

Joint Placement, Phase Shift, and Resource Allocation for Multi-AIRS-Assisted OFDMA Networks

Ahmad Kasaeyan*, Minh Dat Nguyen[†], Wei-Ping Zhu*, and Wessam Ajib[†]

*Concordia University and [†]University of Quebec at Montreal,

*a_kasaey@live.concordia.ca, [†]nguyen.minh_dat@uqam.ca, *weiping@ece.concordia.ca, and [†]ajib.wessam@uqam.ca

Abstract—We investigate a multi-aerial intelligent reflecting surface (AIRS)-assisted orthogonal frequency-division multiple access (OFDMA) downlink under per-user statistical-CSI rate constraints. This paper develops a unified framework that jointly optimizes continuous AIRS placement, reflecting phase shifts, binary subcarrier assignment, and power allocation to maximize the system sum rate under fixed power and bandwidth constraints. The resulting optimization problem is formulated as a coupled non-convex mixed-integer program. To solve it, we propose a multi-start block-coordinate-descent algorithm that alternately optimizes AIRS placement, phase shifts via semidefinite relaxation, user–AIRS association, and OFDMA resource allocation via Lagrangian dual decomposition. Simulation results show that the proposed joint design consistently outperforms baselines that omit placement optimization, omit phase-shift optimization, or omit the AIRS entirely, across varying numbers of reflecting elements, users, and AIRS units. Furthermore, instantaneous-CSI evaluation confirms the effectiveness of the statistical-CSI design, while two-timescale resource adaptation and full per-realization re-design provide benchmarks for the adaptation gain and the achievable upper bound, respectively.

Index Terms—Aerial intelligent reflecting surface (AIRS), resource allocation, block coordinate descent.

I. INTRODUCTION

Future wireless networks are increasingly required to provide high-rate and reliable coverage in locations where terrestrial infrastructure is congested, obstructed, damaged, or unavailable. Two complementary technologies have been widely considered for this purpose: intelligent reflecting surfaces (IRSs) and unmanned aerial vehicle (UAV) base stations. In an IRS-aided link, the propagation environment can be reconfigured by low-cost passive reflecting elements whose phases strengthen desired signals or suppress interference [1], with modeling and optimization methods surveyed in [2]. In parallel, UAV-mounted base stations provide on-demand coverage through controllable three-dimensional deployment, as illustrated by energy-efficient coverage optimization under QoS requirements [3].

Mounting the IRS on a UAV, forming an aerial IRS (AIRS), combines aerial deployment flexibility with passive radio-environment control: the reflecting surface can be repositioned to improve the cascaded base-station–AIRS–user links while avoiding an active radio-frequency chain at each element. The potential of UAV–IRS integration for air-ground networks has been discussed in [4]. For OFDMA systems, sum-rate

maximization with a UAV and a terrestrial IRS has been formulated in [5], where UAV trajectory, IRS scheduling, and communication resource allocation are jointly optimized. Practical limitations caused by imperfect IRS phase compensation and energy constraints have been considered in [6]. UAV-based multi-IRS networks with IRS association, UAV altitude, and phase-shift design have also been studied for sum-rate maximization in [7].

Despite this progress, the joint design of a multi-AIRS OFDMA downlink remains challenging. Because each subcarrier serves at most one user, the per-user rate depends on the binary subcarrier assignment, power, reflection phases, and AIRS placement, all of which are tightly coupled through the cascaded BS–AIRS–user channel and per-user quality-of-service (QoS) constraints; optimizing only a subset of these variables is therefore generally suboptimal.

Existing studies tackle important subsets of this coupled problem. In multi-UAV IRS-enhanced NOMA networks, UAV placement, transmit power, IRS reflection, and decoding orders are jointly optimized [8], but without binary OFDMA scheduling or per-subcarrier power control. A multi-UAV IRS-assisted OFDMA framework with Rician modeling and patch scheduling pursues fair sum rate via deep reinforcement learning [9], whereas joint placement, BS beamforming, and phase-shift design are treated for a single UAV-mounted IRS in [10]. Multiple UAV-carried RIS panels are embedded in a massive-MIMO network in [11], and trajectory–phase design targets IRS–UAV covert communication in [12]. To the best of our knowledge, however, no AIRS study jointly optimizes *binary* OFDMA subcarrier assignment, per-subcarrier power, continuous AIRS placement, and reflection phases under explicit per-user QoS in a multi-AIRS downlink. This paper addresses this gap through the following contributions.

- We formulate a downlink multi-AIRS OFDMA problem that jointly optimizes AIRS placement, reflection phases, binary subcarrier assignment, and power under per-user QoS, coupling aerial deployment, passive beamforming, and OFDMA resource allocation. The resulting formulation uses penalized slack variables to handle residual infeasibility [13].
- We propose a multi-start BCD algorithm whose phase, placement, and resource blocks use penalized SDR, trust-

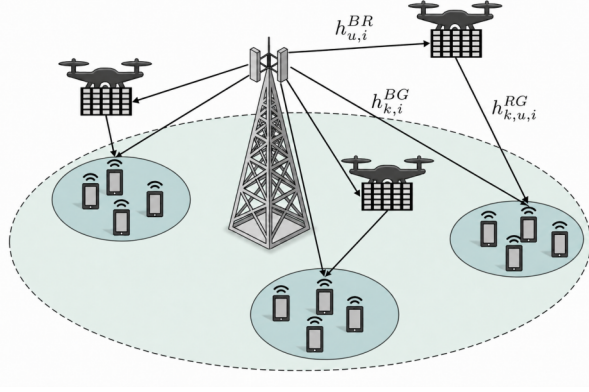


Fig. 1. System model of an AIRS-assisted OFDMA downlink.

region SCA, and a Lagrangian dual method, respectively, together with channel-aware user-AIRS association.

- Through ablation studies against baselines that omit placement optimization, omit phase-shift optimization, or omit the AIRS entirely, together with Monte Carlo evaluations under instantaneous CSI, we quantify the benefits of joint optimization and validate the robustness of the statistical-CSI-based design.

The remainder of the paper is organized as follows. Sec. II presents the system model and problem formulation. Sec. III develops the proposed algorithm. Sec. IV reports the numerical results, and Sec. V concludes the paper.

II. SYSTEM MODEL AND PROBLEM FORMULATION

We consider a multi-AIRS-assisted OFDMA downlink, illustrated in Fig. 1. A single-antenna base station (BS), located at $\mathbf{q}^B = [x_B, y_B, z_B]$, serves K single-antenna ground users at positions $\mathbf{q}_k^G = [x_k, y_k, z_k]$, $k \in \mathcal{K} \triangleq \{1, \dots, K\}$, through the assistance of U aerial IRSs at $\mathbf{q}_u^R = [x_u, y_u, z_u]$, $u \in \mathcal{U} \triangleq \{1, \dots, U\}$, each flying at a common altitude $z_u = H_0$. Each AIRS hosts $M = N^2$ passive reflecting elements arranged in an $N \times N$ uniform planar array (UPA) parallel to the ground, with inter-element spacing d_e . The reflection of the u -th AIRS is modeled by the diagonal phase-shift matrix $\Theta_u = \text{diag}(\theta_u)$, where $\theta_u = [e^{j\vartheta_{u,1}}, \dots, e^{j\vartheta_{u,M}}]^T$ and $\vartheta_{u,m} \in [0, 2\pi)$. The users are initially partitioned into U disjoint clusters via a channel-aware K -means association map $\pi: \mathcal{K} \rightarrow \mathcal{U}$, which is later refined inside the optimization loop.

A. Channel Model

The total bandwidth B is divided into I orthogonal subcarriers of spacing Δf , and each subcarrier is allocated to at most one user. Because the AIRSs operate at high altitude, the BS-AIRS links are modeled as pure LoS, while the BS-user and AIRS-user links are Rician, following [5]. For subcarrier $i \in \mathcal{I} \triangleq \{1, \dots, I\}$, the BS-AIRS channel $\mathbf{h}_{u,i}^{BR} \in \mathbb{C}^{M \times 1}$ is

$$\mathbf{h}_{u,i}^{BR} = \sqrt{\frac{\beta}{(d_u^{BR})^2}} e^{-j2\pi \frac{\Delta f}{c} d_u^{BR}} (\mathbf{a}_N(u_x^{BR}) \otimes \mathbf{a}_N(u_y^{BR})), \quad (1)$$

where $d_u^{BR} = \|\mathbf{q}^B - \mathbf{q}_u^R\|$, β is the channel gain at 1 m, c is the speed of light, and $\mathbf{a}_N(u) =$

$[1, e^{-j2\pi d_e u/\lambda}, \dots, e^{-j2\pi d_e (N-1)u/\lambda}]^T$ is the steering vector with carrier wavelength λ , and $u_x^{BR} = (x_B - x_u)/d_u^{BR}$, $u_y^{BR} = (y_B - y_u)/d_u^{BR}$ are the direction cosines.

The direct BS-user channel $h_{k,i}^{BG} \in \mathbb{C}$ and the AIRS-user channel $\mathbf{h}_{k,u,i}^{RG} \in \mathbb{C}^{M \times 1}$ admit the Rician decompositions

$$h_{k,i}^{BG} = \sqrt{\frac{\beta}{(d_k^{BG})^\alpha}} \left(\sqrt{\frac{K_1}{1+K_1}} \bar{h}_{k,i}^{BG} + \sqrt{\frac{1}{1+K_1}} \tilde{h}_{k,i}^{BG} \right), \quad (2)$$

$$\mathbf{h}_{k,u,i}^{RG} = \sqrt{\frac{\beta}{(d_{k,u}^{RG})^{\alpha'}}} \left(\sqrt{\frac{K_2}{1+K_2}} \bar{\mathbf{h}}_{k,u,i}^{RG} + \sqrt{\frac{1}{1+K_2}} \tilde{\mathbf{h}}_{k,u,i}^{RG} \right), \quad (3)$$

where $d_k^{BG} = \|\mathbf{q}^B - \mathbf{q}_k^G\|$, $d_{k,u}^{RG} = \|\mathbf{q}_u^R - \mathbf{q}_k^G\|$, (α, α') are the BS-user and AIRS-user path-loss exponents, (K_1, K_2) are the corresponding Rician K -factors, where $\tilde{h}_{k,i}^{BG} \sim \mathcal{CN}(0, 1)$ and $\tilde{\mathbf{h}}_{k,u,i}^{RG} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_M)$ are the scalar and vector NLoS components, respectively. The LoS components are $\bar{h}_{k,i}^{BG} = e^{-j2\pi \frac{\Delta f}{c} d_k^{BG}}$ and $\bar{\mathbf{h}}_{k,u,i}^{RG} = e^{-j2\pi \frac{\Delta f}{c} d_{k,u}^{RG}} (\mathbf{a}_N(u_x^{RG}) \otimes \mathbf{a}_N(u_y^{RG}))$, with u_x^{RG}, u_y^{RG} defined in the same manner as u_x^{BR}, u_y^{BR} .

B. Effective Channel and Achievable Rate

For user k served by AIRS $u = \pi(k)$ on subcarrier i , the effective end-to-end channel power is $c_{k,i} = |h_{k,i}^{BG} + (\mathbf{h}_{k,u,i}^{RG})^H \Theta_u \mathbf{h}_{u,i}^{BR}|^2$. Following [8], taking expectation over the small-scale fading yields the deterministic surrogate

$$\mathbb{E}\{c_{k,i}\} = |\hat{h}_{k,i}^{BG} + (\hat{\mathbf{h}}_{k,u,i}^{RG})^H \Theta_u \mathbf{h}_{u,i}^{BR}|^2 + \frac{\beta - \kappa_1}{(d_k^{BG})^\alpha} + \frac{M\beta(\beta - \kappa_2)}{(d_{k,u}^{RG})^{\alpha'} (d_u^{BR})^2}, \quad (4)$$

where the LoS-coherent components are $\hat{h}_{k,i}^{BG} \triangleq \sqrt{\kappa_1 / (d_k^{BG})^\alpha} \bar{h}_{k,i}^{BG}$, $\hat{\mathbf{h}}_{k,u,i}^{RG} \triangleq \sqrt{\kappa_2 / (d_{k,u}^{RG})^{\alpha'}} \bar{\mathbf{h}}_{k,u,i}^{RG}$, with $\kappa_1 = \frac{K_1 \beta}{1+K_1}$ and $\kappa_2 = \frac{K_2 \beta}{1+K_2}$. The corresponding SNR is $\text{SNR}_{k,i} = p_{k,i} \mathbb{E}\{c_{k,i}\} / \sigma^2$, where $p_{k,i}$ is the power allocated to user k on subcarrier i and $\sigma^2 = N_0 \Delta f$ is the per-subcarrier noise power. The achievable rate of user k is

$$R_k = \sum_{i=1}^I s_{k,i} \log_2(1 + \text{SNR}_{k,i}), \quad (5)$$

where $s_{k,i} \in \{0, 1\}$ indicates that subcarrier i is assigned to user k . Reflections from non-associated AIRSs are neglected.

C. Problem Formulation

Let $\mathbf{Q} = \{\mathbf{q}_u^R\}_u$, $\Theta = \{\Theta_u\}_u$, $\mathbf{S} = \{s_{k,i}\}_{k,i}$, and $\mathbf{P} = \{p_{k,i}\}_{k,i}$. The joint sum-rate maximization problem is formulated as

$$(\mathbf{P}): \max_{\mathbf{Q}, \Theta, \mathbf{S}, \mathbf{P}} \sum_{k=1}^K R_k \quad (6)$$

$$\text{s.t. } s_{k,i} \in \{0, 1\}, \forall k, i, \quad \sum_{k=1}^K s_{k,i} \leq 1, \forall i, \quad (6a)$$

$$p_{k,i} \geq 0, \quad \sum_{k,i} p_{k,i} \leq P_{\max}, \quad (6b)$$

$$R_k \geq R_{\min}, \forall k, \quad (6c)$$

$$\vartheta_{u,m} \in [0, 2\pi), \forall u, m. \quad (6d)$$

Constraint (6a) enforces exclusive binary subcarrier allocation; (6b) caps the total transmit power at $P_{\max}$; (6c) imposes a minimum per-user statistical-CSI surrogate-rate target $R_{\min}$; and (6d) restricts each phase shift to $[0, 2\pi)$. To handle potential infeasibility of (6c), we introduce non-negative slacks $\zeta = \{\zeta_k\}_{k=1}^K$ and adopt the slack-penalty reformulation [14]

$$\begin{aligned} (\mathbf{P}'): \quad & \max_{\mathbf{Q}, \boldsymbol{\Theta}, \mathbf{S}, \mathbf{P}, \zeta} \sum_{k=1}^K R_k - \Lambda \sum_{k=1}^K \zeta_k \quad (7) \\ \text{s.t.} \quad & (6a), (6b), (6d), \\ & R_k + \zeta_k \geq R_{\min}, \zeta_k \geq 0, \forall k, \quad (7a) \end{aligned}$$

where $\Lambda > 0$ is geometrically ramped between $\Lambda_{\min}$ and $\Lambda_{\max}$ across outer iterations to drive $\sum_k \zeta_k \rightarrow 0$ at convergence. Problem (7) is a non-convex mixed-integer program due to the binary $\mathbf{S}$ and the nonlinear coupling between $\boldsymbol{\Theta}$ and $\mathbf{Q}$. The next section develops a tractable BCD-based algorithm.

III. PROPOSED ALGORITHM

Problem (7) is jointly non-convex in $(\mathbf{Q}, \boldsymbol{\Theta}, \mathbf{S}, \mathbf{P})$ and contains binary variables. Therefore, we solve it using a multi-start block coordinate descent (BCD) algorithm that alternates among four blocks: phase $\boldsymbol{\Theta}$, placement $\mathbf{Q}$, resource allocation $(\mathbf{P}, \mathbf{S})$, and user-AIRS association π . Each block is replaced by a tractable convex surrogate, while an outer loop gradually increases the penalty on the QoS slack variables to drive them toward zero.

A. Phase-Shift Block ($\boldsymbol{\Theta}$)

With $(\mathbf{Q}, \mathbf{S}, \mathbf{P})$ fixed, the deterministic LoS component of the cascaded BS-AIRS-user channel can be written as $b_{k,u,i} \triangleq (\hat{\mathbf{h}}_{k,u,i}^{\text{RG}})^H \text{diag}(\mathbf{h}_{u,i}^{\text{BR}})$, which gives $|\hat{h}_{k,i}^{\text{BG}} + (\hat{\mathbf{h}}_{k,u,i}^{\text{RG}})^H \boldsymbol{\Theta}_u \mathbf{h}_{u,i}^{\text{BR}}|^2 = |\hat{h}_{k,i}^{\text{BG}} + b_{k,u,i} \boldsymbol{\theta}_u|^2$. Defining $\bar{\mathbf{h}}_{k,u,i}^H = [b_{k,u,i}, \hat{h}_{k,i}^{\text{BG}}]$ and $\mathbf{v}_u = [\boldsymbol{\theta}_u^T, 1]^T$, and introducing the lift $\mathbf{V}_u = \mathbf{v}_u \mathbf{v}_u^H$, the expected channel power (4) becomes

$$\mathbb{E}\{c_{k,i}\} = \text{Tr}(\mathbf{H}_{k,u,i} \mathbf{V}_u) + \bar{\omega}_{k,u}, \quad (8)$$

with $\mathbf{H}_{k,u,i} = \bar{\mathbf{h}}_{k,u,i} \bar{\mathbf{h}}_{k,u,i}^H$ and $\bar{\omega}_{k,u} = \frac{\beta - \kappa_1}{(d_k^{\text{BG}})^\alpha} + \frac{M\beta(\beta - \kappa_2)}{(d_{k,u}^{\text{RG}})^{\alpha'} (d_u^{\text{BR}})^2}$. Each $\mathbf{V}_u \succeq 0$ is Hermitian with unit diagonal and rank one. The phase block then reads

$$\begin{aligned} (\mathbf{P2}): \quad & \min_{\{\mathbf{V}_u\}, \zeta} - \sum_k R_k + \Lambda \sum_{k=1}^K \zeta_k \quad (9) \\ \text{s.t.} \quad & (7a), [\mathbf{V}_u]_{mm} = 1, \mathbf{V}_u \succeq 0, \text{rank}(\mathbf{V}_u) = 1. \end{aligned}$$

For any nonzero $\mathbf{V}_u \succeq 0$, the gap $\|\mathbf{V}_u\|_* - \|\mathbf{V}_u\|_2 \geq 0$ is zero if and only if $\text{rank}(\mathbf{V}_u) = 1$, where $\|\cdot\|_*$ and $\|\cdot\|_2$ denote the nuclear norm (sum of singular values) and the spectral norm (largest singular value), respectively. This yields the penalized problem

$$\begin{aligned} (\mathbf{P2}'): \quad & \min_{\{\mathbf{V}_u\}, \zeta} - \sum_k R_k + \Lambda \sum_{k=1}^K \zeta_k + \zeta_v \sum_{u=1}^U (\|\mathbf{V}_u\|_* - \|\mathbf{V}_u\|_2) \quad (10) \\ \text{s.t.} \quad & (7a), [\mathbf{V}_u]_{mm} = 1, \mathbf{V}_u \succeq 0, \forall u, m. \end{aligned}$$

where $\zeta_v > 0$ is the rank-penalty weight, ramped between $\zeta_{v,\min}$ and $\zeta_{v,\max}$ across inner iterations to drive each $\mathbf{V}_u$ toward rank one. Since $-\|\mathbf{V}_u\|_2$ is concave, it is replaced by its first-order expansion at the current iterate $\mathbf{V}_u^{(n)}$, yielding a convex SDP that is solved by CVX-MOSEK [15] at each inner iteration. The reflection vector is then recovered from the dominant eigenvector of $\mathbf{V}_u^*$, followed by per-element unit-modulus projection.

B. Placement and User-AIRS Reassociation ($\mathbf{Q}, \pi$)

With $(\boldsymbol{\Theta}, \mathbf{S}, \mathbf{P})$ fixed and altitudes fixed at H_0 , only the horizontal coordinates (x_u, y_u) are optimized. To convexify the distance-dependent terms in (4), we introduce squared-distance upper bounds $\mu_{k,u} \geq \|\mathbf{q}_u^{\text{R}} - \mathbf{q}_k^{\text{G}}\|^2$ and $\mu_u \geq \|\mathbf{q}^{\text{B}} - \mathbf{q}_u^{\text{R}}\|^2$, and write a lower bound on the per-subcarrier expected gain as

$$\eta_{k,i}^{\text{lb}} = |\hat{h}_{k,i}^{\text{BG}}|^2 + \frac{C_{k,u,i}^{\text{a}}}{\mu_{k,u}^{\alpha'/2} \mu_u} + \frac{2C_{k,u,i}^{\text{b}}}{\mu_{k,u}^{\alpha'/4} \mu_u^{1/2}} + \frac{\beta - \kappa_1}{(d_k^{\text{BG}})^\alpha}, \quad (11)$$

Inequality (11) is obtained by substituting the squared-distance bounds $\mu_{k,u}$ and μ_u into the expected gain (4) and by locally lower-bounding the resulting distance-dependent terms around the current AIRS locations. Here, following the distance-normalized cascaded-channel convention used in [8], the LoS cascaded gain is evaluated at iteration n as

$$\bar{b}_{k,u,i}^{(n)} \triangleq \sqrt{\kappa_2 \beta} (\bar{\mathbf{h}}_{k,u,i}^{\text{RG}})^H \boldsymbol{\Theta}_u^{(n)} \bar{\mathbf{h}}_{u,i}^{\text{BR}}, \quad (12)$$

where $\bar{\mathbf{h}}_{k,u,i}^{\text{RG}}$ and $\bar{\mathbf{h}}_{u,i}^{\text{BR}}$ denote the corresponding LoS steering vectors without distance-dependent path loss. The constants

$$C_{k,u,i}^{\text{a}} = |\bar{b}_{k,u,i}^{(n)}|^2 + M\beta(\beta - \kappa_2), \quad (13)$$

$$C_{k,u,i}^{\text{b}} = \Re\{(\hat{h}_{k,i}^{\text{BG}})^* \bar{b}_{k,u,i}^{(n)}\} \quad (14)$$

are frozen within a single placement update. The first distance term is linearized by a first-order Taylor expansion at $(\mu_{k,u}^{(n)}, \mu_u^{(n)})$. The second term is handled according to the sign of $C_{k,u,i}^{\text{b}}$: for $C_{k,u,i}^{\text{b}} \geq 0$ the Taylor expansion remains a valid lower bound; for $C_{k,u,i}^{\text{b}} < 0$ we invoke the arithmetic-geometric mean inequality $2\sqrt{ab} \leq \rho a + b/\rho$, with $\rho^{(n)} = \sqrt{(\mu_{k,u}^{(n)})^{\alpha'/2} / \mu_u^{(n)}}$ yielding a convex bound that can be handled by standard convex solvers. The resulting subproblem (P-Q) is convex and solved by CVX-MOSEK [15]. Since the bounds are local surrogates constructed around the current AIRS locations, the placement update is restricted to a neighborhood in which these approximations remain reliable. Specifically, we impose the adaptive trust-region constraint

$$\|\mathbf{q}_u^{\text{R}} - (\mathbf{q}_u^{\text{R}})^{(n)}\| \leq \delta_r.$$

Each candidate is accepted only if the *true* sum rate from (4) does not decrease, in which case δ_r is enlarged; otherwise it is rejected and δ_r reduced. This guarantees a monotonically non-decreasing sequence of true sum rates within the placement block.

After the placement update, the initial K -means association may no longer be channel-optimal. For every pair (k, u) we

evaluate the score $J_{k,u} = \frac{1}{I} \sum_i |\hat{h}_{k,i}^{\text{BG}} + (\hat{\mathbf{h}}_{k,u,i}^{\text{RG}})^H \mathbf{\Theta}_u \mathbf{h}_{u,i}^{\text{BR}}|^2 + \bar{\omega}_{k,u}$, i.e., the mean per-subcarrier channel power user k would see from AIRS u under the current design, and reassign k to $\arg \max_u J_{k,u}$ only if the score improves by more than a fixed margin (5% in our experiments), which prevents association oscillation across outer iterations; an empty-cluster safeguard reassigns one user to any AIRS that would otherwise have no associated users.

C. Subcarrier and Power Block (S, P)

With $(\mathbf{Q}, \mathbf{\Theta})$ fixed and effective gains $g_{k,i} \triangleq \mathbb{E}\{c_{k,i}\}/\sigma^2$, the subcarrier and power block reduces to

$$\begin{aligned} \text{(P3): } \max_{\mathbf{S}, \mathbf{P}, \zeta} \quad & \sum_k R_k - \Lambda \sum_k \zeta_k, \\ \text{s.t. } \quad & \text{(6a), (6b), (7a).} \end{aligned} \quad (15)$$

A direct mixed-integer solve scales poorly with IK , so we exploit the OFDMA decoupling and apply Lagrangian dual decomposition [16]. Dualizing (7a) with multipliers $\lambda \in [0, \Lambda]^K$, where the slack penalty Λ upper-bounds the optimal QoS duals, and (6b) with $\nu \geq 0$, the partial Lagrangian decouples across (i, k) and admits the closed-form water-filling solution

$$p_{k,i}^*(\lambda, \nu) = \left[\frac{1+\lambda_k}{\nu \ln 2} - \frac{1}{g_{k,i}} \right]^+, \quad (16)$$

$$s_{k,i}^*(\lambda, \nu) = \mathbf{1}\{k = \arg \max_{k'} \phi_{k',i}\}, \quad (17)$$

with score $\phi_{k,i} = (1+\lambda_k) \log_2(1+g_{k,i}p_{k,i}^*) - \nu p_{k,i}^*$. The duals are updated by an outer loop: ν by geometric bisection so that $\sum_{k,i} p_{k,i}^* = P_{\max}$, and the QoS duals λ_k are updated by projected subgradient steps over $[0, \Lambda]$. The dual iterations are warm-started from an LP relaxation of (P3), followed by a repair step and a final continuous power-allocation update under the recovered binary assignment. This avoids solving a mixed-integer linear program (MILP) while preserving the OFDMA assignment structure.

D. Overall Algorithm and Complexity Analysis

The proposed method alternately updates the phase shifts, AIRS locations, user–AIRS associations, and OFDMA resource variables while ramping the QoS-slack penalty when needed. Since (P') is non-convex, the resulting BCD procedure is embedded in a multi-start framework: several initial AIRS placements are generated around user-cluster centroids and projected onto the service region, and the best feasible solution is retained. If no initialization yields a feasible point, the best penalized solution is kept as a fallback. The complete procedure is summarized in Algorithm 1.

Complexity. The dominant cost is the phase-shift semidefinite program (SDP). Per BCD iteration, the phase block solves U SDPs of size $(M+1) \times (M+1)$ at the worst-case interior-point cost $\mathcal{O}(T_{\Theta} U (M+1)^{4.5})$ of semidefinite relaxation [17], with T_{Θ} inner penalty iterations; the placement block solves a convex trust-region subproblem at $\mathcal{O}(T_{\text{loc}}(KU+I)^3)$ over T_{loc} trust-region iterations; and the resource block replaces an explicit mixed-integer solve by T_{dual} dual updates of cost

Algorithm 1 Multi-Start BCD for Joint Phase, Placement, Reassociation, and Resource Allocation

```

1: Input:  $N_{\text{st}}, N_{\text{BCD}}, (\Lambda_{\min}, \Lambda_{\max}), (\zeta_{v,\min}, \zeta_{v,\max}),$  tolerance  $\epsilon$ .
2: for  $s = 1$  to  $N_{\text{st}}$  do
3:   Initialize  $\mathbf{Q}_s^{(0)}$ , association map  $\pi^{(0)}$ , and phase vector  $\mathbf{\Theta}^{(0)}$ ; set  $\Lambda \leftarrow \Lambda_{\min}$ .
4:   repeat
5:     Phase: solve (10) with ramped  $\zeta_v$  and recover  $\mathbf{\Theta}^{(n+1)}$ .
6:     Placement: solve the trust-region SCA subproblem in Sec. III-B; accept the candidate only if the true sum rate is non-decreasing.
7:     Reassociation: update  $\pi^{(n+1)}$  using the fixed-margin (5%) score rule in Sec. III-B.
8:     Resource: update  $(\mathbf{S}, \mathbf{P})$  using the dual-decomposition method in Sec. III-C.
9:     if  $\sum_k \zeta_k > 0$  then
10:        $\Lambda \leftarrow \min(2\Lambda, \Lambda_{\max})$ .
11:     until  $|R^{(n)} - R^{(n-1)}| < \epsilon$  or  $n = N_{\text{BCD}}$ 
12:     Record the final sum rate and feasibility status.
13: Output: the feasible start with the highest sum rate; if no start is feasible, output the best penalized solution.
```

$\mathcal{O}(IK \log(1/\epsilon_v))$ each, plus a one-off LP warm start. The overall complexity is $N_{\text{st}} T_{\text{BCD}}$ times the sum of these per-block dominant-order estimates.

IV. NUMERICAL RESULTS

We evaluate Algorithm 1 on an AIRS-assisted OFDMA downlink over a semi-annular service region with inner and outer radii of 500m and 1000m. Users are placed i.i.d. uniformly in this region, and the BS is at (0, 0, 30) m; the remaining parameters are listed in Table I. The design curves are averaged over 30 random trials, each trial corresponds to an independent random drop of the K user locations in the service region, while the instantaneous-CSI evaluations use 50 Monte Carlo (MC) draws per trial.

We compare four schemes: **Joint**, executing all four blocks; **Phase+RA**, which skips the placement update but retains the multi-start AIRS position sampling and optimizes phase, subcarrier assignment, and power; **Loc+RA**, which skips the phase-shift update ($\vartheta_{u,m} = 0$, i.e., $\mathbf{\Theta}_u = \mathbf{I}$) and optimizes placement and resources; and **No-AIRS**, which uses only the direct BS–user link. In Figs. 3–5, solid lines report the statistical-CSI (S-CSI) deterministic design rate, while markers show the MC mean rate of the frozen S-CSI design under instantaneous CSI. A design-line segment is solid when at least 95% of the users, aggregated over all design trials, satisfy the QoS target, and dashed otherwise. For the Joint scheme, we also report two instantaneous-CSI references to quantify the gap between the statistical-CSI design and realization-wise adaptation. The two-timescale (TS) reference freezes the S-CSI optimized AIRS locations, phase shifts, and user–AIRS association, and re-optimizes only the OFDMA resource block for each instantaneous channel realization, in the spirit of two-timescale IRS designs that fix passive beamforming on statistical CSI while adapting the short-term variables on instantaneous CSI [18]. The full-EXACT reference re-runs all four blocks for each realization and, since it solves the complete BCD problem per realization, serves only as an offline upper-reference benchmark.

TABLE I
SIMULATION PARAMETERS.

Param.	Value	Param.	Value
K	10 (8–20)	I	64
U	3 (1–4)	$P_{\max}$	40 dBm
M	36 (16–64)	N_0	−169 dBm/Hz
H_0	120 m	Δf	15 kHz
K_1, K_2	5, 10	f_c	3 GHz
α, α'	4, 2	$R_{\min}$	1 bit/s/Hz
d_e	$\lambda/2$	N_{st}	10
N_{BCD}	20	ϵ	0.05

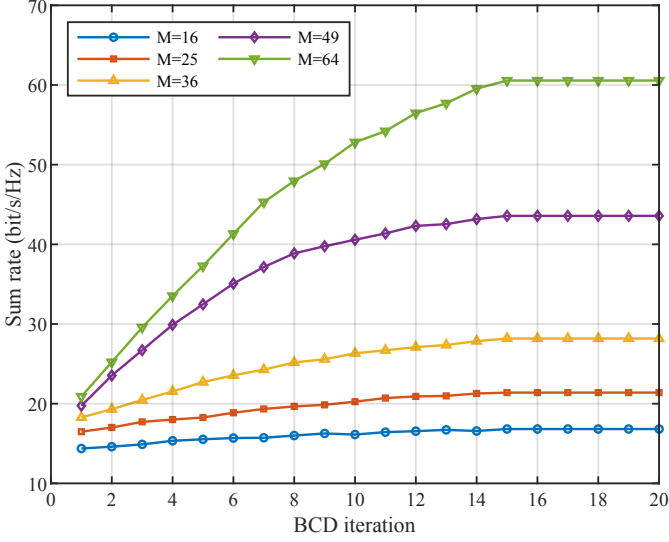


Fig. 2. Joint sum rate vs. outer BCD iteration for several M .

Fig. 2 plots the Joint sum rate versus the outer BCD iteration index. The observed objective is non-decreasing and stabilizes within the BCD iterations for all tested values of M .

Increasing the number of reflecting elements improves the performance of the schemes that optimize the AIRS phase shifts, as shown in Fig. 3 for $K = 10$ and $U = 3$. Among the S-CSI-based designs, the ordering $\text{Joint} > \text{Phase+RA} > \text{Loc+RA} > \text{No-AIRS}$ holds for all tested values of M . The Joint–Phase+RA gap is smaller than the Phase+RA–Loc+RA gap, indicating that phase-shift optimization provides a larger gain than horizontal AIRS repositioning in this setting. Joint and Phase+RA remain QoS-feasible and are plotted with solid curves, whereas Loc+RA and No-AIRS are QoS-infeasible for all tested values of M .

The effect of user loading is summarized in Fig. 4. As K increases, the fixed power and bandwidth must be shared among more QoS-constrained users, and the sum rate decreases for all AIRS-equipped schemes. Joint and Phase+RA remain QoS-feasible for $K \leq 12$ (solid segments) and become infeasible for larger K (dashed segments), as the limited power and bandwidth can no longer support every user. Loc+RA degrades earlier, while No-AIRS violates QoS for at least one user in $\geq 86.7\%$ of trials at every K .

The benefit of deploying multiple AIRSs is shown in Fig. 5.

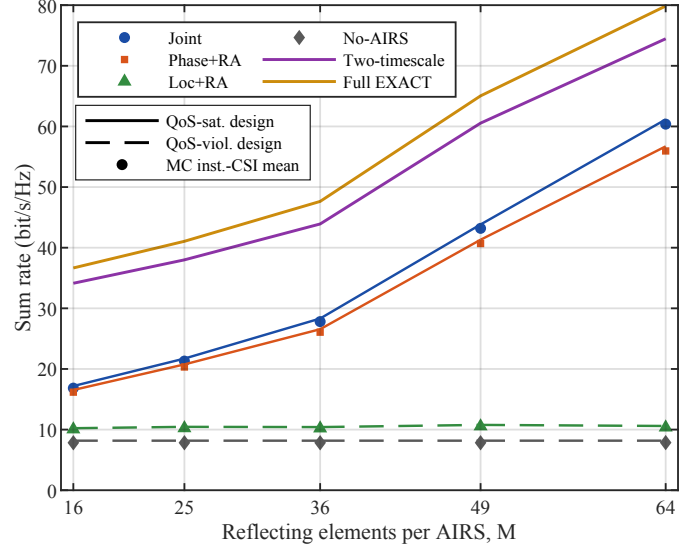


Fig. 3. Sum rate vs. number of passive reflecting elements M ($K = 10$, $U = 3$).

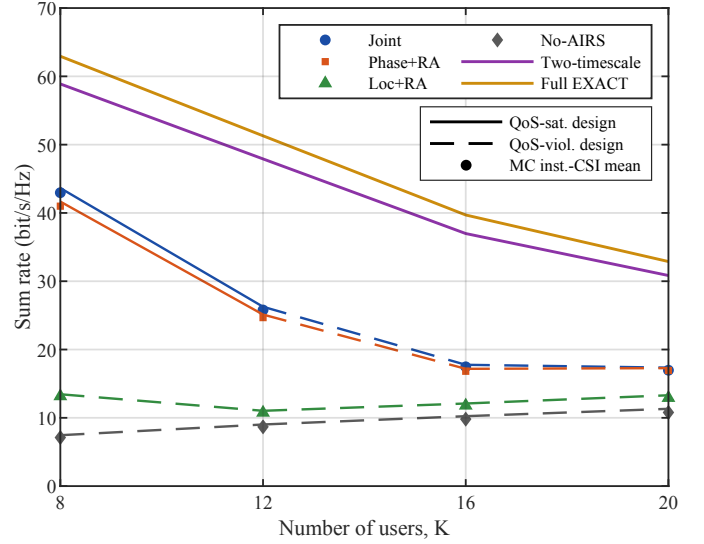


Fig. 4. Sum rate vs. number of users K ($M=36$, $U=3$).

The Joint sum rate increases from about 22.5 bit/s/Hz at $U = 1$ to 35.7 bit/s/Hz at $U = 4$, with the largest gain observed when the fourth AIRS is added. Joint and Phase+RA satisfy the aggregate user-level QoS criterion at every U , whereas Loc+RA and No-AIRS remain below the 95% threshold throughout.

Fig. 6 plots the empirical CDF of the per-user instantaneous rate near the QoS target. The AIRS-assisted schemes shift the distribution to the right of No-AIRS, reducing the per-user outage from approximately 69% for No-AIRS to 47–48% for Joint and Phase+RA (and about 56% for Loc+RA). The residual outage occurs because the S-CSI design satisfies QoS in expectation, not necessarily for every instantaneous realization. Re-optimizing only the resource block in the TS scheme removes the observed instantaneous-QoS violations

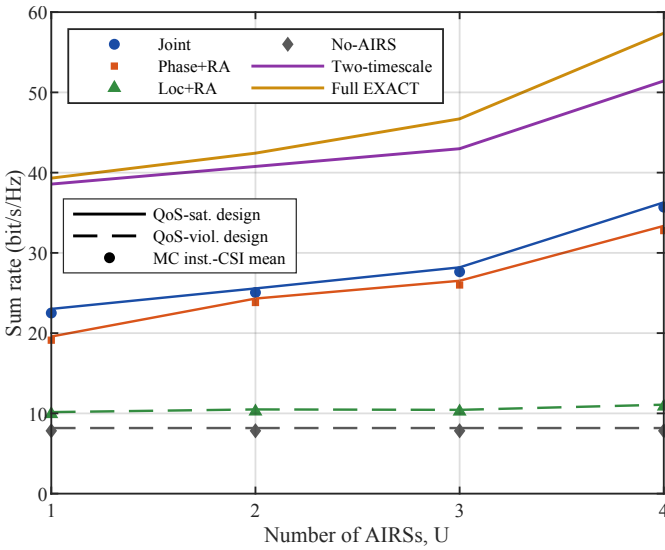


Fig. 5. Sum rate vs. number of AIRS units U ($K=10$, $M=36$).

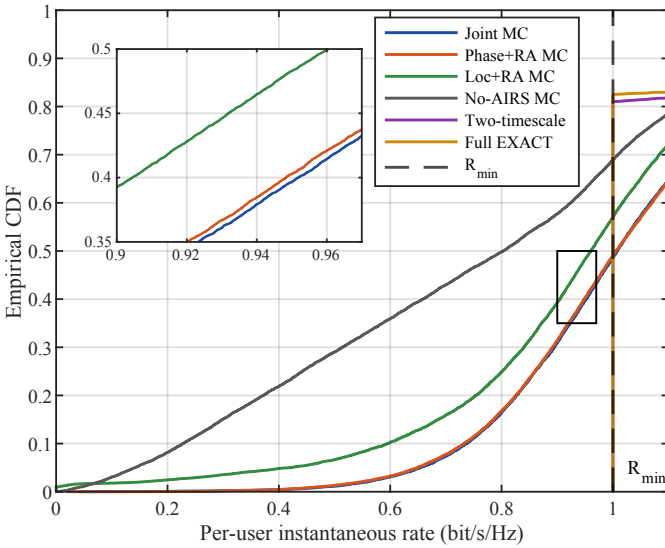


Fig. 6. Empirical CDF of per-user instantaneous rates near the QoS target ($M=36$, $K=10$, $U=3$).

while keeping the phase and placement variables fixed on the slow timescale.

V. CONCLUSION

We studied a multi-AIRS-assisted OFDMA downlink under per-user QoS, jointly optimizing placement, phase shifts, binary subcarrier scheduling, and power via a multi-start BCD combining SDR phase design, trust-region SCA placement, Lagrangian dual resource allocation, and strict-margin re-association. Across varying numbers of reflecting elements, users and AIRSs, the joint design consistently outperforms the Phase+RA, Loc+RA, and no-AIRS baselines, while instantaneous-CSI evaluations confirm the robustness of the statistical-CSI-based design in sum rate. Since QoS is met only in expectation, re-running the closed-form resource block

on instantaneous CSI in a two-timescale regime restores per-realization QoS while keeping the heavier phase and placement variables on the slow timescale. Future work will address imperfect CSI, learning-based phase adaptation, and continuous user-AIRS association.

REFERENCES

- [1] Q. Wu and R. Zhang, "Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network," *IEEE Commun. Mag.*, vol. 58, no. 1, pp. 106–112, Jan. 2020.
- [2] Q. Wu, S. Zhang, B. Zheng, C. You, and R. Zhang, "Intelligent reflecting surface-aided wireless communications: A tutorial," *IEEE Trans. Commun.*, vol. 69, no. 5, pp. 3313–3351, May 2021.
- [3] M. Alzenad, A. El-Keyi, F. Lagum, and H. Yanikomeroglu, "3-D placement of an unmanned aerial vehicle base station (UAV-BS) for energy-efficient maximal coverage," *IEEE Wireless Commun. Lett.*, vol. 6, no. 4, pp. 434–437, Aug. 2017.
- [4] X. Pang, M. Sheng, N. Zhao, J. Tang, D. Niyato, and K.-K. Wong, "When UAV meets IRS: Expanding air-ground networks via passive reflection," *IEEE Wireless Commun.*, vol. 28, no. 5, pp. 164–170, Oct. 2021.
- [5] Z. Wei, Y. Cai, Z. Sun, D. W. K. Ng, J. Yuan, M. Zhou, and L. Sun, "Sum-rate maximization for IRS-assisted UAV OFDMA communication systems," in *Proc. IEEE Global Commun. Conf.*, Dec. 2020, pp. 1–7.
- [6] S. Jangsher, M. Al-Jarrah, A. Al-Dweik, E. Alsusa, and P.-Y. Kong, "Energy constrained sum-rate maximization in IRS-assisted UAV networks with imperfect channel information," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 59, no. 3, pp. 2898–2908, Jun. 2023.
- [7] Y. Li, H. Zhang, K. Long, and A. Nallanathan, "Exploring sum rate maximization in UAV-based multi-IRS networks: IRS association, UAV altitude, and phase shift design," *IEEE Trans. Commun.*, vol. 70, no. 11, pp. 7764–7774, Nov. 2022.
- [8] X. Mu, Y. Liu, L. Guo, J. Lin, and H. V. Poor, "Intelligent reflecting surface enhanced multi-UAV NOMA networks," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 10, pp. 3051–3066, Oct. 2021.
- [9] G. Iacovelli, A. Coluccia, and L. A. Grieco, "Multi-UAV IRS-assisted communications: Multinode channel modeling and fair sum-rate optimization via deep reinforcement learning," *IEEE Internet Things J.*, vol. 11, no. 3, pp. 4470–4482, Feb. 2024.
- [10] H. Nguyen-Kha, H. V. Nguyen, M. T. P. Le, and O.-S. Shin, "Joint UAV placement and IRS phase shift optimization in downlink networks," *IEEE Access*, vol. 10, pp. 111 221–111 231, Oct. 2022.
- [11] M.-H. T. Nguyen, E. Garcia-Palacios, T. Do-Duy, O. A. Dobre, and T. Q. Duong, "UAV-aided aerial reconfigurable intelligent surface communications with massive MIMO system," *IEEE Trans. Cogn. Commun. Netw.*, vol. 8, no. 4, pp. 1828–1838, Dec. 2022.
- [12] Y. Qian, C. Yang, Z. Mei, X. Zhou, L. Shi, and J. Li, "On joint optimization of trajectory and phase shift for IRS-UAV assisted covert communication systems," *IEEE Trans. Veh. Technol.*, vol. 72, no. 10, pp. 12 873–12 883, Oct. 2023.
- [13] C. Qi, H. Chen, Y. Deng, and A. Nallanathan, "Energy efficient multicast precoding for multiuser multibeam satellite communications," *IEEE Wireless Commun. Lett.*, vol. 9, no. 4, pp. 567–570, Apr. 2020.
- [14] O. Mehanna, K. Huang, B. Gopalakrishnan, A. Konar, and N. D. Sidiropoulos, "Feasible point pursuit and successive approximation of non-convex QCQPs," *IEEE Signal Process. Lett.*, vol. 22, no. 7, pp. 804–808, Jul. 2015.
- [15] M. Grant and S. Boyd, "CVX: MATLAB software for disciplined convex programming, version 2.2," [Online]. Available: <http://cvxr.com/cvx/>, Jan. 2020.
- [16] W. Yu and R. Lui, "Dual methods for nonconvex spectrum optimization of multicarrier systems," *IEEE Trans. Commun.*, vol. 54, no. 7, pp. 1310–1322, Jul. 2006.
- [17] Z.-Q. Luo, W.-K. Ma, A. M.-C. So, Y. Ye, and S. Zhang, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 20–34, May 2010.
- [18] M.-M. Zhao, A. Liu, Y. Wan, and R. Zhang, "Two-timescale beamforming optimization for intelligent reflecting surface aided multiuser communication with QoS constraints," *IEEE Trans. Wireless Commun.*, vol. 20, no. 9, pp. 6179–6194, Sep. 2021.