A Joint Precoding Framework for Wideband Reconfigurable Intelligent Surface-Aided Cell-Free Network

Zijian Zhang and Linglong Dai

Abstract—Thanks to the strong ability against the inter-cell interference, cell-free network is considered as a promising technique to improve network capacity. However, further capacity improvement requires to deploy more base stations (BSs) with high cost and power consumption. To address this issue, inspired by the recently developed reconfigurable intelligent surface (RIS) technique, we propose the concept of RIS-aided cell-free network to improve the capacity with low cost and power consumption. The key idea is to replace some of the required BSs by low-cost and energy-efficient RISs. Then, in a wideband RIS-aided cell-free network, we formulate the problem of joint precoding design at BSs and RISs to maximize the network capacity. Due to the non-convexity and high complexity of the formulated problem, we develop an alternating optimization framework to solve this challenging problem. In particular, we decouple this problem via fractional programming, and solve the subproblems alternatively. Note that most of the scenarios considered in existing works are special cases of the general scenario studied in this paper, and the proposed joint precoding framework can serve as a general solution to maximize the capacity in most existing RIS-aided scenarios. Finally, simulation results demonstrate that, compared with the conventional cell-free network, the network capacity under the proposed scheme can be improved significantly.

Index Terms—Cell-free network, reconfigurable intelligent surface (RIS), wideband, joint precoding.

I. INTRODUCTION

NETWORK technique is the most essential technique to increase the capacity of wireless communication systems [2]. Compared with 4G, the capacity of 5G wireless network is expected to be increased by 1000 times [3]. In the currently deployed cellular networks, all users in a cell are mainly served by one base station (BS), thus the users close to the cell boundary usually suffer from the severe inter-cell interference, which is caused by the signals from adjacent cells.

Ultra-dense network (UDN) has been proposed as a promising technique for 5G to further enhance the network capacity [4]. The core idea of UDN is to increase the number of BSs and deploy small cells [5] in the *cell-centric* cellular network. However, as the cell density increases, the inter-cell interference grows larger and larger. The cooperation theory has determined that [6], the upper limit of the network capacity will be bounded by the inter-cell interference as long as the cell-centric network is used. In other words, inter-cell interference becomes the bottleneck for the capacity improvement of UDN. This problem is inherent to the cell-centric network paradigm, and cannot be efficiently solved [7].

To address the issue, a novel *user-centric* network paradigm called cell-free network has been recently proposed [7]. Unlike the classical cell-centric design principle, the cell-free network utilizes the user-centric transmission design, where all BSs in the network jointly serve all users cooperatively without cell boundaries. Due to the efficient cooperation among all distributed BSs [8], the inter-cell interference can be effectively alleviated, and thus the network capacity can be increased accordingly. This promising technique has been considered as a potential candidate for future communication systems [9], and has attracted the increasing research interest such as resource allocation [10], precoding/beamforming [11], channel estimation [12] in recent years.

However, to improve the network capacity further, the deployment of more distributed BSs requires high cost and power consumption in the cell-free network. Fortunately, the emerging new technique called reconfigurable intelligent surface (RIS) is able to provide an energy-efficient alternative to enhance the network capacity. Equipped with a low-cost, energy-efficient and high-gain metasurface, RIS is becoming a promising smart radio technique for future 6G communications [13]. With a large number of low-cost passive elements, RIS is able to reflect the elertromagentic incident signals to any directions with high array gains by adjusting the phase shifts of its elements [14]. Since the wireless environment can be effectively manipulated with low cost and energy consumption [15], RIS can be used to improve channel capacity [16], reduce transmit power [17], enhance transmission reliability [18], and enlarge wireless coverage [19].

A. Prior Works

The existing research works about RIS include antenna design [20], physical model [21], channel estimation [22], joint precoding/beamforming [23], and etc. Particularly, the prototype of RIS-based wireless communication has been recently developed in [24] to demonstrate its functions.

One key guarantee for RIS to improve the network capacity is the joint precoding. Different from the conventional precoding at the BS only, the joint precoding in RIS-based wireless systems refers to the joint design of the beamforming vector at the BS and the phase shifts of the RIS elements. Different RIS-based scenarios have been studied to maximize the capacity in the literature. Specifically, the authors in [25] considered a scenario where one BS and one RIS jointly serve a single user, which was an early attempt to realize the capacity enhancement by using low-cost and energy-efficient RIS. In [26], the authors considered a multi-user scenario and maximized the sum-rate of all users. To obtain the cooperation gain, the multi-BS scenario was considered in [19], while the multi-RIS case was investigated in [18].

Apart from the capacity maximization, the design goal for joint precoding can be different. Specifically, to reduce power consumption, some researchers have proposed the methods to minimize the transmit power [17], and the authors in [15] have developed a method to maximize the energy efficiency. Some researchers have considered the fairness among users, and some methods have been proposed to maximize the minimum signal-to-interference-plus-noise ratio (SINR) [23]. Furthermore, RIS has been combined with different techniques in the literature. For instance, the RIS-aided millimeter-wave (mmWave) system was considered in [18] and [27], where the authors discussed the signal-to-noise ratio (SNR) maximization problem and the joint hybrid precoding design, respectively. In [28] and [29], the application of RIS in wideband orthogonal frequency division multiplexing (OFDM) systems was discussed to maximize the sum-rate.

B. Our Contributions

To address the challenge of cell-free network as mentioned above, in this paper we consider to exploit RISs to realize the improvement of network capacity with low cost and power consumption.[1] Specifically, the contributions of this paper can be summarized as follows.

- We propose the concept of RIS-aided cell-free network to further improve the network capacity of the cell-free network with low cost and power consumption. The key idea is to replace some of the required BSs in cell-free network by the energy-efficient RISs and deploy more RISs in the system for capacity enhancement. In the proposed RIS-aided cell-free network, all BSs and RISs are simultaneously serving all users cooperatively. To the best of our knowledge, this is the first attempt to introduce RIS in cell-free networks.

- For the proposed RIS-aided cell-free network, in a typical wideband scenario, we formulate the problem of joint precoding design at the BSs and RISs to maximize the weighted sum-rate (WSR) of all users to improve the network capacity. Since the considered scenario is very general, i.e., multiple antennas, multiple BSs, multiple RISs, multiple users, and multiple carriers, most of the considered scenarios in existing works, such as single BS, single RIS, single user, and single carrier, or some of them are mutiple, are all special cases of the considered scenario in this paper.

- We propose a joint active and passive precoding framework to solve the formulated problem. Specifically, the proposed framework is an alternating optimization algorithm based on the assumption of fully-known channel state information (CSI), which can gradually approximate a feasible solution to joint precoding design. We first decouple the active precoding at the BSs and the passive precoding at the RISs via Lagrangian dual reformulation and *Multidimensional Complex Quadratic Transform (MCQT)*, and the decoupled problem can be reformulated as two quadratically constrained quadratic program (QCQP) subproblems. Then, by solving the two subproblems alternatively, the system WSR will finally converge to a feasible solution.

- Challenged by the high-dimensional channels introduced by RISs, acquiring all RIS-aided channels constantly is usually unrealistic. To tackle this issue, by exploiting the property that RISs far from users have little contribution to capacity improvement, the proposed joint precoding framework is further extended to a more practical two-timescale scheme, which can reduce the CSI required by joint precoding from the long-term perspective. Specifically, at the beginning of a large timescale, each user is matched with several well-performed RISs by the proposed *linear conic relaxation (LCR)*-based method. Then, in later several small timescales, only the RIS-aided channels of the matched user-RIS pairs are acquired and utilized for joint precoding design, while those of the unmatched pairs are temporarily ignored. Finally, in the next large timescale, the above process will be repeated.

- Simulation results demonstrate that RISs can improve the cell-free network capacity significantly. In particular, with limited CSI knowledge, the two-timescale extension of the proposed framework can improve the capacity efficiently with little performance loss. Besides, it is worth noting that, thanks to the generality of the studied problem, the proposed joint precoding framework can also serve as a general solution to maximize the WSR in most of the existing RIS-aided scenarios in the literature.

C. Organization and Notation

Organization: The rest of the paper is organized as follows. The system model of the proposed RIS-aided cell-free network and corresponding WSR maximization problem formulation of joint precoding design are discussed in Section II. The joint precoding framework to solve the formulated problem is

---

[1]Simulation codes are provided to reproduce the results presented in this article: http://oa.ee.tsinghua.edu.cn/dailinglong/publications/publications.html.
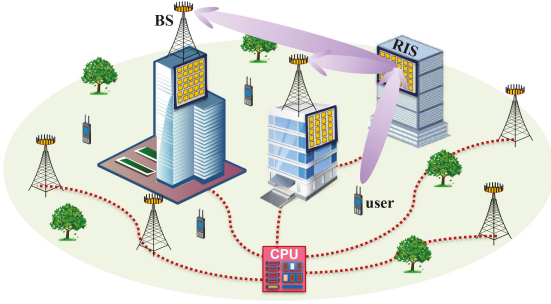
Fig. 1. The proposed concept of RIS-aided cell-free network.

proposed in Section III, and the two-timescale extension of the framework is proposed in Section IV. More framework supplements including convergence and complexity analysis are given in Section V. Simulation results are provided in Section VI to validate the performance of the proposed RIS-aided cell-free network. Finally, in Section VII, conclusions are drawn and future works are discussed.

Notations: $\mathbb{C}$, $\mathbb{R}$, and $\mathbb{R}^+$ denote the set of complex, real, and positive real numbers, respectively; $[\cdot]^{-1}$, $[\cdot]^*$, $[\cdot]^T$, and $[\cdot]^H$ denote the inverse, conjugate, transpose, and conjugate-transpose operations, respectively; $[\cdot]^+$ denotes the negative truncation operation, i.e., $[x]^+ = \max\{x, 0\}$; $[\mathbf{v}]_i$ is the ith element of vector $\mathbf{v}$; $[\mathbf{M}]_{i,j}$ is the element at the ith row and jth column of matrix $\mathbf{M}$; $\|\cdot\|$ denotes the Euclidean norm of its argument; $\mathbb{E}\{\cdot\}$ is the expectation operator; $\mathrm{diag}(\cdot)$ denotes diagonal operation; $\mathfrak{R}\{\cdot\}$ denotes the real part of its argument; $\mathrm{Tr}\{\cdot\}$ denotes the trace of its argument; $\otimes$ denotes the Kronecker product; $\angle[\cdot]$ denotes the angle of its complex argument; $\ln(\cdot)$ denotes natural logarithm; $\mathbf{I}_L$ is an $L \times L$ identity matrix, and $\mathbf{0}_L$ is an $L \times L$ zero matrix; Finally, $\mathbf{e}_l$ is an elementary vector with a one at the l-th position, and $\mathbf{1}_L$ indicates an L-length vector with all elements are 1.

II. SYSTEM MODEL OF THE PROPOSED RIS-AIDED CELL-FREE NETWORK

To improve the network capacity with low cost and power consumption, in this paper we first propose the concept of RIS-aided cell-free network. In this section, the architecture of the proposed RIS-aided cell-free network will be introduced at first. Then, we will discuss the transmitters, channels, and receivers, respectively. Finally, we will formulate the problem of capacity maximization in a wideband RIS-aided cell-free network.

A. System Architecture

In this paper, we consider a wideband RIS-aided cell-free network as shown in Fig. 1, where multiple distributed BSs and RISs are deployed to cooperatively serve all users. A central processing unit (CPU) is deployed for control and planning, to which all BSs are connected by optical cables or wireless backhaul [30]. All RISs are controlled by the CPU or BSs by wired or wireless control. Particularly, the considered network consists of B BSs, R RISs, and K multi-antenna users. The number of antennas at the b-th BS and that at the k-th user are M_b and U_k, respectively. The number of elements at the r-th RIS is N_r. For simplicity but without loss of generality, we assume M_b, U_b and N_r are equal to M, U and N, respectively.
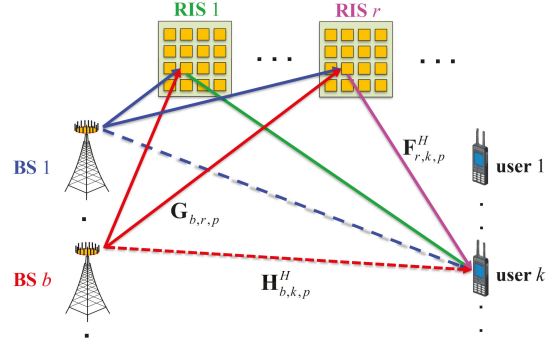

Fig. 2. The downlink channels in the wideband RIS-aided cell-free network.

Finally, the multi-carrier transmission is considered and the number of available subcarriers is P. Let $\mathcal{N} = \{1, \dots, N\}$, $\mathcal{B} = \{1, \dots, B\}$, $\mathcal{R} = \{1, \dots, R\}$, $\mathcal{K} = \{1, \dots, K\}$ and $\mathcal{P} = \{1, \dots, P\}$ denote the index sets of RIS elements, BSs, RISs, users, and subcarriers, respectively.

B. Transmitters

In the proposed RIS-aided cell-free network, all BSs are synchronized, which is necessary to serve all users by coherent joint transmission [8]. Let $\mathbf{s}_p \triangleq [s_{p,1}, \dots, s_{p,K}]^T \in \mathbb{C}^K$, where $s_{p,k}$ denotes the transmitted symbol to the k-th user on the p-th subcarrier. We assume that the transmitted symbols have normalized power, i.e., $\mathbb{E}\{\mathbf{s}_p \mathbf{s}_p^H\} = \mathbf{I}_K, \forall p \in \mathcal{P}$. In the downlink, the frequency-domain symbol $s_{p,k}$ is firstly precoded by the precoding vector $\mathbf{w}_{b,p,k} \in \mathbb{C}^M$ at the b-th BS, so the precoded symbol $\mathbf{x}_{b,p}$ at the b-th BS on the p-th subcarrier can be written as

$$\mathbf{x}_{b,p} = \sum_{k=1}^{K} \mathbf{w}_{b,p,k} s_{p,k}. \tag{1}$$

Then, by inverse discrete Fourier transform (IDFT), the frequency-domain signal $\{\mathbf{x}_{b,p}\}_{p=1}^{P}$ on all P subcarriers at the b-th BS is converted to the time domain. After adding the cyclic prefix (CP), the signal is up-converted to the radio frequency (RF) domain via M RF chains of the b-th BS.

C. Channels

Thanks to the directional reflection supported by R RISs as shown in Fig. 2, the channel between each BS and each user in the proposed RIS-aided cell-free network consists of two parts: the BS-user link and R BS-RIS-user links, where each BS-RIS-user link can be further divided into a BS-RIS link and a RIS-user link. The signal reflection on the RISs can be modeled by multiplying the incident signal with a phase shift matrix and forwarding the phase shifted signal to the user [25]. Therefore, the equivalent channel $\mathbf{h}_{b,k,p}^H \in \mathbb{C}^{U \times M}$ from the b-th BS to the k-th user on the p-th subcarrier can be written as[2] [32]

$$\mathbf{h}_{b,k,p}^H = \underbrace{\mathbf{H}_{b,k,p}^H}_{\text{BS-user link}} + \underbrace{\sum_{r=1}^{R} \mathbf{F}_{r,k,p}^H \mathbf{\Theta}_r^H \mathbf{G}_{b,r,p}}_{\text{BS-RIS-user links}}, \tag{2}$$

[2]The signals reflected by the RISs twice and more are ignored, since they are much weak due to the large path loss of multiple hops [31].

where $\mathbf{H}_{b,k,p}^{H} \in \mathbb{C}^{U \times M}$, $\mathbf{G}_{b,r,p} \in \mathbb{C}^{N \times M}$, and $\mathbf{F}_{r,k,p}^{H} \in \mathbb{C}^{U \times N}$ denote the frequency-domain channel on the subcarrier p from the BS b to the user k, from the BS b to the RIS r, and from the RIS r to the user k, respectively; $\boldsymbol{\Theta}_r \in \mathbb{C}^{N \times N}$ denotes the phase shift matrix at the RIS r, which is written as

$$\boldsymbol{\Theta}_r \triangleq \mathrm{diag}\left(\theta_{r,1}, \ldots, \theta_{r,N}\right), \forall r \in \mathcal{R}, \qquad (3)$$

where $\theta_{r,n} \in \mathcal{F}$. Note that $\mathcal{F}$ is the feasible set of the reflection coefficient (RC) at RIS. To study a generalized model, here we assume $\mathcal{F}$ is the ideal RIS case, i.e., both the amplitude and the phase of $\theta_{r,n}$ associated with the RIS element can be controlled independently and continuously [33], i.e.,

$$\mathcal{F} \triangleq \left\{ \theta_{r,n} \,\Big|\, |\theta_{r,n}| \leq 1 \right\}, \forall r \in \mathcal{R}, \forall n \in \mathcal{N}. \qquad (4)$$

Note that the more practical RIS reflection coefficients such as low-resolution discrete phase shifts will be discussed in Subsection V-A later.

D. Receivers

After passing through the equivalent channel $\mathbf{h}_{b,k,p}^{H}$ as s in (2), the signals will be received by the users. The time-domain signals received by the users are down-converted to the baseband at first. After the CP removal and the discrete Fourier transform (DFT), the frequency-domain symbols can be finally recovered. Let $\mathbf{y}_{b,k,p} \in \mathbb{C}^{U}$ denote the baseband frequency-domain signal, which reaches the user k on the subcarrier p from the BS b. Then, according to the channel model above, $\mathbf{y}_{b,k,p}$ can be expressed by combining (1) and (2) as

$$\begin{aligned} \mathbf{y}_{b,k,p} &= \mathbf{h}_{b,k,p}^{H} \mathbf{x}_{b,p} \\ &= \left(\mathbf{H}_{b,k,p}^{H} + \sum_{r=1}^{R} \mathbf{F}_{r,k,p}^{H} \boldsymbol{\Theta}_r^{H} \mathbf{G}_{b,r,p} \right) \sum_{j=1}^{K} \mathbf{w}_{b,p,j} s_{p,j}. \end{aligned}$$
$$(5)$$

Since there are B BSs serving K users simultaneously, the received signal at user k is the superposition of the signals transmitted by B BSs. Let $\mathbf{y}_{k,p} \in \mathbb{C}^{U}$ denote the received signal at the user k on the subcarrier p. Thereby, considering the additive white Gaussion noise (AWGN) at the receiver, we have the expression of $\mathbf{y}_{k,p}$ as shown in (6) at the bottom of this page, where $\mathbf{z}_{k,p} \triangleq [z_{k,p,1}^{T}, \ldots, z_{k,p,U}^{T}]^{T}$ denotes the AWGN with zero mean $\mathbf{0}_U$ and covariance $\boldsymbol{\Xi}_{k,p} = \sigma^2 \mathbf{I}_U$. Note that the first term on the right-hand side of (6) is the desired signal to user k, while the second term denotes the interference from other users. As

there are P subcarriers available in total, we denote the received signal at user k as $\{\mathbf{y}_{k,p}\}_{p=1}^{P}$.

E. Problem Fomulation

Based on the system model above, we consider to maximize the WSR of the proposed RIS-aided cell-free network subject to the transmit power constraint at BSs and RC constraint RISs in this subsection. At first, the received signal $\mathbf{y}_{b,k,p}$ in (6) can be simplified as

$$\begin{aligned} \mathbf{y}_{k,p} &\overset{(a)}{=} \sum_{b=1}^{B} \sum_{j=1}^{K} \left(\mathbf{H}_{b,k,p}^{H} + \mathbf{F}_{k,p}^{H} \boldsymbol{\Theta}^{H} \mathbf{G}_{b,p} \right) \mathbf{w}_{b,p,j} s_{p,j} + \mathbf{z}_{k,p} \\ &\overset{(b)}{=} \sum_{b=1}^{B} \sum_{j=1}^{K} \mathbf{h}_{b,k,p}^{H} \mathbf{w}_{b,p,j} s_{p,j} + \mathbf{z}_{k,p} \\ &\overset{(c)}{=} \sum_{j=1}^{K} \mathbf{h}_{k,p}^{H} \mathbf{w}_{p,j} s_{p,j} + \mathbf{z}_{k,p}, \end{aligned} \qquad (7)$$

where (a) holds by defining $\boldsymbol{\Theta} = \mathrm{diag}(\boldsymbol{\Theta}_1, \ldots, \boldsymbol{\Theta}_R)$, $\mathbf{F}_{k,p} = [\mathbf{F}_{1,k,p}^{T}, \ldots, \mathbf{F}_{R,k,p}^{T}]^{T}$, and $\mathbf{G}_{b,p} = [\mathbf{G}_{b,1,p}^{T}, \ldots, \mathbf{G}_{b,R,p}^{T}]^{T}$, (b) holds according to (2), and (c) holds by defining $\mathbf{h}_{k,p} = [\mathbf{h}_{1,k,p}^{T}, \ldots, \mathbf{h}_{B,k,p}^{T}]^{T}$ and $\mathbf{w}_{p,k} = [\mathbf{w}_{1,p,k}^{T}, \ldots, \mathbf{w}_{B,p,k}^{T}]^{T}$. Then, the SINR for the transmitted symbol $s_{p,k}$ at the user k on the subcarrier p can be easily calculated as

$$\begin{aligned} &\gamma_{k,p} \\ &= \mathbf{w}_{p,k}^{H} \mathbf{h}_{k,p} \left(\sum_{j=1,j\neq k}^{K} \mathbf{h}_{k,p}^{H} \mathbf{w}_{p,j} \left(\mathbf{h}_{k,p}^{H} \mathbf{w}_{p,j} \right)^{H} + \boldsymbol{\Xi}_{k,p} \right)^{-1} \mathbf{h}_{k,p}^{H} \mathbf{w}_{p,k}. \end{aligned}$$
$$(8)$$

Thereby, the WSR R_{sum} of all K users is given by

$$R_{\mathrm{sum}} = \sum_{k=1}^{K} \sum_{p=1}^{P} \eta_k \log_2 \left(1 + \gamma_{k,p} \right), \qquad (9)$$

where $\eta_k \in \mathbb{R}^{+}$ represents the weight of the user k and $R_{k,p}$ denotes the rate of user k on subcarrier p.

Finally, the WSR maximization optimization problem can be originally formulated as

$$\begin{aligned} \mathbf{y}_{k,p} &= \sum_{b=1}^{B} \mathbf{y}_{b,k,p} + \mathbf{z}_{k,p} = \sum_{b=1}^{B} \sum_{j=1}^{K} \left(\mathbf{H}_{b,k,p}^{H} + \sum_{r=1}^{R} \mathbf{F}_{r,k,p}^{H} \boldsymbol{\Theta}_r^{H} \mathbf{G}_{b,r,p} \right) \mathbf{w}_{b,p,j} s_{p,j} + \mathbf{z}_{k,p} \\ &= \underbrace{\sum_{b=1}^{B} \left(\mathbf{H}_{b,k,p}^{H} + \sum_{r=1}^{R} \mathbf{F}_{r,k,p}^{H} \boldsymbol{\Theta}_r^{H} \mathbf{G}_{b,r,p} \right) \mathbf{w}_{b,p,k} s_{p,k}}_{\text{Desired signal to user } k} + \underbrace{\sum_{b=1}^{B} \sum_{j=1,j\neq k}^{K} \left(\mathbf{H}_{b,k,p}^{H} + \sum_{r=1}^{R} \mathbf{F}_{r,k,p}^{H} \boldsymbol{\Theta}_r^{H} \mathbf{G}_{b,r,p} \right) \mathbf{w}_{b,p,j} s_{p,j}}_{\text{Interference from other users}} + \mathbf{z}_{k,p}. \end{aligned}$$
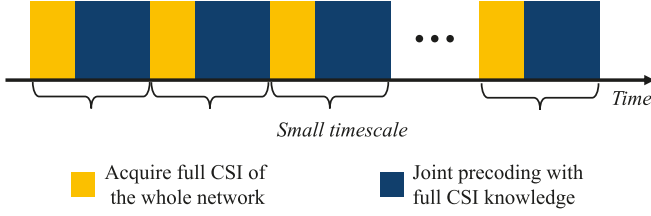$$(6)$$

Fig. 3. The dynamical working process of the proposed joint precoding framework over time.

$$\mathcal{P}^{\mathrm{o}} : \max_{\boldsymbol{\Theta}, \mathbf{W}} \ R_{\mathrm{sum}}(\boldsymbol{\Theta}, \mathbf{W}) = \sum_{k=1}^{K} \sum_{p=1}^{P} \eta_k \log_2 \left(1 + \gamma_{k,p}\right) \quad (10a)$$

$$\text{s.t.} \quad C_1 : \sum_{k=1}^{K} \sum_{p=1}^{P} \|\mathbf{w}_{b,p,k}\|^2 \le P_{b,\max}, \forall b \in \mathcal{B}, \quad (10b)$$

$$C_2 : \theta_{r,n} \in \mathcal{F}, \forall r \in \mathcal{R}, \forall n \in \mathcal{N}, \quad (10c)$$

where $P_{b,\max}$ denotes the maximum transmit power of the BS b and we have defined $\mathbf{W}$ as follows for expression simplicity:

$$\mathbf{W} = \left[\mathbf{w}_{1,1}^T, \mathbf{w}_{1,2}^T, \ldots, \mathbf{w}_{1,K}^T, \mathbf{w}_{2,1}^T, \mathbf{w}_{2,2}^T, \ldots, \mathbf{w}_{P,K}^T\right]^T. \quad (11)$$

Due to the non-convex complex objective function (10a), the joint optimization of the phase shift matrix $\boldsymbol{\Theta}$ and the precoding vector $\mathbf{W}$ is very challenging. Fortunately, inspired by the fractional programming (FP) methods, we propose a joint precoding framework to find a feasible solution to the problem $\mathcal{P}^{\mathrm{o}}$ in the following Section III.

III. PROPOSED JOINT PRECODING FRAMEWORK

In this section, we present the proposed joint precoding framework to solve the WSR optimization problem $\mathcal{P}^{\mathrm{o}}$ in (10). Specifically, the section is summarized as follows. An overview of the proposed framework is first provided in Subsection III-A, where the problem $\mathcal{P}^{\mathrm{o}}$ in (10) is divided into three subproblems. Then, the detailed algorithms to solve these three subproblems are given in Subsection III-B, III-C, and III-F, respectively.

A. Overview of the Proposed Joint Precoding Framework

As the basis of the joint precoding design, we assume that the CSI of the whole RIS-aided cell-free network can be fully acquired by the CPU in advance. Then, based on the fully-known CSI, the joint precoding at the BSs and RISs are further designed and employed. Intuitively, we draw Fig. 3 to show this dynamical process over time. Under the assumption of fully-known CSI in each small timescale, we focus on solving the problem $\mathcal{P}^{\mathrm{o}}$ in (10) to find a feasible precoding design $\mathbf{W}^{\mathrm{opt}}$ and $\boldsymbol{\Theta}^{\mathrm{opt}}$ for each small timescale as follows.

At first, to deal with the complexity of sum-logarithms in the WSR maximization problem $\mathcal{P}^{\mathrm{o}}$ in (10), by utilizing Lagrangian dual reformulation (LDR), a method has been proposed in [34] to decouple the logarithms. Based on this, we have the following *Proposition 1*.

Proposition 1: By introducing an auxiliary variable $\boldsymbol{\rho} \in \mathbb{R}^{PK}$ with $\boldsymbol{\rho} = [\rho_{1,1}, \rho_{1,2}, \ldots, \rho_{1,K}, \rho_{2,1}, \rho_{2,2}, \ldots, \rho_{P,K}]^T$, the

---

Algorithm 1: Proposed Joint Precoding Framework.

Input: All channels $\mathbf{H}_{b,k,p}$, $\mathbf{G}_{b,r,p}$ and $\mathbf{F}_{r,k,p}$ where $\forall b \in \mathcal{B}, k \in \mathcal{K}, p \in \mathcal{P}$.

Output: Optimized active precoding vector $\mathbf{W}$; Optimized passive precoding matrix $\boldsymbol{\Theta}$; Weighted sum-rate R_{sum}.

1:   Initialize $\mathbf{W}$ and $\boldsymbol{\Theta}$;
2:   **while** no convergence of R_{sum} **do**
3:       Update $\boldsymbol{\rho}$ by (15);
4:       Update $\boldsymbol{\xi}$ by (19);
5:       Update $\mathbf{W}$ by solving (24);
6:       Update $\boldsymbol{\varpi}$ by (29);
7:       Update $\boldsymbol{\Theta}$ by solving (35);
8:   **end while**
9:   **return** $\mathbf{W}^{\mathrm{opt}}$, $\boldsymbol{\Theta}^{\mathrm{opt}}$, and R_{sum}.

---

original problem $\mathcal{P}^{\mathrm{o}}$ in (10) is equivalent to

$$\bar{\mathcal{P}} : \max_{\boldsymbol{\Theta}, \mathbf{W}, \boldsymbol{\rho}} \ f(\boldsymbol{\Theta}, \mathbf{W}, \boldsymbol{\rho})$$

$$\text{s.t.} \quad C_1 : \sum_{k=1}^{K} \sum_{p=1}^{P} \|\mathbf{w}_{b,p,k}\|^2 \le P_{b,\max}, \forall b \in \mathcal{B},$$

$$C_2 : \theta_{r,n} \in \mathcal{F}, \forall r \in \mathcal{R}, \forall n \in \mathcal{N}, \quad (12)$$

where the new objective function $f(\boldsymbol{\Theta}, \mathbf{W}, \boldsymbol{\rho})$ is

$$f(\boldsymbol{\Theta}, \mathbf{W}, \boldsymbol{\rho}) = \sum_{k=1}^{K} \sum_{p=1}^{P} \eta_k \ln\left(1 + \rho_{k,p}\right) - \sum_{k=1}^{K} \sum_{p=1}^{P} \eta_k \rho_{k,p}$$

$$+ \sum_{k=1}^{K} \sum_{p=1}^{P} \eta_k (1 + \rho_{k,p}) f_{k,p}(\boldsymbol{\Theta}, \mathbf{W}), \quad (13)$$

wherein the function $f_{k,p}(\boldsymbol{\Theta}, \mathbf{W})$ is denoted by

$$f_{k,p}(\boldsymbol{\Theta}, \mathbf{W}) =$$

$$\mathbf{w}_{p,k}^H \mathbf{h}_{k,p} \left(\sum_{j=1}^{K} \mathbf{h}_{k,p}^H \mathbf{w}_{p,j} \left(\mathbf{h}_{k,p}^H \mathbf{w}_{p,j}\right)^H + \boldsymbol{\Xi}_{k,p} \right)^{-1} \mathbf{h}_{k,p}^H \mathbf{w}_{p,k}. \quad (14)$$

Then, we propose the joint active and passive precoding framework to optimize the variables $\boldsymbol{\rho}$, $\mathbf{W}$, and $\boldsymbol{\Theta}$ in (12) iteratively. After introducing two auxiliary variables $\boldsymbol{\xi}$ and $\boldsymbol{\varpi}$, the proposed joint precoding framework to maximize the WSR is summarized in **Algorithm 1**. In this framework, the variables $\boldsymbol{\rho}, \boldsymbol{\xi}, \mathbf{W}, \boldsymbol{\varpi}$, and $\boldsymbol{\Theta}$ are alternately updated until the convergence of the objective function is achieved.

The optimal solutions to these variables at each step will be introduced in the following three subsections. Specifically, the solution to $\boldsymbol{\rho}^{\mathrm{opt}}$ is firstly present in Subsection III-B. Then, the solutions to $\boldsymbol{\xi}^{\mathrm{opt}}$ and $\mathbf{W}^{\mathrm{opt}}$ for the active precoding design are provided in Subsection III-C. After that, the solutions to $\boldsymbol{\varpi}^{\mathrm{opt}}$ and $\boldsymbol{\Theta}^{\mathrm{opt}}$ for the passive precoding design are finally discussed in Subsection III-F.

B. Fix $(\boldsymbol{\Theta}, \mathbf{W})$ and Solve $\boldsymbol{\rho}^{\text{opt}}$

Given fixed $(\boldsymbol{\Theta}^\star, \mathbf{W}^\star)$, the optimal $\boldsymbol{\rho}$ in (13) can be obtained by solving $\partial f / \partial \rho_{k,p} = 0$ for $\forall k \in \mathcal{K}, \forall p \in \mathcal{P}$. The solution can be written as

$$\rho_{k,p}^{\text{opt}} = \gamma_{k,p}^\star, \forall k \in \mathcal{K}, \forall p \in \mathcal{P}. \tag{15}$$

By substituting $\rho_{k,p}^{\text{opt}}$ in (15) back into f in (13), one can notice that, only the last term in (13) is associated with the variables $\mathbf{W}$ and $\boldsymbol{\Theta}$. Hence, the problem $\bar{\mathcal{P}}$ in (12) can be further solved as shown in the following two subsections.

C. Active Precoding: Fix $(\boldsymbol{\Theta}, \boldsymbol{\rho})$ and Solve $\mathbf{W}^{\text{opt}}$

In the case of given $(\boldsymbol{\Theta}^\star, \boldsymbol{\rho}^\star)$, the equivalent WSR maximization problem $\bar{\mathcal{P}}$ in (12) can be reformulated as the following subproblem $\mathcal{P}_{\text{active}}$ for the active precoding design at BSs:

$$\mathcal{P}_{\text{active}} : \max_{\mathbf{W}} \ g_1(\mathbf{W}) = \sum_{k=1}^{K} \sum_{p=1}^{P} \mu_{k,p} f_{k,p}(\boldsymbol{\Theta}^\star, \mathbf{W})$$
$$\tag{16}$$
$$\text{s.t.} \quad C_1 : \sum_{k=1}^{K} \sum_{p=1}^{P} \|\mathbf{w}_{b,p,k}\|^2 \leq P_{b,\max}, \forall b \in \mathcal{B},$$

where $\mu_{k,p} = \eta_k (1 + \rho_{k,p}^\star)$ holds. However, note that the reformulated subproblem $\mathcal{P}_{\text{active}}$ in (16) is still too difficult to solve due to the high-dimensional non-convex $f_{k,p}$ in (14). Specifically, due to the channels introduced by the RIS-aided cell-free network, this subproblem is actually a special high-dimensional sum-of-fractions problem. Different from the familiar scalar-form fractions, the high-dimensional "fractions" in (14) are the products of matrices and inverse matrices. Thereby, the non-convexity of $f_{k,p}$ in (14) cannot be simply relaxed by adopting the common FP methods such as the Dinkelbach's algorithms [35].

To tackle this issue, we fortunately notice a recently proposed method called *multidimensional complex quadratic transform (MCQT)* [34]. Different from the common FP methods, MCQT extends the common scalar-form fractional programming to matrix-form and can be utilized to address the non-convexity of the high-dimensional "fractions" [34]. Since $f_{k,p}$ in (16) just meets the *concave-convex* conditions required by MCQT, we can apply MCQT to reformulate the subproblem (16) to address its non-convexity. In this way, we obtain *Proposition 2* as below.

Proposition 2: Exploiting the fractional programming method MCQT and by introducing auxiliary variables $\boldsymbol{\xi}_{p,k} \in \mathbb{C}^U$ with $\boldsymbol{\xi} = [\boldsymbol{\xi}_{1,1}, \boldsymbol{\xi}_{1,2}, \ldots, \boldsymbol{\xi}_{1,K}, \boldsymbol{\xi}_{2,1}, \boldsymbol{\xi}_{2,2}, \ldots, \boldsymbol{\xi}_{P,K}]$, the subproblem $\mathcal{P}_{\text{active}}$ in (16) can be further reformulated as

$$\bar{\mathcal{P}}_{\text{active}} : \max_{\mathbf{W}, \boldsymbol{\xi}} \ g_2(\mathbf{W}, \boldsymbol{\xi})$$
$$\text{s.t.} \quad C_1 : \sum_{k=1}^{K} \sum_{p=1}^{P} \|\mathbf{w}_{b,p,k}\|^2 \leq P_{b,\max}, \forall b \in \mathcal{B},$$
$$\tag{17}$$

where

$$g_2(\mathbf{W}, \boldsymbol{\xi}) = \sum_{k=1}^{K} \sum_{p=1}^{P} 2\sqrt{\mu_{k,p}} \, \Re \left\{ \boldsymbol{\xi}_{k,p}^H \mathbf{h}_{k,p}^H \mathbf{w}_{p,k} \right\}$$
$$- \sum_{k=1}^{K} \sum_{p=1}^{P} \boldsymbol{\xi}_{k,p}^H \left(\sum_{j=1}^{K} \mathbf{h}_{k,p}^H \mathbf{w}_{p,j} \left(\mathbf{h}_{k,p}^H \mathbf{w}_{p,j} \right)^H + \boldsymbol{\Xi}_{k,p} \right) \boldsymbol{\xi}_{k,p}.$$
$$\tag{18}$$

Thereby, the updating of $\mathbf{W}$ can be divided into two steps of updating $\boldsymbol{\xi}$ and $\mathbf{W}$ in turn. To achieve this, the reformulated subproblem $\bar{\mathcal{P}}_{\text{active}}$ in (17) can be divided into two subproblems and solved respectively as follows.

1) Fix $\mathbf{W}$ and Solve $\boldsymbol{\xi}^{\text{opt}}$: While fixing $\mathbf{W}$ in $\bar{\mathcal{P}}_{\text{active}}$ in (17), by setting $\partial g_2 / \partial \boldsymbol{\xi}_{k,p}$ to zero, the optimal $\boldsymbol{\xi}$ can be obtained by

$$\boldsymbol{\xi}_{k,p}^{\text{opt}} = \sqrt{\mu_{k,p}} \left(\sum_{j=1}^{K} \mathbf{h}_{k,p}^H \mathbf{w}_{p,j} \left(\mathbf{h}_{k,p}^H \mathbf{w}_{p,j} \right)^H + \boldsymbol{\Xi}_{k,p} \right)^{-1} \mathbf{h}_{k,p}^H \mathbf{w}_{p,k},$$
$$\forall k \in \mathcal{K}, \forall p \in \mathcal{P}. \tag{19}$$

2) Fix $\boldsymbol{\xi}$ and Solve $\mathbf{W}^{\text{opt}}$: While fixing $\boldsymbol{\xi}$ in $\bar{\mathcal{P}}_{\text{active}}$ in (17), for simplification and clarity of (17), we can first define

$$\mathbf{a}_p = \sum_{k=1}^{K} \mathbf{h}_{k,p} \boldsymbol{\xi}_{k,p} \boldsymbol{\xi}_{k,p}^H \mathbf{h}_{k,p}^H, \tag{20a}$$

$$\mathbf{A}_p = \mathbf{I}_K \otimes \mathbf{a}_p, \qquad \mathbf{v}_{k,p} = \mathbf{h}_{k,p} \boldsymbol{\xi}_{k,p}. \tag{20b}$$

Then, by substituting (20) into g_2 in (18), we can rewritten g_2 as

$$g_2(\mathbf{W}) = -\mathbf{W}^H \mathbf{A} \mathbf{W} + \Re \left\{ 2\mathbf{V}^H \mathbf{W} \right\} - Y, \tag{21}$$

where

$$\mathbf{A} = \text{diag}(\mathbf{A}_1, \ldots, \mathbf{A}_P), \quad Y = \sum_{k=1}^{K} \sum_{p=1}^{P} \boldsymbol{\xi}_{k,p}^H \boldsymbol{\Xi}_{k,p} \boldsymbol{\xi}_{k,p}, \tag{22a}$$

$$\mathbf{V} = [\mathbf{v}_{1,1}^T, \mathbf{v}_{1,2}^T, \ldots, \mathbf{v}_{1,K}^T, \mathbf{v}_{2,1}^T, \mathbf{v}_{2,2}^T, \ldots, \mathbf{v}_{P,K}^T]^T. \tag{22b}$$

Therefore, the active precoding problem $\bar{\mathcal{P}}_{\text{active}}$ in (17) can be further simplified as

$$\hat{\mathcal{P}}_{\text{active}} : \min_{\mathbf{W}} \ g_3(\mathbf{W}) = \mathbf{W}^H \mathbf{A} \mathbf{W} - \Re \left\{ 2\mathbf{V}^H \mathbf{W} \right\}$$
$$\text{s.t.} \quad C_1 : \mathbf{W}^H \mathbf{D}_b \mathbf{W} \leq P_{b,\max}, \forall b \in \mathcal{B}, \tag{24}$$

where $\mathbf{D}_b = \mathbf{I}_{PK} \otimes \{ (\mathbf{e}_b \mathbf{e}_b^H) \otimes \mathbf{I}_M \}$ with $\mathbf{e}_b \in \mathbb{R}^B$. Since the matrices $\mathbf{A}$ and $\mathbf{D}_b$ $(\forall b \in \mathcal{B})$ are all positive semidefinite, the simplified subproblem $\hat{\mathcal{P}}_{\text{active}}$ in (24) is a standard QCQP problem, which can be optimally solved by many existing methods such as alternating direction method of multipliers (ADMM) [36].

However, note that the adoption of ADMM in QCQP problem requires the inversion for the matrix $\mathbf{A}$ (along with Lagrange multipliers). Due to the high-dimensional channels of cell-free network, the dimension of $\mathbf{A}$ is usually very high $(BMPK)$. As

$$g_4(\boldsymbol{\Theta}) = \sum_{k=1}^{K} \sum_{p=1}^{P} \sqrt{\mu_{k,p}} \mathbf{Q}_{k,p,k}^H(\boldsymbol{\Theta}) \left(\sum_{j=1}^{K} \mathbf{Q}_{k,p,j}(\boldsymbol{\Theta}) \mathbf{Q}_{k,p,k}^H(\boldsymbol{\Theta}) + \boldsymbol{\Xi}_{k,p} \right)^{-1} \mathbf{Q}_{k,p,k}(\boldsymbol{\Theta}). \tag{23}$$

a result, the inversion for $\mathbf{A}$ has a high computational complexity of about $\mathcal{O}(B^3 M^3 P^3 K^3)$, which may prevent the precoding design from practical application. To avoid the matrix inversion operation thus reducing the complexity, here we provide an inversion-free feasible solution by exploiting the *primal-dual subgradient (PDS)* method [37] to obtain $\mathbf{W}^{\mathrm{opt}}$ in Appendix A.

D. Passive Precoding: Fix $(\boldsymbol{\rho}, \mathbf{W})$ and Solve $\boldsymbol{\Theta}^{\mathrm{opt}}$

Based on the given $(\boldsymbol{\rho}^\star, \mathbf{W}^\star)$, for the equivalent WSR maximization problem $\bar{\mathcal{P}}$ in (12), the subproblem of the RIS precoding design at RISs can be equivalently rewritten as

$$\mathcal{P}_{\mathrm{passive}} : \max_{\boldsymbol{\Theta}} \quad g_4(\boldsymbol{\Theta}) = \sum_{k=1}^{K} \sum_{p=1}^{P} \mu_{k,p} f_{k,p}(\boldsymbol{\Theta}, \mathbf{W}^\star) \quad (25)$$

$$\text{s.t.} \quad C_2 : \theta_{r,n} \in \mathcal{F}, \forall r \in \mathcal{R}, \forall n \in \mathcal{N},$$

where $\mu_{k,p} = \eta_k(1 + \rho_{k,p}^\star)$. Similarly, to reduce the complexity, we wish to simplify the expression of g_4 in (25). Firstly, by defining a new auxiliary function with respect to $\boldsymbol{\Theta}$ as

$$\mathbf{Q}_{k,p,j}(\boldsymbol{\Theta}) = \sum_{b=1}^{B} \left(\mathbf{H}_{b,k,p}^H + \mathbf{F}_{k,p}^H \boldsymbol{\Theta}^H \mathbf{G}_{b,p} \right) \mathbf{w}_{b,p,j}, \quad (26)$$

we can rewrite g_4 in (25) as (23) at the bottom of previous page. However, this subproblem is still hard to solve due to the mutidimensional fractions in $f_{k,p}$ in (14). Notice that the subproblem (25) satisfies the *concave-convex* conditions [34], again we exploit the MCQT to address this issue again by using the following *Proposition 3*.

Proposition 3: With the fractional programming method MCQT, by introducing an auxiliary variable $\boldsymbol{\varpi}_{p,k} \in \mathbb{C}^U$ and $\boldsymbol{\varpi} = [\boldsymbol{\varpi}_{1,1}, \boldsymbol{\varpi}_{1,2}, \ldots, \boldsymbol{\varpi}_{1,K}, \boldsymbol{\varpi}_{2,1}, \boldsymbol{\varpi}_{2,2}, \ldots, \boldsymbol{\varpi}_{P,K}]$, the passive precoding subproblem $\mathcal{P}_{\mathrm{passive}}$ in (25) can be reformulated as

$$\bar{\mathcal{P}}_{\mathrm{passive}} : \max_{\boldsymbol{\Theta}} \quad g_5(\boldsymbol{\Theta}, \boldsymbol{\varpi}) = \sum_{k=1}^{K} \sum_{p=1}^{P} g_{k,p}(\boldsymbol{\Theta}, \boldsymbol{\varpi}) \quad (27a)$$

$$\text{s.t.} \quad C_2 : \theta_{r,n} \in \mathcal{F}, \forall r \in \mathcal{R}, \forall n \in \mathcal{N}, \quad (27b)$$

where

$$g_{k,p}(\boldsymbol{\Theta}, \boldsymbol{\varpi}) = 2\sqrt{\mu_{k,p}} \Re \left\{ \boldsymbol{\varpi}_{k,p}^H \mathbf{Q}_{k,p,k}(\boldsymbol{\Theta}) \right\}$$

$$- \boldsymbol{\varpi}_{k,p}^H \left(\sum_{j=1}^{K} \mathbf{Q}_{k,p,j}(\boldsymbol{\Theta}) \mathbf{Q}_{k,p,j}^H(\boldsymbol{\Theta}) + \boldsymbol{\Xi}_{k,p} \right) \boldsymbol{\varpi}_{k,p}. \quad (28)$$

Next, similar to the previous processing of solving the subproblem $\bar{\mathcal{P}}_{\mathrm{active}}$ in (17), we consider to optimize two variables $\boldsymbol{\varpi}$ and $\boldsymbol{\Theta}$ in (27) in turn. The reformulated subproblem $\bar{\mathcal{P}}_{\mathrm{passive}}$ in (27) can be further divided into two subproblems and respectively solved as follows.

1) Fix $\boldsymbol{\Theta}$ and Solve $\boldsymbol{\varpi}^{\mathrm{opt}}$: For given fixed $\boldsymbol{\Theta}$ in $\bar{\mathcal{P}}_{\mathrm{passive}}$ in (27), by solving $\partial g_5 / \partial \boldsymbol{\varpi}_{k,p} = 0$ for $\forall k \in \mathcal{K}$ and $\forall p \in \mathcal{P}$, we can obtain the optimal $\boldsymbol{\varpi}_{k,p}$ for all $k \in \mathcal{K}$ and $p \in \mathcal{P}$ by

$$\boldsymbol{\varpi}_{k,p}^{\mathrm{opt}} =$$

$$\sqrt{\mu_{k,p}} \left(\sum_{j=1}^{K} \mathbf{Q}_{k,p,j}(\boldsymbol{\Theta}) \mathbf{Q}_{k,p,j}^H(\boldsymbol{\Theta}) + \boldsymbol{\Xi}_{k,p} \right)^{-1} \mathbf{Q}_{k,p,k}(\boldsymbol{\Theta}). \quad (29)$$

1) Fix $\boldsymbol{\varpi}$ and Solve $\boldsymbol{\Theta}^{\mathrm{opt}}$: While fixing $\boldsymbol{\varpi}$ in g_5 in (27), due to the complexity of $\bar{\mathcal{P}}_{\mathrm{passive}}$ in (27), we first consider to simplify the expression of g_5 by using the new auxiliary function $\mathbf{Q}_{k,p,j}(\boldsymbol{\Theta})$ with respect to $\boldsymbol{\Theta}$ in (26) as follows:

$$\boldsymbol{\varpi}_{k,p}^H \mathbf{Q}_{k,p,j}(\boldsymbol{\Theta})$$

$$\overset{(a)}{=} \sum_{b=1}^{B} \left(\boldsymbol{\varpi}_{k,p}^H \mathbf{H}_{b,k,p}^H \mathbf{w}_{b,p,j} + \boldsymbol{\varpi}_{k,p}^H \mathbf{F}_{k,p}^H \boldsymbol{\Theta}^H \mathbf{G}_{b,p} \mathbf{w}_{b,p,j} \right)$$

$$\overset{(b)}{=} \sum_{b=1}^{B} \boldsymbol{\varpi}_{k,p}^H \mathbf{H}_{b,k,p}^H \mathbf{w}_{b,p,j} + \boldsymbol{\theta}^H \sum_{b=1}^{B} \mathrm{diag}\left(\boldsymbol{\varpi}_{k,p}^H \mathbf{F}_{k,p}^H \right) \mathbf{G}_{b,p} \mathbf{w}_{b,p,j}$$

$$\overset{(c)}{=} c_{k,p,j} + \boldsymbol{\theta}^H \mathbf{g}_{k,p,j}, \quad (30)$$

where (a) holds according to (23), (b) is obtained by defining $\boldsymbol{\theta} = \boldsymbol{\Theta} \mathbf{1}_{RN}$, and (c) is achieved by defining

$$c_{k,p,j} = \sum_{b=1}^{B} \boldsymbol{\varpi}_{k,p}^H \mathbf{H}_{b,k,p}^H \mathbf{w}_{b,p,j}, \quad (31a)$$

$$\mathbf{g}_{k,p,j} = \sum_{b=1}^{B} \mathrm{diag}\left(\boldsymbol{\varpi}_{k,p}^H \mathbf{F}_{k,p}^H \right) \mathbf{G}_{b,p} \mathbf{w}_{b,p,j}. \quad (31b)$$

By substituting (30) into (28), we obtain:

$$g_{k,p}(\boldsymbol{\Theta}) = 2\sqrt{\mu_{k,p}} \Re \left\{ c_{k,p,k} + \boldsymbol{\theta}^H \mathbf{g}_{k,p,k} \right\}$$

$$- \sum_{j=1}^{K} \left(c_{k,p,j} + \boldsymbol{\theta}^H \mathbf{g}_{k,p,j} \right) \left(c_{k,p,j}^* + \mathbf{g}_{k,p,j}^H \boldsymbol{\theta} \right) - \boldsymbol{\varpi}_{k,p}^H \boldsymbol{\Xi}_{k,p} \boldsymbol{\varpi}_{k,p}. \quad (32)$$

Then, we can further substitute (32) into (27a), so that g_5 in (27a) can be simplified as

$$g_5(\boldsymbol{\Theta}) = -\boldsymbol{\theta}^H \boldsymbol{\Lambda} \boldsymbol{\theta} + \Re \left\{ 2\boldsymbol{\theta}^H \boldsymbol{\nu} \right\} - \zeta, \quad (33)$$

where

$$\boldsymbol{\Lambda} = \sum_{k=1}^{K} \sum_{p=1}^{P} \sum_{j=1}^{K} \mathbf{g}_{k,p,j} \mathbf{g}_{k,p,j}^H, \quad (34a)$$

$$\boldsymbol{\nu} = \sum_{k=1}^{K} \sum_{p=1}^{P} \sqrt{\mu_{k,p}} \mathbf{g}_{k,p,k} - \sum_{k=1}^{K} \sum_{p=1}^{P} \sum_{j=1}^{K} c_{k,p,j}^* \mathbf{g}_{k,p,j}, \quad (34b)$$

$$\zeta = \sum_{k=1}^{K} \sum_{p=1}^{P} \sum_{j=1}^{K} |c_{k,p,j}|^2 + \sum_{k=1}^{K} \sum_{p=1}^{P} \boldsymbol{\varpi}_{k,p}^H \boldsymbol{\Xi}_{k,p} \boldsymbol{\varpi}_{k,p}$$

$$- 2 \sum_{k=1}^{K} \sum_{p=1}^{P} \sqrt{\mu_{k,p}} \Re \left\{ c_{k,p,k} \right\}. \quad (34c)$$

Therefore, the reformulated passive precoding subproblem $\bar{\mathcal{P}}_{\text{passive}}$ in (27) can be further simplified as

$$\hat{\mathcal{P}}_{\text{passive}} : \min_{\boldsymbol{\Theta}} \quad g_6(\boldsymbol{\Theta}) = \boldsymbol{\theta}^H \boldsymbol{\Lambda} \boldsymbol{\theta} - \Re\left\{2\boldsymbol{\theta}^H \boldsymbol{\nu}\right\}$$

$$\text{s.t.} \quad C_2 : \theta_{r,n} \in \mathcal{F}, \forall r \in \mathcal{R}, \forall n \in \mathcal{N}. \tag{35}$$

This simplified subproblem $\hat{\mathcal{P}}_{\text{passive}}$ is similar to those in [19], [38]. Since the matrix $\boldsymbol{\Lambda}$ is positive semidefinite, the objective function is convex. Besides, since we have $\mathcal{F} \triangleq \{\theta_{r,j} \big| |\theta_{r,j}| \leq 1\}$ according to (4), the constraint C_2 is also convex. Thereby, this subproblem $\hat{\mathcal{P}}_{\text{passive}}$ can be solved by ADMM [36].

However, similar to the optimization of active precoding vector $\mathbf{W}$, note that the adoption of ADMM [36] in this QCQP problem requires the matrix inversion for $\boldsymbol{\Lambda}$ with high computational complexity of $\mathcal{O}(R^3 N^3)$. Since the RIS element number N is usually very large, the complexity of using ADMM is very high. Again, to reduce the complexity, here we also provide a potential method based on the PDS to obtain the optimal solution $\boldsymbol{\Theta}^{\text{opt}}$, which will be detailedly introduced in Appendix B.

IV. TWO-TIMESCALE EXTENSION OF THE PROPOSED JOINT PRECODING FRAMEWORK

Up to now, we have provided a joint precoding framework for the proposed RIS-aided cell-free network under the assumption of fully-known CSI in every small timescale, as shown in Fig. 3. However, due to the inherent high-dimensional channels introduced by RISs, acquiring all RIS-aided channels so frequently is usually unrealistic [22]. In particular, this issue is greatly exacerbated by the high-density of users, BSs, and RISs in RIS-aided cell-free networks, which prevents many existing joint precoding schemes based on fully-known CSI from practical adopting.

To tackle this issue, in this section, the proposed joint precoding framework is further extended to a more practical two-timescale scheme, which can serve as a trade-off scheme between overhead and performance. The key idea is to match each user with several well-performed RISs at the beginning of a large timescale. Then, in later several small timescales, only the RIS-aided channels of the matched user-RIS pairs are acquired and utilized for joint precoding design, while those of the unmatched pairs are temporarily ignored to relieve the pressure of RIS channel acquisitions.

Specifically, this section is summarized as follows. Firstly, in Subsection IV-A, the two-timescale extension of the proposed framework is overviewed. Then, the user-RIS matching problem is formulated in Subsection IV-B. Finally, a LCR-based method is proposed to solve this problem in Subsection IV-C.

A. Overview of the Proposed Two-Timescale Extension

In this subsection, the two-timescale extension of the proposed joint precoding framework is proposed as a more practical scheme to address the challenge of high-dimensional channels introduced by RISs in cell-free networks. Intuitively, we present how this two-timescale scheme works in Fig. 4, which illustrates its dynamical working process over time in a large timescale.
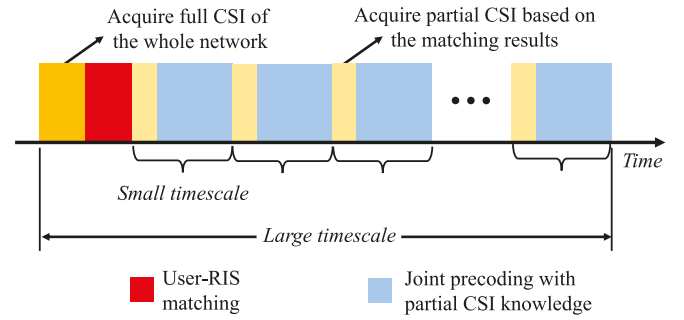

Fig. 4. The dynamical working process of the two-timescale extension of the proposed joint precoding framework.

Particularly, the proposed two-timescale scheme exploits the properties that RISs far from users have little contribution to system capacity and the user's mobile scope is limited in a large timescale, thus the RISs far from users can be ignored temporarily. Specifically, at the beginning of a large timescale, all BS-RIS-user channels are fully acquired once. Then, under the principle of matching relatively strong user-RIS pairs, each user is matched with several well-performed RISs. Subsequently, in the next several small timescales, only the RIS-aided channels of the matched user-RIS pairs and the direct-link channels are acquired and utilized for precoding design while those unmatched are ignored. Finally, the above process will be repeated in the next large timescale.

With this two-timescale scheme, only the limited CSI is acquired and utilized in small timescales, and it is unnecessary to acquire all RIS-aided channels so frequently as before. As the cost of the incomplete CSI, the two-timescale scheme will lead to performance loss, while we will show this loss is actually very limited in later Section VI. Consequently, compared with the original scheme in Fig. 3, the two-timescale scheme in Fig. 4 is more practical to be applied, especially for the networks with a large number of distributed RISs.

B. User-RIS Matching: Problem Formulation

To determine which RIS-aided channels should be acquired and utilized in small timescales as shown in Fig. 4, in this subsection, the user-RIS matching problem is formulated.

Firstly, let $u_{k,r} \in \{0, 1\}$ denote the indicator variable that indicates whether the user k and the RIS r is matched. Particularly, $u_{k,r}$ being one/zero means that the channels between user k and RIS r will/won't be acquired and utilized in later small timescales. To simplify the expression, we define $\mathbf{u}_k = [u_{k,1}, \ldots, u_{k,R}]^T$ and $\mathbf{u} = [\mathbf{u}_1^T, \ldots, \mathbf{u}_K^T]^T$. To determine which user-RIS pairs should be matched, we define the following virtual sum-rate $\hat{R}_{\text{sum}}$ as the bonus function[3] of the user-RIS

---

[3]This bonus function is chosen for the expected result that, by maximizing $\hat{R}_{\text{sum}}$, the user-RIS pairs with stronger links are more preferred to be matched, while those unmatched pairs have relatively little contribution to capacity improvement so that they can be temporarily ignored in later small timescales to relieve the pressure of RIS channel acquisitions.

matching problem:

$$\hat{R}_{\text{sum}} = \sum_{k=1}^{K} \sum_{p=1}^{P} \eta_k \log_2 \left(1 + \hat{\gamma}_{k,p}\right) \tag{36}$$

where the virtual SINR $\hat{\gamma}_{k,p}$ is equivalent to the real SINR $\gamma_{k,p}$ in (8) in which the end-to-end channels $\mathbf{h}_{b,k,p}^H$ are replaced by

$$\hat{\mathbf{h}}_{b,k,p}^H = \mathbf{H}_{b,k,p}^H + \sum_{r=1}^{R} u_{k,r} \mathbf{F}_{r,k,p}^H \boldsymbol{\Theta}_r^H \mathbf{G}_{b,r,p}. \tag{37}$$

Next, let R_{match} denote the maximum number of RISs that each user can be matched with ($R_{\text{match}} \leq R$), thus the user-RIS matching problem $\mathcal{P}_{\text{match}}$ can be formulated as

$$\mathcal{P}_{\text{match}} : \max_{\mathbf{u}, \mathbf{W}, \boldsymbol{\Theta}} \quad \hat{R}_{\text{sum}}(\mathbf{u}, \mathbf{W}, \boldsymbol{\Theta}) = \sum_{k=1}^{K} \sum_{p=1}^{P} \eta_k \log_2 \left(1 + \hat{\gamma}_{k,p}\right)$$

$$\text{s.t.} \quad C_1 : \sum_{k=1}^{K} \sum_{p=1}^{P} \|\mathbf{w}_{b,p,k}\|^2 \leq P_{b,\max}, \forall b \in \mathcal{B},$$

$$C_2 : \theta_{r,n} \in \mathcal{F}, \forall r \in \mathcal{R}, \forall n \in \mathcal{N},$$

$$C_3 : u_{k,r} \in \{0,1\}, \forall k \in \mathcal{K}, \forall r \in \mathcal{R},$$

$$C_4 : \sum_{r=1}^{R} u_{k,r} \leq R_{\text{match}}, \forall k \in \mathcal{K}, \tag{38}$$

where C_3 is zero-one constraint of the indicator variable $\mathbf{u}$, and C_4 ensures that each user can only be matched with at most R_{match} RISs.

Our goal is to find a best-possible matching vector $\mathbf{u}^{\text{opt}}$ to maximize the virtual sum-rate[4] $\hat{R}_{\text{sum}}$ in user-RIS matching problem $\mathcal{P}_{\text{match}}$ in (38). However, due to the non-convexity of zero-one constraint C_3, the problem $\mathcal{P}_{\text{match}}$ is challenging. Again, we consider to solve the problem via alternative optimization, which will be detailedly introduced in the next subsection.

C. User-RIS Matching: LCR-Based Method

We consider to solve the matching problem $\mathcal{P}^{\text{m}}$ in (38) by optimizing $\mathbf{u}$, $\mathbf{W}$, and $\boldsymbol{\Theta}$ alternately. Note that, when $\mathbf{u}$ is fixed, the problem $\mathcal{P}_{\text{match}}$ in (38) has the same form as the original problem $\mathcal{P}^{\text{o}}$ in (10). Therefore, the desired algorithm to solve $\mathcal{P}_{\text{match}}$ can be realized by inserting a new step of optimizing $\mathbf{u}$ into **Algorithm 1**, and here we summarize it as **Algorithm 2** for clarity. In the rest of this subsection, we will focus on the step 8, i.e., optimizing $\mathbf{u}$ while fixing the other variables, to complete this algorithm.

Firstly, we consider to reformulate $\mathcal{P}_{\text{match}}$ as a solvable convex problem. Similar to the optimization of $\boldsymbol{\Theta}$, by applying LDR in *Proposition 1* and MCQT in *Proposition 3* and then fixing all auxiliary variables, the user-RIS matching problem

---

[4]The virtual WSR $\hat{R}_{\text{sum}}$ is just the bonus function for user-RIS matching instead of the real WSR, while the real WSR can be evaluated by substituting the optimized $\mathbf{W}^{\text{opt}}$ and $\boldsymbol{\Theta}^{\text{opt}}$ into R_{sum} in (9).

Algorithm 2: Proposed User-RIS Matching Method.

Input: All channels $\mathbf{H}_{b,k,p}, \mathbf{G}_{b,r,p}$ and $\mathbf{F}_{r,k,p}$ where $\forall b \in \mathcal{B}, k \in \mathcal{K}, p \in \mathcal{P}$.
Output: Optimized user-RIS matching vector $\mathbf{u}$.
1:    Initialize $\boldsymbol{\Theta}$, $\mathbf{W}$, and $\mathbf{u}$;
2:    **while** no convergence of $\hat{R}_{\text{sum}}$**do**
3:      Update $\boldsymbol{\rho}$ by (15);
4:      Update $\boldsymbol{\xi}$ by (19);
5:      Update $\mathbf{W}$ by solving (24);
6:      Update $\boldsymbol{\varpi}$ by (29);
7:      Update $\boldsymbol{\Theta}$ by solving (35);
8:      Update $\mathbf{u}$ by solving (46);
9:    **end while**
10:   **return** $\mathbf{u}^{\text{opt}}$.

$\mathcal{P}_{\text{match}}$ can be reformulated as:

$$\bar{\mathcal{P}}_{\text{match}} : \max_{\mathbf{u}} \quad \hat{g}(\mathbf{u}) = \sum_{k=1}^{K} \sum_{p=1}^{P} \hat{g}_{k,p}(\mathbf{u})$$

$$\text{s.t.} \quad C_3 : u_{k,r} \in \{0,1\}, \forall k \in \mathcal{K}, \forall r \in \mathcal{R},$$

$$C_4 : \sum_{r=1}^{R} u_{k,r} \leq R_{\text{match}}, \forall k \in \mathcal{K}, \tag{39}$$

where $\hat{g}_{k,p}(\mathbf{u})$ is defined as

$$\hat{g}_{k,p}(\mathbf{u}) = 2\sqrt{\mu_{k,p}} \mathfrak{R}\left\{\boldsymbol{\varpi}_{k,p}^H \hat{\mathbf{Q}}_{k,p,k}(\mathbf{u})\right\} -$$

$$\boldsymbol{\varpi}_{k,p}^H \left(\sum_{j=1}^{K} \hat{\mathbf{Q}}_{k,p,j}(\mathbf{u}) \hat{\mathbf{Q}}_{k,p,j}^H(\mathbf{u}) + \boldsymbol{\Xi}_{k,p}\right) \boldsymbol{\varpi}_{k,p}, \tag{40}$$

in which the auxiliary functions $\hat{\mathbf{Q}}_{k,p,j}(\mathbf{u}) \in \mathbb{C}^U$ satisfy

$$\hat{\mathbf{Q}}_{k,p,j}(\mathbf{u}) = \sum_{b=1}^{B} \left(\mathbf{H}_{b,k,p}^H + \sum_{r=1}^{R} u_{k,r} \mathbf{F}_{r,k,p}^H \boldsymbol{\Theta}_r^H \mathbf{G}_{b,r,p}\right) \mathbf{w}_{b,p,j}. \tag{41}$$

Then, we wish to simplify the expression of the objective function $\hat{g}$ by exploiting the following definitions:

$$\beta_{k,p,j,r} = \sum_{b=1}^{B} \boldsymbol{\varpi}_{k,p}^H \mathbf{F}_{r,k,p}^H \boldsymbol{\Theta}_r^H \mathbf{G}_{b,r,p} \mathbf{w}_{b,p,j}, \tag{42a}$$

$$\boldsymbol{\beta}_{k,p,j} = [\beta_{k,p,j,1}, \beta_{k,p,j,2}, \dots, \beta_{k,p,j,R}]^T. \tag{42b}$$

After substituting (31a) and (42) into the objective function $\hat{g}$ in (39) and then removing the terms unrelated to $\mathbf{u}$, $\hat{g}$ in (39) can be rewritten as

$$\hat{g}_1(\mathbf{u}) = -\mathbf{u}^T \boldsymbol{\Omega} \mathbf{u} + 2\boldsymbol{\zeta}^T \mathbf{u} \tag{43}$$

where

$$\boldsymbol{\Omega} = \text{diag}(\boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2, \dots, \boldsymbol{\Omega}_K), \boldsymbol{\zeta} = \left[\boldsymbol{\zeta}_1^T, \boldsymbol{\zeta}_2^T, \dots, \boldsymbol{\zeta}_K^T\right]^T, \tag{44}$$

in which $\boldsymbol{\Omega}_k$ and $\boldsymbol{\zeta}_k$ are given by

$$\boldsymbol{\Omega}_k = \Re\left\{\sum_{p=1}^{P}\sum_{j=1}^{K}\boldsymbol{\beta}_{k,p,j}\boldsymbol{\beta}_{k,p,j}^H\right\}, \tag{45a}$$

$$\boldsymbol{\zeta}_k = \Re\left\{\sum_{p=1}^{P}\sqrt{\mu_{k,p}}\boldsymbol{\beta}_{k,p,k} - \sum_{p=1}^{P}\sum_{j=1}^{K}c_{k,p,j}^*\boldsymbol{\beta}_{k,p,j}\right\}. \tag{45b}$$

Thus, $\bar{\mathcal{P}}_{\text{match}}$ in (39) can be equivalently reorganized as

$$\hat{\mathcal{P}}_{\text{match}} : \min_{\mathbf{u}} \quad \hat{g}_2(\mathbf{u}) = \mathbf{u}^T\boldsymbol{\Omega}\mathbf{u} - 2\boldsymbol{\zeta}^T\mathbf{u}$$

$$\text{s.t.} \quad C_3 : [\mathbf{u}]_j^2 = [\mathbf{u}]_j, \forall j \in \{1,\ldots,KR\},$$

$$C_4 : \left(\mathbf{I}_K \otimes \mathbf{1}_R^T\right)\mathbf{u} \leq R_{\text{match}}\mathbf{1}_K, \tag{46}$$

where the constraints C_3 and C_4 are also equivalently rewritten for clearer expression.

Note that, the equivalent problem $\hat{\mathcal{P}}_{\text{match}}$ in (46) is actually a zero-one quadratic programming problem with linear constrains, which is proved to be NP-hard [39]. Since the dimension of the matching vector $\mathbf{u}$ is KR, when the number of users K and RISs R is not too large, Bruto-Force method can be directly used to search for the optimal solution to $\hat{\mathcal{P}}_{\text{match}}$. However, when the densities of users and RISs in cell-free network are high, it is unpractical to adopt Bruto-Force due to the large search space (2^{KR} candidates). In this case, we can apply linear conic relaxation (LCR) [40] to relax the zero-one programming problem $\hat{\mathcal{P}}_{\text{match}}$ and try to find a sub-optimal solution as a substitute. In this way, we obtain the following proposition.

Proposition 4: Applying LCR [40] for the zero-one programming problem $\hat{\mathcal{P}}_{\text{match}}$ in (46) and by introducing an symmetric auxiliary matrix $\mathbf{U} \in \mathbb{R}^{KR \times KR}$ to replace the matrix $\mathbf{u}\mathbf{u}^T$, $\hat{\mathcal{P}}_{\text{match}}$ can be reformulated as

$$\tilde{\mathcal{P}}_{\text{match}} : \min_{\mathbf{u},\mathbf{U}} \hat{g}_2(\mathbf{u},\mathbf{U}) = \text{Tr}(\boldsymbol{\Omega}\mathbf{U}) - 2\boldsymbol{\zeta}^T\mathbf{u}$$

$$\text{s.t.} \quad C_3 : [\mathbf{U}]_{j,j} = [\mathbf{u}]_j, \forall j \in \{1,\ldots,KR\},$$

$$C_4 : \left(\mathbf{I}_K \otimes \mathbf{1}_R^T\right)\mathbf{u} \leq R_{\text{match}}\mathbf{1}_K,$$

$$C_5 : [\mathbf{U}]_{i,j} = [\mathbf{U}]_{j,i}, \forall i,j \in \{1,\ldots,KR\},$$

$$C_6 : \left(\mathbf{I}_K \otimes \mathbf{1}_R^T\right)\mathbf{U} \leq R_{\text{match}}\mathbf{1}_K\mathbf{u}^T,$$

$$C_7 : \begin{bmatrix} 1 & \mathbf{u}^T \\ \mathbf{u} & \mathbf{U} \end{bmatrix} \succcurlyeq 0, \tag{47}$$

where the newly added constraints C_5, C_6, C_7 are the guarantees for $\mathbf{U}$ to approach $\mathbf{u}\mathbf{u}^T$ as close as possible, and the rank-one constraint of $\mathbf{U}$ is relaxed.

Notice that, the constraint C_3 in the problem $\tilde{\mathcal{P}}_{\text{match}}$ in (47) is now convex, and thus $\tilde{\mathcal{P}}_{\text{match}}$ is actually a solvable semidefinite programming (SDP) problem. Thereby, the optimal solution $\mathbf{u}^{\text{opt}}$ to $\tilde{\mathcal{P}}_{\text{match}}$ can be efficiently obtained by adopting the modern SDP solvers such as SDPT3 [41], of which the computational complexity is about $\mathcal{O}(K^2R^2(KR+1)^{2.5})$ in the worst case [42].

V. FRAMEWORK SUPPLEMENTS

In this section, more supplements of the proposed joint precoding framework are provided. Specifically, limited by the practical hardware implementation, we discuss the non-ideal RIS case in Subsection V-A. Then, we analyze the convergency and computational complexity of the proposed algorithms in Subsection V-D and Subsection V-E, respectively.

A. Extension to Non-Ideal RIS Cases

Up to now, we have provided a complete joint precoding framework for the proposed RIS-aided cell-free network, where ideal RIS case mentioned in Subsection II-C is considered. In this subsection, we will extend this framework to the more practical non-ideal RIS cases.

1) Non-Ideal RIS Cases: According to (4) in Subsection II-C, $\mathcal{F}$ is defined as the ideal RIS case, where both the amplitude and phase of $\theta_{r,n}$ associated with the RIS element can be controlled independently and continuously. However, limited by the hardware implementation of metamaterials, the RISs in practice are usually non-ideal. For consistent discussion, here we redefine the ideal RIS case $\mathcal{F}$ as $\mathcal{F}_1$. We also define $\mathcal{F}_2$ as the general case where only the phase of $\theta_{r,j}$ can be controlled continuously [32], and $\mathcal{F}_3$ as the practical case where the low-resolution phase of $\theta_{r,n}$ is discrete [24], i.e.,

$$\mathcal{F}_1 \triangleq \left\{\theta_{r,n}\,\middle|\,|\theta_{r,n}| \leq 1\right\}, \tag{48a}$$

$$\mathcal{F}_2 \triangleq \left\{\theta_{r,n}\,\middle|\,|\theta_{r,n}| = 1\right\}, \tag{48b}$$

$$\mathcal{F}_3 \triangleq \left\{\theta_{r,n}\,\middle|\,\theta_{r,n} \in \left\{1, e^{j\frac{2\pi}{L}}, \ldots, e^{j\frac{2\pi(L-1)}{L}}\right\}\right\}, \tag{48c}$$

where L indicates that $\mathcal{F}_3$ contains L discrete phase shifts.

According to the simplified subproblem $\hat{\mathcal{P}}_{\text{passive}}$ in (35), for the passive precoding design at the RISs, when the phase shift constraint $\mathcal{F}$ is the ideal case $\mathcal{F}_1$ as discussed above, the subproblem $\hat{\mathcal{P}}_{\text{passive}}$ can be directly solved due to the convexity of $\mathcal{F}_1$. However, when the constraint $\mathcal{F}$ in (35) is non-ideal $\mathcal{F}_2$ or $\mathcal{F}_3$, the simplified subproblem $\hat{\mathcal{P}}_{\text{passive}}$ in (35) becomes non-convex. To address such problems, we provide the feasible solutions in the following two parts, respectively.

2) Optimal Solution in Continuous Phase Shift Case $\mathcal{F}_2$: When $\mathcal{F} = \mathcal{F}_2$, the subproblem (35) can be equivalently rewritten as

$$\hat{\mathcal{P}}_{\text{passive-con}} : \min_{\boldsymbol{\Theta}} \quad g_6(\boldsymbol{\Theta}) = \boldsymbol{\theta}^H\boldsymbol{\Lambda}\boldsymbol{\theta} - \Re\left\{2\boldsymbol{\theta}^H\boldsymbol{\nu}\right\}$$

$$\text{s.t.} \quad C_2 : |\theta_{r,n}| = 1, \forall r \in \mathcal{R}, \forall n \in \mathcal{N}, \tag{49}$$

which is non-convex due to the constant modulus constraint. Fortunately, a similar subproblem has been discussed in [19, Eqn. 41], while it has been proved that its optimal solution can be obtained by adopting the majorization-minimization (MM) algorithm [43] or complex circle manifold (CCM) algorithm [44].

3) Approximation in Discrete Phase Shift Case $\mathcal{F}_3$: When $\mathcal{F} = \mathcal{F}_3$, the common solution to address the non-convex constraint of discrete space is approximation projection [17], [23], [28]. Specifically, we can first relax $\mathcal{F}_3$ to $\mathcal{F}_2$, and obtain the

TABLE I
COMPUTATIONAL COMPLEXITY FOR UPDATING EACH VARIABLE

Variable	Computational complexity $\mathcal{O}(\cdot)$
$\boldsymbol{\rho}$	$PK\left(KUBM + U^3 + (K+1)U^2 + U\right)$
$\boldsymbol{\xi}$	$PK\left(KUBM + U^3 + (K+1)U^2\right)$
$\mathbf{W}$	$I_{\mathrm{a}}\left(B^2M^2P^2K^2 + BMPK\right)$
$\boldsymbol{\varpi}$	$PK\left(U^3 + (K+1)U^2\right)$
$\boldsymbol{\Theta}$	$I_{\mathrm{p}}\left(R^2N^2 + RN\right)$
$\mathbf{u}$	$K^2R^2\left(KR+1\right)^{2.5}$

optimal solution $\boldsymbol{\theta}^{\mathrm{opt}}$ by solving (49). Then, following the proximity principle, we can simply project the solved $\boldsymbol{\theta}^{\mathrm{opt}}$ to the elements in $\mathcal{F}_3$ by a approximation projection, written as

$$\angle\theta_{r,j}^{\mathrm{sub}} = \underset{\phi\in\mathcal{F}_3}{\arg\min}\left|\angle\theta_{r,j}^{\mathrm{opt}} - \angle\phi\right|, \forall r\in\mathcal{R}, \forall j\in\mathcal{N}, \quad (50)$$

where $\theta_{r,j}^{\mathrm{sub}}$ denotes the approximated sub-optimal solution to the phase shift $\theta_{r,j}$. Thus, we can finally obtain the sub-optimal solution $\boldsymbol{\Theta}^{\mathrm{sub}} \in \mathcal{F}_3$ to the subproblem (35).

B. Algorithm Convergency

In the ideal case $\mathcal{F} = \mathcal{F}_1$ and general case $\mathcal{F} = \mathcal{F}_2$, the proposed joint precoding framework, has strict convergency,[5] since each step of the iteration, i.e., (15), (19), (24), (29) and (35), can be easily proved to be monotonous. However, in the non-ideal case $\mathcal{F} = \mathcal{F}_3$, the convergency of the proposed framework can not be proved strictly, since the update of $\boldsymbol{\Theta}$ has no guarantee of monotony in some cases due to the approximation operation in (50). Besides, for **Algorithm 2**, how the relaxation operation of optimizing the user-RIS matching vector $\mathbf{u}$ will influence the global convergency is also unpredictable [40]. Fortunately, since the rest of iterative steps are all monotonous, actually the loss caused by the approximation operation and relaxation operation have little adverse effect on the global convergence, which will be verified through simulations in the next Section VI.

C. Computational Complexity

The overall computational complexities of the proposed joint precoding framework and its two-timescale extension are mainly introduced by the updates of the variables $\boldsymbol{\rho}, \boldsymbol{\xi}, \mathbf{W}, \boldsymbol{\varpi}, \boldsymbol{\Theta}$, and $\mathbf{u}$, as shown in (15), (19), Appendix A, (29), Appendix B, and (47), respectively. To make it clear, we summarize their computational complexities in Tab. I, wherein I_{a} and I_{p} are the required iteration numbers for the convergence of the subproblems $\hat{\mathcal{P}}_{\mathrm{active}}$ in (24) and $\hat{\mathcal{P}}_{\mathrm{passive}}$ in (35), respectively.

From Table I, we can observe that the updates of $\boldsymbol{\rho}, \boldsymbol{\xi}$, and $\boldsymbol{\varpi}$ have much lower computational complexity than $\mathbf{W}, \boldsymbol{\Theta}$, and $\mathbf{u}$, since $\boldsymbol{\rho}, \boldsymbol{\xi}$, and $\boldsymbol{\varpi}$ have close-form update formula while $\mathbf{W}^{\mathrm{opt}}$, $\boldsymbol{\Theta}^{\mathrm{opt}}$, and $\mathbf{u}^{\mathrm{opt}}$ can only be obtained by solving QCQP and SDP problems via iterative optimization methods. Furthermore, since

---

[5]Note that, the proposed joint precoding scheme converges to a feasible solution, while only the convergence rate is fully analyzed can its local optimality be strictly proved.
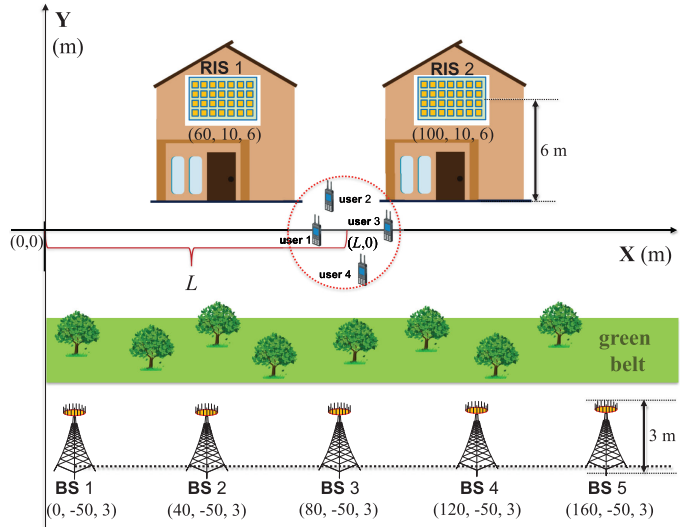

Fig. 5. The simulation scenario where five BSs assisted by two RISs serve four users.

the density of BSs and users in cell-free network is high [9] and the number of RIS elements is usually large [18], we can reasonably assume that $BMPK \gg U$ and $RN \gg U$. Under this assumption, the overall computational complexity of **Algorithm 1** can be approximated by $\mathcal{O}(I_{\mathrm{o}}(I_{\mathrm{a}}B^2M^2P^2K^2 + I_{\mathrm{p}}R^2N^2))$ and that of **Algorithm 2** can be similarly approximated by $\mathcal{O}(I_{\mathrm{o}}(I_{\mathrm{a}}B^2M^2P^2K^2 + I_{\mathrm{p}}R^2N^2 + K^2R^2(KR+1)^{2.5}))$, wherein I_{o} denotes the required number of iterations for the algorithm convergence.

VI. SIMULATION RESULTS

In this section, we provided extensive simulation results under different conditions to validate the performance of the proposed concept of RIS-aided cell-free network.

A. Simulation Setup

For the simulation of the proposed RIS-aided cell-free network, we draw on the potential deployment schemes for cell-free network from the existing works [8] to construct simulation scenarios. For simplicity but without loss of generality, we consider a 3-D scenario with the topology shown in Fig. 5. In this setup, a cell-free network with five BSs serve four users simultaneously, while the network capacity is limited due to the obstruction of a green belt. To improve the capacity, two RISs are separately deployed on two distant building surfaces, which are high enough to construct extra reflection links. We assume the i-th BS is located at $(40(i-1)\,\mathrm{m}, -50\,\mathrm{m}, 3\,\mathrm{m})$, and two RISs are located at $(60\,\mathrm{m}, 10\,\mathrm{m}, 6\,\mathrm{m})$ and $(100\,\mathrm{m}, 10\,\mathrm{m}, 6\,\mathrm{m})$, respectively. Since the BS in cell-free network is usually small with few antennas and low transmit power [8], we further assume the maximum transmit power at the BS is set as $P_{b,\max} = 0$ dBm, while the number of antennas at each BS is set as $M = 2$ and that at each user is set as $U = 2$. The number of RIS elements is set as $N = 100$. The number of subcarriers is set as $P = 4$, and the noise power is set as $\sigma^2 = -80$ dBm [32].

For the channel model, we consider the same settings as those in [32]. For the large-scale fading model, firstly we define d_{Bu}, d_{BR}, d_{Ru} as the distance between BS and user, BS and RIS, RIS and user, respectively. Thus the distance-dependent path loss model is given by

$$L(d) = C_0 \left(\frac{d}{d_0} \right)^{-\kappa}, d \in \{d_{\mathrm{Bu}}, d_{\mathrm{BR}}, d_{\mathrm{Ru}}\} \qquad (51)$$

where C_0 is the path loss at the reference distance $d_0 = 1\,\mathrm{m}$ and κ denotes the path loss exponent. Here we assume $C_0 = -30\,\mathrm{dB}$, and the path loss exponents of the BS-RIS link, RIS-user link, and BS-user link are set as $\kappa_{\mathrm{BR}} = 2.2$, $\kappa_{\mathrm{Ru}} = 2.8$, and $\kappa_{\mathrm{Bu}} = 3.5$ [32], respectively. To account for the small-scale fading, we further consider a Rician fading channel model, thus the BS-user channel $\mathbf{H}$ is obtained by

$$\mathbf{H} = \sqrt{\frac{\omega_{\mathrm{Bu}}}{1 + \omega_{\mathrm{Bu}}}} \mathbf{H}^{\mathrm{LoS}} + \sqrt{\frac{1}{1 + \omega_{\mathrm{Bu}}}} \mathbf{H}^{\mathrm{NLoS}}, \qquad (52)$$

where ω_{Bu} denotes the Rician factor, and $\mathbf{H}^{\mathrm{LoS}}$ and $\mathbf{H}^{\mathrm{NLoS}}$ denote the LoS and Rayleigh fading components, respectively. Note that $\mathbf{H}$ is equivalent to a LoS channel when $\omega_{\mathrm{Bu}} \to \infty$, and a Rayleigh fading channel when $\omega_{\mathrm{Bu}} = 0$. Then, $\mathbf{H}$ is multiplied by the square root of the distance-dependent path loss $L(d_{\mathrm{Bu}})$ in (51). Similarly, the BS-RIS and RIS-user channels can also be generated by the above procedure, and let ω_{BR} and ω_{Ru} denote the Rician factors of them, respectively. We further assume $\omega_{\mathrm{BR}} \to \infty$, $\omega_{\mathrm{Bu}} = 0$ and, $\omega_{\mathrm{Ru}} = 0$ [32].

At last, for the joint precoding framework, we set the weights of users as $\eta_k = 1$. As an alternating algorithm, $\mathbf{\Theta}$ is initialized by random values in $\mathcal{F}$, $\mathbf{W}$ is initialized by identical power and random phases, and $\mathbf{u}$ is initialized by setting all of its elements to one.

B. Weighted Sum-Rate of the RIS-Aided Cell-Free Network

The WSR of the proposed RIS-aided cell-free network is evaluated in this subsection. Firstly, we assume that four users are randomly distributed in a circle centered at $(L, 0)$ with radius $1\,\mathrm{m}$. The height of these users is set as $1.5\,\mathrm{m}$, and we also assume the ideal RIS case ($\mathcal{F} = \mathcal{F}_1$). Then, we plot the WSR against the distance L in Fig. 6, in which the five curves are defined as follows:

- *Ideal RIS case:* Based on the fully-known CSI, **Algorithm 1** is performed to maximize the WSR.
- *Two-timescale scheme:* This curve is realized in three steps. First, **Algorithm 2** is performed to match each user with no more than $R_{\mathrm{match}} = 1$ RISs. Then, only the RIS-aided channels of the matched user-RIS pairs and the direct-link channels are utilized for joint precoding design through **Algorithm 1**. Finally, WSR is obtained by substituting the optimized $\mathbf{W}^{\mathrm{opt}}$ and $\mathbf{\Theta}^{\mathrm{opt}}$ from **Algorithm 1** into R_{sum} in (9).
- *Without direct link:* Assume that all direct links between BSs and users are completely obstructed ($\mathbf{H} = 0$), and then **Algorithm 1** is performed to maximize the WSR.
- *Random phase shift:* All the phase shifts of RIS elements are randomly set to the values in $\mathcal{F}_1$. Then, based on the
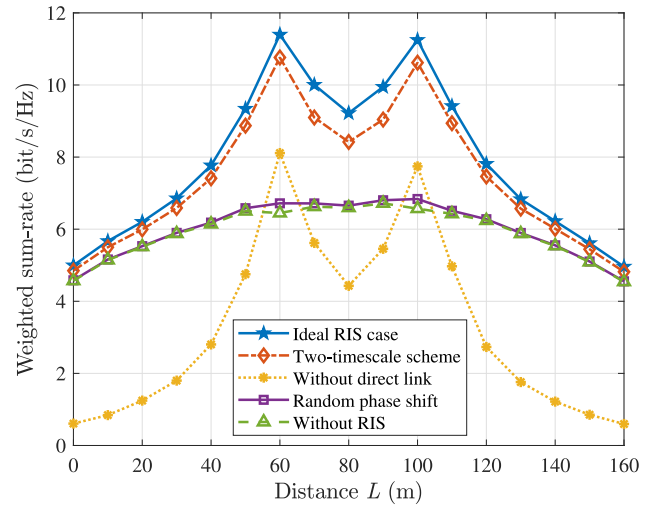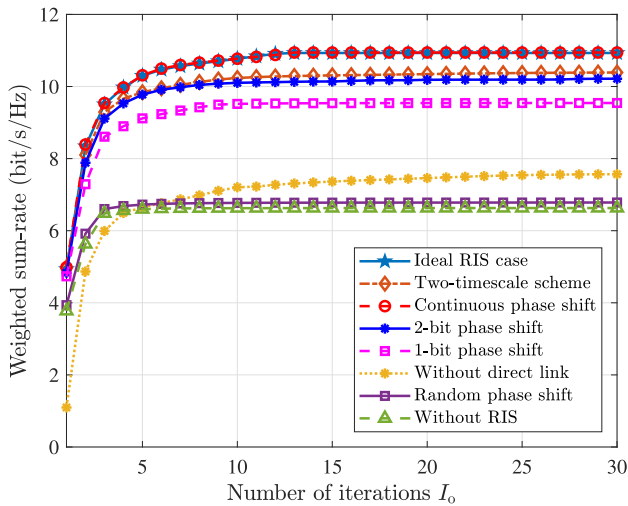

Fig. 6. Weighted sum-rate against the distance L.

combined channels, **Algorithm 1** is only performed at BSs to maximize the WSR.
- *Without RIS:* The conventional cell-free network without RIS. Based on the CSI of BS-user channels, the multi-user precoding method in [34] is performed at BSs to maximize the WSR.

From Fig. 6, we have three observations. First, for the schemes with RIS deployed, we can see two obvious peaks at $L = 60\,\mathrm{m}$ and $L = 100\,\mathrm{m}$. It indicates that the WSR rises when the users approach one of the two RISs, since the users can receive strong signals reflected from the RISs. While for the conventional scheme without RIS, these two peaks will not appear. Thus, we can conclude that the network capacity can be substantially increased by deploying RISs in the network, and the signal coverage can be accordingly extended. Second, we notice that the performance of "*Random phase shift*" has very limited gain compared with the scheme without RIS. The reason is that, without passive beamforming at RISs, the signals reaching RISs cannot be accurately directed to the users, which demonstrates the necessity of passive precoding. At last, we notice that, compared with "*Ideal RIS case*", "*Two-timescale scheme*" suffers an averaged performance loss of about 10%. The reason is that, "*Ideal RIS case*" acquires and utilizes all RIS-aided channels ($KR = 8$) for joint precoding design, while "*Two-timescale scheme*" only acquires and utilizes the RIS-aided channels of the matched user-RIS pairs ($KR_{\mathrm{match}} = 4$). In later Section VI-E, we will analyze the system WSR with different R_{match} in detail.

C. Convergency of the Joint Precoding Framework

To show the convergence of the proposed algorithms, we plot the WSR against the number of iterations I_o in Fig. 7 by running algorithms once with random initializations. The simulation setups are same as those used in Subsection VI-B, and we fix the distance L as $L = 65\,\mathrm{m}$. To evaluate the convergency in non-ideal RIS phase shift cases, we add the curves "*Continuous phase shift*", "*1-bit phase shift*", and "*2-bit phase shift*" to denote the cases $\mathcal{F}_2$ and $\mathcal{F}_3$. The results in Fig. 7 illustrate that, when the convergence error is no more than 1%, the proposed joint

Fig. 7. Weighted sum-rate against the number of iterations I_o.

Fig. 8. Weighted sum-rate against the CSI error parameter δ.

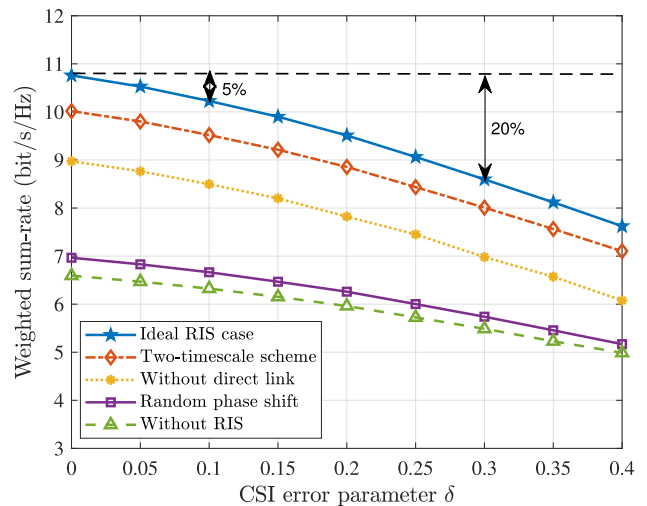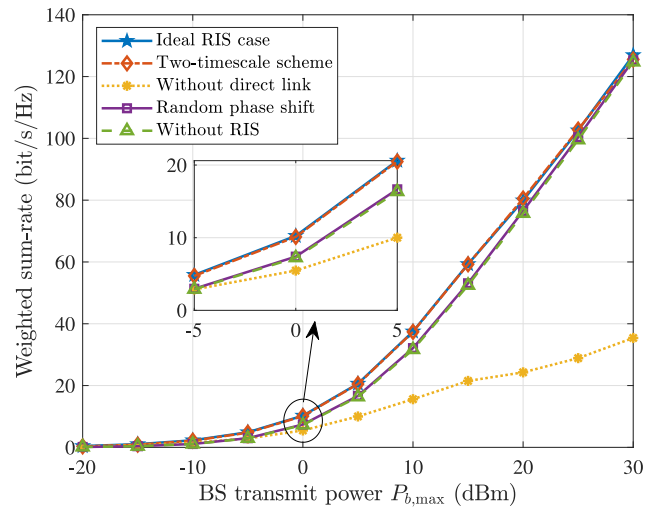Fig. 9. Weighted sum-rate against the BS transmit power $P_{b,\max}$.

precoding framework can converge within 15 iterations. To be more specific, the ideal RIS case, two-timescale scheme, continuous phase shift case, and the case without direct links can converge within 15 iterations, while the discrete phase shift case can converge within 10 iterations. Since the conventional cell-free network without RIS and the scheme "*Random phase shift*" do not need to address the RIS precoding, the WSR in these cases can converge within 5 iterations. The results also indicate that, although the approximation operation in (50) causes the uncertainty of convergence, the proposed framework still enjoys a fast global convergence. More importantly, we find that the curves "*Ideal RIS case*" and "*Continuous phase shift*" are very close. The essential reason is that, for "Ideal RIS case", actually most of the optimized RIS elements have an amplitude equal to 1. It indicates that, for the precoding design of the amplitude-uncontrollable RIS, we can simply relax its constant-modulus constraint and optimize its precoding as an ideal RIS. Then, we can directly project the optimized precoding matrix to the constant-modulus and obtain the final design. In this way, the optimization of RIS precoding becomes simple and low-complexity. Besides, it also implies that there is no need to design an amplitude-controllable RIS for wireless communications.

D. Robustness of the Joint Precoding to CSI Error

Generally, in the proposed RIS-aided cell-free system, the channel estimation is very challenging due to the high-dimensional channels. Here we analyze the robustness of the proposed joint precoding scheme to CSI error. To generate the imperfect channels, we model the practically estimated channel $\hat{h}$ as [45]

$$\hat{h} = h + e, \tag{53}$$

where h denotes the real channel and e represents the estimation error with Gaussian distribution and zero mean, i.e. $e \sim \mathcal{CN}(0, \sigma_e^2)$. We assume the variance σ_e^2, i.e. the error power, satisfies $\sigma_e^2 \triangleq \delta |h|^2$ wherein δ denotes the ratio of the error power σ_e^2 to the channel gain $|h|^2$, which characterizes the level

of CSI error. Then, using the same setups in Subsection VI-C, we plot the WSR against the CSI error parameter δ in Fig. 8. From this figure, we can observe that, the performance loss grows with the increasing of δ. Particularly, for the "*Ideal RIS case*", compared with the perfect CSI without error (i.e. $\delta = 0$), the system performance suffers a loss of 5% when the error power is 10% of the channel gain (i.e. $\delta = 0.1$), and a loss of 20% when $\delta = 0.3$. Thereby, the proposed joint precoding scheme shows strong robustness to CSI error.

E. The Impact of Key System Parameters

To reveal more insights of the proposed RIS-aided cell-free network with different system parameters, we consider a new simulation scenario with more dispersed users and reset the locations of the four users in Fig. 5 as $(L - 60\,\mathrm{m}, 0)$, $(L - 20\,\mathrm{m}, 0)$, $(L + 20\,\mathrm{m}, 0)$, and $(L + 60\,\mathrm{m}, 0)$, respectively.

1) WSR Against BS Transmit Power: By fixing $L = 80\,\mathrm{m}$, we present the average weighted sum-rate against the BS transmit power in Fig. 9. From this figure we can observe that, with the increase of the BS transmit power, the weighted sum-rate
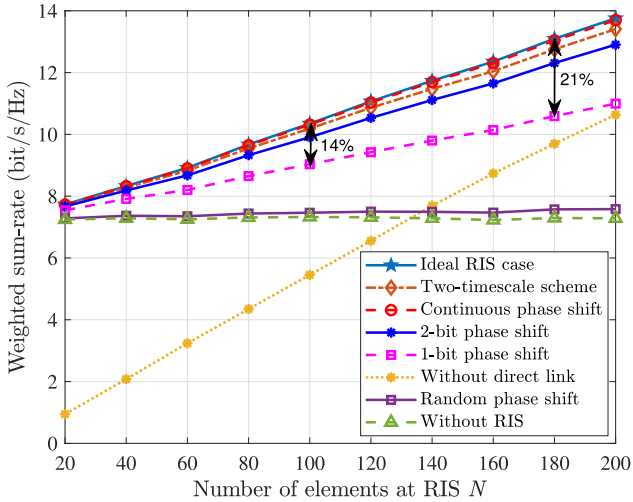
Fig. 10. Weighted sum-rate against the number of RIS elements.

Fig. 11. Weighted sum-rate against the distance L with different allowable matched user-RIS pairs.

rises rapidly in all cases. Particularly, "*Two-timescale scheme*" achieves very similar performance to "*Ideal RIS case*". It indicates that, the RIS-aided channels of the unmatched user-RIS pairs indeed have negligible contribution to capacity improvement, which demonstrates the effectiveness of the proposed two-timescale joint precoding scheme. Besides, we have an important finding that, the performance gain brought by RISs is significant only when the BS transmit power is moderate (e.g., form 0 dBm to 20 dBm), while the performance gain becomes negligible when the BS transmit power is too low (e.g. -20 dBm) or too high (e.g. 30 dBm). We explain this phenomenon in detail as follows.

Intuitively, when the BS transmit power is too low, the reflected signals by RIS are so weak that RISs have little contribution to performance improvement. While, when the BS transmit power is too high, the BSs prefer to allocate most of the power to the beams towards the BS-user links rather than the beams towards the BS-RIS-user links, which makes the role of RISs less obvious. To be specific, as we can observe from Fig. 9, "*Without direct link*" can be roughly regarded as the scheme when the BSs allocate all power to the BS-RIS-user links, while "*Without RIS*" denotes the scheme when the BSs allocate all power to the BS-user links [32]. With the increasing of BS transmit power, the performance gap between these two schemes becomes larger and larger. It indicates that, compared with the first scheme (i.e. BSs allocate all power to the BS-user links), the benefit of applying the second scheme (i.e. BSs allocate all power to the BS-RIS-user links) is relatively lower and lower. When the BS transmit power is large enough, the performance can obtain much more gain from BS allocating power to BS-user link than to BS-RIS-user link, since the performance gap between them is too large. In this case, the BS will tend to allocate most of its power to the beam towards the BS-user link, which weakens the role of RIS.

2) WSR Against Number of RIS Elements: Using the same setups as above, we present the average weighted sum-rate against the number of RIS elements in Fig. 10. We can observe that, the WSR of the proposed RIS-aided cell-free network increases
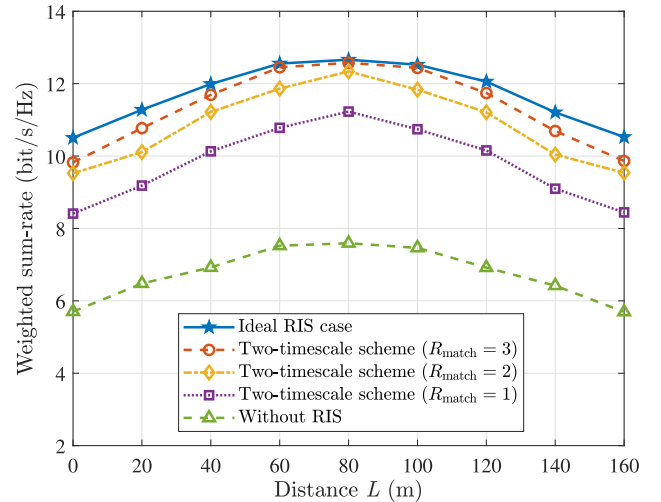
as the number of RIS elements rises. More importantly, we find that, with the increasing of RIS elements, the approximation loss of the low-resolution phase shift becomes larger. For instance, when $N = 100$, the approximation loss in the case "*1-bit phase shift*" is about 14% of the ideal case, and this loss grows to 21% when $N = 180$. The observation implies that, when the number of RIS is large, is is necessary to use more accurate phase shifts for passive precoding, so that the signals reflected from RISs can reach the users more precisely. In addition, since the dimension of $\mathbf{\Theta}$ is RN, the complexity of solving (35) grows with the number of RIS elements N, and too more RIS elements will also make the channel estimations more challenging. Thus, it is essential to choose the RIS element number reasonably.

3) WSR With Different Allowable Matched user-RIS Pairs: To evaluate the performance of the proposed two-timescale extension of the joint precoding scheme, in this part, we analyze the system WSR with different R_{match}, which denotes the maximum number of RISs that each user can be matched with. To show more insights, in the same simulation scenario shown in Fig. 5, we reset the total number of RISs as $R = 7$ and the RIS elements as $N = 50$. We further assume that the i-th RIS is located at $(20 \times i \,\mathrm{m}, 10 \,\mathrm{m}, 6 \,\mathrm{m})$. Then, we plot the WSR against the distance L with different R_{match} in Fig. 11.

From this figure, we can observe that, the WSR increases as each user can be matched with more RISs. Compared with the ideal joint precoding design which acquires and utilizes all $KR = 28$ RIS-aided channels, the proposed two-timescale scheme suffers performance but consumes much less overhead for CSI acquisitions. For example, when $R_{\mathrm{match}} = 3$, i.e. each user can be matched with at most 3 RISs, compared with the perfect joint precoding design to some extent, the proposed two-timescale scheme acquires and utilizes no more than $KR_{\mathrm{match}} = 12$ RIS-aided channels for joint precoding design but only suffers an averaged capacity loss of about 5%. Particularly, at the positions from $L = 60 \,\mathrm{m}$ to $L = 100 \,\mathrm{m}$, the two-timescale scheme ($R_{\mathrm{match}} = 3$) can almost achieve the same performance as the ideal joint precoding design. In this
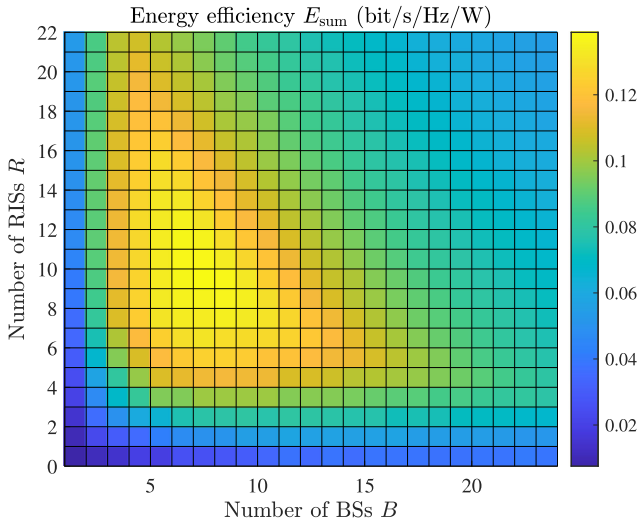
Fig. 12.   System energy efficiency against BS number B and RIS number R.

way, the proposed two-timescale scheme consumes much less overhead for acquiring RIS-aided channels from the long-term perspective [22]. Thus, we can conclude that, the proposed two-timescale scheme can serve as a efficient scheme for balancing the performance and overhead for CSI acquisitions, which is especially suitable to employ in the cell-free networks with a large number of distributed RISs.

F. Trade-Off Between the Number of BSs and RISs

Since deploying more RISs can improve the network capacity with low cost and power consumption, in this subsection we analyze the trade-off between the number of BSs and RISs.

To fairly evaluate the losses and gains, here we take the energy efficiency as the performance metric. According to the power consumption model developed in [15], the system energy efficiency is given by

$$E_{\text{sum}} = \frac{R_{\text{sum}}}{\tau_{\text{r}}\|\mathbf{W}\|^2 + BP_{\text{BS}} + KP_{\text{UE}} + RNP_{\text{RIS}}}, \quad (54)$$

where τ_{r}^{-1} is the efficiency of the transmit power amplifier, while P_{BS}, P_{UE}, and P_{RIS} denote the hardware power consumption at each BS, each user, and each RIS element, respectively.

For the simulation setups, we consider the same settings as those in [15], i.e. $\tau_{\text{r}} = 1.2$, $P_{\text{BS}} = 9$ dBW, $P_{\text{UE}} = 10$ dBm and $P_{\text{RIS}} = 10$ dBm. To find more essential insights, we reset some parameters as $P = M = U = 1$ and $N = 20$. We fix the distance-dependent effects by setting the related distances as $d_{\text{Bu}} = d_{\text{BR}} = 110$ m and $d_{\text{Ru}} = 15$ m. By performing **Algorithm 1** in ideal RIS case, we plot the energy efficiency against the number of BSs B and that of RISs R in Fig. 12. From this figure, we can observe that, given B, the energy efficiency can be significantly improved by increasing the number of RISs R. Particularly, when $B = 7$ and $R = 9$, the maximal energy efficiency can be achieved, i.e. $E_{\text{sum}} = 0.138$ bit/s/Hz/W. However, when R is too large, the energy efficiency will decline after deploying more RISs. For example, given $B = 7$, the system energy efficiency decreases to $E_{\text{sum}} = 0.09$ bit/s/Hz/W when

$R = 21$, which is about a 34.8% loss. The reason is that, when the network capacity is large enough, the capacity improvement via deploying more RISs cannot make up for the extra power consumption. Therefore, carefully designing the trade-off between the number of BSs and RISs is the guarantee for balancing the capacity and power consumption in practical systems.

VII. CONCLUSION

In this paper, we first propose the concept of RIS-aided cell-free network, which aims to improve the network capacity with low cost and power consumption. Then, for the proposed RIS-aided cell-free network, in a typical wideband scenario, we formulate the joint precoding design problem to maximize the weighted sum-rate to optimize the network capacity, subject to the transmit power constraint of BS and the phase shift constraint of RIS. Finally, we propose a joint precoding framework to solve this problem. Since most of the considered scenarios in existing works are special cases of the general scenario in this paper, the proposed joint precoding framework can also serve as a general solution to maximize the capacity in most of existing RIS-aided scenarios. Simulation results demonstrate that, with the assist of low-cost and energy-efficient RISs, the proposed RIS-aided cell-free network can realize higher capacity than the conventional cell-free network.

Compared with the traditional cell-free network, the proposed RIS-aided cell-free network can achieve wider signal coverage [13], higher spectrum [32], and higher energy efficiency [15]. For future works, some open problems are still worth further investigations. For example, some other important performance metrics such as the energy efficiency [15], BS transmit power [17], and user fairness [23] are left for future works. Besides, in some specific scenarios such as the internet of vehicles, how to select the replaced BSs and optimize the placements of RISs is also an interesting research direction.

Another promising technique benefiting from the advance of meta-materials is large intelligent surface (LIS) [46]–[48]. Different from the passive RIS, LIS is considered as a quasi-continuous surface [46] which can actively achieve surface beamforming [47]. Moreover, the signal model of LIS is in continuous-integral form [46]–[48]. The introduction of LIS into cell-free network will bring many open problems, such as how to bridge the gap between discrete model and continuous model, which are also left for future works.

APPENDIX A
PDS-BASED METHOD FOR SOLVING SUBPROBLEM (24)

We consider to solve (24) via the PDS method [37]. Firstly, by introducing Lagrange multipliers $\boldsymbol{\lambda} = [\lambda_1, \ldots, \lambda_B]^T \in \mathbb{R}^B$ and $\vartheta > 0$, the Lagrangian for the augmented problem of (24) can be written as

$$\Gamma_1(\mathbf{W}, \boldsymbol{\lambda}) = g_3(\mathbf{W}) + \boldsymbol{\lambda}^T \mathbf{F}(\mathbf{W}) + \frac{\vartheta}{2}\|\mathbf{F}(\mathbf{W})\|^2. \quad (55)$$

where the function vector $\mathbf{F}(\mathbf{W}) \in \mathbb{R}^B$ is defined by $\mathbf{F}(\mathbf{W}) = [F_1^+(\mathbf{W}), \ldots, F_B^+(\mathbf{W})]^T$ with $F_b(\mathbf{W})$ being $F_b(\mathbf{W}) = \mathbf{W}^H \mathbf{D}_b \mathbf{W} - P_{b,\max}$.

Next, we introduce a superscript t to each variable as the iteration index, and we have the iterative formulas of $\mathbf{W}^t$ and $\boldsymbol{\lambda}^t$ written as

$$
\begin{bmatrix} \mathbf{W}^{t+1} \\ \boldsymbol{\lambda}^{t+1} \end{bmatrix} = \begin{bmatrix} \mathbf{W}^t \\ \boldsymbol{\lambda}^t \end{bmatrix} - \alpha^t \begin{bmatrix} \partial\Gamma_1/\partial\mathbf{W}|_{\mathbf{W}=\mathbf{W}^t} \\ -\partial\Gamma_1/\partial\boldsymbol{\lambda}|_{\boldsymbol{\lambda}=\boldsymbol{\lambda}^t} \end{bmatrix} \quad (56)
$$

where the step length $\alpha^t > 0$ is a sufficiently small positive value and the iteration directions $\partial\Gamma_1/\partial\mathbf{W}$ and $\partial\Gamma_1/\partial\lambda$ are given by

$$
\frac{\partial\Gamma_1}{\partial\mathbf{W}} = (\mathbf{A}\mathbf{W})^* - \mathbf{V}^* + \sum_{b=1}^{B}\left(\lambda_b + \vartheta F_b^+(\mathbf{W})\right)\boldsymbol{\Pi}_b^*(\mathbf{W}), \quad (57a)
$$

$$
\frac{\partial\Gamma_1}{\partial\lambda} = \mathbf{F}(\mathbf{W}), \quad (57b)
$$

wherein the auxiliary function $\boldsymbol{\Pi}_b(\mathbf{W}) \in \mathbb{C}^{BMPK}$ is denoted by

$$
\boldsymbol{\Pi}_b(\mathbf{W}) = \begin{cases} \mathbf{D}_b\mathbf{W}, & F_b(\mathbf{W}) > 0 \\ \mathbf{0}_{BMPK}, & F_b(\mathbf{W}) \leq 0 \end{cases}, \forall b \in \mathcal{B}. \quad (58)
$$

By simultaneously optimizing $\mathbf{W}$ and λ via (56) until convergence, the optimal $\mathbf{W}^{\text{opt}}$ can be finally obtained without need of the inversion operation for the high-dimensional matrix $\mathbf{A}$.

Remark 1: We discuss the computational complexity of solving (24). Since $\mathbf{D}_b$ is a fixed diagonal matrix with MPK elements being 1, we can simply derive that the complexity of computing $F_b^+(\mathbf{W})$ is $\mathcal{O}(MPK)$. Then, we can further obtain that the computational complexity of updating $\mathbf{W}$ and λ via (56) are $\mathcal{O}(B^2M^2P^2K^2)$ and $\mathcal{O}(BMPK)$, respectively. Let I_a denote the required iteration number for convergence. Then the overall complexity of the applied PDS-based method is about $\mathcal{O}(I_a(B^2M^2P^2K^2 + BMPK))$.

APPENDIX B
PDS-BASED METHOD FOR SOLVING SUBPROBLEM (35)

We consider to solve (35) via the PDS method [37]. Similar to Appendix A, by introducing Lagrange multipliers $\boldsymbol{\chi} = [\chi_1, \ldots, \chi_{RN}]^T \in \mathbb{R}^{RN}$ and $\varphi > 0$, the Lagrangian for the augmented problem of (35) can be derived as

$$
\Gamma_2(\boldsymbol{\Theta}, \boldsymbol{\chi}) = g_6(\boldsymbol{\Theta}) + \boldsymbol{\chi}^T\mathbf{G}(\boldsymbol{\Theta}) + \frac{\varphi}{2}\|\mathbf{G}(\boldsymbol{\Theta})\|^2, \quad (59)
$$

where the function vector $\mathbf{G}(\boldsymbol{\Theta})$ is defined by $\mathbf{G}(\boldsymbol{\Theta}) = [G_1^+(\boldsymbol{\Theta}), \ldots, G_{RN}^+(\boldsymbol{\Theta})]^T$ with $G_j(\boldsymbol{\Theta}) = \boldsymbol{\theta}^H\mathbf{E}_j\boldsymbol{\theta} - 1$ and $\mathbf{E}_j \triangleq \mathbf{e}_j\mathbf{e}_j^H$. Then, we have the iterative formulas of $\boldsymbol{\theta}^t$ and $\boldsymbol{\chi}^t$ written as

$$
\begin{bmatrix} \boldsymbol{\theta}^{t+1} \\ \boldsymbol{\chi}^{t+1} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\theta}^t \\ \boldsymbol{\chi}^t \end{bmatrix} - \beta^t \begin{bmatrix} \partial\Gamma_2/\partial\boldsymbol{\theta}|_{\boldsymbol{\theta}=\boldsymbol{\theta}^t} \\ -\partial\Gamma_2/\partial\boldsymbol{\chi}|_{\boldsymbol{\chi}=\boldsymbol{\chi}^t} \end{bmatrix} \quad (60)
$$

wherein $\beta^t > 0$ denotes the sufficiently small step length and the iteration directions $\partial\Gamma_2/\partial\boldsymbol{\theta}$ and $\partial\Gamma_2/\partial\boldsymbol{\chi}$ are obtained by

$$
\frac{\partial\Gamma_2}{\partial\boldsymbol{\theta}} = (\boldsymbol{\Lambda}\boldsymbol{\theta})^* - \boldsymbol{\nu}^* + \sum_{j=1}^{RN}\left(\chi_j + \varphi G_j^+(\boldsymbol{\Theta})\right)\boldsymbol{\Psi}_j^*(\boldsymbol{\theta}), \quad (61a)
$$

$$
\frac{\partial\Gamma_2}{\partial\boldsymbol{\chi}} = \mathbf{G}(\boldsymbol{\Theta}). \quad (61b)
$$

Note that, the auxiliary function $\boldsymbol{\Psi}_j(\boldsymbol{\theta}) \in \mathbb{C}^{RN}$ is defined as

$$
\boldsymbol{\Psi}_j(\boldsymbol{\theta}) = \begin{cases} \mathbf{E}_j\boldsymbol{\theta}, & \boldsymbol{\theta}^H\mathbf{E}_j\boldsymbol{\theta} > 0 \\ \mathbf{0}_{RN}, & \boldsymbol{\theta}^H\mathbf{E}_j\boldsymbol{\theta} \leq 0 \end{cases}, \forall j \in \{1, \ldots, RN\} \quad (62)
$$

By simultaneously optimizing $\boldsymbol{\theta}$ and $\boldsymbol{\chi}$ via (60) until convergence, the optimal $\boldsymbol{\Theta}^{\text{opt}}$ can be finally obtained without need of the inversion operation for the high-dimensional matrix $\boldsymbol{\Lambda}$.

Remark 2: We study the computational complexity of solving (35). Similar to **Remark 1**, we can easily derive that the complexity of updating $\boldsymbol{\Theta}$ and $\boldsymbol{\chi}$ by (60) are $\mathcal{O}(R^2N^2)$ and $\mathcal{O}(RN)$, respectively. Therefore, the overall computational complexity is $\mathcal{O}(I_p(R^2N^2 + RN))$ where I_p denotes the required iteration number for convergence.

REFERENCES

[1] Z. Zhang and L. Dai, "Capacity improvement in wideband reconfigurable intelligent surface-aided cell-free network," in *Proc. IEEE 21st Int. Workshop Signal Process. Adv. Wireless Commun.*, 2020, pp. 1–5.

[2] D. López-Pérez, M. Ding, H. Claussen, and A. H. Jafari, "Towards 1 Gbps/UE in cellular systems: Understanding ultra-dense small cell deployments," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 4, pp. 2078–2101, Oct.-Dec. 2015.

[3] F. E. Idachaba, "5 G networks: Open network architecture and densification strategies for beyond 1000x network capacity increase," in *Proc. Future Technol. Conf.*, 2016, pp. 1265–1269.

[4] M. Kamel, W. Hamouda, and A. Youssef, "Ultra-dense networks: A survey," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 4, pp. 2522–2545, Oct.-Dec. 2016.

[5] J. G. Andrews, X. Zhang, G. D. Durgin, and A. K. Gupta, "Are we approaching the fundamental limits of wireless network densification?," *IEEE Commun. Mag.*, vol. 54, no. 10, pp. 1558–1896, Oct. 2016.

[6] A. Lozano, R. W. Heath, and J. G. Andrews, "Fundamental limits of cooperation," *IEEE Trans. Inf. Theory*, vol. 59, no. 9, pp. 5213–5226, Sep. 2013.

[7] E. Nayebi, A. Ashikhmin, T. L. Marzetta, and H. Yang, "Cell-free massive MIMO systems," in *Proc. 49th Asilomar Conf. Signals Syst. Comput.*, 2015, pp. 695–699.

[8] G. Interdonato, E. Björnson, H. Q. Ngo, P. Frenger, and E. G. Larsson, "Ubiquitous cell-free massive MIMO communications," *EURASIP J. Wireless Commun. Netw.*, pp. 197–206, 2019.

[9] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, "Cell-free massive MIMO versus small cells," *IEEE Trans. Wireless Commun.*, vol. 16, no. 3, pp. 1834–1850, Mar. 2017.

[10] S. Mosleh, H. Almosa, E. Perrins, and L. Liu, "Downlink resource allocation in cell-free massive MIMO systems," in *Proc. Int. Conf. Comput. Netw. Commun.*, 2019, pp. 883–887.

[11] M. Attarifar, A. Abbasfar, and A. Lozano, "Modified conjugate beamforming for cell-free massive MIMO," *IEEE Commun. Lett.*, vol. 8, no. 2, pp. 616–619, Apr. 2019.

[12] Y. Jin, J. Zhang, S. Jin, and B. Ai, "Channel estimation for cell-free mmwave massive MIMO through deep learning," *IEEE Trans. Veh. Technol.*, vol. 68, no. 10, pp. 10325–10329, Oct. 2019.

[13] Y. Liang, R. Long, Q. Zhang, J. Chen, H. V. Cheng, and H. Guo, "Large intelligent surface/antennas (LISA): Making reflective radios smart," *J. Commun. Inf. Netw.*, vol. 4, no. 2, pp. 40–50, Jun. 2019.

[14] E. Basar, M. Di Renzo, J. De Rosny, M. Debbah, M. Alouini, and R. Zhang, "Wireless communications through reconfigurable intelligent surfaces," *IEEE Access*, vol. 7, pp. 116753–116773, Aug. 2019.

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

[16] M. Di Renzo *et al.*, "Reconfigurable intelligent surfaces vs. relaying: Differences, similarities, and performance comparison," *IEEE Open J. Commun. Soc.*, vol. 1, pp. 798–807, Jun. 2020.

[17] Q. Wu and R. Zhang, "Beamforming optimization for wireless network aided by intelligent reflecting surface with discrete phase shifts," *IEEE Trans. Commun.*, vol. 68, no. 3, pp. 1838–1851, Mar. 2020.

[18] P. Wang, J. Fang, X. Yuan, Z. Chen, and H. Li, "Intelligent reflecting surface-assisted millimeter wave communications: Joint active and passive precoding design," *IEEE Trans. Veh. Technol.*, vol. 69, no. 12, pp. 14 960–14 973, Dec. 2020.

[19] C. Pan *et al.*, "Multicell MIMO communications relying on intelligent reflecting surfaces," *IEEE Trans. Wireless Commun.*, vol. 19, no. 8, pp. 5218–5233, Aug. 2020.

[20] C. Huang, C. Zhang, J. Yang, B. Sun, B. Zhao, and X. Luo, "Reconfigurable metasurface for multifunctional control of electromagnetic waves," *Adv. Opt. Mater.*, vol. 5, no. 22, p. 1700485, Sep. 2017.

[21] Ö. Özdogan, E. Björnson, and E. G. Larsson, "Intelligent reflecting surfaces: Physics, propagation, and pathloss modeling," *IEEE Wireless Commun. Lett.*, vol. 9, no. 5, pp. 581–585, May 2020.

[22] C. Hu, L. Dai, S. Han, and X. Wang, "Two-timescale channel estimation for reconfigurable intelligent surface aided wireless communications," *IEEE Trans. Commun.*, to be published, doi: 10.1109/TCOMM.2021.3072729.

[23] Q.-U.-A. Nadeem, A. Kammoun, A. Chaaban, M. Debbah, and M.-S. Alouini, "Asymptotic max-min SINR analysis of reconfigurable intelligent surface assisted MISO systems," *IEEE Trans. Commun.*, vol. 19, no. 12, pp. 7748–7764, Dec. 2020.

[24] L. Dai *et al.*, "Reconfigurable intelligent surface-based wireless communications: Antenna design, prototyping, and experimental results," *IEEE Access*, vol. 8, pp. 45913–45923, Mar. 2020.

[25] Q. Wu and R. Zhang, "Intelligent reflecting surface enhanced wireless network: Joint active and passive beamforming design," in *Proc. IEEE Global Commun. Conf.*, 2018, pp. 1–6.

[26] C. Huang, A. Zappone, M. Debbah, and C. Yuen, "Achievable rate maximization by passive intelligent mirrors," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, 2018, pp. 3714–3718.

[27] B. Di, H. Zhang, L. Song, Y. Li, Z. Han, and H. V. Poor, "Hybrid beamforming for reconfigurable intelligent surface based multi-user communications: Achievable rates with limited discrete phase shifts," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 8, pp. 1809–1822, Aug. 2020.

[28] H. Li, R. Liu, M. Liy, Q. Liu, and X. Li, "IRS-enhanced wideband MU-MISO-OFDM communication systems," in *Proc. IEEE Wireless Commun. Netw. Conf.*, 2020, pp. 1–6.

[29] Y. Yang, S. Zhang, and R. Zhang, "IRS-enhanced OFDM: Power allocation and passive array optimization," in *Proc. IEEE Global Commun. Conf.*, 2019, pp. 1–6.

[30] U. Siddique, H. Tabassum, and E. Hossain, "Downlink spectrum allocation for in-band and out-band wireless backhauling of full-duplex small cells," *IEEE Trans. Commun.*, vol. 65, no. 8, pp. 3538–3554, Aug. 2017.

[31] Y. Han, S. Zhang, L. Duan, and R. Zhang, "Cooperative double-IRS aided communication: Beamforming design and power scaling," *IEEE Wireless Commun. Lett.*, vol. 9, no. 8, pp. 1206–1210, Aug. 2020.

[32] Q. Wu and R. Zhang, "Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming," *IEEE Trans. Wireless Commun.*, vol. 18, no. 11, pp. 5394–5409, Nov. 2019.

[33] Q. Zheng *et al.*, "Ultra-wideband side-lobe level suppression using amplitude-adjustable metasurfaces," *J. Phys. D: Appl. Phys.*, vol. 52, no. 6, Dec. 2018, Art. no. 065102.

[34] K. Shen and W. Yu, "Fractional programming for communication systems—Part I: Power control and beamforming," *IEEE Trans. Signal Process.*, vol. 66, no. 10, pp. 2616–2630, May 2018.

[35] W. Dinklebach, "On nonlinear fractional programming," *Manage. Sci.*, vol. 13, no. 7, pp. 492–498, Oct. 2014.

[36] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "*Distributed optimization and statistical learning via the alternating direction method of multipliers*," pp. 1–122, Nov. 2014. [Online]. Available: https://stanford.edu/ boyd/papers/pdf/admm_distr_stats.pdf

[37] S. Boyd, "*Subgradient methods*," May 2014. [Online]. Available: https://web.stanford.edu/class/ee364b/lectures/subgrad_method_notes.pdf

[38] H. Guo, Y. Liang, J. Chen, and E. G. Larsson, "Weighted sum-rate maximization for reconfigurable intelligent surface aided wireless networks," *IEEE Trans. Wireless Commun.*, vol. 19, no. 5, pp. 3064–3076, May 2020.

[39] K. G. Murty and S. N. Kabadi, "Some NP-complete problems in quadratic and nonlinear programming," *Math. Program.*, vol. 39, pp. 117–129, Jun. 1987.

[40] T. Nie, S.-C. Fang, Z. Deng, and J. E. Lavery, "On linear conic relaxation of discrete quadratic programs," *Optim. Methods Softw.*, vol. 31, pp. 737–754, Jan. 2016.

[41] K. C. Toh, M. J. Todd, and R. H. Tütüncü, "SDPT3 - a matlab software package for semidefinite programming, version 1.3," *Optimiz. Methods Softw.*, vol. 11, no. 1-4, pp. 545–581, Jan. 2008.

[42] Z. Yang, L. Xie, and C. Zhang, "A discretization-free sparse and parametric approach for linear array signal processing," *IEEE Trans. Signal Process.*, vol. 62, no. 19, pp. 4959–4973, Oct. 2014.

[43] Y. Sun, P. Babu, and D. P. Palomar, "Majorization-minimization algorithms in signal processing, communications, and machine learning," *IEEE Trans. Signal Process.*, vol. 65, no. 3, pp. 794–816, Aug. 2017.

[44] K. Alhujaili, V. Monga, and M. Rangaswamy, "Transmit MIMO radar beampattern design via optimization on the complex circle manifold," *IEEE Trans. Signal Process.*, vol. 67, no. 13, pp. 3561–3575, May 2019.

[45] P. Ubaidulla and A. Chockalingam, "Relay precoder optimization in MIMO-relay networks with imperfect CSI," *IEEE Trans. Signal Process.*, vol. 59, no. 11, pp. 5473–5484, Aug. 2011.

[46] C. Huang *et al.*, "Holographic MIMO surfaces for 6G wireless networks: Opportunities, challenges, and trends," *IEEE Trans. Wireless Commun.*, vol. 27, no. 5, pp. 118–125, Oct. 2020.

[47] D. Dardari, "Communicating with large intelligent surfaces: Fundamental limits and models," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2526–2537, Nov. 2020.

[48] J. Yuan, H. Q. Ngo, and M. Matthaiou, "Towards large intelligent surface (LIS)-based communications," *IEEE Trans. Commun.*, vol. 68, no. 10, pp. 6568–6582, Oct. 2020.

Zijian Zhang received the B.E. degree in 2020 in electronic engineering from Tsinghua University, Beijing, China, where he is currently working toward the Ph.D. degree in electronic engineering. His research interests include physical-layer algorithms for massive MIMO and reconfigurable intelligent surfaces. He was the recipient of the National Scholarship in 2019 and the Excellent Thesis Award of Tsinghua University in 2020.

Linglong Dai received the B.S. degree from Zhejiang University, Hangzhou, China, in 2003, the M.S. degree (with the highest honor) from the China Academy of Telecommunications Technology, Beijing, China, in 2006, and the Ph.D. degree (with the highest honor) from Tsinghua University, Beijing, China, in 2011. From 2011 to 2013, he was a Postdoctoral Research Fellow with the Department of Electronic Engineering, Tsinghua University, where he was an Assistant Professor from 2013 to 2016 and has been an Associate Professor since 2016.

He has coauthored the book *MmWave Massive MIMO: A Paradigm for 5G* (Academic Press, 2016). He has authored or coauthored more than 60 IEEE journal papers and more than 40 IEEE conference papers. He also holds 19 granted patents. His current research interests include reconfigurable intelligent surface, massive MIMO, millimeter-wave or Terahertz communications, and machine learning for wireless communications. He was the recipient of the five IEEE Best Paper Awards at the IEEE ICC 2013, the IEEE ICC 2014, the IEEE ICC 2017, the IEEE VTC 2017-Fall, and the IEEE ICC 2018. He was also the recipient of the Tsinghua University Outstanding Ph.D. Graduate Award in 2011, the Beijing Excellent Doctoral Dissertation Award in 2012, the China National Excellent Doctoral Dissertation Nomination Award in 2013, the URSI Young Scientist Award in 2014, the IEEE Transactions on Broadcasting Best Paper Award in 2015, the Electronics Letters Best Paper Award in 2016, the National Natural Science Foundation of China for Outstanding Young Scholars in 2017, the IEEE ComSoc Asia-Pacific Outstanding Young Researcher Award in 2017, the IEEE ComSoc Asia-Pacific Outstanding Paper Award in 2018, the China Communications Best Paper Award in 2019, the IEEE Access Best Multimedia Award in 2020, and the IEEE Communications Society Leonard G. Abraham Prize in 2020. He was listed as a Highly Cited Researcher by Clarivate Analytics in 2020.

He is an Area Editor of the IEEE COMMUNICATIONS LETTERS, and the Editor of the IEEE TRANSACTIONS ON COMMUNICATIONS and IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY. Particularly, he is dedicated to reproducible research and has made a large amount of simulation codes publicly available.