Two-Timescale Transmission Design for RIS-Aided Cell-Free Massive MIMO Systems

Jianxin Dai\textsuperscript{a}, Member, IEEE, Jin Ge\textsuperscript{b}, Kangda Zhi\textsuperscript{b}, Cunhua Pan\textsuperscript{a}, Senior Member, IEEE, Zaichen Zhang\textsuperscript{a}, Senior Member, IEEE, Jiangzhou Wang\textsuperscript{c}, Fellow, IEEE, and Xiaohu You\textsuperscript{a}, Fellow, IEEE

Abstract—This paper investigates the performance of a two-timescale transmission design for uplink reconfigurable intelligent surface (RIS)-aided cell-free massive multiple-input multiple-output (CF-mMIMO) systems. We consider the Rician channel model and design the passive beamforming of RISs based on the long-time statistical channel state information (CSI), while the central processing unit (CPU) utilizes the maximum ratio combining (MRC) technology to perform fully centralized processing based on the instantaneous overall channel, which is the superposition of the direct and RIS-reflected channels. Firstly, we derive the closed-form approximate expression of the uplink achievable rate for arbitrary numbers of access point (AP) antennas and RIS reflecting elements, which can be used to obtain energy efficiency through the proposed total power model. Relying on the derived expressions, we theoretically analyze the impact of important system parameters on the rate and draw explicit insights into the benefits of RISs. Then, based on the rate expression under statistical CSI, we optimize the phase shifts of RISs by using the genetic algorithm (GA) to maximize the sum rate and minimum rate of users, respectively. Finally, the numerical results demonstrate the correctness of our expressions and the benefits of deploying large-size RISs into cell-free mMIMO systems. Also, we investigate the optimality and convergence behaviors of the GA to verify its effectiveness. To give a more beneficial analysis, we present numerical results to show the high energy efficiency of the system with the help of RISs. Besides, our results have revealed the benefits of distributed deployment of APs and RISs in the RIS-aided mMIMO system with cell-free networks.

Index Terms—Reconfigurable intelligent surface (RIS), cell-free (CF), massive MIMO, two-timescale design, achievable rate, statistical CSI.

I. INTRODUCTION

MASSIVE multiple-input multiple-output (mMIMO) technology has been widely envisioned as an essential technique to achieve high spectral efficiency and network throughput in current and future wireless communication systems [1]. However, in cell-centric multi-cell MIMO systems, all users in one cell are mainly managed by a dedicated access point (AP). Thus, the users near the cell boundary are susceptible to severe inter-cell interference, which results in the limited throughput of the system. Therefore, a novel user-centric network paradigm called cell-free network has been recently proposed to solve the issue [2]. Unlike the concept of classical cell-centric design, cell-free networks have been proposed as a user-centric implementation that can enable all APs to coordinate with each other to provide services for all users without cell boundaries [3]. As a result, the inter-cell interference can be effectively alleviated, and the network capacity can be improved accordingly. In cell-free mMIMO systems, the APs are connected to the central processing unit (CPU) by optical cables or wireless fronthaul, which serve all users in the network with a certain level of cooperation. This promising technique is considered as one of the key technologies in the next generation wireless communication, which can provide uniform quality of service (QoS) to multiple users through simple signal processing [4], [5], [6], [7]. One of the advantages of cell-free mMIMO systems is the lack of cell boundaries, which results in a good QoS for users. Nevertheless, conventional cell-free mMIMO systems require large-scale deployment of APs, leading to poor energy efficiency due to the high costs of both hardware and power consumption.

In this regard, reconfigurable intelligent surface (RIS) has emerged as a cost-efficient revolutionary technology to support high data rate transmission [8], [9], [10]. RIS is a passive array composed of a large number of passive reflecting elements, which can intelligently tune the phase shift or amplitude of the incident signal with the help of a controller, thus strengthening...
the desired signal power or weakening the interference signal. Different from the APs, base station (BSs), and relays, the RIS does not need active radio frequency (RF) chains and power amplifiers, introduces no additive noise, and operates in a full-duplex (FD) mode without self-interference [11]. Due to these attractive benefits, RIS is an efficient and cost-effective solution to increase network capacity, improve transmission reliability [12], reduce transmitted power [13], and enlarge wireless coverage [14]. RISs can also bring gains to various emerging systems, such as RIS-aided massive MIMO systems [15], non-orthogonal multiple access (NOMA) networks [16], secure communication systems [17], device-to-device (D2D) communications [18], and millimeter-wave systems [19]. All these studies provide insightful analysis of the improved performance while exhibiting lower cost and higher efficiency than existing systems.

Motivated by the above background, RISs have also been integrated into cell-free mMIMO systems, and to maximize the energy efficiency, Zhang et al. [20] proposed a hybrid beamforming (HBF) scheme that decomposed the original optimization problem into two subproblems, i.e., the digital beamforming subproblem and the RIS-based analog beamforming subproblem. For the spatially correlated RIS-aided cell-free mMIMO systems, the authors studied the uplink spectral efficiency (SE) of the practical system over Rician fading channels [21]. Furthermore, the impact of the generalized maximum ratio (GMR) combining was investigated in [22], which showed that the GMR could double the data rate over the maximum ratio (MR). The authors of [23] investigated the secure communication in a RIS-aided cell-free mMIMO system in the presence of active eavesdropping, and a RIS-based downlink (DL) transmission scheme was proposed in [24] to minimize the information leakage to eavesdropper while maintaining certain QoS requirements for legitimate users. In addition to the downlink, the system performance of RIS-aided cell-free mMIMO uplink was studied in [25]. Furthermore, Ge et al. [26] proposed a generalized superimposed channel estimation scheme for the uplink cell-free mMIMO system, which is aided by several RISs to improve the SE.

However, most of the above-mentioned contributions considered the design of the passive beamforming at the RIS based on instantaneous channel state information (CSI). In fact, there are two problems associated with the instantaneous CSI-based scheme. The first one is the pilot overhead for the knowledge of instantaneous CSI, which is proportional to the number of RIS elements in most of the existing channel estimation schemes [27], [28]. However, the number of reflecting elements should be large enough to serve an excessive number of users, which results in a prohibitively high pilot overhead. Secondly, the beamforming calculation and information feedback need to be performed in each channel coherence interval for the instantaneous CSI-based scheme, which incurs a high computational complexity, power consumption, and feedback overhead. To tackle these two problems, some authors have proposed a novel RIS design based on the two-timescale scheme, which applies the two-timescale transmission design to the cellular MIMO system assisted by only one RIS [29], [30], [31], [32]. Specifically, the two-timescale scheme designs the phase shifts of RISs only based on slowly-varying statistical CSI, which means that the phase shift of RISs does not need to be redesigned for a long time. The phase shifts of the RISs need to be redesigned only when the statistical CSI changes. Therefore, compared with the instantaneous CSI-based scheme that needs to redesign the phase shifts of RIS in each channel coherence interval, the statistical CSI-based strategy can significantly decrease power consumption and feedback overhead required by the RIS.

This two-timescale design scheme has recently been adopted in RIS-aided cell-free mMIMO systems [33]. Due to multiple APs and RISs, the cell-free mMIMO system requires more signal processing for system control and transmission scheduling, resulting in a more challenging design for beamforming at the AP and the phase shift design at RIS. Also, due to the distributed deployment of multiple RISs and APs in a cell-free network, the RIS-aided cascaded channels are complicated and strongly coupled between different APs, RISs, and users. As a result, the derivation of the rate expression is much more complex than that of the RIS-aided cellular MIMO system with the centralized deployment of RIS and AP. Therefore, it is challenging to derive the closed-form analytical expressions of RIS-aided cell-free mMIMO systems and theoretically characterize the benefits of RISs in the two-timescale transmission design. Specifically, the authors of [33] introduced an aggregated channel estimation approach with low pilot overhead for the RIS-aided cell-free mMIMO systems over spatially-correlated channels. However, the phase shifts of the RIS are not optimized in [33], and the line-of-sight (LoS) link of the RIS-aided cascaded channel is not considered in the proposed channel model. Also, the authors of [34] proposed a two-timescale design scheme of RIS-aided cell-free mMIMO systems with imperfect CSI, where the AP processes and detects the signal locally and then sends it to the CPU for final decoding. Nevertheless, [34] focuses on proposing optimization algorithms and thus lacks analytical insights into the benefits and properties of RIS-aided cell-free systems, such as the power scaling laws. Meanwhile, [34] only considers the sum user rate performance without investigating minimum user rate and energy efficiency.

Against the above background, we propose a two-timescale design of the uplink RIS-aided cell-free mMIMO system with the general Rician fading model. Due to the heights of APs and RISs, both LoS and non-LoS (NLoS) channel components would exist in RIS-aided cell-free mMIMO systems. By adjusting the ratio between LoS and NLoS components, we can gain some critical insights into the spatial diversity of multi-user communication, which also enables us to provide guidelines for RIS deployment. Besides, the CPU utilizes a low-complexity maximum-ratio combining (MRC) receiver for fully centralized processing based on instantaneous overall CSI, while the phase shifts of RISs are designed using statistical CSI. In this paper, we first derive the closed-form approximate expression of the achievable rate and obtain the energy efficiency by combining the proposed total power consumption model. Also, we draw some insights to provide guidelines for deploying RIS into cell-free mMIMO systems. Then, we design the RIS phase shifts relying only on statistical
CSI to solve the sum rate and minimum user rate maximization problems. Moreover, we present numerical simulations to validate our analysis and demonstrate the benefits of deploying RISs into cell-free mMIMO systems. The main contributions of this paper are summarized as follows:

- First, we derive the closed-form approximate expression of the uplink achievable rate based on the general Rician fading model in cell-free mMIMO systems. Under Rician fading models, the derivation becomes much more complicated due to the coupled random variables and the extensive expanded terms. We successfully propose some decoupling methods and an effective framework to complete the calculation of the expectation. Note that this analytical rate expression holds for multiple APs and RISs and arbitrary numbers of AP antennas and RIS elements. Based on the analytical rate expression, we draw some useful insights to analyze the impact of various system parameters on the achievable rate and the asymptotic behaviors of the rate. The insights can serve as clear guidelines for the benefits of the proposed RIS-aided cell-free mMIMO systems. Besides, we propose the total power consumption model to obtain the closed-form expression of energy efficiency, which sheds light on the gain of RISs on the energy efficiency performance.

- Then, using the rate expression, we design the phase shifts of RISs based on statistical CSI. To guarantee fairness among different users and improve system capacity, we exploit the GA-based method to optimize the sum user rate and the minimum user rate, respectively.

- Finally, we validate the correctness of our derived expressions and conduct extensive simulations based on the closed-form approximate rate expression. To verify the effectiveness of the proposed GA method in optimizing the RIS-aided cell-free mMIMO system, we investigate the optimality and convergence behaviors of the GA. Also, numerical results reveal the impacts of both the array size of RISs and the large-scale path loss on the performance of RIS-aided cell-free mMIMO systems. As a result, we validate the promising benefits of integrating large-size RISs into cell-free mMIMO systems. We also showcase promising gains brought by RISs in enhancing energy efficiency. Besides, we show the benefits of distributed deployment of APs and RISs in the RIS-aided mMIMO system with cell-free network.

The remainder of this paper is organized as follows. Section II describes the system model of the uplink RIS-aided cell-free mMIMO systems based on the two-timescale design. Section III derives the closed-form approximate expression of uplink achievable rate and gives some useful insights. Section IV solves the sum rate maximization and the minimum user rate maximization problems by exploiting the GA-based method. Section V provides extensive numerical results and Section VI concludes this paper.

**Notations:** Vectors and matrices are denoted by bold lowercase and uppercase letters, respectively. $\mathbf{A}^H$ and $\mathbf{A}^T$ respectively denote the conjugate transpose and transpose. $|a|$ denotes the modulus of the complex number and $\|a\|$ denotes $l_2$-norm of the vector. $[a]_{m,n}$ denotes the $m$-th entry of the vector $a$. $[\mathbf{A}]_{m,n}$ denotes the $((m - 1)N + n)$-th entry of the $MN \times 1$ vector $\mathbf{A}$. $\Phi$ denotes the $(m,n)$-th entry of matrix $\Phi$. $\mathbb{C}^{M \times N}$ denotes the space of $M \times N$ complex matrix. Besides, $x \sim \mathcal{CN}(a,b)$ is a complex Gaussian distributed random variable with mean $a$ and variance $b$. Operation $\lfloor |n| \rfloor$ and $\lceil |n| \rceil$ respectively denote the nearest integer smaller than and greater than $n$, and operation mod means returning the remainder after division.

**II. System Model**

As shown in Fig. 1, we consider an RIS-aided cell-free mMIMO system, where multiple distributed APs and RISs are deployed to serve all users cooperatively. A CPU is deployed for system control and planning, and decides the transmission scheduling based on the locations of APs, RISs, and users. Specifically, we consider the uplink transmission of the proposed system, where $M$ APs equipped with $M_b$ antennas simultaneously communicate with $K$ single-antenna users with the aid of $N$ RISs equipped with $N_r$ elements.

**A. Channel Model**

The LoS paths between users and the AP could be blocked due to a large number of environmental blocking objects in the communication area. As in [14], [29], and [35], we adopt the Rayleigh fading model, and then the direct channel $\mathbf{D} \in \mathbb{C}^{MM_b \times K}$ can be expressed as $\mathbf{D} = [\mathbf{d}_1, \mathbf{d}_2, \ldots, \mathbf{d}_K]$, $\mathbf{d}_k^T = [d_{1,k}, d_{2,k}, \ldots, d_{M_b,k}]$, $\mathbf{d}_{m,k} = \sqrt{\gamma_{m,k}} \mathbf{d}_{m,k}$, where $\gamma_{m,k}$ is the large-scale path loss, and $\mathbf{d}_{m,k}$ denotes the NLoS of the direct channel between AP $m$ and user $k$. The entries of $\mathbf{d}_{m,k}$ are independent and identically distributed (i.i.d.) complex Gaussian random variables, i.e., $\mathbf{d}_{m,k} \sim \mathcal{CN}(0, \mathbf{I}_{M_b})$.

With the help of RIS, we provide additional paths to assist the communications for users in the rich scattering region. Besides, we consider only the signal reflected by the RIS the first time and ignore the signals reflected by the RIS two or more times [36]. The overall configuration matrix of the RISs can be written as $\mathbf{\Phi} = \text{diag}\{\mathbf{\Phi}_1, \mathbf{\Phi}_2, \ldots, \mathbf{\Phi}_N\} \in \mathbb{C}^{N_r \times NN_r}$, $\mathbf{\Phi}_n = \text{diag}\{e^{j\theta_{n,1}}, e^{j\theta_{n,2}}, \ldots, e^{j\theta_{n,N_r}}\} \in \mathbb{C}^{N_r \times N_r}$, where $\theta_{n,r} \in [0, 2\pi)$ represents the phase shift of the
By contrast, \( \phi \) represents the wavelength, respectively.\( Z \) components of the user and LoS components of the RIS and the AP under the uniform square planar array (USPA) model [42]. Therefore, the LoS components\( \bar{h}_{n,k} \) and\( n,k \) are the corresponding non-line-of-sight (NLoS) channel components, whose elements are i.i.d. complex Gaussian random variables following the distribution of \( CN(0, 1) \) [40],[41]. Furthermore, we characterize the LoS paths of the cascaded channel between the RIS and the AP under the uniform square planar array (USPA) model [42]. Therefore, the LoS components\( \bar{h}_{n,k} \) and\( \bar{Z}_{m,n} \) can be respectively modeled as follows
\[
\bar{h}_{n,k} = a_{N_n} (\varphi_{n,k}^a, \varphi_{m,n}^c) \triangleq a_{N_n}(n, k), \\
\bar{Z}_{m,n} = a_{M_m} (\varphi_{n,k}^b, \varphi_{m,n}^c) \triangleq a_{M_m}(m, n),
\]
where\( a_{X}(\vartheta^a, \vartheta^c) \in \mathbb{C}^{X \times 1} \) is the array response vector and its \( x \)-th entry is
\[
[a_{X}(\vartheta^a, \vartheta^c)]_x = \exp \left[ j2\pi \frac{d}{\lambda} \left( \frac{(x-1)}{\sqrt{X}} \right) \sin \vartheta^a \sin \vartheta^c \\
+ \left( (x-1) \mod \sqrt{X} \right) \cos \vartheta^c \right],
\]
where\( d \) and\( \lambda \) denote the element spacing and carrier wavelength, respectively.\( \varphi_{n,k}^a \) and\( \varphi_{n,k}^c \) are respectively the azimuth and elevation angles of arrival (AoA) of the incident signal at RIS \( n \) from user \( k \).\( \varphi_{m,n}^a \) and\( \varphi_{m,n}^c \) represent the azimuth and elevation angles of departure (AoD) reflected by RIS \( n \) to AP \( m \), respectively.\( \varphi_{m,n}^c \) and\( \varphi_{m,n}^c \) respectively denote the AoA at AP \( m \) from RIS \( n \).

B. Uplink Transmission

The collective signal received by all APs can be written as
\[
y = (G + D)Px + n = (Z\Phi + D)Px + n,
\]
where\( P = \text{diag}\left\{ \sqrt{p_1}, \sqrt{p_2}, \ldots, \sqrt{p_K} \right\} \) and\( p_k \) is transmit power of user \( k \).\( x = [x_1, x_2, \ldots, x_K]^T \in \mathbb{C}^{K \times 1} \) represents the transmit symbols of \( K \) users, where\( \mathbb{E}\{|x_k|^2\} = 1 \) and\( x_k \sim \mathcal{CN}(0, \sigma^2 I_{M_m}) \) denotes the receiver noise vector, which is additional white Gaussian noise (AWGN).

We assume all APs send their received data and pilot signals to the CPU, while the CPU perfectly estimates all channels and utilizes the MRC technology for fully centralized processing [43],[44]. Thus, with perfect CSI, the CPU performs MRC by multiplying the received signal \( y \) with\( (G + D) \) as follows
\[
r = (G + D)^{H}y \\
= (G + D)^{H}(G + D)Px + (G + D)^{H}n,
\]
and the received signal corresponding to user \( k \) can be expressed as
\[
r_k = \sum_{i=1, i \neq k}^{K} \sqrt{p_i} (g_{i,k} + d_k)(g_i + d_i)x_i \\
+ \sqrt{p_k} (g_k + d_k)(g_k + d_k)x_k + (g_k + d_k)^Hn_k.
\]

In general, the total power consumed by an RIS-aided multi-cell-free MIMO system, we first define the energy efficiency\( \eta_k \) as
\[
\eta_k = \frac{N_{R_k} - N_{R_k}^{c}}{N_{R_k}^{c}}
\]
where\( \eta_k = \frac{N_{R_k} - N_{R_k}^{c}}{N_{R_k}^{c}} \) and\( N_{R_k}^{c} \) represents the total user rate [46],[47]. Based on the above considerations, we characterize the total power consumed by RIS-aided cell-free MIMO systems as follows:
\[
P_{\text{total}} = \sum_{k=1}^{K} \left( \frac{1}{\xi_k}p_k + P_{c,k} \right).
\]
where \(0 < \xi_k \leq 1\) represents the efficiency of the transmit power amplifier deployed at the user \(k\), constants \(P_{ap,m}\) and \(P_{c,k}\) are the circuit hardware power consumed by each antenna of AP \(m\) and user \(k\), respectively. \(P_{th,m}\) is the power consumption of the fronthaul link connecting the CPU and the AP \(m\), given by \(P_{th,m} = P_{0,m} + P_{th,m} W \sum_{k=1}^{K} R_k\), where \(P_{0,m}\) is a fixed power consumption of each fronthaul, \(P_{ft,m}\) is the traffic-dependent power (in Watt per bit/s), and \(W\) is the system bandwidth. Moreover, the term \(N_r P_{ris,n}(b)\) represents the power consumption of the RIS \(n\) with a \(b\)-bit resolution phase shifter, proportional to the number of RIS elements \(N_r\), where \(P_{ris,n}(b)\) is the hardware power consumed by each element.

### III. UPLINK ACHIEVABLE RATE ANALYSIS

In this section, we first derive the closed-form approximate expression of the uplink rate and then draw some useful insights to serve as clear guidelines for the benefits of the RIS-aided cell-free mMIMO systems. The expression of energy efficiency is also derived and analyzed based on the power consumption model.

**Theorem 1:** In the RIS-aided cell-free mMIMO system, the closed-form expression of the uplink achievable rate of user \(k\) can be approximated as

\[
R_k \approx \log_2 \left( 1 + \frac{p_k E_k^{(signal)}(\Phi)}{\sum_{i=1, i \neq k}^K p_i I_{ki}(\Phi) + \sigma^2 E_k^{(noise)}(\Phi)} \right),
\]

where \(E_k^{(noise)}(\Phi)\), \(E_k^{(signal)}(\Phi)\), and \(I_{ki}(\Phi)\) represent the terms of noise, desired signal, and multi-user interference, respectively, and their detailed expressions are shown in (46), (47), as shown at the bottom of page 15, and (48), as shown at the bottom of page 16, at the end of Appendix A. Besides, some parameters and functions in \(E_k^{(noise)}(\Phi)\), \(E_k^{(signal)}(\Phi)\), and \(I_{ki}(\Phi)\) are defined as follows

\[
c_{m,n,k} \triangleq \beta_{m,n,k} \alpha_{n,k},
\]

\[
f_{m,n,k}(\Phi) \triangleq a_{m,n} \sum_{r=1}^{N_r} e^{j(c_{m,n,k} + \theta_{m,n,r})},
\]

where

\[
c_{m,n,k} = 2 \pi \vartheta_{m,n} \lambda \left( (r-1)/\sqrt{N_r} \right) \sin \varphi_{n,k} \sin \varphi_{m,n} - \sin \varphi_{n,k} \times \sin \varphi_{m,n} + (r-1) \mod \sqrt{N_r} \times \cos \varphi_{n,k} \cos \varphi_{m,n}.
\]

**Proof:** Please refer to Appendix A.

**Corollary 1:** The achievable rate for RIS-free cell-free mMIMO systems can be given by \(R_k^{(w)} = \log_2 \left( 1 + SINR_k^{(w)} \right)\) with

\[
\begin{align*}
SINR_k^{(w)} &= p_k \left( \sum_{m=1}^{M} \gamma_{m,k} M_b + \sum_{m=1}^{M} \gamma_{m,k} M_b \right) - \sum_{i=1, i \neq k}^K p_i \gamma_{i,k} \gamma_{m,k} M_b + \sigma^2 \sum_{m=1}^{M} \gamma_{m,k} M_b.
\end{align*}
\]

**Proof:** We can ignore the zero terms by substituting \(N = 0\) or \(c_{m,n,k} = 0, \forall m, n, k\) into the rate expression in (13) and then complete the proof after some simplifications.

**Corollary 2:** Corollary 1 shows that in the conventional cell-free mMIMO system, i.e., the RIS-free cell-free mMIMO system, when the number of AP antennas \(M_b\) is large, the interference power could be ignored compared with the desired signal. However, this feature no longer holds for the RIS-aided cell-free mMIMO system. We can find that both the signal term in (47) and the interference term in (48) scale as \(O(M_b^2)\), which means that a large \(M_b\) will make the SINR converge to a constant. Therefore, a question arises as to whether this additional interference will limit the gains of RISs. To answer
this question and get more insights, we consider some special cases to explore the gains of deploying RISs.

**Corollary 2:** When the RIS-aided cascaded channels are pure NLoS (i.e., \( \epsilon_{m,k} = 0 \) and \( \delta_{m,n} = 0 \), \( \forall m, n, k \)), the achievable rate is \( R_k^{(NL)} \) with

\[
\text{SINR}_k^{(NL)} \approx \frac{p_k E_k^{(NL)}}{\sum_{i=1,i \neq k}^K p_i F_i^{(NL)} + \sigma^2 E_k^{(noise,NL)}},
\]

where

\[
E_k^{(noise,NL)} = \sum_{m=1}^M \sum_{n=1}^N \epsilon_{m,n,k} M_b N \gamma_m M_b + \sum_{m=1}^M \sum_{n=1}^N \sum_{k=1}^N \gamma_{m,n,k} M_b N^2 \]

\[
E_k^{(NL)} = \sum_{m=1}^M \sum_{n=1}^N \sum_{k=1}^N \epsilon_{m,n,k}^2 M_b^2 N^2 + \sum_{m=1}^M \sum_{n=1}^N \sum_{k=1}^N \gamma_{m,n,k} M_b N^2 \]

\[
I_{k_{(NL)}}^{(NL)} = \sum_{m=1}^M \sum_{n=1}^N \sum_{k=1}^N \left( \epsilon_{m,n,k}^2 \gamma_{m,n,k} + \gamma_{m,n,k} \gamma_{m,n,k} \right) + \sum_{m=1}^M \sum_{n=1}^N \gamma_{m,n,k} M_b N^2,
\]

Corollary 2 represents the scenario where the cascaded channels are in a rich scattering environment, and the Rician channel reduces to the Rayleigh channel. It can be seen that the rate expression \( R_k^{(NL)} \) does not contain the parameter \( \Phi \), which means the rate performance is not affected by the phase shifts of RISs. Therefore, there is no need to design the phase shifts of RISs when the communication area around RIS is located in a rich scattering environment.

Besides, when \( M_b \to \infty \), we can get the result by ignoring the insignificant terms which do not scale with \( M_b \). Note that \( \alpha_{M_b}(m_1,n_1) \alpha_{M_b}(m_2,n_2) = 0 \) when \( m_1 = m_2 \) and \( n_1 = n_2 \). Then, as \( M_b \to \infty \), we have

\[
R_k^{(NL)} \to R_k^{(NL)}(1) \approx \log_2 \left( 1 + \frac{p_k E_k^{(NL)}}{\sum_{i=1,i \neq k}^K p_i F_i^{(NL)}} \right),
\]

where

\[
E_k^{(NL)}(1) = \sum_{m=1}^M \sum_{n=1}^N \sum_{i=1,i \neq k}^K c_{m,n,k} \gamma_{m,n,k} + c_{m,n,k} \gamma_{m,n,k} N^2
\]

\[
nr_{k_{(NL)}}^{(NL)}(1) = \sum_{m=1}^M \sum_{n=1}^N \sum_{k=1}^N \left( \epsilon_{m,n,k}^2 \gamma_{m,n,k} + \gamma_{m,n,k} \gamma_{m,n,k} \right) + \sum_{m=1}^M \sum_{n=1}^N \gamma_{m,n,k} M_b N^2.
\]

We can find that \( R_k^{(NL)}(1) \to \infty \) when \( N_r \to \infty \), which means the rate \( R_k^{(NL)} \) will increase without limit when both \( N_r \to \infty \) and \( M_b \to \infty \). Therefore, even though there are no LoS paths in the cascaded channels, significant performance benefits can be obtained by deploying RISs with large numbers of reflecting elements in cell-free mMIMO systems.

**Corollary 3:** When the cascaded channels are only LoS (i.e., \( \epsilon_{m,n} \to \infty \) and \( \delta_{m,n} \to \infty \), \( \forall m, n, k \)), and \( M_b \to \infty \), the achievable rate is \( R_k^{(OL)} \) with

\[
\text{SINR}_k^{(OL)} \approx \frac{p_k E_k^{(NL)}}{\sum_{i=1,i \neq k}^K p_i F_i^{(OL)}} \Phi,
\]

where

\[
E_k^{(NL)}(\Phi) = \sum_{m=1}^M \sum_{n=1}^N \sum_{k=1}^N \left( \epsilon_{m,n,k}^2 \gamma_{m,n,k} + \gamma_{m,n,k} \gamma_{m,n,k} \right) + \sum_{m=1}^M \sum_{n=1}^N \gamma_{m,n,k} M_b N^2
\]

\[
nr_{k_{(NL)}}^{(NL)}(\Phi) = \sum_{m=1}^M \sum_{n=1}^N \sum_{k=1}^N \left( \epsilon_{m,n,k}^2 \gamma_{m,n,k} + \gamma_{m,n,k} \gamma_{m,n,k} \right) + \sum_{m=1}^M \sum_{n=1}^N \gamma_{m,n,k} M_b N^2.
\]

Authorized licensed use limited to the terms of the applicable license agreement with IEEE. Restrictions apply.
Proof: As all the Rician factors go to infinity, we have $c_{m,n,k} = 0$ but $c_{m,n,k}\delta_{m,n}\varepsilon_{n,k} = \alpha_{n,k}\beta_{m,n}, \forall m,n,k$, and the zero terms can be ignored. When $M_b \to \infty$, we can ignore the insignificant terms which do not scale with $M_b$. Note that $H_{M_b}(m_1,n_1)\mathbf{a}_{M_b}(m_2,n_2) = M_b$ when $m_1 = m_2$ and $n_1 = n_2$. Then, we can complete the proof by some simplifications.

Corollary 3 represents the scenario where no obstacles exist and only LoS paths exist. We can find that in the RIS-aided cell-free mMIMO systems with a low-complexity MRC scheme, when the number of AP antennas $M_b$ is large, the achievable rate in (26) will converge to a constant due to the existence of interference. However, this rate degradation can be compensated by properly designing phase shifts $\Phi$. For example, when the phase shifts are aligned to user $k$, the interference imposed on user $k$ can be compensated by properly designing phase shifts $\Phi$. The existence of interference. However, this rate degradation can be compensated by properly designing phase shifts $\Phi$. For example, when the phase shifts are aligned to user $k$, the interference imposed on user $k$ becomes negligible compared with the desired signal received by user $k$. This observation emphasizes the importance of optimizing the phase shifts $\Phi$.

Corollary 4: Assume that the phase shifts of RISs are set randomly in each coherence interval. When $M_b \to \infty$ and $N_r \to \infty$, the achievable rate is given by $R_{\text{TM}}(\text{ram})$ with

$$\text{SINR}_k \approx \frac{p_k E^{(\text{signal,rm})}_k}{\sum_{i=1,i\neq k}^{K} p_i I_{ki}^{(\text{rm})}},$$

(29)

where

$$E^{(\text{signal,rm})}_k = \sum_{m=1}^{M} \sum_{n=1}^{N} c_{m,n,k}\varepsilon_{n,k}^2,$$

$$I_{ki}^{(\text{rm})} = \sum_{m=1}^{M} \sum_{n=1}^{N} c_{m,n,k}\varepsilon_{n,k}(\varepsilon_{n,k} + 1).$$

Proof: Please refer to Appendix B.

Corollary 4 shows that when the antenna number $M_b$ and element number $N_r$ are large, the achievable rate is still limited in random phase shifts-based RIS-aided systems. This conclusion emphasizes the necessity of optimizing the phase shifts $\Phi$ in the RIS-aided cell-free mMIMO systems. Nevertheless, it can be found that as the numbers of APs and RISs increase (i.e., as $M$ and $N$ increase), the rate in (29) could be improved effectively, demonstrating the gains of applying the cell-free and multi-RIS structures. Besides, unlike the case of $\delta_{m,n} = 0$ in Corollary 2, we can find that the rate in (29) decreases as $\delta_{m,n}$ increases. This is because when the phase shifts are designed randomly in each coherence interval, it tends to distribute the passive beamforming gain equally among all users.

Corollary 5: According to the power consumption model in (12) and the closed-form rate expression in (13), the expression of the ergodic energy efficiency can be given by

$$EE_{\text{ris}} \triangleq \frac{W}{P_{\text{total}}},$$

(32)

where $P_{\text{total}} = \sum_{k=1}^{K} (\frac{1}{\xi_k} p_k + P_{c,k}) + \sum_{m=1}^{M} (M_b P_{ap,m} + P_{0,m} + P_{ft,m}W \sum_{k=1}^{K} R_k) + \sum_{n=1}^{N_r} N_r P_{\text{ris},n}(b)$.

It can be seen that the denominator of (32) contains a term proportional to the rate, which means that the energy efficiency will increase limitedly as the rate increases. We consider a case where all users’ transmit power is equal to $p$, i.e., $p_k = p, \forall k$. Based on the rate expression (13), we can find that the rate will increase as $p$ increases. However, even when $P_{b,m}$ is small enough to approach zero power, the energy efficiency is limited as $p$ increases due to the term $\sum_{k=1}^{K} \frac{1}{\xi_k} p$ proportional to $p$.

The corresponding energy efficiency in the RIS-free cell-free mMIMO system is $EE_{\text{d}} \triangleq W \sum_{k=1}^{K} R_k / P_{\text{total}}$, where $P_{\text{total}} = \sum_{k=1}^{K} (\frac{1}{\xi_k} p_k + P_{c,k}) + \sum_{m=1}^{M} (M_b P_{ap,m} + P_{0,m} + P_{ft,m}W \sum_{k=1}^{K} R_k)$, and $R_k^{(w)}$ has been given in Corollary 1. We consider that the rate of RIS-aided systems is greater than that of RIS-free systems due to the performance improvement brought by the RIS. In the case of the same power consumption, the RIS-aided system has a higher energy efficiency than the RIS-free system. However, compared to $P_{\text{total}}$, the RIS-free system’s power consumption $P_{\text{total}}^{(w)}$ lacks the hardware power $\sum_{n=1}^{N_r} N_r P_{\text{ris},n}(b)$ consumed at RISs. Although the power consumption at RISs is insignificant, the energy efficiency of RIS-aided systems is not necessarily higher than that of RIS-free systems given a large number of RIS reflecting elements. Therefore, it is valuable to evaluate the energy efficiency of RIS-aided systems.

IV. PHASE SHIFTS DESIGN

In this section, we design the phase shifts of RISs to maximize the achievable rate derived in Theorem 1. Since the derived rate expression only depends on long-term statistical CSI, we only need to redesign the phase shifts of RISs when the long-term CSI changes, which could effectively decrease the channel estimation overhead and the update frequency at RISs. To make a more general investigation into the RIS-aided cell-free mMIMO system, we consider two optimization problems with different objective functions, i.e., the sum rate maximization and the minimum user rate maximization problems. The sum rate maximization problem, which improves the system capacity, is formulated as

$$\max_{\Phi} \sum_{k=1}^{K} R_k(\Phi),$$

(33a)
where \( R_k(\Phi) \) is given by (13). Then, the minimum user rate maximization problem, which could guarantee user fairness and represent system spatial multiplexing, is formulated as

\[
\max_k \min \Phi R_k(\Phi), \\
\text{s.t.} \quad (33b).
\]

Considering the complex structure of the achievable rate expressions, we can conclude that the objective function is non-convex. Meanwhile, the unit modulus constraint is also non-convex. Therefore, the optimization problems (33) and (34) are non-convex. Due to the non-convex feature and the complex objective function and the tightly coupled phase shifts \( \Phi_n \) from different RISs, the two problems are NP-hard and difficult to obtain globally optimal solutions by exploiting conventional optimization methods, such as semi-definite programming (SDP), the majorization-minimization (MM), and the gradient ascent-based algorithms. Therefore, we solve the two optimization problems using the GA method to obtain the near-optimal solution, whose effectiveness in optimizing RIS-aided systems has been validated in [48].

The GA-based method simulates the evolution of a population [49], and its main idea is to regard the RIS phase shifts as the gene of a population. GA is initialized by generating a population with \( S \) individuals, where \( S_c \) individuals will be selected as elites, 2\( S_c \) individuals will be selected as crossover parents to generate offspring, and the remaining \( S_m = S - S_c - S_e \) individuals will be used for mutation operation. The detailed process of the GA-based optimization method is given in Algorithm 1, where \( N_r \) is the number of RIS elements, and \( N \) is the number of RISs.

Specifically, we describe the implementation details of the GA for both problems (33) and (34) in the following five parts. It is worth noting that the unit modulus constraint is transformed to the constraints of the angles \( \theta_{n,r} \) between 0 and \( 2\pi \) in the GA.

1) Initial Population: GA generates an initial population with \( S \) individuals, where each individual contains \( N N_r \) chromosomes, and the \( ((n-1)N_r+r) \)-th chromosome corresponds to the phase shift \( \theta_{n,r} \) of the \( r \)-th reflecting element of RIS \( n \). In general, the chromosomes of individuals in the population are randomly generated from \([0, 2\pi]\) in the GA.

2) Fitness Evaluation and Scaling: We first calculate the initial fitness of each individual in the current population based on the objective function in (33a) or (34a). Next, we will scale the initial fitness value of individuals in terms of their rank in the population. Individuals with large fitness may reproduce their chromosomes too frequently and have premature convergence in the population. Therefore, we convert initial fitness values to a more suitable range by utilizing a rank scaling method to avoid premature. We sort the initial fitness of individuals and compute their scaled fitness as follows

\[
f_i = 2S_c \frac{\text{rank}_i^{-0.5}}{\sum_{s=1}^{S_c} \text{rank}_s^{-0.5}}, \quad 1 \leq s \leq S_c,
\]

where \( f_i \) is the scaled fitness of individual \( i \), \( \text{rank}_i \) is the index of initial fitness of individual \( i \) after the descending sort, and \( S_c \) is the number of crossover parents to be used in the next operation.

3) Selection: Then, some individuals are selected as elites from the current population, and some are chosen as parents to generate offspring. Specifically, we select \( S_c \) individuals with larger \( f_i \) as elite individuals, which will be directly passed on to the next generation. Then, 2\( S_c \) individuals will be selected as crossover parents based on stochastic universal sampling. To perform this method, we form a roulette wheel with 2\( S_c \) slots corresponding to the crossover parents, and the size of slot \( i \) is proportional to \( f_i \) as follows

\[
\text{slot}_i = \frac{f_i}{2S_c},
\]

with \( \sum_{i=1}^{S_c} \text{slot}_i = 1 \). Next, we rotate the roulette wheel 2\( S_c \) times in an equal step \( \frac{1}{2S_c} \) and select the individual to which the pointer points as a parent. Finally, we will use the remaining \( S_m = S - S_c - S_e \) individuals for mutation operation.

4) Crossover: Based on previously selected 2\( S_c \) parents, we perform the crossover operation that can extract and recombine the best chromosome from parents to generate superior \( S_c \) offspring. Specifically, we adopted the
two-point crossover method, and its processes are sketched in Algorithm 2.

Algorithm 2 Crossover Algorithm
1: Set $a_1 = 1$, $a_2 = 2$;
2: for $i = 1 : S_c$ do
3: Select the $a_1$-th and the $a_2$-th parents from the $2S_c$ combination;
4: Generate two different integer crossover points $i_1$ and $i_2$ randomly from $[1, NN_r - 1]$;
5: if $i_1 > i_2$ then
6: Swap $i_1$ and $i_2$, swap parents $a_1$ and $a_2$;
7: end if
8: Generate the $i$-th offspring by $[a_1(1 : i_1), a_2(i_1 + 1 : i_2), a_1(i_2 + 1 : NN_r)];$
9: $a_1 = 2 + a_1$, $a_2 = 2 + a_2$;
10: end for

5) Mutation: To increase the diversity of the population and improve the possibility of generating offspring with better fitness, we use the remaining $S_m$ individuals to conduct the mutation operation with probability $p_m$ and produce $S_m$ offspring. In this paper, we use the uniform mutation method and sketch the process in Algorithm 3.

Algorithm 3 Mutation Algorithm
1: for $i = 1 : S_m$ do
2: for $n = 1 : NN_r$ do
3: if $p_m > \text{rand}(1)$ then
4: The $n$-th chromosome $\theta_n$ of parent $i$ mutates to $2\pi \times \text{rand}(1)$;
5: end if
6: end for
7: end for

V. NUMERICAL RESULTS

In this section, we present several numerical simulations to study the performance of the RIS-aided cell-free mMIMO system. By exploiting Theorem 1 and the GA, we can solve these two optimization problems in Section IV and obtain the analytical results. Unless otherwise specified, we consider a cell-free network with the topology shown in Fig. 2. In this setup, a cell-free mMIMO system with $M = 3$ APs serves $K = 4$ users, while the system throughput is limited due to the obstruction of the woods. To address the issue, we deploy $N = 2$ RISs separately on two building surfaces, which are high enough to construct extra reflection links. Similar to [50], we assume that four users are randomly distributed in a circle centered at $(75, 0, 0)$ m with a radius of $3$ m and height of $0$ m. The AoA and AoD of APs, RISs, and users are generated randomly from $0, 2\pi$ [14], [51], and these angles will be fixed after the initial generation. Large-scale path-loss is calculated as $\alpha_{n,k} = 10^{-3}(d_{m,k}^{\text{UR}})^{-\beta_{\text{UR}}}$, $\beta_{m,n} = 10^{-3}(d_{m,n}^{\text{RA}})^{-\beta_{\text{RA}}}$, and $\gamma_{m,k} = 10^{-3}(d_{m,k}^{\text{UA}})^{-\beta_{\text{UA}}}$ [14], where $d_{m,k}^{\text{UR}}$, $d_{m,n}^{\text{RA}}$, and $d_{m,k}^{\text{UA}}$ are respectively the distances of user $k$-RIS $n$, RIS $n$-AP $m$, and user $k$-AP $m$, and the path-loss exponents are $\beta_{\text{UR}} = 2$, $\beta_{\text{RA}} = 2.5$, and $\gamma_{\text{UA}} = 4$, $\forall m, n, k$ [52], [53]. Also, other simulation parameters (unless otherwise stated) are defined in Table I. The MC simulations are obtained by averaging $10^5$ random channel realizations.

Firstly, we verify the accuracy of the derivations of the closed-form expressions in (46), (47), and (48). Fig. 3 shows the desired signal power $p_1\mathbb{E}\{||g_1 + d_1||^4\}$, the sum power of multi-user interference $\sum_{i=2}^{4} p_i \mathbb{E}\{||g_i^H + d_i^H||^2\}$ and noise $\sigma^2\mathbb{E}\{||g_1 + d_1||^4\}$ for user 1 under two independent random channel realizations. For each number of RIS elements or AP antennas, we employ random phase shifts to adequately verify the accuracy of our derivations. The theoretical results of random channel realization are obtained from our derived expressions. Similarly, based on the same channel realization and phase shift, we can get the results from the MC simulation. It can be observed that the results obtained from our derived expressions exactly match the MC simulations, which verifies the accuracy of our derivations and the correctness of the expressions.

Next, we utilize the GA to optimize the phase shifts of RISs based on the closed-form approximate expression of the achievable rate in (13), and all of the following MC simulations are presented to compare further the approximations of the rate obtained from the expression. To this

3The curves are non-monotonic since random (unoptimized) phase shifts are employed in plotting each point leading to random performance.
TABLE I
SIMULATION PARAMETERS

| AP antennas | $M_b = 9$ | RIS elements | $N_r = 49$ | Amplifier efficiency | $\xi_k = 0.3$ | Bandwidth | $W = 20$ MHz |
|-------------|-----------|--------------|------------|----------------------|--------------|-----------|--------------|
| Antenna spacing | $d = \Lambda/2$ | Noise power | $\sigma^2 = -104$ dBm | Hardware consumed power | $P_{k, m} = 20$ dBm, $P_k = 10$ dBm, $\forall m, k$ |           |             |
| Transmit power | $p_k = P = 30$ dBm, $\forall k$ | Pronthail consumed power | $P_{0, m} = 23$ dBm, $\forall m$ | Traffic-dependent power | $P_{0, m} = 24$ dB/m(Gbits/s), $\forall m$ |           |             |
| Ricean factors | $\delta_{m, n} = 1, \epsilon_{n, k} = 10, \forall m, n, k$ | Per-element static power at RIS | $P_{\text{pse, } n}(b) = 25$ dBm for $b \to \infty, \forall n$ |           |             |           |             |
| GA parameters | $S_c = 10, S_c = 160, S_m = 30, \rho_m = 0.2$ |           |             |           |             |           |             |

Fig. 4. The optimality and convergence behaviors of the proposed GA method.

(a) Convergence behavior of the GA method for different population sizes and RIS elements.

(b) Performance comparison of the GA method and the exhaustive search method for $N_r = 4$.

end, we respectively solve the optimization Problem (33) and Problem (34) to obtain their optimized phase shifts $\Phi_{\text{sum}}^*$ and $\Phi_{\text{min}}^*$, and then calculate two kinds of performance metrics, i.e., the sum rate $\sum_{k=1}^{K} R_k(\Phi_{\text{sum}}^*)$ and the minimum user rate $\min_k R_k(\Phi_{\text{min}}^*)$.

A. The Optimality and Convergence Analysis of the Proposed GA Method

In this subsection, based on numerical simulations similar to [54] and [55], we investigate the optimality and convergence behaviors of the GA to verify the effectiveness of the proposed GA method. In Fig. 4(a), we verify the convergence behavior of the GA for the different values of population size $S = \{100, 150, 200\}$ having the number of RIS elements $N_r = \{16, 25\}$, where the sum rate is calculated by the optimized phase shift $\Phi_{\text{sum}}^*$ based on the GA. From this figure, we observe that although the population size $S$ of the GA can affect the convergence speed of this method, the achievable rate is almost unaffected when $S$ is large enough. Furthermore, both the convergence speed and the achievable rate are affected by the number of RIS elements $N_r$. The results show that the GA method converges within 1000 iterations, and the maximum number of iterations decreases as the RIS element number $N_r$ decreases or the population size $S$ increases. Generally, we cannot guarantee that the GA achieves the optimal design. However, to showcase the effectiveness of the proposed GA, we consider comparing it with the exhaustive search method. Constrained by the computational complexity, we consider the case with discrete phase shifts and small $N_r$. Note that the proposed algorithm can be easily extended to the case with discrete phase shifts. The phase shift of the $r$-th reflecting element of the $r$-th RIS with $b$ bits quantization precision can be expressed as $\theta_{n,r} \in \left\{ \frac{2\pi}{2^b}, \ldots, \left(2^b - 1\right) \frac{2\pi}{2^b} \right\}, \forall n, r$. In Fig. 4(b), we test 20 random channel realizations under different numbers of AP antennas, where “GA” refers to the sum rate optimized by the GA, and “exhaustive search” refers to the sum rate obtained by the exhaustive search method [48]. It can be seen that the GA has almost the same rate performance as the globally optimal solution obtained by the exhaustive search method under different channel realizations and AP antennas, which demonstrates the optimality behavior of the GA.

B. The Interplay Between RIS and Cell-Free Massive MIMO

In this subsection, we aim to investigate the performance gains of deploying RISs in cell-free mMIMO systems. To achieve a larger system capacity and guarantee user fairness, we consider the sum rate maximization problem in (33) and the minimum user rate maximization problem in (34) in the following simulations. We consider the cell-free mMIMO systems with the optimized phase shifts-based RISs, where the sum rate calculated by the optimized phase shift $\Phi_{\text{sum}}^*$ is called “sum rate by max-sum” or “sum rate by optimized phase”, the sum rate calculated by the optimized phase shift $\Phi_{\text{min}}^*$ is called “sum rate by max-min”, the minimum user rate calculated by the optimized phase shift $\Phi_{\text{sum}}^*$ is called “min rate by max-sum”, and the minimum user rate calculated by the optimized phase shift $\Phi_{\text{min}}^*$ is called “min rate by max-min” or “min rate by optimized phase”. For comparison, we adopt the following benchmark systems:

1) The Cell-Free Massive MIMO System With RISs Under the Random Phase Shifts: We consider the random phase shifts-based design of RISs, where the sum rate and minimum user rate obtained by averaging $10^5$ random phase shifts are
2) The RIS-Free Cell-Free Massive MIMO System: We consider the RIS-free cell-free mMIMO system as the benchmark system [4], which means that the number of RISs is zero and represents the conventional cell-free mMIMO system. The sum and minimum user rates of the RIS-free system are called “sum rate by RIS-free” and “min rate by RIS-free”, respectively.

Fig. 5 plots the achievable rate versus transmit power \( P \). As \( P \) increases, there is a small gap between the MC simulation and the analytical result due to the approximate rate expression. In general, the approximated rate expression (13) matches well with the MC simulations, validating the correctness of our derived expressions. Besides, we can see that deployment of the RISs can still effectively improve the rate performance in cell-free mMIMO systems in the low signal-to-noise ratio (SNR) region. However, as SNR increases, the RIS-free systems will gradually exceed random phase shifts-based RIS systems. This is because the rate will be limited by multi-user interference in the high SNR region, which aggravates the negative impacts of additional interference caused by RISs. Moreover, with the help of optimized phase shift design, the achievable rate of the RIS-aided system is always higher than that of the RIS-free system for arbitrary SNR, which shows the benefits of deploying RISs and emphasizes the importance of optimizing phase shifts.

In Fig. 6, we evaluate the achievable rate as a function of the number of RIS reflecting elements \( N_r \). The RIS-free system arises for ease of comparison with the RIS-aided cell-free mMIMO system to show the gains of the optimized phase-based RIS. It can be seen that as the number of RIS reflecting elements \( N_r \) increases, the optimized phase shifts-based RIS can bring a significant performance improvement to the cell-free mMIMO systems, which supports the motivation of deploying the RIS in cell-free mMIMO systems. By contrast, the rate improvement over the random phase shifts-based RIS design is smaller and approaches saturation when \( N_r \to \infty \).

The result shows that the rate gap caused by the optimized phase shift increases with the increase of RIS elements, which again emphasizes the importance of optimizing phase shifts.

In Fig. 7, we show the achievable rate versus the number of AP antennas \( M_b \). Specifically, Fig. 7(a) compares the achievable rate under different phase shift designs at RISs. It can be seen that as \( M_b \) increases, the rate could be significantly improved and will approach saturation when \( M_b \to \infty \) due to the multi-user interference. Furthermore, when \( M_b \) is large, the RIS-free cell-free mMIMO systems will outperform the random phase shifts-based RIS-aided systems. Then, in the RIS-free cell-free mMIMO systems, a large number of AP antennas are required to serve a large number of users. However, increasing the number of active AP antennas requires a large-sized array,
high power consumption, and high hardware cost. To tackle this problem, we show the achievable rate under the optimized phase shift-based RIS with $N_r = 25$ and $N_r = 49$ in Fig. 7(b). It can be observed that with the help of the RIS, we can achieve the same throughput as the RIS-free cell-free mMIMO systems with a smaller number of antennas. In particular, the rate achieved by the system with 25 AP antennas and 49 RIS elements is equal to that achieved by the system with 100 AP antennas and 25 RIS elements. Due to the RIS element’s lower cost and energy consumption, RIS-aided cell-free mMIMO systems are promising to maintain the network capacity requirement with much-reduced hardware cost and power consumption in future wireless communication systems.

The above simulations are carried out for $\beta_{RA} = 2.5$ under the assumption that the RIS-AP channels are less blocked. However, this scenario may not hold in more complex environments. Hence, it is essential to investigate the impact of path-loss exponent $\beta_{RA}$ on the system performance, as shown in Fig. 8. We can find that both the max-sum problem (33) and max-min problem (34) lead to similarly good performance when $\beta_{RA}$ is small. Secondly, as the path-loss exponent $\beta_{RA}$ increases, if we want to guarantee user fairness, the achievable rate will reduce and ultimately approach the rate obtained by random phase shifts. Based on these findings, we could conclude that if we want to maintain high network capacity and guarantee user fairness, the path-loss exponent should be as small as possible, corresponding to a shorter distance. Therefore, we should carefully select the locations of RISs to find a less blocked scenario associated with a low $\beta_{RA}$.

C. The Comparison of Cell-Free and Single-Cell Networks

In this subsection, we evaluate the achievable rate performance in cell-free and single-cell networks. As shown in Fig. 2, we consider the single-cell mMIMO system with the optimized phase shift-based RIS as a new benchmark system [29], where the centralized RIS with $2N_r$ elements and the centralized AP with $3M_b$ antennas are deployed respectively at (75 m, 28 m, 5 m) and (75 m, -803 m, 8 m). The single-cell network is compared to the cell-free network with the distributed deployment of APs and RISs, where the latter represents the cell-free mMIMO systems with RISs under the optimized phase shifts. Fig. 9 shows the rate performance of RIS-aided mMIMO systems in single-cell and cell-free networks. We can see that the achievable rate in the cell-free network is higher than that in the single-cell network. This is because different distances from distributed RISs to distributed APs offer extra distance diversity compared to centralized deployment. Besides, it is found that the performance gap between the two networks is more significant for the larger number of reflecting elements $N_r$, which indicates that the RIS-aided cell-free mMIMO system has great development potential for achieving high system capacity.

D. The Energy Efficiency of RIS-Aided Cell-Free Massive MIMO Systems

In this subsection, we consider the energy efficiency maximization design of the RIS-aided cell-free mMIMO system by utilizing the GA method, where the phase shifts of RISs are optimized to maximize the energy efficiency $EE_{	ext{ris}}$ in (32). For comparison, we consider the RIS-free cell-free mMIMO system as the benchmark system [46], [47]. The corresponding energy efficiency $EE_{	ext{rf}}$ can be found in Corollary 5. In Fig. 10, we present the energy efficiency performance versus transmit power $P$ under different numbers of RIS elements. We consider the continuous phase shift case where the RIS reflecting elements have infinite phase resolutions,
i.e., $b \to \infty$. Specifically, the relevant parameters are defined in Table I. We can find a unique transmit power $P_{\text{max}}$ that maximizes the energy efficiency of the systems. The energy efficiency increases with transmit power $P$ when $P$ is smaller than the threshold $P_{\text{max}}$ but declines rapidly when $P$ is larger than this threshold. This behavior is explained by the reason that the achievable rate of systems reaches saturation as the transmit power $P$ increases, and the corresponding energy efficiency will reach a peak or even decrease due to the increase in total power consumption. Moreover, we can see that the RIS can effectively improve the rate performance of the cell-free mMIMO system. It is worth noting that the rate improvement brought by RIS is more significant with the increase of RIS elements, even though the RIS with infinite phase shift accuracy requires a large circuit hardware power consumption. This observation shows the high energy efficiency of RIS-aided systems, meaning that the RIS is a power-effective solution to support high data rate transmission.

VI. CONCLUSION

This paper has investigated and optimized the achievable rate performance in the uplink RIS-aided cell-free mMIMO systems based on the two-timescale scheme. We have considered the general Rician channel model and applied RISs to provide additional channels to the users in the rich scattering region. The CPU has utilized the MRC technique based on the short-time instantaneous CSI for fully centralized processing, while the phase shifts of RISs have been designed based on long-time statistical CSI, which could reduce the channel estimation overhead and computational complexity. On this basis, we first derived the closed-form approximate expression of the achievable rate and provided analytical insights. These valuable insights revealed the impact of various system parameters on the achievable rate and the asymptotic behaviors of the rate, which can serve as clear guidelines for the benefits of the RIS-aided cell-free mMIMO systems. Besides, we have proposed a total power consumption model to obtain the closed-form expression of energy efficiency, and made an analysis based on it. Then, we have designed the optimized phase shifts to maximize users’ sum and minimum rates by exploiting the GA method. Finally, we have provided numerical results to validate the effectiveness and the benefits of integrating RISs into the cell-free mMIMO systems. Our results have demonstrated the correctness of our derived expressions and verified the proposed GA method’s effectiveness by showing the GA’s optimality and convergence behaviors. Also, we have investigated the benefits of the optimized phase shift-based RISs in the cell-free mMIMO system, and revealed that the smaller path-loss exponent of the RIS-AP channels is more conducive to simultaneously maintaining high network capacity and guaranteeing user fairness. Compared to the AP antenna with high power consumption and hardware cost, the passive reflecting elements of RISs are promising to provide the same rate increase for cell-free mMIMO systems at a lower cost. To give a more beneficial analysis, we have presented numerical results to show the high energy efficiency achieved by RIS-aided systems, which implies that the RIS will be power effective. Besides, we have shown the benefits of distributed deployment of APs and RISs for the RIS-aided mMIMO network.

APPENDIX A

As known in Section II, by applying [44, Lemma 1], the uplink achievable rate expression of user $k$ can be approximated as

$$R_k \approx \log_2 \left( 1 + \frac{p_k E_k^\text{(signal)}(\Phi)}{\sum_{i=1,i \neq k}^K p_i I_{ki}(\Phi) + \sigma^2 E_k^\text{(noise)}(\Phi)} \right), \quad (37)$$

where $E_k^\text{(noise)}(\Phi) = \mathbb{E}\left\{ ||g_k + d_k||^2 \right\}$, $E_k^\text{(signal)}(\Phi) = \mathbb{E}\left\{ ||g_k + d_k||^4 \right\}$, $I_{ki}(\Phi) = \mathbb{E}\left\{ ||(g_i + d_i)^H (g_i + d_i)||^2 \right\}$.

To derive the closed-form expression of (37), we can derive the preliminary expression by expanding and simplifying the three mathematical expectation terms. Specially, note that $d_k$, $d_i$ and $g_k$ are independent of each other, $\forall i \neq k$, and $d_k$ is composed of i.i.d. random variables with zero mean. Firstly, the noise term $\mathbb{E}\left\{ ||g_k + d_k||^4 \right\}$ can be written as

$$\mathbb{E}\left\{ ||g_k + d_k||^2 \right\} = \mathbb{E}\left\{ ||g_k + d_k||^4 \right\} = \mathbb{E}\left\{ ||g_k + d_k||^2 \right\} + \mathbb{E}\left\{ ||d_k||^2 \right\} = \mathbb{E}\left\{ ||g_k||^2 \right\} + \sum_{m=1}^M \gamma_{m,k} M_b. \quad (38)$$

Next, we can derive the signal term $\mathbb{E}\left\{ ||g_k + d_k||^4 \right\}$ as

$$\mathbb{E}\left\{ ||g_k + d_k||^4 \right\} = \mathbb{E}\left\{ ||g_k||^4 \right\} + 4 \mathbb{E}\left\{ (\text{Re}\{d_k^H g_k\})^2 \right\} + 2 \mathbb{E}\left\{ ||g_k||^2 ||d_k||^2 \right\}. \quad (39)$$

Assuming that $[g_k]_{m,b} = [g_k]_{m,b} = v_{m,b} + j w_{m,b}$ and $[d_k]_{m,b} = [d_k]_{m,b} = s_{m,b} + j t_{m,b}$, where both $s_{m,b}$ and $t_{m,b}$ independently follow $\mathcal{N}(0, \frac{\sigma^2}{2})$, $\forall b$, we have

$$\mathbb{E}\left\{ (\text{Re}\{d_k^H g_k\})^2 \right\} = \mathbb{E}\left\{ \left( \sum_{m=1}^M \sum_{b=1}^{M_b} s_{m,b} v_{m,b} - t_{m,b} w_{m,b} \right)^2 \right\} = \mathbb{E}\left\{ \sum_{m=1}^M \sum_{b=1}^{M_b} (s_{m,b} v_{m,b} - t_{m,b} w_{m,b})^2 \right\} = \sum_{m=1}^M \sum_{b=1}^{M_b} (s_{m,b} v_{m,b})^2 + (t_{m,b} w_{m,b})^2 \right\}$$
Finally, the interference term can be expanded as

$$
\begin{align*}
&= \sum_{m=1}^{M} \gamma_{m,k} \mathbb{E} \left\{ \sum_{b=1}^{M_b} (v_{m,b})^2 + (w_{m,b})^2 \right\} \\
&= \sum_{m=1}^{M} \gamma_{m,k} \mathbb{E} \left\{ \|g_{m,k}\|^2 \right\},
\end{align*}
$$

(40)

Then, the remaining two terms in (39) can be obtained as

$$
\begin{align*}
\mathbb{E} \left\{ \|d_k\|^2 \right\} &= \mathbb{E} \left\{ \left( \sum_{m=1}^{M} M_b \sum_{b=1}^{M_b} |d_{k,m,b}|^2 \right)^2 \right\} \\
&= \mathbb{E} \left\{ \sum_{m=1}^{M} M_b \left( \sum_{m=1}^{M} \gamma_{m,k} \mathbb{E} \left\{ \|d_{k,m,b}\|^2 \right\} \right)^2 \right\} \\
&= 2 \sum_{m=1}^{M} \gamma_{m,k} \mathbb{E} \left\{ \sum_{m=1}^{M} \gamma_{m,k} \mathbb{E} \left\{ \|d_{k,m,b}\|^2 \right\} \right\} \\
&= \sum_{m=1}^{M} \gamma_{m,k} \mathbb{E} \left\{ \|g_k\|^2 \right\},
\end{align*}
$$

(41)

and

$$
\begin{align*}
\mathbb{E} \left\{ \|g_k\|^2 \|d_k\|^2 \right\} &= \mathbb{E} \left\{ \|g_k\|^2 \right\} \mathbb{E} \left\{ \|d_k\|^2 \right\} \\
&= \sum_{m=1}^{M} \gamma_{m,k} \mathbb{E} \left\{ \|g_k\|^2 \right\}.
\end{align*}
$$

(42)

Substituting (40), (41) and (42) into (39), we obtain the expression of signal term as follows

$$
\begin{align*}
\mathbb{E} \left\{ \|g_k + d_k\|^4 \right\} &= \mathbb{E} \left\{ \|g_k\|^4 \right\} + \sum_{m=1}^{M} \gamma_{m,k} \mathbb{E} \left\{ \|g_k\|^2 \right\} \\
&+ M_b \mathbb{E} \left\{ \|g_k\|^2 \right\} + \sum_{m=1}^{M} \gamma_{m,k} \mathbb{E} \left\{ \|g_k\|^2 \right\} \\
&= \sum_{m=1}^{M} \gamma_{m,k} \mathbb{E} \left\{ \|g_k\|^2 \right\} + \sum_{m=1}^{M} \gamma_{m,k} \mathbb{E} \left\{ \|g_k\|^2 \right\}.
\end{align*}
$$

(43)

Finally, the interference term can be expanded as

$$
\begin{align*}
\mathbb{E} \left\{ (g_k^H + d_k^H) (g_i + d_i) \right\}^2 &= \mathbb{E} \left\{ |b_{k,i} |^2 g_k^H g_i + b_{k,i}^2 d_k^H d_i + c_{k,i}^2 g_k^H d_i + c_{k,i}^2 d_k^H g_i \right\} \\
&= \mathbb{E} \left\{ |b_{k,i}|^2 g_k^H g_i \right\} + \mathbb{E} \left\{ |b_{k,i}|^2 d_k^H d_i \right\} \\
&+ \mathbb{E} \left\{ |c_{k,i}|^2 g_k^H g_i \right\} + \mathbb{E} \left\{ |c_{k,i}|^2 d_k^H d_i \right\} \\
&= \mathbb{E} \left\{ |b_{k,i}|^2 g_k^H g_i \right\} + \mathbb{E} \left\{ |c_{k,i}|^2 d_k^H d_i \right\} \\
&+ \mathbb{E} \left\{ |b_{k,i}|^2 g_k^H d_i \right\} + \mathbb{E} \left\{ |c_{k,i}|^2 g_k^H d_i \right\}.
\end{align*}
$$

(44)

The derivation of the achievable rate in (37) follows by calculating the expectations of \(|g_{m,k}|^2\), \(|g_k|^2\), \(|g_k|^4\), and \(|g_k^H g_i|^2\) in (38) ~ (45) using the decoupling methods, the properties of independence, moments of Gaussian variables, and some algebraic manipulations. Due to the complex form of the channel \(g_k\), the detailed process is very long, cumbersome, and complicated. Therefore, to save space, we only present the final results in the following, and interested readers can refer to our extended version in [56, Appendix A]. Then, the term \(E_k^{(\text{noise})}(\Phi) = \mathbb{E} \{ \|g_k + d_k\|^2 \}\) can be given by

$$
\begin{align*}
\mathbb{E} \left\{ \|g_k + d_k\|^2 \right\} &= \sum_{m=1}^{M} \sum_{n=1}^{N} \sum_{k=1}^{N} \mathbb{E} \left\{ \|g_k^H d_k \|^2 \right\} \\
&= \sum_{m=1}^{M} \gamma_{m,k} \mathbb{E} \left\{ \|g_k\|^2 \right\} + \sum_{m=1}^{M} \gamma_{m,k} \mathbb{E} \left\{ \|g_k\|^2 \right\}. \\
&= \sum_{m=1}^{M} \gamma_{m,k} \mathbb{E} \left\{ \|g_k\|^2 \right\}.
\end{align*}
$$

(45)

and the terms \(E_k^{(\text{signal})}(\Phi) = \mathbb{E} \{ \|g_k + d_k\|^4 \}\) and \(I_{k,i}(\Phi) = \mathbb{E} \{ (g_k + d_k)^H (g_i + d_i) \}^2 \) are respectively given by (47) and (48).

**APPENDIX B**

We assume that the phase shift of each element is randomly generated from \([0, 2\pi]\) and adjusted independently in each updated interval. To obtain the rate expression based on random phase shifts \(\Phi\), we present some properties that will be utilized in the following derivation. Based on the symmetry of the odd function \(\sin(\theta_{n,r})\) with \(\theta_{n,r}\) and the probability density function of \(\theta_{n,r}\), we have

$$
\mathbb{E} \{ e^{\pm j \theta_{n,r}} \} = \mathbb{E} \{ \cos(\theta_{n,r}) \} = 0, \ \forall n, r.
$$

For the expectation

$$
\sum_{m=1}^{M} \sum_{n=1}^{N} \sum_{k=1}^{N} \mathbb{E} \{ f_{m,n,k}^H f_{m,n,k} \}(\Phi) f_{m,n,k}^H f_{m,n,k} \}
$$

with respect to random phase shifts \(\Phi\), using the same method as [32, Corollary 4], we have

$$
\sum_{m=1}^{M} \sum_{n=1}^{N} \sum_{k=1}^{N} \mathbb{E} \{ f_{m,n,k}^H f_{m,n,k} \}(\Phi) f_{m,n,k}^H f_{m,n,k} \}
$$

(49)
with

\[
E \{ f_{m_1,n_1,k}^{H}(\Phi) f_{m_2,n_2,k}(\Phi) \} = E \{ f_{m_1,n_1,k}^{H}(\Phi) \} E \{ f_{m_2,n_2,k}(\Phi) \} \]

\[= E \left\{ \sum_{r_1=1}^{N_r} e^{-j(\zeta_{m_1,n_1,k}^{c} + \theta_{n_1,r_1})} \right\} \times E \left\{ \sum_{r_2=1}^{N_r} e^{j(\zeta_{m_2,n_2,k}^{c} + \theta_{n_2,r_2})} \right\} = \sum_{r_1=1}^{N_r} e^{-j\zeta_{m_1,n_1,k}^{c}} E \{ e^{-j\theta_{n_1,r_1}} \} \sum_{r_2=1}^{N_r} e^{j\zeta_{m_2,n_2,k}^{c}} E \{ e^{j\theta_{n_2,r_2}} \} = 0, \quad (50)
\]

\[E \{ f_{m_1,n_1,k}^{H}(\Phi) f_{m_2,n_2,k}(\Phi) \} = E \left\{ \sum_{r_1=1}^{N_r} e^{-j(\zeta_{m_1,n_1,k}^{c} + \theta_{n_1,r_1})} \sum_{r_2=1}^{N_r} e^{j(\zeta_{m_2,n_2,k}^{c} + \theta_{n_2,r_2})} \right\} = E \left\{ \sum_{r_1=1}^{N_r} e^{-j(\zeta_{m_1,n_1,k}^{c} + \theta_{n_1,r_1})} \right\} E \left\{ \sum_{r_2=1}^{N_r} e^{j(\zeta_{m_2,n_2,k}^{c} + \theta_{n_2,r_2})} \right\} = 0, \quad (51)\]
where the derivations exploit the independence and the zero-mean properties of $\Phi_n$ and $\theta_n$, $\forall n, r$.

For the expectation $\sum_{m_1=1}^{M} \sum_{m_2=1}^{N} \sum_{n_1=1}^{N} \sum_{n_2=1}^{N} \mathbb{E}\{f_{m_1,n_1,k}(\Phi) f_{m_2,n_2,i}(\Phi)\}$, we have

$$
\sum_{m_1=1}^{M} \sum_{m_2=1}^{N} \sum_{n_1=1}^{N} \sum_{n_2=1}^{N} \mathbb{E}\{f_{m_1,n_1,k}(\Phi) f_{m_2,n_2,i}(\Phi)\} = \sum_{m_1=1}^{M} \sum_{m_2=1}^{N} \sum_{n_1=1}^{N} \sum_{n_2=1}^{N} \mathbb{E}\{f_{m_1,n_1,k}(\Phi) f_{m_2,n_2,r}(\Phi)\},
$$

(52)

where

$$
\mathbb{E}\{f_{m_1,n_1,k}(\Phi) f_{m_2,n_2,i}(\Phi)\} = \mathbb{E}\{f_{m_1,n_1,k}(\Phi)\} \mathbb{E}\{f_{m_2,n_2,i}(\Phi)\} = \mathbb{E}\{f_{m_1,n_1,k}(\Phi)\} \mathbb{E}\{f_{m_2,n_2,i}(\Phi)\} = 0,
$$

(54)

$$
\mathbb{E}\{f_{m_1,n_1,k}(\Phi) f_{m_2,n_2,i}(\Phi)\} = \sum_{r=1}^{N_r} e^{j(\zeta_{m_2,n_2,i} - \zeta_{m_1,n_1,k})} = a_{N_r}^H(m_2, n_2) \text{diag}\{\Pi_{n_1}\} \text{diag}\{\Gamma_{n_1}^H\} a_{N_r}(m_1, n_1).
$$

(55)

Based on the independence and the zero-mean properties of $\Phi_n$ and $\theta_n$, $\forall n, r$, we can ignore the zero term in the following derivation. For the expectation $\sum_{m_1=1}^{M} \sum_{m_2=1}^{N} \sum_{n_1=1}^{N} \sum_{n_2=1}^{N} \mathbb{E}\{f_{m_1,n_1,k}(\Phi) f_{m_2,n_2,i}(\Phi)\}$, we have

$$
\sum_{m_1=1}^{M} \sum_{m_2=1}^{N} \sum_{n_1=1}^{N} \sum_{n_2=1}^{N} \mathbb{E}\{f_{m_1,n_1,k}(\Phi) f_{m_2,n_2,i}(\Phi)\} = \sum_{m_1=1}^{M} \sum_{m_2=1}^{N} \sum_{n_1=1}^{N} \sum_{n_2=1}^{N} \mathbb{E}\{f_{m_1,n_1,k}(\Phi) f_{m_2,n_2,i}(\Phi)\},
$$

(53)

$$
\mathbb{E}\{f_{m_1,n_1,k}(\Phi) f_{m_2,n_2,i}(\Phi)\} = \mathbb{E}\{f_{m_1,n_1,k}(\Phi)\} \mathbb{E}\{f_{m_2,n_2,i}(\Phi)\} = \mathbb{E}\{f_{m_1,n_1,k}(\Phi)\} \mathbb{E}\{f_{m_2,n_2,i}(\Phi)\} = 0.
$$

(54)

$$
\mathbb{E}\{f_{m_1,n_1,k}(\Phi) f_{m_2,n_2,i}(\Phi)\} = \sum_{r=1}^{N_r} e^{j(\zeta_{m_2,n_2,i} - \zeta_{m_1,n_1,k})} = a_{N_r}^H(m_2, n_2) \text{diag}\{\Pi_{n_1}\} \text{diag}\{\Gamma_{n_1}^H\} a_{N_r}(m_1, n_1).
$$

(55)
\[
\begin{align*}
&\times \sum_{m=1}^{M} \sum_{n=1}^{N} \sum_{k=1}^{N} \mathbb{E} \left\{ |f_{m,n,k}(\Phi)|^2 |f_{m,n,i}(\Phi)|^2 \right\} \\
&= \sum_{m=1}^{M} \sum_{n=1}^{N} \sum_{k=1}^{N} \mathbb{E} \left\{ |f_{m,n,k}(\Phi)|^2 |f_{m,n,i}(\Phi)|^2 \right\} \\
&\times \sum_{m=1}^{M} \sum_{n=1}^{N} \sum_{k=1}^{N} \mathbb{E} \left\{ f_{m,n,k}(\Phi) f_{m,n,i}(\Phi) \right\} \\
&\times f_{m,n,k}(\Phi) f_{m,n,i}(\Phi) \},
\end{align*}
\]

where

\[
\begin{align*}
&\mathbb{E} \left\{ |f_{m,n,k}(\Phi)|^2 |f_{m,n,i}(\Phi)|^2 \right\} \\
&= \mathbb{E} \left\{ |f_{m,n,k}(\Phi)|^2 \right\} \mathbb{E} \left\{ |f_{m,n,i}(\Phi)|^2 \right\} = N_r^2,
\end{align*}
\]

\[
\begin{align*}
&\mathbb{E} \left\{ f_{m,n,k}(\Phi) f_{m,n,i}(\Phi) \right\} \\
&= \sum_{r=1}^{N_r} \mathbb{E} \left\{ e^{j\left(\gamma_{m,n,k} - \gamma_{m,n,i}\right)} \right\} + N_r^2 - N_r
\end{align*}
\]

and all the terms

\[
\begin{align*}
&\mathbb{E} \left\{ f_{m,n,k}(\Phi) f_{m,n,i}(\Phi) f_{m,n,k}(\Phi) f_{m,n,i}(\Phi) \right\} \\
&= \mathbb{E} \left\{ f_{m,n,k}(\Phi) f_{m,n,i}(\Phi) \right\} \mathbb{E} \left\{ f_{m,n,k}(\Phi) f_{m,n,i}(\Phi) \right\} \\
&= \sum_{r=1}^{N_r} \mathbb{E} \left\{ e^{j\left(\gamma_{m,n,k} - \gamma_{m,n,i}\right)} \right\} + N_r^2 - N_r
\end{align*}
\]

Similarly, the expectation \( \sum_{m=1}^{M} \sum_{n=1}^{N} \sum_{k=1}^{N} \mathbb{E} \left\{ f_{m,n,k}(\Phi) f_{m,n,i}(\Phi) \right\} \) can be derived as

\[
\begin{align*}
&\sum_{m=1}^{M} \sum_{n=1}^{N} \sum_{k=1}^{N} \sum_{m=1}^{M} \sum_{n=1}^{N} \sum_{k=1}^{N} \mathbb{E} \left\{ f_{m,n,k}(\Phi) f_{m,n,i}(\Phi) \right\} \\
&\times f_{m,n,k}(\Phi) f_{m,n,i}(\Phi) \\
&= \sum_{m=1}^{M} \sum_{n=1}^{N} \sum_{k=1}^{N} \mathbb{E} \left\{ f_{m,n,k}(\Phi) f_{m,n,i}(\Phi) \right\} \\
&\times f_{m,n,k}(\Phi) f_{m,n,i}(\Phi) \\
&\times f_{m,n,k}(\Phi) f_{m,n,i}(\Phi) \},
\end{align*}
\]

are not on the order of \( \mathcal{O}(N_r^2) \).

(61)
Then, based on (49) \( \sim (66) \), we can substitute terms involving \( \Phi \) in (13) with their expectation. When \( N_r \to \infty \) or \( M_b \to \infty \), we can ignore the insignificant terms which are not on the order of \( O(N_r^2) \) or \( O(M_b^2) \). After some direct simplifications, we can derive the rate expression in (29).

**REFERENCES**

[1] E. Björnson, J. Hoydis, and L. Sanguinetti, “Massive MIMO networks: Spectral, energy, and hardware efficiency,” Found. Trends® Signal Process., vol. 11, nos. 3–4, pp. 154–655, 2017.

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

[3] Z. H. Shaik, E. Björnson, and E. G. Larsson, “Cell-free massive MIMO with radio stripes and sequential uplink processing,” in Proc. IEEE Int. Conf. Commun. Workshops, Jun. 2020, pp. 1–6.

[4] S. Elhoushy, M. Ibrahim, and W. Hamouda, “Cell-free massive MIMO: A survey,” IEEE Commun. Surveys Tuts., vol. 24, no. 1, pp. 492–523, 1st Quart., 2022.

[5] 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. (ICNC), Feb. 2019, pp. 883–887.

[6] M. Attarifar, A. Abbassaf, and A. Lozano, “Modified conjugate beamforming for cell-free massive MIMO,” IEEE Wireless Commun. Lett., vol. 8, no. 6, pp. 616–619, Apr. 2019.

[7] 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.

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

[9] M. Di Renzo et al., “Smart radio environments empowered by reconfigurable intelligent surfaces: How it works, state of research, and the road ahead,” IEEE J. Sel. Areas Commun., vol. 38, no. 11, pp. 2450–2525, Nov. 2020.

[10] M. D. Renzo et al., “Smart radio environments empowered by reconfigurable AI meta-surfaces: An idea whose time has come,” EURASIP J. Wireless Commun. Netw., vol. 2019, no. 1, pp. 1–20, Dec. 2019.

[11] 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, 2020.

[12] 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. 14960–14973, Dec. 2020.

[13] 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.

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

[15] K. Zhi, C. Pan, H. Ren, and K. Wang, “Ergodic rate analysis of reconfigurable intelligent surface-aided massive MIMO systems with ZF detectors,” IEEE Commun. Lett., vol. 26, no. 2, pp. 264–268, Feb. 2022.

[16] F. Yang, J. Dai, C. Pan, S. Hong, H. Ren, and K. Wang, “Robust beamforming design for RIS-aided NOMA networks with imperfect channels,” in Proc. IEEE 95th Veh. Technol. Conf., Jun. 2022, pp. 1–6.

[17] S. Hong, C. Pan, H. Ren, K. Wang, K. K. Chai, and A. Nallanathan, “Robust transmission design for intelligent reflecting surface-aided secure communication systems with imperfect cascaded CSI,” IEEE Trans. Wireless Commun., vol. 20, no. 4, pp. 2487–2501, Apr. 2021.

[18] Z. Peng, T. Li, C. Pan, H. Ren, and J. Wang, “RIS-aided D2D communications relying on statistical CSI with imperfect hardware,” IEEE Commun. Lett., vol. 26, no. 2, pp. 473–477, Feb. 2022.

[19] G. Zhou, C. Pan, H. Ren, K. Wang, M. Elkashlan, and M. D. Renzo, “Stochastic learning-based robust beamforming design for RIS-aided millimeter-wave systems in the presence of random blockages,” IEEE Trans. Veh. Technol., vol. 70, no. 1, pp. 1057–1061, Jan. 2021.

[20] Y. Zhang et al., “Beyond cell-free MIMO: Energy efficient reconfigurable surface aided cell-free MIMO communications,” IEEE Trans. Cognit. Commun. Netw., vol. 7, no. 2, pp. 412–426, Jun. 2021.

[21] E. Shi et al., “Spatially correlated reconfigurable intelligent surface-aided cell-free massive MIMO systems,” IEEE Trans. Veh. Technol., vol. 71, no. 8, pp. 9073–9077, Aug. 2022.

[22] E. Shi, J. Zhang, Y. Du, Z. Wang, B. Ai, and D. W. K. Ng, “Spatially correlated RIS-aided CF massive MIMO systems with generalized MR combining,” IEEE Trans. Veh. Technol., vol. 71, no. 10, pp. 11245–11250, Oct. 2022.

[23] X. Zhang, T. Liang, K. An, H. Yang, and C. Niu, “Secure transmission in RIS-aided cell-free massive MIMO system with low resolution ADCs/DACs,” in Proc. IEEE Wireless Commun. Netw. Conf. (WCNC), Apr. 2022, pp. 339–344.

[24] S. Elhoushy, M. Ibrahim, and W. Hamouda, “Exploiting RIS for limiting information leakage to active eavesdropper in cell-free massive MIMO,” IEEE Wireless Commun. Lett., vol. 11, no. 3, pp. 443–447, Mar. 2022.

[25] M. Bashar, K. Cumanan, A. G. Burr, P. Xiao, and M. Di Renzo, “On the performance of reconfigurable intelligent surface-aided cell-free massive MIMO uplink,” in Proc. IEEE Global Commun. Conf., Dec. 2020, pp. 1–6.

[26] H. Ge, N. Garg, and T. Ratnarajah, “Generalized superimposed channel estimation for uplink RIS-aided cell-free massive MIMO systems,” in Proc. IEEE Wireless Commun. Netw. Conf. (WCNC), Apr. 2022, pp. 405–410.

[27] Z. Wang, L. Liu, and S. Cui, “Channel estimation for intelligent reflecting surface assisted multiuser communications: Framework, algorithms, and analysis,” IEEE Trans. Wireless Commun., vol. 19, no. 10, pp. 6607–6620, Oct. 2020.

[28] Q.-U.-A. Nadeem, H. Alwazani, A. Kammoun, A. Chaaban, M. Debbah, and M.-S. Alouini, “Intelligent reflecting surface-aided multi-user MISO communication: Channel estimation and beamforming design,” IEEE Open J. Commun. Soc., vol. 1, pp. 661–680, 2020.

[29] K. Zhi, C. Pan, H. Ren, and K. Wang, “Statistical CSI-based design for reconfigurable intelligent surface-aided massive MIMO systems with direct links,” IEEE Wireless Commun. Lett., vol. 10, no. 5, pp. 1128–1132, May 2021.

[30] K. Zhi et al., “Two-timescale design for reconfigurable intelligent surface-aided massive MIMO systems with imperfect CSI,” IEEE Trans. Inf. Theory, vol. 69, no. 5, pp. 3001–3033, May 2023.

[31] M.-M. Zhao, A. Liu, Y. Wan, and R. Zhang, “Two-timescale beamforming optimization for intelligent reflecting surface aided multiuser communication with QoS constraints,” IEEE Trans. Wireless Commun., vol. 20, no. 9, pp. 6179–6194, Sep. 2021.

[32] K. Zhi, C. Pan, H. Ren, and K. Wang, “Power scaling law analysis and phase shift optimization of RIS-aided massive MIMO systems with statistical CSI,” IEEE Trans. Commun., vol. 70, no. 5, pp. 3558–3574, May 2022.

[33] T. Van Chien, H. Q. Ngo, S. Chatzinotas, M. Di Renzo, and B. Ottersten, “Reconfigurable intelligent surface-assisted cell-free massive MIMO systems over spatially-correlated channels,” IEEE Trans. Wireless Commun., vol. 21, no. 7, pp. 5106–5128, Jul. 2022.
Y. Han, W. Tang, S. Jin, C.-K. Wen, and X. Ma, “Large intelligent surface-assisted wireless communication exploiting statistical CSI,” *IEEE Trans. Veh. Technol.*, vol. 68, no. 8, pp. 8238–8242, Aug. 2019.

Y. Zhang, J. Zhang, M. D. Renzo, H. Xiao, and B. Ai, “Performance analysis of RIS-assisted systems with practical phase shift and amplitude response,” *IEEE Trans. Veh. Technol.*, vol. 70, no. 5, pp. 4501–4511, May 2021.

Z. Huang, X. Cheng, and X. Yin, “A general 3D non-stationary 6G channel model with time-space consistency,” *IEEE Trans. Commun.*, vol. 70, no. 5, pp. 3436–3450, May 2022.

Y. Zheng et al., “Ultra-massive MIMO channel measurements at 5.3 GHz and a general 6G channel model,” *IEEE Trans. Veh. Technol.*, vol. 72, no. 1, pp. 20–34, Jan. 2023.

I. Yildirim, A. Uyrys, and E. Basar, “Modeling and analysis of reconfigurable intelligent surfaces for indoor and outdoor applications in future wireless networks,” *IEEE Trans. Commun.*, vol. 69, no. 2, pp. 1290–1301, Feb. 2021.

H. Zhu and J. Wang, “Chunk-based resource allocation in OFDMA systems—Part I: Chunk allocation,” *IEEE Trans. Commun.*, vol. 57, no. 9, pp. 2734–2744, Sep. 2009.

H. Zhu and J. Wang, “Chunk-based resource allocation in OFDMA systems—Part II: Joint chunk, power and bit allocation,” *IEEE Trans. Commun.*, vol. 60, no. 2, pp. 499–509, Mar. 2012.

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

E. Björnson and L. Sanguinetti, “Making cell-free massive MIMO competitive with MMSE processing and centralized implementation,” *IEEE Trans. Wireless Commun.*, vol. 19, no. 1, pp. 77–90, Jan. 2020.

Q. Zhang, S. Jin, K.-K. Wong, H. Zhu, and M. Matthaiou, “Power scaling of uplink massive MIMO systems with arbitrary-rank channel means,” *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 966–981, Oct. 2014.

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.

M. Bashar et al., “Uplink spectral and energy efficiency of cell-free massive MIMO with optimal uniform quantization,” *IEEE Trans. Commun.*, vol. 69, no. 1, pp. 223–245, Jan. 2021.

H. Q. Ngo, L.-N. Tran, T. Q. Duong, M. Matthaiou, and E. G. Larsson, “On the total energy efficiency of cell-free massive MIMO,” *IEEE Trans. Green Commun. Netw.*, vol. 2, no. 1, pp. 25–39, Mar. 2018.

Z. Peng, T. Li, C. Pan, H. Ren, W. Xu, and M. D. Renzo, “Analysis and optimization for RIS-aided multi-pair communications relying on statistical CSI,” *IEEE Trans. Veh. Technol.*, vol. 70, no. 4, pp. 3897–3901, Apr. 2021.

M. Mitchell, *An Introduction to Genetic Algorithms*. Cambridge, MA, USA: MIT Press, 1998.

M. Esksandari, K. Zhi, H. Zhu, C. Pan, and J. Wang, “Two-timescale design for RIS-aided cell-free massive MIMO systems with imperfect CSI,” 2023, arXiv:2304.02606.

J. Dai et al., “Two-timescale transmission design for RIS-aided cell-free massive MIMO systems,” 2022, arXiv:2210.08514.

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.

Jianxin Dai (Member, IEEE) received the B.S. degree from the Mathematics Department, Nanjing Normal University, Nanjing, China, in 1995, the M.S. degree in communications science from the Nanjing University of Posts and Telecommunications, Nanjing, in 2007, and the Ph.D. degree in electronic engineering from the National Mobile Communications Research Laboratory, Southeast University, Nanjing, in 2014. From 2015 to 2017, he held a post-doctoral position with the Nanjing University of Posts and Telecommunications.

From 2016 to 2017, he was an Academic Visitor with the University of Kent, U.K. Since 2009, he has been an Associate Professor with the Nanjing University of Posts and Telecommunications. His current research interests include reconfigurable intelligent surfaces (RISs), mm-wave communications, massive MIMO systems, and cognitive radio networks.

Jing Ge received the B.E. degree from the School of Communication Engineering, East China Jiaotong University, China, in 2021. He is currently pursuing the M.E. degree with the College of Telecommunications and Information Engineering, Nanjing University of Posts and Telecommunications, China. His research interests include reconfigurable intelligent surface (RIS) and cell-free massive multiple-input–multiple-output (mMIMO).

M. Mitchell, *An Introduction to Genetic Algorithms*. Cambridge, MA, USA: MIT Press, 1998.

Kangda Zhi received the B.E. degree from the School of Communication and Information Engineering, Shanghai University (SHU), Shanghai, China, in 2017, the M.E. degree from the School of Information Science and Technology, University of Science and Technology of China (USTC), Hefei, China, in 2020, and the Ph.D. degree from the School of Electronic Engineering and Computer Science, Queen Mary University of London, U.K., in 2023. His research interests include reconfigurable intelligent surface (RIS), massive MIMO, and near-field communications. He received the Exemplary Reviewer Certificate of the IEEE WIRELESS COMMUNICATIONS LETTERS in 2021 and 2022.
mainly include reconfigurable intelligent surfaces (RISs), intelligent reflection surfaces (IRSs), ultra-reliable low latency communication (URLLC), machine learning, UAV, the Internet of Things, and mobile edge computing. He serves as a TPC Member for numerous conferences, such as IEEE ICC and IEEE GLOBECOM, and the Student Travel Grant Chair for IEEE ICC 2019. He received the IEEE ComSoc Leonard G. Abraham Prize in 2022 and the IEEE ComSoc Asia–Pacific Outstanding Young Researcher Award in 2022. He is the Workshop Organizer of the IEEE ICC 2021 on the topic of Reconfigurable Intelligent Surfaces for Next Generation Wireless Communications (RIS for 6G Networks) and the IEEE GLOBECOM 2021 on the topic of Reconfigurable Intelligent Surfaces for Future Wireless Communications. He is currently the Workshops and Symposia Officer for Reconfigurable Intelligent Surfaces Emerging Technology Initiative. He is the Workshop Chair for IEEE WCNC 2024 and the TPC Co-Chair for IEEE ICCT 2022. He is currently an Editor of IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, IEEE WIRELESS COMMUNICATION LETTERS, IEEE COMMUNICATIONS LETTERS, and IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS. He serves as the Guest Editor for IEEE JOURNAL ON SELECTED TOPICS IN SIGNAL PROCESSING Special Issue on “xURLLC in 6G: Next Generation Ultra-Reliable and Low-Latency Communications.” He also serves as a Leading Guest Editor for IEEE JOURNAL ON SELECTED TOPICS IN COMMUNICATIONS Special Issue on “Advanced Signal Processing for Reconfigurable Intelligent Surface-Aided 6G Networks,” IEEE VEHICULAR TECHNOLOGY Magazine Special Issue on “Backscatter and Reconfigurable Intelligent Surface Empowered Wireless Communications in 6G,” IEEE OPEN JOURNAL OF VEHICULAR TECHNOLOGY Special Issue on “Reconfigurable Intelligent Surface Empowered Wireless Communications in 6G and Beyond,” and IEEE ACCESS Special Issue on “Reconfigurable Intelligent Surface Aided Communications for 6G and Beyond.”

Zaichen Zhang (Senior Member, IEEE) was born in Nanjing, China, in 1975. He received the B.S. and M.S. degrees in electrical and information engineering from Southeast University, Nanjing, in 1996 and 1999, respectively, and the Ph.D. degree in electrical and electronic engineering from The University of Hong Kong, Hong Kong, China, in 2002. From 2002 to 2004, he was a Post-Doctoral Fellow at the National Mobile Communications Research Laboratory, Southeast University. He joined the School of Information Science and Engineering, Southeast University, in 2004, where he is currently a Chair Professor. He has published over 300 papers and issued 100 patents. His current research interests include 6G wireless communications, optical wireless communications, and quantum information.