Statistical CSI-Based Joint Optimization for Passive RIS-Aided Cell-Free MIMO Systems

Ghadir Dogim
German University in Cairo
ghader.mahmoud@guc.edu.eg

Ahmed El-Mahdy
German University in Cairo
ahmed.elmahdy@guc.edu.eg

Falko Dressler
Technische Universität Berlin
dressler@ccs-labs.org

***Abstract*—This paper addresses the problem of maximizing the downlink sum-rate in passive Reconfigurable Intelligent Surface (RIS)-assisted Cell-Free MIMO (CF-MIMO) networks. To circumvent the prohibitive signaling overhead inherent in instantaneous Channel State Information (CSI) acquisition, a statistical CSI-based framework is proposed, effectively minimizing estimation burdens. A tractable expression for the sum spectral efficiency is first derived, based on statistics obtained from Minimum Mean Squared Error (MMSE) channel estimation. The resulting optimization problem is non-convex and involves the joint design of Access Point (AP) transmit power and RIS phase shifts. To efficiently solve this problem, an Alternating Optimization (AO) scheme is developed to decompose it into two subproblems: power allocation, addressed via the Weighted Minimum Mean Squared Error (WMMSE) algorithm, and phase-shift optimization, tackled using a model-driven Projected Gradient Ascent (PGA) approach. Simulation results demonstrate that the proposed framework significantly enhances both spectral and energy efficiency compared to existing baseline methods, while ensuring rapid convergence and improved system reliability, thereby validating the effectiveness of the statistical CSI-driven design.**

***Index Terms*—Cell-Free MIMO, Reconfigurable Intelligent Surface (RIS), Alternating Optimization (AO), WMMSE, Projected Gradient Ascent (PGA), Statistical CSI.**

I. Introduction

Future 5G and 6G demand massive connectivity and high throughput which are severely constrained by inter-cell interference in cellular architectures. CF-MIMO overcomes this limitation through a user-centric structure where many distributed Access Points (APs) jointly serve all users, providing a uniformly high quality of service [1]. However, the practical deployment of CF-MIMO faces challenges in harsh propagation environments, where signal blockages and deep fading severely attenuate the direct links, leading to an unpredictable degradation in achievable data rates and coverage uniformity [2].

To address these issues, RISs are introduced as a low-cost, energy-efficient solution [3]. Passive RISs introduce multiplicative fading, making random phase configurations ineffective [1]. Unlocking the network potential requires complex joint optimization of AP power and RIS phases. However, acquiring the instantaneous CSI required for such optimization incurs prohibitive signaling, scaling poorly with the number of RIS elements.

To overcome the dual challenges of high Channel State Information (CSI) overhead and algorithmic complexity, this work proposes a statistical CSI-based approach. The primary contributions of this work are summarized below:

- We formulate the downlink sum-rate maximization problem for a passive RIS-assisted CF-MIMO system, relying solely on statistical CSI. This approach significantly reduces the channel estimation overhead compared to methods requiring instantaneous CSI.
- A tractable expression for the sum spectral efficiency is derived based on MMSE channel estimates and statistical CSI. This expression accounts for the effects of beamforming uncertainty and inter-user interference, enabling efficient optimization.
- We propose a low-complexity (a computationally efficient) AO framework to solve the non-convex joint optimization problem. The framework decouples the problem into two sub-problems: power control optimization solved using the WMMSE algorithm and RIS phase-shift optimization solved using a model-based PGA method.

These contributions collectively address the gap in existing literature by providing a low-overhead, model-based optimization framework that efficiently enhances both spectral and energy efficiency in RIS-aided cell-free systems using only statistical CSI.

II. Literature review

Existing research in CF-MIMO has diverged. Authors in [4] propose a two-timescale scheme, employing instantaneous CSI for APs and statistical CSI for the RIS, which still requires significant estimation overhead for the AP-to-user links. Others operating solely on statistical CSI have resorted to heuristic or complex algorithms. For example, in [5], authors tackle sum Spectral Efficiency (SE) maximization using a successive-QT algorithm combined with a computationally intensive and complex Particle Swarm Optimization (PSO) heuristic for the phase optimization.

In contrast, several recent works focus primarily on performance evaluation. For instance, the authors in [6] analyzed the Hybrid Relay-Reflecting Intelligent Surface (HR-RIS) system, but relied on simplified equal power allocation. Similarly, in [7], the authors derived closed-form expressions for the spectral efficiency of wireless-powered RIS-aided CF-MIMO networks using two-layer decoding. While providing valuable analytical insights into these complex architectures, these

studies do not propose a joint optimization framework to actively enhance the system performance.

Recent efforts addressing robustness and fairness often incur high computational overhead. For instance, [8] and [9] employed complex iterative schemes, such as the Penalty Convex-Concave Procedure (P-CCP) and Successive Convex Approximation (SCA), to maximize worst-case sum rate and fairness-based energy efficiency (EE), respectively; however, these methods require solving sequences of convex problems, introducing significant delays. Similarly, [10] tackled EE maximization via Quadratic and Lagrangian transforms but explicitly resorted to a Deep Learning alternative to mitigate the high complexity of the iterative approach.

Recent works exploit statistical CSI to reduce pilot overhead, yet gaps remain. Studies like [11] and [12] are limited to multi-cell architectures, missing the macro-diversity gains of cell-free systems. While [13] addresses the cell-free context, their work is tailored to STAR-RIS hardware and relies on meta-heuristic algorithms. Consequently, analytically tractable joint optimization for standard RIS-aided cell-free massive MIMO remains largely unexplored.

Consequently, recent approaches in CF-MIMO resort to data-driven learning methods to bypass the optimization complexity. For instance, in [14], the authors employ a Multi-Agent Deep Reinforcement Learning (MA-DRL) framework to jointly design beamforming and reflection codebooks. While effective for distributed implementation, such DRL-based solutions suffer from high training latency, lack of interpretability, and significant computational overhead during the learning phase compared to direct model-based optimization.

Based on the above survey, there is still a need for developing a model based optimization framework that has a low computational complexity and depends only on statistical CSI. In this paper, we bridge this gap by proposing a joint optimization framework based on AO to solve the resulting highly non-convex problem under statistical channel knowledge alone.

III. System and Channel Models

A. System Model

A downlink Cell-Free Multiple-Input Multiple-Output (CF-MIMO) network is considered as shown in Fig.1, where B APs, each equipped with M antennas cooperatively serve K single-antenna User Equipments (UEs). In the downlink data phase, all B APs are centrally coordinated to coherently and simultaneously transmit the data streams intended for all K UEs, enabling system-wide spatial processing. The system is assisted by R passive RISs, each consisting of N-elements. Adopting Time-Division Duplex (TDD) operation, the analysis concentrates on two distinct phases: uplink pilot training and downlink data transmission. All nodes are uniformly distributed over a $D \times D$ m^2 square area.

B. Channel Model

The effective uplink channel $\mathbf{g}_{bk} \in \mathbb{C}^{M\times 1}$ from UE k to AP b is modeled as the superposition of the direct path and the R reflected paths:

$$\mathbf{g}_{bk} = \mathbf{h}_{bk} + \sum_{r=1}^{R} \mathbf{G}_{br}\mathbf{\Theta}_r\mathbf{f}_{kr} \tag{1}$$

where $\mathbf{h}_{bk} \in \mathbb{C}^{M\times 1}$ is the direct channel from UE k to AP b, $\mathbf{f}_{kr} \in \mathbb{C}^{N\times 1}$ is the channel from UE k to RIS r, $\mathbf{G}_{br} \in \mathbb{C}^{M\times N}$ is the channel from RIS r to AP b, and $\mathbf{\Theta}_r = \text{diag}(\theta_{r1}, \ldots, \theta_{rN})$ is the diagonal matrix of complex reflection coefficients of the r-th RIS. Each coefficient can be written as $\theta_{rn} = |\theta_{rn}|e^{j\phi_{rn}}$. For a passive RIS, each element must satisfy the unit modulus constraint, i.e., $|\theta_{rn}| = 1$. Accordingly, the phase given by $\phi_{rn} = \arg(\theta_{rn})$ is the sole optimization variable.

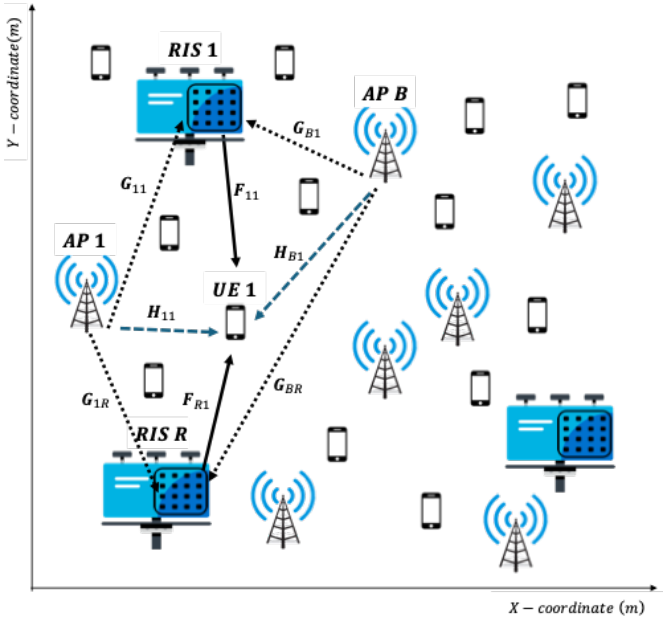

Fig. 1. CF-MIMO system model.

A Rician fading model is assumed for all links in (1), comprising a deterministic Line-of-Sight (LoS) mean component and a stochastic Non-Line-of-Sight (NLoS) scattering term.

$$\mathbf{h}_{bk} = \overline{\mathbf{h}}_{bk} + \tilde{\mathbf{h}}_{bk} \tag{2}$$

$$\mathbf{f}_{kr} = \overline{\mathbf{f}}_{kr} + \tilde{\mathbf{f}}_{kr} \tag{3}$$

$$\mathbf{G}_{br} = \overline{\mathbf{G}}_{br} + \tilde{\mathbf{G}}_{br} \tag{4}$$

where $\overline{\mathbf{h}}_{bk} \triangleq \mathbb{E}\{\mathbf{h}_{bk}\}$, $\overline{\mathbf{f}}_{kr} \triangleq \mathbb{E}\{\mathbf{f}_{kr}\}$, and $\overline{\mathbf{G}}_{br} \triangleq \mathbb{E}\{\mathbf{G}_{br}\}$ are the deterministic LoS components. The NLoS components $\tilde{\mathbf{h}}_{bk}$, $\tilde{\mathbf{f}}_{kr}$, and $\tilde{\mathbf{G}}_{br}$ are modeled as zero-mean Rayleigh fading random variables.

The total average power of the channel, determined by the large-scale fading, is ζ_{bk}. This power is split between the LoS and NLoS components based on the Rician factor κ_{bk}. The average power of the LoS component is:

$$\mathbb{E}\{||\overline{\mathbf{h}}_{bk}||^2\} = \zeta_{bk}\left(\frac{\kappa_{bk}}{\kappa_{bk}+1}\right) \tag{5}$$

The NLoS component $\tilde{\mathbf{h}}_{bk}$ has a covariance matrix $\text{Cov}(\tilde{\mathbf{h}}_{bk}) = \beta_{bk}\mathbf{I}_M$, where β_{bk} is the average power (variance) of each NLoS element:

$$\beta_{bk} = \zeta_{bk}\left(\frac{1}{\kappa_{bk}+1}\right) \tag{6}$$

The same model applies to the $\mathbf{f}_{kr}$ and $\mathbf{G}_{br}$ channels with their respective ζ and κ values.

C. Assumptions and Signaling Overhead

The system operates in TDD mode, relying on channel reciprocity. Each coherence interval of τ_c samples is divided into an uplink pilot training phase of τ_p samples and a downlink data transmission phase of $\tau_d = \tau_c - \tau_p$ samples. We assume a quasi-static block fading model where large-scale statistics remain constant over multiple intervals.

Unlike instantaneous CSI-based schemes that require estimating the cascaded channel for every RIS element (scaling as $\tau_p \geq K + N$), our framework relies solely on slowly varying statistical covariance matrices for joint optimization. This decouples the pilot overhead from the number of RIS elements, requiring only $\tau_p \geq K$. Furthermore, distinct from conventional two-timescale approaches [1], [4] that often overlook channel estimation errors in their optimization objectives, our formulation explicitly incorporates the MMSE estimation error (beamforming uncertainty) and pilot contamination into the sum-rate maximization. This ensures robustness against imperfect CSI while maximizing the data transmission duration τ_d for large RIS deployments.

IV. Problem Formulation

A. Uplink Channel Estimation

During the uplink training phase, all K UEs simultaneously transmit their assigned pilot sequences. Let $\boldsymbol{\varphi}_k \in \mathbb{C}^{\tau_p \times 1}$ be the pilot sequence for UE k. The pilot sequences are assumed to be mutually orthogonal, such that $\boldsymbol{\varphi}_k^H \boldsymbol{\varphi}_j = \delta_{kj}$, where δ_{kj} is the Kronecker delta. This orthogonality requires $\tau_p \geq K$.

Let ρ_p be the transmit power of each pilot symbol. The received signal matrix at AP b, $\mathbf{Y}_b \in \mathbb{C}^{M \times \tau_p}$, is the superposition of all UE signals and additive noise:

$$\begin{aligned}\mathbf{Y}_b &= \sum_{k=1}^{K} \sqrt{\tau_p \rho_p}\mathbf{g}_{bk}\boldsymbol{\varphi}_k^H + \mathbf{N}_b \\ &= \underbrace{\sqrt{\tau_p \rho_p}\mathbf{g}_{bj}\boldsymbol{\varphi}_j^H}_{\text{Desired Pilot Signal}} + \underbrace{\sum_{k \neq j}^{K} \sqrt{\tau_p \rho_p}\mathbf{g}_{bk}\boldsymbol{\varphi}_k^H}_{\text{Inter-User Pilot Interference}} + \mathbf{N}_b\end{aligned} \tag{7}$$

where $\mathbf{N}_b \in \mathbb{C}^{M \times \tau_p}$ is the additive white Gaussian noise (AWGN) matrix at the AP, with i.i.d. entries $\sim \mathcal{CN}(0, \sigma_n^2)$.

To estimate the channel $\mathbf{g}_{bj}$ for UE j, AP b projects the received signal matrix $\mathbf{Y}_b$ onto the pilot sequence $\boldsymbol{\varphi}_j$. Accordingly, the projected received signal at AP b, used for estimating the channel of UE j, is given as:

$$\mathbf{y}_{bj} = \mathbf{Y}_b \boldsymbol{\varphi}_j = \sqrt{\tau_p \rho_p}\mathbf{g}_{bj} + \mathbf{n}_{bj} \tag{8}$$

where the orthogonality of the pilots removes the inter-user interference. The effective noise vector $\mathbf{n}_{bj} = \mathbf{N}_b \boldsymbol{\varphi}_j$ has a covariance matrix $\text{Cov}(\mathbf{n}_{bj}) = \sigma_n^2 \mathbf{I}_M$.

Equation (8) provides a standard linear model for the channel $\mathbf{g}_{bj}$. MMSE estimation necessitates that each AP utilizes the first and second-order channel statistics, specifically the mean and covariance of g_{bk}.

The mean of the effective channel $\mathbf{g}_{bk}$ is calculated as the sum of the deterministic LoS components. Thus, using (2), (3) and (4), it is given by:

$$\boldsymbol{\mu}_{bk}(\boldsymbol{\theta}) \triangleq \mathbb{E}\{\mathbf{g}_{bk}\} = \overline{\mathbf{h}}_{bk} + \sum_{r=1}^{R} \overline{\mathbf{G}}_{br}\boldsymbol{\Theta}_r \overline{\mathbf{f}}_{kr} \tag{9}$$

The covariance matrix of the effective channel, $\mathbf{C}_{bk}(\boldsymbol{\theta}) \triangleq \text{Cov}(\mathbf{g}_{bk})$, is derived by summing the covariances of the independent NLoS components in (1). This results in:

$$\mathbf{C}_{bk}(\boldsymbol{\theta}) = \beta_{bk}\mathbf{I}_M + \sum_{r=1}^{R}\sum_{n=1}^{N} |\theta_{rn}|^2 \boldsymbol{\Psi}_{brn,krn} \tag{10}$$

where $\boldsymbol{\Psi}_{brn,krn} \triangleq \mathbb{C}\{\mathbf{g}_{brn} f_{krn}\}$ is the covariance of the single reflected path through the (r,n)-th element, given by.

$$\begin{aligned}\boldsymbol{\Psi}_{brn,krn} = (\beta_{brn}\beta_{krn} + \beta_{brn}|\overline{f}_{krn}|^2)\mathbf{I}_M \\ + \beta_{krn}\overline{\mathbf{g}}_{brn}\overline{\mathbf{g}}_{brn}^H\end{aligned} \tag{11}$$

where $\overline{\mathbf{g}}_{brn}$ and $\overline{f}_{krn}$ are the LoS components and β_{brn} and β_{krn} are the NLoS variances of the n-th element's channels. The derivation of $\boldsymbol{\Psi}_{brn,krn}$ is given in **Appendix A**.

Given these statistics, the AP computes the linear Minimum Mean Squared Error (MMSE) estimate of the channel $\mathbf{g}_{bk}$ from the projected pilot signal $\mathbf{y}_{bk}$ as:

$$\begin{aligned}\hat{\mathbf{g}}_{bk} = \boldsymbol{\mu}_{bk} + \sqrt{\tau_p \rho_p}\mathbf{C}_{bk}(\boldsymbol{\theta})(\tau_p \rho_p \mathbf{C}_{bk}(\boldsymbol{\theta}) + \sigma_n^2 \mathbf{I}_M)^{-1} \\ \times(\mathbf{y}_{bk} - \sqrt{\tau_p \rho_p}\boldsymbol{\mu}_{bk})\end{aligned} \tag{12}$$

The derivation of (12) is given in **Appendix B**. The resulting estimate $\hat{\mathbf{g}}_{bk}$ is uncorrelated with the estimation error $\tilde{\mathbf{g}}_{bk} \triangleq \mathbf{g}_{bk} - \hat{\mathbf{g}}_{bk}$. The estimate has the same mean as the true channel, i.e., $\mathbb{E}\{\hat{\mathbf{g}}_{bk}\} = \boldsymbol{\mu}_{bk}(\boldsymbol{\theta})$. The statistical properties of the estimate and the error are characterized by their respective covariance matrices, $\hat{\mathbf{C}}_{bk}(\boldsymbol{\theta})$ and $\tilde{\mathbf{C}}_{bk}(\boldsymbol{\theta})$, the derivation which is detailed in **Appendix B**. These statistical matrices are fundamental for characterizing the downlink performance.

B. Achievable downlink data rate

Relying on TDD channel reciprocity, the downlink channel from AP b to UE k is the transpose of its uplink counterpart, given by $\mathbf{g}_{bk}^T \in \mathbb{C}^{1 \times M}$. During the downlink data phase, the APs cooperatively transmit data to the UEs. Let s_k be the data symbol intended for UE k, with $\mathbb{E}\{|s_k|^2\} = 1$. The APs use the channel estimates to perform Conjugate Beamforming (CB) [15]. The signal vector transmitted from AP b is:

$$\mathbf{x}_b = \sum_{j=1}^{K} \sqrt{p_{bj}}\hat{\mathbf{g}}_{bj}^* s_j \tag{13}$$

where $\hat{\mathbf{g}}_{bj}^*$ is the beamforming vector for UE j at AP b, and p_{bj} is the corresponding power control coefficient.

The received signal at UE k, r_k, is the sum of transmissions from all APs plus noise:

$$r_k = \sum_{b=1}^{B} \mathbf{g}_{bk}^T \mathbf{x}_b + n_k \tag{14}$$

where $n_k \sim \mathcal{CN}(0, \sigma_d^2)$ is the additive white Gaussian noise at the UE, with variance σ_d^2.

1) Downlink SINR Analysis: By substituting $\mathbf{x}_b$ from (13) into (14), and separating the desired user term ($j = k$) from the interference, the received signal r_k is expressed as:

$$\begin{aligned} r_k = & \underbrace{\mathbb{E}\left\{\sum_{b=1}^{B} \sqrt{p_{bk}} \mathbf{g}_{bk}^T \hat{\mathbf{g}}_{bk}^*\right\} s_k}_{\text{Desired Signal (DS)}} + \underbrace{n_k}_{\text{Noise}} \\ & + \underbrace{\left(\sum_{b=1}^{B} \sqrt{p_{bk}} \mathbf{g}_{bk}^T \hat{\mathbf{g}}_{bk}^* - \mathbb{E}\left\{\sum_{b=1}^{B} \sqrt{p_{bk}} \mathbf{g}_{bk}^T \hat{\mathbf{g}}_{bk}^*\right\}\right) s_k}_{\text{Beamforming Uncertainty (BU)}} \\ & + \underbrace{\sum_{j \neq k}^{K} \left(\sum_{b=1}^{B} \sqrt{p_{bj}} \mathbf{g}_{bk}^T \hat{\mathbf{g}}_{bj}^*\right) s_j}_{\text{Inter-User Interference (IUI)}} \end{aligned} \tag{15}$$

where the Beamforming Uncertainty (BU) term represents the self-interference power of the desired signal s_k that arises due to the mismatch between the true channel ($\mathbf{g}_{bk}$) and the estimated channel ($\hat{\mathbf{g}}_{bk}$), which the user's receiver cannot coherently decode. To derive a tractable expression for the achievable sum rate, the powers of these components must be computed. First, based on the channel statistics derived in Section IV, the following key terms are defined:

$$u_{bk}(\boldsymbol{\theta}) \triangleq \mathbb{E}\left\{\mathbf{g}_{bk}^T \hat{\mathbf{g}}_{bk}^*\right\} = \text{trace}(\hat{\mathbf{C}}_{bk}(\boldsymbol{\theta})) + ||\boldsymbol{\mu}_{bk}(\boldsymbol{\theta})||^2 \tag{16}$$

$$v_{b,kj}(\boldsymbol{\theta}) \triangleq \mathbb{E}\left\{\mathbf{g}_{bk}^T \hat{\mathbf{g}}_{bj}^*\right\} = \boldsymbol{\mu}_{bk}^T(\boldsymbol{\theta}) \boldsymbol{\mu}_{bj}^*(\boldsymbol{\theta}), \quad \text{for } k \neq j \tag{17}$$

$$d_{b,kj}^2(\boldsymbol{\theta}) \triangleq \mathbb{E}\left\{|\mathbf{g}_{bk}^T \hat{\mathbf{g}}_{bj}^*|^2\right\} \tag{18}$$

where $u_{bk}(\boldsymbol{\theta})$ represents the average desired gain, $v_{b,kj}(\boldsymbol{\theta})$ is the average interference coupling, and $d_{b,kj}^2(\boldsymbol{\theta})$ is the average squared magnitude of the cross-gain. The detailed derivation of the cross-gain term $d_{b,kj}^2(\boldsymbol{\theta})$ in (18) is provided in **Appendix B**. This derivation utilizes the fourth-order moment expansion of complex Gaussian vectors. Using these definitions and assuming uncorrelated user symbols with unit power, the effective SINR for user k is given by $\text{SINR}_k = \mathcal{S}_k / (\mathcal{I}_k + \sigma_d^2)$

where $\mathcal{S}_k$ is the signal power and $\mathcal{I}_k$ is the total interference power. As derived in Appendix B, these are given by:

$$\mathcal{S}_k = \left|\sum_{b=1}^{B} \sqrt{p_{bk}} u_{bk}(\boldsymbol{\theta})\right|^2 \tag{19}$$

$$\begin{aligned} \mathcal{I}_k = & \underbrace{\sum_{b=1}^{B} p_{bk} \left(d_{b,kk}(\boldsymbol{\theta})^2 - |u_{bk}(\boldsymbol{\theta})|^2\right)}_{\text{Beamforming Uncertainty Power}} \\ & + \underbrace{\sum_{j \neq k}^{K} \left(\left|\sum_{b=1}^{B} \sqrt{p_{bj}} v_{b,kj}(\boldsymbol{\theta})\right|^2 + \sum_{b=1}^{B} p_{bj} \sigma_{b,kj}^2(\boldsymbol{\theta})\right)}_{\text{Inter-User Interference Power}} \end{aligned} \tag{20}$$

where $\sigma_{b,kj}^2(\boldsymbol{\theta}) \triangleq d_{b,kj}^2(\boldsymbol{\theta}) - |v_{b,kj}(\boldsymbol{\theta})|^2$ represents the variance of the interference link. The derivation of (20) is provided in **Appendix B**.

Finally, the achievable downlink data rate for UE k is:

$$R_k = \frac{\tau_d}{\tau_c} \log_2(1 + \text{SINR}_k) \tag{21}$$

V. PROPOSED JOINT OPTIMIZATION FRAMEWORK

Our objective is to maximize the sum rate $R_{\text{sum}} = \sum_k R_k$, by jointly optimizing the downlink AP power control coefficients, $\mathbf{P} \triangleq \{p_{bj}\} \geq 0$ for all b, j, and the passive phase-shift matrices of all RISs, $\boldsymbol{\Theta} \triangleq \{\boldsymbol{\Theta}_r\}$ for all r.

A. Problem Formulation

The sum rate maximization problem can be formulated as:

$$\mathcal{P}_1: \quad \max_{\mathbf{P}, \boldsymbol{\Theta}} \quad \sum_{k=1}^{K} \log_2(1 + \text{SINR}_k(\mathbf{P}, \boldsymbol{\Theta})) \tag{22}$$

$$\text{s.t.} \quad \sum_{j=1}^{K} p_{bj} u_{bj}(\boldsymbol{\theta}) \leq P_b^{\max}, \quad \forall b \tag{23}$$

$$|\theta_{rn}| = 1, \quad \forall r, n \tag{24}$$

where $\log_2(1 + \text{SINR}_k)$ is the achievable data rate for UE k, and $P_b^{\max}$ is the maximum transmit power of AP b. The MMSE estimate satisfies $\mathbb{E}\{\|\hat{\mathbf{g}}_{bj}\|^2\} = u_{bj}(\boldsymbol{\theta})$, due to the orthogonality between the estimate and the estimation error. Hence, the power constraint (23) is expressed in terms of $u_{bj}(\boldsymbol{\theta})$. Problem $\mathcal{P}_1$ is a highly non-convex due to two factors: the objective function (22) is non-convex due to the coupling of the optimization variables $\mathbf{P}$ and $\boldsymbol{\Theta}$, and the RIS phase constraint (24) is a non-convex unit-modulus constraint. To address this, a joint AO framework is proposed to decompose the original problem into two manageable sub-problems.

B. Power Control Optimization

For a fixed set of RIS phases $\boldsymbol{\Theta} = \boldsymbol{\Theta}^{(i-1)}$ (from the previous iteration), the problem reduces to finding the optimal power coefficients $\mathbf{P}$. This sub-problem, denoted by $\mathcal{P}_2$, is formulated as:

$$\mathcal{P}_2: \quad \max_{\mathbf{P} \geq 0} \quad \sum_{k=1}^{K} \log_2(1 + \text{SINR}_k(\mathbf{P}, \boldsymbol{\Theta}^{(i-1)})) \tag{25}$$

$$\text{s.t.} \quad (23)$$

Problem $\mathcal{P}_2$ is still non-convex. This is solved using the WMMSE algorithm, which finds a locally optimal solution for

TABLE I
COMPUTATIONAL COMPLEXITY COMPARISON

Algorithm	Total Complexity
Proposed AO	$\mathcal{O}\big(I_{max}(K^{3.5} + I_{GA}KLN^2)\big)$
PSO Benchmark	$\mathcal{O}\big(I_{PSO} \cdot S \cdot KLN^2\big)$

sum rate maximization [1]. The convex sub-problem in each WMMSE iteration is solved using CVX tool. The algorithm is initialized with an Equal Power Control (EPC) solution where the power at each AP is divided equally among all UEs.

C. RIS Phase-Shift Optimization

For a fixed power allocation $\mathbf{P} = \mathbf{P}^{(i)}$ (from the previous step), the problem reduces to finding the optimal RIS phases $\mathbf{\Theta}$. Then, the optimization problem $\mathcal{P}_3$ is formulated as:

$$\mathcal{P}_3: \quad \max_{\mathbf{\Theta}} \quad \sum_{k=1}^{K} \log_2(1 + \text{SINR}_k(\mathbf{P}^{(i)}, \mathbf{\Theta})) \qquad (26)$$
$$\text{s.t.} \quad (24)$$

Problem $\mathcal{P}_3$ remains non-convex due to the objective function and the unit-modulus constraint in (24). This is solved using the computationally efficient PGA method. The algorithm iteratively updates the phases in two steps:

1) **Ascent Step:** The gradient of the sum rate with respect to the phases is computed using the chain rule, this is given by:
$$\mathbf{G}_\theta = \nabla_{\boldsymbol{\theta}^*} \mathrm{R}_{\text{sum}} \qquad (27)$$
The proof of (27) is shown in **Appendix C**. A step is then taken in this direction using a step size η to get an unconstrained update $\mathbf{\Theta}_{\text{unproj}}$:
$$\mathbf{\Theta}_{\text{unproj}} = \mathbf{\Theta}^{(i)} + \eta \cdot \mathbf{G}_\theta$$
2) **Projection Step:** The unconstrained update $\mathbf{\Theta}_{\text{unproj}}$ is projected back onto the non-convex feasible set by forcing its magnitude to 1 while preserving its phase:
$$\theta_{rn}^{(i+1)} = \exp(j \cdot \arg(\theta_{\text{unproj},rn}^{(i+1)}))$$

D. Overall Algorithm and Convergence

The overall algorithm, summarized in Algorithm 1, alternates between solving $\mathcal{P}_2$ and $\mathcal{P}_3$. The sum-rate is guaranteed to be non-decreasing in each step, and thus the algorithm is guaranteed to converge to a locally optimal solution.

Complementing this theoretical stability, Table I highlights the proposed framework's efficiency. While the PSO benchmark scales with the swarm size S as $\mathcal{O}(S \cdot KLN^2)$, the proposed AO is dominated by gradient calculations, scaling as $\mathcal{O}(I_{GA}KLN^2)$. Since typically $S \gg I_{GA}$, the deterministic AO offers significantly lower computational overhead.

Proposition 1: The sequence of achievable sum-rates $\{R_{\text{sum}}(\mathbf{P}^{(i)}, \mathbf{\Theta}^{(i)})\}$ generated by the proposed Joint AO Algorithm is non-decreasing and converges to a locally optimal solution.

Proof: Convergence is guaranteed by the monotonic improvement of the alternating sub-problems. First, the WMMSE step ($\mathcal{P}_2$) ensures a non-decreasing power allocation via block coordinate descent [1]. Second, the PGA step ($\mathcal{P}_3$) guarantees ascent for the phase configuration. Consequently, the combined sequence $\{R_{\text{sum}}^{(i)}\}$ is non-decreasing and, being upper-bounded by finite transmit power limits, must converge to a stationary point. ■

Algorithm 1 Joint AO for Sum-Rate Maximization

Initialize: Set $\mathbf{\Theta}^{(0)}$ with random phases.
Calculate EPC $\mathbf{P}^{(0)}$ based on $\mathbf{\Theta}^{(0)}$.
Set iteration $i = 1$, max iterations I_{max}, tolerance R_{tol}.
repeat
  // Fix RIS, Optimize Power (Solve $\mathcal{P}_2$ in (25))
  Calculate EPC $\mathbf{P}_{\text{epc}}$ using $\mathbf{\Theta}^{(i-1)}$.
  $\mathbf{P}^{(i)} \leftarrow$ Solve $\mathcal{P}_2$ via WMMSE, initialized with $\mathbf{P}_{\text{epc}}$.
  // Fix Power, Optimize RIS (Solve $\mathcal{P}_3$ in (26))
  Calculate gradient $\mathbf{G}_\theta = \nabla_\theta R_{\text{sum}}$ using Eq. (27).
  Update phases: $\mathbf{\Theta}_{\text{step}} = \mathbf{\Theta}^{(i-1)} + \eta \cdot \mathbf{G}_\theta$.
  Project to unit-modulus: $\mathbf{\Theta}^{(i)} \leftarrow \exp(j\angle\mathbf{\Theta}_{\text{step}})$.
  Calculate $R_{\text{sum}}^{(i)} = R_{\text{sum}}(\mathbf{P}^{(i)}, \mathbf{\Theta}^{(i)})$.
  if $i > 1$ and $|R_{\text{sum}}^{(i)} - R_{\text{sum}}^{(i-1)}| < R_{\text{tol}}$ **then**
    break
  end if
  $i \leftarrow i + 1$.
until $i > I_{\text{max}}$

VI. SIMULATION RESULTS

In this section, numerical results are provided to evaluate the performance of the proposed AO algorithm. All results are averaged over 200 random user, AP, and RIS layouts.

Simulations are conducted over a 200×200 m^2 area, where all APs, UEs, and RISs are uniformly distributed. To create a challenging benchmark, a favorable downlink scenario is simulated using the 3GPP UMi LoS Rician model [16]. The K-factor is defined by $\kappa_{\text{dB}} = 13 - 0.03d$, where d is the distance between the transmitter and receiver for the corresponding link. This dominant LoS component establishes a rigorous baseline, compelling the algorithm to improve upon an already robust system configuration.

The average sum-rate of three main schemes is compared, each with both optimized WMMSE power control and standard EPC:

- **No RIS:** The baseline cell-free system.
- **Random RIS:** The RIS elements have random phases with unit amplitude.
- **Optimized RIS:** The proposed joint AO in Algorithm 1, which jointly optimizes $\mathbf{P}$ and $\mathbf{\Theta}$.

The key simulation parameters are summarized in Table II.

Fig. 2 plots the average sum-rate versus transmit power $P_b^{\max}$, comparing the proposed AO with a standard PSO benchmark [5], [13]. The AO scheme consistently yields superior rates; at $P_b^{\max} = 10$ dBm, it achieves ≈ 550 Mbit/s versus ≈ 520 Mbit/s for PSO, indicating that gradient-based

TABLE II
SIMULATION PARAMETERS

Parameter	Value
Simulation area	200×200 m^2
Number of APs (B)	30
Number of UEs (K)	5
Number of RISs (R)	10 (for RIS schemes)
Number of RIS elements (N)	Varied (10 to 120)
Antennas per AP (M)	2
System bandwidth (B_0)	20 MHz
Noise figure (NF)	9 dB
Noise power (σ_n^2, σ_d^2)	-92 dBm
Downlink Tx power ($P_b^{\max}$)	Varied (-50 to 30 dBm)
Uplink pilot power (ρ_p)	1.0 W (30 dBm)
Coherence interval (τ_c)	200 samples
Pilot length (τ_p)	5 samples (K)
Max AO iterations ($I_{\max}$)	20
AO tolerance (R_{tol})	10^{-3}

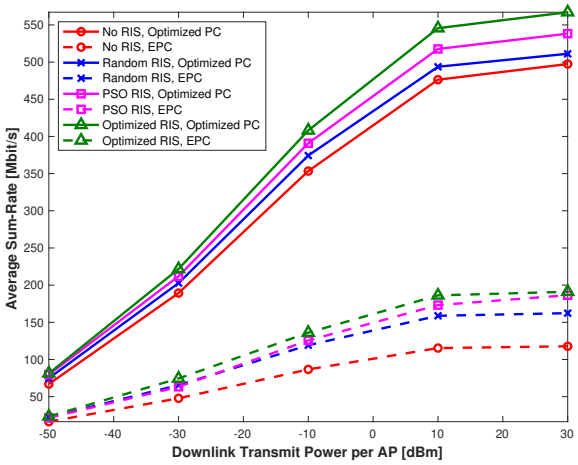

Fig. 2. Average Sum-Rate vs. AP Transmit Power $P_b^{\max}$ ($N = 40$).

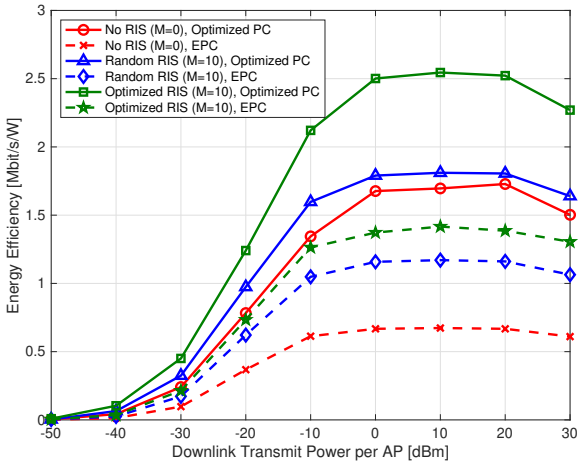

Fig. 3. Energy Efficiency vs. AP Transmit Power $P_b^{\max}$ ($N = 50$).

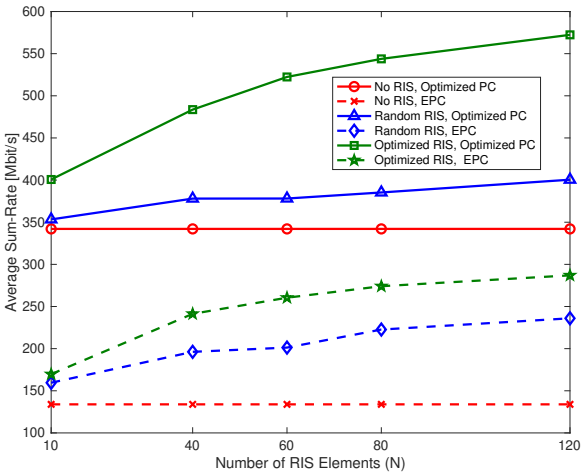

Fig. 4. Average Sum-Rate vs. N ($P_b^{\max} = -10$ dBm).

optimization locates better optima than heuristic search. Additionally, both schemes significantly outperform the 'Random' and 'No RIS' baselines. Finally, saturation occurs for $P_b^{\max} > 10$ dBm aas the system transitions from a noise-limited to an interference-limited regime.

Fig. 3 shows the Energy Efficiency (EE), defined as $\text{EE} = \frac{\sum_{k=1}^{K} R_k}{P_{\text{total}}}$ [Mbit/s/W] where P_{total} includes both the AP transmit power and the static circuit power. The EE for all schemes increases at low power, peaks, and then decreases. This is because at high $P_b^{\max}$, the sum-rate saturates due to interference, while the linear increase in power consumption causes the EE to drop. The proposed optimized RIS scheme achieves the highest EE peak, reaching ≈ 2.5 Mbit/s/W at 10 dBm. This is about $\approx 47\%$ over the 'Random RIS' peak (≈ 1.7 Mbit/s/W) and $\approx 56\%$ over the 'No RIS' peak, demonstrating that the algorithm improves both spectral and energy efficiency.

Fig. 4 investigates the sum-rate scalability as the number of RIS elements N increases. The No RIS scheme (solid red line) is constant at ≈ 340 Mbit/s, as it is independent of N. The 'Random RIS' scheme (solid blue line) shows only marginal gains, increasing from ≈ 350 to ≈ 400 Mbit/s as N grows from 10 to 120. In contrast, the proposed optimized RIS scheme (solid green line) scales exceptionally well, rising from ≈ 400 Mbit/s to over 570 Mbit/s. This demonstrates that the proposed algorithm effectively deploys the additional elements to create coherent beamforming gains.

Fig. 5 compares the fairness index of the different schemes using Jain's Fairness Index [17], defined as:

$$\mathcal{J} = \frac{\left(\sum_{k=1}^{K} R_k\right)^2}{K \sum_{k=1}^{K} R_k^2}, \tag{28}$$

The 'No RIS' and 'Random RIS' baselines exhibit poor fairness performance, yielding indices of ≈ 0.58 and ≈ 0.59,

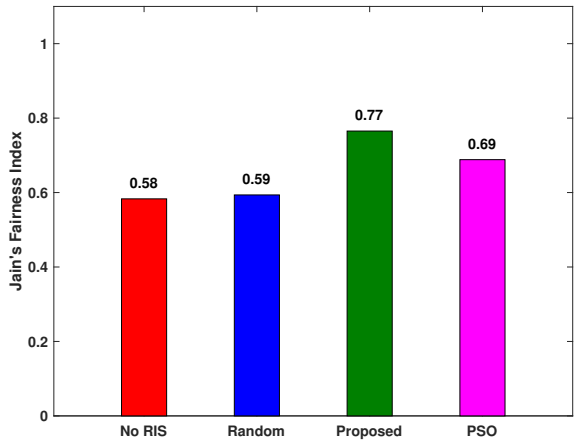

Fig. 5. Comparison of Jain's Fairness Index across different schemes.

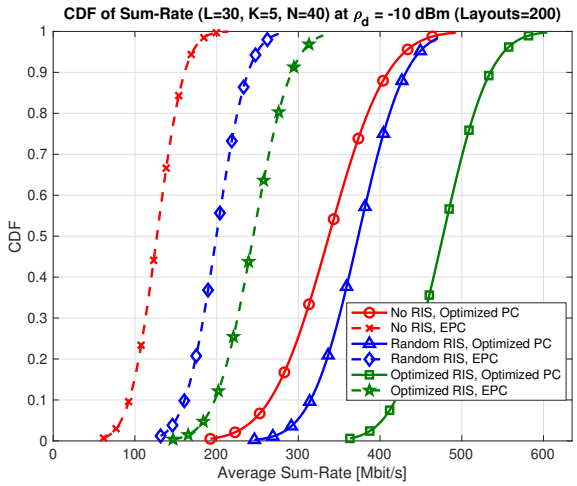

Fig. 6. CDF of Sum-Rate ($N = 40$) at $P_b^{\max} = -10$ dBm.

respectively. This low performance indicates a significant disparity in user data rates, where users suffer from a weak signal strength. In contrast, the proposed AO algorithm significantly improves the fairness index to ≈ 0.77, demonstrating its ability to effectively mitigate path loss and enhance the user experience. Crucially, the proposed method outperforms the PSO benchmark (≈ 0.69), confirming that the AO provides a more robust solution for balancing system throughput and user equity compared to the meta-heuristic approach.

Finally, the reliability and stability of the proposed algorithm are validated. Fig. 6 plots the Cumulative Distribution Function (CDF) of the sum-rate. The proposed optimized RIS scheme (solid green line) is stochastically dominant, providing a higher sum-rate at all probability levels. For example, the median (50-percentile) sum-rate for the Optimized RIS scheme is ≈ 470 Mbit/s, compared to ≈ 370 Mbit/s for Random RIS and ≈ 310 Mbit/s for No RIS. More importantly, the 10-percentile rate shows a massive improvement, rising from ≈ 100 Mbit/s (No RIS, EPC) to ≈ 380 Mbit/s with AO.

Fig. 7 validates the rapid convergence of the proposed AO algorithm. The 'Optimized RIS (AO)' scheme starts at ≈ 250 Mbit/s, which is the performance of random phases before any power optimization. After just one iteration, the algorithm achieves a sum-rate of ≈ 440 Mbit/s. The rate increases monotonically, achieving over 90% of its total gain within 3-4 iterations and converging by iteration 10, confirming the non-decreasing rate guarantee from Section V.

VII. Conclusion

In this paper, the downlink sum-rate of a passive RIS-assisted Cell-Free MIMO system was investigated. A framework was proposed to solve the non-convex problem of jointly optimizing the AP power control and RIS phase shifts. The algorithm efficiently optimizes power using WMMSE and RIS phases using PGA. Numerical results demonstrated that the proposed joint optimization scheme provides substantial gains in spectral efficiency, energy efficiency, and system reliability

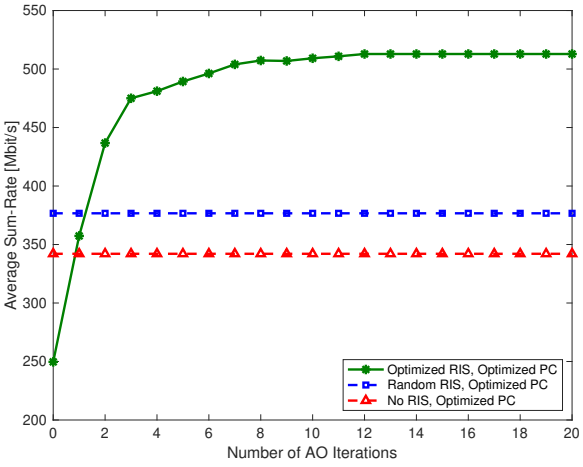

Fig. 7. Average Sum-Rate vs. number of iterations at $P_b^{\max} = -10$ dBm, $N = 40$.

over the benchmarks. Future work will address strict minimum rate constraints and the mitigation of non-ideal hardware impairments.

Appendix A
Proof of Channel Statistics

A. *Proof of RIS Channel Covariance* (11)

The term $\mathbf{\Psi}_{brn,krn}$ is the covariance of a single reflected path, $\mathbf{x} = \mathbf{g}_{brn} f_{krn}$, where $\mathbf{g}_{brn} \sim \mathcal{CN}(\overline{\mathbf{g}}_{brn}, \beta_{brn}\mathbf{I}_M)$ and $f_{krn} \sim \mathcal{CN}(\overline{f}_{krn}, \beta_{krn})$ are independent. The covariance is defined as $\mathbb{C}\{\mathbf{x}\} = \mathbb{E}\{\mathbf{x}\mathbf{x}^H\} - \mathbb{E}\{\mathbf{x}\}\mathbb{E}\{\mathbf{x}\}^H$.

The mean term is $\mathbb{E}\{\mathbf{x}\} = \mathbb{E}\{\mathbf{g}_{brn}\}\mathbb{E}\{f_{krn}\} = \overline{\mathbf{g}}_{brn}\overline{f}_{krn}$.

The correlation term $\mathbb{E}\{\mathbf{x}\mathbf{x}^H\}$ is:

$$\begin{aligned}\mathbb{E}\{\mathbf{g}_{brn}\mathbf{g}_{brn}^H\}\mathbb{E}\{|f_{krn}|^2\} &= (\mathbb{E}\{\tilde{\mathbf{g}}_{brn}\tilde{\mathbf{g}}_{brn}^H\} + \overline{\mathbf{g}}_{brn}\overline{\mathbf{g}}_{brn}^H) \\ &\quad \times (\mathbb{E}\{|\tilde{f}_{krn}|^2\} + |\overline{f}_{krn}|^2) \\ &= (\beta_{brn}\mathbf{I}_M + \overline{\mathbf{g}}_{brn}\overline{\mathbf{g}}_{brn}^H) \\ &\quad \times (\beta_{krn} + |\overline{f}_{krn}|^2) \end{aligned} \tag{29}$$

Finally, substituting these two results back into the covariance definition and simplifying gives the result in (11):

$$\begin{aligned}\mathbf{\Psi}_{brn,krn} &= (\beta_{brn}\mathbf{I}_M + \overline{\mathbf{g}}_{brn}\overline{\mathbf{g}}_{brn}^H)(\beta_{krn} + |\overline{f}_{krn}|^2) \\ &\quad - (\overline{\mathbf{g}}_{brn}\overline{f}_{krn})(\overline{\mathbf{g}}_{brn}\overline{f}_{krn})^H \\ &= (\beta_{brn}\beta_{krn} + \beta_{brn}|\overline{f}_{krn}|^2)\mathbf{I}_M + \beta_{krn}\overline{\mathbf{g}}_{brn}\overline{\mathbf{g}}_{brn}^H \end{aligned} \tag{30}$$

B. Proof of MMSE Statistics

The objective of the MMSE channel estimation is to find the estimator $\hat{\mathbf{g}}_{bk}$ of the true channel $\mathbf{g}_{bk}$ from the received signal $\mathbf{y}_{bk}$ that minimizes the Mean Squared Error (MSE). To facilitate a tractable derivation, we assume that the pilot sequences are mutually orthogonal ($\tau_p \geq K$) to eliminate inter-user interference, and that the NLoS scattering components are independent. Furthermore, we approximate the aggregate effective channel $\mathbf{g}_{bk}$ as a complex Gaussian vector, which is a standard approach in RIS analysis to enable closed-form LMMSE formulations [18]. The minimization problem is defined as:

$$\min_{\hat{\mathbf{g}}_{bk}} \quad \mathbb{E}\left\{\|\mathbf{g}_{bk} - \hat{\mathbf{g}}_{bk}\|^2\right\} \tag{31}$$

The resulting optimal MMSE estimator is the conditional mean $\hat{\mathbf{g}}_{bk} = \mathbb{E}\{\mathbf{g}_{bk} \mid \mathbf{y}_{bk}\}$. Since the channel components follow a Rician distribution, which is based on complex Gaussian processes, this optimal MMSE estimator is **equivalent to a linear estimator** (LMMSE). This allows us to express the estimate in the following tractable closed form:

$$\hat{\mathbf{g}}_{bk} = \mathbb{E}\{\mathbf{g}_{bk}\} + \text{Cov}(\mathbf{g}_{bk}, \mathbf{y}_{bk})\text{Cov}(\mathbf{y}_{bk})^{-1}(\mathbf{y}_{bk} - \mathbb{E}\{\mathbf{y}_{bk}\}) \tag{32}$$

To verify the final expression in (12), the required statistical terms are computed as follows:

1) The mean of the channel:

$$\mathbb{E}\{\mathbf{g}_{bk}\} = \boldsymbol{\mu}_{bk}(\boldsymbol{\theta}) \tag{33}$$

2) The mean of the received signal vector:

$$\mathbb{E}\{\mathbf{y}_{bk}\} = \mathbb{E}\{\sqrt{\tau_p\rho_p}\mathbf{g}_{bj} + \mathbf{n}_{bj}\} = \sqrt{\tau_p\rho_p}\boldsymbol{\mu}_{bk}(\boldsymbol{\theta}) \tag{34}$$

3) The cross-covariance between the channel and the received signal:

$$\begin{aligned}\text{Cov}(\mathbf{g}_{bk}, \mathbf{y}_{bk}) &= \mathbb{E}\left\{\mathbf{g}_{bk}(\mathbf{y}_{bk} - \mathbb{E}\{\mathbf{y}_{bk}\})^H\right\} \\ &= \sqrt{\tau_p\rho_p}\mathbf{C}_{bk}(\boldsymbol{\theta}) \end{aligned} \tag{35}$$

4) The covariance of the received signal vector:

$$\begin{aligned}\text{Cov}(\mathbf{y}_{bk}) &= \mathbb{E}\left\{(\mathbf{y}_{bk} - \mathbb{E}\{\mathbf{y}_{bk}\})(\mathbf{y}_{bk} - \mathbb{E}\{\mathbf{y}_{bk}\})^H\right\} \\ &= \tau_p\rho_p\mathbf{C}_{bk}(\boldsymbol{\theta}) + \sigma_n^2\mathbf{I}_M \end{aligned} \tag{36}$$

Substituting these terms back into (32) gives (12).

The covariance matrices for the estimate ($\hat{\mathbf{C}}_{bk}$) and the estimation error ($\tilde{\mathbf{C}}_{bk}$) are then derived. The covariance of the estimate, $\hat{\mathbf{C}}_{bk}(\boldsymbol{\theta})$, is given by the general LMMSE identity:

$$\hat{\mathbf{C}}_{bk}(\boldsymbol{\theta}) = \text{Cov}(\hat{\mathbf{g}}_{bk}(\boldsymbol{\theta})) = \text{Cov}(\mathbf{g}_{bk}, \mathbf{y}_{bk})\text{Cov}(\mathbf{y}_{bk})^{-1}\text{Cov}(\mathbf{y}_{bk}, \mathbf{g}_{bk}) \tag{37}$$

Substituting the statistical terms from (35) and (36) into this definition yields:

$$\begin{aligned}\hat{\mathbf{C}}_{bk}(\boldsymbol{\theta}) &\triangleq \text{Cov}(\hat{\mathbf{g}}_{bk}) \\ &= \tau_p\rho_p\mathbf{C}_{bk}(\boldsymbol{\theta})(\tau_p\rho_p\mathbf{C}_{bk}(\boldsymbol{\theta}) + \sigma_n^2\mathbf{I}_M)^{-1}\mathbf{C}_{bk}(\boldsymbol{\theta})^H \end{aligned} \tag{38}$$

$$\tilde{\mathbf{C}}_{bk}(\boldsymbol{\theta}) \triangleq \text{Cov}(\tilde{\mathbf{g}}_{bk}) = \mathbf{C}_{bk}(\boldsymbol{\theta}) - \hat{\mathbf{C}}_{bk}(\boldsymbol{\theta}) \tag{39}$$

APPENDIX B
DERIVATION OF TRACTABLE SINR EXPRESSIONS

The derivations for the signal and interference powers in (19)-(20) are provided here.

1) Desired Signal Power: We first derive $u_{bk}(\boldsymbol{\theta})$:

$$\begin{aligned}\mathbb{E}\left\{\mathbf{g}_{bk}^T\hat{\mathbf{g}}_{bk}^*\right\} &= \mathbb{E}\left\{(\hat{\mathbf{g}}_{bk}^T + \tilde{\mathbf{g}}_{bk}^T)\hat{\mathbf{g}}_{bk}^*\right\} \\ &= \mathbb{E}\left\{\hat{\mathbf{g}}_{bk}^T\hat{\mathbf{g}}_{bk}^*\right\} = \mathbb{E}\left\{\text{trace}(\hat{\mathbf{g}}_{bk}\hat{\mathbf{g}}_{bk}^H)\right\} \\ &= \text{trace}(\mathbb{E}\{\hat{\mathbf{g}}_{bk}\hat{\mathbf{g}}_{bk}^H\}) \\ &= \text{trace}(\hat{\mathbf{C}}_{bk}(\boldsymbol{\theta}) + \boldsymbol{\mu}_{bk}(\boldsymbol{\theta})\boldsymbol{\mu}_{bk}(\boldsymbol{\theta})^H) \\ &= \text{trace}(\hat{\mathbf{C}}_{bk}(\boldsymbol{\theta})) + ||\boldsymbol{\mu}_{bk}(\boldsymbol{\theta})||^2 \triangleq u_{bk}(\boldsymbol{\theta}) \end{aligned} \tag{40}$$

By assuming independent channels across APs, the signal power $\mathcal{S}_k$ becomes the squared coherent sum of u_{bk} terms, as given in (19).

2) Beamforming Uncertainty Power: This term is $\mathbb{E}\left\{\left|\sum_{b=1}^B \sqrt{p_{bk}}q_{bk}\right|^2\right\}$, where $q_{bk} \triangleq \mathbf{g}_{bk}^T\hat{\mathbf{g}}_{bk}^* - u_{bk}(\boldsymbol{\theta})$ is a zero-mean variable.

$$\begin{aligned}\mathbb{E}\{|\text{BU}_k|^2\} &= \mathbb{E}\left\{\left|\sum_{b=1}^B \sqrt{p_{bk}}q_{bk}\right|^2\right\} \\ &= \sum_{b=1}^B p_{bk}\mathbb{E}\{|q_{bk}|^2\} + \sum_{b\neq b'}^B \sqrt{p_{bk}p_{b'k}}\mathbb{E}\{q_{bk}q_{b'k}^*\} \end{aligned} \tag{41}$$

Assuming q_{bk} and $q_{b'k}$ are uncorrelated for $b \neq b'$, the cross-terms are zero. The remaining term is:

$$\mathbb{E}\{|q_{bk}|^2\} = \mathbb{E}\{|\mathbf{g}_{bk}^T\hat{\mathbf{g}}_{bk}^*|^2\} - |u_{bk}(\boldsymbol{\theta})|^2 \tag{42}$$

For the first term, $\mathbf{x} = \mathbf{g}_{bk} \sim \mathcal{CN}(\boldsymbol{\mu}_{bk}, \mathbf{C}_{bk})$ and $\mathbf{y} = \hat{\mathbf{g}}_{bk} \sim \mathcal{CN}(\boldsymbol{\mu}_{bk}, \hat{\mathbf{C}}_{bk})$ are approximated as independent complex Gaussian vectors. Using the 4^{th} order moment identity:

$$\begin{aligned}\mathbb{E}\{|\mathbf{x}^T\mathbf{y}^*|^2\} &= |\mathbb{E}\{\mathbf{x}^T\mathbf{y}^*\}|^2 + \text{trace}(\mathbb{C}\{\mathbf{x}\}\mathbb{C}\{\mathbf{y}\}^*) \\ &\quad + \mathbb{E}\{\mathbf{x}\}^T\mathbb{C}\{\mathbf{y}\}^*\mathbb{E}\{\mathbf{x}\}^* + \mathbb{E}\{\mathbf{y}\}^H\mathbb{C}\{\mathbf{x}\}\mathbb{E}\{\mathbf{y}\} \end{aligned} \tag{43}$$

Substituting the specific channel statistics yields:

$$\begin{aligned}\mathbb{E}\{|\mathbf{g}_{bk}^T\hat{\mathbf{g}}_{bk}^*|^2\} &= |u_{bk}(\boldsymbol{\theta})|^2 + \text{trace}(\mathbf{C}_{bk}(\boldsymbol{\theta})\hat{\mathbf{C}}_{bk}(\boldsymbol{\theta})^*) \\ &\quad + \boldsymbol{\mu}_{bk}(\boldsymbol{\theta})^T\hat{\mathbf{C}}_{bk}^*(\boldsymbol{\theta})\boldsymbol{\mu}_{bk}^*(\boldsymbol{\theta}) \\ &\quad + \boldsymbol{\mu}_{bk}(\boldsymbol{\theta})^H\mathbf{C}_{bk}(\boldsymbol{\theta})\boldsymbol{\mu}_{bk}(\boldsymbol{\theta}) \triangleq d_{b,kk}^2(\boldsymbol{\theta}) \quad (44)\end{aligned}$$

Substituting this back gives $\mathbb{E}\{|q_{bk}|^2\} = d_{b,kk}^2(\boldsymbol{\theta}) - |u_{bk}(\boldsymbol{\theta})|^2$, which matches the first term in (20).

3) Inter-User Interference Power: From the IUI term:

$$\begin{aligned}\mathbb{E}\left\{\left|\sum_{b=1}^{B}\sqrt{p_{bj}}\mathbf{g}_{bk}^T\hat{\mathbf{g}}_{bj}^*\right|^2\right\} &= \sum_{b=1}^{B}p_{bj}\mathbb{E}\left\{|\mathbf{g}_{bk}^T\hat{\mathbf{g}}_{bj}^*|^2\right\} \\ + \sum_{b\neq b'}^{B}\sqrt{p_{bj}p_{b'j}}\mathbb{E}\left\{(\mathbf{g}_{bk}^T\hat{\mathbf{g}}_{bj}^*)(\mathbf{g}_{b'k}^T\hat{\mathbf{g}}_{b'j}^*)^*\right\} \quad (45)\end{aligned}$$

Assuming independence across APs:

$$\begin{aligned}\mathbb{E}\left\{(\mathbf{g}_{bk}^T\hat{\mathbf{g}}_{bj}^*)(\mathbf{g}_{b'k}^T\hat{\mathbf{g}}_{b'j}^*)^*\right\} &= \mathbb{E}\left\{\mathbf{g}_{bk}^T\hat{\mathbf{g}}_{bj}^*\right\}\mathbb{E}\left\{(\mathbf{g}_{b'k}^T\hat{\mathbf{g}}_{b'j}^*)^*\right\} \\ &= v_{b,kj}(\boldsymbol{\theta})\cdot v_{b',kj}(\boldsymbol{\theta})^* \quad (46)\end{aligned}$$

The first term, $\mathbb{E}\left\{|\mathbf{g}_{bk}^T\hat{\mathbf{g}}_{bj}^*|^2\right\}$, is computed using the same fourth-order moment identity as above, yielding $d_{b,kj}^2(\boldsymbol{\theta})$. Rearranging the sums yields the final IUI expression in (20).

Appendix C
PGA Gradient Derivation

The gradient $\mathbf{G}_\theta = \nabla_{\boldsymbol{\theta}^*}R_{\text{sum}}$ is derived using the chain rule. The sum-rate is $R_{\text{sum}} = \sum_k \frac{\tau_d}{\tau_c}\log_2(1+\text{SINR}_k)$. Let θ_{rn}^* be the conjugate of the (r,n)-th phase shift. The element-wise gradient is:

$$\frac{\partial R_{\text{sum}}}{\partial\theta_{rn}^*} = \sum_{k=1}^{K}\frac{\tau_d}{\tau_c\ln(2)(1+\text{SINR}_k)}\frac{\partial\text{SINR}_k}{\partial\theta_{rn}^*} \quad (47)$$

Using the quotient rule on $\text{SINR}_k = \mathcal{S}_k/(\mathcal{I}_k+\sigma_d^2)$:

$$\frac{\partial\text{SINR}_k}{\partial\theta_{rn}^*} = \frac{(\mathcal{I}_k+\sigma_d^2)\frac{\partial\mathcal{S}_k}{\partial\theta_{rn}^*} - \mathcal{S}_k\frac{\partial\mathcal{I}_k}{\partial\theta_{rn}^*}}{(\mathcal{I}_k+\sigma_d^2)^2} \quad (48)$$

The terms $\mathcal{S}_k$ and $\mathcal{I}_k$ depend on $u_{bk}(\boldsymbol{\theta})$, $v_{b,kj}(\boldsymbol{\theta})$, and $d_{b,kj}^2(\boldsymbol{\theta})$. We derive the derivatives of these building blocks. Recall that the channel mean can be expanded element-wise as $\boldsymbol{\mu}_{bk}(\boldsymbol{\theta}) = \overline{\mathbf{h}}_{bk} + \sum_{r,n}\theta_{rn}\overline{\mathbf{g}}_{brn}\overline{f}_{krn}$, where $\overline{\mathbf{g}}_{brn}$ and $\overline{f}_{krn}$ are the deterministic LoS components associated with the n-th element of the r-th RIS.

The derivatives with respect to θ_{rn}^* are:

$$\begin{aligned}\frac{\partial u_{bk}(\boldsymbol{\theta})}{\partial\theta_{rn}^*} &= \text{trace}\left(\frac{\partial\hat{\mathbf{C}}_{bk}(\boldsymbol{\theta})}{\partial\theta_{rn}^*}\right) + \frac{\partial(\boldsymbol{\mu}_{bk}^H(\boldsymbol{\theta})\boldsymbol{\mu}_{bk}(\boldsymbol{\theta}))}{\partial\theta_{rn}^*} \\ &= \theta_{rn}\text{trace}(\boldsymbol{\Psi}_{brn,krn}) + (\overline{\mathbf{g}}_{brn}\overline{f}_{krn})^H\boldsymbol{\mu}_{bk}(\boldsymbol{\theta}) \quad (49)\end{aligned}$$

$$\begin{aligned}\frac{\partial d_{b,kj}^2(\boldsymbol{\theta})}{\partial\theta_{rn}^*} &= v_{b,kj}(\boldsymbol{\theta})\frac{\partial v_{b,kj}^*(\boldsymbol{\theta})}{\partial\theta_{rn}^*} + \boldsymbol{\mu}_{bk}^T(\boldsymbol{\theta})\hat{\mathbf{C}}_{bj}^*(\boldsymbol{\theta})(\overline{\mathbf{g}}_{brn}\overline{f}_{krn})^* \\ &\quad + (\overline{\mathbf{g}}_{brn}\overline{f}_{jrn})^H\mathbf{C}_{bk}(\boldsymbol{\theta})\boldsymbol{\mu}_{bj}(\boldsymbol{\theta}) \\ &\quad + \theta_{rn}\Big(\text{trace}(\boldsymbol{\Psi}_{brn,krn}\hat{\mathbf{C}}_{bj}^*(\boldsymbol{\theta})) \\ &\qquad + \text{trace}(\mathbf{C}_{bk}(\boldsymbol{\theta})\boldsymbol{\Psi}_{brn,jrj}^*)\Big) \quad (50)\end{aligned}$$

Substituting these partial derivatives back into the expressions for $\mathcal{S}_k$ and $\mathcal{I}_k$ results in $\mathbf{G}_\theta$ required for the PGA step.

References

[1] Z. Zhang and L. Dai, "A joint precoding framework for wideband reconfigurable intelligent surface-aided cell-free network," *IEEE Transactions on Signal Processing*, vol. 69, pp. 4085–4101, 2021.

[2] E. Shi, J. Zhang, H. Du, B. Ai, C. Yuen, D. Niyato, K. B. Letaief, and X. Shen, "Ris-aided cell-free massive mimo systems for 6g: Fundamentals, system design, and applications," *Proceedings of the IEEE*, vol. 112, no. 4, pp. 331–364, 2024.

[3] G. Dogim, E. A. Maher, A. El-Mahdy, and F. Dressler, "Joint active precoding for ris-aided cell-free mimo networks," in *2023 12th IFIP/IEEE International Conference on Performance Evaluation and Modeling in Wired and Wireless Networks (PEMWN)*, 2023, pp. 1–6.

[4] M. Mohammadi, H. Q. Ngo, and M. Matthaiou, "Phase-shift and transmit power optimization for ris-aided massive mimo swipt iot networks," *IEEE Transactions on Communications*, vol. 73, no. 1, pp. 631–647, 2025.

[5] M. Chakraborty, E. Sharma, H. A. Suraweera, and H. Quoc Ngo, "Analysis and optimization of ris-assisted cell-free massive mimo noma systems," *IEEE Transactions on Communications*, vol. 73, no. 4, pp. 2631–2647, 2025.

[6] N. T. Nguyen, V.-D. Nguyen, H. V. Nguyen, H. Q. Ngo, S. Chatzinotas, and M. Juntti, "Downlink throughput of cell-free massive mimo systems assisted by hybrid relay-reflecting intelligent surfaces," in *ICC 2022 - IEEE International Conference on Communications*, 2022, pp. 1475–1480.

[7] Q. Sun, X. Yu, F. Li, X. Chen, M. Xu, and J. Zhang, "Performance analysis of ris-aided wireless-powered cell-free iot networks with imperfect statistical csi," *IEEE Transactions on Communications*, vol. 73, no. 11, pp. 12 456–12 472, 2025.

[8] J. Yao, J. Xu, W. Xu, D. W. K. Ng, C. Yuen, and X. You, "Robust beamforming design for ris-aided cell-free systems with csi uncertainties and capacity-limited backhaul," *IEEE Transactions on Communications*, vol. 71, no. 8, pp. 4636–4649, 2023.

[9] W. Zhou, W. Jiao, L. Suo, and C. Li, "Max-min energy efficient optimization for ris-aided cell-free mimo systems with statistical csi," *IEEE Wireless Communications Letters*, vol. 13, no. 12, pp. 3518–3522, 2024.

[10] S.-N. Jin, D.-W. Yue, Y.-L. Chen, and Q. Hu, "Energy efficiency maximization in irs-aided cell-free massive mimo system," *IEEE Wireless Communications Letters*, vol. 12, no. 10, pp. 1652–1656, 2023.

[11] X. Li, L. Jiang, C. Luo, Y. Han, M. Matthaiou, and S. Jin, "Ris-enhanced multi-cell downlink transmission using statistical channel state information," *Science China Information Sciences*, vol. 66, no. 11, p. 212301, 2023.

[12] D. Gunasinghe and G. A. Aruma Baduge, "Achievable Rate Analysis for Multi-Cell RIS-Aided Massive MIMO With Statistical CSI-Based Optimizations," *IEEE Transactions on Wireless Communications*, vol. 23, no. 8, pp. 8117–8135, Jan. 2024.

[13] Z. Sui, H. Q. Ngo, M. Matthaiou, and L. Hanzo, "Performance analysis and optimization of star-ris-aided cell-free massive mimo systems relying on imperfect hardware," *Trans. Wireless. Comm.*, vol. 24, no. 4, p. 2925–2939, Apr. 2025. [Online]. Available: https://doi.org/10.1109/TWC.2025.3526563

[14] A. Abdallah, A. Celik, M. M. Mansour, and A. M. Eltawil, "Multi-agent drl for distributed codebook design in ris-aided cell-free massive mimo networks," *IEEE Transactions on Communications*, vol. 73, no. 5, pp. 3283–3297, 2025.

[15] T. L. Marzetta, E. G. Larsson, and H. Yang, *Fundamentals of Massive MIMO*. Cambridge University Press, 2016.

[16] "Spatial channel model for multiple input multiple output (mimo) simulations," 3rd Generation Partnership Project (3GPP), Tech. Rep. TR 25.996, 2003.

[17] R. Jain, D. M. Chiu, and W. R. Hawe, "A quantitative measure of fairness and discrimination for resource allocation in shared computer systems," *ArXiv*, vol. cs.NI/9809099, 1998. [Online]. Available: https://api.semanticscholar.org/CorpusID:1105820

[18] Özdogan, E. Björnson, and E. G. Larsson, "Massive mimo with spatially correlated rician fading channels," *IEEE Transactions on Communications*, vol. 67, no. 5, pp. 3234–3250, 2019.