Toward Energy Efficient Multiuser IRS-Assisted URLLC Systems: A Novel Rank Relaxation Method

Jalal Jalali[†], Filip Lemic[♯], Hina Tabassum[⋆], Rafael Berkvens[†], and Jeroen Famaey[†]

[†]IDLab, University of Antwerp - imec, Sint-Pietersvliet 7, 2000 Antwerp, Belgium

[♯]AI-Driven Systems Lab, i2Cat Foundation, Barcelona, Spain

[⋆]Department of Electrical Engineering and Computer Science, York University, Toronto, ON, Canada

Email: jalal.jalali@uantwerpen.be

Abstract—This paper proposes an energy efficient resource allocation design algorithm for an intelligent reflecting surface (IRS)-assisted downlink ultra-reliable low-latency communication (URLLC) network. This setup features a multi-antenna base station (BS) transmitting data traffic to a group of URLLC users with short packet lengths. We maximize the total network's energy efficiency (EE) through the optimization of active beamformers at the BS and passive beamformers (a.k.a, phase shifts) at the IRS. The main non-convex problem is divided into two sub-problems. An alternating optimization (AO) approach is then used to solve the problem. Through the use of the successive convex approximation (SCA) with a novel iterative rank relaxation method, we construct a concave-convex objective function for each sub-problem. The first sub-problem is a fractional program that is solved using the Dinkelbach method and a penalty-based approach. The second sub-problem is then solved based on semi-definite programming (SDP) and the penalty-based approach. The iterative solution gradually approaches the rank-one for both the active beamforming and unit modulus IRS phase-shift sub-problems. Our results demonstrate the efficacy of the proposed solution compared to existing benchmarks.

Index Terms—Alternating optimization (AO), intelligent reflecting surface (IRS), semi-definite programming (SDP), ultra-reliable low-latency communication (URLLC).

I. INTRODUCTION

Intelligent reflecting surface (IRS) has recently emerged as a cost-effective and energy efficient technology to reconfigure wireless environments through passive reflections. Unlike conventional relays, IRSs do not require signal decoding or amplification [1]. The integration of IRSs into the communication network offers enhanced reliability, reduced packet retransmissions, and latency. Consequently, IRSs can efficiently realize ultra-reliable low-latency communications (URLLC) with finite block length regime [2]. Unlike infinite blocklength regime, the achievable rate in the finite blocklength regime is a complex function of the signal-to-interference-plus-noise ratio (SINR), blocklength, and the probability of decoding errors [3].

Recently, a handful of research studies investigated resource allocation in IRS-assisted URLLC networks. For instance, Hashemi *et al.* presented the performance analysis of the average achievable data-rate and error probability over an IRS-aided URLLC transmission with and without phase noise [4]. Considering non-linear energy harvesting, the end-to-end performance of the IRS-assisted wireless system was analyzed in [5] for industrial URLLC applications, and the approximate closed-form expression of the block error rate was derived. Xie *et al.* studied an IRS-assisted downlink multiuser URLLC system and jointly optimized the user grouping and the blocklength allocation at the base station (BS), as well as the reflective beamforming at the IRS for latency minimization [6]. Li *et al.* investigated maximizing uplink data-rate in a URLLC system by jointly optimizing user transmit power, BS receive beamformer, and IRS reflection phase shifts under given error probability and blocklength conditions [7]. None of the aforementioned research works considered an *energy efficient design of IRS-assisted URLLC* networks.

In this paper, we design an energy efficient beamforming and phase-shift optimization algorithm for downlink IRS-assisted URLLC services, wherein a multi-antenna BS serves multiple single-antenna URLLC receivers with the support of an IRS. To this end, we formulate an energy efficiency maximization problem, subject to the minimum required data-rate (in the finite blocklength regime) for each URLLC user and unit-modulus constraints at the IRS. We decompose the problem into two sub-problems and apply an alternating optimization (AO) approach to optimize both active and passive beamformers at the BS and IRS, respectively. Using an innovative iterative rank relaxation and successive convex approximation (SCA) method, we construct a concave-convex objective function for each sub-problem. Then, the first sub-problem (which is also a fractional program) is solved by applying the Dinkelbach method and a penalty-based approach; whereas, the second sub-problem is solved by applying the penalty-based approach and semi-definite programming (SDP). The iterative solution gradually approaches the rank-one for both the active beamforming and unit modulus IRS phase-shift sub-problems. Our simulation results indicate the benefits of the IRS in enhancing the energy efficiency of URLLC systems.

Notations: In this work, matrices and vectors are denoted by boldface capital letters, like $\mathbf{A}$, and lower-case letters, such as $\mathbf{a}$, respectively. If we consider a square matrix $\mathbf{A}$, $\mathbf{A}^T$, $\mathbf{A}^H$, $\text{rank}(\mathbf{A})$, $\text{Tr}(\mathbf{A})$, and $||\mathbf{A}||_*$ denote its transpose, Hermitian
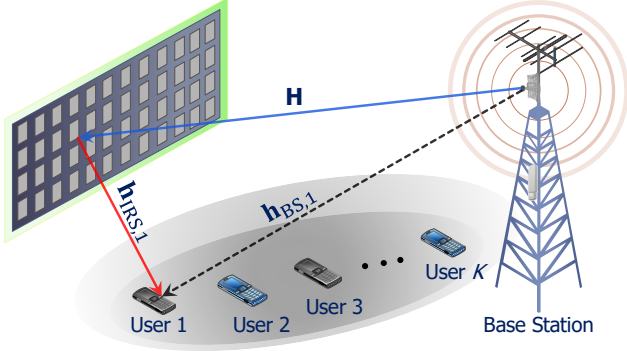
Fig. 1. Illustration of a multi-user IRS-assisted downlink network comprising one BS and K URLLC users employing finite block length transmissions. $\mathbf{H}$ is the channel matrix between the BS and the IRS, $\mathbf{h}_{\mathrm{IRS},1}$ denotes the channel response vector from the IRS to user 1, and $\mathbf{h}_{\mathrm{BS},1}$ indicates the channel between the BS and user 1.

conjugate transpose, rank, trace, and norm, respectively. $\mathbf{I}_N$ represents the $N \times N$ identity matrix. The diagonalization operation is represented by $\mathrm{diag}(\cdot)$. $\Re\{\cdot\}$ is used to denote the real part of a complex function. We express the absolute value of a complex scalar and the Euclidean norm of a complex vector by $|\cdot|$. The distribution of a circularly symmetric complex Gaussian (CSCG) random vector with mean $\boldsymbol{\mu}$ and covariance matrix $\mathbf{C}$ is denoted as $\mathcal{CN}(\boldsymbol{\mu}, \mathbf{C})$. $Q^{-1}(\cdot)$ represents the inverse of the Gaussian Q-function. $\mathbb{C}^{M \times N}$ denotes an $M \times N$ dimensional complex matrix, and $\nabla_{\mathbf{x}}$ expresses the gradient vector with respect to $\mathbf{x}$. Finally, we express $N \times N$ positive semi-definite matrices as $\mathbf{A} \in \mathbb{S}_+^N$ and read $\mathbf{A} \succeq \mathbf{0}$.

II. SYSTEM MODEL AND PROBLEM FORMULATION

We consider a downlink multiuser URLLC-enabled IRS-assisted system, represented in Fig. I, which comprises an IRS with N elements, a BS equipped with M antennas, and K users, each with a single antenna. The group of IRS elements, BS antennas, and users are denoted by the sets $\mathcal{N} = \{1, ..., N\}$, $\mathcal{M} = \{1, ..., M\}$, and $\mathcal{K} = \{1, ..., K\}$, respectively. Additionally, we suppose that B_k bits of information are assigned to user k. In this setup, the BS transforms these information bits into a block code in time slot l, characterized by a length of m_d symbols. We decode the k-th user block code as $x_{k,l}$, with l being part of the set $\mathcal{L} = \{1, 2, ..., m_d\}$.

The transmission signal emanating from the BS can be expressed as $\mathbf{s}[l] = \sum_{k \in \mathcal{K}} \boldsymbol{\omega}_{k,l} x_{k,l}$, where $\boldsymbol{\omega}_{k,l} \in \mathbb{C}^{M \times 1}$ represents the beamforming vector for user k. The channel links exhibit time invariance (i.e., slow fading). Additionally, we assume that the BS possesses full access to the channel state information (CSI) and URLLC users' delay requirements [8]–[10]. The baseband equivalent channel responses for BS-to-IRS, IRS-to-user k, and BS-to-user k are denoted as $\mathbf{H} \in \mathbb{C}^{N \times M}$, $\mathbf{h}_{\mathrm{IRS},k} \in \mathbb{C}^{N \times 1}$, and $\mathbf{h}_{\mathrm{BS},k} \in \mathbb{C}^{M \times 1}$, respectively. Also, we define the matrix of reflection coefficients at the IRS as $\boldsymbol{\Psi}_l = \mathrm{diag}(\alpha_{1,l} e^{j\phi_{1,l}}, \alpha_{2,l} e^{j\phi_{2,l}}, ..., \alpha_{N,l} e^{j\phi_{N,l}})$, where $\alpha_{n,l} \in [0, 1]$ and $\phi_{n,l} \in (0, 2\pi]$, $\forall n \in \mathcal{N}, \forall l \in \mathcal{L}$ are the reflection amplitude and phase shift of the n-th reflection coefficient at the IRS during time slot l, respectively.

By defining the combined channel link seen by the k-th URLLC user as $\mathbf{h}_k^H \triangleq \mathbf{h}_{\mathrm{IRS},k}^H \boldsymbol{\Psi}_l \mathbf{H} + \mathbf{h}_{\mathrm{BS},k}^H$, $\forall k \in \mathcal{K}, \forall l \in \mathcal{L}$, the SINR of that user can be expressed as:

$$\Gamma_{k,l} = \frac{\left|\mathbf{h}_k^H \boldsymbol{\omega}_{k,l}\right|^2}{\sum_{i \neq k, i \in \mathcal{K}} \left|\mathbf{h}_k^H \boldsymbol{\omega}_{i,l}\right|^2 + \sigma_k^2}, \forall k \in \mathcal{K}, \forall l \in \mathcal{L}, \quad (1)$$

where σ_k^2 represents the noise variance received at user k. In the context of URLLC systems, an accurate estimate of the achievable data-rate for each user can be delineated as follows [11]:

$$\mathcal{R}_k(\boldsymbol{\omega}_{k,l}, \boldsymbol{\Psi}_l) = \mathcal{U}_k(\boldsymbol{\omega}_{k,l}, \boldsymbol{\Psi}_l) - \mathcal{V}_k(\boldsymbol{\omega}_{k,l}, \boldsymbol{\Psi}_l), \ \forall k \in \mathcal{K}, \quad (2)$$

where

$$\mathcal{U}_k(\boldsymbol{\omega}_{k,l}, \boldsymbol{\Psi}_l) = \sum_{l \in \mathcal{L}} \log_2(1 + \Gamma_{k,l}), \qquad \forall k \in \mathcal{K}, \quad (3)$$

$$\mathcal{V}_k(\boldsymbol{\omega}_{k,l}, \boldsymbol{\Psi}_l) = \sum_{l \in \mathcal{L}} \beta_k \sqrt{\Delta_{k,l}}, \qquad \forall k \in \mathcal{K}, \quad (4)$$

$$\beta_k = \frac{Q^{-1}(\epsilon_k)}{\sqrt{m_d}}, \qquad \forall k \in \mathcal{K}. \quad (5)$$

Here, ϵ_k refers to the decoding error probability, m_d, as defined earlier, is representative of the blocklength, and $\Delta_{k,l}$ the channel dispersion, computed as:

$$\Delta_{k,l} = (\log_2 e)^2 \left(1 - \frac{1}{(1 + \Gamma_{k,l})^2}\right), \forall k \in \mathcal{K}, \forall l \in \mathcal{L}. \quad (6)$$

To ensure users' quality of service (QoS) regarding the received number of bits, the reliability, and the latency, a minimum data-rate denoted by $R_{\mathrm{min},k}$ should be satisfied for each user as follows:

$$\mathcal{R}_k(\boldsymbol{\omega}_{k,l}, \boldsymbol{\Psi}_l) \geq R_{\mathrm{min},k}, \ \forall k \in \mathcal{K}. \quad (7)$$

Next, we describe the network's energy efficiency (EE) as the ratio of the total system data-rate over the associated network power consumption in [bits/Joule/Hz]:

$$\eta_{\mathrm{eff}}(\boldsymbol{\omega}_{k,l}, \boldsymbol{\Psi}_l) = \frac{\sum_{k \in \mathcal{K}} \mathcal{R}_k(\boldsymbol{\omega}_{k,l}, \boldsymbol{\Psi}_l)}{\sum_{k \in \mathcal{K}} \sum_{l \in \mathcal{L}} \|\boldsymbol{\omega}_{k,l}\|^2 + P_{\mathrm{IRS}} + N P_d + P_c^{\mathrm{BS}}}, \quad (8)$$

where P_{IRS} represents the static power consumption necessary for sustaining the basic circuit operations of the IRS, while P_d is the dynamic power expended per reflecting component, and P_c^{BS} denotes the circuit power at the BS.

Our objective is to maximize the total EE by optimizing the active beamformers at the BS and the phase shifts at the IRS. Consequently, the problem can be articulated as follows:

$$\mathrm{P}_1 : \max_{\boldsymbol{\omega}_{k,l}, \boldsymbol{\Psi}_l} \quad \eta_{\mathrm{eff}}(\boldsymbol{\omega}_{k,l}, \boldsymbol{\Psi}_l)$$

$$s.t. : \mathcal{R}_k(\boldsymbol{\omega}_{k,l}, \boldsymbol{\Psi}_l) \geq R_{\mathrm{min},k}, \forall k \in \mathcal{K}, \quad (9a)$$

$$|\boldsymbol{\Psi}_{nn,l}| = 1, \qquad \forall n \in \mathcal{N}, \forall l \in \mathcal{L}, \quad (9b)$$

$$\sum_{l \in \mathcal{L}} \sum_{k \in \mathcal{K}} \|\boldsymbol{\omega}_{k,l}\|^2 \leq p_{\mathrm{max}}, \quad (9c)$$

$$\boldsymbol{\omega}_{k,l} = 0, \qquad \forall k \in \mathcal{K}, \forall l \geq \mathrm{T}_k, \quad (9d)$$

where (9a) is the minimum data-rate requirement ($R_{\mathrm{min},k}$) of the k-th URLLC user. Constraint (9b) seeks that the diagonal phase shift matrix contains N unit-modulus elements along its main diagonal. The constraint (9c) defines the limitation of the BS transmission power budget, with p_{max} signifying the maximum permissible BS transmission power. Constraint

(9d) is imposed to safeguard the real-time URLLC service functionality, ensuring that user k receives service within the first T_k time slots to satisfy its delay requirements.

Given the non-convex nature of the objective function and constraints in P_1, in the following, we propose an AO-based low-complexity solution.

III. ENERGY EFFICIENT DESIGN FOR IRS-ASSISTED URLLC NETWORK

In this section, we decompose P_1 into two sub-problems. The first sub-problem focuses on determining the optimal active beamformers at the BS, while the second sub-problem concentrates on optimizing the phase shifts at the IRS. We adopt the SCA technique to initially derive a lower bound for $\mathcal{R}_k(\boldsymbol{\omega}_{k,l}, \boldsymbol{\Psi}_l)$, which will subsequently be utilized when addressing each sub-problem.

Lemma 1: Let (j) denote the subscript associated with the feasible solution procured in the j-th iteration of the SCA algorithm. A lower bound approximation of $\mathcal{R}_k(\boldsymbol{\omega}_{k,l}, \boldsymbol{\Psi}_l)$ is given by:

$$\mathcal{R}_k(\boldsymbol{\omega}_{k,l}, \boldsymbol{\Psi}_l) \geq \tilde{\mathcal{R}}_k(\boldsymbol{\omega}_{k,l}, \boldsymbol{\Psi}_l) = \sum_{l \in \mathcal{L}} \tilde{\mathcal{R}}_{k,l}^{(j)}(\boldsymbol{\omega}_{k,l}, \boldsymbol{\Psi}_l), \quad (10)$$

$$\tilde{\mathcal{R}}_{k,l}^{(j)}(\boldsymbol{\omega}_{k,l}, \boldsymbol{\Psi}_l) \triangleq -\gamma_{k,l}\mathfrak{b}_{k,l} + 2\rho_{k,l}\Re\left\{\sum_{i \neq k, i \in \mathcal{K}} \mathfrak{a}_{i,k,l}^{(j)}\mathfrak{a}_{i,k,l}^H + \sigma_k^2\right\}$$

$$+ 2\Re\left\{\frac{\mathfrak{a}_{k,k,l}^{(j)}\mathfrak{a}_{k,k,l}^H}{\mathfrak{b}_{k,l}^{(j)}(\mathfrak{b}_{k,l}^{(j)} - |\mathfrak{a}_{k,k,l}^{(j)}|^2)}\right\} + \xi_{k,l},$$

$$\forall k \in \mathcal{K}, \forall l \in \mathcal{L}, \quad (11)$$

where

$$\mathfrak{a}_{k,i,l} = \mathbf{h}_k^H \boldsymbol{\omega}_{i,l}, \quad (12)$$

$$\mathfrak{b}_{k,l} = \sum_{i \in \mathcal{K}} \left|\mathbf{h}_k^H \boldsymbol{\omega}_{i,l}\right|^2 + \sigma_k^2, \quad (13)$$

$$\rho_{k,l} = \frac{\beta_k(\log_2 e)^2}{\mathfrak{b}_{k,l}^{(j)}\sqrt{(\Gamma_{k,l}^{(j)})^2 + 2\Gamma_{k,l}^{(j)}}} \quad (14)$$

$$\gamma_{k,l} = \frac{|\mathfrak{a}_{k,k,l}^{(j)}|^2}{\mathfrak{b}_{k,l}^{(j)}(\mathfrak{b}_{k,l}^{(j)} - |\mathfrak{a}_{k,k,l}^{(j)}|^2)} + \frac{\rho_{k,l}(\mathfrak{b}_{k,l}^{(j)} - |\mathfrak{a}_{k,k,l}^{(j)}|^2)}{\mathfrak{b}_{k,l}^{(j)}}, \quad (15)$$

$$\xi_{k,l} = \log_2(1 + \Gamma_{k,l}^{(j)}) - \frac{|\mathfrak{a}_{k,k,l}^{(j)}|^2}{\mathfrak{b}_{k,l}^{(j)}(\mathfrak{b}_{k,l}^{(j)} - |\mathfrak{a}_{k,k,l}^{(j)}|^2)}$$

$$- \frac{\beta_k(2\Delta_{k,l}^{(j)} + (\log_2 e)^2)}{2\sqrt{\Delta_{k,l}^{(j)}}} - \frac{\rho_{k,l}\mathfrak{b}_{k,l}^{(j)}}{2(1 + \Gamma_{k,l}^{(j)})}. \quad (16)$$

In the $(j+1)$-th iteration, the data-rate function $\tilde{\mathcal{R}}_{k,l}(\boldsymbol{\omega}_{k,l}, \boldsymbol{\Psi}_l)$ exhibits concavity with respect to each variable, reaching its boundary at the point $(\boldsymbol{\omega}_{k,l}^{(j)}, \boldsymbol{\Psi}_l^{(j)})$.

Proof 1: See Appendix A. ∎

The lower bound presented is more manageable compared to the original data-rate function $\tilde{\mathcal{R}}_k(\boldsymbol{\omega}_{k,l}, \boldsymbol{\Psi}_l)$ in (2). However, this bound still involves coupled optimization variables. To address this, we implement the AO approach. Specifically, we optimize $\boldsymbol{\omega}_{k,l}$ and $\boldsymbol{\Psi}_l$ by alternatively refining each variable while keeping the others constant.

A. Sub-problem $^{(1)}$: Optimizing $\boldsymbol{\omega}_{k,l}$ with given $\boldsymbol{\Psi}_l$

In this sub-problem, we presume that the passive reflecting elements at the IRS, i.e., $\boldsymbol{\Psi}_l$ are fixed, and we proceed with designing the active beamformers, $\boldsymbol{\omega}_{k,l}$, at the BS. By employing the principles of semi-definite programming (SDP), we derive the following: $\boldsymbol{\Omega}_{k,l} = \boldsymbol{\omega}_{k,l}\boldsymbol{\omega}_{k,l}^H \in \mathbb{S}_+^M$ and $\mathbf{H}_k = \mathbf{h}_k\mathbf{h}_k^H \in \mathbb{S}_+^M$, $\forall k \in \mathcal{K}$. Leveraging on the aforementioned SDP relaxation, we can rewrite the SINR in (1) as follows:

$$\Gamma_{k,l} = \frac{\text{Tr}(\mathbf{H}_k^H \boldsymbol{\Omega}_{k,l})}{\sum_{i \in \mathcal{K}, i \neq k} \text{Tr}(\mathbf{H}_k^H \boldsymbol{\Omega}_{i,l}) + \sigma_k^2}, \forall k \in \mathcal{K}, \forall l \in \mathcal{L}, \quad (17)$$

Consequently, the data-rate function defined in (2) can be reshaped as $\mathcal{R}_k(\boldsymbol{\Omega}_{k,l}) = \mathcal{U}_k(\boldsymbol{\Omega}_{k,l}) - \mathcal{V}_k(\boldsymbol{\Omega}_{k,l})$. With the aid of the SDP transformations, the EE optimization problem P_1 can be reformulated as follows:

$$P_2 : \max_{\boldsymbol{\Omega}_{k,l}} \quad \eta_{\text{eff}}(\boldsymbol{\Omega}_{k,l})$$

$$s.t. : \mathcal{R}_k(\boldsymbol{\Omega}_{k,l}) \geq R_{\min,k}, \forall k \in \mathcal{K}, \quad (18a)$$

$$\boldsymbol{\Omega}_{k,l} \succeq \mathbf{0}, \qquad \forall k \in \mathcal{K}, \forall l \in \mathcal{L}, \quad (18b)$$

$$\text{rank}(\boldsymbol{\Omega}_{k,l}) \leq 1, \qquad \forall k \in \mathcal{K}, \forall l \in \mathcal{L}, \quad (18c)$$

$$\sum_{l \in \mathcal{L}} \sum_{k \in \mathcal{K}} \text{Tr}(\boldsymbol{\Omega}_{k,l}) \leq p_{\max}, \quad (18d)$$

$$\text{Tr}(\boldsymbol{\Omega}_{k,l}) = 0, \qquad \forall k \in \mathcal{K}, \forall l \geq T_k. \quad (18e)$$

It is important to note that the constraint (18a) in P_2 does not exhibit concavity. To circumvent this non-concavity, we employ the outcome from *Lemma 1*. A proposed surrogate lower bound of $\mathcal{R}_k(\boldsymbol{\Omega}_{k,l})$ is introduced, where it is guaranteed that $\mathcal{R}_k(\boldsymbol{\Omega}_{k,l}) \geq \tilde{\mathcal{R}}_k(\boldsymbol{\Omega}_{k,l})$. Hence, we can re-express P_2 as:

$$P_3 : \max_{\boldsymbol{\Omega}_{k,l}} \quad \tilde{\eta}_{\text{eff}}(\boldsymbol{\Omega}_{k,l}) = \frac{\sum_{k \in \mathcal{K}} \tilde{\mathcal{R}}_k(\boldsymbol{\Omega}_{k,l})}{\mathcal{E}_{\text{tot}}(\boldsymbol{\Omega}_{k,l})}$$

$$s.t. : \tilde{\mathcal{R}}_k(\boldsymbol{\Omega}_{k,l}) \geq R_{\min,k}, \qquad \forall k \in \mathcal{K}, \quad (19a)$$

$$(18b) - (18e).$$

where $\mathcal{E}_{\text{tot}}(\boldsymbol{\Omega}_{k,l}) = \sum_{l \in \mathcal{L}} \sum_{k \in \mathcal{K}} \text{Tr}(\boldsymbol{\Omega}_{k,l}) + P_{\text{IRS}} + NP_d + P_c^{\text{BS}}$. The numerator of the objective function and constraint (19a) exhibit the form of convex-concave functions. Nevertheless, the overall problem remains non-convex due to the presence of the non-convex rank-one constraint (18c) and the fractional objective function. Our forthcoming focus firstly centers on resolving the non-convex rank-one constraint (18c), which emerges within the context of our optimization problem. Fortunately, an innovative method has been proposed in [12], specifically devised to address these types of rank constraints. We hereby direct our attention to the ensuing proposition.

Proposition 1: Consider a nonzero positive semi-definite beamforming matrix, denoted as $\boldsymbol{\Omega}_{k,l} \in \mathbb{S}_+^M$, where M represents the dimensionality of the matrix. We assert that $\boldsymbol{\Omega}_{k,l}$ is a rank one matrix if and only if the inequality $\varpi_{k,l}\mathbf{I}_{M-1} - \mho^T\boldsymbol{\Omega}_{k,l}\mho \succeq \mathbf{0}$ holds true, where $\varpi_{k,l} = 0$, $\mathbf{I}_{M-1}$ denotes an identity matrix with dimension $M-1$, and $\mho \in \mathbb{R}^{M \times (M-1)}$ corresponds to the eigenvectors corresponding to the $M-1$ smallest eigenvalues of $\boldsymbol{\Omega}_{k,l}$.

Proof 2: See Appendix B. ∎

According to *proposition 1*, we can replace the rank constraint (19a) with a positive semi-definite constraint:

$$\varpi_{k,l}\mathbf{I}_{M-1} - \mho^{(q)^T}\mathbf{\Omega}_{k,l}\mho^{(q)} \succeq \mathbf{0}, \forall k \in \mathcal{K}, \forall l \in \mathcal{L}. \quad (20)$$

Given the unavailability of $\mho^{(q)^T}$, we resort to the SCA method and employ the smallest eigenvectors of $\mathbf{\Omega}_{k,l}$. To achieve the goal of ultimately having $\varpi_{k,l} = 0$ while facilitating the attainment of an initial feasible point, we introduce a penalty term for $\varpi_{k,l}$ in the objective function. Thus, at SCA iteration (q), the following convex problem is addressed:

$$\mathrm{P}_4 : \max_{\mathbf{\Omega}_{k,l}, \varpi_{k,l}} \quad \tilde{\eta}_{\mathrm{eff}}(\mathbf{\Omega}_{k,l}) - \varsigma^{(q)}\sum_{l \in \mathcal{L}}\sum_{k \in \mathcal{K}}\varpi_{k,l}$$

$$s.t. : (18b), (18d), (18e), (19a), (20).$$

Note that the penalty factor's updates follow an interior iterative sub-algorithm with the formula $\varsigma^{(q)} = \min(\nu\varsigma^{(q-1)}, \vartheta_{\max})$, where ν is a positive coefficient factor, and $\vartheta_{\max}$ is the maximum allowable penalty factor.

The objective function in P_4 is a fractional function, rendering the optimization problem to be still non-convex. To tackle this issue, we adopt Dinkelbach method [13], with (d) as its iteration, to address the fractional form of the optimization problem. Denoting $\tilde{\eta}_{\mathrm{eff}}^*(\mathbf{\Omega}_{k,l})$ as the optimal energy efficiency point within the feasible solution set defined by its constraints, we then transform the problem into a non-fractional optimization problem for further resolution, i.e.,

$$\mathrm{P}_5 : \max_{\mathbf{\Omega}_{k,l}, \varpi_{k,l}} \sum_{k \in \mathcal{K}}\tilde{\mathcal{R}}_k(\mathbf{\Omega}_{k,l}^{(d)}) - \varrho^{(d)}\mathcal{E}_{\mathrm{tot}}(\mathbf{\Omega}_{k,l}^{(d)}) - \varsigma^{(d)}\sum_{l \in \mathcal{L}}\sum_{k \in \mathcal{K}}\varpi_{k,l}$$

$$s.t. : (18b), (18d), (18e), (19a), (20),$$

where $\varrho^{(d)} = \max_{\mathbf{\Omega}_{k,l}^{(d)}} \tilde{\eta}_{\mathrm{eff}}^{(d)}(\mathbf{\Omega}_{k,l}^{(d)})$, and $\varsigma^{(d)}$ is the new penalty factor update rules $\varsigma^{(d)} = \min(\nu\varsigma^{(d-1)}, \vartheta_{\max}) \times \mathcal{E}_{\mathrm{tot}}(\mathbf{\Omega}_{k,l}^{(d)})$. Finally, the convexity of the objective function in P_5 with respect to the active beamforming variables can be demonstrated through a formal proof based on the following proposition.

Proposition 2: The optimal energy efficiency, denoted as $\tilde{\eta}_{eff}^*(\mathbf{\Omega}_{k,l}^*)$, serves as a means to derive the resource allocation policy if and only if

$$\max_{\mathbf{\Omega}_{k,l}} \sum_{k \in \mathcal{K}}\tilde{\mathcal{R}}_k(\mathbf{\Omega}_{k,l}^{(d)}) - \varrho^{(d)}\mathcal{E}_{\mathrm{tot}}(\mathbf{\Omega}_{k,l}^{(d)} - \varsigma^{(q)}\sum_{l \in \mathcal{L}}\sum_{k \in \mathcal{K}}\varpi_{k,l} =$$

$$\sum_{k \in \mathcal{K}}\tilde{\mathcal{R}}_k(\mathbf{\Omega}_{k,l}^*) - \varrho^* \mathcal{E}_{\mathrm{tot}}(\mathbf{\Omega}_{k,l}^*) - \varsigma^{(q)}\sum_{l \in \mathcal{L}}\sum_{k \in \mathcal{K}}\varpi_{k,l} = 0,$$

$$(23)$$

for $\sum_{k \in \mathcal{K}}\tilde{\mathcal{R}}_k(\mathbf{\Omega}_{k,l}^*) \geq 0$ and $\mathcal{E}_{\mathrm{tot}}(\mathbf{\Omega}_{k,l}^*) \geq 0$, where $\mathbf{\Omega}_{k,l}^*$ provides the optimal solution to P_5.

Proof 3: See Appendix C. ∎

The optimization problem P_5 is now a convex problem and can be efficiently solved by standard convex optimization solvers such as CVX [14], [15]. Ultimately, we detail our proposed algorithm in **Algorithm 1**.

B. Sub-problem $^{(2)}$: Optimizing $\mathbf{\Psi}_l$ with given $\boldsymbol{\omega}_{k,l}$

In the second sub-problem, equipped with the optimal active beamforming matrices $\boldsymbol{\omega}_{k,l}$ from the preceding sub-problem, we progress with the optimization of the passive reflecting elements at the IRS, denoted by $\mathbf{\Psi}_l$. Given that the semi-definite matrix $\mathbf{\Omega}_{k,l}$ is provided, the optimization problem in P_1 shifts

---

Algorithm 1 Iterative SCA-based Resource Allocation Algorithm

Input: Set $d = 0$ for the Dinkelbach procedure, set the maximum number of iteration $D_{\max}$, initialize the beamformer matrix $\mathbf{\Omega}_{k,l} = \mathbf{\Omega}_{k,l}^0$, initialize ς^0, set $\vartheta_{\max} \gg 1$, $\nu > 1$, and set the tolerance $\varepsilon = 10^{-3}$.

1: **repeat**
2:     Calculate $\tilde{\mathcal{R}}_k(\boldsymbol{\omega}_{k,l}, \mathbf{\Psi}_l)$ for a given $\mathbf{\Psi}_l$.
3:     Calculate the rank-one relaxation constraint in (20).
4:     Solve P_5 for $\varrho^{(d-1)}$.
5:     **if** $\big| \sum_{k \in \mathcal{K}}\tilde{\mathcal{R}}_k(\mathbf{\Omega}_{k,l}^{(d)}) - \varrho^{(d-1)}\mathcal{E}_{\mathrm{tot}}(\mathbf{\Omega}_{k,l}^{(d)})$
                     $-\varsigma^{(d)}\sum_{l \in \mathcal{L}}\sum_{k \in \mathcal{K}}\varpi_{k,l}\big| \leq \varepsilon$
6:         **return** $\mathbf{\Omega}_{k,l} = \mathbf{\Omega}_{k,l}^{(d)}$, $\varrho^* = \varrho^{(d-1)}$.
7:     **else** $\varrho^{(d)} = \tilde{\eta}_{eff}^{(d)}(\mathbf{\Omega}_{k,l}^{(d)})$, **end if**.
8:     Update $\varsigma^{(d)} = \min(\nu\varsigma^{(d-1)}, \vartheta_{\max})$.
9:     $d \leftarrow d + 1$.
10:    **until** $d = D_{\max}$.
11: **return** $\mathbf{\Omega}_{k,l} = \mathbf{\Omega}_{k,l}^*$.

---

towards the maximization of the data-rate. The major hurdle in optimizing the phase shifts at the IRS arises due to the constraint (9b). Particularly, the constraint (9b) imposes a unit-modulus constraint, which presents a significant challenge in the quest to solve the problem. As such, we start by defining $\boldsymbol{e} = (e^{j\phi_1}, ..., e^{j\phi_N})^H \in \mathbb{C}^{N \times 1}$ and $\tilde{\boldsymbol{e}} = [\boldsymbol{e}^T \psi]^T \in \mathbb{C}^{(N+1) \times 1}$, where $\psi \in \mathbb{C}$ is a dummy variable with $|\psi| = 1$. To aid in crafting the solution, we further define $\mathbf{E} = \tilde{\boldsymbol{e}}\tilde{\boldsymbol{e}}^H \in \mathbb{C}^{(N+1) \times (N+1)}$. Consequently, we derive the following:

$$\big|(\mathbf{h}_{\mathrm{IRS,k}}^H\mathbf{\Psi}_l\mathbf{H} + \mathbf{h}_{\mathrm{BS,k}}^H)\mathbf{\Omega}_{k,l}\big|^2 \triangleq \mathrm{Tr}(\mathbf{E}\mathbf{Z}_k\mathbf{\Omega}_{k,l}\mathbf{Z}_k^H)$$

$$= \mathrm{Tr}(\mathbf{\Omega}_{k,l}\mathbf{Y}_k), \forall k \in \mathcal{K}, \quad (24)$$

where $\mathbf{Z}_k = [(\mathrm{diag}(\mathbf{h}_{\mathrm{IRS,k}}^H)\mathbf{H})^T \quad \mathbf{h}_{\mathrm{BS,k}}^*]^T$, $\mathbf{Y}_k = \mathbf{Z}_k^H\mathbf{E}\mathbf{Z}_k$. In a manner akin to P_3, we address the non-convex constraint (9a) and the objective function. To achieve this, we employ *Lemma 1* and subsequently rewrite the data-rate function as $\tilde{\mathcal{R}}_k(\mathbf{\Psi}_l) = \mathcal{U}_k(\mathbf{\Psi}_l) - \mathcal{V}_k(\mathbf{\Psi}_l)$. Now, we restate the optimization problem P_1 as follows:

$$\mathrm{P}_6 : \max_{\mathbf{E}, \mathbf{\Psi}_l} \sum_{k \in \mathcal{K}}\tilde{R}_k(\mathbf{\Psi}_l)$$

$$s.t. : \tilde{\mathcal{R}}_k(\mathbf{\Psi}_l) \geq R_{\min,k}, \forall k \in \mathcal{K}, \quad (25a)$$

$$\mathrm{diag}(\mathbf{\Psi}_l) = \mathbf{1}_{N+1}, \forall l \in \mathcal{L}, \quad (25b)$$

$$\mathrm{rank}(\mathbf{E}) \leq 1, \quad (25c)$$

$$\mathbf{\Psi}_l \succeq 0, \quad \forall l \in \mathcal{L}, \quad (25d)$$

$$\mathbf{E} \succeq \mathbf{0}. \quad (25e)$$

Following a similar approach as in P_6, and drawing on the insights from *proposition 1*, we can substitute the rank constraint (25c) with a positive semi-definite constraint:

$$\zeta_l\mathbf{I}_N - \Upsilon^{(g)^T}\mathbf{\Psi}_l\Upsilon^{(g)} \succeq \mathbf{0}, \forall l \in \mathcal{L}, \quad (26)$$

where $\Upsilon^{(g)}$ is an $N \times (N+1)$ matrix at the (g)-th SCA iteration with its columns corresponding to the smallest N eigenvectors of $\mathbf{\Psi}_l$. Furthermore, to sustain a rank of one for $\mathbf{\Psi}_l$, equation (26) must hold true, assuming $\zeta_l = 0$. Hence, by incorporating ζ_l as a penalty term into the objective function of P_6, we write

Algorithm 2 Proposed Iterative AO algorithm

Input: Set $s = 0$, maximum number of iteration $S_{\max}$, and initialize the beamformer matrix and phase shifts as $\boldsymbol{\Omega}_{k,l} = \boldsymbol{\Omega}_{k,l}^0$ and $\boldsymbol{\Psi}_l = \boldsymbol{\Psi}_l^0$, respectively.

1: **repeat**
2:   Solve problem P$_5$ for given $\boldsymbol{\Psi}_l^s$, and obtain the optimal solution $\boldsymbol{\Omega}_{k,l}^s$ based on **Algorithm 1**.
3:   Solve problem P$_7$ for given $\boldsymbol{\Omega}_{k,l}^s$, and obtain the optimal solution $\boldsymbol{\Psi}_l^s$.
4:   $s \leftarrow s + 1$.
5: **until** $s = S_{\max}$
6: **return** $\{\boldsymbol{\Omega}_{k,l}^*, \boldsymbol{\Psi}_l^*\} = \{\boldsymbol{\Omega}_{k,l}^s, \boldsymbol{\Psi}_l^s\}$.

---

the subsequent convex problem:

$$P_7 : \max_{\mathbf{E}, \boldsymbol{\Psi}_l} \sum_{k \in \mathcal{K}} \tilde{\mathcal{R}}_k(\boldsymbol{\Psi}_l) - \varkappa^{(g)} \sum_{l \in \mathcal{L}} \zeta_l$$

$$s.t. : (25a), (25b), (25d), (25e), (26),$$

where $\varkappa^{(g)}$ is a sequences of increasing weights. The optimization problem P$_7$ now can be efficiently solved just as P$_5$. The solution of these two sub-problems yields the suboptimal solution of P$_1$.

Proposition 3: The objective function of P$_1$ is ensured to be monotonically non-decreasing throughout the iterations of the proposed algorithm.

Proof 4: See Appendix D. ∎

Ultimately, the final iterative AO algorithm that solves sub-problems is presented in **Algorithm 2**.

IV. NUMERICAL RESULTS AND DISCUSSIONS

This section provides simulation results to investigate the efficiency of the proposed algorithm in IRS-supported downlink URLLC systems, utilizing finite blocklength codes. A rectangular region with dimensions of $(100, 100)$ meters is considered. The BS is positioned at $(0, 0)$ meters, the IRS is located at $(50, 0)$ meters, and all users are scattered randomly within the boundaries of this rectangle region. The path loss model is expressed as $35.3 + 37.6 \log_{10}(d_k)$ [dBm], where d_k refers to the distance in kilometers between the BS and the k-th URLLC user [9]. The convergence tolerance for the proposed AO algorithm, which is based on rank relaxation, SCA, and the Dinkelback method, is set at 10^{-2}. It is assumed that the thermal noise density stands at -174 [dBm/Hz]. In addition, the maximum probability of decoding error for URLLC user k, represented as ϵ_k, is set to be 10^{-7}. Furthermore, the simulation parameters are configured with $K = 4$, $M = 5$, and $R_{\min,k} = 1.6$ [bits/Sec/Hz] , in accordance with the methodology outlined in references [8] and [9].

Fig. 2 depicts the average EE in correlation with diverse maximum decoding error probability values, denoted as $\epsilon_{k,\max}$, keeping the block code length constant at 250 symbols, represented as $m_d = 250$. The plotted data brings to light an augmenting pattern in EE as the permissible decoding error broadens. This increment is attributable to the inverse proportionality of $Q^{-1}(\epsilon_k)$ and ϵ_k. The dominating contributor to this rise is a diminished rate loss $\mathcal{V}_k(\boldsymbol{\omega}_{k,l}, \boldsymbol{\Psi}_l)$ of (2) correlating
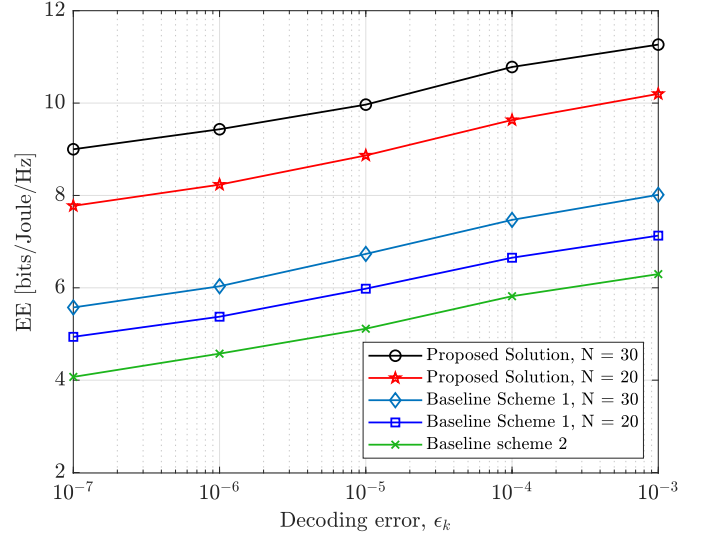

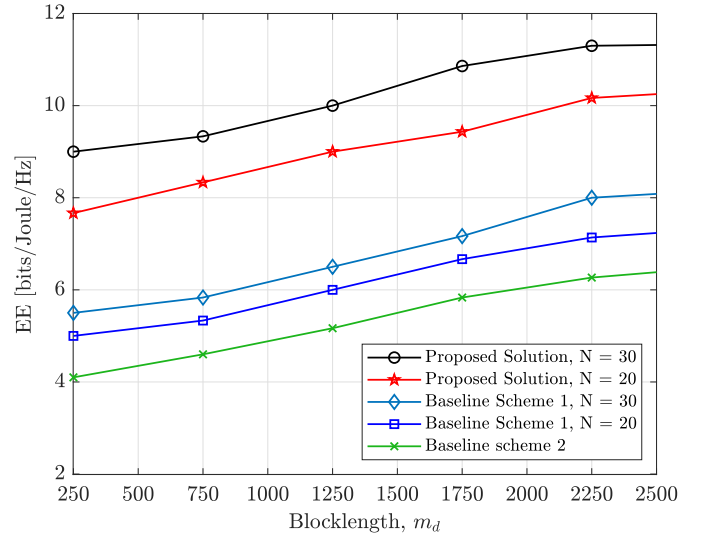Fig. 2. Impact of tolerable decoding error, ϵ_k, on the EE.

Fig. 3. Impact of the blocklength, m_d, on the EE.

with an enlargement in the maximum acceptable decoding error probability. This correlation ensures that the minimal data-rate obligation is met at reduced transmission power, thereby amplifying EE. The figure also signifies an upsurge in EE when the number of reflecting elements is increased.

Moreover, the figure also includes a comparative study with two basic models. *Baseline scheme 1* operates under a fixed beamforming policy at the IRS [9], whereas *Baseline scheme 2* considers a system without IRS. Our proposed methodology's performance, as demonstrated by the figure, surpasses that of baseline scheme 2 by 78.9%. This superior performance results from the incorporation of IRS and the harmonized optimization of active and passive beamforming matrices at the BS and IRS. This strategy displays an improved performance when compared to baseline scheme 1, 51.7% and 48.8% for $N = 30$ and $N = 20$, respectively, thereby underscoring the substantial benefits of integrating IRS into our system.

Fig. 3 illustrates the impact of the blocklength size, denoted as m_d, on the achievable EE at a constant $\epsilon_k = 10^{-7}$. With an increase in m_d, the EE initially experiences a moderate boost, followed by a slow elevation until it plateaus. In comparison to the scenario with an IRS, the scenario without it consistently exhibits lower EE, independent of the quantity of reflecting elements. Smaller N values necessitate higher transmit power to meet QoS standards, leading to a reduction in EE. This reduction arises from the amplified interference in the data-rate function due to multi-user interference in the SINR function, resulting in an increase in transmit power, a decrease in data-rate, and consequently negatively pruned EE.

V. Conclusion

This paper investigated the resource allocation for a downlink multiuser intelligent reflecting surface (IRS)-aided ultra-reliable low-latency communication (URLLC) system. In particular, the resource allocation design with active/passive beamforming was formulated to maximize the energy efficiency (EE) while considering URLLC Quality of Service (QoS) requirements for each user with adopting the short packet transmission. We proposed a novel iterative rank minimization and adopted the successive convex approximation (SCA) approach, as well as a penalty-based method to solve each sub-problem. Simulation results demonstrated the efficacy of the proposed method compared to a system without or with unoptimized IRSs.

Appendix A
Proof of the Lemma 1

According to the data-rate equation in (2), we can drive a lower bound on $\mathcal{U}_k(\boldsymbol{\omega}_{k,l}, \boldsymbol{\Psi}_l)$ and an upper bound on $\mathcal{V}_k(\boldsymbol{\omega}_{k,l}, \boldsymbol{\Psi}_l)$. In order to approximate the non-convex logistic function, we utilize the Majorization Minimization (MM) algorithm [16], [17]. This is achieved by constructing a surrogate function using a first-order Taylor approximation, which is expressed as follows:

$$\mathcal{U}_k(\mathbf{a}) \simeq \mathcal{U}_k(\mathbf{a}^{(j)}) + \nabla_{\mathbf{a}}\mathcal{U}_k(\mathbf{a}^{(j)}).(\mathbf{a} - \mathbf{a}^{(j)}) \triangleq \tilde{\mathcal{U}}_k(\mathbf{a}), \quad (28)$$

$$\mathcal{V}_k(\mathbf{b}) \simeq \mathcal{V}_k(\mathbf{b}^{(j)}) + \nabla_{\mathbf{b}}\mathcal{V}_k(\mathbf{b}^{(j)}).(\mathbf{b} - \mathbf{b}^{(j)}) \triangleq \tilde{\mathcal{V}}_k(\mathbf{b}), \quad (29)$$

where $\mathbf{a}, \mathbf{b} = \{\boldsymbol{\omega}_{k,l}, \boldsymbol{\Psi}_l\}$ and j is the iteration number. Moreover, $\mathbf{a}^{(j)}$ and $\mathbf{b}^{(j)}$ denote the solutions of the problem at $(j)^{th}$ iteration. Based on (12)-(16), we first rewrite the SINR formula as follows:

$$\Gamma_{k,l} = \frac{|\mathfrak{a}_{k,l}|^2}{\mathfrak{b}_{k,l} - |\mathfrak{a}_{k,l}|^2}, \forall k \in \mathcal{K}, \forall l \in \mathcal{L}. \quad (30)$$

Equivalently, we can rewrite the channel dispersion as:

$$\Delta_{k,l} = (\log_2 e)^2 \left(1 - \left(\frac{\mathfrak{b}_{k,l} - |\mathfrak{a}_{k,l}|^2}{\mathfrak{b}_{k,l}}\right)^2\right),$$
$$\forall k \in \mathcal{K}, \forall l \in \mathcal{L}. \quad (31)$$

Subsequently, we have:

$$\mathcal{U}_k(\boldsymbol{\omega}_{k,l}, \boldsymbol{\Psi}_l) = -\sum_{l \in \mathcal{L}} \log_2 \left(1 - \frac{|\mathfrak{a}_{k,l}|^2}{\mathfrak{b}_{k,l}}\right)$$
$$= \sum_{l \in \mathcal{L}} \mathcal{U}_{k,l}(\boldsymbol{\omega}_{k,l}, \boldsymbol{\Psi}_l), \qquad \forall k \in \mathcal{K}, \quad (32)$$

$$\mathcal{V}_k(\boldsymbol{\omega}_{k,l}, \boldsymbol{\Psi}_l) = \sum_{l \in \mathcal{L}} \beta_k \sqrt{\Delta_{k,l}}$$
$$= \sum_{l \in \mathcal{L}} \mathcal{V}_{k,l}(\boldsymbol{\omega}_{k,l}, \boldsymbol{\Psi}_l), \qquad \forall k \in \mathcal{K}. \quad (33)$$

Thus, employing the MM technique allows us to derive the following lower limit for $\mathcal{U}_{k,l}(\boldsymbol{\omega}_{k,l}, \boldsymbol{\Psi})$:

$$\mathcal{U}_{k,l}(\boldsymbol{\omega}_{k,l}, \boldsymbol{\Psi}) \geq - \frac{\left|\mathfrak{a}_{k,k,l}^{(j)}\right|^2}{\mathfrak{b}_{k,l}^{(j)}\left(\mathfrak{b}_{k,l}^{(j)} - \left|\mathfrak{a}_{k,k,l}^{(j)}\right|^2\right)} \mathfrak{b}_{k,l}$$
$$+ 2\Re \left\{\frac{\mathfrak{a}_{k,k,l}^{(j)}\mathfrak{a}_{k,k,l}^H}{\mathfrak{b}_{k,l}^{(j)} - \left|\mathfrak{a}_{k,k,l}^{(j)}\right|^2}\right\} + \xi_{k,l}^{\mathcal{U}},$$
$$\forall k \in \mathcal{K}, \forall l \in \mathcal{L}, \quad (34)$$

where

$$\xi_{k,l}^{\mathcal{U}} = \log_2(1 + \Gamma_{k,l}^{(j)}) - \frac{|\mathfrak{a}_{k,k,l}^{(j)}|^2}{\mathfrak{b}_{k,l}^{(j)}(\mathfrak{b}_{k,l}^{(j)} - |\mathfrak{a}_{k,k,l}^{(j)}|^2)}. \quad (35)$$

Similarly, we use the first-order Taylor approximation to derive the following upper limit for $\mathcal{V}_{k,l}(\boldsymbol{\omega}_{k,l}, \boldsymbol{\Psi})$:

$$\mathcal{V}_{k,l}(\boldsymbol{\omega}_{k,l}, \boldsymbol{\Psi}) \leq \beta_k \sqrt{\Delta_{k,l}^j} + \frac{(\log_2 e)^2 \beta_k}{2\sqrt{\Delta_{k,l}^j}} \left(1 - \frac{1}{(1 + \Gamma_{k,l})^2}\right),$$
$$\forall k \in \mathcal{K}, \forall l \in \mathcal{L}. \quad (36)$$

Now, by defining $\Xi_{k,l} \triangleq 1/(1 + \Gamma_{k,l})$ and according to [9], we can rewrite:

$$\frac{1}{(1 + \Gamma_{k,l})^2} = \Xi_{k,l}^2 \geq 2\Xi_{k,l}^{(j)}\Xi_{k,l} - (\Xi_{k,l}^{(j)})^2$$
$$= \frac{2}{\left(1 + \Gamma_{k,l}^{(j)}\right)(1 + \Gamma_{k,l})} - \frac{1}{\left(1 + \Gamma_{k,l}^{(j)}\right)^2},$$
$$\forall k \in \mathcal{K}, \forall l \in \mathcal{L}, \quad (37)$$

where the inequality holds due to convexity of $(\Xi_{k,l})^2$. The term $1/(1 + \Gamma_{k,l})$ in (37) is non-convex due to its fractional form. We rewrite this term as follows and then apply MM:

$$\frac{1}{1 + \Gamma_{k,l}} = \frac{\mathfrak{b}_{k,l} - |\mathfrak{a}_{k,k,l}|^2}{\mathfrak{b}_{k,l}} \geq - \frac{\left(\mathfrak{b}_{k,l}^{(j)} - |\mathfrak{a}_{k,k,l}^{(j)}|^2\right)\mathfrak{b}_{k,l}}{\left|\mathfrak{b}_{k,l}^{(j)}\right|^2}$$
$$+ 2\Re \left\{\frac{\sum_{i \neq k, i \in \mathcal{K}} \mathfrak{a}_{i,k,l}^{(j)}\mathfrak{a}_{i,k,l}^H + \sigma_k^2}{\mathfrak{b}_{k,l}^{(j)}}\right\}, \forall k \in \mathcal{K}, \forall l \in \mathcal{L}, \quad (38)$$

where, again, the inequality holds due to the convexity with respect to the $\mathfrak{a}_{i,k,l}$ and $\mathfrak{b}_{k,l}$. Thus, we derive an upper bound as:

$$\mathcal{V}_{k,l}(\boldsymbol{\omega}_{k,l}, \boldsymbol{\Psi}) \leq \frac{(\log_2 e)^2 \beta_k}{\sqrt{(\Gamma_{k,l}^{(j)})^2 + 2\Gamma_{k,l}^{(j)}}} \left(\frac{\left(\mathfrak{b}_{k,l}^{(j)} - |\mathfrak{a}_{k,k,l}^{(j)}|^2\right)\mathfrak{b}_{k,l}}{\left|\mathfrak{b}_{k,l}^{(j)}\right|^2}\right.$$
$$\left. - 2\Re \left\{\frac{\sum_{i \neq k, i \in \mathcal{K}} \mathfrak{a}_{i,k,l}^{(j)}\mathfrak{a}_{i,k,l}^H + \sigma_k^2}{\mathfrak{b}_{k,l}^{(j)}}\right\}\right) + \xi_{k,l}^{\mathcal{V}},$$
$$\forall k \in \mathcal{K}, \forall l \in \mathcal{L}, \quad (39)$$

where
$$\xi_{k,l}^{\mathcal{V}} = \frac{(2\Delta_{k,l}^{(j)} + (\log_2 e)^2)\beta_k}{2\sqrt{\Delta_{k,l}^{(j)}}}$$
$$+ \frac{(\log_2 e)^2\beta_k}{2(1+\Gamma_{k,l}^{(j)})\sqrt{(\Gamma_{k,l}^{(j)})^2 + 2\Gamma_{k,l}^{(j)}}}, \forall k \in \mathcal{K}, \forall l \in \mathcal{L}. \quad (40)$$

Ultimately, the lower bound of the data-rate function in (2) can be approximated as:
$$\tilde{\mathcal{R}}_k(\boldsymbol{\omega}_{k,l}, \boldsymbol{\Psi}_l) \triangleq \sum_{l \in \mathcal{L}} \left(\mathcal{U}_{k,l}(\boldsymbol{\omega}_{k,l}, \boldsymbol{\Psi}) - \mathcal{V}_{k,l}(\boldsymbol{\omega}_{k,l}, \boldsymbol{\Psi}) \right), \forall k \in \mathcal{K}. \quad (41)$$

This completes the proof. ∎

APPENDIX B
PROOF OF THE PROPOSITION 1

Assuming the nonnegative eigenvalues of $\boldsymbol{\Omega}_{k,l}$ are arranged in descending order as $[\kappa_M; \kappa_{M-1}; ...; \kappa_1]$, we can exploit the relationship between an eigenvector's Rayleigh quotient and its associated eigenvalue. Consequently, the matrix $\varpi_{k,l}\mathbf{I}_{M-1} - \mathcal{U}^T\boldsymbol{\Omega}_{k,l}\mathcal{U}$, $\forall k \in \mathcal{K}, \forall l \in \mathcal{L}$, takes on the form of a diagonal matrix, with its diagonal elements set as $[\varpi_{k,l} - \kappa_{M-1}; \varpi_{k,l} - \kappa_{M-2}; ...; \varpi_{k,l} - \kappa_1]$. In light of this, we observe that $\boldsymbol{\Omega}_{k,l}$ possesses $M-1$ smallest eigenvalues all being zero if and only if the conditions $\varpi_{k,l}\mathbf{I}_{M-1} - \mathcal{U}^T\boldsymbol{\Omega}_{k,l}\mathcal{U} \succeq \mathbf{0}$ and $\varpi_{k,l} = 0$, $\forall k \in \mathcal{K}, \forall l \in \mathcal{L}$, are satisfied simultaneously. Consequently, $\boldsymbol{\Omega}_{k,l}$ qualifies as a rank one matrix under these conditions [12]. ∎

APPENDIX C
PROOF OF THE PROPOSITION 2

Let's denote ϱ^* be the optimal solution, corresponding to the optimal resource allocation policy $\boldsymbol{\Omega}_{k,l}^*$ of the objective function in P_5, that is:
$$\varrho^* = \max_{\boldsymbol{\Omega}_{k,l}} \frac{\sum_{k \in \mathcal{K}} \tilde{\mathcal{R}}_k(\boldsymbol{\Omega}_{k,l})}{\mathcal{E}_{\text{tot}}(\boldsymbol{\Omega}_{k,l})} - \varsigma^{(q)} \sum_{l \in \mathcal{L}} \sum_{k \in \mathcal{K}} \varpi_{k,l}. \quad (42)$$
Therefore, the optimal EE must satisfy the following inequality:
$$\varrho^* = \frac{\sum_{k \in \mathcal{K}} \tilde{\mathcal{R}}_k^*(\boldsymbol{\Omega}_{k,l}^*)}{\mathcal{E}_{\text{tot}}^*(\boldsymbol{\Omega}_{k,l}^*)} - \varsigma^{(q)} \sum_{l \in \mathcal{L}} \sum_{k \in \mathcal{K}} \varpi_{k,l}$$
$$\geq \frac{\sum_{k \in \mathcal{K}} \tilde{\mathcal{R}}_k(\boldsymbol{\Omega}_{k,l})}{\mathcal{E}_{\text{tot}}^{(d)}(\boldsymbol{\Omega}_{k,l})} - \varsigma^{(q)} \sum_{l \in \mathcal{L}} \sum_{k \in \mathcal{K}} \varpi_{k,l}. \quad (43)$$
It becomes readily apparent that:
$$\sum_{k \in \mathcal{K}} \tilde{\mathcal{R}}_k(\boldsymbol{\Omega}_{k,l}) - \varrho^*\mathcal{E}_{\text{tot}}(\boldsymbol{\Omega}_{k,l}) - \varsigma^{(q)} \sum_{l \in \mathcal{L}} \sum_{k \in \mathcal{K}} \varpi_{k,l}\mathcal{E}_{\text{tot}}(\boldsymbol{\Omega}_{k,l}) \leq 0, \quad (44)$$
Thus, it can be inferred that:
$$\sum_{k \in \mathcal{K}} \tilde{\mathcal{R}}_k(\boldsymbol{\Omega}_{k,l}^*) - \varrho^*\mathcal{E}_{\text{tot}}(\boldsymbol{\Omega}_{k,l}^*) - \varsigma^{(q)} \sum_{l \in \mathcal{L}} \sum_{k \in \mathcal{K}} \varpi_{k,l}\mathcal{E}_{\text{tot}}(\boldsymbol{\Omega}_{k,l}^*) = 0, \quad (45)$$
Therefore, we have:
$$\max_{\boldsymbol{\Omega}_{k,l}} \sum_{k \in \mathcal{K}} \tilde{\mathcal{R}}_k(\boldsymbol{\Omega}_{k,l}) - \varrho^* \mathcal{E}_{\text{tot}}(\boldsymbol{\Omega}_{k,l})$$
$$- \varsigma^{(q)} \sum_{l \in \mathcal{L}} \sum_{k \in \mathcal{K}} \varpi_{k,l}\mathcal{E}_{\text{tot}}(\boldsymbol{\Omega}_{k,l}) = 0, \quad (46)$$
and this can be attained through the resource allocation policy. This concludes the proof. ∎

APPENDIX D
PROOF OF THE PROPOSITION 3

Let's denote the objective functions of P_1, P_5, and P_7 respectively as $\daleth_{P_1}$, $\daleth_{P_5}$, and $\daleth_{P_7}$. Also, in consideration of $\{\boldsymbol{\Omega}_{k,l}^s, \boldsymbol{\Psi}_l^s\}$ and $\{\boldsymbol{\Omega}_{k,l}^{s-1}, \boldsymbol{\Psi}_l^{s-1}\}$ as the feasible solutions of P_1 in the s-th and $(s-1)$-th iterations respectively, we can establish the following inequalities:
$$\daleth_{P_1}(\boldsymbol{\Omega}_{k,l}^s, \boldsymbol{\Psi}_l^s) = \daleth_{P_7}(\boldsymbol{\Omega}_{k,l}^*, \boldsymbol{\Psi}_l^s)$$
$$\geq \daleth_{P_7}(\boldsymbol{\Psi}_l^{s-1}) = \daleth_{P_1}(\boldsymbol{\Omega}_{k,l}^s, \boldsymbol{\Psi}_l^{s-1}), \quad (47)$$
$$\daleth_{P_1}(\boldsymbol{\Omega}_{k,l}^s, \boldsymbol{\Psi}_l^{s-1}) = \daleth_{P_5}(\boldsymbol{\Omega}_{k,l}^s, \boldsymbol{\Psi}_l^*)$$
$$\geq \daleth_{P_5}(\boldsymbol{\Omega}_{k,l}^{s-1}) = \daleth_{P_1}(\boldsymbol{\Omega}_{k,l}^{s-1}, \boldsymbol{\Psi}_l^s). \quad (48)$$
Utilizing inequalities (47) and (48), we can guarantee the improvement in the value of $\daleth_{P_1}$ after every iteration. ∎

REFERENCES

[1] Q. Wu and R. Zhang, "Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network," *IEEE Commun. Mag.*, vol. 58, pp. 106–112, Jan. 2020.

[2] M. Bennis, M. Debbah, and H. V. Poor, "Ultrareliable and low-latency wireless communication: Tail, risk, and scale," *Proc. IEEE*, vol. 106, pp. 1834–1853, Oct. 2018.

[3] Y. Polyanskiy, H. V. Poor, and S. Verdu, "Channel coding rate in the finite blocklength regime," *IEEE Trans. Inf. Theory*, vol. 56, pp. 2307–2359, May 2010.

[4] R. Hashemi, S. Ali, N. H. Mahmood, and M. Latva-aho, "Average rate and error probability analysis in short packet communications over RIS-aided URLLC systems," *IEEE Trans. Veh. Tech.*, vol. 70, pp. 10320–10334, Aug. 2021.

[5] S. Dhok, P. Raut, P. K. Sharma, K. Singh, and C.-P. Li, "Non-linear energy harvesting in RIS-Assisted URLLC networks for industry automation," *IEEE Trans. Commun.*, vol. 69, pp. 7761–7774, Nov. 2021.

[6] H. Xie, J. Xu, Y.-F. Liu, L. Liu, and D. W. K. Ng, "User grouping and reflective beamforming for IRS-Aided URLLC," *IEEE Wirel. Commun. Lett.*, vol. 10, pp. 2533–2537, Nov. 2021.

[7] Z. Li, H. Shen, W. Xu, P. Zhu, and C. Zhao, "Resource allocation for irs-assisted uplink URLLC systems," *IEEE Commun. Lett.*, vol. 27, pp. 1540–1544, Jun. 2023.

[8] J. Jalali, A. Rezaei, A. Khalili, and J. Famaey, "Power-efficient joint resource allocation and decoding error probability for multiuser downlink MISO with finite block length codes," in *2022 Proc. IEEE 25th Int. Symp. Wirel. Personal Multimed. Commun. (WPMC)*, pp. 232–237, Oct. 2022.

[9] J. Jalali, A. Khalili, A. Rezaei, R. Berkvens, M. Weyn, and J. Famaey, "IRS-based energy efficiency and admission control maximization for IoT users with short packet lengths," *IEEE Trans. Veh. Technol.*, vol. 72, pp. 12379–12384, Sept. 2023.

[10] J. Jalali, A. Khalili, R. Berkvens, and J. Famaey, "Joint offloading policy and resource allocation in IRS-aided MEC for IoT users with short packet transmission," in *2023 Proc. IEEE Veh. Technol. Conf. (VTC2023-Fall)*, pp. 1–7, Oct. 2023.

[11] Y. Polyanskiy, H. V. Poor, and S. Verdu, "Channel coding rate in the finite blocklength regime," *IEEE Trans. Inf. Theory*, vol. 56, pp. 2307–2359, May 2010.

[12] C. Sun and R. Dai, "An iterative rank penalty method for nonconvex quadratically constrained quadratic programs," *SIAM J. Control Optim.*, vol. 57, pp. 3749–3766, Jan. 2019.

[13] W. Dinkelbach, "On nonlinear fractional programming," *J. Manag. Sci.*, vol. 13, no. 3, p. 492–498, 1967.

[14] J. Jalali, *Resource allocation for SWIPT in multi-service wireless networks*. M.S. thesis, Dept. Telecommun. Inf. Process., TELIN/IMEC, Ghent Univ., Ghent, Belgium, Jun. 2020.

[15] J. Jalali, A. Khalili, H. Tabassum, R. Berkvens, J. Famaey, and W. Saad, "On the energy efficiency of THz-NOMA enhanced UAV cooperative network with SWIPT," Sept. 2023, [Online]. Available: https://arxiv.org/pdf/2309.13836.pdf.

[16] J. Lee and S. Leyffer, *Mixed Integer Nonlinear Programming*. Springer Science & Business Media, Dec. 2011.

[17] Y. Sun, P. Babu, and D. P. Palomar, "Majorization-minimization algorithms in signal processing, communications, and machine learning," *IEEE Trans. Signal Process.*, vol. 65, pp. 794–816, Feb. 2017.