Energy-Aware Resource Optimization for Improved URLLC in Multi-hop Integrated Aerial Terrestrial Networks

Muhammad Awais, *Graduate Student Member, IEEE,* Haris Pervaiz, *Member, IEEE,* Muhammad Ali Jamshed, *Senior Member, IEEE,* Wenjuan Yu, *Member, IEEE,* and Qiang Ni, *Senior Member, IEEE,*

***Abstract*—The development of futuristic wireless infrastructure necessitates low power consumption, high reliability, and massive connectivity. One of the most promising solutions to address these requirements is the integration of aerial base station (ABS) based communication systems that employ both in the air (aerial) and on the ground (terrestrial) components. This integration enhances line of sight connections, enabling the fulfillment of escalating quality-of-service (QoS) demands. This article examines the problem of resource allocation in ABS assisted multi-hop wireless networks. We investigate a joint optimization problem that involves subcarrier (SC) assignment, power allocation, and blocklength allocation, subject to delay, reliability, and QoS constraints to improve the sum-rate under the finite blocklength (FBL) regime. We propose an approach for SC allocation and selection of cooperative ABSs based on matching theory. Subsequently, we employ an alternating optimization method to propose a novel bisection-based low-complexity adaptation (BLCA) algorithm to optimize the resource allocation policy. This algorithm includes a two-step projected gradient descent-based strategy to optimize the power allocation on each SC using dynamic and geometric programming. Furthermore, we examine flexible blocklength and power allocation use cases under the next generation of multiple access techniques. Monte-Carlo simulations validate that the proposed algorithmic solution significantly achieves a near-optimal solution while requiring 1600 times less computational cost compared to benchmarks in its counterparts.**

***Index Terms*—URLLC, ABSs, Resource allocation, Integrated Aerial Terrestrial Networks, blocklength optimization.**

I. Introduction

THE emergence of advanced wireless infrastructure has transformed how information is generated, disseminated, received, and perceived [1]. The capacity is expected to increase by up to 1000 times to support the growing number of wireless users and internet of things (IoTs) devices [2]. Therefore, a few novel communication paradigms are needed to address three key connectivity types that align with the new technical requirements: enhanced mobile broadband (eMBB) to provide high throughput for demanding clients, massive machine-type communication (mMTC) to support low-cost, low-power IoT devices, and ultra-reliable low-latency communication (URLLC) to support mission-critical IoT devices, such as tactile internet and autonomous vehicles, which require stringent quality of service (QoS) requirements to achieve a delay of less than one millisecond and reliability greater than 99.9999% [3].

In the context of 5G and beyond, aerial base stations (ABSs) have recently emerged as highly adaptable airborne wireless technology to ensure massive connectivity with minimal human intervention. They provide high data rates and a wide range of services, including monitoring of the Internet of agricultural things, surveillance during natural disasters, and data offloading in different hotspots [4]. Properties like low infrastructure expenditures, controlled mobility, and flexible deployment also make them a great choice [5]. Non-orthogonal multiple access (NOMA) is emerging as a new paradigm in cellular networks and beyond applications to address the problem of scarcity in shared spectrum resources [6]. It helps maintain high link quality ubiquitously and increases spectral efficiency by exploiting available resources more efficiently. Therefore, the integration of NOMA allows ABSs to provide a practical pathway to meet the demand for massive connectivity and improved energy efficiency [7].

The emergence of short packets with finite blocklength (FBL) is considered a key enabler for supporting emerging technologies such as intelligent transportation systems and virtual reality [8]. Given that advanced wireless networks require reliable and efficient transmission, studying communication systems in the FBL regime becomes crucial. However, the total achievable rate cannot be approximated using the Shannon capacity formula, which demands an alternative solution [9]. Prior work not only highlights some inherent problems, such as channel estimation errors and successive interference cancellation (SIC), but also brings further challenges, such as the spatial distribution of ABSs which poses challenges in the security domain, and flexible mobility which makes the channel more complex [10]. Therefore, there is a strong need to explore the potential applications of integrated aerial terrestrial communication in the FBL regime by utilizing next generation multiple access techniques[1].

Part of this paper [1] is presented at Smart Computing for Smart Cities-IEEE WoWMoM 2022 workshop.

Muhammad Awais, Haris Pervaiz, Wenjuan Yu, and Qiang Ni are with the School of Computing and Communications, Lancaster University, Lancaster LA1 4YW, UK. e-mails: {m.awais11, h.b.pervaiz, w.yu8, q.ni}@lancaster.ac.uk. Muhammad Ali Jamshed is with James Watt School of Engineering, University of Glasgow, G12 8QQ, UK. e-mail: muhammadali.jamshed@glasgow.ac.uk.

Manuscript received March 27, 2023; revised September 28, 2023.

[1]It refers to innovative techniques that surpass conventional approaches. In the context of 5G and beyond, these techniques explore novel methodologies, i.e., NOMA to optimize resource allocation, resulting an improved URLLC characterized by higher data rates, reduced latency, and enhanced connectivity.

In the leading study of [11], the potential of intelligent reflecting surfaces (IRS) based ABS is investigated. It is tested in real-world settings including constrained QoS requirements. However, the proposed method is not suitable for need based networks specifically in situations like disaster response and mission-critical applications (where the existing infrastructure may be severely damaged or unavailable). Hence, a small malfunction or failure in reflective elements can have a significant impact on overall network performance and reliability. The author in [12] proposes a low-complexity algorithm to position the ABS and plan an efficient route for data collection, resulting in improved performance subject to delay. The researchers observed that the placement of ABSs is complex and requires attention. Therefore, the authors in [13] developed a framework that uses Markov chain and Gibbs sampling. The study is extended to use clustering for deploying ABSs and user association using NOMA. However, the number of covered users served by a single ABS is limited [14].

Considering these factors, it is better to rely on more traditional and reliable communication technologies to align better with the requirements and objectives of a need based network. Motivated by the benefits of ABS and NOMA, their integration is analyzed and investigated. In [15], opportunistic channel gain disparities against each IoT device are identified, and the positions of ABSs are optimized, which becomes more challenging when power limitations constrain the problem. To maximize the minimum rate, [16] investigates aerial jamming and power allocation to enhance security and reliability in ABS-assisted NOMA communication. Additionally, in the latter work, a relay selection strategy is explored to optimize the power allocation of ABSs and to maximize energy efficiency under the NOMA scheme [17].

To meet the increasing requirements of URLLC, the use of NOMA is being investigated in the FBL regime with reliability constraints [18]. Due to the benefits of FBL, efficient bandwidth allocation schemes that consider delay constraints have also been developed. However, the use of multiple hops adds complexity to both the resource allocation and decision-making processes. In [19], the authors optimize the amount of information transmitted from the control station of an ABS-aided system by concurrently optimizing blocklength and transmit power. This work is further extended to jointly optimize ABS placement and transmit power to reduce decoding error probability (DEP) [20]. The research is then expanded to an optimal resource allocation technique for heterogeneous communication links that use both orthogonal multiple access (OMA) and NOMA [21]. However, it should be noted that the proposed approach is limited to throughput maximization and does not provide a closed-form expression. Hence, developing a low-complexity and connectivity-aware optimal resource allocation policy is crucial for enhancing network performance. However, the strong coupling between optimization variables poses a challenge, especially when grouping IoT devices in multi-carrier transmission. In this work, we investigate the development of a resource allocation strategy for multi-carrier communication. It involves a fast network formation where integrated aerial terrestrial networks[2] can be used to ensure IoT device connectivity in a signal dead zone. The primary contributions of this work are outlined below.

- We examine a problem in an integrated aerial terrestrial network that involves mixed-integer non-linear programming. To address this issue, we reframe the problem of maximizing the sum-rate by utilizing its decomposition property and jointly optimizing the channel allocation, power allocation, and blocklength allocation for both OMA and NOMA systems.
- We utilize the alternating optimization method to present an iterative bisection-based low-complexity adaption (BLCA) algorithm for optimizing the resource allocation problem subject to delay, reliability, and QoS constraints. The formulated problem is solved in three steps. Firstly, we employ the matching theory to allocate subcarriers (SCs) and select the best cooperative ABSs. Secondly, we compute the optimal blocklength using the bisection algorithm. Finally, we use dynamic and geometric programming to perform power distribution by optimizing the power budget on each SC with a two-tier projected gradient descent-based algorithm.
- We evaluate the optimality of the BLCA algorithm against a high complexity benchmark scheme, namely lagrangian duality and dynamic programming (LDDP), which employs Lagrangian dual to relax the individual power constraint [22]. To this end, Monte Carlo simulations are conducted to compare the performance of the proposed algorithmic solution against the LDDP scheme. Moreover, for comparative analysis, two benchmark algorithms are also implemented, where the worst SC is avoided being assigned to a transmitting node [23], [24]. Additionally, the proposed algorithmic solution is analyzed against random and fixed blocklength approaches using legacy OMA and NOMA with different power allocation schemes. Results show that the proposed algorithmic solution significantly achieves a near-optimal solution and outperforms the LDDP scheme.

The remaining paper is structured as follows. Section II presents the case of interest and a mathematical framework for the proposed work. Section III formulates the problem and Section IV presents the proposed solution. Section V provides simulation results with in-depth analysis. Finally, Section VI concludes the paper.

II. System Model

Fig. 1 illustrates the conceptual architecture of an integrated aerial terrestrial multi-hop downlink network. The architecture consists of a single macro base station (MBS) that utilizes next-generation multiple access schemes to establish communication with multiple IoT devices. The system comprises a set of ABSs denoted by $u \in \mathcal{U} = \{1, 2, ..., |\mathrm{U}|\}$. This set

[2]It refers to the integration and convergence of communication systems that employ both in the air (aerial) on the ground (terrestrial). This concept helps to establish a unified network infrastructure that inherits the capabilities of aerial platforms such as ABSs and airborne systems with conventional terrestrial communication systems.

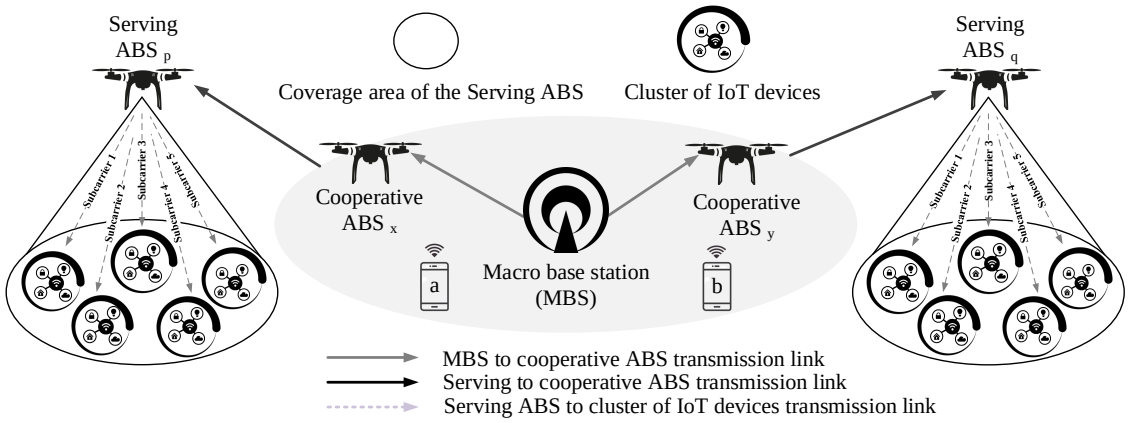

Fig. 1. Considered scenario: $\mathcal{U}_h = \{x, y\}$, where $(x, y) \in \mathcal{U}_h$, $\mathcal{U}_u = \{p, q\}$, where $(p, q) \in \mathcal{U}_u$, $|\mathcal{I}|$ IoT devices are grouped into $\mathcal{L}$ communities located within the coverage of serving ABSs, and $\mathcal{M} = \{a, b\}$, where $(a, b) \in \mathcal{M}$ located within the coverage of MBS.

$\mathcal{U}$ is further categorized into two subsets based on different QoS: the set of cooperative ABSs and the set of serving ABSs denoted by $\mathcal{U}_u$ and $\mathcal{U}_h$, respectively. The cooperative ABSs[3] act as relay nodes to facilitate the successful transmission of messages to the i^{th} IoT device. The set of $\mathcal{U}_h$ is represented by $u_h \in \mathcal{U}_h = \{x, y, ..., |\mathrm{U}_h|\}$ and set of $\mathcal{U}_u$ is expressed as $u_u \in \mathcal{U}_u = \{p, q, ..., |\mathrm{U}_u|\}$. The IoT devices within the coverage of MBS are denoted by the set $m \in \mathcal{M} = \{a, b, ..., |\mathrm{M}|\}$ and IoT devices within the coverage of each serving ABS are represented by set $i \in \mathcal{I} = \{u_k^{(1)}, u_k^{(2)}, ..., |\mathrm{I}|\}$. It is important to note that all IoT devices are positioned on the ground and can be served either directly from MBS or through ABS using multi-hop communication.

As it is a need-based network, therefore it is reasonable to consider that the network is resource-constrained and has limited bandwidth. We devide the total bandwidth W into $|C|$ orthogonal SCs, denoted by $c \in \mathcal{C} = \{1, 2, ..., |C|\}$, i.e., $\sum_{c \in \mathcal{C}} w_c = W$. Moreover, all devices (IoT devices and ABSs) are grouped into $\mathcal{L}$ communities. The set $\mathcal{L}$ is expressed as $l \in \mathcal{L} = \{1, 2, ..., |L_c|\}$, where $c \in \mathcal{C}$ and L_c denotes the maximum number of devices that can be served on the given SC c. If $L_c = 1$, it means incorporating the novel concepts of the orthogonal scheme, whereas $1 < L_c \leq |S|$ means the incorporation of the NOMA scheme. Hence we define two sets, $n \in \mathcal{N} = \{c | c \in \mathcal{C}, L_c = 1\}$ and $k \in \mathcal{K} = \{c | c \in C, 1 < L_c \leq |S|\}$ containing the indexes of the OMA SCs and NOMA SCs, respectively. The notation $|S|$ presents the threshold value for the maximum number of devices to combine multiple devices on a given SC.

Assume the set $\mathcal{U}_k$ and $\mathcal{U}_n$ containing the indexes of the devices assigned to NOMA SC ($k \in \mathcal{K}$) and OMA SC ($n \in \mathcal{N}$). We define a set $U_{k,[u_k]} = \{\cup\ i, \|h_{u_u,[i]}\|^2 \geq \|h_{u_u,[u_k]}\|^2, i, u_k \in \mathcal{U}_k\}$ containing the indexes for IoT devices imposing interference on IoT device u_k allocated to the SC k within the same community, where $h_{u_u,[i]}$ denotes the channel of i^{th} IoT device served by serving ABS u_u. The priority to provide fairness among each IoT device is given by $\sum_{u_k \in \mathcal{U}_k} \omega_{k,[u_k]} = 1$. The SIC decoding order is also important for the power domain NOMA. We consider the optimal decoding error, where we decode the IoT device's signals from highest to lowest normalized noise power [25]. If decoding order on the given SC is $\pi_k : \{U_{k,[u_k]}, u_k \in \mathcal{U}_k\}$. For $i \in \{1, 2, ..., |U_k|\}$, $\pi_k(i)$ gives the location of the i^{th} decoded device on the k^{th} SC while $\pi_k^{-1}(i)$ gives its decoding order. The IoT device $\pi_k(i)$ first decodes the signals from IoT devices $\pi_k(1)$ to $\pi_k(i-1)$ and before decoding the needed signal, it subtracts them from the overlaid signal. The intervention from the IoT devices $\pi_k(\bar{i})$ for $\bar{i} > i$ is considered as noise.

It is assumed that SCs belonging to set $\mathcal{N}$ can serve at-most single cooperative ABS or IoT devices (within the coverage of MBS). Still, one SC can serve multiple links simultaneously, i.e., the link between MBS to the IoT device and between cooperative to serving ABS. However, each transmission link can be allocated to no more than one SC between the MBS to IoT devices (within its coverage) and the link between MBS to the cooperative ABSs. The SCs belonging to the set $\mathcal{K}$ can be used by a maximum one serving ABS, and that SC can only be allocated among |S| IoT devices within the community, where $s \in \mathcal{S} = \{2, 3, ..., |\mathrm{S}|\}$. A matrix with size $(|\mathrm{U}_h| + |\mathrm{M}|) \times \mathrm{N}$ is defined to describe the SC allocation indicator for MBS to cooperative ABSs, and MBS to IoT devices within its vicinity. It is denoted by $\boldsymbol{\psi} = \left[\psi^n_{\mathrm{mbs},j}\right]$. For $j \leq |\mathrm{U}_h|$, $\psi^n_{\mathrm{mbs},j} = 1$ means a SC n is assigned to u_h^{th} cooperative ABS, otherwise $\psi^n_{\mathrm{mbs},j} = 0$. Whereas, for $j > |\mathrm{U}_h|$, $\psi^n_{\mathrm{mbs},j} = 1$ means a SC n is assigned to a IoT device m, otherwise $\psi^n_{\mathrm{mbs},j} = 0$. We define a matrix with size $|\mathrm{U}_u| \times \mathrm{K}$ to describe the SC allocation indicator for serving ABS to IoT devices within its vicinity, shown by $\boldsymbol{\varphi} = \left[\varphi^k_{u_u,[u_k]}\right]$. Therefore, $\varphi^k_{u_u,[u_k]} = 1$ means a SC k is assigned to IoT device u_k, otherwise $\varphi^k_{u_u,[u_k]} = 0$.

The physical locations of the MBS, m^{th} IoT device, u_h^{th} cooperative ABS, u_u^{th} serving ABS, and u_k^{th} IoT device are denoted as $(x = 0, y = 0, z = 0)$, (x_m, y_m, z_m), $(x_{u_h}, y_{u_h}, z_{u_h})$, $(x_{u_u}, y_{u_u}, z_{u_u})$ and $(x_{u_k}, y_{u_k}, z_{u_k})$, respec-

[3]We assume all cooperative ABSs are situated within the coverage of MBS, and the most optimal cooperative ABSs among them will be elected as relay nodes.

tively[4]. We have used different channel models due to the different channel characteristics in line of sight (LoS) and non-LoS (NLoS) probabilities for air to ground, ground to ground, and ground to air propagation models [26]. The distance between MBS and u_h^{th} cooperative ABS and the distance between MBS and m^{th} IoT device are computed as $d_{\text{mbs},u_h} = \sqrt{(x-x_{u_h})^2+(y-y_{u_h})^2+(z-z_{u_h})^2}$ and $d_{\text{mbs},m} = \sqrt{(x-x_m)^2+(y-y_m)^2+(z-z_m)^2}$, respectively. The pathloss for the given link is given by $l_{\text{mbs},m} = 55.9 + 38\log(d_{\text{mbs},m}) + (24.5 + 1.5f/925)\log(f)$, where f represents carrier frequency. The distance between ${u_h}^{th}$ cooperative ABS and ${u_u}^{th}$ serving ABS is given as $d_{u_h,u_u} = \sqrt{(x_{u_h}-x_{u_u})^2+(y_{u_h}-y_{u_u})^2+(z_{u_h}-z_{u_u})^2}$, where $d_{u_h,u_u}^{-\alpha_{\text{mbs}}}$ represents the pathloss, where α_{mbs} represents the pathloss exponent.

The LoS probability between IoT device u_k and serving ABS u_u is given as $p^{\text{los}}_{u_u,[u_k]} = \frac{1}{1+a\,\exp\left(-b\,\left[(\frac{180}{\pi})\left(\sin^{-1}(\frac{z_{u_u}}{d_{u_u,u_k}})\right)-a\right]\right)}$, where d_{u_u,u_k} is the distance between the given IoT device and serving ABS, computed as $d_{u_u,u_k} = \sqrt{(x_{u_u}-x_{u_k})^2+(y_{u_u}-y_{u_k})^2+(z_{u_u}-z_{u_k})^2}$, and a and b are the constant values depending on environmental factors. The probability of establishing non-LoS link is $p^{\text{nlos}}_{u_u,[u_k]} = 1 - p^{\text{los}}_{u_u,[u_k]}$. The pathloss between IoT device u_k and serving ABS u_u for LoS and N-LoS connection is given by $l^{\text{los}}_{u_u,[u_k]} = l_{\text{fsp}} + 20\log_{10}(d_{u_u,u_k}) + \eta_{\text{los}}$, and $l^{\text{nlos}}_{u_u,[u_k]} = l_{\text{fsp}} + 20\log_{10}(d_{u_u,u_k}) + \eta_{\text{nlos}}$, respectively. The free space pathloss is given as $l_{\text{fsp}} = 20\log_{10}(f) + 20\log_{10}\left(\frac{4\pi}{c}\right)$, where c represents the speed of light, η_{los} and η_{nlos} presents the attenuation due to LoS and N-LoS connection, respectively. Thus, the average pathloss is given by $pl^{\text{avg}}_{u_u,[u_k]} = p^{\text{los}}_{u_u,[u_k]} l^{\text{los}}_{u_u,[u_k]} + p^{\text{nlos}}_{u_u,[u_k]} l^{\text{nlos}}_{u_u,[u_k]}$ [26].

Let h^n_{mbs,u_h} be the channel between the MBS and cooperative ABS u_h. It is computed as $h^n_{\text{mbs},u_h} = \frac{g^n_{\text{mbs},u_h}}{\left[(x-x_{u_h})^2+(y-y_{u_h})^2+(z-z_{u_h})^2\right]}$, where g^n_{mbs,u_h} is the channel power gain on the given SC. The signal-to-noise ratio (SNR) at the u_h^{th} cooperative ABS on the given SC is computed as $\varrho^n_{\text{mbs},u_h} = \frac{\psi^n_{\text{mbs},u_h} p^n_{\text{mbs},u_h} \|h^n_{\text{mbs},u_h}\|^2}{\delta^2}$, where δ^2 is the noise spectral density, and p^n_{mbs,u_h} shows the power allocated to the given cooperative ABS. The achievable rate for the given link is computed by normalizing over the SC's bandwidth w_n [27].

$$r^n_{\text{mbs},u_h} = \log_2\left(1+\varrho^n_{\text{mbs},u_h}\right) - \sqrt{\frac{V^n_{\text{mbs},u_h}}{b_{\text{mbs},u_h}}}\frac{Q^{-1}(\epsilon_{\text{mbs},u_h})}{\ln 2}, \quad (1)$$

where, b_{mbs,u_h} is the adopted blocklength, and Q is the Gaussian Q-function, i.e., $Q(x) = \frac{1}{2\pi}\int_x^\infty \exp(-\frac{t^2}{2})\,dt$ [4]. The DEP for the link between MBS to cooperative ABS u_h is approximated as $\epsilon_{\text{mbs},u_h} \approx Q\left(f(\varrho^{\min}_{\text{mbs},u_h}, r^{u_h}_{\min}, b_{\text{mbs},u_h})\right),\ \forall\, u_h \in \mathcal{U}_h$, where $\varrho^{\min}_{\text{mbs},u_h}$ is the minimum received SNR across all the allocated SCs for the link between MBS and cooperative ABS u_h and $r^{u_h}_{\min}$ is the minimum achievable rate. The overall DEP for this hop is given by $\epsilon_{\text{mbs},u_h}, \forall u_h \in \mathcal{U}_h$. The channel dispersion for the given link is computed by $V^n_{\text{mbs},u_h} = 1 - (1+\varrho^n_{\text{mbs},u_h})^{-2}$. The sum-rate for the given cooperative ABS is computed as $r_{\text{mbs},u_h} = \sum_{n\in\mathcal{N}} \psi^n_{\text{mbs},u_h} r^n_{\text{mbs},u_h},\ \forall\, u_h \in \mathcal{U}_h$.

Let $h^n_{\text{mbs},m}$ be the channel between the MBS and m^{th} IoT device, which is defined as $h^n_{\text{mbs},m} = \frac{g^n_{\text{mbs},m}}{l_{\text{mbs},m}}$, where $g^n_{\text{mbs},m}$ is the channel power gain. The received signal to interference plus noise ratio (SINR) at the m^{th} device is computed as $\varrho^n_{\text{mbs},m} = \frac{\psi^n_{\text{mbs},m} p^n_{\text{mbs},m} \|h^n_{\text{mbs},m}\|^2}{\delta^2 + I^n_{u_h,m}}$. where $p^n_{\text{mbs},m}$ denotes the transmitted power for the m^{th} device, and $I^n_{u_h,m}$ is the interference power caused by the re-used link between cooperative to serving ABSs. It is defined as $I^n_{u_h,m} = p^n_{u_h,u_u} h^n_{u_h,m}$, where $p^n_{u_h,u_u}$ is the allocated power for the link between cooperative to serving ABS and $h^n_{u_h,m}$ is the channel between u_h^{th} cooperative ABS and m^{th} IoT device. It is given by $h^n_{u_h,m} = g^n_{u_h,m} l^{-\alpha_{\text{mbs}}}_{u_h,m}$, where $g^n_{u_h,m}$ is the channel gain and $l^{-\alpha_{\text{mbs}}}_{u_h,m}$ is the pathloss for the following channel with pathloss exponent α_{mbs}. The achievable rate for the given link is computed by normalizing over the SC's bandwidth w_n.

$$r^n_{\text{mbs},m} = \log_2\left(1+\varrho^n_{\text{mbs},m}\right) - \sqrt{\frac{V^n_{\text{mbs},m}}{b_{\text{mbs},m}}}\frac{Q^{-1}(\epsilon_{\text{mbs},m})}{\ln 2}, \quad (2)$$

where, $b_{\text{mbs},m}$ is the adopted blocklength. The DEP for the link between MBS and IoT device m is approximated as $\epsilon_{\text{mbs},m} \approx Q\left(f(\varrho^{\min}_{\text{mbs},m}, r^m_{\min}, b_{\text{mbs},m})\right),\ \forall\, m \in \mathcal{M}$, where $\varrho^{\min}_{\text{mbs},m}$ is the minimum received SNR across all the allocated SCs for the link between MBS to the IoT device m and $r^m_{\min}$ is the minimum achievable rate. The overall DEP for this hop is given by $\epsilon_{\text{mbs},m}, \forall m \in \mathcal{M}$. The channel dispersion for the given link is computed by $V^n_{\text{mbs},m} = 1 - (1+\varrho^n_{\text{mbs},m})^{-2}$. The sum-rate for the given IoT device is computed as $r_{\text{mbs},m} = \sum_{n\in\mathcal{N}} \psi^n_{\text{mbs},m} r^n_{\text{mbs},m},\ \forall\, m \in \mathcal{M}$.

Let $h^n_{u_h,u_u}$ be the channel between cooperative ABS u_h and serving ABS u_u, which is given by $h^n_{u_h,u_u} = g^n_{u_h,u_u} d^{-\alpha_{\text{mbs}}}_{u_h,u_u}$, where $g^n_{u_h,u_u}$ represents the channel gain. The SINR is computed as $\varrho^n_{u_h,u_u} = \frac{\psi^n_{u_h,u_u} p^n_{u_h,u_u} \|h^n_{u_h,u_u}\|^2}{\delta^2 + I^n_{u_h,u_u}}$ at the u_u^{th} serving ABS. The interference power caused by IoT devices in set $\mathcal{M}$ is defined as $I^n_{u_h,u_u} = \sum_{m\in\mathcal{M}} p^n_{\text{mbs},m} h^n_{\text{mbs},u_u}$, where h^n_{mbs,u_u} represents the channel between the MBS to serving ABS u_u. It is computed as $h^n_{\text{mbs},u_u} = g^n_{\text{mbs},u_u} \times l^{-\alpha_{\text{mbs}}}_{\text{mbs},u_u}$, where the terms g^n_{mbs,u_u} and $l^{-\alpha_{\text{mbs}}}_{\text{mbs},u_u}$ represents gain and pathloss between MBS and the given serving ABS, respectively. The achievable rate for the given link is computed as [27]

$$r^n_{u_h,u_u} = \log_2\left(1+\varrho^n_{u_h,u_u}\right) - \sqrt{\frac{V^n_{u_h,u_u}}{b_{u_h,u_u}}}\frac{Q^{-1}(\epsilon_{u_h,u_u})}{\ln 2}, \quad (3)$$

where, b_{u_h,u_u} is the allocated blocklength. We define $\epsilon_{u_h,u_u} \approx Q(f(\varrho^{\min}_{u_h,u_u}, r^{u_u}_{\min}, b_{u_h,u_u})),\ \forall\, u_u \in \mathcal{U}_u$ as the DEP for the link between cooperative ABS u_h and serving ABS u_u, where $\varrho^{\min}_{u_h,u_u}$ is the minimum received SNR across all the allocated SCs for the link between cooperative ABS u_h and serving ABS u_u and $r^{u_u}_{\min}$ is the minimum achievable rate. The overall DEP for the link between MBS and serving ABS u_u

[4]ABSs have diverse applications beyond communication. However, the scope of the current work primarily focuses on communication aspects that do not facilitate device to device communication. The locations of IoT devices are predetermined by the ABSs, whereas the pilot signals are used to determine the channel state information (CSI).

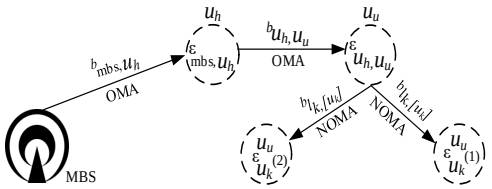

Fig. 2. Systematic diagram depicting blocklength and DEPs for the considered scenario setting.

is given by $\epsilon_2 \approx \epsilon_{\text{mbs},u_h} + (1-\epsilon_{\text{mbs},u_h}).\epsilon_{u_h,u_u}, \forall u_u \in \mathcal{U}_u$. The channel dispersion at the given link is computed as $V^n_{u_h,u_u} = \left(1-(1+\varrho^n_{u_h,u_u})^{-2}\right)$. The sum-rate for the link between cooperative ABS u_h and serving ABS u_u is computed as $r_{u_h,u_u} = \sum_{n\in\mathcal{N}} \psi^n_{u_h,u_u} r^n_{u_h,u_u},\ \forall\ u_u \in \mathcal{U}_u$.

Let $h_{u_u,[u_k]}$ is the channel between serving ABS u_u and IoT device u_k. It is defined as $h_{u_u,[u_k]} = \frac{g^k_{u_u,[u_k]}}{pl^{\text{avg}}_{u_u,[u_k]}}$, where $g^k_{u_u,[u_k]}$ is the channel gain for the given SC, and $pl^{\text{avg}}_{u_u,[u_k]}$ is the average pathloss. The SINR computed at u^{th}_k IoT device is expressed as $\varrho^k_{u_u,[u_k]} = \frac{\varphi^k_{u_u,[u_k]} p^k_{u_u,[u_k]} \|h_{u_u,[u_k]}\|^2}{\delta^2 + I^k_{u_u,[u_k]}}$, where $p^k_{u_u,[u_k]}$ is the transmitted power. The interference power caused by other IoT devices is given by $I^k_{u_u,[u_k]} = \sum_{u_k\in\mathcal{U}_k, u_k\neq u_k} g^k_{u_u,[u_k]} p^k_{u_u,[u_k]}$. The rate for the IoT device u_k is computed by normalizing over the given SC's bandwidth w^k.

$$r^k_{u_u,u_k} = \omega_{[u_k]} \log_2(1+\varrho^k_{u_u,[u_k]}) - \sqrt{\frac{V^k_{u_u,[u_k]}}{b_{l_{k,[u_k]}}}} \frac{Q^{-1}(\epsilon_{u_u,[u_k]})}{\ln 2}, \tag{4}$$

where, $\omega_{[u_k]}$ represents the priority of the given IoT device, and $b_{l_{k,[u_k]}}$ denotes the blocklength allocated to l^{th} community served by k^{th} SC. The channel dispersion for the given link is computed as $V^k_{u_u,[u_k]} = (1-(1+\varrho^k_{u_u,[u_k]})^{-2})$. The sum-rate for the given IoT device is computed as $r_{u_u,[u_k]} = \sum_{k\in\mathcal{K}} \varphi^k_{u_u,[u_k]} r^k_{u_u,u_k},\ \forall\ u_k \in \mathcal{U}_k$. The energy efficiency of the system is defined as the ratio of the system's achievable rate to the total system consumed power including flexible transmit power and circuit power (CP) [28].

A. DEP for the NOMA Phase

Considering NOMA |S|= 2 in Fig. 2, where two IoT devices namely: $u^{(1)}_k$ and $u^{(2)}_k$ are allocated to the SC k within the same community within the coverage of serving ABS u_u. The device $u^{(1)}_k$ is considered as a stronger user and device $u^{(2)}_k$ is considered weaker user. IoT device $u^{(1)}_k$ as a stronger device can perform SIC and first decodes the message of IoT device $u^{(2)}_k$ while treating its own message as interference. If this is successful, then it decodes its own message. Therefore, the total DEP of $u^{(1)}_k$ depends on the DEP of previous transmission links and successful SIC at $u^{(1)}_k$. Whereas, the IoT device $u^{(2)}_k$ directly decodes its signal while treating the message of IoT device $u^{(1)}_k$ as noise. Therefore, the total DEP of $u^{(2)}_k$ only depends on the DEP of previous transmission phases. The DEP for detecting the data of IoT device $u^{(2)}_k$ at IoT device $u^{(1)}_k$ is approximated as

$$\epsilon^{u_u}_{\left[u^{(1)}_k,u^{(2)}_k\right]} \approx Q\left(f\left(\varrho^{\min}_{\left[u^{(1)}_k,u^{(2)}_k\right]}, r^{[u_k]}_{\min}, b_{l_{k,[u_k]}}\right)\right), \tag{5}$$

where $r^{[u_k]}_{\min}$ is the minimum achievable rate of the IoT device and the minimum received SINR across all the allocated SCs for the IoT device $u^{(1)}_k$ related to detecting data of IoT device $u^{(2)}_k$ is computed as $\varrho^{\min}_{\left[u^{(1)}_k,u^{(2)}_k\right]} = \min\left(\varrho^1_{\left[u^{(1)}_k,u^{(2)}_k\right]}, \varrho^2_{\left[u^{(1)}_k,u^{(2)}_k\right]}, ..., \varrho^{|\mathcal{U}_k|}_{\left[u^{(1)}_k,u^{(2)}_k\right]}\right)$, where SIC is applied at the receiver end and $\varrho^k_{\left[u^{(1)}_k,u^{(2)}_k\right]} = \frac{p^k_{u_u,\left[u^{(2)}_k\right]} \left\| h_{u_u,\left[u^{(1)}_k\right]} \right\|^2}{\left(p^k_{u_u,\left[u^{(1)}_k\right]} \times \left\| h_{u_u,\left[u^{(1)}_k\right]} \right\|^2\right) + \delta^2}, \forall k \in \mathcal{K}$. The DEP for detecting the data of IoT device $u^{(1)}_k$ at the IoT device $u^{(1)}_k$ is given by

$$\epsilon^{u_u}_{\left[u^{(1)}_k,u^{(1)}_k\right]} \approx Q\left(f\left(\varrho^{\min}_{\left[u^{(1)}_k,u^{(1)}_k\right]}, r^{\left[u_k\right]}_{\min}, b_{l_{k,[u_k]}}\right)\right), \tag{6}$$

where, $\varrho^{\min}_{\left[u^{(1)}_k,u^{(1)}_k\right]}$ is the minimum received SNR across all allocated SC for the IoT device $u^{(1)}_k$ related to detecting the data of IoT device $u^{(1)}_k$, and $\varrho^k_{\left[u^{(1)}_k,u^{(1)}_k\right]} = \frac{p^k_{u_u,\left[u^{(1)}_k\right]} \left\| h_{u_u,\left[u^{(1)}_k\right]} \right\|^2}{\delta^2}$. Similarly, the DEP for detecting the data of IoT device $u^{(2)}_k$ at IoT device $u^{(2)}_k$ is given by

$$\epsilon^{u_u}_{\left[u^{(2)}_k,u^{(2)}_k\right]} \approx Q\left(f\left(\varrho^{\min}_{\left[u^{(2)}_k,u^{(2)}_k\right]}, r^{\left[u_k\right]}_{\min}, b_{l_{k,[u_k]}}\right)\right), \tag{7}$$

where, $\varrho^{\min}_{\left[u^{(2)}_k,u^{(2)}_k\right]}$ is the minimum received SNR across all allocated SCs for the IoT device $u^{(2)}_k$ related to detecting the data of IoT device $u^{(2)}_k$, and $\varrho^k_{\left[u^{(2)}_k,u^{(2)}_k\right]} = \frac{p^k_{u_u,\left[u^{(2)}_k\right]} \left\| h_{u_u,\left[u^{(2)}_k\right]} \right\|^2}{\left(p^k_{u_u,\left[u^{(1)}_k\right]} \times \left\| h_{u_u,\left[u^{(2)}_k\right]} \right\|^2\right) + \delta^2}$ after employing successful SIC. The overall DEPs for both IoT devices from MBS are given by

$$\epsilon^{u_u}_{\left[u^{(1)}_k\right]} = \epsilon_2 + (1-\epsilon_2).\left(\epsilon^{u_u}_{\left[u^{(1)}_k,u^{(2)}_k\right]} + \left(1-\epsilon^{u_u}_{\left[u^{(1)}_k,u^{(2)}_k\right]}\right).\epsilon^{u_u}_{\left[u^{(1)}_k,u^{(1)}_k\right]}\right), \tag{8}$$

$$\epsilon^{u_u}_{\left[u^{(2)}_k\right]} = \epsilon_2 + (1-\epsilon_2).\epsilon^{u_u}_{\left[u^{(2)}_k,u^{(2)}_k\right]}. \tag{9}$$

III. Problem Formulation

We aim to optimize the SC allocation, blocklength allocation, and power allocation to maximize the minimum feasible rates while ensuring that the delay, reliability, and

QoS constraints are met. The proposed optimization problem does not consider the transmission link between the MBS and IoT device $m \in \mathcal{M}$, as it only focuses on maximizing the minimum rate across each hop involved in transmitting information from the MBS to the IoT device $u_k \in \mathcal{U}_k$ [5]. The notation $\boldsymbol{b} = \left\{b_{\text{mbs},u_h}, b_{\text{mbs},m}, b_{u_h,u_u}, b_{l_{k,[u_k]}}, \forall u_h \in \mathcal{U}_h, \forall m \in \mathcal{M}, \forall u_u \in \mathcal{U}_u, \forall l \in \mathcal{L}\right\}$ represent the variable of blocklengths, $\boldsymbol{\phi} = \left\{\psi^n_{\text{mbs},j}, \varphi^k_{u_u,[u_k]}, \ \forall j \in \left\{(|\text{U}_h|+|\text{M}|) \times \text{N}\right\} \ \forall k \in \left\{|(\text{U}_u)|\times\text{K}\right\}\right\}$ represent the variable of SC associations, and the variable for allocated transmit powers is notated as $\boldsymbol{p} = \left\{P^n_{\text{mbs},m}, p^n_{\text{mbs},u_h}, p^n_{\text{mbs},m}, p^n_{u_h,u_u}, p^k_{u_u,[u_k]}, \forall m \in \mathcal{M} \ \forall u_k \in \mathcal{U}_k \ \forall u_h \in \mathcal{U}_h \ \forall u_u \in \mathcal{U}_u\right\}$. The term D_{max} denotes then maximum tolerable delay, T_{block} represents the duration of the time required to convey one unit of blocklength, $\propto^k_{u_u,u_k}$ is an association based binary variable, where $\propto^k_{u_u,u_k}= 1$ means IoT device u_k is served by the ABS u_u on the given SC k, otherwise 0. The optimization problem can be formulated as follows: 2

$$
\begin{aligned}
&\max_{\boldsymbol{p},\boldsymbol{b},\boldsymbol{\phi}} \ \min\left(r^n_{\text{mbs},u_h}, r^n_{u_h,u_u}, r^k_{u_u,u_k}\right) \\
&\text{subject to} \\
&C_1 : \sum_{j\in\mathcal{M}} \psi^n_{\text{mbs},j} \leq 1, \sum_{j\in\mathcal{U}_h} \psi^n_{\text{mbs},j} \leq 1, \sum_{k\in\mathcal{K}} \varphi^k_{u_u,[u_k]} \leq |\text{S}| \\
&C_2 : \max\left(b_{\text{mbs},m}, \left(b_{\text{mbs},u_h} + b_{u_h,u_u} + b_{l_{k,[u_k]}}\right)\right) \leq M_{\text{max}}, \\
&\quad \text{where, } M_{\text{max}} = \left(D_{\text{max}}/T_{\text{block}}\right), \\
&\quad \text{and } \left(b_{\text{mbs},m}, b_{\text{mbs},u_h}, b_{u_h,u_u}, b_{l_{k,[u_k]}}\right) \in \mathbb{Z}^+ \\
&C_3 : \left(x^2_{u_h} + y^2_{u_h}\right) \leq r^2_{\text{max}} \\
&C_4 : 0 < p_{\text{mbs},m} \leq p^{\text{min}}_{\text{mbs},m}, 0 < p_{\text{mbs},u_h} \leq p^{\text{min}}_{\text{mbs},u_h}, \\
&\quad 0 < p_{u_h,u_u} \leq p^{\text{min}}_{u_h,u_u},, \forall m \in \mathcal{M}, \forall u_h \in \mathcal{U}_h, \forall u_u \in \mathcal{U}_u \\
&C_5 : 0 \leq \sum_{u_k\in\mathcal{U}_k} \propto^k_{u_u,u_k} p^k_{u_u,[u_k]} \leq \bar{p}^k, \forall k \in \mathcal{K} \\
&C_6 : \left(\epsilon_{\text{mbs},m}, \epsilon_{\text{mbs},u_h}, \epsilon_{u_h,u_u}, \epsilon^{u_u}_{u_k}\right) \leq \epsilon_{\text{threshold}} \\
&C_7 : r_{\text{mbs},m} \geq r^m_{\text{min}}, r_{\text{mbs},u_h} \geq r^{u_h}_{\text{min}}, r_{u_h,u_u} \geq r^{u_u}_{\text{min}}, \\
&\quad r_{u_u,[u_k]} \geq r^{[u_k]}_{\text{min}}, \forall m \in \mathcal{M}, \forall u_h \in \mathcal{U}_h \ \forall u_u \in \mathcal{U}_u \ \forall u_k \in \mathcal{U}_k \\
&C_8 : \left(\frac{g^k_{u_u,[u_k]} p^k_{u_u,[u_k]}}{I^k_{u_u,[u_k]}}\right) \geq \hbar, \ \forall u_k \in \mathcal{U}_k \forall k \in \mathcal{K} \\
&C_9 : \left\|h_{u_u,[i]}\right\|^2 \geq \left\|h_{u_u,[u_k]}\right\|^2, i, u_k \in \mathcal{U}_k.
\end{aligned} \tag{1}
$$

We define the constraints in (10) as follows. Constraint C_1 assures the maximum number of devices multiplexed on each sub-channel. Constraint C_2 restricts the blocklength and satisfies the end-to-end transmission delay for a single communication link. Constraint C_3 ensures that given cooperative ABSs lie within the radius of the MBS, denoted as r_{max}. Constraint C_4 encompasses the minimum power requirements for various entities. Specifically, it represents the minimum power required by IoT device m to meet the minimum rate requirement r^m_{min}, the minimum power required by cooperative ABS u_h to satisfy the minimum rate requirement $r^{u_h}_{\text{min}}$, and the minimum power required for the link between cooperative ABS u_h and serving ABS u_u to fulfill the minimum rate requirement $r^{u_u}_{\text{min}}$. Constraint C_5 ensures that the power allocated to all the IoT devices within the same community should not be more than the total power $\bar{p}^k$ allocated to the given SC for that community. Constraint C_6 guarantees that the DEP of each user (i.e., IoT devices and ABSs) should not violate their threshold $\epsilon_{\text{threshold}}$. Constraint C_7 ensures that each device's achievable rate should be more than or equal to its minimum rate requirement. Constraint $C_8 - C_9$ ensures that the SIC decoding is done successfully.

Our objective function is a mixed-integer non-linear programming problem; therefore, it cannot be solved in polynomial time [25]. It results from the non-convexity of the non-convex normal approximation and the combinatorial constraint C_2. The problem (10) can be resolved by combining a penalty technique with monotonic optimization at a significant computational cost [29]. Alternatively, it can be resolved by leveraging the problem's decomposition property. Therefore, we use a common relaxation strategy to divide the maximization problem into two sub-problems [26]. The detailed explanation of the proposed solution is clearly explained in the following section.

IV. Proposed Solution

A. SC Allocation and Selection of Cooperative ABSs within the Coverage of MBS

To obtain the subsequent iterative solution of SC allocation, i.e., $(\boldsymbol{\phi})^{i+1}$, we solve the problem (10) with fixed values of $\left(\boldsymbol{b}^{(i)}, \boldsymbol{p}^{(i)}\right)$.

$$
\begin{aligned}
&\max_{\boldsymbol{\phi}} \ \min \quad \left(r^n_{\text{mbs},u_h}, r^n_{u_h,u_u}, r^k_{u_u,u_k} \ \right) \\
&\text{subject to} \quad C_1 - C_3 \text{ and } C_6 - C_7.
\end{aligned} \tag{11}
$$

1) SC Allocation using Stable Matching: A traditional way to compute the best solution to the above sub-problem is to exhaustively search for every potential combination of SCs and IoT devices. However, it is time-consuming and computationally expensive. Alternatively, we can reformulate it using matching theory with a low-complexity algorithm. The basic concepts are given below.

Definition 1 *(Two Way Matching):* The problem (11) is a two-way matching problem because a maximum of one IoT device should be allocated to SC from its priority order based on their rate[6] values. For better understanding, preference order introduced for given IoT device m with any two SCs (in its preference order) $j, j' \in (|\text{U}_h|+|\text{M}|) \times \text{N}$, $j, j' > |\text{U}_h|$, the two matchings τ and $\tau^{'}$ are defined as

$$
(j,\tau) \succ_m (j^{'},\tau^{'}) \Leftrightarrow r^n_{\text{mbs},m}(\tau) > r^n_{\text{mbs},m}(\tau^{'}), \tag{12}
$$

[5]We define $\overbrace{\sum_{m\in\mathcal{M}} b_{\text{mbs},m}}^{b_0}, \overbrace{\sum_{u_h\in\mathcal{U}_h} b_{\text{mbs},u_h}}^{b_1}, \overbrace{\sum_{u_u\in\mathcal{U}_u} b_{u_h,u_u}}^{b_2}, \overbrace{\sum_{l\in\mathcal{L}} b_{l_{k,[u_k]}}}^{b_3}$, where the notations b_0, b_1, b_2, and b_3 denote the sum of the blocklengths of IoT devices belonging to set $\mathcal{M}$, cooperative ABSs in set $\mathcal{U}_h$, serving ABSs in set $\mathcal{U}_u$, and IoT devices in community l assigned to SC set $\mathcal{K}$, respectively. It is important to note that all devices within the same community share the same blocklength.

[6]This rate is calculated based on initial power allocation, which is to be optimized later to achieve better rates.

which implies that if m^{th} IoT device achieves a higher rate than SC j', then device m prioritizes SC j in τ against SC j' in $\tau^{'}$. The terms swap matching and swap blocking pair are introduced and defined below to demonstrate the impact of externalities (peer effects).

Definition 2 *(Swap Matching):* Considering two IoT devices (m, m') and two SCs (j, j'), we denote the current matching state as $\tau(m) = j$ and $\tau(m') = j'$. A swap matching will be performed between (m, m') and (j, j') if m prefers SC j' over its current match, and SC j' also prefers m over its current match. We define the swap matching operation as follows:

$$\tau_m^{m'} = \Big\{\tau \backslash \big\{(m,j),(m',j')\big\} \cup \big\{(m,j'),(m',j)\big\}\Big\}, \quad (13)$$

where $\tau_m^{m'}$ represents the updated matching state, indicating that IoT device m is now matched with SC j', and vice versa. Therefore, we can define the swap-blocking pair as follows.

Definition 3 *(Swap Blocking Pair):* From the given matching state $\tau(m) = j$, $\tau(m') = j'$, an IoT pair of devices (m, m') is a swap pair if there exists

1) $\forall\ q \in \big\{m, m', j, j'\big\}, \tau_m^{m'}(q) \succeq_q \tau(q)$
2) $\exists\ q \in \big\{m, m', j, j'\big\}, \tau_m^{m'}(q) \succ_q \tau(q)$

where q shows the involved player (either SC or IoT device). It means that swap matching $\tau_m^{m'}$ is approved, and both IoT devices (m, m') can switch their SCs in τ by following these two conditions: 1) rate should not reduce after swapping and 2) rate of the at least one IoT device increases. The process continues until the not swap-blocking pair exists, resulting in a globally converged solution. However, if optimal matching is $\big\{(m,j),(m',j')\big\}$ and current matching is $\big\{(m,j'),(m',j)\big\}$, then the solution may not converge and stuck to a local optimum. The same procedure is adopted for the SCs allocation to each cooperative ABS.

2) Selection of the Best Cooperative ABSs: The selection of the best cooperative ABS from the set $\mathcal{U}_h$ (to relay the information to neighbor serving ABS) is based on the maximum achievable rate, which is given by

$$u_h^{opt} = \arg \max_{u_h \in \mathcal{U}_h} r_{\text{mbs},u_h}. \quad (14)$$

B. SC Allocation and Selection of IoT Devices within the Coverage of Serving ABS

The concept of dynamic programming is utilized to compute the SC's allocation and selection of IoT devices under multiplexing constraint C_1, power constraint C_5, and SIC constraints $C_8 - C_9$. The idea is to recursively compute three auxiliary vectors to keep the record of the current value of power, optimal solution, and backtracking, i.e., V, Q and T, respectively. Assuming $\bar{p}^k$ as fixed power budget for the SC k, if $s \in \mathcal{S} = \{1, 2, ..., |\mathrm{S}|\}$, $u_k \in \mathcal{U}_k = \{1, 2, ..., |\mathrm{U}_k|\}$ and $f \geq u_k$, we compute $V[s, u_k, f]$ as an optimal power value after satisfying the constraints as mentioned earlier. The recurrence relation is given by $V[s, u_k, f] = \max(v_a, v_b, v_c)$,

Algorithm 1 Computing optimal power on [0, $\bar{p}^k$]

1: **Input:** $u_k, \pi_k, (g^k_{u_u,[u_k]})_{u_k \in \mathcal{U}_k}, \bar{p}^k$
2: **Output:** p^{opt}
3: $s \leftarrow \pi_k\ (u_k)$, $t \leftarrow \pi_k\ (u_k - 1)$
4: **if** $u_k = 1$ or $\omega_{[s]} \geq \omega_{[t]}$ **then**
5: **return** $\bar{p}^k$
6: **else**
7: **return** $\max\Big(0, \min\big(\frac{\omega_{[t]} g^k_{u_u,[s]} - \omega_{[s]} g^k_{u_u,[t]}}{\omega_{[s]} - \omega_{[t]}}, \bar{p}^k\big)\Big)$
8: **end if**

where (v_a, v_b, v_c) represents power allocations. These are defined as follows.

$$\begin{cases} v_a = V[s, u_k, f] \\ v_b = \begin{cases} V[s-1, u_k - 1, u_k - 1] + \bar{A} - \bar{B}, \\ \text{if } 0 < p^{opt} < Q[s-1, u_k - 1, u_k - 1] \\ 0, \text{ otherwise} \end{cases} \\ v_c = V[s, u_k - 1, f]. \end{cases} \quad (15)$$

The variable $\bar{A} = \sum_{f \in u_k} r^k_{u_u,f}(p^{opt})$ and $\bar{B} = \sum_{f \in u_k} r^k_{u_u,f}(0)$. The pseudocode for computing the optimal power p^{opt} within the range of $[0, \bar{p}^k]$ is provided in Algorithm (1). The algorithm first assigns the variables s and t with the values of $\pi_k(u_k)$ and $\pi_k(u_k - 1)$, respectively. It then checks whether $u_k = 1$ or if the value of $\omega[s]$ is greater than or equal to $\omega[t]$. If either condition is true, it returns the value of $\bar{p}^k$ as the optimal power (line 4). Otherwise, the optimal power is computed using the formula specified in line 7. Algorithm (1) performs a fixed number of basic operations; therefore, its complexity is $\mathcal{O}(1)$.

C. Joint Blocklength and Power Optimization

For clarity, the sub-problem (11) can be explicitly articulated by sub-problems (16) and (17), which implies that the power allocation and blocklength allocation are done solely. To obtain the next best value of $\boldsymbol{b}^{(i+1)}$, we first solve the sub-problem (16) with fixed values of $\Big(\boldsymbol{\phi}^{(i+1)}, \boldsymbol{p}^{(i)}\Big)$.

$$\begin{aligned} \max_{\mathbf{b}} \quad & \min \quad \big(r^n_{\text{mbs},u_h}, r^n_{u_h,u_u}, r^k_{u_u,u_k}\ \big) \\ \text{subject to} \quad & C_1 - C_9. \end{aligned} \quad (16)$$

In relation to blocklength constraint C_2, the bisection-based optimal value of blocklength is computed to minimize the complexity of the proposed solution. We assume $b^k_{\text{lb}} = 1, b^k_{\text{ub}} = M_{\max} - B$, where B is a fixed value calculated as $B = b1 + b_2$. Subsequently, the optimal value of blocklength $b_{l_{k,[u_k]}}$ is computed, defined as $b_{opt} = \arg \max_{\left\{\lfloor b^k_{\text{mid}} \rfloor, b^k_{\text{mid}}\right\}} \big(r^k_{u_u,u_k}\big)$, which is upper bounded by a threshold value $\bar{\xi}$. Initially, we set the initial value of $b^k_{\text{mid}} = \frac{\left(b^k_{\text{ub}} + b^k_{\text{lb}}\right)}{2}$ and then update the value of $b^k_{\text{ub}} = b^k_{\text{mid}}$ if $r^k_{u_u,u_k}(b_{opt})|_{b_{opt} = b^k_{\text{mid}}} > \bar{\xi}$. Otherwise, we set $b^k_{\text{lb}} = b^k_{\text{mid}}$. This process continues until $b^k_{\text{ub}} - b^k_{\text{lb}} > \bar{\sigma}$ is achieved. The complexity of these steps is $\mathcal{O}(\log_2(M_{\max}/\bar{\sigma}))$, where $\bar{\sigma} = 0.01$. Subsequent analysis reveals that the worst-case computing complexity of the exhaustive search method

is $\mathcal{O}(\mathcal{K}^3)$, which is significantly higher compared to our proposed steps.

Afterwards, the problem (17) is solved with the fixed values of $\left(\boldsymbol{\phi}^{(i+1)}, \boldsymbol{b}^{(i+1)}\right)$ to determine the next best value of $\boldsymbol{p}^{(i+1)}$.

$$\begin{aligned} \max_{\mathbf{p}} \quad & \min \quad \left(r^n_{\text{mbs},u_h}, r^n_{u_h,u_u}, r^k_{u_u,u_k}\right) \\ \text{subject to} \quad & C_1 - C_9. \end{aligned} \tag{17}$$

To solve the above-mentioned sub-problem, we compute the minimum power required by the given device ϑ on the given SC n to achieve its minimum rate requirement under constraints C_4 and C_7, where $\vartheta \in \{m, u_h\}$ like that in [30]. We define the overall minimum power required to the given device as $p^{\min}_{\text{mbs},\vartheta} = \sum_{n\in\mathcal{N}} p^{\min,n}_{\text{mbs},\vartheta}$. We compute the minimum power on SC n as below.

$$p^{\min,n}_{\text{mbs},\vartheta} = \left(\mu_\vartheta - \frac{1}{g^n_{\text{mbs},\vartheta}}\right)^+, \forall\ n \in \mathcal{N},\ \vartheta \in \{m, u_h\}, \tag{18}$$

$$\sum_{n\in\left\{n\in N | p^{\min,n}_{\text{mbs},\vartheta}>0\right\}} w_n \log_2\left(\mu_\vartheta g^n_{\text{mbs},\vartheta}\right) = r^\vartheta_{\min},\ \vartheta \in \{m, u_h\}, \tag{19}$$

where $(x)^+$ represents $\max(x, 0)$, μ and μ_ϑ are the intermediate variables. The optimal powers for the device such as IoT device m and cooperative AB u_h on given SC can be computed using the water-filling algorithm [30].

$$p^n_{\text{mbs},\vartheta} = p^{\min,n}_{\text{mbs},\vartheta} + \left(\mu - \frac{1}{g^n_{\text{mbs},\vartheta}} - p^{\min,n}_{\text{mbs},\vartheta}\right)^+,\ \vartheta \in \{m, u_h\}, \tag{20}$$

$$\sum_{\vartheta\in\vartheta} \sum_{n\in\left\{n\in N | p^n_{\text{mbs},\vartheta} > p^{\min,n}_{\text{mbs},\vartheta}\right\}} \left(\mu - \frac{1}{g^n_{\text{mbs},\vartheta}} - p^{\min,n}_{\text{mbs},\vartheta}\right) = P_{\text{mbs}} - \sum_{\vartheta\in\vartheta}\sum_{n\in N} p^{\min,n}_{\text{mbs},\vartheta}, \tag{21}$$

where, P_{mbs} is the sum of all the powers allocated to the given SCs defined as $\sum_{n\in\mathcal{N}} \psi^n_{\text{mbs},\vartheta} p^n_{\text{mbs},\vartheta} = P_{\text{mbs}}$, where $\psi^n_{\text{mbs},\vartheta}$ is a binary indicator for SC allocation. So $\psi^n_{\text{mbs},\vartheta} = 1$ if the given SC is allocated to the device ϑ; otherwise $\psi^n_{\text{mbs},\vartheta} = 0$. We guarantee adherence to the minimum QoS criteria, ensuring that every communication link satisfies its specific minimum rate requirement Therefore, the received SINR of the IoT device m from the MBS should be greater than or equivalent to its minimum SINR threshold $\varrho^{\min,n}_{\text{mbs},m}$ for the following link. It is given by

$$\left(\frac{\psi^n_{\text{mbs},m} p^n_{\text{mbs},m} h^n_{\text{mbs},m}}{\delta^2 + p^n_{u_h,u_u} h^n_{u_h,m}}\right) \geq \varrho^{\min,n}_{\text{mbs},m}. \tag{22}$$

Hence, the total achievable rate of the IoT device m computed across all the allocated SCs should be greater than or equal to $r^m_{\min}$. The maximum power allocated to the cooperative to the serving ABS communication link must also be restricted to achieve the minimum QoS criteria for the IoT devices within the coverage of MBS. Hence, the power allocated to the link between the cooperative to the serving ABS should be subject to the following constraints.

$$p^{\min,n}_{u_h,u_u} \leq p^n_{u_h,u_u} \leq \left(\frac{p^n_{\text{mbs},m} h^n_{\text{mbs},m}}{\varrho^{\min,n}_{\text{mbs},m} h^n_{u_h,m}} - \frac{\delta^2}{h^n_{u_h,m}}\right) \leq p^{\max}_{u_u}, \tag{23}$$

The allocated power for the link between the given cooperative and serving ABS should also meet its minimum QoS requirement as given below.

$$\left(\frac{\psi^n_{\text{mbs},m} p^{\min,n}_{u_h,u_u} g^n_{u_h,u_u} d^{-\alpha_{\text{mbs}}}_{u_h,u_u}}{\delta^2 + \sum_{m\in\mathcal{M}} \psi^n_{\text{mbs},m} p^n_{\text{mbs},m} h^n_{u_u,m}}\right) \geq \varrho^{\min,n}_{u_h,u_u}, \tag{24}$$

$$p^{\min,n}_{u_h,u_u} \leq \frac{\left(\delta^2 + \sum_{m\in\mathcal{M}} \psi^n_{\text{mbs},m} p^n_{\text{mbs},m} h^n_{u_u,m}\right) \varrho^{\min,n}_{u_h,u_u}}{g^n_{u_h,u_u} d^{-\alpha_{\text{mbs}}}_{u_h,u_u}}, \tag{25}$$

where $\varrho^{n,opt}_{u_h,u_u}$ can be computed by setting $p^n_{u_h,u_u} = p^{n,opt}_{u_h,u_u}$.

$$p^{n,opt}_{u_h,u_u} = \begin{cases} 0, & \text{if } p^{\min,n}_{u_h,u_u} > p^n_{u_h,u_u} \\ p^n_{u_h,u_u}, & \text{if } p^{\min,n}_{u_h,u_u} < p^n_{u_h,u_u} \\ p^{\max}_{u_u}, & \text{if } p^{\max}_{u_u} \in \left[p^{\min,n}_{u_h,u_u}, p^n_{u_h,u_u}\right] \\ \min\left(p^n_{u_h,u_u}, \max\left(p^{\max}_{u_u}, p^{\min,n}_{u_h,u_u}\right)\right), & \text{Otherwise.} \end{cases} \tag{26}$$

Thus, the total achievable rate for the link between cooperative ABS to serving ABS computed across all the allocated SCs should be greater than or equal to $r^{u_u}_{\min}$. Basically, the idea is to divide the minimum rate requirement for each device across all the allocated SCs to ensure that the total power allocated across all the allocated SCs to the given device should result in a rate better than the minimum rate requirement for that device [7]. The overall minimum power required for the following link is given below.

$$p^{\min}_{u_h,u_u} = \sum_{n\in\mathcal{N}} p^{\min,n}_{u_h,u_u}. \tag{27}$$

Relevant to constraint C_5, we distribute the power to the given number of IoT devices within the serving ABS allocated to SC k within the same community. It is worth mentioning that the sum of the allocated powers to each IoT device within a community assigned to a SC k must be less than or equal to $\bar{p}^k$. It is given by

$$\bar{p}^k \geq \sum_{u_k\in\mathcal{U}_k} \propto^k_{u_u,u_k} p^k_{u_u,[u_k]}, \forall\ k \in \mathcal{K} \text{ and } \forall\ u_u \in \mathcal{U}_u. \tag{28}$$

The feasible set containing the feasible powers for these devices is given as

$$\mathcal{R} = \left\{\bar{P} : \sum_{k\in\mathcal{K}} \bar{p}^k \leq p^{\max}_{u_u} \text{ and } 0 \leq \bar{p}^k \leq p^k, \forall\ k \in \mathcal{K}\right\}, \tag{29}$$

where the set $\mathcal{R}$ can also be expressed as Cartesian's product of all the user's feasible sets, and p^k represents the power limit to the given SC. determines the optimal value at line 4 by employing a for loop for each IoT device. If constraint C_5 is satisfied, then $p^k_{u_u,u_k[u_k]} = p^{opt}$. Otherwise,

[7]We consider the SC allocated to the device with the lowest SINR to calculate the minimum power requirement per SC for that device.

Algorithm 2 Power Distribution on Given IoT Devices

1: **Input:** $(\mathcal{U}_k)_{u_k\in\mathcal{U}_k}, (\mathcal{U}_k)_{u_k\in\mathcal{U}_k}, (g^k_{u_u,[u_k]})_{u_k\in\mathcal{U}_k}, \bar{p}^k$
2: **Output:** $p^k_{u_u,[1]}, p^k_{u_u,[2]}, ..., p^k_{u_u,[\mathrm{U}_k]}$
3: **for** $u_k \in |\mathrm{U}_k|$ **do**
4: $p^{opt} \leftarrow$ OptimalPower$(u_k, u_k, w_c, g^k_{u_u,[u_k]}, \bar{p}^k)$
5: $i \leftarrow (u_k - 1)$
6: **while** $i \geq 1$ $p^k_{u_u,[i]} < p^{opt}$ **do**
7: $p^{opt} \leftarrow$ OptimalPower$(i, u_k, w_c, g^k_{u_u,[u_k]}, \bar{p}^k)$
8: $i \leftarrow (i - 1)$
9: **end while**
10: $p^k_{u_u,[i+1]}, ..., p^k_{u_u,[u_k]} \leftarrow p^{opt}$
11: **end for**

the algorithm backtracks and finds the highest index i such that $p^k_{u_u,u_k[i]} \geq p^{opt}$. In this way, the optimal vector containing the power values for each IoT device can be retrieved, i.e., $p^k_{u_u,[i+1]}, ..., p^k_{u_u,[u_k]} \leftarrow p^{opt}$ in line 10. Consequently, the complexity of the algorithm is $\mathcal{O}(S^2)$. However, if the optimal power is computed over D different power budgets, the complexity will be $\mathcal{O}(S^2 + DS)$.

D. Proposed BLCA Algorithm

Algorithm (3) is designed to perform SC allocation utilizing matching theory with fixed values of power and blocklength in line 5. The best cooperative ABSs are selected based on the derived results in line 6. To determine the optimal blocklengths for the subsequent iteration, a bisection search is conducted within the specified range, as indicated in line 7. Subsequently, the available power is distributed using water-filling, while adhering to the power constraints C_4 and C_5, to meet the minimum QoS requirements, as stated in line 9. In line 10, the power is allocated to the links between cooperative ABSs and serving ABSs using equations (22) to (26).

The power distribution is achieved by optimizing the power budget on each SC through dynamic and geometric programming. This process involves a two-tier projected gradient descent-based algorithm that distributes the power among the devices. The algorithm iterates for each SC to optimize the power budget until the condition $\|\bar{P}' - \bar{P}\|_2^2 \leq \lambda$ is satisfied, as described in lines 11-19. The search direction in line 14 is computed using the exact gradient method, and the step size is determined by backtracking using the exact line search method. In line 16, the projection of $\bar{P}$ onto the feasible set $\mathcal{R}$ is calculated, as presented in [31]. The power distribution among the devices within a given SC is performed from line 17 to line 19, while the rates are computed in line 20. The formulated problem is solved iteratively until $i > t_{\max}$. The proposed algorithm converges within $\mathcal{O}(\log_2(1/\epsilon))$ iterations, where ϵ represents the error tolerance.

V. Results and Discussion

A comparative analysis is conducted to evaluate the BLCA (blocklength-constrained algorithm). Three power allocation use cases are also analyzed, i.e., minimum power allocation, where each IoT device satisfies its minimum rate, and dynamic power allocation, where low-priority IoT devices first fulfill their minimum rate requirements compared to high-priority IoT devices. Any remaining power is then distributed optimally among the high-priority IoT devices.

Algorithm 3 Proposed BLCA Algorithm with Optimal Resource Allocation

1: **Input:** σ,$(\mathcal{U}_k)_{u_k\in\mathcal{U}_k}$,$(\mathcal{N})_{n\in\mathcal{N}}$,$(\mathcal{K})_{k\in\mathcal{K}}$,$\lambda$, $|\mathrm{S}|$, $p^{\max}_{u_u}$, recursive index $i = 1$, highest amount of iterations possible $t_{\max}$, and randomly choose feasible values p_c, $\boldsymbol{\phi}^{(0)}$, $\boldsymbol{b}^{(0)}$ and $\boldsymbol{p}^{(0)}$.
2: **Output:** $\boldsymbol{p}^*, \boldsymbol{b}^*, \boldsymbol{\phi}^*$
3: Suppose the starting point $\bar{P} = 0$
4: **while** Convergence or $i > t_{\max}$ **do**
5: Solve (11) for fixed $\left(\boldsymbol{b}^{(i)}, \boldsymbol{p}^{(i)}\right)$ to find $(\boldsymbol{\phi})^{i+1}$
6: Selection of best cooperative ABSs (14)
7: Solve (16) using bisection-based algorithmic steps
8: Solve (17) for fixed $\left(\boldsymbol{\phi}^{(i+1)}, \boldsymbol{b}^{(i+1)}\right)$ to find $\boldsymbol{p}^{(i+1)}$
9: Power distribution using water-filling $(18-21)$
10: Power allocation for cooperative to the serving ABS communication link $(22-26)$
11: Power distribution to IoT devices on the SC k within the same community $(28-29)$
12: **while** $\|\bar{P}' - \bar{P}\|_2^2 \leq \lambda$ **do**
13: $\bar{P}' \leftarrow \bar{P}$ saving previous power vector
14: $\Delta = \Delta \sum_{u_k\in\mathcal{U}_k} r^k_{u_u,u_k}(\bar{p}^k)$ and update step size σ
15: $\bar{P} =$ Projection of $\bar{P} + \left(\sigma\Delta \text{ on } \mathcal{R}\right)$
16: **end while**
17: **for** $k \in \mathcal{K}$ **do**
18: Allocate power to the IoT device u_k by algorithm (2)
19: **end for**
20: Compute rates using (1), (3) and (4)
21: Set $i : i + 1$
22: **end while**

A. Simulations Setup

We configured the MBS to transmit at a power of 40 watts with a coverage radius of 500 meters [32]. Within this setup, we deployed a total of $\mathcal{M} = 5$ IoT devices at a minimum distance of 30 meters, $\mathcal{U}_h = 5$ cooperative ABSs at a minimum distance of 350 meters, and $\mathcal{U}_u = 2$ serving ABSs at a minimum distance of 80 meters [33]. The serving ABSs contain $\mathcal{I}_q = 9$ and $\mathcal{I}_q = 11$ IoT devices, respectively. The circular coverage area of the serving ABSs has a radius of 50 meters, and their maximum transmit power is limited to 1 watt [34]. The cooperative ABS is approximately 100 meters away from the serving ABS. We consider a maximum of $\mathcal{C} = 20$ SCs. The maximum delay considered for the analysis is set at 1 millisecond. Additionally, we assume the minimum time required to convert one unit of blocklength to be 0.01 milliseconds [4]. Unless specifically stated otherwise, we assume the following parameter values: the threshold for the DEP is $\epsilon_{\text{threshold}} = 10^{-5}$, the path loss exponent is $\alpha_{\text{mbs}} = 2$, the speed of light is $c = 3 \times 10^8$ meters per second, the circuit power is 10 watts, and the noise power density is $\delta^2 = -174$ dBm/Hz [4].

TABLE I
SIMULATION PARAMETERS.

Parameters	Values
Altitude of the serving ABS and cooperative ABS (z_{u_u}, z_{u_h}) [4]	(50, 50) Meters
Coefficients for LoS and N-LoS $(\eta_{los}, \eta_{nlos})$	(1dB, 20dB)
Density and height of building $(a,\ b)$ [35]	(12, 0.135)
Minimum rate requirement for each device $r_{\min}$ [34]	2 bits/s/Hz
Number of cooperative and serving ABSs $(\mathcal{U}_h, \mathcal{U}_u)$ [33]	(5, 2)
Number of IoT devices within the coverage of cooperative ABS p and q $(\mathcal{I}_p, \mathcal{I}_q)$	(9,11)
Number of SCs and IoT devices within the coverage of MBS $(\mathcal{C}, \mathcal{M})$ [25]	(20, 5)
Noise power density [4]	-174 dBm/Hz
Power of the serving ABS $p_{u_u}^{\max}$ [34]	1 Watt
Power of the MBS P_{mbs} [32]	40 Watts
Pathloss exponent (α_{mbs}) [26]	2
Radius of the MBS $r_{\max}$ [4]	500 Meters
Time needed to convey one unit of block-length T_{block} [4]	0.01 Millisecond
Transmission delay $D_{\max}$ [4]	1 Millisecond

The altitude of both serving and cooperative ABSs is set to $z_{u_u} = 50$ meters and $z_{u_h} = 50$ meters, respectively [26]. The attenuation for LoS and N-LoS connection is assumed to be 1 dB and 20 dB, respectively. The channel parameters, including the building density and height, are $a = 12$ and $b = 0.135$, respectively [35]. The minimum rate requirement for each link is set as $r_{\min} = r_{\min}^{m} = r_{\min}^{u_h} = r_{\min}^{u_u} = r_{\min}^{[u_k]} = 2$ bits/s/Hz [34]. We utilize the radio propagation channel model provided in [1]. For simplicity, we assume the sum of the blocklengths of each communication link within each hop is equal to the blocklength of individual links. We compare our proposed scheme BLCA (Blocklength constrained algorithm) under two multiple access techniques, i.e., OMA and NOMA named as BLCA-OMA and BLCA-NOMA, respectively. Additionally, we investigate them under two distinct scenarios of finite blocklength, i.e., fixed and random blocklength approaches. In the fixed blocklength approach, we select a fixed value of $b_{opt} \in \big[1, 2, ..., (M_{\max} - B)\big]$. In the random blocklength approach, we randomly select a value of $b_{opt} \in \big[1, 2, ..., (M_{\max} - B)\big]$. The simulation parameters are summarized in Table I.

B. Performance Comparison

In Fig. 3, we analyze the impact of heterogeneous delay on the time required to transmit a unit blocklength on the system sum-rate. We compare the proposed BLCA algorithm with two baseline resource allocation schemes, namely random matching [23] and WSA matching [24]. We observe that the achievable rate increases with an increase in $M_{\max}$ because it depends on the maximum transmission delay. Moreover, the proposed scheme demonstrates superior performance over benchmark schemes and the performance gap between the proposed scheme and the WSA matching and random matching schemes widens as the value of $M_{\max}$ increases. The enhanced throughput in the proposed scheme can be attributed to the significant improvement in both channel qualities and achievable SNR per SC achieved through stable matching. In contrast,

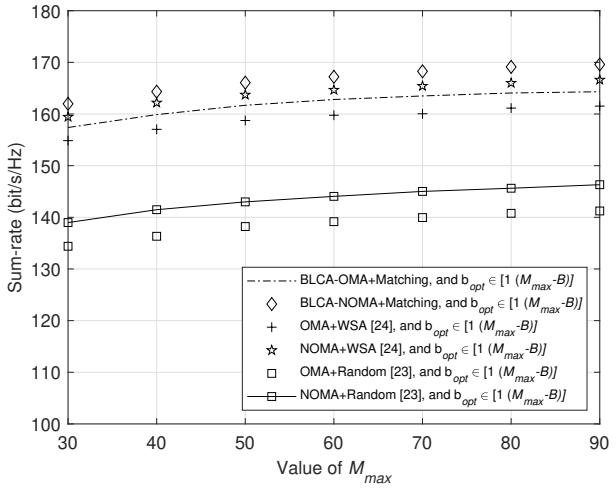

Fig. 3. Sum-rate versus $M_{\max}$ with optimal blocklength and power allocation, where $b_1 = k \times b_2$, $b_2 = 3$, $M_{\max} = k \times 10$, and $k = \{3, 4, ..., 9\}$.

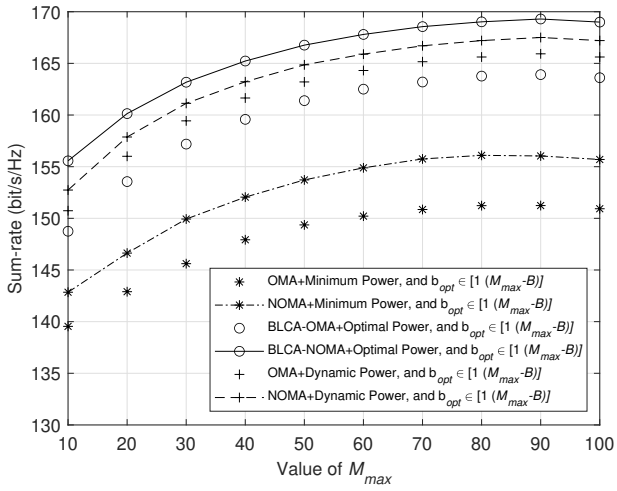

Fig. 4. Sum-rate versus $M_{\max}$ using matching-based SC allocation and optimal blocklength allocation with different power allocation strategies, where $b_1 = k \times b_2$, $b_2 = 3$, $M_{\max} = k \times 10$, and $k = \{1, 2, ..., 10\}$.

the random matching approach [23] involves devices selecting SCs randomly, potentially resulting in SC assignments with inferior channel qualities. Similarly, the WSA scheme [24] may assign SCs to devices that have lower channel qualities. The effectiveness of the proposed scheme is evaluated using two different multiplexing techniques: OMA with $(|\mathrm{S}| = 1)$ and NOMA with $(|\mathrm{S}| = 2)$. The superior performance of the NOMA scheme can be attributed to its fundamental principles, such as superposition coding at the transmitter (multiplexing two IoT devices per SC) and SIC at the receiver (demultiplexing based on power levels). Simulation results demonstrate that initially, when $M_{\max}$ is set to 30, the system sum-rate is relatively low. However, it gradually increases to 169.59 bits/s/Hz, representing a 3.21% improvement (for BLCA-NOMA), after which it remains relatively constant.

Fig. 4 illustrates the trade-off between heterogeneous delays over blocklength and different power allocation techniques

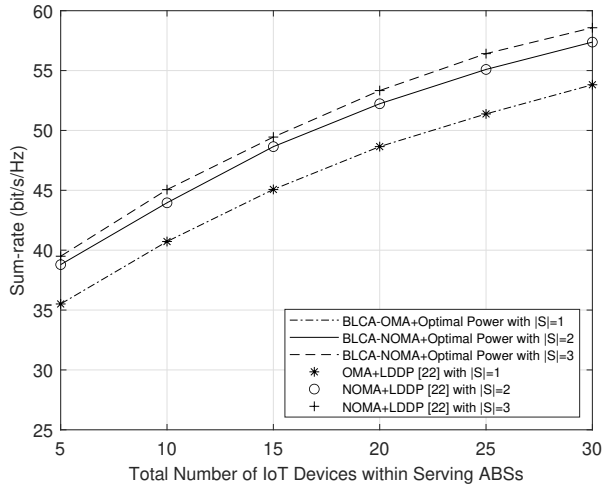

Fig. 5. Sum-rate versus different number of IoT devices with matching-based SC allocation against LDDP [22], where $M_{\max} = 100$.

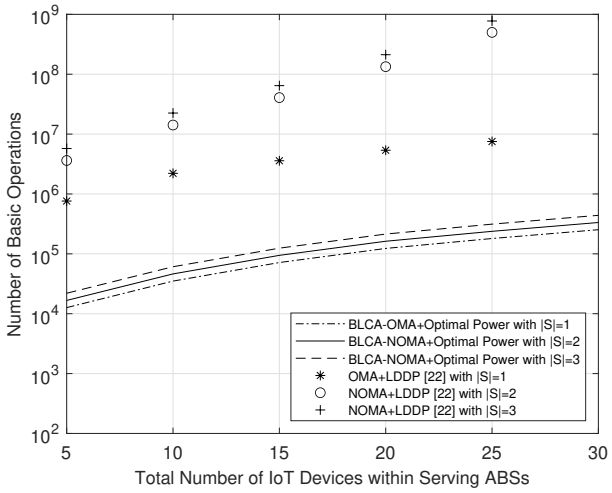

Fig. 6. Complexity of the BLCA scheme and LDDP [22] with stable matching-based SC allocation, considering the basic number of operations, where $M_{\max} = 100$.

for both OMA and NOMA systems. The following observations can be made: 1) the system throughput increases with an increase in the value of $M_{\max}$ for all power allocation approaches, and 2) the proposed scheme outperforms both other power allocation techniques (OMA and NOMA with minimum power and dynamic power allocation) for both OMA and NOMA systems. The effectiveness of the NOMA system is significantly higher than that of the OMA system. This increase in effectiveness can be attributed to superposition coding at the transmitting node and SIC at the receiving node in conventional NOMA. The results indicate that the sum-rate for NOMA with optimal power is 4.58% higher than that for legacy OMA with optimal power and 2.68% higher than that for NOMA with dynamic power allocation.

The impact of an increasing number of IoT devices on the sum-rate is analyzed in Fig. 5. The proposed BLCA algorithm is compared against the near-optimal high-complexity

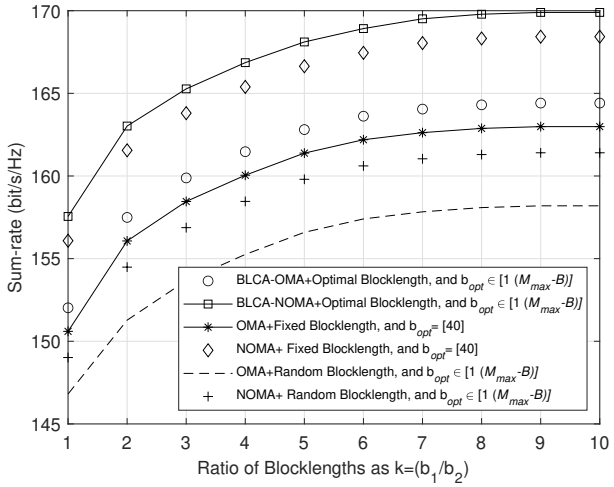

Fig. 7. Sum-rate versus the ratio of blocklengths with stable matching based SC allocation and optimal power allocation, where $b_1 = k \times b_2$, $b_2 = 3$, and $M_{\max} = 100$.

benchmark scheme, LDDP [22]. Both schemes are simulated by varying the number of IoT devices in the vicinity of each serving ABS from 5 to 30 due to high computational run-time. As expected, the following observations are made: the throughput increases with an increase in the number of IoT devices, and greater device participation further elevates system throughput. The throughput gain of NOMA (with two and three IoT devices multiplexed per SC) is greater than OMA (with one IoT device per SC). There is a constant gap between both NOMA with $|\mathrm{S}|$=2 and $|\mathrm{S}|$=3. Furthermore, the performance gain of both BLCA and LDDP is almost the same for any number of IoT devices, indicating that the proposed BLCA algorithm is near-optimal. It is worth noting that the proposed BLCA algorithm runs within seconds on a computer with specifications such as a core i5, 8th generation for $\mathcal{I} \leq 30$. In contrast, LDDP [22] requires 1600 times more operations for $\mathcal{I}$=20 and $|\mathrm{S}|$=2 (as shown in Fig. 6), validating its low computational cost towards an optimal solution.

Fig. 7 shows the impact of the ratio of blocklengths, $k = (b_1/b_2)$, on the achievable system sum-rate. The results demonstrate that an increase in the ratio of blocklengths corresponds to a higher system throughput. This effect is because the degree of freedom to transmit data packets depends mainly on the blocklength. Consequently, greater blocklength values lead to enhanced system sum rates. The proposed BLCA algorithm employing an optimal blocklength consistently yields better results compared to scenarios involving fixed or random blocklengths, which emphasizes the importance of efficient blocklength allocation to maximize performance. When combining the advantages of the NOMA scheme with optimal blocklength, it emerges as the optimal choice, surpassing NOMA with fixed or random blocklengths. Hence, the NOMA scheme outperforms OMA due to its efficient utilization of spectrum resources, thereby accommodating multiple devices within resource constraints. It is important to note that NOMA with optimal blocklength surpasses OMA with optimal block-

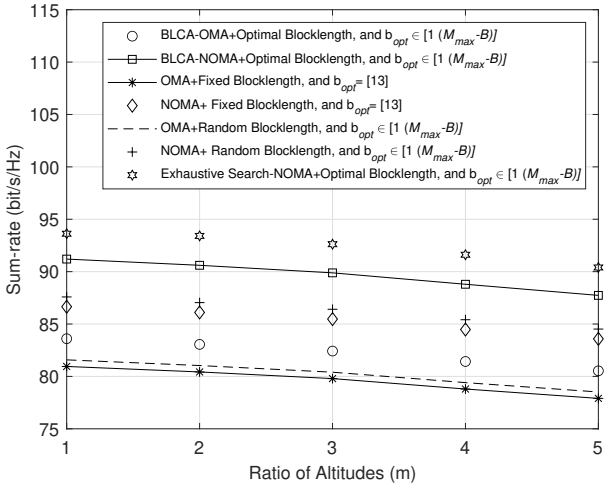

Fig. 8. Sum-rate versus altitudes using stable matching based SC allocation and optimal power allocation, where $b_1 = 13$, $b_2 = 13$, $z_{u_u} = 50$, $z_{u_h} = \{50, 100, ..., 250\}$ and $M_{\max} = 100$.

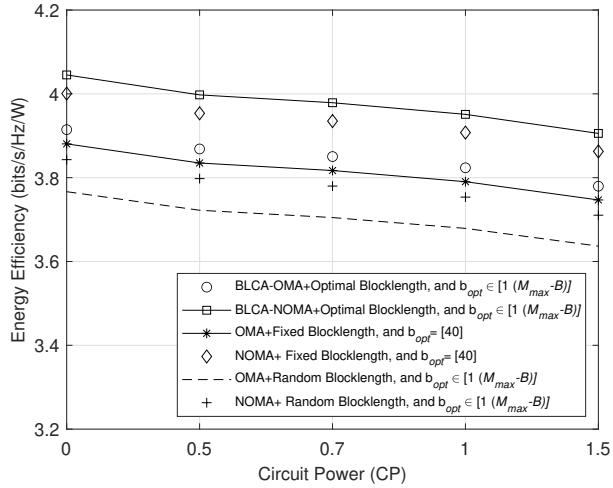

Fig. 9. Spectral efficiency versus CP using stable matching based SC allocation and optimal power allocation, where $b_1 = 13$, $b_2 = 13$, and $M_{\max} = 100$.

length. Similarly, NOMA with fixed or random blocklengths outperforms their respective OMA counterparts in their corresponding scenarios. The results validate that the throughput of BLCA-NOMA with optimal blocklength is 3.63% higher than that of BLCA-OMA with optimal blocklength, across all scenarios.

Fig. 8 investigates the impact of ratio of altitudes, i.e., $H = (z_{u_h}/z_{u_u})$ on the achievable system sum-rate. Our assumption posits that the cooperative ABSs are strategically positioned at higher altitudes compared to the serving ABSs. Notably, the achievable rate of the proposed scheme decreases by increasing the altitude, which is due to higher channel fading and increased LoS interference. Regardless of the considered blocklength scenario whether it is optimal, random, or fixed, the NOMA scheme consistently outperforms OMA. It is evident from the fact that the NOMA curve maintains a higher position than the OMA curve across all cases. In addition, the scheme employing an optimal blocklength consistently yields better results when compared to scenarios involving fixed or random blocklengths. NOMA with fixed or random blocklengths outperforms their respective OMA counterparts within their respective scenarios. Simulation results solidify the observation that the sum-rate for BLCA-NOMA, employing optimal blocklength allocation, exceeds that of BLCA-OMA with optimal blocklength allocation by a margin of 9.09%. It is noteworthy that the curve for NOMA with a fixed blocklength allocation is lower than that for NOMA with a random blocklength allocation. This difference arises from our choice of a higher random blocklength value compared to the fixed blocklength. Monte-Carlo simulations are also conducted to compute the best possible solution. To validate our plotted curves, we have included an upper-bound solution curve (computed using the exhaustive search method, also known as the brute force method). This curve demonstrates the proximity of our proposed solution to the optimal one.

Fig. 9 evaluates the current energy efficiency values of the proposed solution by analyzing the total energy efficiency of the system against CP. In our evaluation, we emphasize the distinction between the BLCA scheme under two distinct multiple-access techniques: OMA and NOMA. We evaluate the efficacy of our proposed algorithm under varying blocklength scenarios. Simulation and results illustrate that increasing the value of CP results in a decrease in the total energy efficiency of the system. It is important to note that NOMA with optimal blocklength surpasses OMA with optimal blocklength within their respective scenarios. Similarly, NOMA with fixed or random blocklengths outperforms their respective OMA counterparts in their corresponding scenarios. Comparative analysis shows that the proposed BLCA-NOMA achieves a 5.25% (resp. BLCA-OMA 3.39%) improvement in energy efficiency for NOMA with random blocklength and a 1.12% improvement for NOMA with fixed blocklength (resp. OMA 0.87%). The fundamental reason behind this minimal increase is the selection of a fixed blocklength value closer to its optimal value.

VI. Conclusion and Future Work

This study explores a mixed-integer non-linear programming problem for optimizing joint resource allocation in an integrated aerial-terrestrial wireless network to maximize the system sum-rate. A novel BLCA (blocklength constrained) algorithm is proposed, which utilizes alternating optimization and a two-step projected gradient descent-based strategy to optimize the resource allocation policy while considering delay, reliability, and QoS constraints through dynamic and geometric programming. We compared the proposed algorithm with different benchmark algorithms that avoid allocating the worst SC to transmitting devices under various techniques. The study concludes that NOMA with optimal blocklength surpasses OMA with optimal blocklength and NOMA with fixed or random blocklengths outperforms their respective OMA counterparts in their corresponding scenarios. Simulation results demonstrate the efficacy of the proposed algorithm, which requires 1600 times less computational cost than

baseline approaches. Future work will explore the concept of digital twins to further improve the system.

ACKNOWLEDGMENTS

This should be a simple paragraph before the References to thank those individuals and institutions who have supported your work on this article.

REFERENCES

[1] M. Awais, H. Pervaiz, M. Jamshed, W. Yu, and Q. Ni, "Enhancing URLLC in Integrated Aerial Terrestrial Networks: Design Insights and Performance Trade-offs," in *IEEE WowMom 2022*, Belfast, Uk, Jun 14-17, 2022.

[2] P. Rahimi, C. Chrysostomou, H. Pervaiz, V. Vassiliou, and Q. Ni, "Joint Radio Resource Allocation and Beam forming Optimization for Industrial Internet of Things in Software-Defined Networking-Based Virtual Fog-Radio Access Network 5G-and-Beyond Wireless Environments," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 6, pp. 4198-4209, 2022.

[3] W. Gao, Z. Zhao, G. Min, Q. Ni, and Y. Jiang, "Resource Allocation for Latency-Aware Federated Learning in Industrial Internet of Things," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 12, pp. 8505-8513, 2021.

[4] Q. Wu, J. Xu, Y. Zeng, D.W.K. Ng, N. Al-Dhahir, R. Schober, and A.L. Swindlehurst, "A Comprehensive Overview on 5G-and-Beyond Networks With UAVs: From Communications to Sensing and Intelligence," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 10, pp. 2912-2945, 2021.

[5] A. Bhowmick, S. Dhar Roy, and S. Kundu, "Throughput Maximization of a UAV Assisted CR Network With NOMA-Based Communication and Energy-Harvesting," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 1, pp. 362-374, 2022.

[6] W. New, C.Y. Leow, K. Navaie, and Z. Ding, "Aerial-Terrestrial Network NOMA for Cellular-Connected UAVs," *IEEE Transactions on Vehicular Technology*, pp.1-1, 2022.

[7] L. Qian, Y. Wu, F. Jiang, N. Yu, W. Lu, and B. Lin, "NOMA Assisted Multi-Task Multi-Access Mobile Edge Computing via Deep Reinforcement Learning for Industrial Internet of Things," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 8, pp. 5688-5698, 2021.

[8] F. Salehi, N. Neda, M.H. Majidi, and H. Ahmadi, "Cooperative NOMA-Based User Pairing for URLLC: A Max–Min Fairness Approach," *IEEE Systems Journal*, pp.1-11, 2021.

[9] M. Amjad, L. Musavian, and S. Aissa, "NOMA Versus OMA in Finite Blocklength Regime: Link-Layer Rate Performance," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 12, pp. 16253-16257, 2020.

[10] H. Lei, C. Zhu, K. Park, W. Lei, I. Ansari, and T. Tsiftsis, "On Secure NOMA-Based Terrestrial and Aerial IoT Systems," *IEEE IoTs Journal*, vol. 9, no. 7, pp. 5329-5343, 2022.

[11] L. Haiquan, Y. Zeng, S. Jin, and R. Zhang. "Enabling panoramic full-angle reflection via aerial intelligent reflecting surface," *In 2020 IEEE ICC Workshops*, pp. 1-6. IEEE, 2020.

[12] D. Van Huynh, T. Do-Duy, L. D. Nguyen, M.-T. Le, N.-S. Vo, and T. Q. Duong, "Real-time optimized path planning and energy consumption for data collection in unmanned Ariel Vehicles-aided intelligent wireless sensing," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 4, pp. 2753–2761, 2022.

[13] Z. Kang, C. You, and R. Zhang, "Learning for multi-UAV relaying: A Gibbs sampling approach," *IEEE ICC,* 2020.

[14] J. Wang, M. Liu, J. Sun, G. Gui, H. Sari, and F. Adachi, "Multiple unmanned-aerial-vehicles deployments and user pairing for Non-orthogonal Multiple Access Schemes," *IEEE IoTs Journal*, vol. 8, no. 3, pp. 1883–1895, 2021.

[15] S. Arzykulov, A. Celik, G. Nauryzbayev, and A. Eltawil, "UAV-Assisted Cooperative & amp; Cognitive NOMA: Clustering, and Resource Allocation," *IEEE Transactions on Cog. Comm. and Net.*, vol. 8, no. 1, pp. 263-281, 2022.

[16] D. Diao, B. Wang, K. Cao, R. Dong, and T. Cheng, "Enhancing Reliability and Security of UAV-enabled NOMA Communications with Power Allocation and Aerial Jamming," *IEEE Transactions on Vehicular Technology*, pp. 1-1, 2022.

[17] S. Mirbolouk, M. Valizadeh, M. Amirani, and S. Ali, "Relay Selection and Power Allocation for Energy Efficiency Maximization in Hybrid Satellite-UAV Networks With CoMP-NOMA Transmission," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 5, pp. 5087-5100, 2022.

[18] T. Vu, T. Nguyen, T. Nguyen, B. Vo-Nguyen, and S. Kim, "Short-Packet Communications in NOMA-CDRT IoT Networks With Cochannel Interference and Imperfect SIC," *IEEE Transactions on Vehicular Technology*, vol 71, no. 5, pp. 5552-5557, 2022.

[19] L. Yuan, N. Yang, F. Fang, and Z. Ding, "Performance analysis of UAV-assisted short-packet cooperative communications," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 4, pp. 4471–4476, 2022.

[20] B. Yu, X. Guan, and Y. Cai, "Joint blocklength and power optimization for half duplex unmanned aerial vehicle relay system with short packet communications," *Wireless Communications and Signal Processing (WCSP)*, 2020.

[21] Y. Hu, G. Sun, G. Zhang, M. C. Gursoy, and A. Schmeink, "Optimal Resource Allocation in-ground wireless networks supporting unmanned aerial vehicle transmissions," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 8, pp. 8972–8984, 2020.

[22] L. Lei, D. Yuan, C. Ho, and S. Sun, "Power and Channel Allocation for Non-Orthogonal Multiple Access in 5G Systems: Tractability and Computation," *IEEE Transactions on Wireless Communications*, vol. 15, no. 12, pp. 8580-8594, 2016.

[23] B. Su, Z. Qin, and Q. Ni, "Energy Efficient Uplink Transmissions in LoRa Networks," *IEEE Transactions on Communications*, vol. 68, no. 8, pp. 4960-4972, 2020.

[24] J. Shi, and L. Yang, "Novel Subcarrier-Allocation Schemes for Downlink MC DS-CDMA Systems," *IEEE Transactions on Wireless Communications*, vol. 13, no. 10, pp. 5716-5728, 2014.

[25] S. Lou, C. S. Chen, and M. Coupechoux. "Optimal joint subcarrier and power allocation in NOMA is strongly NP-hard," *In 2018 IEEE International Conference on Communications (ICC)*, pp. 1-7. IEEE, 2018.

[26] S. Zhang, H. Zhang, B. Di, and L. Song, "Cellular UAV-to-X communications: Design and optimization for Multi-UAV networks," *IEEE Transactions on Wireless Communications*, vol. 18, no. 2, pp. 1346–1359, 2019.

[27] V. H. Dang, V. Nguyen, V. Sharma, O. A. Dobre, and T. Q. Duong. "Digital twin empowered ultra-reliable and low-latency communications-based edge networks in industrial IoT environment," *In IEEE ICC*, pp. 5651-5656, 2022.

[28] K. Aryan, J. Thompson, E. Vlachos, C. Tsinos, and S. Chatzinotas, "Dynamic RF chain selection for energy efficient and low complexity hybrid beamforming in millimeter wave MIMO systems." *IEEE Transactions on Green Communication and Networking*, no. 4, pp. 886-900, 2019.

[29] W. Ghanem, V. Jamali, Y. Sun, and R. Schober, "Resource Allocation for Multi-User Downlink MISO OFDMA-URLLC Systems," *IEEE Transactions on Communications*, vol. 68, no. 11, pp. 7184-7200, 2020.

[30] J. Zhang, Y. Jiang, and X. Li, "Energy-efficient resource allocation in multiuser relay-based OFDMA networks," *Concurrency and Computation: Practice and Experience*, vol. 25, no. 9, pp. 1113-1125, 2012.

[31] S. Boyd, and L. Vandenberghe, "Convex optimization," *Cambridge: Cambridge Univ. Pr.*, 2011.

[32] Y. Chen, L. Xun, and S. Zhang, "The energy efficiency of heterogeneous cellular networks based on the poisson hole process," *Future Internet*, vol. 15, no. 2, pp. 56, 2023.

[33] A. Khalili, E. M. Monfared, S. Zargari, M. R. Javan, M. N. Yamchi, E. A. Jorswieck, "Resource Management for Transmit Power Minimization in UAV-assisted RIS HETNETS supported by dual connectivity," *IEEE Transactions on Wireless Communications*, vol. 21, no. 3, pp. 1806–1822, 2021.

[34] J. Chakareski, S. Naqvi, N. Mastronarde, J. Xu, F. Afghah, A. Razi, "An energy efficient framework for UAV-assisted millimeter wave 5G heterogeneous cellular networks," *IEEE Transactions on Green Communications and Networking*, vol. 3, no. 1, pp. 37–44, 2019.

[35] K. Minsu, and J. Lee. "Outage probability of UAV communications," *IEEE GLOBECOM*, pp. 1-6. IEEE, 2018.