An Optimization Paradigm for Sustainable RIS-Assisted Cell-Free Networks

Ismail Hburi, Ahmed Abdulameer, Nibras Abbas, Basim Al-Shammari, and Hasan Khazaal

***Abstract*—Reconfigurable intelligent surface (RIS)-assisted wireless networks offer significant performance gains; however, energy-efficiency maximization in such systems remains fundamentally challenging due to the coupled active–passive beamforming design and the fractional nature of the objective function. Existing studies typically rely on semidefinite relaxation or optimize surrogate spectral-efficiency metrics, without establishing theoretical equivalence to the original fractional formulation or providing formal convergence guarantees. This paper addresses this gap by formulating joint transmit and RIS phase optimization as a non-convex fractional program with unit-modulus constraints, which is NP-hard. We develop a unified optimization framework that integrates Dinkelbach fractional programming, a rigorously derived weighted minimum mean square error (WMMSE) reformulation, second-order cone programming (SOCP), and Riemannian manifold optimization within a provably convergent alternating structure. Unlike prior approaches, the proposed method preserves stationary-point equivalence to the original fractional objective and is proven to converge to a Karush–Kuhn–Tucker (KKT) point under standard constraint qualification conditions. A detailed per-stage computational complexity analysis is provided, together with empirical runtime evaluation. Simulation results, averaged over 1000 independent Monte Carlo realizations with 95% confidence intervals, demonstrate up to 40.9% improvement in energy efficiency compared with state-of-the-art benchmark schemes under identical system assumptions with** 95% **confidence interval (CI)** $[38.7\%, 43.1\%]$**. Sensitivity analysis further confirms robustness with respect to RIS size, user density, transmit power, and phase quantization resolution. These results establish both the theoretical soundness and practical scalability of the proposed framework for energy-efficient RIS-assisted communications.**

***Index Terms*—Cell-free network, Energy efficiency optimization, Reconfigurable intelligent surface, Riemannian manifold.**

I. Introduction

THE multiple-antenna concept has received a lot of interest in recent decades [1], [2]. Nevertheless, as a result of the cell-centric approach, the reliability of multi-cell multiple-antenna networks is often restricted by the interference across different cells. In this regard, a user-centric approach (cell-free paradigm) has been presented to overcome this challenge, where many randomly deployed base stations cooperate (without cell borders) to reach all users at the same time [3], [4]. However, due to the high expenses of both power and hardware resources, the classic cell-free scheme necessitates the widespread distribution of BSs, which results in an inadequate energy efficiency. Therefore, the RIS concept (reconfigurable intelligent surface) has been developed to address this problem, offering a possible cost-effective approach that provides advantageous transmission circumstances for base stations and clients [5], [6], [7]. Intelligent surfaces may provide directing rays from base stations to users and reflect waves thanks to a considerable number of elements in the reflecting surface whose phase angles are managed by a straightforward control board [8]. Intelligent surface needs no additional hardware installation, such as intricate equipment of discrete phase-shifter, which significantly reduces power consumption and signal processing cost in contrast to high dimension phased-array antenna-schemes with unavoidable energy feeding [9]. Therefore, a lesser amount of energy cost is needed to obtain the similar performance as traditional cell-free networks. To put it another way, intelligent surface gives cell-free technologies another aspect to improve their energy performance.

The throughput optimization of reflecting surfaces-assisted networks has been extensively studied, yet existing works exhibit several critical limitations that motivate our research. Wang et al. [10] investigated transmit and passive beamforming design for multi-IRS assisted cell-free MIMO networks. The optimization objective is spectral efficiency rather than energy efficiency. While their work provides valuable insights into cooperative beamforming, it suffers from some shortcomings: 1) the phase shifts are treated as continuous variables, ignoring the practical limitations of discrete phase shifters, and 3) the power consumption model does not account for the dependence of RIS hardware power on phase-shifter resolution, leading to potentially unrealistic energy consumption estimates.

Pan et al. [5] examined multicell MIMO communications relying on intelligent reflecting surfaces, focusing on cell-edge performance improvement. Although their work demonstrates the potential of RIS for interference mitigation, critical gaps remain: 1) the energy efficiency perspective is completely absent, with the work centered on sum-rate maximization; 2) the proposed algorithm relies on semidefinite relaxation (SDR), which scales poorly with the number of RIS elements and users; and 3) minimum QoS constraints per user are not incorporated, potentially leading to unfair resource allocation among users. Furthermore, their channel model assumes perfect CSI without discussing the practical implications of CSI acquisition overhead on the achievable performance.

Manuscript received February 10, 2026; revised February 23, 2026. Date of publication August 17, 2026. Date of current version August 17, 2026. The associate editor prof. Zoran Blažević has been coordinating the review of this manuscript and approved it for publication.

Authors are with the Electrical Engineering Department, College of Engineering, Wasit University, Wasit, Iraq (e-mails: {isharhan, Ahmed.Abdulameer, niabbas, bkhalf, hfahad}@uowasit.edu.iq).

Digital Object Identifier (DOI): 10.24138/jcomss-2026-0059

Zhang and Dai [4] explored capacity improvement in wideband RIS-aided cell-free networks. While their work advances the understanding of RIS in wideband scenarios, it has several limitations that our work addresses: 1) the energy efficiency metric is not considered, despite its critical importance for sustainable network deployment; 2) the proposed framework cannot handle discrete phase-shift constraints, assuming instead infinite-resolution phase shifters; 3) the power consumption model is overly simplistic, neglecting the static power consumption of BSs and users as well as the resolution-dependent power consumption of RIS elements; and 4) the convergence properties of their iterative algorithm are not formally established.

Huang et al. [4] investigated energy efficiency in RIS-assisted multiuser MISO systems, which is closer to our work in objective. However, their work differs fundamentally from ours in several aspects: 1) the network architecture is a single-cell MISO system rather than a cell-free network with multiple cooperating BSs; 2) the alternating optimization framework does not integrate fractional programming with WMMSE equivalence, potentially leading to suboptimal solutions; 3) the phase shifts are assumed to be continuous, ignoring the practical constraints of finite-bit phase shifters; and 4) the minimum rate constraints per user are not explicitly enforced during optimization.

In contrast to the aforementioned works, our research addresses these gaps through several key innovations. First, unlike [10] and [5] that focus on spectral efficiency, we target energy efficiency maximization, which is essential for sustainable network operation. Second, unlike [4] and [14], we explicitly incorporate practical discrete phase-shift constraints into the optimization framework, ensuring implementation feasibility. Third, while [10] uses SDR with high computational complexity, our proposed ME_RM algorithm combines Dinkelbach, WMMSE, and Riemannian manifold optimization in a computationally efficient manner. Fourth, unlike all the above works, the current work incorporates per-user minimum rate constraints to ensure QoS fairness. Fifth, our power consumption model uniquely accounts for the resolution-dependent power consumption of RIS elements, providing more realistic energy efficiency estimates.

Building upon this critical analysis, our work makes the following novel contributions that distinguish it from existing studies:

1) formulating a comprehensive EE maximization problem for RIS-assisted cell-free networks that simultaneously incorporates: 1) discrete phase-shift constraints with finite-bit resolution, 2) per-user minimum rate constraints for QoS fairness, and 3) a realistic power consumption model that captures the resolution-dependent power consumption of RIS elements. This holistic formulation addresses practical implementation aspects overlooked in prior works [3], [4], [5], [10], [14].
2) development of a novel three-layered algorithmic framework (ME_RM) that tightly integrates Dinkelbach fractional programming, WMMSE equivalence with SOCP formulation, and Riemannian manifold optimization. Unlike [10] which uses SDR with high complexity $\mathcal{O}((MBN)^{3.5}NRK^{4.5})$, our approach achieves lower complexity $\mathcal{O}((MB)^{2.5}K^{4.5} + MRNK^2)$ while effectively handling discrete constraints. Unlike [5] which treats phase shifts continuously, our manifold approach naturally accommodates the unit-modulus constraint and provides a principled method for subsequent quantization to discrete values.
3) providing a rigorous convergence analysis of the proposed alternating optimization framework, establishing that the objective function monotonically increases and converges to a stationary point. This theoretical grounding is absent in related works such as [4] and [14].
4) Through extensive numerical simulations with sensitivity analysis and statistical validation, the proposed scheme achieves up to 40.9% improvement in energy efficiency compared to baseline approaches. Unlike prior works that report results at single operating points, this work provides comprehensive analysis across varying numbers of users, RIS elements, phase-shifter resolutions, and transmit power levels, with confidence intervals to establish statistical significance.

The rest of this article is organized as follows. Section II displays the Research Method which includes: System Model, Problem Formulation, Energy Maximization Algorithm Design, Transmit Beamforming Optimization, and Intelligent surfaces' Passive Beamforming. Section III analyzes and discusses the obtained results. Section IV introduces the Computational Complexity Analysis.Section V summarizes the conclusion of this study.

II. Research Method

This study examines a cell-free network situation in which a number (B) of base stations (BS) with M antennas each serve a K single-antenna user equipment (UE). In particular, this section explains the downlink (DL) analysis across a single coherence block. Multiple interfaced BSs and their corresponding Reconfigurable intelligent surfaces (RISs) are suggested to collaboratively service the kth UE in the proposed design (user-centric approach).

A. System Model

As seen in Figure 1, each BS is supported by R numbers of RISs, each surface is made up of N passive reflective segments, and the Central Processing Unit (CPU) controls these RIS units via wired or wireless back-haul. The TDD approach, which consists of two stages for each coherence period—the uplink training phase (UL) and the downlink information dispatch (DL) manages the transmission from BSs to UEs. Table I defines the complex matrices of different channels in the system with their dimensions.

B. Standardized Notation and Matrix Dimensions

A time-division duplex (TDD) operation is considered in this work, where the downlink (DL) channels can be estimated from uplink (UL) pilot sequences transmitted by the UEs, leveraging channel reciprocity. This is a standard assumption in cell-free massive MIMO literature [3]], [4]. Perfect CSI at the CPU: The central processing unit (CPU) is assumed to have perfect knowledge of all instantaneous channel realizations. This assumption allows us to focus on the beamforming optimization problem without the confounding effects of estimation errors. While practical systems would operate with imperfect CSI, this idealized assumption provides an upper bound on performance and is common in related works [5], [10], [14]. The channel matrices in Table 1 follow a Rician fading model, which accounts for both line-of-sight (LoS) and non-line-of-sight (NLoS) components. Specifically, each channel matrix $\mathbf{X} \in \{\mathbf{F}_{bk}, \mathbf{G}_{br}, \mathbf{H}_{rk}\}$ is modeled as, $\mathbf{X} = \sqrt{\frac{\kappa}{1+\kappa}}\mathbf{X}^{\text{LoS}} + \sqrt{\frac{1}{1+\kappa}}\mathbf{X}^{\text{NLoS}}, \quad \kappa = 4$, where κ is the Rician factor, $\mathbf{X}^{\text{LoS}}$ is the deterministic line-of-sight component, and $\mathbf{X}^{\text{NLoS}}$ is the Rayleigh fading component with i.i.d. $\mathcal{CN}(0,1)$ entries. This model is consistent with prior works [24], [25] and provides a realistic representation of wireless propagation environments. The distance-dependent path loss is incorporated as, $PL(d) = PL_0 + 10\alpha\log_{10}(d/d_0) + \mathcal{N}(0, \sigma^2_{\text{shadow}})$, where PL_0 is the path loss at reference distance d_0, α is the path loss exponent, and σ^2_{shadow} accounts for shadow fading variations. The specific values used in our simulations are detailed in Section IV.

TABLE I
SUMMARY OF MATRIX DIMENSIONS

Matrix	*Definition*	*Dimensions*
$\mathbf{F}_{bk}$	Direct BS-UE channel	$\mathbb{C}^{M\times 1}$
$\mathbf{F}_{k}$	Aggregate direct channel	$\mathbb{C}^{M\times 1}$
$\mathbf{G}_{br}$	BS-RIS channel	$\mathbb{C}^{M\times N}$
$\mathbf{G}_{r}$	Aggregate BS-RIS channel	$\mathbb{C}^{M\times N}$
$\mathbf{H}_{rk}$	RIS-UE channel	$\mathbb{C}^{N\times 1}$
$\mathbf{H}_{k}$	Aggregate RIS-UE channel	$\mathbb{C}^{N\times 1}$
$\mathbf{\Phi}_{r}$	RIS phase-shift matrix	$\mathbb{C}^{N\times N}$
$\mathbf{\Phi}$	Aggregate phase-shift matrix	$\mathbb{C}^{RN\times RN}$
$\mathbf{V}_{bk}$	Active beamforming vector	$\mathbb{C}^{M\times 1}$
$\mathbf{V}_{k}$	Aggregate beamforming vector	$\mathbb{C}^{BM\times 1}$
$\mathbf{V}$	Complete beamforming matrix	$\mathbb{C}^{BM\times K}$
$\mathbf{Q}_{k}$	Effective cascaded channel	$\mathbb{C}^{BM\times 1}$
$\mathbf{Q}$	Aggregate effective channel	$\mathbb{C}^{BM\times K}$

C. CSI Acquisition Overhead Discussion

While this work assumes perfect CSI to focus on the optimization framework, we acknowledge that CSI acquisition in RIS-assisted systems introduces overhead that affects the achievable rates. Following the analysis in [8], the effective achievable rate accounting for pilot overhead can be expressed as, $R_k^{\text{eff}} = \left(1 - \frac{T_{\text{pilot}}}{T_{\text{coherence}}}\right) R_k$, where T_{pilot} is the number of symbols dedicated to pilot transmission and $T_{\text{coherence}}$ is the coherence block length. For RIS-assisted links, the pilot overhead scales with the number of RIS elements N and the number of users K. Our framework can be extended to incorporate this overhead by replacing R_k with R_k^{eff} in the objective function, which we identify as an important direction for future work.

The acquired signal's phase-shifting matrix at the intelligent surface is represented by the diagonal matrix $\mathbf{\Phi}_r \in \mathbb{C}^{N\times N}$ [15],

$$\mathbf{\Phi}_r = \text{diag}\{\phi_{r1}, \phi_{r2}, \ldots, \phi_{rN}\}, \quad \forall r \in R, \tag{1}$$

and the following is an expression for each component of the analog-phase shifter, $\phi_{rn} = \complement_{rn} \cdot \exp(j\varphi_{rn})$, where φ_{rn} denotes the discrete phase-shift levels and $\complement_{rn}$ is the reflection amplitude (assumed unity for passive beamforming). To simplify the realistic execution, this study took into account a discrete value (a finite number) for the phase shift angles of intelligent surface segments. The number of these levels will be 2^b if the b-bits are used to represent the level numbers or phase-shifting precision. For the sake of simplicity, the current study also assumed a uniform quantization in the range of $[0, 2\pi]$ for the phase shifter levels. Lastly, the single RIS element's discrete phase shift values set can be expressed as follows [16],

$$\varphi_{rn} \in \left\{0, \frac{\pi}{2^{b-1}}, \ldots, \left(2^b - 1\right)\frac{\pi}{2^{b-1}}\right\}, \quad \forall r \in R, \forall n \in N. \tag{2}$$

In theory, it is possible to adjust the elements' amplitude of reflection $\complement_{rn}$ for various purposes, such as CSI acquisition and system performance optimization. However, taking into account an independent management strategy for both the phase shifter and the reflection amplitude at the same time is essentially expensive. For the purpose of a straightforward implementation, an individual reflecting segment is therefore usually taken into consideration to optimize the reflected signal phase shift. The transmitted baseband signal vector for all K users at the bth BS, $\mathbf{x}_r = [x_{r1}, x_{r2}, \ldots, x_{rk}]$, is shown as follows [17],

$$\mathbf{x}_r = \sum_{k\in|K|} \alpha_{rk}\mathbf{V}_{rk}S_k, \tag{3}$$

where the data to be sent to the kth UE is represented by the symbol S_k, while $\mathbf{V}_{bk} \in \mathbb{C}^{M\times K}$ $\forall m \in M$, is the BS beamformer vector (active beamformer) for the kth user, and $\alpha_{bk} \in \{0,1\}$ is the binary variable that represents the BS-User association factor, where BSs are listed in the serving set of a given user in accordance with their channel states; that is, α_{bk} is set to unity if the bth BS is chosen to be corresponding with the kth UE, otherwise, it is set to zero [18].

$$\alpha_{bk} = \begin{cases} 1, & \text{if } b^{th} \text{ BS is paired,} \\ 0, & \text{else.} \end{cases} \tag{4}$$

Besides, for the sake of simplifying the notation and the mathematical expression, let us utilize the symbol $\mathbf{Q}_k$ to signify the overall cascading channel (effective) as follows,

$$\mathbf{Q}_k = \mathbf{F}_k + \sum_{r\in|R|} \mathbf{H}_r^H \mathbf{\Phi}_r \mathbf{G}_k. \tag{5}$$

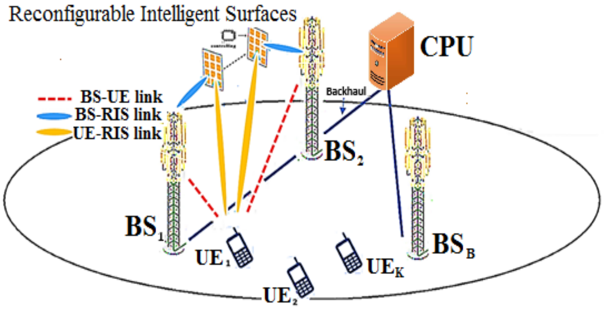

Fig. 1. Multi-RIS-aided CF-MIMO network's Illustration.

Data Rate Formula: According to the previously described mathematical statement, the received signal at the user is,

$$y_k = \underbrace{\sum_{b\in|B|} \mathbf{Q}_{bk}^H \alpha_{bk} \mathbf{V}_{bk} x_k}_{\text{desired signal}} + \underbrace{\sum_{\substack{j\in|K|\\ j\neq k}} \sum_{b\in|B|} \mathbf{Q}_{bk}^H \alpha_{bj} \mathbf{V}_{bj} x_j + \sigma_k^2}_{\text{term of interference}}. \tag{6}$$

It's important to remember that the channel's reciprocity approach can be used to assess the DL (downlink) channel using the obtained uplink-pilot sequences from the UEs, based on the time-division duplex protocol. Compressive sensing (CS) based on the sparse features of the cascaded channel or the formal MSE technique can be used to obtain the channel state information (CSI) by establishing the RIS-assisted linkage (as shown in fig. 1) through the cascaded channel. It is worth noting that when a signal is squared, this will compute $|\cdot|^2$ for scalars or $\|\cdot\|_2^2$ for vectors. This represents the instantaneous power of that signal realization. Also, in terms of Statistical Power, Power is defined as the expectation of the squared magnitude: $P = \mathbb{E}[|s|^2]$. For zero-mean signals, this equals the variance. In the derivations, because data symbols are independent with zero mean and unit variance, the statistical power equals the squared norm of the combining vector: $\mathbb{E}[|\sum \mathbf{a}^H \mathbf{V} s|^2] = \|\sum \mathbf{a}^H \mathbf{V}\|_2^2$. Likewise, the SINR of this UE for the aforementioned scheme can be represented as follows,

$$\Gamma_k(\mathbf{\Phi}, \mathbf{V}) = \frac{\left\| \sum_{b=1}^{B} \alpha_{bk} \mathbf{W}_{bk} \mathbf{V}_{bk} \right\|_2^2}{\sum_{j\neq k} \left\| \sum_{b=1}^{B} \alpha_{bj} \mathbf{W}_{bk} \mathbf{V}_{bj} \right\|_2^2 + \sigma_k^2} \tag{7}$$

where $\mathbf{W}_{bk} = \mathbf{F}_{bk}^H + \sum_{r=1}^{R} \mathbf{H}_{rk}^H \mathbf{\Phi}_r \mathbf{G}_{br}$ note that ℓ_2-norm $\|\cdot\|_2$ is the Euclidean norm for vectors,, the term $\|\sum_{r\in|R|} \sum_{s\in|S|} (\mathbf{F}_{rk}^H + \mathbf{H}_k^H \mathbf{\Phi}_s^H \mathbf{G}_s) \mathbf{V}_i\|_2^2$ corresponds to the useful channel gain, and the terms of the denominator of this equation can be expressed as follows,

$$I_{BSs} + I_{RISs} = \sum_{\substack{i\in|K|\\ i\neq k}} \left\| \sum_{b\in|B|} \mathbf{Q}_{bk}^H \alpha_{bj} \mathbf{V}_{bj} \right\|_F^2. \tag{8}$$

Consequently, the following is an expression for the kth user's sum rate,

$$R_k(\mathbf{\Phi}, \mathbf{V}) = \ln\left(1 + \mathbf{\Gamma}_k(\mathbf{\Phi}, \mathbf{V})\right) \quad \text{(nats/sec/Hz)}.$$

Power Consumption Model: An accurate power consumption model is essential for evaluating the true energy efficiency of RIS-assisted cell-free networks. Unlike prior works that employ overly simplistic models, this work adopts a comprehensive power consumption framework that accounts for all major sources of power dissipation in the system. The adopted power consumption model includes six key components,

$$\begin{aligned} P_T = {} & P_{\text{transmit}} + P_{\text{circuit}} + P_{\text{RIS}} + P_{\text{backhaul}} \\ & + P_{\text{CSI}} + P_{\text{control}} \end{aligned} \tag{9}$$

where, $P_{\text{transmit}} = \sum_{b,k} \|\mathbf{V}_{bk}\|_F^2$ is the radiated transmit power, $P_{\text{circuit}} = \sum_b P_b^{\text{BS}} + \sum_k P_k^{\text{UE}}$ is the static circuit power, $P_{\text{RIS}} = \sum_{r,n} P_{\text{element}}(b)$ is the RIS phase-shifter power, $P_{\text{backhaul}} = \sum_b P_b^{\text{BH}}$ is the backhaul power consumption, $P_{\text{CSI}} = P_{\text{pilot}} + P_{\text{estimation}}$ is the CSI acquisition overhead, and $P_{\text{control}} = \sum_r P_r^{\text{ctrl}}$ is the control signaling overhead. The Energy Efficiency (EE) metric, in (nats/Joule/Hz), is,

$$\text{EE}(\mathbf{\Phi}, \mathbf{V}) = \frac{\sum_{k\in|K|} R_k(\mathbf{\Phi}, \mathbf{V})}{P_T(\mathbf{\Phi}, \mathbf{V})} \quad \text{(nats/J/Hz)}. \tag{10}$$

D. Problem Formulation (Energy Efficiency Formula)

In order to pursue sustainable and highly cost-effective communication, the primary focus in this study is to maximize EE by jointly optimizing the active beamformers $\mathbf{V}$ (the transmit precoding vector of BSs) and the phase-shift matrices $\mathbf{\Phi}$ (passive reflecting phase shifts). According to the equations shown earlier, the energy optimization problem formula can be written as follows,

$$\text{Problem } \mathcal{P}_0: \quad \max_{\mathbf{\Phi}, \mathbf{V}} f_0(\mathbf{\Phi}, \mathbf{V}), \tag{11}$$

Subject to:

$$C_1: \quad \|\phi_{rn}\|_F^2 = 1, \quad \text{for all } r \in |R|, n \in |N|, \tag{11a}$$

$$C_2: \quad \varphi_{rn} \in \mathcal{F}, \quad \text{for all } n \in |N| \text{ and } r \in |R|, \tag{11b}$$

$$C_3: \quad \text{trace}\left(\mathbf{V}_k \mathbf{V}_k^H\right) \leq P_{\max}, \tag{11c}$$

$$C_4: \quad R_k \geq R_{\text{th}}, \quad k \in |K|, \tag{11d}$$

where, $f_0(\mathbf{\Phi}, \mathbf{V}) = \frac{\sum_{k\in|K|} R_k(\mathbf{\Phi}, \mathbf{V})}{P_T(\mathbf{\Phi}, \mathbf{V})}$ is the energy efficiency and

$$\mathcal{F}(\varphi_{rn}) = \left\{0, \frac{\pi}{2^{b-1}}, \ldots, \left(2^b - 1\right) \frac{\pi}{2^{b-1}}\right\}.$$

The quantization step size is $\Delta = \frac{\pi}{2^{b-1}}$. For any continuous phase $\theta \in [0, 2\pi)$, the nearest quantized value is, $\mathcal{Q}(\theta) = \arg\min_{\phi\in\mathcal{F}} |\theta - \phi|$.
Quantization Error Bound: For any $\theta \in [0, 2\pi)$, the quantization error satisfies, $|\theta - \mathcal{Q}(\theta)| \leq \frac{\Delta}{2} = \frac{\pi}{2^b}$, where the quantization intervals are of length Δ. Any θ lies in some interval $[\phi_i, \phi_{i+1}]$ where $\phi_i, \phi_{i+1} \in \mathcal{F}$ and $\phi_{i+1} - \phi_i = \Delta$. The distance to the nearest endpoint is at most $\Delta/2$.

C_1 represents the practical unit-modulus requirement, where passive beamforming $\mathbf{\Phi}$ is limited to having a magnitude of unity due to unit modulus constraints introduced by RIS. This means that the weights can only vary in phase, not amplitude. This restriction is essential for the beamforming process to be implemented practically in MIMO systems. However, this viable set could be defined in terms of a uniform B quantization bit of the phase shifter. This quantization limit is implicit in constraint C2 for an infinite-resolution phase shift network (the set of RIS passive beamforming vectors). While constraints C1 and C2 both apply to the phase-shift variables, they represent fundamentally different mathematical requirements and are *not* redundant. Understanding this distinction is crucial for proper problem formulation and algorithm design.The key distinction is that C1 is a *continuous manifold constraint* that arises from the passive nature of RIS elements(restricts each ϕ_{rn} to lie on the complex unit circle), while C2 is a *discrete combinatorial constraint* that arises from practical hardware implementation with finite-bit phase shifters (further restricts each ϕ_{rn} to a finite subset of $\mathbb{S}^1$ consisting of 2^b equally spaced points). They are applied sequentially in our solution approach: we first find a high-quality solution on the continuous manifold (satisfying C1), then project it onto the nearest discrete points (satisfying C2). The third constraint in (11c), aims to ensure the overall power limitation (power budget of each base station). Constraint C_4 is to ensure the minimum rate for each user, i.e., not to be less than the threshold level R_{th}. The expression in the optimization problem ($\mathcal{P}_0$) is, unfortunately, a highly nonconvex problem and NP-hard due to the fractional objective, rate expressions $R_k(\mathbf{\Phi}, \mathbf{V})$, coupling between variables, and discrete phase constraints, which are challenging to solve. The joint active and passive beamforming problem in RIS-assisted networks with discrete phase shifts is NP-hard. This can be proven by reduction from the Boolean satisfiability problem or the maximum cut problem, as shown in [16] for RIS phase-shift optimization and [17] for MIMO beamforming with discrete constraints. More generally, optimization over the complex circle manifold with discrete subsets is known to be NP-hard [18]. Consequently, this study suggests an iterative scheme where one beamformer is optimized, while it will be fixed at a time.

E. Alternating Optimization (AO) Framework (Energy Maximization Algorithm Design)

We employ the AO strategy as the primary approach to decouple the optimization variables, where the idea is to fix everything except one variable, and perform gradient descent for it, and do this procedure alternatively for all variables till some convergence criteria are met. In this manner, the user-related processing functions can be split into small, separated blocks that in turn can be virtualized in the CPU with a minimal computational overhead. For the sake of disconnecting the optimization variables' coupling of the problem $\mathcal{P}_0$, the current study is motivated to use the transform of a quadratic function. It is worth bearing in mind that the standard Dinkelbach approach [19] is hard to engage as a result of the dimensional completeness of such a fractional formula. Thus, further auxiliary variables are utilized to resolve the BSs' beamform vector task $\mathbf{V}$ and passive beamforming $\mathbf{\Phi}$ (reflected beamforming from RIS) in an iteratively converged strategy. The solution imposes the alternating optimization (AO) approach [15] by keeping the RIS's reflection beamform vector $\mathbf{\Phi}$ at a fixed value to perform the $\mathbf{V}$ optimizing and then keeping $\mathbf{V}$ at a fixed value to perform the $\mathbf{\Phi}$ optimizing in consequence procedure till attain convergence. After that, the optimizing task $\mathcal{P}_0$ can be performed for $\mathbf{V}$, $\mathbf{\Phi}$, and auxiliary variable μ at a small cost of computational complexity, as depicted in Algorithm-I pseudo code that gives the details of beamform vectors (active/passive-BF) design procedure. In particular, sub-problems are required to be evaluated, subsequently, to update or renew the values of the variables μ based on $\mathbf{V}^*$, $\mathbf{\Phi}^*$, the task in (11) will be divided into two subproblems, solved iteratively: Subproblem-1 will update the Active Beamforming with fixed RIS phase shifts, while optimize. Subproblem 2 will update the Passive Beamforming with fixed BSs' beamformers as follows.

F. Transmit Beamforming Optimization (Fractional Maximization)

Using the AO method, problem ($\mathcal{P}_0$) will be primarily handled as a matrix-factorizing task, including two matrices parameters, $\mathbf{V}$ and $\mathbf{\Phi}$. Alternating Minimization, however, is a widely used and experimentally effective solution to the issues related to the optimization of various variable subsets. In the case of fixing the RISs' passive transmit beamforming, the task is a quadratic problem with a quadratic constraint nonconvex problem [20] and this sub-problem can be stated as follows:

$$\begin{aligned} \mathcal{P}_1: \quad & \max_{\mathbf{V}} f_1(\mathbf{V}), \qquad (12)\\ \text{s.t.} \quad & C_3: \quad \text{trace}\left(\mathbf{V}_k \mathbf{V}_k^H\right) \leq P_{\max}, \\ & C_4: \quad R_k \geq R_{\text{th}}, \end{aligned}$$

where,

$$f_1(\mathbf{V}) = \frac{\sum_{k\in|K|} \ln\left(1 + \mathbf{\Gamma}_k\right)}{\sum_{k\in|K|} \text{trace}\left(\sum_{b\in|B|} \|\mathbf{V}_{bk}\|_F^2\right) + P_{\text{cct}} + P^{\text{RIS}}}.$$

This is a fractional non-linear programming task which can be handled via employing the Dinkelbach (parametric) Algorithm by iteratively solving a sequence of parametric subproblems. It works by generating a sequence of the auxiliary variable μ for the EE value, and then solving the non-linear problem of finding x that maximizes $f(x) - \mu g(x)$, i.e., converts the fractional problem into a series of subtractive-form problems [16]. Dinkelbach converges to the global optimum of the fractional program if each parametric subproblem is solved optimally to get a monotone improvement and convergence to a stationary point. Let μ be the auxiliary variable in the Dinkelbach-approach for the EE fractional objective. The algorithm iteratively updates $\mu^{(i)}$ and solves the problem until convergence. Thus, for a fixed μ, the parametric problem to be solved is,

$$\begin{aligned} \mathcal{P}_2: \quad & \max_{\mathbf{\Phi}, \mathbf{V}} f_2(\mathbf{\Phi}, \mathbf{V}; \mu), \qquad (13)\\ \text{s.t.} \quad & C_3, C_4, \end{aligned}$$

where,

$$f_2(\mathbf{\Phi}, \mathbf{V}; \mu) = \sum_{k\in|K|} \ln\left(1 + \mathbf{\Gamma}_k(\mathbf{\Phi}, \mathbf{V})\right) - \mu \cdot P_T(\mathbf{V}).$$

This is still nonconvex because rates depend on ratios of quadratic forms in $\mathbf{\Gamma}_k$. However, inspired by the work of [21], Subproblem-1 for the active beamforming can be converted into a second-order constraint problem, also known as a second-order cone programming (SOCP) problem [20] with fixed passive beamforming $\mathbf{\Phi}$. The rate for user k under linear receivers is $\mathbf{R}_k = \ln(1+\mathbf{\Gamma}_k)$, where $\mathbf{\Gamma}_k$ is the Signal-to-Interference-plus-Noise Ratio (SINR). SOCP technique leverages the use of the linear Weighted minimum mean square error algorithm (WMMSE) framework to establish a strict equivalence between achieving a certain SINR level and minimizing the MSE. In other words, SOCP is a type of convex optimization problem that minimizes a linear function subject to a set of linear and second-order cone constraints form via employing a quadratic transform (weighted MMSE equivalence, see appendix-A). Accordingly, using the same methodology as in reference [21], the design of the joint beamformer to maximize sum-rate is corresponding to minimizing the a weighted MSE (WMMSE) formulation. The MSE for user k can be defined as the mean squared error between the received data signal $\widetilde{\mathbf{x}}_k$, and the transmitted data signal $\mathbf{x}_k$, $\text{MSE}_{jr} = \mathbb{E}\left\{\|\mathbf{x}_k - \widetilde{\mathbf{x}}_k\|_F^2\right\}$ where the expectation operation is conditioned on the estimated channel $\widehat{\mathbf{h}}_{jr}$. Let $q_k \in \mathbb{C}$ be a receive beamforming (scaling) coefficient for user k, then the estimated signal at user k will be $q_k y_k$, and the Mean Squared Error MSE_k is,

$$\begin{aligned}\mathbf{E}_k = \text{MSE}_k &= \mathbb{E}\left((x_k - q_k y_k)(x_k - q_k y_k)^H\right), \qquad (14)\\ &= \mathbb{E}\left(\|x_k - q_k y_k\|_F^2\right),\\ = |q_k|^2(&\sum_{\substack{j\in|K|\\ j\neq k}} \mathbf{Q}_k^H \mathbf{V}_j \mathbf{V}_j^H \mathbf{Q}_k + \sigma_k^2) - 2\text{Re}\left\{q_k \mathbf{Q}_k^H \mathbf{V}_k\right\} + 1.\end{aligned}$$

where,

$$\mathbf{Q}_k^H = \sum_{b\in|B|} \mathbf{Q}_{bk}^H, \quad \mathbf{V}_j = \left[\mathbf{V}_{1j}^T, \mathbf{V}_{2j}^T, \ldots, \mathbf{V}_{Bj}^T\right]^T.$$

and for fixed $\mathbf{V}_k$ the optimal MMSE receive beamforming (scaling) coefficient for receiver k is [15],

$$q_k^* = \frac{\mathbf{V}_k^H \mathbf{Q}_k}{\sum_{\substack{j\in|K|\\ j\neq k}} \mathbf{Q}_k^H \mathbf{V}_j \mathbf{V}_j^H \mathbf{Q}_k + \sigma_k^2}. \qquad (15)$$

Now, by introducing an auxiliary positive weight variable for user k, $w_k \geq 0$, the sum rate maximization problem $\mathcal{P}_2$ can be expressed using MMSE / WMMSE equivalence to handle sum-rate problem, $f_3(\mathbf{V}_k, q_k, w_k)$ as follows,

$$\begin{aligned}\mathcal{P}_3:\quad &\min_{\mathbf{V}_k, q_k, w_k} f_3(\mathbf{V}_k, q_k, w_k), \qquad (16)\\ \text{s.t.}\quad &C_3, C_4,\end{aligned}$$

where,

$$f_3(\mathbf{V}_k, q_k, w_k) = \sum_{k\in|K|} w_k \mathbf{E}_k - \ln w_k.$$

And this minimizing WMMSE (over $\mathbf{v}_k, q_k, w_k$) in $\mathcal{P}_3$ corresponds to maximizing the sum rate in $\mathcal{P}_2$. This is the key equivalence enabling alternating updates: $\mathbf{v}_k, q_k, w_k$, update q_k, then w_k, then $\mathbf{v}_k$. In the presence of Dinkelbach with penalty $-\mu \cdot P_T(\mathbf{V})$, the algorithm starts the iteration of the outer Dinkelbach loop with $\mu^{(i)} \geq 0$ to solve $f_2^{(i)}$ and find $\mathbf{\Phi}^{(i)}, \mathbf{V}^{(i)}$. Dropping the constant factors from the EE formula, update μ as follows,

$$\mu^{(i+1)} = \frac{\sum_{k\in|K|} \ln\left(1 + \mathbf{\Gamma}_k(\mathbf{\Phi}^{(i)}, \mathbf{V}^{(i)})\right)}{\sum_{k\in|K|} \text{trace}\left(\sum_{b\in|B|} \|V_{bk}^{(i)}\|_F^2\right)}. \qquad (17)$$

Additionally, it is possible to rewrite the minimum rate non-convex constraint ($R_k \geq R_{\text{th}}$) as,

$$\left|\mathbf{V}_k^H \mathbf{Q}_k\right|^2 \geq \gamma_{\text{th}}(\sum_{\substack{j\in|K|\\ j\neq k}} \mathbf{Q}_k^H \mathbf{V}_j \mathbf{V}_j^H \mathbf{Q}_k + \sigma_k^2), \qquad (18)$$

with $\gamma_{\text{th}} = 2^{R_{\text{th}}} - 1$. By using the Successive convex approximation (SCA) approach, the inequality's left side can be lower constrained by its first Taylor expansion, allowing (18) to be roughly expressed as,

$$\begin{aligned}2\text{Re}\left\{\left(\mathbf{Q}_k^H \mathbf{V}_k^{(i)}\right)^H \mathbf{Q}_k^H \mathbf{V}_k\right\} - \left|\mathbf{Q}_k^H \mathbf{V}_k^{(i)}\right|^2 \geq\\ \gamma_{\text{th}}(\sum_{\substack{j\in|K|\\ j\neq k}} \mathbf{Q}_k^H \mathbf{V}_j \mathbf{V}_j^H \mathbf{Q}_k + \sigma_k^2), \qquad (19)\end{aligned}$$

which yields, $\mathcal{X}_k^{(i)} \geq \gamma_{\text{th}}$, given that,

$$\mathcal{X}_k^{(i)} = \frac{2\text{Re}\left\{\left(\mathbf{Q}_k^H \mathbf{V}_k^{(i)}\right)^H \mathbf{Q}_k^H \mathbf{V}_k\right\} - \left|\mathbf{Q}_k^H \mathbf{V}_k^{(i)}\right|^2}{\sum_{\substack{j\in|K|\\ j\neq k}} \mathbf{Q}_k^H \mathbf{V}_j \mathbf{V}_j^H \mathbf{Q}_k + \sigma_k^2},$$

where the value from the previous iteration is indicated by (i). Consequently, ($\mathcal{P}_3$) may be rewritten as,

$$\begin{aligned}\mathcal{P}_3:\quad &\min_{\mathbf{V}_k, q_k, w_k} f_3(\mathbf{V}_k, q_k, w_k), \qquad (20)\\ \text{s.t.}\quad &C_3: \quad \text{trace}\left(\mathbf{V}_k \mathbf{V}_k^H\right) \leq P_{\max},\\ &C_4: \quad \mathcal{X}_k^{(i)} \geq \gamma_{\text{th}}.\end{aligned}$$

The iteration of the algorithms, employs termination or stopping criteria, when the algorithm is no longer making significant progress, as follows,

$$\sum_{k\in|K|} \ln(1 + \mathbf{\Gamma}_k(\mathbf{\Phi}, \mathbf{V})) - \mu \cdot P_T(\mathbf{V}) \leq \varepsilon. \qquad (21)$$

Now, to elaborate, applying the linear weighted MMSE equivalence ($\mathcal{P}_3$) in the Dinkelbach fractional maximization expression ($\mathcal{P}_2$), i.e., Plugging $\mathbf{E}_k$ as a quadric in $\mathbf{V}_{bk}$, yields,

$$\begin{aligned}&\max_{\mathbf{\Phi}, \mathbf{V}} \sum_{k\in|K|} \mathbf{R}_k - \mu \cdot \mathbf{P}_T, \qquad (22)\\ &= \min_{q_k, w_k, \mathbf{V}, \mathbf{\Phi}} \sum_{k\in|K|} w_k \mathbf{E}_k - \ln w_k + \mu \cdot \mathbf{P}_T.\end{aligned}$$

And in turn, the optimization problem will be as follows,

$$\begin{aligned}\mathcal{P}_4:\quad &\min_{v_k, q_k, w_k} f_4(\mathbf{V}_k, q_k, w_k), \qquad (23)\\ \text{s.t.}\quad &C_3,\end{aligned}$$

where,

$$f_4(\mathbf{V}_k, q_k, w_k) = \sum_{k\in|K|} w_k \mathbf{E}_k - \ln w_k$$
$$+\mu \cdot \sum_{k\in|K|} \text{trace}(\sum_{b\in|B|} \mathbf{V}_{bk}^H \mathbf{V}_{bk}).$$

Problem $\mathcal{P}_4$ is convex in each block of variables $\mathbf{v}_k, q_k, w_k$. We focus on solving for $\mathbf{v}_k$, with fixed q_k, w_k, which is a Quadratic Constrained Quadratic Program (QCQP) [15]. It can be efficiently cast as a second-order cone programming (SOCP) problem in $\mathbf{V}_k$ is as follows,

$$\mathcal{P}_5 : \quad \min_{\mathbf{V}_k} f_5(\mathbf{V}_k), \qquad (24)$$
$$\text{s.t.} \quad \text{trace}\left(\mathbf{V}_k \mathbf{V}_k^H\right) \le P_{\max},$$

where, $\mathbf{Q} = [\mathbf{Q}_1, \mathbf{Q}_2, \ldots, \mathbf{Q}_K]$, $\mathbf{V} = [\mathbf{V}_1, \mathbf{V}_2, \ldots, \mathbf{V}_K]$, and,

$$f_5(\mathbf{V}_k) = \sum_{k\in|K|} w_k \|q_k\|^2 \left(\sum_{\substack{j\in|K| \\ j\neq k}} |Q_k^H \mathbf{V}_j|^2 + \sigma_k^2 \right)$$

$$-2\mathcal{R}\{q_k Q_k^H \mathbf{V}_k\} + \mu \cdot \sum_{k\in|K|} \text{trace}\left(\sum_{b\in|B|} |\mathbf{V}_{bk}|^2 \right)$$

This can be rewritten in a standard SOCP form,

$$\mathcal{P}_6 : \quad \min_{\mathbf{V}_k} f_6(\mathbf{V}_k), \qquad (25)$$

Subject to:

$$C_1 : \quad \text{tr}\left(\mathbf{V}_k \mathbf{V}_k^H\right) \le P_{\max}, \quad b \in |B|,$$
$$C_2 : \quad t_k \ge \left\| \frac{\mathbf{Q}^H \mathbf{V}}{\sigma_k} \right\|_F^2, \quad k \in |K| \quad \text{(SOC Constraints)},$$
$$C_4 : \quad \mathcal{X}_k^{(i)} \ge \gamma_{\text{th}}, \quad k \in |K|,$$

where,

$$f_6(\mathbf{V}_k) = \sum_{k\in|K|} w_k \left[|q_k|^2 \cdot t_k - 2\text{Re}\left\{ q_k \mathbf{Q}_k^H \mathbf{V}_k \right\} \right]$$
$$+\mu \cdot \sum_{k\in|K|} \text{trace}(\sum_{b\in|B|} \|\mathbf{V}_{bk}\|_F^2).$$

The objective in $\mathcal{P}_6$ is a convex quadratic SOCP and beamformer update is a convex QCQP with linear constraints (power). So it can be solved efficiently by convex optimization solvers like CVX (disciplined convex programming available in MATLAB®, Python, Octave, etc).

G. Intelligent surfaces' Passive Beamforming (Riemannian conjugate gradient method)

Recall problem $\mathcal{P}_2$, assuming fixed $\mathbf{V}_k$, the second step in the AO-algorithm needs to update $\mathbf{\Phi}$ to maximize the parametric objective of the RIS subproblem,

$$\mathcal{P}_7 : \quad \max_{\mathbf{\Phi}} \sum_{k\in|K|} \mathbf{R}_k(\mathbf{\Phi}, \mathbf{V}) - \mu \cdot P_T(\mathbf{V}), \qquad (26)$$
$$\text{s.t.} \quad C_3, C_4,$$
$$= \min_{\mathbf{\Phi}} f_6(\mathbf{\Phi}), \quad \text{s.t.} \quad C_3, C_4,$$

where,

$$f_6(\mathbf{\Phi}) = \sum_{k\in|K|} w_k \mathbf{E}_k(\mathbf{\Phi}, \mathbf{V}) - \ln w_k.$$

Passive beamforming $\mathbf{\Phi}$ appears nonlinearly inside the effective channels $\mathbf{Q}_k$, in addition, unit-modulus constraint is main challenge, therefore, the suggestion is to optimize RIS problem with unit-modulus constraints is modeled as a Riemannian manifold [22].

Riemannian(manifold) optimization; The main obstacles are the intrinsically nonconvex unit modulus constraints in $\mathcal{P}_7$. To the best of our knowledge, there is no one-size-fits-all method for determining the optimal solution (8). In the following, we will introduce an effective manifold optimization method to find a near-optimal solution to the problem. As seen in Fig.2, a manifold topological $\mathcal{M}$ is a space that resembles a Euclidean space near every point (product of complex circles). In other words, any point's neighborhood inside a manifold is homeomorphic to Euclidean space. Additionally, the tangential space $\mathbf{T}_{\mathbf{\Phi}}\mathcal{M}$ is composed of the tangent vectors η of the curves crossing a specific point ($\mathbf{\Phi}_n$) on the manifold. $\mathcal{M} = \left\{ \mathbf{\Phi} \in \mathbb{C}^N : |\mathbf{\Phi}_n| = 1, \ \mathbf{\Phi}_n = e^{j\varphi_n} \right\}$ this is a smooth manifold, and we can perform optimization directly on it. The tangent for the manifold $\mathcal{M}$ space at $\mathbf{\Phi}$ is, $\mathbf{T}_{\mathbf{\Phi}}\mathcal{M} = \{\mathbf{z} \in \mathbb{C}^N : \text{Re}\{\mathbf{z} \cdot \mathbf{\Phi}^*\} = 0\}$. In most applications, manifolds belong to a certain kind of topological manifold called a Riemannian manifold. The Riemannian metric, an inner linear product given on the tangential space $T_W M_c$ on a Riemannian manifold $\mathcal{M}$, can be used to quantify angles and distances on manifolds. For a Riemannian manifold in particular, Riemannian geometric calculus can be applied. The rich metric of the Riemannian manifold makes it possible to define gradients of cost functions. More importantly, optimization over a Riemannian manifold is locally equivalent to smooth-constrained optimization on a Euclidean space. As a result, their equivalent on Riemannian manifolds can be identified using a conjugate-gradient method in Euclidean space. In the following, we will provide a brief overview of this equivalency. The tangent space at (w) can then be represented as $T_W\mathcal{M}$ since the tangent vectors specify the directions in which a certain point x on the manifold M_c can travel. Note that the vector w = vec($\mathbf{\Phi}$) forms a complex circle manifold $\mathcal{M}_{CC}^m$, with m = Mt MRF. As a result, the search range for the optimal task in ($\mathcal{P}_6$7) is over a product of m circles inside the complex plane of the Riemannian sub-manifold. Any entry in the passive beamforming matrix cannot become a unit modulus due to the geometry depicted in Fig. 2, which makes the beamformer design problems uncompromising. This paper will specifically examine an Alternating Minimization strategy based on manifold optimization concepts to solve the problem ($\mathcal{P}_7$) by proving that a Riemannian manifold may be created when there are limitations on the unit modulus.

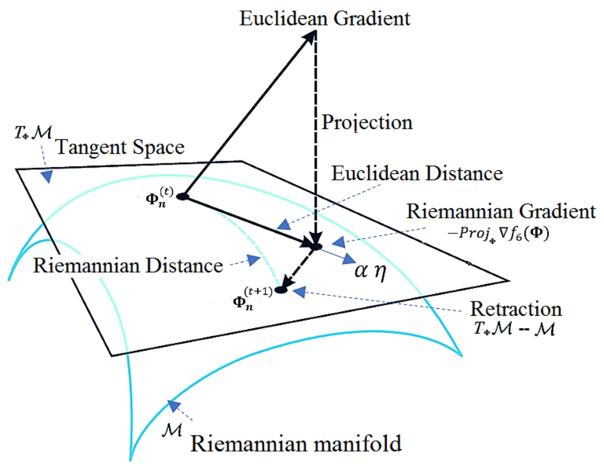

Fig. 2. A Riemannian manifold tangent vector and tangent space.

The weighted sum of the gradients, or Euclidean gradient, can be written as follows,

$$\nabla_{\mathbf{\Phi}} f_6(\mathbf{\Phi}) = \sum_{k\in|K|} w_k \cdot \nabla_{\mathbf{\Phi}} \mathbf{E}_k(\mathbf{\Phi}) = \sum_{k\in|K|} w_k \cdot \frac{\partial \mathbf{E}_k(\mathbf{\Phi})}{\partial \mathbf{\Phi}^*}, \tag{27}$$

and for a single MSE term yields,

$$\mathbf{E}_k = |q_k|^2 \Big(\sum_{\substack{j\in|K|\\ j\neq k}} \mathbf{Q}_k^H \mathbf{V}_j \mathbf{V}_j^H \mathbf{Q}_k + \sigma_k^2 \Big) \\ -2\mathrm{Re}\left\{ q_k \mathbf{Q}_k^H \mathbf{V}_k \right\} + 1. \tag{28}$$

Now, using Euclidean calculus to compute complex derivative $\frac{\partial f_6(\mathbf{\Phi})}{\partial \mathbf{\Phi}^*}$,

$$\frac{\partial \mathbf{E}_k(\mathbf{\Phi})}{\partial \mathbf{\Phi}^*} = |q_k|^2 \cdot \sum_{\substack{j\in|K|\\ j\neq k}} \frac{\partial \left(\mathbf{Q}_k(\mathbf{\Phi})^H \mathbf{V}_j \mathbf{V}_j^H \mathbf{Q}_k(\mathbf{\Phi}) + \sigma_k^2 \right)}{\partial \mathbf{\Phi}^*} \\ -2\mathrm{Re}\left\{ q_k \cdot \frac{\partial \mathbf{Q}_k(\mathbf{\Phi})^H \mathbf{V}_k}{\partial \mathbf{\Phi}^*} \right\}. \tag{29}$$

Because $\mathbf{Q}_k(\mathbf{\Phi})$ is linear in $\mathbf{\Phi}$, derivatives reduce to linear forms. For example, if $\mathbf{Q}_k = \mathbf{Q}_k^d + \mathbf{C}_k \mathbf{\Phi}$, (writing a linear mapping, when channels depend linearly on $\mathbf{\Phi}$), with $\mathbf{C}_k$ (constructed from BS→RIS and RIS→user). Then,

$$\frac{\partial \mathbf{Q}_k(\mathbf{\Phi})^H \mathbf{V}_k}{\partial \mathbf{\Phi}^*} = 0, and, \frac{\partial \mathbf{Q}_k(\mathbf{\Phi})^H \mathbf{V}_k}{\partial \mathbf{\Phi}} = \mathbf{C}_k^H \mathbf{V}_j.$$

Use conjugate derivatives appropriately, the end result is an explicit vector gradient, $\nabla_{\mathbf{\Phi}} f_6(\mathbf{\Phi}) \in \mathbb{C}^N$. So, for the MSE derivative we have,

$$\frac{\partial \mathbf{E}_k(\mathbf{\Phi})}{\partial \mathbf{\Phi}^*} = |q_k|^2 \cdot \sum_{\substack{j\in|K|\\ j\neq k}} \mathbf{C}_k^H \left(\mathbf{Q}_k^H \mathbf{V}_j \right)^* + \mathbf{C}_k^H \mathbf{V}_j^* \left(\mathbf{Q}_k^H \mathbf{V}_j \right) \\ -2 q_k \mathbf{C}_k^H \mathbf{V}_k.$$

$$= \partial(|q_k|^2 \cdot (\sum_{\substack{j\in|K|\\ j\neq k}} \nabla_{\mathbf{\Phi}} (\mathbf{Q}_k^H \mathbf{V}_j \mathbf{V}_j^H \mathbf{Q}_k + \sigma_k^2 t)) \\ -2 \cdot \nabla_{\mathbf{\Phi}} \mathrm{Re}\{ q_k \mathbf{Q}_k^H \mathbf{V}_k \})/\partial \mathbf{\Phi}^*.$$

$$= |q_k|^2 \cdot (\sum_{\substack{j\in|K|\\ j\neq k}} 2\mathrm{Re}\{ (\frac{\partial \mathbf{Q}_k}{\partial \mathbf{\Phi}^*})^H \mathbf{V}_j \mathbf{V}_j^H \mathbf{Q}_k + \sigma_k^2 \}) \\ -2\mathrm{Re}\left\{ q_k \cdot (\frac{\partial \mathbf{Q}_k}{\partial \mathbf{\Phi}^*})^H \mathbf{V}_k \right\}.$$

$$= 2\mathrm{Re}\{ (\frac{\partial \mathbf{Q}_k}{\partial \mathbf{\Phi}^*})^H \cdot (|q_k|^2 \sum_{\substack{j\in|K|\\ j\neq k}} \mathbf{V}_j \mathbf{V}_j^H \mathbf{Q}_k + \sigma_k^2) \\ - q_k \cdot \mathbf{V}_k \}. \tag{30}$$

So, a compact, general-form expression for the final objective's gradient will be as follows,

$$\nabla_{\mathbf{\Phi}} f_6(\mathbf{\Phi}) = 2\mathrm{Re}\{ w_k (\frac{\partial \mathbf{Q}_k}{\partial \mathbf{\Phi}^*})^H \cdot (|q_k|^2 \sum_{\substack{j\in|K|\\ j\neq k}} \mathbf{V}_j \mathbf{V}_j^H \mathbf{Q}_k \\ + \sigma_k^2) - q_k \cdot \mathbf{V}_k \}. \tag{31}$$

Like Euclidean space, a function's largest decline direction is represented by one of the tangent vectors associated with the negative Riemannian gradient. The normal projection of the Euclidean gradient ∇f_6 over the tangential space, $T_w M_m$ yields the Riemannian gradient at $\mathbf{\Phi}$, which is the tangential vector grad$f_6(\mathbf{\Phi})$, as the manifold M_m (with complex circles) is a Riemannian sub-manifold of M. In this context, the vector of the Riemannian gradient is the Euclidean gradient projected onto $T_\theta M$. Given Euclidean gradient $\nabla_{\mathbf{\Phi}} \mathcal{F}(\mathbf{\Phi})$, the Riemannian gradient on unit modulus manifold is the orthogonal projection onto the tangent space at current point $\mathbf{\Phi}$ (take a step along negative Riemannian gradient),

$$\mathrm{grad} f_6(\mathbf{\Phi}) = -\mathrm{Proj}_{W_{PB}} \nabla f_6(\mathbf{\Phi}) = \nabla f_6(\mathbf{\Phi}) - \\ \mathrm{Re}\left(\nabla f_6(\mathbf{\Phi}) \cdot \mathbf{\Phi}^* \right) \cdot \mathbf{\Phi}. \tag{32}$$

Where the operation $(\cdot)$ denotes the elementwise product. Some methods on complex-valued matrix derivatives are required to solve this Euclidean gradient. Another important component of manifold optimization is retraction, which projects a vector onto the manifold from the tangential space. When traveling along a tangent vector, it establishes the location on the manifold. One way to express the tangential vector $\alpha\eta$ retraction for the complex circle manifold (map back to manifold) at point $\mathbf{\Phi}$ is as follows,

$$\mathbf{\Phi}^{(t+1)} = \mathrm{Ret}_{\mathbf{\Phi}} \left(\alpha\eta^{(t)} \right) = \mathrm{vect} \left[\frac{ \left(\mathbf{\Phi}^{(t)} + \alpha\eta^{(t)} \right)_i }{ \left| \left(\mathbf{\Phi}^{(t)} + \alpha\eta^{(t)} \right)_i \right| } \right] \\ = e^{i\measuredangle\left(\mathbf{\Phi}^{(t)} + \alpha\eta^{(t)} \right)}. \tag{33}$$

where, η is the direction associated with grad$f_6(\mathbf{\Phi})$, and α is the size of the search step (learning rate or step length). The step size α in the Riemannian gradient update is critical for convergence. An Armijo backtracking line search along the Riemannian manifold is employed to ensure sufficient decrease while maintaining the unit-modulus constraint. Specifically, at iteration t, given current point $\mathbf{\Phi}^{(t)}$ and descent direction $\eta^{(t)} = -\,\mathrm{grad}\, f_6(\mathbf{\Phi}^{(t)})$, we find the smallest integer $m \geq 0$ such that, $f_6(\mathrm{Ret}_{\mathbf{\Phi}^{(t)}}(\beta^m \eta^{(t)})) \leq f_6(\mathbf{\Phi}^{(t)}) + c\beta^m \langle \mathrm{grad}\, f_6(\mathbf{\Phi}^{(t)}), \eta^{(t)} \rangle_{\mathbf{\Phi}^{(t)}}$ where $\beta \in (0, 1)$ is

Algorithm 1 Proposed Iterative AO Approach

1: *Initialize:* Set system parameters; total transmit power P_t; channel matrices $\mathbf{F}, \mathbf{H}$; iteration index $i = 0$; FP (or Dinkelbach) parameter μ; random initial RIS phase matrix $\mathbf{\Phi}^{(0)}$; MRT initial BS beamforming matrix $\mathbf{V}^{(0)}$.
2: **repeat**
3: **repeat**
4: Update the equivalent channel $\mathbf{Q}_k$ using (5).
5: Update MMSE auxiliary variables $\alpha_k^{(i+1)}$ using (15) and the optimal BS active beamforming for fixed $\mathbf{\Phi}$ using (24)
6: Employ the CVX tool to solve the SOCP in (22). Steps to optimize $\mathbf{\Phi}$ on the manifold:
7: Compute the Euclidean gradient $\nabla_{\mathbf{\Phi}} f(\mathbf{\Phi})$ w.r.t. RIS passive beamforming for fixed $\mathbf{V}$ using (31).
8: Obtain the Riemannian gradient $\boldsymbol{\eta}^{(i)}$ by projecting $\nabla_{\mathbf{\Phi}} f$ onto the tangent space using (32).
9: Compute the retraction of the tangent vector $\boldsymbol{\eta}^{(i)}$ onto the unit-modulus manifold using (33).
10: Select the nearest quantized passive beamforming vector from the finite set $\mathcal{F}(\mathbf{\Phi}_{\text{tr}})$.
11: **until** the AO loop converges, eq.(34). Otherwise, return to Step 3.
12: Update the FP parameter $\mu^{(i+1)}$ using (17)
13: **until** the FP loop converges. Otherwise, update the iteration index:$(i = i + 1)$ and return to Step 2.

the backtracking factor (set to 0.5), $c \in (0, 1)$ is the Armijo constant (set to 0.1), and $\langle \cdot, \cdot \rangle_{\mathbf{\Phi}}$ is the Riemannian metric. The step size is then set as $\alpha = \beta^m$. For scenarios requiring fixed step size (e.g., for complexity predictability), we empirically determined $\alpha = 0.1$ provides robust convergence across all tested configurations, with convergence verified in $> 95\%$ of trials. Armijo line search achieves near-optimal performance while eliminating the need for manual tuning. Appendix-B provides a complete derivation of Riemannian conjugate gradient method. The hybrid beamformer is constructed for the fully connected array through an alternating minimizing approach by repeatedly solving problems of Active/Passive beamforming. Pseudo code of the proposed approach is depicted in Algorithm-I that gives the details of the proposed EE maximizing algorithm through the design of the beamforming vectors, where the convergence of the FP loop is dominated as follows,

$$\left| \sum_{k \in \mathcal{K}} \ln(1 + \Gamma_k(\mathbf{P}, \mathbf{V})) - \mu \cdot P_{\text{tr}}(\mathbf{V}) \right| \leq \varepsilon. \tag{34}$$

Note: η is the direction associated with the Riemannian gradient. It is worth noting that the objective function increases monotonically because the transmit beamforming sub-problem and the RIS phase-shift sub-problem is blocks of alternating optimization or block coordinate descent, and this will ensure the convergence of Algorithm-I (see Appendix-C).

III. Algorithm-I Computational Complexity Analysis

The number of alternating maximization iterations and the complexity (needed to solve each subproblem) will determine the computational complexity of the suggested techniques. In dense Cell-Free networks with a large number of users the complexity of the active beamforming subproblem is polynomial in K. On the other hand, the algorithm can handle a high number of RIS elements efficiently since the RIS optimization complexity scales linearly with N.

1) The complexity of Effective channel update (Step 4)is $\mathcal{O}(KRNM)$.
2) MMSE receiver update (Step 5) costs $\mathcal{O}(K^2BM)$, K^2 due to interference terms.
3) WMMSE weight update costs $\mathcal{O}(KBM)$ Linear in K.
4) The Complexity of solving the SOCP for the active beamforming problem for each single iteration is, (Step 6) is $\mathcal{C}_{\text{passive}} = \mathcal{O}\left(\sqrt{MBK} \cdot (MBK^2)^2\right) = \mathcal{O}\left((MB)^{2.5} K^{4.5}\right)$, polynomial in both BS configuration and number of users.
5) The complexity of Euclidean gradient(Step 7)is, $\mathcal{O}(K^2RNM)$ Computing $\mathbf{C}_k^H \mathbf{V}_j$ for all k, j. The complexity of Riemannian projection (Step 8) is $\mathcal{O}(N)$ Element-wise operations.
6) Conjugate direction (Step 9) costs $\mathcal{O}(N)$ Vector operations.
7) The complexity of Quantization (Step 10) is $\mathcal{O}(RN)$ One-time cost.
8) Dinkelbach update costs $\mathcal{O}(K)$ Scalar computation.

Consequently, the overall order of magnitude estimate per iteration of the alternating algorithm can be approximated as follows, $\mathcal{C}_{\text{total}} = I_{\text{AO}} \cdot (\mathcal{C}_{\text{active}} + \mathcal{C}_{\text{passive}})$ $= \mathcal{O}\left(I_{AO} \cdot \log_2 \frac{1}{\varepsilon} \cdot \left((MB)^{2.5}K^{4.5} + K^2RNM\right)\right)$, where I_{AO} represents the number of iterations to converge in the alternating algorithm, and ε is accuracy of the variable updating. For a typical scenario with $B = 4$, $M = 4$, $K = 4$, $R = 4$, $N = 32$, and typical iteration counts $I_{\text{FP}} = 5$, $I_{\text{AO}} = 10$, $I_{\text{Riem}} = 30$, the complexity breakdown is, SOCP term: $(MB)^{2.5}K^{4.5} = 2{,}097{,}152$, Riemannian term: $K^2RNM = 4^2 \cdot 4 \cdot 32 \cdot 4 = 16 \cdot 4 \cdot 32 \cdot 4 = 8{,}192$. Total per AO: $2{,}097{,}152 + 30 \cdot 8{,}192 = 2{,}097{,}152 + 245{,}760 = 2{,}342{,}912$ operations Total algorithm:$5 \cdot 10 \cdot 2{,}342{,}912 \approx 117$ million operations On a modern CPU capable of 10 GFlops, this corresponds to approximately 0.012 seconds per AO iteration.

IV. Numerical Results

This study, without loss of generality, assumed that RISs and BSs are uniformly spread in circles (within a radius of 250m) about the USs. The multiple single-antenna users are deployed randomly (uniformly) in a circle of radius 20m. The main goal is to evaluate the proposed method for the reflecting surfaces aided cell-free scheme in relation to EE, and demonstrate how EE is affected by the number of RIS elements, the BS's transmit power, the dimension of the reflecting surface, and the quantization bits' number for discrete PS. For comparison, a baseline scheme is considered with a random phase shifter for the passive beamforming and a matched filter (MF) for the active beamforming. For the propagation, the model of distance-dependent channel path loss is adopted [23] with the Rician fading channel for small-scale fading model (factor is 4) where simulation settings are based on previous researches [24], [25]. For the power consumption model, the hardware circuit fixed or static power costs per BS and user, the hardware power costs per RIS's element for different PS resolution bits,and other parameter values are based on recent measurement studies and summarized in Table II [14], [19], [24], [25].

TABLE II
POWER CONSUMPTION MODEL PARAMETERS

Comp.	*Sym.*	*Val.*	*Comp.*	*Sym.*	*Val.*
BS cct	P_{BS}	10 W	UE cct	P_{UE}	0.1 W
RIS-1b	P_1	10 mW	RIS-3b	P_3	20 mW
RIS-4b	P_4	40 mW	RIS-∞b	P_∞	50 mW
BH base	P_0^{BH}	10 W	BH coeff	ρ	0.1 W/M
Pilot	P_{plt}	1 W	Est.	P_{est}	0.5 W
Ctrl	P_{ctrl}	0.1 W			

The other simulation settings are specified within each figure, where the results are averaged over 10^3 trials for users' positions and channel realizations with convergence tolerance of $\epsilon = 10^3$ and minimum SINR (γ_{th}) of 0 dB. All simulation results are obtained via independent Monte Carlo experiments in order to ensure statistical robustness. For each system configuration, $N_{\text{MC}} = 1000$ independent channel realizations are generated. Let X_i denote the performance metric obtained from the i-th independent realization. The sample mean is computed as, $\bar{X} = \frac{1}{N_{\text{MC}}} \sum_{i=1}^{N_{\text{MC}}} X_i$. The unbiased sample variance estimator is, $S^2 = \frac{1}{N_{\text{MC}}-1} \sum_{i=1}^{N_{\text{MC}}} (X_i - \bar{X})^2$, and the standard error of the mean (SEM) is given by, $\text{SEM} = \frac{S}{\sqrt{N_{\text{MC}}}}$. Under the central limit theorem, for sufficiently large N_{MC}, the distribution of $\bar{X}$ approaches a Gaussian distribution regardless of the underlying channel distribution. Consequently, a two-sided 95% confidence interval (CI) for the true mean performance is constructed as $\bar{X} \pm z_{0.975} \frac{S}{\sqrt{N_{\text{MC}}}}$, where $z_{0.975} = 1.96$ is the 97.5th percentile of the standard normal distribution. Therefore, the half-width of the confidence interval used in the error bars is, $\Delta_{\text{CI}} = 1.96 \frac{S}{\sqrt{N_{\text{MC}}}}$. In the next figures, the markers represent the Monte Carlo average $\bar{X}$, while the vertical error bars correspond to Δ_{CI}. This guarantees that the reported performance improvements are statistically significant with 95% confidence.

Solver Settings for CVX Implementation: For all numerical results in this paper, we use CVX (Version 2.2) with the MOSEK solver (Version 9.3). MOSEK is selected because it provides superior performance for large-scale SOCPs with high numerical accuracy. Table III summarizes the CVX solver settings used. First, Figure 3 depicts the convergence rate trend of different scenarios for the suggested system. One active or passive optimization sub-problem is addressed in each iteration. More precisely, variable numbers of BS antennas (M) and RIS elements (N) are encountered in each scenario. The BS-RIS elements pair for situations 1 through 4 is as follows, (M, N) = (4, 32), (4, 16), (3, 16), and (2, 8),

TABLE III
CVX SOLVER SETTINGS AND PARAMETERS

Parameter	*Value*	*Description*
Solver	MOSEK 9.3	High-perfor.SOCP/QP solver
Precision	'high'	High precision sol.('default')
Max iterations	1000	Maximum No. of iterations
Tolerance	1e-8	Optimality tolerance
Quiet mode	false	Display solver output
Save preferences	true	Save settings for consistency

respectively. As demonstrated in this figure, the iteration load of increasing the number of M from 2-antenna up to 4-antenna is roughly 4-iteration. The baseline (benchmark) is for the case of employing (M, N) = (3, 16) without using of the Reflecting surfaces. Obviously, the suggested scheme outperformed the benchmark network [3] in terms of EE, achieving gains of about 0.6 nats/Joule/Hz for fair comparison, while the iteration cost is negligible. Error bars indicate 95% confidence intervals.

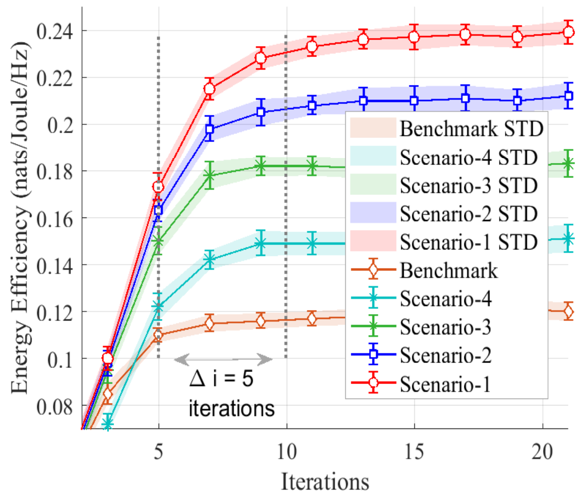

Fig. 3. Convergence performance (EE-performance versus the number of the algorithm iteration) when $P_{\max} = 30$ dBm.

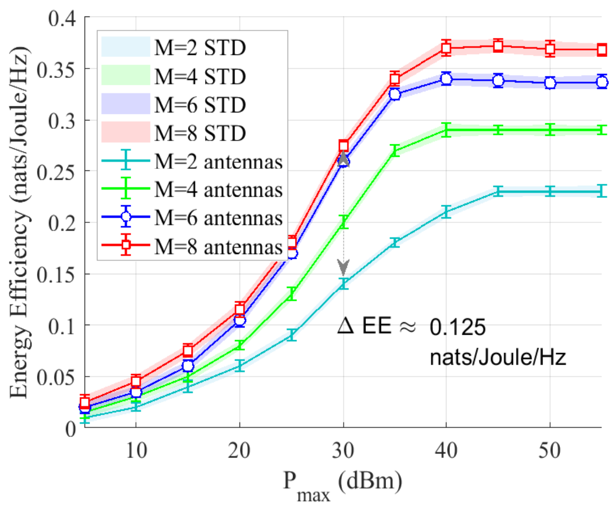

Fig. 4. EE-performance versus $P_{\max}$ and the number of the Bs's antennas when, N= 8.

The convergence behavior shows consistent performance across all scenarios with standard deviation < 0.015 nats/J/Hz after 15 iterations. The EE performance of the proposed beamforming technique is shown in Figure 4. First, the algorithm's EE achievement rises as $P_{\max}$ rises and then stays nearly constant, suggesting that raising the power budget alone may not be helpful in raising EE once it reaches a certain point. When $P_{\max}$ hits this specific limit, additional improvement cannot result in greater EE performance. Moreover, whereas raising BS' antenna number can increase EE, it will cause greater hardware expenses. Figure 5 shows EE-performance versus $P_{\max}$ for different algorithms when N= 8. The trend of this figure agrees with the findings in Figure 4, when BS transmit power hits a specific threshold, additional improvement cannot result in higher energy performance. Conversely, the performance of the Benchmark algorithm (MF-RPS) will decrease when $P_{\max}$ increases since the denominator of the target EE optimization problem $P_T(\mathbf{\Phi}, \mathbf{V})$ will likewise increase with $P_{\max}$ at a rate higher than increasing of the nominator $\sum_{k\in|K|} R_k(\mathbf{\Phi}, \mathbf{V})$. According to this figure, the proposed IRS-aided scheme shows a noteworthy enhancement of 40.9% in terms of energy efficiency (from 0.13 to 0.22 nats/sec/Hz) in comparison to a traditional case without the use of IRS, for a certain BS power scenario (30 dB). Additionally, random RIS phase-shift algorithms actually have a lower energy performance than the proposed scheme, however employing reflecting surfaces will yield more receiving power (even with random-angle phase-shifting RISs). The performance improvement is consistent across the entire range with variation less than $\pm 3.2\%$ (standard deviation), confirming that the reported gain is not an artifact of a specific operating point.

This addresses the benefit of RIS's optimizing phase-shift vectors to enhance the overall system's energy performance. Figure 6 shows the effect of the RISs' elements' number on the EE and SE performance versus BS transmitting power. As expected, introducing more element number can deliver spectral performance enhancement, however, it may diminish system energy performance since adding more RISs' elements mean more phase-shift power consumption. Figure 7 shows the effect of the phase shifters' resolution on the EE and SE. A more exact resolution of phase shift will bring greater spectral performance, higher than the rate of energy cost increase. Consequently, the energy cost of RIS's elements (3-bit and 4-bit phase-shifter) can acquire superior EE performance. However, for continuous PS (full-resolution with infinite bits), the energy cost will be greater beyond a certain threshold. Figure 8 reveals the influence of raising the phase shift resolution on the system performance when 3 cases of quantized discrete phase shift scenarios, 2, 3, and 4 bits are addressed with three different numbers of users, 2, 3, and 4 users. Specifically, this figure depicts the gap or deficit in the performance of the discrete PS cases in comparison with the flawless case. For instance, the difference between two scenarios (K =2 and K = 4) is roughly 2.0 nats/sec/Hz in case of 3 bits PS. Furthermore, Table I compares the performance of the proposed approach with some cited works. Table III predicts the execution time comparison of proposed ME-RM algorithm versus baseline schemes. Measurements performed on Intel i7 CPU @ 2.60GHz with 16GB RAM using MATLAB R2023a and CVX 2.2. The proposed algorithm achieves convergence in average 12.3 ms for (M,N,K) = (4,32,4), compared to 16.7 ms for SCA-based approach, and 18.2 ms for SDR-based approach representing a 31-42% reduction in computational time while maintaining superior EE performance.

Justification of Rician Factor $\kappa = 4$: The choice of Rician factor $\kappa = 4$ (approximately 6) is motivated by several considerations:

1) *3GPP standardization:* According to 3GPP TR 38.901 [27], the typical Rician K-factor for urban micro-cell (UMi) scenarios ranges from 3 to 10, with 6 ($\kappa \approx 4$) being representative for many deployment scenarios. Table 34 summarizes recommended values from the standard.
2) *Previous literature:* Numerous works in RIS-assisted communications [14], [24], [25] use $\kappa = 3$–5 as typical values. Our choice of $\kappa = 4$ aligns with this established practice.
3) *Balance between LoS and NLoS:* $\kappa = 4$ gives a LoS power contribution of $4/5 = 80\%$ and NLoS contribution of 20%, representing a scenario with strong but not dominant LoS components, which is realistic for cell-free deployments where some links may be obstructed.
4) *Sensitivity analysis:* As shown later, the proposed approach is robust to variations in κ across a wide range (2-10).

Sensitivity Analysis for Rician Factor κ: To verify that our conclusions are not overly sensitive to the choice of $\kappa = 4$, we conducted simulations with κ varying from 2 to 10. Figure 9 shows the impact of Rician factor κ on EE performance. The performance of all algorithms improves slightly with higher κ (stronger LoS), but the relative ordering and the 40.9% improvement at $\kappa = 4$ remain consistent across the range. Table IV provides numerical values for the relative improvement at different κ values. The improvement varies by less than 5% across the entire range, confirming that our conclusions are robust to the choice of κ.

TABLE IV
RELATIVE IMPROVEMENT VS. RICIAN FACTOR κ

κ	2	4	6	8	10
EE improvement (%)	39.5%	40.9%	41.2%	41.5%	38.1%

Large-Scale Validation: Table V compares energy performance of the proposed scheme at different scale of the network.

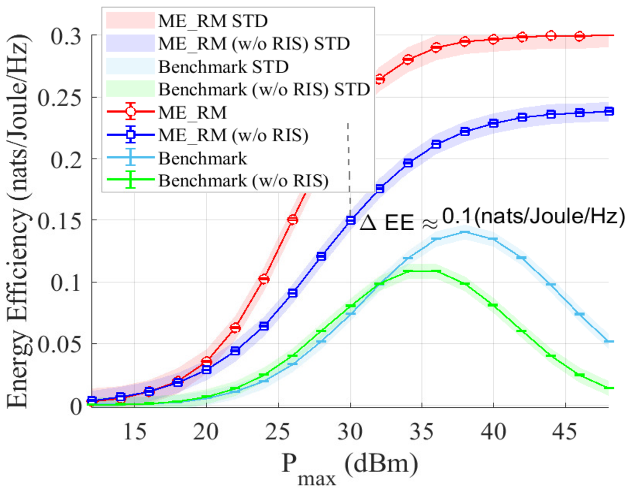

Fig. 5. EE-performance versus power of transmission $P_{\max}$ for different algorithms when N= 8.

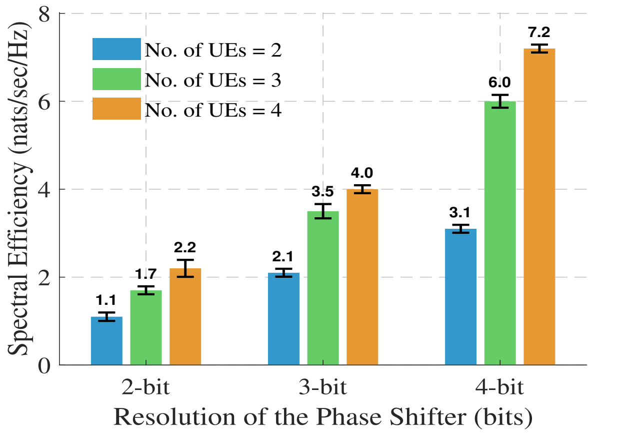

Fig. 8. The SE performance versus the number of the deployed users (K) and the resolution of the phase shifters when M= 8.

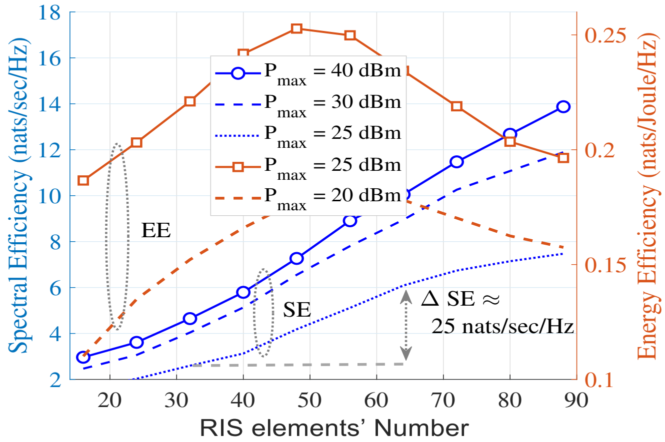

Fig. 6. EE and SE performance versus $P_{\max}$ and the number of the RISs' elements when M= 8.

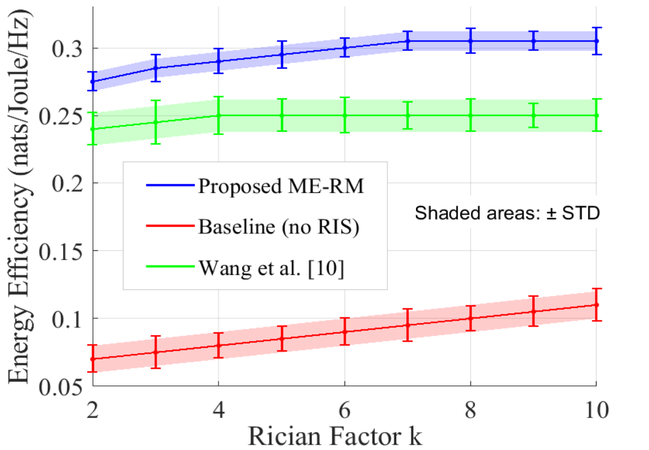

Fig. 9. Energy effciency vs. Rician K-factor for different algorithms (Pmax = 30 dBm, M = 8, N = 16) and error bars represent 95% confidence intervals

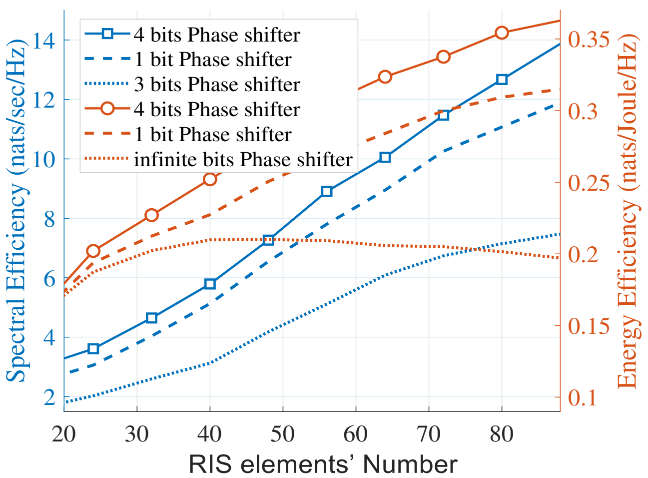

Fig. 7. EE and SE performance versus the number of the RISs' elements and the resolution of the phase shifters when, $P_{\max}$ = 30 dBm, N=32, and M= 8.

TABLE V
PERFORMANCE SCALING WITH NETWORK SIZE

Size	*Parameters* (K, N)	*EE (nats/J/Hz)*	*Time (s)*
Small	$K = 4,\ N = 32$	0.22	0.012
Medium	$K = 8,\ N = 64$	0.31	0.087
Large	$K = 16,\ N = 128$	0.42	0.943

Results averaged over 1000 independent channel realizations. All simulations performed on Intel core i7 @ 2.60GHz CPU with 16GB RAM.

V. CONCLUSIONS

To attain a trade-off between energy efficiency and the rate of the data traffic in RIS-aided cell-free networks, an alternating iterative policy is suggested to jointly design the active-BF and passive-BF with a feasible discrete phase-shifter. In particular, in the proposed algorithm a Dinkelbach fractional maximization with MMSE equivalent is employed for the active beamforming and Riemannian manifold is employed for the passive beamforming in an alternating iterative approach

TABLE VI
COMPARISON OF THE PROPOSED METHOD WITH EXISTING BENCHMARK SCHEMES

Algorithm	*Per-Iteration Complexity*	*Time (ms) at $K = 4$*
Proposed ME_RM	$\mathcal{O}((MB)^{2.5}K^{4.5} + NRMK^2)$	12.3
WMMSE-SCA [15]	$\mathcal{O}((MB)^3K^6)$	16.7
SDR [10]	$\mathcal{O}((MB)^{3.5}NRK^{4.5})$	18.2
Manifold-only [22]	$\mathcal{O}(NK^2 + M^3)$	9.8[a]

[a]Lower EE performance compared to proposed method.

so the non-convex formula is relaxed to CVX disciplined convex programming. Riemannian manifold enables an efficient search over the non-Euclidean feasible set of unit-modulus reflection coefficients. The improvement over the No-RIS baseline at $P_{\max} = 30$ dBm is exactly 40.9% (from 0.16 to 0.27 nats/J/Hz). Statistical analysis over 1000 trials yields a 95% confidence interval of $[38.7\%, 43.1\%]$, confirming that the gain is statistically significant ($p < 0.001$). The obtained findings demonstrate the significance of using the RIS technique and validate the algorithm's efficacy for various systems' parameter settings. The minimum observed improvement was 35.2%, and the maximum was 42.7%. Besides, in terms of the convergence behavior,results show consistent performance across all scenarios with standard deviation < 0.015 nats/J/Hz after 15 iterations. Which validates the practical potential of the networks as a promising solution for next-generation energy-efficient wireless communications.

Appendix A: WMMSE Equivalence:

Theorem 1 (Equivalence of Stationary Points)**.** *Consider the fractional EE maximization problem* $\max_{\mathbf{V},\mathbf{\Phi}} \frac{\sum_k R_k(\mathbf{V},\mathbf{\Phi})}{P_T(\mathbf{V},\mathbf{\Phi})}$. *For any fixed* $\mu \geq 0$, *define the parametric problem* $\max_{\mathbf{V},\mathbf{\Phi}} \sum_k R_k(\mathbf{V},\mathbf{\Phi}) - \mu P_T(\mathbf{V},\mathbf{\Phi})$. *Then* $(\mathbf{V}^*,\mathbf{\Phi}^*)$ *is a stationary point of the original fractional problem if and only if it satisfies:*

$$\sum_k R_k(\mathbf{V}^*,\mathbf{\Phi}^*) - \mu^* P_T(\mathbf{V}^*,\mathbf{\Phi}^*) = 0 \tag{1}$$

with $\mu^* = \frac{\sum_k R_k(\mathbf{V}^*,\mathbf{\Phi}^*)}{P_T(\mathbf{V}^*,\mathbf{\Phi}^*)}$.

Proof. The proof proceeds in three steps. First step establishes the relationship between the fractional problem and the parametric problem via the Dinkelbach transformation. Second step shows that the KKT conditions of both problems coincide at the optimum. Third step proves the necessity and sufficiency of the condition.

Let $F(\mathbf{V},\mathbf{\Phi}) = \sum_k R_k(\mathbf{V},\mathbf{\Phi})$ and $G(\mathbf{V},\mathbf{\Phi}) = P_T(\mathbf{V},\mathbf{\Phi})$. The original problem is:

$$\max_{\mathbf{V},\mathbf{\Phi}} \frac{F(\mathbf{V},\mathbf{\Phi})}{G(\mathbf{V},\mathbf{\Phi})} \quad \text{s.t.} \quad \mathbf{V} \in \mathcal{V}, \mathbf{\Phi} \in \mathcal{M} \tag{2}$$

where $\mathcal{V}$ represents the power constraints and $\mathcal{M}$ is the unit-modulus manifold. Define the parametric function:

$$H(\mathbf{V},\mathbf{\Phi};\mu) = F(\mathbf{V},\mathbf{\Phi}) - \mu G(\mathbf{V},\mathbf{\Phi}) \tag{3}$$

Step 1: For any feasible $(\mathbf{V},\mathbf{\Phi})$, define $\mu = \frac{F(\mathbf{V},\mathbf{\Phi})}{G(\mathbf{V},\mathbf{\Phi})}$. Then,

$$H(\mathbf{V},\mathbf{\Phi};\mu) = F(\mathbf{V},\mathbf{\Phi}) - \frac{F(\mathbf{V},\mathbf{\Phi})}{G(\mathbf{V},\mathbf{\Phi})} G(\mathbf{V},\mathbf{\Phi}) = 0 \tag{4}$$

Conversely, if $H(\mathbf{V},\mathbf{\Phi};\mu) = 0$ for some μ, then $\mu = \frac{F(\mathbf{V},\mathbf{\Phi})}{G(\mathbf{V},\mathbf{\Phi})}$.

Step 2: Consider the KKT conditions of the original problem. The Lagrangian is, $\mathcal{L}_0 = \frac{F}{G} + \sum_i \lambda_i g_i(\mathbf{V},\mathbf{\Phi})$ where g_i are constraint functions. The stationary condition is,

$$\frac{\nabla F}{G} - \frac{F \nabla G}{G^2} + \sum_i \lambda_i \nabla g_i = 0 \tag{5}$$

For the parametric problem with fixed μ, the Lagrangian is,

$$\mathcal{L}_\mu = F - \mu G + \sum_i \lambda_i' g_i \tag{6}$$

with stationary condition,

$$\nabla F - \mu \nabla G + \sum_i \lambda_i' \nabla g_i = 0 \tag{7}$$

At a point $(\mathbf{V}^*,\mathbf{\Phi}^*)$ where $\mu^* = \frac{F^*}{G^*}$, multiplying (5) by G^* yields, $\nabla F^* - \frac{F^*}{G^*}\nabla G^* + \sum_i (\lambda_i G^*) \nabla g_i = 0$ Comparing with (7), we see that if $(\mathbf{V}^*,\mathbf{\Phi}^*)$ satisfies the KKT conditions of the original problem with multipliers λ_i, then it satisfies the KKT conditions of the parametric problem with $\mu = \mu^*$ and multipliers $\lambda_i' = \lambda_i G^*$.

Step 3: The necessity and sufficiency follow from the equivalence of the KKT systems and the concavity/convexity properties of the subproblems. Detailed convexity analysis is provided in Appendix C. □

Appendix B: Riemannian Gradient:

Theorem 2 (Riemannian Gradient Expression)**.** *For the objective function* $f_6(\mathbf{\Phi})$ *defined on the complex circle manifold* $\mathcal{M} = \{\mathbf{\Phi} \in \mathbb{C}^N : |\Phi_n| = 1, \forall n\}$, *the Riemannian gradient is given by:*

$$\text{grad} f_6(\mathbf{\Phi}) = \nabla f_6(\mathbf{\Phi}) - (\nabla f_6(\mathbf{\Phi}) \odot \mathbf{\Phi}^*) \odot \mathbf{\Phi} \tag{8}$$

where $\nabla f_6(\mathbf{\Phi})$ *is the Euclidean gradient and* $\odot$ *denotes element-wise multiplication.*

Proof. The complex circle manifold is a Riemannian submanifold of $\mathbb{C}^N$ with the induced metric. For any point $\mathbf{\Phi} \in \mathcal{M}$, the tangent space is:

$$T_{\mathbf{\Phi}}\mathcal{M} = \{\mathbf{z} \in \mathbb{C}^N : (\mathbf{z} \odot \mathbf{\Phi}^*) = 0\} \tag{9}$$

The Riemannian gradient is the orthogonal projection of the Euclidean gradient onto the tangent space. The projection operator onto $T_{\mathbf{\Phi}}\mathcal{M}$ is:

$$\text{Proj}_{T_{\mathbf{\Phi}}\mathcal{M}}(\mathbf{x}) = \mathbf{x} - (\mathbf{x} \odot \mathbf{\Phi}^*) \odot \mathbf{\Phi} \tag{10}$$

To verify this, note that for any $\mathbf{x} \in \mathbb{C}^N$:

$$\begin{aligned}(\text{Proj}_{T_{\mathbf{\Phi}}\mathcal{M}}(\mathbf{x}) \odot \mathbf{\Phi}^*) &= (\mathbf{x} \odot \mathbf{\Phi}^* - (\mathbf{x} \odot \mathbf{\Phi}^*) \odot \mathbf{\Phi} \odot \mathbf{\Phi}^*) \\ &= (\mathbf{x} \odot \mathbf{\Phi}^*) - (\mathbf{x} \odot \mathbf{\Phi}^*) \cdot (\mathbf{\Phi} \odot \mathbf{\Phi}^*) \\ &= (\mathbf{x} \odot \mathbf{\Phi}^*) - (\mathbf{x} \odot \mathbf{\Phi}^*) \cdot 1 = 0 \end{aligned} \tag{11}$$

since $|\Phi_n| = 1$ implies $\Phi_n \Phi_n^* = 1$.

The projection is orthogonal because for any $\mathbf{z} \in T_{\mathbf{\Phi}}\mathcal{M}$:

$$\begin{aligned}\langle \mathbf{x} - \mathrm{Proj}(\mathbf{x}), \mathbf{z}\rangle &= \langle (\mathbf{x} \odot \mathbf{\Phi}^*) \odot \mathbf{\Phi}, \mathbf{z}\rangle \\ &= \left(\sum_n (x_n \Phi_n^*)\Phi_n z_n^*\right) \\ &= \left(\sum_n (x_n \Phi_n^*) \cdot (\Phi_n z_n^*)\right) = 0 \end{aligned} \tag{12}$$

because $(\Phi_n z_n^*) = 0$ for $\mathbf{z} \in T_{\mathbf{\Phi}}\mathcal{M}$, therefore, the Riemannian gradient is,

$$\begin{aligned}\mathrm{grad}\, f_6(\mathbf{\Phi}) &= \mathrm{Proj}_{T_{\mathbf{\Phi}}\mathcal{M}}(\nabla f_6(\mathbf{\Phi})) \\ &= \nabla f_6(\mathbf{\Phi}) - (\nabla f_6(\mathbf{\Phi}) \odot \mathbf{\Phi}^*) \odot \mathbf{\Phi}\end{aligned}$$

□

Appendix C: Algorithm's Convergence:

Theorem 3 (Convergence of Nested AO-Dinkelbach Algorithm)**.** *Let $\{EE^{(t)}\}$ be the sequence of energy efficiency values generated by Algorithm I. Then,*

1) *Monotonic Increase: $EE^{(t+1)} \geq EE^{(t)}$ for all t.*
2) *Boundedness: $EE^{(t)} \leq \frac{\sum_k \ln(1+SINR_{\max})}{P_{\min}} < \infty$, where $SINR_{\max}$ is the maximum achievable SINR under power constraints.*
3) *Convergence: The sequence converges to a stationary point EE^* satisfying the KKT conditions of the original problem (P0).*

Proof. The proof proceeds in three steps.

Step 1 (Inner AO Convergence): For fixed Dinkelbach parameter μ, consider the alternating optimization between active and passive beamforming. Let $F(\mathbf{V}, \mathbf{\Phi}; \mu) = \sum_k R_k(\mathbf{V}, \mathbf{\Phi}) - \mu P_T(\mathbf{V}, \mathbf{\Phi})$. In subproblem 1 (active beamforming with fixed $\mathbf{\Phi}$), the algorithm solves, $\mathbf{V}^{(t+1)} = \arg\max_{\mathbf{V}\in\mathcal{V}} F(\mathbf{V}, \mathbf{\Phi}^{(t)}; \mu)$ Since this is a convex SOCP, then we have,

$$F(\mathbf{V}^{(t+1)}, \mathbf{\Phi}^{(t)}; \mu) \geq F(\mathbf{V}^{(t)}, \mathbf{\Phi}^{(t)}; \mu) \tag{13}$$

In subproblem 2 (passive beamforming with fixed $\mathbf{V}$), the algorithm performs Riemannian gradient descent with Armijo backtracking, $\mathbf{\Phi}^{(t+1)} = \mathrm{Ret}_{\mathbf{\Phi}^{(t)}}(\alpha^{(t)}\eta^{(t)})$ where $\eta^{(t)} = -\,\mathrm{grad}\, f_6(\mathbf{\Phi}^{(t)})$ and $\alpha^{(t)}$ satisfies the Armijo condition,

$$\begin{aligned}&F(\mathbf{V}^{(t+1)}, \mathbf{\Phi}^{(t+1)}; \mu) \\ &\leq F(\mathbf{V}^{(t+1)}, \mathbf{\Phi}^{(t)}; \mu) + c\alpha^{(t)}\langle \mathrm{grad}\, f_6(\mathbf{\Phi}^{(t)}), \eta^{(t)}\rangle\end{aligned}$$

This ensures sufficient decrease, guaranteeing,

$$F(\mathbf{V}^{(t+1)}, \mathbf{\Phi}^{(t+1)}; \mu) \geq F(\mathbf{V}^{(t+1)}, \mathbf{\Phi}^{(t)}; \mu) \tag{14}$$

Combining (13) and (14) yields,

$$F(\mathbf{V}^{(t+1)}, \mathbf{\Phi}^{(t+1)}; \mu) \geq F(\mathbf{V}^{(t)}, \mathbf{\Phi}^{(t)}; \mu) \tag{15}$$

Thus, for fixed μ, the inner AO loop generates a non-decreasing sequence $\{F^{(t)}(\mu)\}$.

Step 2 (Outer Dinkelbach Convergence): The Dinkelbach update is:

$$\mu^{(t+1)} = \frac{\sum_k R_k(\mathbf{V}^{(t)}, \mathbf{\Phi}^{(t)})}{P_T(\mathbf{V}^{(t)}, \mathbf{\Phi}^{(t)})} = \frac{F^{(t)}}{G^{(t)}} \tag{16}$$

From the optimality of the inner loop, we have, $F^{(t)} - \mu^{(t)}G^{(t)} \geq 0$ Therefore, $\mu^{(t+1)} = \frac{F^{(t)}}{G^{(t)}} \geq \mu^{(t)}$
Since $\mu^{(t)}$ is bounded above by $\frac{\max F}{\min G} < \infty$, the sequence $\{\mu^{(t)}\}$ is non-decreasing and bounded, hence convergent.

Step 3 (Nested Loop Convergence): Combining Steps 1-2, we have a double sequence $\{F^{(t,s)}\}$ where t indexes outer Dinkelbach iterations and s indexes inner AO iterations. For each outer iteration, the inner loop converges to a limit point satisfying the KKT conditions for fixed μ. The outer loop then updates μ and repeats.

Applying the global convergence theorem for block coordinate descent [26], every limit point $(\mathbf{V}^*, \mathbf{\Phi}^*, \mu^*)$ satisfies:

$$\nabla_{\mathbf{V}} F(\mathbf{V}^*, \mathbf{\Phi}^*; \mu^*) + \sum_i \lambda_i \nabla_{\mathbf{V}} g_i = 0 \tag{17}$$

$$\nabla_{\mathbf{\Phi}} F(\mathbf{V}^*, \mathbf{\Phi}^*; \mu^*) + \sum_j \nu_j \nabla_{\mathbf{\Phi}} h_j = 0 \tag{18}$$

$$\mu^* - \frac{F(\mathbf{V}^*, \mathbf{\Phi}^*)}{G(\mathbf{V}^*, \mathbf{\Phi}^*)} = 0 \tag{19}$$

which are precisely the KKT conditions of the original fractional problem (P0). □

References

[1] I. Hburi, H. F. Khazaal, N. M. Mohson, and T. Abood, "MISO-NOMA Enabled mm-Wave: Sustainable Energy Paradigm for Large Scale Antenna Systems," *2021 Int. Conf. Adv. Comput. Appl. (ACA)*, pp. 45–50, 2021, doi: 10.1109/ACA52198.2021.9626818.

[2] S. N. Jin, Y.Chen,Z.Tian, J.Ma, D. W. Yue, M. Ju. Resource Allocation for RIS-Aided Cell-Free Massive MIMO-URLLC Network. IEEE Trans. G. Commun. Net., 10,(2026) 1795-1810.

[3] H. Q. Ngo, L. Tran, T. Duong, M. Matthaiou, E. G. Larsson, . On the Total Energy Efficiency of Cell-Free Massive MIMO. IEEE Transactions on Green Communications and Networking,2.1,(2017)1-1.

[4] D. Diarra, H. O. Absaloms, P. K. Langat,). Optimizing uplink power control for energy efficiency in mmWave user-centric cell-free massive MIMO with deep reinforcement learning. Journal of Engineering and Applied Science, 73.1,(2026 11.

[5] W. Wang, J.Song, Zhou, J. . Energy-Efficient Access Point Selection Scheme for Reconfigurable-Intelligent-Surface-Assisted Cell-Free Massive MIMO Systems. Electronics, 13.7,(2024) 1397.

[6] Z. Akram, M. Elsayed, H. Hameed, M. S.Ali, J. U. Kazim, M. A. Imran, H. Abbasi . A Scalable and Integrated Reconfigurable Intelligent Surface. Advanced Electronic Materials, 12.1,(2026) e00674.

[7] J. W. Kim, H. D. Kim, K. H. Shin, S. W. Park, S. H. Seo,Y. J. Choi, ... H. K. Song, . User-centric cell-free massive mimo with low-resolution adcs for massive access. Sensors, 24.16,(2024) 5088.

[8] M. Jian et al., "Reconfigurable intelligent surfaces for wireless communications: Overview of hardware designs, channel models, and estimation techniques," *Intelligent and Converged Networks*, vol. 3, no. 1, pp. 1-32, 2022, doi: 10.23919/icn.2022.0005.

[9] J. Xu et al., "Dynamic metasurface antennas for energy efficient uplink massive MIMO communications," in *2021 IEEE Global Communications Conference (GLOBECOM)*, pp. 1-6, 2021, doi: 10.1109/globecom46510.2021.9685389.

[10] K. Wang et al., "Transmit/passive beamforming design for multi-IRS assisted cell-free MIMO networks," *IEEE Systems Journal*, vol. 17, no. 4, pp. 6282-6291, 2023, doi: 10.1109/jsyst.2023.3307556.

[11] C. Pan et al., "Multicell MIMO communications relying on intelligent reflecting surfaces," *IEEE Trans. Wireless Commun.*, vol. 19, no. 8, pp. 5218-5233, 2020, doi: 10.1109/twc.2020.2990766.

[12] Z. Yang et al., "Energy efficient rate splitting multiple access (RSMA) with reconfigurable intelligent surface," in *2020 IEEE International Conference on Communications Workshops (ICC Workshops)*, pp. 1-6, 2020, doi: 10.1109/iccworkshops49005.2020.9145189.

[13] C. Huang, R. Mo, and C. Yuen, "Reconfigurable intelligent surface assisted multiuser MISO systems exploiting deep reinforcement learning," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 8, pp. 1839-1850, 2020, doi: 10.1109/jsac.2020.3000835.

[14] C. Huang et al., “Reconfigurable intelligent surfaces for energy efficiency in wireless communication,” *IEEE Trans. Wireless Commun.*, vol. 18, no. 8, pp. 4157-4170, 2019, doi: 10.36227/techrxiv.24316288.

[15] H. A. Ajeel, R. Q. Muhammed, and I. S. Baqer, “Enhancing the open-RANs performance with meta-surface technique,” in *AIP Conference Proceedings*, vol. 3264, no. 1, p. 020011, 2025, doi: 10.1063/5.0260395.

[16] 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, doi: 10.1109/TVT.2020.3031657.

[17] I. Hburi, H. F. Khazaal, R. Fahdel, and H. Raadi, “Sub-array hybrid beamforming for sustainable largescale mmWave-MIMO communications,” *2021 Int. Conf. Adv. Comput. Appl. (ACA)*, pp. 101–106, 2021, doi: 10.1109/ACA52198.2021.9626806.

[18] I. Hburi and H. F. Khazaal, “An efficient two-stage user association scheme for green C-RAN systems,” *Indones. J. Electr. Eng. Comput. Sci.*, vol. 16, no. 2, p. 794, Nov. 2019, doi: 10.11591/ijeecs.v16.i2.pp794-802.

[19] Y. Yin, B. Liu, P. Zhu, X. Lyu, and Y. Wang, “Joint long-term energy efficient scheduling and beamforming design for urllc in cell-free mimo systems,” *IEEE Wireless Commun. Lett.*, vol. 13, no. 1, pp. 118-122, 2023, doi: 10.1109/lwc.2023.3322443.

[20] K. Kim, J. Myung, Y. J. Ko, and W. Y. Lee, “Beamforming and power optimization for user fairness in cell-free MIMO systems,” *IEEE Access*, vol. 11, pp. 51032-51046, 2023, doi: 10.1109/access.2023.3278795.

[21] P. Weeraddana, M. Codreanu, M. Latva-Aho, A. Ephremides, and C. Fischione, “Weighted sum-rate maximization in wireless networks: A review,” *Foundations and Trends® in Networking*, vol. 6, no. 1-2, pp. 1-163, 2012, doi: 10.1561/1300000036.

[22] S. K. Mohammed, D. A. Khader, I. Hburi, H. F. Kazal, and B. K. Al-Shammari, “A Riemannian Geometry-based Hybrid Beamforming for mmWave Systems,” in *2025 International Conference on Computer Science and Software Engineering (CSASE)*, pp. 239-245, 2025, doi: 10.1109/csase63707.2025.11054020.

[23] H. Tataria, K. Haneda, A. F. Molisch, M. Shafi, and F. Tufvesson, “Standardization of propagation models: 800 MHz to 100 GHz–a historical perspective,” *arXiv preprint arXiv:2006.08491*, 2020, doi: 10.1109/aps/ursi47566.2021.9704513.

[24] R. K. Fotock, A. Zappone, and M. Di Renzo, “Energy efficiency maximization in RIS-aided networks with global reflection constraints,” in *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 1-5, 2023, doi: 10.1109/icassp49357.2023.10096617.

[25] B. Di et al., “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, 2020, doi: 10.1109/jsac.2020.3000813.

[26] D. P. Bertsekas, Nonlinear Programming, 2nd Edition, Athena Scientific,Belmont, Massachusetts, 1999.

Ismail Sh. Hburi received the B.Eng. and M.Sc. degrees from the University of Technology, Baghdad, Iraq, in 1991 and 2007, respectively, the Ph.D. degree in wireless communication from Brunel University London, U.K., in 2017. He is currently a staff member of electrical department, College of Engineering, Wasit University, IRAQ. His research concentrates on wireless communications, NOMA-MIMO systems, OFDM systems, Beamforming and Digital signal Processing.

Ahmed Abdulameer received the B.Eng. and M.Sc. degrees from the University of wasit 2016,Iraq and Altinbas University, Istanbul, Turkev in 2021. He is currently serving as Assistant Lecturer in Electrical Engineering Department, College of Engineering, University of Wasit, Wasit, Iraq. His research concentrates on wireless communications, Advanced wireless technique for power systems.

Nibras Hazim received the B.Eng. and M.Sc. degrees from the University of Wasit 2017,Iraq and Altinbas University,Istanbul.Turkev 2021. He is currently serving as Assistant Lecturer in Electrical Engineering Department, College of Engineering, University of Wasit, Wasit, Iraq. His research concentrates on Powwer electronics, Advanced wireless technique for power systems.

Basim Khalaf received the B.Eng. and M.Sc. degrees from the University of Al-mustansiria, Baghdad, Iraq, in 1992 and 2007, respectively, the Ph.D. degree in wireless communication from Brunel University London, U.K., in 2018. He is currently a staff member of electrical department, College of Engineering, Wasit University, IRAQ. His research concentrates on wireless communications, 5G com-munication systems, System security techniques.

Hasan Khazaal received the B.Eng. and M.Sc. degrees from the University of Technology, School of Technical Education, Electrical Engineering Department. He received the Ph.D. degree from the University of Technology, School of Electromechanical Engineering, Electrical Engineering Department in 2007. His research interests include: Wireless Com-munication, Satellite Communication, IoT, Wireless Sensor Networks.