There are N_h columns and N_v rows, and, thus, a total number of $N = N_h \times N_v$ reflecting elements on the RIS. The antenna array at each BS has M antenna elements

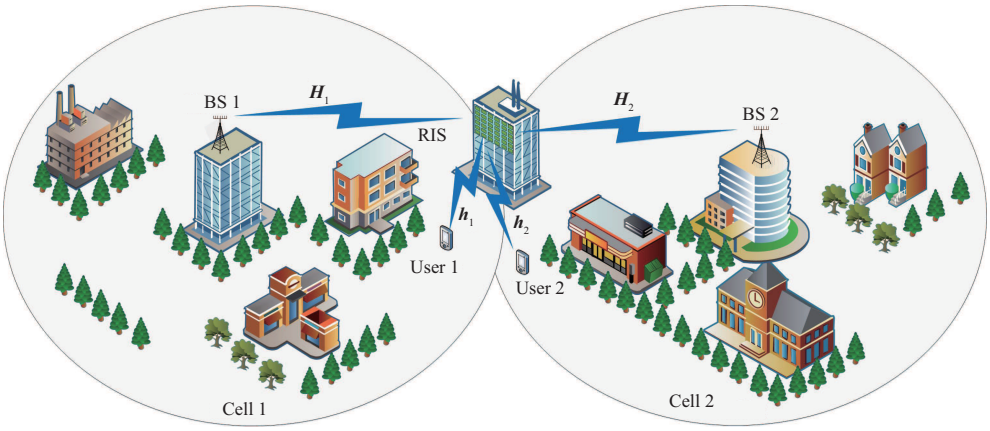

Figure 1 (Color online) Multi-cell downlink transmission system enhanced by an RIS.

arranged as a uniform linear array (ULA). There exists a single-antenna user in each cell [28–30][1)], whose direct links with the BSs are blocked or very weak and, hence, it has been served through the RIS. We assume that the user in each cell is served by its own BS, and each BS can only get its own user's statistical CSI. Here, we assume that the RIS is electronically configured by a controller. The controller can get access to the users' statistical CSI in its covered region by information exchange with the BSs. With this information, the controller calculates the phase shift parameters of the RIS. The RIS obtains the phase shift parameters from the controller through either a wired or wireless link, and then adjusts its elements correspondingly.

For this system, the signal received at cell k's user (henceforth denoted as user k for the simplicity of representation) can be formulated as

$$y_k = \sqrt{p_k \rho_k \varsigma_k} \boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{\Phi} \boldsymbol{H}_k \boldsymbol{w}_k s_k + \sum_{j=1, j \neq k}^{K} \sqrt{p_j \rho_j \varsigma_k} \boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{\Phi} \boldsymbol{H}_j \boldsymbol{w}_j s_j + n_k, \tag{1}$$

where $p_k \leqslant P_{k,\max}$ is the transmit power of BS k with $P_{k,\max}$ being the maximum value, ρ_k and $\boldsymbol{H}_k \in \mathbb{C}^{N \times M}$ are the large-scale fading coefficient and small-scale fading channel matrix between BS k and the RIS, ς_k and $\boldsymbol{h}_k \in \mathbb{C}^{N \times 1}$ are the large-scale fading coefficient and small-scale fading vector between the RIS and user k, s_k represents the transmit symbol for user k, $\mathbb{E}\{|s_k|^2\} = 1$, $\boldsymbol{\Phi} = \mathrm{diag}\{\phi_1, \phi_2, \ldots, \phi_N\}$ is the RIS phase shift matrix with the diagonal elements $\phi_n = \mathrm{e}^{\mathrm{j}\theta_n}$ [25], the beamforming vector of BS k to user k is represented by $\boldsymbol{w}_k \in \mathbb{C}^{M \times 1}$ and it satisfies $\|\boldsymbol{w}_k\|^2 = 1$, while $n_k \sim \mathcal{CN}(0, \sigma_k^2)$ is the noise, respectively.

We consider uncorrelated Rician fading [25,31], i.e.,

$$\boldsymbol{H}_k = \sqrt{\frac{\alpha_k}{\alpha_k + 1}} \bar{\boldsymbol{H}}_k + \sqrt{\frac{1}{\alpha_k + 1}} \tilde{\boldsymbol{H}}_k, \tag{2}$$

$$\boldsymbol{h}_k = \sqrt{\frac{\beta_k}{\beta_k + 1}} \bar{\boldsymbol{h}}_k + \sqrt{\frac{1}{\beta_k + 1}} \tilde{\boldsymbol{h}}_k, \tag{3}$$

where α_k and β_k are the Rician factors, $\tilde{\boldsymbol{h}}_k$ and $\tilde{\boldsymbol{H}}_k$ are respectively an uncorrelated complex Gaussian distributed vector and matrix with zero mean and unit variance elements, $\bar{\boldsymbol{H}}_k$ and $\bar{\boldsymbol{h}}_k$ are the line-of-sight (LoS) parts expressed as

$$\bar{\boldsymbol{H}}_k = \boldsymbol{a}_N\left(\varphi_k^{\mathrm{AoA}}, \theta_k^{\mathrm{AoA}}\right) \boldsymbol{b}_M^{\mathrm{H}}\left(\varpi_k\right), \tag{4}$$

$$\bar{\boldsymbol{h}}_k = \boldsymbol{a}_N\left(\varphi_k^{\mathrm{AoD}}, \theta_k^{\mathrm{AoD}}\right), \tag{5}$$

where

$$\boldsymbol{a}_N(\psi, \vartheta) = \left[1, \ldots, \mathrm{e}^{\mathrm{j}2\pi(N_v - 1)\frac{d_R}{\lambda_c}\sin(\psi)}\right]^{\mathrm{H}} \otimes \left[1, \ldots, \mathrm{e}^{\mathrm{j}2\pi(N_h - 1)\frac{d_R}{\lambda_c}\cos(\psi)\sin(\vartheta)}\right]^{\mathrm{H}}, \tag{6}$$

1) In practice, there could be multiple users in a cell to be served by the RIS. In this case, we can select one user in each cell according to certain scheduling algorithms [28–30], or just serve them in turn in different time slots.

$$\boldsymbol{b}_M(\phi) = \left[1, \mathrm{e}^{\mathrm{j}2\pi\frac{d_B}{\lambda_c}\sin(\phi)}, \ldots, \mathrm{e}^{\mathrm{j}2\pi\frac{d_B}{\lambda_c}(M-1)\sin(\phi)}\right]^{\mathrm{H}}, \tag{7}$$

where d_R is the adjacent elements' spacing on the RIS, λ_c is the carrier wavelength, d_B is the adjacent antennas spacing at each BS, ϖ_k is the azimuth angle-of-departure (AoD) of BS k relative to the RIS, θ_k^{AoA} and φ_k^{AoA} denote the azimuth and elevation angle-of-arrival (AoA) at the RIS relative to BS k, θ_k^{AoD} and φ_k^{AoD} are the azimuth and elevation AoD of the RIS relative to user k, respectively.

The SINR of user k can be formulated as

$$\mathrm{SINR}_k = \frac{p_k\rho_k\varsigma_k\left|\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{\Phi}\boldsymbol{H}_k\boldsymbol{w}_k\right|^2}{\sum_{j=1,j\neq k}^{K}p_j\rho_j\varsigma_k\left|\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{\Phi}\boldsymbol{H}_j\boldsymbol{w}_j\right|^2+\sigma_k^2}, \tag{8}$$

and thus, the user's ergodic SE is obtained as

$$R_k = \mathbb{E}\{\log(1+\mathrm{SINR}_k)\}. \tag{9}$$

Since both total SE and minimum user SE are important metrics to evaluate the system performance, we henceforth formulate two optimization problems under these two different metrics. For the metric of total system throughput maximization, with γ_k denoting the weight of user k, the following WSSE maximization problem is formulated:

$$\begin{aligned}(\mathrm{P1})\ \max_{\boldsymbol{\Phi},\{\boldsymbol{w}_k\},\{p_k\}}\ & \sum_{k=1}^{K}\gamma_k R_k,\\ \text{s.t.}\quad & p_k\leqslant P_{k,\max},\ k=\{1,\ldots,K\},\\ & |\phi_n|=1,\ n\in\{1,\ldots,N\},\\ & \|\boldsymbol{w}_k\|^2=1,\ k\in\{1,\ldots,K\}.\end{aligned} \tag{10}$$

From a different aspect, to guarantee the QoS of each user and fairness, we also consider the maximization of the minimum user SE, which is formulated as the following max-min problem:

$$\begin{aligned}(\mathrm{P2})\ \max_{\boldsymbol{\Phi},\{\boldsymbol{w}_k\},\{p_k\}}\ & \min_k\ R_k,\\ \text{s.t.}\quad & p_k\leqslant P_{k,\max},\ k\in\{1,\ldots,K\},\\ & |\phi_n|=1,\ n\in\{1,\ldots,N\},\\ & \|\boldsymbol{w}_k\|^2=1,\ k\in\{1,\ldots,K\}.\end{aligned} \tag{11}$$

Regarding this metric, if one user has a lower SE, the algorithm will try to increase this user's SE. Thus, the optimal solution will be that all users achieve almost the same SE, which leads to universal fairness.

Note that, both problems (P1) and (P2) are non-convex optimization problems. We then try to develop effective algorithms for them in the following two sections, respectively.

3 WSSE transmission design

This section focuses on solving problem (P1), i.e., the WSSE optimization problem. Note that the derivation of optimal BS beamforming and RIS phase shift in (P1) requires a closed-form expression of (9), which is very challenging to obtain. To overcome this problem, we first get a closed-form analytical approximation of (9). Then, we try to maximize this approximation other than the original expression.

From [31], we can have the following approximation of (9):

$$R_k \approx \hat{R}_k = \log\left(1+\frac{p_k\rho_k\varsigma_k\mathbb{E}\{\left|\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{\Phi}\boldsymbol{H}_k\boldsymbol{w}_k\right|^2\}}{\sum_{j=1,j\neq k}^{K}p_j\rho_j\varsigma_k\mathbb{E}\{\left|\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{\Phi}\boldsymbol{H}_j\boldsymbol{w}_j\right|^2\}+\sigma_k^2}\right) = \log\left(1+\overline{\mathrm{SINR}}_k\right), \tag{12}$$

where

$$\overline{\mathrm{SINR}}_k = \frac{p_kA_{k,k}|\bar{\boldsymbol{h}}_k^{\mathrm{H}}\boldsymbol{\Phi}\bar{\boldsymbol{H}}_k\boldsymbol{w}_k|^2+p_kB_{k,k}\|\boldsymbol{w}_k\|^2+p_kC_{k,k}\|\bar{\boldsymbol{H}}_k\boldsymbol{w}_k\|^2}{\sum_{j=1,j\neq k}^{K}p_j(A_{j,k}|\bar{\boldsymbol{h}}_k^{\mathrm{H}}\boldsymbol{\Phi}\bar{\boldsymbol{H}}_j\boldsymbol{w}_j|^2+B_{j,k}\|\boldsymbol{w}_j\|^2+C_{j,k}\|\bar{\boldsymbol{H}}_j\boldsymbol{w}_k\|^2)+\sigma_k^2}, \tag{13}$$

$$A_{j,k} \triangleq \frac{\rho_j \varsigma_k}{(\beta_k+1)(\alpha_j+1)}\beta_k\alpha_j, \tag{14}$$

$$B_{j,k} \triangleq \frac{\rho_j \varsigma_k N(\beta_k+1)}{(\beta_k+1)(\alpha_j+1)}, \tag{15}$$

$$C_{j,k} \triangleq \frac{\rho_j \varsigma_k \alpha_j}{(\beta_k+1)(\alpha_j+1)}. \tag{16}$$

With this approximation $\hat{R}_k$, we convert problem (P1) to the following one:

$$\begin{aligned} \text{(P1-1)} \max_{\{\boldsymbol{w}_k\},\boldsymbol{\Phi},\{p_k\}} \quad & \sum\nolimits_{k=1}^{K} \gamma_k \hat{R}_k, \\ \text{s.t.} \quad & p_k \leqslant P_{k,\max}, k \in \{1,\ldots,K\}, \\ & |\phi_n| = 1, n \in \{1,\ldots,N\}, \\ & \|\boldsymbol{w}_k\|^2 = 1, k \in \{1,\ldots,K\}. \end{aligned} \tag{17}$$

Next, we consider the optimization of $\{\boldsymbol{w}_k\}$, $\{p_k\}$, and $\boldsymbol{\Phi}$ to solve the problem (P1-1).

3.1 Design of beamforming vector and transmission power

First, consider the design of $\{\boldsymbol{w}_k\}$ and $\{p_k\}$. Note that we assume that each BS has no information of the users in other cells. Under this condition, for each user, the signals transmitted by other BSs are actually unknown interference. In this case, what can be done at each BS to improve the performance is to maximize its own user's useful signal, i.e., to maximize the numerator of (13). Thus, we decompose problem (P1-1) into the following K independent problems:

$$\begin{aligned} \text{(P1-2)} \max_{\boldsymbol{w}_k,p_k} \quad & \hat{z}_k, \\ \text{s.t.} \quad & \|\boldsymbol{w}_k\|^2 = 1,\ p_k \leqslant P_{k,\max}, \end{aligned} \tag{18}$$

where $k = 1,\ldots,K$ and

$$\hat{z}_k \triangleq p_k A_{k,k}\left|\bar{\boldsymbol{h}}_k^{\mathrm{H}}\boldsymbol{\Phi}\bar{\boldsymbol{H}}_k\boldsymbol{w}_k\right|^2 + p_k B_{k,k}\|\boldsymbol{w}_k\|^2 + p_k C_{k,k}\left\|\bar{\boldsymbol{H}}_k\boldsymbol{w}_k\right\|^2. \tag{19}$$

Since the above expression is increasing with p_k, we have that the optimal p_k is

$$p_k^{\mathrm{opt}} = P_{k,\max}. \tag{20}$$

Then, with fixed p_k, problem (P1-2) is converted to

$$\begin{aligned} \text{(P1-3)} \quad & \max_{\boldsymbol{w}_k} p_k A_{k,k}\left|\bar{\boldsymbol{h}}_k^{\mathrm{H}}\boldsymbol{\Phi}\bar{\boldsymbol{H}}_k\boldsymbol{w}_k\right|^2 + p_k C_{k,k}\left\|\bar{\boldsymbol{H}}_k\boldsymbol{w}_k\right\|^2, \\ & \text{s.t.} \quad \|\boldsymbol{w}_k\|^2 = 1. \end{aligned} \tag{21}$$

Applying (4)–(7), the above objective function is further written as

$$\begin{aligned} & p_k A_{k,k}\left|\bar{\boldsymbol{h}}_k^{\mathrm{H}}\boldsymbol{\Phi}\bar{\boldsymbol{H}}_k\boldsymbol{w}_k\right|^2 + p_k C_{k,k}\left\|\bar{\boldsymbol{H}}_k\boldsymbol{w}_k\right\|^2 \\ & = p_k C_{k,k}\left\|\boldsymbol{a}_N\left(\varphi_k^{\mathrm{AoA}},\theta_k^{\mathrm{AoA}}\right)\right\|^2\left|\boldsymbol{b}_M^{\mathrm{H}}\left(\varpi_k\right)\boldsymbol{w}_k\right|^2 \\ & \quad + p_k A_{k,k}\left|\boldsymbol{a}_N^{\mathrm{H}}\left(\varphi_k^{\mathrm{AoD}},\theta_k^{\mathrm{AoD}}\right)\boldsymbol{\Phi}\boldsymbol{a}_N\left(\varphi_k^{\mathrm{AoA}},\theta_k^{\mathrm{AoA}}\right)\right|^2\left|\boldsymbol{b}_M^{\mathrm{H}}\left(\varpi_k\right)\boldsymbol{w}_k\right|^2. \end{aligned} \tag{22}$$

Maximizing (22), the optimal $\boldsymbol{w}_k$ is obtained as

$$\boldsymbol{w}_k^{\mathrm{opt}} = M^{-\frac{1}{2}}\boldsymbol{b}_M\left(\varpi_k\right). \tag{23}$$

Thus, the approximation $\hat{R}_k$ is further obtained as

$$\hat{R}_k = \log\left(\frac{c_k + \boldsymbol{\eta}^{\mathrm{H}}\boldsymbol{A}_k\boldsymbol{\eta}}{\boldsymbol{\eta}^{\mathrm{H}}\boldsymbol{B}_k\boldsymbol{\eta} + d_k + \sigma_k^2} + 1\right), \tag{24}$$

where

$$\boldsymbol{\eta}^{\mathrm{H}} \triangleq [\phi_1, \phi_2, \ldots, \phi_N], \tag{25}$$

$$\boldsymbol{A}_k \triangleq \frac{p_k \rho_k \varsigma_k M \alpha_k \beta_k}{(\alpha_k+1)(\beta_k+1)} \mathrm{diag}\left\{\bar{\boldsymbol{h}}_k^{\mathrm{H}}\right\} \boldsymbol{a}_N\left(\varphi_k^{\mathrm{AoA}}, \theta_k^{\mathrm{AoA}}\right) \boldsymbol{a}_N^{\mathrm{H}}\left(\varphi_k^{\mathrm{AoA}}, \theta_k^{\mathrm{AoA}}\right) \mathrm{diag}\left\{\bar{\boldsymbol{h}}_k\right\}, \tag{26}$$

$$\boldsymbol{B}_k \triangleq \sum_{j=1, j\neq k}^{K} \left(\frac{p_j \rho_j \varsigma_k M \alpha_j \beta_k}{(\alpha_j+1)(\beta_k+1)} \mathrm{diag}\left\{\bar{\boldsymbol{h}}_k^{\mathrm{H}}\right\} \boldsymbol{a}_N\left(\varphi_j^{\mathrm{AoA}}, \theta_j^{\mathrm{AoA}}\right) \boldsymbol{a}_N^{\mathrm{H}}\left(\varphi_j^{\mathrm{AoA}}, \theta_j^{\mathrm{AoA}}\right) \mathrm{diag}\left\{\bar{\boldsymbol{h}}_k\right\} \right), \tag{27}$$

$$c_k \triangleq \frac{p_k \rho_k \varsigma_k N}{(\alpha_k+1)(\beta_k+1)} (M\alpha_k + \beta_k + 1), \tag{28}$$

$$d_k \triangleq \sum_{j=1, j\neq k}^{K} \frac{p_j \rho_j \varsigma_k N (M\alpha_j + \beta_k + 1)}{(\alpha_j+1)(\beta_k+1)}. \tag{29}$$

Then, we only need to obtain the RIS phase shift, which will be considered next.

3.2 Design of phase shift

In this subsection, two algorithms are proposed for the RIS phase shift design under the WSSE criterion.

3.2.1 *FP-CCM algorithm*

With the obtained approximated ergodic SE (24), the optimization problem (P1-1) degrades to

$$\begin{aligned} \text{(P1-4)} \quad & \max_{\eta} \ \sum_{k=1}^{K} \gamma_k \log\left(\frac{c_k + \boldsymbol{\eta}^{\mathrm{H}} \boldsymbol{A}_k \boldsymbol{\eta}}{\boldsymbol{\eta}^{\mathrm{H}} \boldsymbol{B}_k \boldsymbol{\eta} + d_k + \sigma_k^2} + 1 \right), \\ & \text{s.t.} \quad |\phi_n| = 1, \ n \in \{1, \ldots, N\}. \end{aligned} \tag{30}$$

Define the objective function of (30) as the following function:

$$f(\boldsymbol{\eta}) \triangleq \sum_{k=1}^{K} \gamma_k \log\left(\frac{c_k + \boldsymbol{\eta}^{\mathrm{H}} \boldsymbol{A}_k \boldsymbol{\eta}}{\boldsymbol{\eta}^{\mathrm{H}} \boldsymbol{B}_k \boldsymbol{\eta} + d_k + \sigma_k^2} + 1 \right). \tag{31}$$

Exploiting an auxiliary $\boldsymbol{q} \in \mathbb{C}^{K\times 1}$, whose k-th element is q_k, $f(\boldsymbol{\eta})$ is converted to the following equation:

$$f^{CF}(\boldsymbol{\eta}, \boldsymbol{q}) \triangleq \sum_{k=1}^{K} \left(\gamma_k \log(q_k+1) - \gamma_k q_k \right) + \sum_{k=1}^{K} \frac{\gamma_k (q_k+1)\left(c_k + \boldsymbol{\eta}^{\mathrm{H}} \boldsymbol{A}_k \boldsymbol{\eta}\right)}{\boldsymbol{\eta}^{\mathrm{H}} (\boldsymbol{A}_k + \boldsymbol{B}_k) \boldsymbol{\eta} + c_k + d_k + \sigma_k^2}. \tag{32}$$

It can be obtained that the optimal $\boldsymbol{q}$ maximizing (32) is

$$q_k^{\mathrm{opt}} = \frac{\boldsymbol{\eta}^{\mathrm{H}} \boldsymbol{A}_k \boldsymbol{\eta} + c_k}{\boldsymbol{\eta}^{\mathrm{H}} \boldsymbol{B}_k \boldsymbol{\eta} + d_k + \sigma_k^2}, \tag{33}$$

and the function $f^{CF}(\boldsymbol{\eta}, \boldsymbol{q})$ equals $f(\boldsymbol{\eta})$ under this case. Thus, problem (P1-4) can be transformed to

$$\begin{aligned} \text{(P1-5)} \quad & \max_{\eta, \boldsymbol{q}} \ f^{CF}(\boldsymbol{\eta}, \boldsymbol{q}), \\ & \text{s.t.} \quad |\phi_n| = 1, \ n \in \{1, \ldots, N\}. \end{aligned} \tag{34}$$

The above problem can be solved via an alternate optimization of $\boldsymbol{\eta}$ and $\boldsymbol{q}$, and Eq. (33) is the optimal $\boldsymbol{q}$ under given $\boldsymbol{\eta}$. Under a given $\boldsymbol{q}$, the optimal $\boldsymbol{\eta}$ is the solution of the following problem:

$$\begin{aligned} \text{(P1-6)} \quad & \max_{\eta} \ \sum_{k=1}^{K} \frac{\gamma_k (q_k+1)\left(c_k + \boldsymbol{\eta}^{\mathrm{H}} \boldsymbol{A}_k \boldsymbol{\eta}\right)}{c_k + d_k + \sigma_k^2 + \boldsymbol{\eta}^{\mathrm{H}} (\boldsymbol{A}_k + \boldsymbol{B}_k) \boldsymbol{\eta}}, \\ & \text{s.t.} \quad |\phi_n| = 1, \ n \in \{1, \ldots, N\}. \end{aligned} \tag{35}$$

We now define

$$\boldsymbol{F}_k \triangleq \frac{c_k + d_k + \sigma_k^2}{N}\boldsymbol{I}_N + \boldsymbol{A}_k + \boldsymbol{B}_k, \tag{36}$$

$$\boldsymbol{U}_k \triangleq \gamma_k \left(q_k + 1\right)\left(\frac{c_k}{N}\boldsymbol{I}_N + \boldsymbol{A}_k\right). \tag{37}$$

Since $\boldsymbol{U}_k$ is positive-definite Hermitian, there exists an invertible matrix $\boldsymbol{E}_k \in \mathbb{C}^{N\times N}$ that admits the following decomposition:

$$\boldsymbol{U}_k = \boldsymbol{E}_k^{\mathrm{H}}\boldsymbol{E}_k. \tag{38}$$

Then, substituting (36)–(38) into (35), problem (P1-6) becomes

$$\begin{aligned}(\text{P1-7})\ \max_{\eta}\ & \sum_{k=1}^{K}\frac{\boldsymbol{\eta}^{\mathrm{H}}\boldsymbol{E}_k^{\mathrm{H}}\boldsymbol{E}_k\boldsymbol{\eta}}{\boldsymbol{\eta}^{\mathrm{H}}\boldsymbol{F}_k\boldsymbol{\eta}},\\ \text{s.t.}\ & |\phi_n| = 1,\ n \in \{1,\ldots,N\}.\end{aligned} \tag{39}$$

Applying the quadratic transform [32], we convert the above problem to

$$\begin{aligned}(\text{P1-8})\ \min_{\eta,\{\boldsymbol{y}_k\}}\ & \sum_{k=1}^{K}\boldsymbol{y}_k^{\mathrm{H}}\boldsymbol{y}_k\boldsymbol{\eta}^{\mathrm{H}}\boldsymbol{F}_k\boldsymbol{\eta} - 2\,\mathrm{Re}\left\{\boldsymbol{y}_k^{\mathrm{H}}\boldsymbol{E}_k\boldsymbol{\eta}\right\},\\ \text{s.t.}\ & |\phi_n| = 1,\ n \in \{1,\ldots,N\}.\end{aligned} \tag{40}$$

Then, we try to solve (P1-8) via an alternate optimization of $\{\boldsymbol{y}_k\}$ and $\boldsymbol{\eta}$. Note that, in (P1-8), the optimal value of $\{\boldsymbol{y}_k\}$ for a given $\boldsymbol{\eta}$ is

$$\boldsymbol{y}_k^{\mathrm{opt}} = \boldsymbol{E}_k\boldsymbol{\eta}/\boldsymbol{\eta}^{\mathrm{H}}\boldsymbol{F}_k\boldsymbol{\eta}. \tag{41}$$

Under a given $\{\boldsymbol{y}_k\}$, we define the objective function of (P1-8) as follows:

$$\boldsymbol{J}\left(\boldsymbol{\eta}\right) \triangleq \sum_{k=1}^{K}\boldsymbol{y}_k^{\mathrm{H}}\boldsymbol{y}_k\left(\boldsymbol{\eta}^{\mathrm{H}}\boldsymbol{F}_k\boldsymbol{\eta}\right) - 2\,\mathrm{Re}\left\{\boldsymbol{y}_k^{\mathrm{H}}\boldsymbol{E}_k\boldsymbol{\eta}\right\}. \tag{42}$$

Thus, we have the following optimization problem of $\boldsymbol{\eta}$ under a given $\{\boldsymbol{y}_k\}$:

$$\begin{aligned}(\text{P1-9})\ \min_{\eta}\ & \boldsymbol{J}\left(\boldsymbol{\eta}\right),\\ \text{s.t.}\ & |\phi_n| = 1,\ n \in \{1,\ldots,N\}.\end{aligned} \tag{43}$$

It can be seen that, the searching space of problem (P1-9) is equivalent to

$$\bar{\mathcal{C}}^N \triangleq \left\{\bar{\boldsymbol{c}} \in \mathbb{C}^{N\times 1} : |\bar{c}_i| = 1, i = 1,\ldots,N\right\} \tag{44}$$

with $\bar{c}_i$ being the i-th element of the vector $\bar{\boldsymbol{c}}$, actually a complex circle manifold. Now, as a convex problem on $\bar{\mathcal{C}}^N$, problem (P1-9) can be iteratively tackled by manifold optimization. Letting $\boldsymbol{\eta}^t$ denote the result of the t-th iteration, we list the key steps of the $(t+1)$-th iteration as follows:

(1) Calculate the Euclidean gradient of $\boldsymbol{J}(\boldsymbol{\eta}^t)$ given by

$$\nabla\boldsymbol{J}(\boldsymbol{\eta}^t) = 2\boldsymbol{y}_k^{\mathrm{H}}\boldsymbol{y}_k\boldsymbol{F}_k\boldsymbol{\eta} - 2\boldsymbol{E}_k^{\mathrm{H}}\boldsymbol{y}_k. \tag{45}$$

(2) Obtain the corresponding Riemannian gradient given by [33]

$$P_{\nabla\boldsymbol{J}_{\boldsymbol{\eta}}}\left(\boldsymbol{\eta}^t\right) = \nabla\boldsymbol{J}_{\boldsymbol{\eta}}\left(\boldsymbol{\eta}^t\right) - \mathrm{Re}\left\{\nabla\boldsymbol{J}_{\boldsymbol{\eta}}(\boldsymbol{\eta}^t)^* \odot \boldsymbol{\eta}^t\right\} \odot \boldsymbol{\eta}^t. \tag{46}$$

(3) With the Armijo-Goldstein step size τ_c, obtain $\bar{\boldsymbol{\eta}}^{t+1}$ as follows:

$$\bar{\boldsymbol{\eta}}^{t+1} = \boldsymbol{\eta}^t - \tau_c P_{\nabla\boldsymbol{J}_{\boldsymbol{\eta}}}\left(\boldsymbol{\eta}^t\right). \tag{47}$$

(4) Retract $\bar{\boldsymbol{\eta}}^{t+1}$ into $\bar{\mathcal{C}}^N$, i.e.,

$$\boldsymbol{\eta}^{t+1} = \boldsymbol{u}\big(\bar{\boldsymbol{\eta}}^{t+1}\big), \tag{48}$$

where $\boldsymbol{u}(\cdot)$ represents normalizing the modulus of each element.

The above FP-CCM algorithm is detailed in Algorithm 1.

Algorithm 1 FP-CCM algorithm tackling (P1-1)

Calculate $\{p_k\}$ and $\{\boldsymbol{w}_k\}$ using (20) and (23), respectively;
Setting $r = 0$, initialize $\boldsymbol{\Phi}^0$;
repeat
  Obtain $\{q_k^{\rm opt}\}$ using (33);
  Compute $\{\boldsymbol{y}_k^{\rm opt}\}$ utilizing (41);
  Set $t = 0$, initialize $\boldsymbol{\eta}^0$ by $\boldsymbol{\Phi}^r$ and (25);
  repeat
    Obtain $\nabla \boldsymbol{J}(\boldsymbol{\eta}^t)$ by (45);
    Calculate $P_{\nabla \boldsymbol{J}_{\boldsymbol{\eta}}}\left(\boldsymbol{\eta}^t\right)$ using (46);
    Compute $\bar{\boldsymbol{\eta}}^{t+1}$ by (47);
    Obtain $\boldsymbol{\eta}^{t+1}$ by (48);
    Set $t = t + 1$;
  until The objective function of problem (P1-9) converges;
  Set $r = r + 1$;
  Let $\boldsymbol{\Phi}^r = \mathrm{diag}\{(\boldsymbol{\eta}^t)^{\rm H}\}$;
until The objective function of problem (P1-4) converges;
Output $\boldsymbol{\Phi}^{\rm opt} = \boldsymbol{\Phi}^r$.

Complexity and convergence analysis. It can be calculated that the complexity is given by $\mathcal{O}\big(N\big)$ for obtaining $\{p_k\}$ and $\{\boldsymbol{w}_k\}$, $\mathcal{O}\big(N^2\big)$ for computing $\{q_k^{\rm opt}\}$, and $\mathcal{O}\big(N^3\big)$ for calculating $\{\boldsymbol{y}_k^{\rm opt}\}$. Moreover, the complexity is $\mathcal{O}\big(N^2\big)$ for computing the Euclidean gradient $\nabla \boldsymbol{J}(\boldsymbol{\eta}^t)$, $\mathcal{O}\big(N^2\big)$ for calculating $P_{\nabla \boldsymbol{J}_{\boldsymbol{\eta}}}\left(\boldsymbol{\eta}^t\right)$, and $\mathcal{O}\big(N^3\big)$ for calculating the step size τ_c [33]. Thus, the overall complexity of the FP-CCM algorithm is $\mathcal{O}\big(N^3\big)$. Then, we consider its convergence. The convergence from problem (P1-4) to problem (P1-9) has been verified in [32]. For problem (P1-9), the gradient method will converge to a stationary point [33] while its step size satisfies the Armijo-Goldstein rule. Thus, the convergence of the FP-CCM algorithm is guaranteed, leastwise to a local optimum.

3.2.2 *CCM algorithm*

In Subsection 3.2.1, two auxiliaries, i.e., $\boldsymbol{q}$ and $\{\boldsymbol{y}_k\}$, are needed for the proposed FP-CCM algorithm. To reduce complexity, we now introduce a CCM algorithm to work out problem (P1-4) directly. Although sub-optimal, it can avoid the calculation of the auxiliaries $\boldsymbol{q}$ and $\{\boldsymbol{y}_k\}$.

Let us define $h\left(\boldsymbol{\eta}\right) = -f\left(\boldsymbol{\eta}\right)$. Problem (P1-4) is then converted to the following problem:

$$\begin{aligned} \text{(P1-10)} \quad & \min_{\eta} \ \ h\left(\boldsymbol{\eta}\right), \\ & \text{s.t.} \ \ |\phi_n| = 1, \ n \in \{1, \ldots, N\}. \end{aligned} \tag{49}$$

The gradient of $h\left(\boldsymbol{\eta}\right)$ is obtained as

$$\begin{aligned} \nabla h\left(\boldsymbol{\eta}\right) = & -\sum_{k=1}^{K} \frac{2\gamma_k \boldsymbol{A}_k \boldsymbol{\eta}}{\boldsymbol{\eta}^{\rm H} \boldsymbol{A}_k \boldsymbol{\eta} + c_k + \boldsymbol{\eta}^{\rm H} \boldsymbol{B}_k \boldsymbol{\eta} + \sigma_k^2 + d_k} \\ & + \sum_{k=1}^{K} \frac{2\gamma_k \left(\boldsymbol{\eta}^{\rm H} \boldsymbol{A}_k \boldsymbol{\eta} + c_k\right) \boldsymbol{B}_k \boldsymbol{\eta}}{\left(\boldsymbol{\eta}^{\rm H} \boldsymbol{A}_k \boldsymbol{\eta} + c_k + \boldsymbol{\eta}^{\rm H} \boldsymbol{B}_k \boldsymbol{\eta} + \sigma_k^2 + d_k\right) \left(\boldsymbol{\eta}^{\rm H} \boldsymbol{B}_k \boldsymbol{\eta} + \sigma_k^2 + d_k\right)}. \end{aligned} \tag{50}$$

Then, the CCM algorithm can be applied to (P1-10) to calculate $\boldsymbol{\Phi}$. The key steps are similar to those described in Subsection 3.2.1.

The detailed CCM algorithm is delineated in Algorithm 2.

Convergence and complexity analysis of CCM algorithm. As discussed above, the CCM algorithm will converge eventually. Compared with the FP-CCM algorithm, although the CCM algorithm still has a complexity of $\mathcal{O}\big(N^3\big)$ [33], it avoids the calculation of the auxiliaries $\boldsymbol{q}$ and $\{\boldsymbol{y}_k\}$, which reduces the algorithmic complexity.

Algorithm 2 CCM algorithm tackling (P1-1)

1: Calculate $\{p_k\}$ and $\{\boldsymbol{w}_k\}$ using (20) and (23), respectively;
2: Setting $r=0$, initialize $\boldsymbol{\eta}^0$;
3: **repeat**
4: Obtain $\nabla h\left(\boldsymbol{\eta}^r\right)$ applying (50);
5: Compute $P_{\nabla h_{\boldsymbol{\eta}}}\left(\boldsymbol{\eta}^r\right)$ via (46);
6: Obtain $\bar{\boldsymbol{\eta}}^{r+1}$ using (47);
7: Calculate $\boldsymbol{\eta}^{r+1}$ exploiting (48);
8: Set $r=r+1$;
9: **until** The objective function of problem (P1-10) converges;
10: Output $\boldsymbol{\Phi}^{\mathrm{opt}}=\mathrm{diag}\{\boldsymbol{\eta}^r\}$.

4 Fairness transmission design

Apart from the ergodic sum SE, we point out that the minimum user SE is also a vital performance metric in many scenarios [34,35], so as to ensure fairness. Next, we will consider the optimization problem (P2), which aims to maximize the minimum user SE.

Similar with the optimization problem (P1), problem (P2) is also quite tough to tackle in the absence of a closed-form objective function. In this section, we also utilize the approximation obtained in (12), and try to solve the following reformulated approximated ergodic max-min problem:

$$\begin{aligned}(\text{P2-1}) \quad & \max_{\{\boldsymbol{w}_k\},\boldsymbol{\Phi},\{p_k\}} \min_k \ \hat{R}_k, \\ & \text{s.t.} \quad p_k \leqslant P_{k,\max},\ k\in\{1,\ldots,K\}, \\ & \qquad |\phi_n|=1,\ n\in\{1,\ldots,N\}, \\ & \qquad \|\boldsymbol{w}_k\|^2=1,\ k\in\{1,\ldots,K\}.\end{aligned} \tag{51}$$

Note that each BS has only the CSI within its own cell; thus, each BS's transmit strategy, i.e., beamforming vector and transmit power, can be optimized independently. Since each cell contains one user, the WSSE maximization is equivalent to the minimum SE maximization. Thus, when $\boldsymbol{\Phi}$ is fixed, problem (P2-1) can be transformed to (P1-2), and the optimal BS beamforming vector and transmission power are therefore the same as the results in Subsection 3.1.

Substituting the obtained $\{\boldsymbol{w}_k\}$ and $\{p_k\}$ into the objective function of (P2-1), we have that

$$\begin{aligned}(\text{P2-2}) \quad & \max_{\eta} \min_k \log\left(1+\frac{\boldsymbol{\eta}^{\mathrm{H}}\boldsymbol{A}_k\boldsymbol{\eta}+c_k}{\boldsymbol{\eta}^{\mathrm{H}}\boldsymbol{B}_k\boldsymbol{\eta}+d_k+\sigma_k^2}\right), \\ & \text{s.t.} \quad |\phi_n|=1,\ n\in\{1,\ldots,N\}.\end{aligned} \tag{52}$$

Because of the non-convex unit module constraint, the optimal $\boldsymbol{\eta}$ of (P2-2) is hard to obtain. Next, we propose an SCA and a PSG algorithm to solve this problem.

4.1 SCA algorithm

Note that the approximation (24) is an increasing function of the effective SINR, i.e., $\frac{\boldsymbol{\eta}^{\mathrm{H}}\boldsymbol{A}_k\boldsymbol{\eta}+c_k}{\boldsymbol{\eta}^{\mathrm{H}}\boldsymbol{B}_k\boldsymbol{\eta}+d_k+\sigma_k^2}$. Thus, maximizing the minimum ergodic user SE approximation equals maximizing the minimum effective SINR. Then, problem (P2-2) is equivalent to

$$\begin{aligned}(\text{P2-3}) \quad & \max_{\eta} \min_k \frac{\boldsymbol{\eta}^{\mathrm{H}}\boldsymbol{A}_k\boldsymbol{\eta}+c_k}{\boldsymbol{\eta}^{\mathrm{H}}\boldsymbol{B}_k\boldsymbol{\eta}+d_k+\sigma_k^2}, \\ & \text{s.t.} \quad |\phi_n|=1,\ n\in\{1,\ldots,N\}.\end{aligned} \tag{53}$$

Exploiting the Dinkelbach's transform [36], we can decouple the numerator and denominator of (53) through the introduction of an auxiliary variable τ, and the problem (P2-3) is reformulated as

$$\begin{aligned}(\text{P2-4}) \quad & \min_{\eta,\tau} \max_k \tau\left(\boldsymbol{\eta}^{\mathrm{H}}\boldsymbol{B}_k\boldsymbol{\eta}+d_k+\sigma_k^2\right)-\boldsymbol{\eta}^{\mathrm{H}}\boldsymbol{A}_k\boldsymbol{\eta}-c_k, \\ & \text{s.t.} \quad |\phi_n|=1,\ n\in\{1,\ldots,N\}.\end{aligned} \tag{54}$$

The above problem can be resolved through an alternative optimization of $\boldsymbol{\eta}$ and τ. When $\boldsymbol{\eta}$ is fixed, the optimal τ is [36]

$$\tau^{\mathrm{opt}} = \min_k \frac{\boldsymbol{\eta}^{\mathrm{H}}\boldsymbol{A}_k\boldsymbol{\eta} + c_k}{\boldsymbol{\eta}^{\mathrm{H}}\boldsymbol{B}_k\boldsymbol{\eta} + d_k + \sigma_k^2}. \tag{55}$$

Note that, with the obtained τ, it is still difficult to solve (P2-4). To obtain $\boldsymbol{\eta}$, we define

$$\mathcal{F}_k(\tau, \boldsymbol{\eta}) \triangleq \tau\left(\boldsymbol{\eta}^{\mathrm{H}}\boldsymbol{B}_k\boldsymbol{\eta} + d_k + \sigma_k^2\right) - \boldsymbol{\eta}^{\mathrm{H}}\boldsymbol{A}_k\boldsymbol{\eta} - c_k. \tag{56}$$

Utilizing the first order Taylor expansion, we can get that $\boldsymbol{\eta}^{\mathrm{H}}\boldsymbol{A}_k\boldsymbol{\eta} \approx -\tilde{\boldsymbol{\eta}}^{\mathrm{H}}\boldsymbol{A}_k\tilde{\boldsymbol{\eta}} + 2\tilde{\boldsymbol{\eta}}^{\mathrm{H}}\boldsymbol{A}_k\boldsymbol{\eta}$, where $\tilde{\boldsymbol{\eta}}$ is a constant point in the feasible domain. Therefore, Eq. (56) can be transformed to

$$\mathcal{F}_k(\tau, \boldsymbol{\eta}) \approx \mathcal{F}_k^U(\tau, \boldsymbol{\eta}) \triangleq \tau\left(\boldsymbol{\eta}^{\mathrm{H}}\boldsymbol{B}_k\boldsymbol{\eta} + d_k + \sigma_k^2\right) + \tilde{\boldsymbol{\eta}}^{\mathrm{H}}\boldsymbol{A}_k\tilde{\boldsymbol{\eta}} - 2\tilde{\boldsymbol{\eta}}^{\mathrm{H}}\boldsymbol{A}_k\boldsymbol{\eta} - c_k. \tag{57}$$

Next, let us define the maximum value of $\{\mathcal{F}_k^U(\tau, \boldsymbol{\eta})\}$ as $z \triangleq \max_k \mathcal{F}_k^U(\tau, \boldsymbol{\eta})$. Thus, with the obtained τ, problem (P2-4) is further transformed to

$$\begin{aligned}
\text{(P2-5)}\ \min_{\eta}\ & z, && \text{(58a)}\\
\text{s.t.}\ \ & \mathcal{F}_k^U(\tau, \boldsymbol{\eta}) \leqslant z,\ k \in \{1, \ldots, K\}, && \text{(58b)}\\
& |\phi_n| = 1,\ n \in \{1, \ldots, N\}. && \text{(58c)}
\end{aligned}$$

In (P2-5), only Eq. (58c) is a non-convex constraint. To effectively solve (P2-5), we can relax (58c) to a convex constraint, which leads to

$$\begin{aligned}
\text{(P2-6)}\ \min_{\eta}\ & z,\\
\text{s.t.}\ \ & \mathcal{F}_k^U(\tau, \boldsymbol{\eta}) \leqslant z,\ k \in \{1, \ldots, K\},\\
& |\phi_n| \leqslant 1,\ n \in \{1, \ldots, N\}.
\end{aligned} \tag{59}$$

The problem (P2-6) is convex; hence, we can first obtain the optimal $\boldsymbol{\eta}$ exploiting the CVX toolbox, and then normalize $\boldsymbol{\eta}$ to satisfy the constraint $|\phi_n| = 1, n \in \{1, \ldots, N\}$. The details of this SCA algorithm are illustrated in Algorithm 3.

Algorithm 3 SCA algorithm solving (P2-1)

1: Obtain $\{p_k\}$ and $\{\boldsymbol{w}_k\}$ exploiting (20) and (23), respectively;
2: Set $r = 0$, initialize $\boldsymbol{\eta}^r$;
3: **repeat**
4: Set $\tilde{\boldsymbol{\eta}} = \boldsymbol{\eta}^r$;
5: Compute τ^r using (55);
6: Calculate $\mathcal{F}_k^U(\boldsymbol{\eta}, \tau)$ utilizing (57);
7: Obtain $\boldsymbol{\eta}^{r+1}$ via sloving (P2-6);
8: Set $r = r + 1$;
9: **until** The objective function of problem (P2-4) converges;
10: Output $\boldsymbol{\Phi}^{\mathrm{opt}} = \mathrm{diag}\{\boldsymbol{\eta}^r\}$.

Convergence and complexity analysis. Note that $\mathcal{F}_k^U(\boldsymbol{\eta}, \tau)$ is a non-increasing function in each iteration. Moreover, $\mathcal{F}_k^U(\boldsymbol{\eta}, \tau)$ is a convex function. Thus, the SCA-based algorithm will converge eventually. Then, let us consider the complexity. In the SCA-based algorithm, we use the CVX toolbox to solve (P2-5), whose complexity for updating $\boldsymbol{\eta}$ is $\mathcal{O}\big((K+N)^{1/2}N\big(N^2+K^3\big)\big)$ [37].

4.2 PSG algorithm

Note that in the SCA algorithm, we apply the CVX toolbox to solve (P2-5), whose computational complexity is $\mathcal{O}\big(N^{3.5}\big)$. To further reduce the computational complexity, we apply the PSG algorithm alternatively. Notice that, in Subsection 4.1, with the obtained τ, we change the problem (P2-4) to

$$\begin{aligned}
\text{(P2-7)}\ \min_{\eta}\ & \max_k \mathcal{F}_k^U(\tau, \boldsymbol{\eta})\\
\text{s.t.}\ \ & |\phi_n| = 1,\ n \in \{1, \ldots, N\}.
\end{aligned} \tag{60}$$

Let us define $\mathcal{G}(\boldsymbol{\eta}) \triangleq \max_k \mathcal{F}_k^U(\tau, \boldsymbol{\eta})$. The subdifferential of $\mathcal{G}(\boldsymbol{\eta})$ can be obtained as

$$\partial \mathcal{G}(\boldsymbol{\eta}) \triangleq \operatorname{Conv}\left\{\nabla \mathcal{F}_k^U(\tau, \boldsymbol{\eta}) \,|k = 1, \ldots, K\right\}, \tag{61}$$

with $\nabla \mathcal{F}_k^U(\boldsymbol{\eta}, \tau) = 2\tau \boldsymbol{B}_k \boldsymbol{\eta} - 2\boldsymbol{A}_k \tilde{\boldsymbol{\eta}}$. We assume that $\boldsymbol{g}(\boldsymbol{\eta})$ is one subgradient of (61), i.e., $\boldsymbol{g}(\boldsymbol{\eta}) \in \partial \mathcal{G}(\boldsymbol{\eta})$, and $l = \arg\max_k \mathcal{F}_k^U(\tau, \boldsymbol{\eta})$. Thus, we can have the following subgradient of $\mathcal{G}(\boldsymbol{\eta})$:

$$\boldsymbol{g}(\boldsymbol{\eta}) = 2\tau \boldsymbol{B}_l \boldsymbol{\eta} - 2\boldsymbol{A}_l \tilde{\boldsymbol{\eta}}. \tag{62}$$

Note that the PSG algorithm is generally not a strict descent method. When we obtain $\boldsymbol{\eta}^r$ of the r-th iteration, we need to search a series of points along the negative subgradient direction. Furthermore, we need to keep track of the best point among them. Then, we can choose the best point as the phase shift matrix $\boldsymbol{\eta}^{r+1}$.

In detail, with the obtained $\boldsymbol{\eta}^r$, we need to search T steps along the $-\boldsymbol{g}(\boldsymbol{\eta})$ direction in total. Then, assuming that the t-th $(0 \leqslant t < T)$ iteration point $\hat{\boldsymbol{\eta}}^t$ has been obtained, we can use this iteration formula to obtain the next iteration point, i.e.,

$$\hat{\boldsymbol{\eta}}^{t+1} = \operatorname{units}\big(\hat{\boldsymbol{\eta}}^t - \lambda_t \boldsymbol{g}\big(\hat{\boldsymbol{\eta}}^t\big)\big), \tag{63}$$

where $\{\hat{\boldsymbol{\eta}}^t \mid 0 \leqslant t < T\}$ is the T search points, $\lambda_t = \varepsilon / \|\boldsymbol{g}\big(\hat{\boldsymbol{\eta}}^t\big)\|^2$ is a constant step length, ε is a positive constant, and $\hat{\boldsymbol{\eta}}^0 = \boldsymbol{\eta}^r$. Let $\mathcal{G}(\hat{\boldsymbol{\eta}}_{\text{best}})$ denote the minimum objective function value in the T iterations; that is

$$\mathcal{G}(\hat{\boldsymbol{\eta}}_{\text{best}}) = \min\left\{\mathcal{G}\left(\hat{\boldsymbol{\eta}}^0\right), \ldots, \mathcal{G}\left(\hat{\boldsymbol{\eta}}^{\mathrm{T}}\right)\right\}. \tag{64}$$

Then, the next point $\boldsymbol{\eta}^{r+1}$ is set to $\hat{\boldsymbol{\eta}}_{\text{best}}$.

In summary, the PSG algorithm is given in Algorithm 4.

Algorithm 4 PSG algorithm solving (P2-1)

```
Compute {p_k} and {w_k} exploiting (20) and (23), respectively;
Initialize r = 0, η^r, T;
repeat
    Set t = 0, η̂^t = η^r;
    while t < T do
        Compute τ^t using (55);
        Calculate g(η̂^t) utilizing (62);
        Obtain η̂^{t+1} via (63);
        Set t = t + 1;
    end while
    Obtain η̂_best exploiting (64);
    Update η^{r+1} = η̂_best;
    Set r = r + 1;
until The objective function of problem (P2-7) converges;
Output Φ^opt = diag{η^r}.
```

Convergence and complexity analysis. Note that $\mathcal{F}_k^U(\boldsymbol{\eta}, \tau)$ is a non-increasing function after each iteration. Moreover, $\mathcal{F}_k^U(\boldsymbol{\eta}, \tau)$ is a convex function. It means that there exists a minimum value of $\mathcal{F}_k^U(\boldsymbol{\eta}, \tau)$, which can guarantee the convergence of $\mathcal{F}_k^U(\boldsymbol{\eta}, \tau)$. Meanwhile, the PSG algorithm's complexity mainly lies on the calculation of the Euclidean gradient $\boldsymbol{g}(\boldsymbol{\eta})$. Thus, the overall complexity is $\mathcal{O}\big(N^2\big)$.

5 Numerical results

This section provides the numerical results to assess the proposed algorithms' performance. Note that in this paper we consider an RIS with a planar architecture. The RIS can only reflect the signal on one side. In this case, the RIS can only serve the cells on the reflecting side. Therefore, for practical consideration, we elaborate on a downlink two-cell transmission system, i.e., $K = 2$, as illustrated in Figure 2. In this system, there is a single RIS installed at the cell-edge, such that only the two cells that are nearest to the RIS and both situated on the RIS's reflecting side are taken into account. The distance between the RIS and the BS k is d_k^{BI}. In cell k, the serving user is randomly located within a circle at

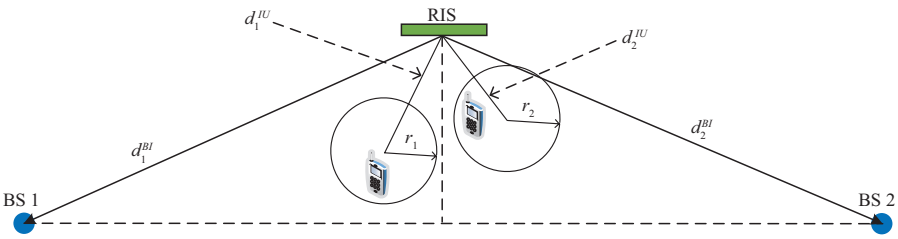

Figure 2 (Color online) The simulated transmission scenario.

the cell-edge. The radius of the circle is r_k, and the distance between its center and the RIS is d_k^{IU}. We set $d_1^{BI} = d_2^{BI} = 200$ m, $r_1 = r_2 = 5$ m, while d_1^{IU} and d_2^{IU} are uniformly distributed between 15 and 20 m. For the large scale-fading coefficients in (1), we only take into account the path loss, which is given by

$$L = L_0 - 10\hat{\alpha}\log_{10}\big(d/d_0\big), \tag{65}$$

where L is given in dB, L_0 is the path loss at the reference distance d_0, $\hat{\alpha}$ is the path loss exponent, while d is the link distance. Here, we take $d_0 = 1$ m and $L_0 = -30$ dB in all the simulations. The path loss exponent $\tilde{\alpha}$ usually ranges from 2 to 4 [14,15,38]. In this paper, we set the BS-RIS link's path loss exponent to $\tilde{\alpha}_{BI} = 2.5$ and the RIS-user link's path loss exponent to $\tilde{\alpha}_{IU} = 3$. For all the simulations, we set that $T = 100$, $M = 8$, $\alpha_k = \alpha$, $\beta_k = \beta$, $\gamma_k = 1$, $\sigma_k^2 = -90$ dBm, and $P_{k,\max} = P_{\max}$, $k \in \{1, \ldots, K\}$. The simulation results here are obtained by averaging over 500 user drops.

We compare the WSSE and the fairness performance of the four proposed algorithms, as well as the following benchmarks:

(1) **Random phase (RP).** The RIS phase shift matrix is set randomly satisfying (10c), whilst the beamforming vectors are acquired using (23).

(2) **Random beamforming (RB).** Randomly generated beamforming vectors are used at each BS, whilst the RIS phase shifts are acquired utilizing the CCM algorithm.

(3) **BestUser/WorstUser Phase.** In these two algorithms, the beamforming vectors at each BS are acquired through (23). For given $\{\boldsymbol{w}_k\}$ and $\{p_k\}$, we define the numerator of the approximated ergodic SINR of user k in (13) as

$$e_k \triangleq p_k A_{k,k}\big|\bar{\boldsymbol{h}}_k^{\mathrm{H}}\boldsymbol{\Phi}\bar{\boldsymbol{H}}_k\boldsymbol{w}_k\big|^2 + p_k B_{k,k}\|\boldsymbol{w}_k\|^2 + p_k C_{k,k}\big\|\bar{\boldsymbol{H}}_k\boldsymbol{w}_k\big\|^2. \tag{66}$$

When the RIS phase shift matrix is set to $\tilde{\boldsymbol{\Phi}}_k \triangleq \mathrm{diag}\{\big(\mathrm{diag}\{\bar{\boldsymbol{h}}_k^{\mathrm{H}}\}\,\bar{\boldsymbol{H}}_k\boldsymbol{w}_k\big)^{\mathrm{H}}\}$, e_k can yield its maximum value (defined as $\tilde{e}_k$). In the BestUser Phase scheme, the RIS phase shift matrix is set to $\boldsymbol{\Phi} = \tilde{\boldsymbol{\Phi}}_i$, $i = \arg\max_k \tilde{e}_k$, while in the WorstUser Phase scheme, it is set to $\boldsymbol{\Phi} = \tilde{\boldsymbol{\Phi}}_j$, $j = \arg\min_k \tilde{e}_k$. In other words, the BestUser Phase method tries to increase the received signal power of the user with better channel condition, while the WorstUser Phase method tries to increase the received signal power of the user with worse channel condition.

In Figures 3(a) and (b), the approximated ergodic WSSE and minimum user SE of these four proposed algorithms are compared with their Monte Carlo results simulated by (12). In Figures 3(a) and (b), $N = 100$, $\alpha = 7$ dB, and $\beta = 13$ dB. From Figures 3(a) and (b), we can observe that for all these algorithms the approximated SEs are almost the same with their corresponding Monte Carlo results. This shows that optimizing the transmit powers, phase shift matrix, and beamforming vectors using the proposed approximation are reasonable. Moreover, the FP-CCM algorithm and CCM algorithm have approximately comparable performance, while the PSG algorithm and SCA algorithm also have almost identical performance.

Figures 4(a) and (b) illustrate the ergodic WSSE and ergodic minimum user SE performance with respect to each BS's power budget $P_{\max}$ for the proposed and benchmark algorithms, respectively. The parameter settings of these figures are the same as in Figures 3(a) and (b). It can be observed in Figure 4(a) that the FP-CCM and CCM algorithms perform similarly over the entire range of transmit power, and they both significantly outperform the other algorithms. However, no alternating optimization is involved in the CCM algorithm, thus, its complexity is relatively lower. It can be observed that to ensure satisfactory performance for both users, the WSSE achieved by the SCA and PSG algorithms is inferior compared with the CCM and FP-CCM algorithms. In addition, the WSSE performance of BestUser Phase is close to that of the CCM and FP-CCM algorithms in the low transmission power

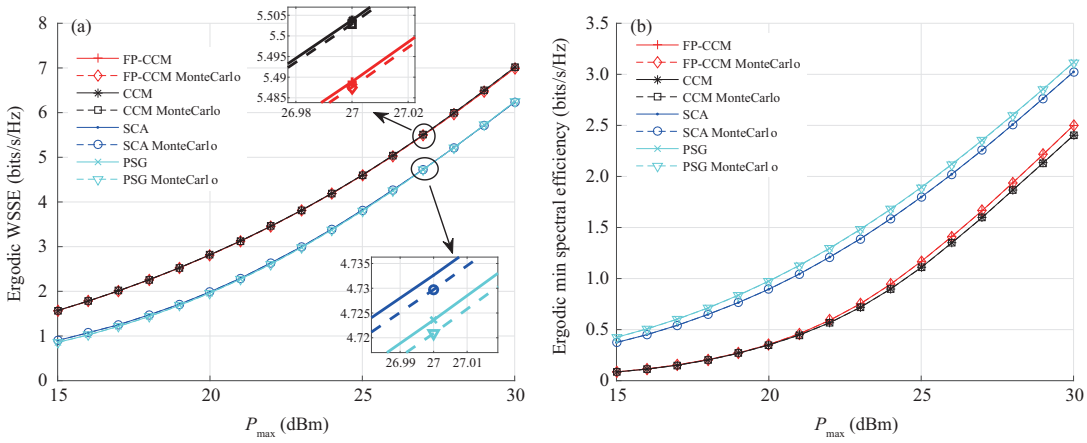

Figure 3 (Color online) (a) Approximated ergodic WSSE versus $P_{\max}$; (b) approximated ergodic minimum SE versus $P_{\max}$.

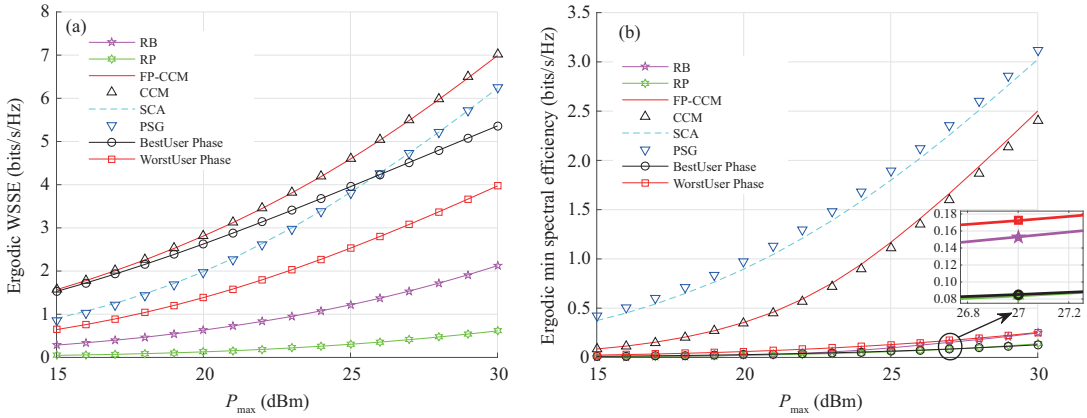

Figure 4 (Color online) (a) Ergodic WSSE versus $P_{\max}$; (b) ergodic minimum SE versus $P_{\max}$.

region, while the WSSE performance of the CCM and FP-CCM algorithms significantly outperforms that of the BestUser Phase scheme in the high transmission power region. These are consistent with the fact that the system is noise-limited in the low SNR region, thus steering the reflected signal to the user with better channel conditions will lead to high WSSE. However, the system is interference-limited in the high SNR region. Steering the signal to one user might also strengthen the interference to this user, thus, leading to low WSSE.

In Figure 4(b), we compare the ergodic minimum user SE with respect to the power budget $P_{\max}$. We can see that the ergodic minimum SE of all schemes increases with $P_{\max}$. The PSG algorithm performs slightly better than the SCA algorithm. Both of them significantly outperform the other algorithms, which validate their superiority in terms of fairness. Additionally, the PSG algorithm has lower calculation complexity compared with the SCA algorithm [39]. Moreover, we can observe that the performance of CCM and FP-CCM algorithms is degraded compared with the SCA and PSG algorithms, which reveals that the sum SE optimization metric may compromise the fairness performance.

Figures 5(a) and (b) compare the ergodic WSSE and ergodic minimum user SE of different algorithms with respect to N, i.e., the number of reflection elements, respectively. In these figures, $P_{\max} = 27$ dBm, $\alpha = 7$ dB, and $\beta = 13$ dB. From Figure 5(a), we can observe that, for all these algorithms the achieved ergodic WSSEs increase with N. Also, the ergodic WSSE achieved by the CCM algorithm and the FP-CCM algorithm is almost the same, and both outperform the other schemes significantly. Equally importantly, in Figure 5(b), the ergodic minimum user SE performance of CCM and FP-CCM algorithms is compromised compared with the SCA and PSG algorithms, which perform best in the ergodic minimum user SE.

Figures 6(a) and (b) show the ergodic WSSEs and ergodic minimum user SEs performance of different algorithms for different Rician factor β. In Figures 6(a) and (b), we set $P_{\max} = 27$ dBm, $N = 100$, and $\alpha = 7$ dB. They demonstrate that the ergodic WSSEs and ergodic minimum user SEs of these four

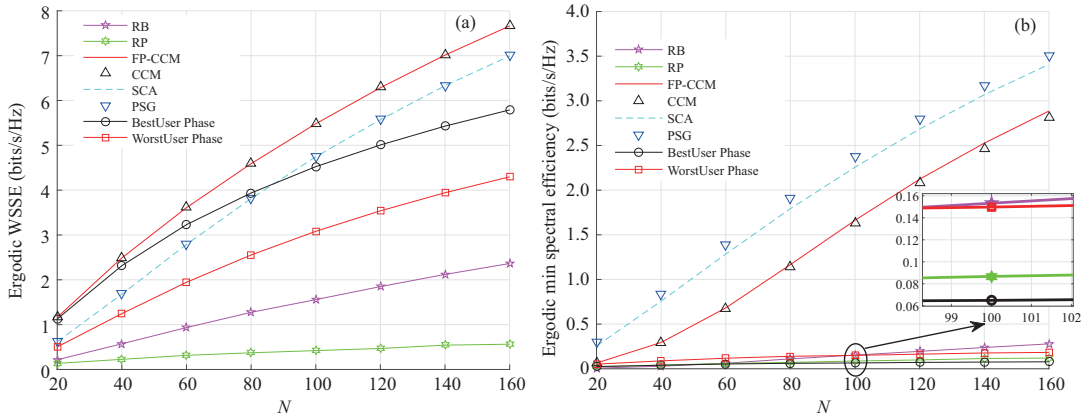

Figure 5 (Color online) (a) Ergodic WSSE versus N; (b) ergodic minimum SE versus N.

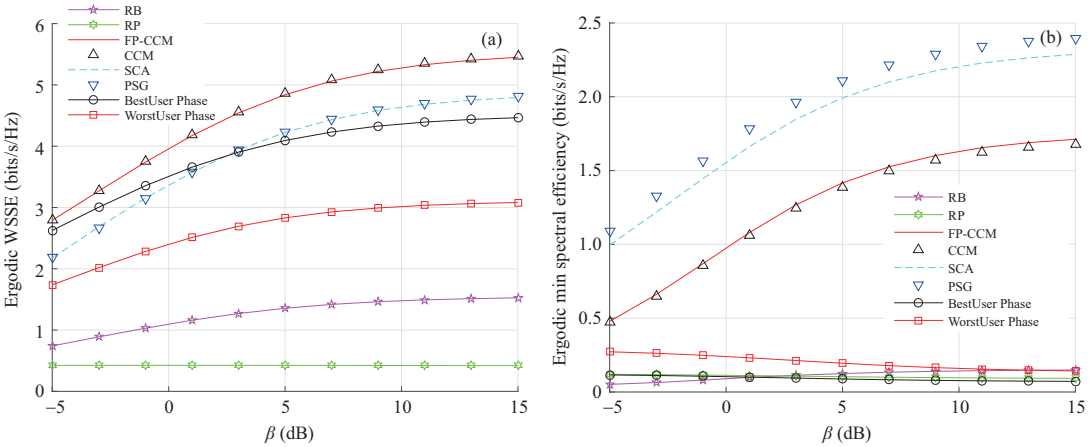

Figure 6 (Color online) (a) Ergodic WSSE versus β; (b) ergodic minimum SE versus β.

proposed algorithms grow with the Rician factor in the low Rician factor region, and they eventually reach the saturation values in the high Rician factor region. From Figure 6(b), we can see that the PSG and SCA algorithms perform best in terms of fairness, and the PSG algorithm performs slightly superior to the SCA algorithm. These are similar with the results in Figures 4(b) and 5(b). Interestingly, in Figure 6(b), we see that the ergodic minimum SEs of the "best-user" and "worst-user" algorithm is reduced with the increase of β. The reason for this is that, in these two algorithms, the RIS phase shift matrix is designed to reflect the signal towards one certain user. In this case, the other user, which is the one with the lower ergodic SE, can hardly receive any signal power through the LoS path to the RIS, since the two users are at different directions to the RIS. The majority of the signal power this second user receives is through the scattering part of the channel. Thus, the ergodic minimum SE decreases with an increasing β, since the scattering part of the channel gets weaker, which causes the decrease in the received signal power. Additionally, from Figures 4–6, we can see that the FP-CCM and CCM algorithms are superior in terms of the WSSE performance, while the PSG and SCA algorithms are superior in terms of the ergodic minimum user SE performance. This is consistent with their design objectives.

6 Conclusion

This paper investigated the downlink multi-cell system enhanced by a single RIS. The BSs could only get the statistical CSI of the user in their own cells, while the RIS could have the statistical CSI of all serving users. A closed-form approximation of the ergodic WSSE was derived. Then, we proposed a FP-CCM and a CCM algorithm to maximize the approximation of the WSSE, in order to improve the system throughput. Next, to enhance fairness, a SCA and a PSG algorithm were proposed. Through simulations, we verified that the derived ergodic SE approximation matches well with the actual SE, and

we demonstrated that the proposed algorithms work effectively.

Acknowledgements The work of X. LI was supported in part by National Natural Science Foundation of China (Grant Nos. 62231009, 61971126), National Natural Foundation of Jiangsu Province (Grant No. BK20211511), and Jiangsu Province Frontier Leading Technology Basic Research Project (Grant No. BK20212002). The work of S. JIN was supported by National Natural Science Foundation of China (Grant No. 61921004). The work of M. MATTHAIOU was supported by Research Grant from Department for the Economy Northern Ireland under US-Ireland R&D Partnership Programme.

References

1 Marzetta T L. Noncooperative cellular wireless with unlimited numbers of base station antennas. IEEE Trans Wireless Commun, 2010, 9: 3590–3600
2 Zhang J, Bjornson E, Matthaiou M, et al. Prospective multiple antenna technologies for beyond 5G. IEEE J Sel Areas Commun, 2020, 38: 1637–1660
3 Li X, Jin S, Gao X, et al. Three-dimensional beamforming for large-scale FD-MIMO systems exploiting statistical channel state information. IEEE Trans Veh Technol, 2016, 65: 8992–9005
4 Li X, Jin S, Suraweera H A, et al. Statistical 3-D beamforming for large-scale MIMO downlink systems over rician fading channels. IEEE Trans Commun, 2016, 64: 1529–1543
5 Li X, Li C, Jin S, et al. Interference coordination for 3-D beamforming-based HetNet exploiting statistical channel-state information. IEEE Trans Wireless Commun, 2018, 17: 6887–6900
6 Bjornson E, Matthaiou M, Debbah M. Massive MIMO with non-ideal arbitrary arrays: hardware scaling laws and circuit-aware design. IEEE Trans Wireless Commun, 2015, 14: 4353–4368
7 Matthaiou M, Yurduseven O, Ngo H Q, et al. The road to 6G: ten physical layer challenges for communications engineers. IEEE Commun Mag, 2021, 59: 64–69
8 Wu Q, Zhang S, Zheng B, et al. Intelligent reflecting surface-aided wireless communications: a tutorial. IEEE Trans Commun, 2021, 69: 3313–3351
9 Renzo M D, Zappone A, Debbah M, et al. Smart radio environments empowered by reconfigurable intelligent surfaces: how it works, state of research, and the road ahead. IEEE J Sel Areas Commun, 2020, 38: 2450–2525
10 Wu Q, Zhang R. Towards smart and reconfigurable environment: intelligent reflecting surface aided wireless network. IEEE Commun Mag, 2020, 58: 106–112
11 Gacanin H, Renzo M D. Wireless 2.0: toward an intelligent radio environment empowered by reconfigurable meta-surfaces and artificial intelligence. IEEE Veh Technol Mag, 2020, 15: 74–82
12 Tang W, Chen M Z, Chen X, et al. Wireless communications with reconfigurable intelligent surface: path loss modeling and experimental measurement. IEEE Trans Wireless Commun, 2020, 20: 421–439
13 Wang J, Tang W, Liang J C, et al. Reconfigurable intelligent surface: power consumption modeling and practical measurement validation. 2022. arXiv.2211.00323
14 Wu Q, Zhang R. Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming. IEEE Trans Wireless Commun, 2019, 18: 5394–5409
15 Feng K, Wang Q, Li X, et al. Deep reinforcement learning based intelligent reflecting surface optimization for MISO communication systems. IEEE Wireless Commun Lett, 2020, 9: 745–749
16 Huang C, Mo R, Yuen C. Reconfigurable intelligent surface assisted multiuser MISO systems exploiting deep reinforcement learning. IEEE J Sel Areas Commun, 2020, 38: 1839–1850
17 Hua M, Wu Q, Ng D W K, et al. Intelligent reflecting surface-aided joint processing coordinated multipoint transmission. IEEE Trans Commun, 2021, 69: 1650–1665
18 Pan C, Ren H, Wang K, et al. Multicell MIMO communications relying on intelligent reflecting surfaces. IEEE Trans Wireless Commun, 2020, 19: 5218–5233
19 Xie H, Xu J, Liu Y F. Max-Min fairness in IRS-Aided multi-cell MISO systems with joint transmit and reflective beamforming. IEEE Trans Wireless Commun, 2021, 20: 1379–1393
20 Feng K, Li X, Han Y, et al. Physical layer security enhancement exploiting intelligent reflecting surface. IEEE Commun Lett, 2021, 25: 734–738
21 Chu Z, Hao W, Xiao P, et al. Intelligent reflecting surface aided multi-antenna secure transmission. IEEE Wireless Commun Lett, 2020, 9: 108–112
22 Shafique T, Tabassum H, Hossain E. Optimization of wireless relaying with flexible UAV-borne reflecting surfaces. IEEE Trans Commun, 2021, 69: 309–325
23 Li S, Duo B, Yuan X, et al. Reconfigurable intelligent surface assisted UAV communication: joint trajectory design and passive beamforming. IEEE Wireless Commun Lett, 2020, 9: 716–720
24 Shi E, Zhang J, Chen S, et al. Wireless energy transfer in RIS-aided cell-free massive MIMO systems: opportunities and challenges. IEEE Commun Mag, 2022, 60: 26–32
25 Han Y, Tang W, Jin S, et al. Large intelligent surface-assisted wireless communication exploiting statistical CSI. IEEE Trans Veh Technol, 2019, 68: 8238–8242
26 Zhao M M, Wu Q, Zhao M J, et al. Intelligent reflecting surface enhanced wireless networks: two-timescale beamforming optimization. IEEE Trans Wireless Commun, 2021, 20: 2–17

27 Zhang J, Liu H, Wu Q, et al. RIS-aided next-generation high-speed train communications: challenges, solutions, and future directions. IEEE Wireless Commun, 2021, 28: 145–151

28 Jiang L, Li X, Matthaiou M, et al. Joint user scheduling and phase shift design for RIS assisted multi-cell MISO systems. IEEE Wireless Commun Lett, 2023, 12: 431–435

29 Li X, Sun T T, Qin N N, et al. User scheduling for downlink FD-MIMO systems under Rician fading exploiting statistical CSI. Sci China Inf Sci, 2018, 61: 082302

30 Yu X X, Guo J J, Li X, et al. Deep learning based user scheduling for massive MIMO downlink system. Sci China Inf Sci, 2021, 64: 182304

31 Zhang Q, Jin S, Wong K K, et al. Power scaling of uplink massive MIMO systems with arbitrary-rank channel means. IEEE J Sel Top Signal Process, 2014, 8: 966–981

32 Shen K, Yu W. Fractional programming for communication systems-Part I: power control and beamforming. IEEE Trans Signal Process, 2018, 66: 2616–2630

33 Absil P-A, Absil R, Sepulchre R. Optimization Algorithms on Matrix Manifolds. Princeton: Princeton University Press, 2008

34 Khaleel A, Basar E. A novel NOMA solution with ris partitioning. IEEE J Sel Top Signal Process, 2022, 16: 70–81

35 Nadeem Q U A, Kammoun A, Chaaban A, et al. Asymptotic Max-Min SINR analysis of reconfigurable intelligent surface assisted MISO systems. IEEE Trans Wireless Commun, 2020, 19: 7748–7764

36 Dinkelbach W. On nonlinear fractional programming. Manage Sci, 1967, 13: 492–498

37 Ben-Tal A, Nemirovski A. Lectures on modern convex optimization. In: Proceedings of MPS/SIAM Series on Optimization, 2001

38 Zhao M M, Wu Q, Zhao M J, et al. Exploiting amplitude control in intelligent reflecting surface aided wireless communication with imperfect CSI. IEEE Trans Commun, 2021, 69: 4216–4231

39 Boyd S, Xiao L, Mutapcic A. Subgradient Methods. Lecture notes of EE392o. Stanford: Stanford University, 2003