

Dynamic Control for IAB Systems with Mixture of Latency- and Throughput-Sensitive Traffic

Natalia Yarkina, Dmitri Moltchanov, Yevgeni Koucheryavy

Abstract—The Integrated Access and Backhaul (IAB) architecture proposed by 3GPP is expected to provide a cost-efficient deployment option for 5G New Radio (NR) systems and should support all the services defined in 5G. However, IAB relies on multi-hop wireless communications, which makes the packet latency a critical metric for such systems. Latency optimization in IAB is further complicated by the half-duplex constraint forbidding IAB nodes from transmitting and receiving simultaneously over multiple antennas. As a result, active links must be dynamically scheduled to avoid conflicts and satisfy requirements of all supported traffic types in both uplink and downlink directions. In this paper, by utilizing the tools of Markov decision processes (MDP), we develop a reduced-complexity control policy for 5G self-backhauled systems simultaneously carrying latency- and throughput-sensitive traffic. The proposed policy minimizes the latency metric under low and moderate traffic conditions operating in ultra-low-complexity mode and switches to a more computationally demanding throughput-oriented regime when the offered load increases. Our results show that the core part of the policy – the proposed ultra-low-complexity Tree level switching control (TLS) – minimizes the latency for high-priority latency-sensitive traffic over the whole capacity region of the system but also for dominant throughput-oriented traffic over 75–85% of the capacity region.

Index Terms—5G, IAB, half-duplex, Markov decision process, latency.

I. INTRODUCTION

5G New Radio (NR) operating in microwave (μ Wave) and millimeter-wave (mmWave) bands is nowadays ready for commercial deployment. However, network operators still restrain from deploying mmWave NR systems. One of the reasons is the nature of mmWave links affected by blockages caused by both static and dynamic objects in the channel [1], [2]. These impairments result in small coverage areas of single mmWave NR base stations (BS) limited to just a few hundred meters [3]. A ubiquitous coverage at these frequencies thus requires very dense deployments drastically increasing operators' capital expenditures (CAPEX) [4].

To address the above challenge, 3GPP has standardized the Integrated Access and Backhaul (IAB) technology [5], [6, Clause 4.7], which employs low-cost relay nodes, called IAB-nodes, to extend the coverage of the BS, in this context referred to as the IAB-donor. IAB-nodes are connected to each

other and to the IAB-donor via wireless backhaul links. User equipment (UEs) can connect to either the IAB-node or IAB-donor without changing their behavior [6] and should be able to perceive the same quality of service (QoS) no matter the access point.

As the number of hops between the UE and the IAB-donor increases, the throughput and latency performance tend to degrade [5], [7]. To ensure the best possible performance, resource allocation at various levels of IAB operation should be coordinated among the nodes and optimized. The half-duplex constraint forbidding IAB-nodes/donor to simultaneously transmit and receive on their different interfaces [8] significantly complicates this task [9], [10] by making careful time allocation between subsets of links, known as link scheduling, a key factor of network performance [11].

Wireless backhauling becoming an integral part of the 5G network, IAB should support all services 5G has to offer [12] and adequately serve latency- and throughput-sensitive traffic. However, a latency-oriented network control may not allow to fully exploit the capacity of the system [13], whereas a throughput-optimal control may induce otherwise avoidable packet latencies and is characterized by high implementation complexity [14], [11]. There is thus the need for an IAB link scheduling policy capable of accommodating latency- and throughput-sensitive traffic and adaptive to varying traffic conditions.

The aim of this paper is to propose a reduced-complexity link scheduling policy for 5G IAB networks delivering a mixture of latency- and throughput-sensitive traffic. To this aim, we first formalize the IAB link scheduling optimization problem in a latency-focused manner, formulate the corresponding Markov decision process (MDP), and solve it directly using latency as the optimization criterion. We then identify a scalable fixed link scheduling pattern that ensures low latency for high-priority latency-sensitive traffic. Finally, we design a hybrid control policy that dynamically switches the operational regime of the system between latency- and throughput-oriented depending on the network conditions and whose latency-complexity trade-off can be customized.

The main contributions of our work are:

- a mathematical framework for latency-focused link scheduling optimization in IAB systems, built in terms of an MDP with finite states' and actions' sets which can be solved directly or via reinforcement learning;
- an ultra-low-complexity scalable fixed control policy called Tree level switching (TLS) that minimizes the latency for sparse high-priority latency-oriented traffic over the whole capacity region of the system, but also

N. Yarkina, D. Moltchanov and Y. Koucheryavy are with Tampere University, Finland. Email: firstname.lastname@tuni.fi.

This work is supported by the Academy of Finland under the projects "Machine Learning Methods and Algorithms for 6G Terahertz Cellular Access" (HARMONIOUS), "Enabling Mobile Terahertz Communication for 6G Cellular Networks" (EMERGENT), and "Machine Learning Algorithms for Energy Efficient and QoS Aware Communications in Heterogeneous 6G mmWave/sub-THz Networks" (ML6GThz).

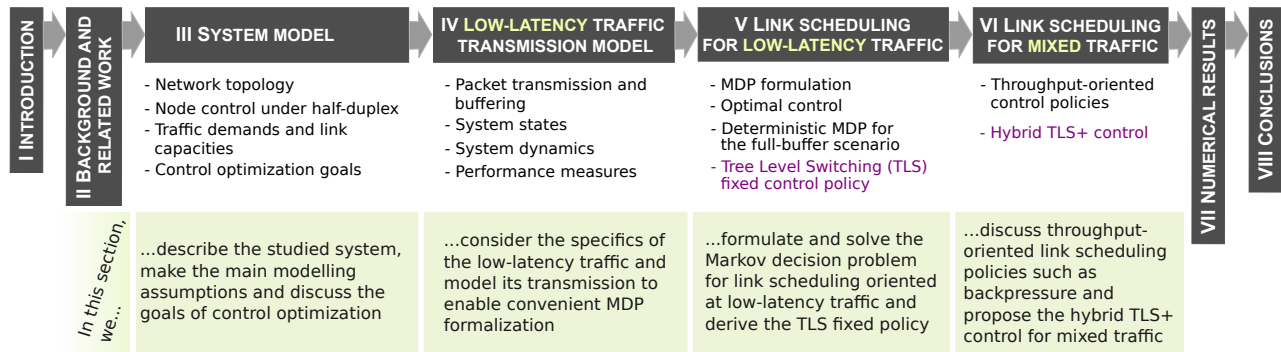


Fig. 1: The paper's structure and focuses.

for dominant throughput-oriented traffic over 75–85% of the capacity region;

- a hybrid TLS+ dynamic control policy for IAB networks with latency- and throughput-sensitive traffic adaptively switching between latency- and throughput-oriented regimes, which allows to fully utilize the throughput capacity of the system and has a customizable latency-complexity trade-off.

The paper is organized as follows. In Section II we introduce the 3GPP IAB architecture and review related studies. Section III presents the system model and assumptions. Sections IV and V deal with latency-sensitive traffic: in Section IV we formalize mathematically its transmission model, and in Section V formulate the associated MDP and derive a simple latency-oriented link scheduling policy not requiring explicit signaling between the IAB-nodes and donor. Then, in Section VI we address the throughput-oriented control and extend the proposed policy to the case of mixed traffic. A detailed numerical study of the proposed policies is provided in Section VII. Conclusions are drawn in the last section. For the reader's convenience, the structure and focuses of the core paper sections are portrayed in Fig. 1.

II. BACKGROUND AND RELATED WORK

In this section we first outline the IAB architecture and capabilities as defined by 3GPP and then proceed discussing the related work.

A. The IAB Technology

According to the IAB architecture defined by 3GPP in TS 39.401 [15], the IAB-donor is a gNB with additional functionality to support self-backhauling. It combines a gNB Central Unit (CU) and one or more gNB Distributed Units

(DU), and performs centralized resource, topology and route management for its controlled IAB network. An IAB-node implements gNB-DU to provide access connectivity to UEs and backhaul to its next-hop (children) IAB-nodes. The functional split and protocols in an example stand-alone (SA) IAB deployment are depicted in Fig. 2. For connecting to the DU of its parent node, which can be another IAB-node or the IAB-donor, the IAB-node supports a subset of UE functionality called IAB Mobile Termination (MT), which terminates the Uu interface of the upstream (towards the donor) backhaul link. Such an architecture results in hierarchical tree-like IAB topologies rooted at the donor [5], [16]. To enable packet routing over multiple hops, the IP layer in the IAB backhaul is carried over the IAB-specific Backhaul Adaptation Protocol (BAP) sub-layer defined in 3GPP TS 38.340 [17].

Details of the NR IAB operation are specified in 3GPP TS 38.174 [8]. Access and backhaul links of the IAB network can operate in either the same or different frequency bands, although the in-band option is considered more efficient as it allows the network operator to fully utilize the available costly spectrum [18]. IAB backhaul is mainly expected to use mmWave [5], [19], yet several sub-6 GHz bands are also supported [8]. Both SA and non-SA IAB deployments are possible. In the former, all the signalling is performed by the IAB system, whereas the latter implies LTE support. Multiple paths between an IAB-node and the donor are supported in SA mode by means of the NR dual connectivity mechanism, however, recent research suggests that the impact of meshed communications on increasing route redundancy may be negligible if the network is well planned [19].

Being a major challenge in multi-hop wireless networks, interference imposes restrictions on the use of interfaces at IAB-nodes/donor, in particular the half-duplex constraint [7], [16], which is illustrated in Fig. 3. Although in mmWave NR IAB, directional transmissions along with physical antenna separation and smart scheduling at IAB-nodes can substantially alleviate the interference problem [20], [21], [22], simultaneous transmission and reception by the same node is still out of reach in practical IAB deployments [19]. Full-duplex radios are expected to address this challenge in the future, but the technology is not yet mature for commercial market [23]. This is why the IAB specification [8] demands DU and MT of an IAB-node both either transmit or receive when operating simultaneously and indicates time division

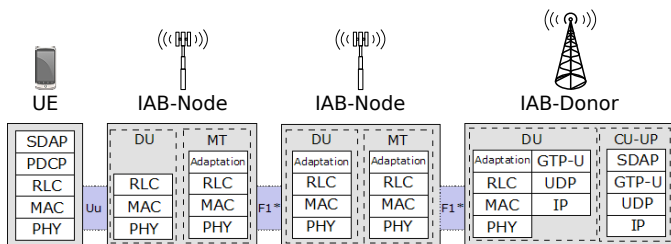


Fig. 2: Protocols in the IAB architecture, stand-alone mode.

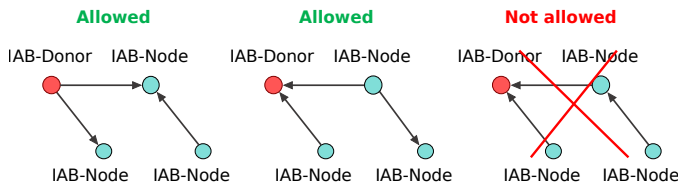


Fig. 3: The half-duplex constraint in IAB.

duplexing (TDD) as the main duplexing option.

B. Performance Optimization of IAB Systems

Performance optimization of IAB networks has been investigated along several major axes including user association and topology management [24], [25], route selection [26], [27], spacial multiplexing and beamforming [22], [28], resource allocation [20], [29], and link scheduling [9], [10], [13], [30], [31]. The latter is particularly relevant to 3GPP-compliant IAB deployments subject to the half-duplex constraint, which profoundly impacts the system performance [32]. A comprehensive survey of previous IAB-related research can be found in [19]. A recent survey on latency-oriented link scheduling in IAB networks is provided in [11].

The problem of link scheduling optimization has been studied in generic multi-hop wireless networks with interference-induced constraints. The seminal works [33], [34] by Tassiulas and Ephremides yielded the well-known backpressure (BP) link scheduling and routing algorithm along with a proof of its throughput optimality using the queuing-theoretic framework and the Lyapunov drift technique. BP provided the basis for a number of control policies aimed at remedying its two major shortcomings: poor latency performance [35], [36] and prohibitive implementation complexity [14], [37]. The latter covers not only the computational complexity per control decision, but also such characteristics as the signaling overhead/delay [36] or the need to maintain per-flow queues [37], [38]. Study [36] revealed a connection between BP and an MDP-based latency-minimizing control and exploited it to develop a throughput-optimal policy with improved latency performance, however it is still characterized by high complexity and the necessity for real-time signaling.

The majority of the aforementioned studies viewed the network's throughput capacity as the primary control optimization criterion, with the packet latency or implementation complexity coming next as secondary optimization goals. However, having a multi-hop network instead of just a single link at the last mile in 5G makes packet latency a critical performance metric. This has been observed, for instance, in multi-hop vehicle-to-vehicle (V2V) communications systems [39], [40], [41]. Therefore, although throughput remains among the key criteria of IAB link scheduling optimization [30], [42], the end-to-end packet latency improvement often plays an equally important [7], [32] and sometimes central role [9], [13], [31].

The MDP framework [43], along with reinforcement learning (RL) techniques, provides a systematic approach to latency-oriented control. Thus, the authors of [9] considered the problem of link scheduling in self-backhauled systems with the half-duplex constraint as an MDP, solved it formally by utilizing dynamic programming and then obtained a practical

solution via RL approximation. The reported results illustrate the superiority of the proposed approach for topologies where IAB-nodes/donor have a single radio interface. However, only the downlink transmission direction was considered, whereas a particular challenge in half-duplex IAB control is to accommodate both downlink and uplink traffic.

Although in some special cases the MDP framework can result in a simple optimal solution [13], [44], typically, the control policy it yields is associated with the particular workload conditions under which it has been obtained and is centralized, i.e., relies on the full knowledge of the network state by the controller [9], [29], which may induce signaling delays in SA IAB deployments. Throughput-oriented distributed link scheduling policies for downlink-only IAB networks were proposed in [30], [45]. While the solution in [45] stems from the queueing-theoretic framework, that in [30] relies on a multi-agent RL architecture where control agents are located at IAB-donor/nodes' antenna panels. In the training phase, the agents are coordinated by centralized critics located at the IAB-donor, but once the training is completed, control is fully distributed with no information exchange between the donor and nodes. A latency-oriented distributed RL-based control was proposed for half-duplex IAB with uplink traffic in [31]. The authors aim their policy at reliable latency performance by making the loss function combine the value of packet latency with its conditional value at risk (CVaR) metric, which provides an estimate of the expected latency in the tail of its distribution. Link scheduling decisions are taken locally at IAB-nodes by multi-armed bandit-type agents, however, a centralized consensus mechanism is used for coordination.

Although an IAB deployment should satisfy requirements of both throughput- and latency-sensitive traffic in the downlink and uplink directions, a limited number of optimization results deliver an IAB-oriented compromise between latency performance, throughput and control complexity. In this work, we aim to fill this gap and propose a reduced-complexity approach to link scheduling in a half-duplex IAB system delivering a mixture of latency- and throughput-sensitive traffic.

III. SYSTEM MODEL

In this section we describe the considered system model and introduce our key assumptions. Similarly to [9], [10], [32], we focus on the network-related aspects of the system, namely the multi-hop transmission under time division multiplexing (TDM) and the half-duplex constraint as predominant factors affecting the packet latency.

A. Network Topology

We consider an IAB network with one IAB-donor and $N-1 > 0$ IAB-nodes forming a tree topology with the donor at its root, as shown in Fig. 4. Spanning-tree topologies are suggested for IAB by 3GPP in [16] and imply that only one route exists between the IAB-donor and each IAB-node. It is assumed that each IAB-node and the donor have a designated interface for backhauling and the backhaul topology is fixed, e.g., at the network planning phase.

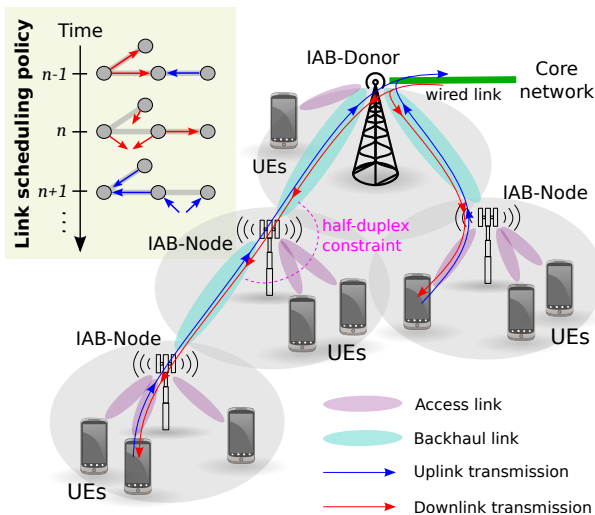


Fig. 4: System model illustration. The IAB-donor implements the control functionality and has a wired link to the core network. The IAB-nodes act as access points and wireless relays carrying traffic from and to the IAB-donor. The link scheduling policy determines the subsets of wireless links to be activated at each time.

Let the IAB network be represented by an undirected graph $\langle \mathcal{N}, \mathcal{E}_{\mathcal{N}} \rangle$ that consists of $|\mathcal{N}| = N$ vertices corresponding to all network nodes including the donor. Hereafter, we use the terms “node” and “vertex” interchangeably to refer to IAB-nodes and the IAB-donor. The set $\mathcal{E}_{\mathcal{N}}$ includes the edge between two vertices whenever backhaul transmission between the corresponding network nodes is allowed by the topology.

The vertices are labeled from 1 to N so that: (i) vertex 1 corresponds to the IAB-donor, and (ii) if $i < j \leq N$ then the distance from vertex 1 to vertex i (in terms of the number of edges) is not greater than from 1 to j . The graph $\langle \mathcal{N}, \mathcal{E}_{\mathcal{N}} \rangle$ is specified by the adjacency matrix $\mathbf{T} = (T_{i,j})_{i,j \in \mathcal{N}}$, in which $T_{i,j} = 1$ if there is an edge between vertices i and j , i.e., $(i, j) \in \mathcal{E}_{\mathcal{N}}$, and $T_{i,j} = 0$ otherwise.

Data packets are transmitted in the uplink and downlink directions. The uplink direction corresponds to transmission from the IAB-nodes (UEs) to the IAB-donor, and the downlink from the IAB-donor to the IAB-nodes (UEs). Downlink packets arrive to the IAB-donor from the core network (CN) over a wired link, so do uplink packets leaving the IAB-donor for CN. The superdiagonal entries of \mathbf{T} define the downlink transmission, while its subdiagonal elements describe the uplink. The diagonal entries of \mathbf{T} are assumed zero.

B. Node Control under Half-Duplex

Due to the half-duplex constraint, it is assumed that simultaneous transmission and reception over different interfaces are not allowed at the IAB-donor and IAB-nodes. In other words, an IAB network node can be either sending or receiving data.

We assume that the network is controlled in a semi-centralized manner. The model time is slotted and the time step duration corresponds to the transmission time interval (TTI) in 5G NR. In each time step, first, a backhaul transmission takes place during a fixed time interval of duration $h > 0$.

Here, transmit/receive modes are determined by an IAB-donor-located controller following a control policy. The mode of each node can be switched at the beginning of each time step and remains constant during the backhaul transmission phase. Then, until the beginning of the next time step, access is scheduled and provided to UEs independently by all IAB network nodes. The centralized control action applies to the first backhaul phase only.

Let $a_i(n) = 0$ if, at time step n , node i is receiving, and $a_i(n) = 1$ if it is transmitting, $\mathbf{a}(n) = (a_i(n))_{i \in \mathcal{N}}$. Then the actual transmission topology at time steps $n = 0, 1, 2, \dots$ is determined by \mathbf{T} and the control vector $\mathbf{a}(n)$. In order for a packet to travel from one node to another, the source node must be transmitting while the destination node receiving. Thus, for any $i, j \in \mathcal{N}$, data transmission from i to j is possible only if the following three conditions are met: (i) node i is transmitting, (ii) node j is receiving, and (iii) $T_{i,j} = 1$. Then, given that $\mathbf{a}(n) = \mathbf{a} \in \{0, 1\}^N$, the elements of an adjacency matrix $\mathbf{R}(n) = \mathbf{R}(\mathbf{a}(n))$ defining the actual transmission topology graph at time n are

$$R_{i,j}(\mathbf{a}) = T_{i,j} \mathbf{1}_{a_i - a_j = 1}, \quad i, j \in \mathcal{N}, \quad (1)$$

where $\mathbf{1}_x$ denotes the indicator function that equals 1 if x is true, and 0 otherwise.

Fig. 5 shows an example five-node IAB backhaul network with an applied control $\mathbf{a}(n) = (1, 1, 0, 0, 0)$ and the corresponding adjacency matrix of the transmission topology.

C. Traffic Demands and Link Capacities

We consider two services provisioned over the IAB network producing two types of traffic: high-priority low-latency (LL) and voluminous conventional data traffic referred to as broadband (BB). Three mechanisms to comply with the latency and reliability constraints of the LL service are utilized [46]. First, to decrease its latency at a single backhaul link, we assume that LL traffic receives queuing priority over BB, i.e., BB packets are transmitted only when there are no LL packets queued for the same link. Second, to account for the reliability constraints, a conventional target block error rate (BLER) of $p_{\text{BB}} = 0.1$ is utilized for BB packets, while a conceptually stronger modulation and coding (MCS) scheme enabling BLER of $p_{\text{LL}} = 0.01$ is used for LL packets. Finally, to reach the target error probability of 10^{-6} , for LL packets replication coding with parameter $r = 3$ is assumed [47].

We let the IAB-donor and IAB-nodes all have groups of associated UEs sending and receiving data packets of both services to/from the CN. Thus, the IAB backhaul serves $2(N - 1)$ traffic flows, $N - 1$ uplink and $N - 1$ downlink, which we index by the number of the network node to which the UE group is associated. Node 1 is the source node of all downlink flows and the destination node of all uplink flows. Let $\mathcal{N}_k \subset \mathcal{N}$ denote the set of nodes through which packets of flow k travel and let $\mathcal{N}_1 = \{1\}$. The sets \mathcal{N}_k , $k \in \mathcal{N} \setminus \{1\}$, can be easily obtained from \mathbf{T} and uniquely determine the routes of the flows and, by consequence, the order of visited nodes, since the nodes are indexed according to their distance from the root node. Denote $N_k = |\mathcal{N}_k|$, $k \in \mathcal{N}$, and notice that $N_k - 1$ is

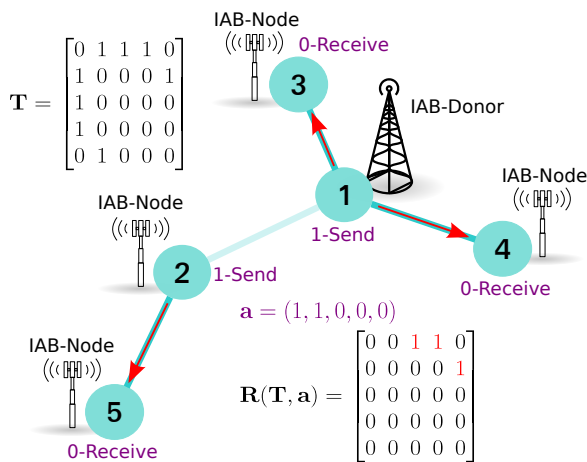


Fig. 5: An example IAB backhaul network with an applied control and the corresponding connectivity matrix.

the distance from node k to node 1. The longest route in the network is thus of length $N_N - 1$.

The number of packets exogenously arriving to the source node of each flow in each time step is assumed Poisson-distributed with means λ_k^{UL} and λ_k^{DL} for uplink and downlink flows, respectively. An arriving packet belongs to the LL type with probability q_{LL} , otherwise it is of BB type.

The IAB-donor and nodes operate in the mmWave band according to 5G NR. To account for specifics of radio transmission including directional antennas, code rate, numerology as well as MCS schemes, we utilize the model specified in 3GPP TS 38.306 [48, Clause 4.1.2]. Accordingly, the radio link capacity can be approximated as

$$C_{[\text{Mbps}]} = 10^{-6} \nu Q_m R_{\text{max}} 12 N_{\text{PRB}}^{B, \mu} (1 - \omega) / T_{\text{smb1}}^{\mu}, \quad (2)$$

where ν is the number of multiplexed layers, Q_m the modulation order, R_{max} the code rate, $N_{\text{PRB}}^{B, \mu}$ the number of available primary resource blocks for the given bandwidth B and numerology μ , w the overhead, and $T_{\text{smb1}}^{\mu} = 10^{-3} / (14 \times 2^{\mu})$ the average symbol duration for numerology μ .

We assume that the same amount of link capacity, $\sigma_{[\text{bit}]}$, is needed to transmit one LL and BB packet because the former are replicated to achieve a greater transmission reliability.

D. Control Optimization Goals

For the considered system, we are interested in an efficient control policy for the IAB network nodes. However, the QoS requirements of the considered traffic types differ. The key performance measure for LL traffic is the end-to-end packet latency. For BB traffic, on the other hand, the central metric of interest is the network throughput, i.e. the amount of traffic that can be transmitted by the network. Moreover, especially for LL traffic, the control policy should not induce important additional computational and/or signaling delays.

To reconcile these contradictory needs, the link scheduling policy should (i) minimize the end-to-end latency of packet delivery in the IAB network at least for the prioritized LL packets, (ii) allow to exploit the full capacity region of the system, and (iii) induce as low control overhead/delay as possible. The latter requirement implies that the policy should

not rely on the full knowledge of the system state at any decision epoch, as it is the case for throughput-optimal link scheduling policies [11].

Our response to this challenge is a hybrid control, referred to as TLS+, that by default operates in low-complexity latency-minimizing mode, but switches to a more computationally challenging throughput-maximizing mode upon a customizable trigger. In the next two sections we focus on LL traffic and derive the latency-oriented component, namely the TLS policy. Then, in Section VI we turn to BB traffic, discuss throughput-oriented link scheduling, and propose the hybrid TLS+ policy, which can be customized to balance the needs of both traffic types.

IV. LOW-LATENCY TRAFFIC TRANSMISSION MODEL

To derive a control policy that addresses the needs of LL traffic, we first build a model of LL traffic transmission in a half-duplex IAB backhaul and then use it to formulate a Markov decision problem in Section V.

A. Packet Transmission and Buffering

We let all vertices in \mathcal{N} have per-neighbor buffers of unlimited capacity. Node 1 has several downlink buffers – one for each incident edge – that receive external arrivals. These represent packets arriving from the CN over a wired link. At the same time, node 1 is the sink for all uplink flows. Each node $i \in \mathcal{N} \setminus \{1\}$ has one uplink buffer receiving external arrivals and $\sum_{j=i}^N T_{i,j}$ downlink buffers. Nodes in $\mathcal{N} \setminus \{1\}$ are sinks for the corresponding downlink flows.

Upon arrival to node i , a downlink packet of flow k joins the downlink buffer corresponding to the outbound edge through which vertex k is accessible. An uplink packet joins the uplink buffer. If the packet is of BB-type, it is placed at the end of the queue. An LL packet is placed before the first BB packet.

Now, if $R_{i,j}(n) = 1$ for some $i < j$, then in time step n the first $\lfloor h_{[s]} C_{[\text{bps}]} / \sigma_{[\text{b}]} \rfloor$ packets from downlink buffer (i, j) of node i attempt transmission, which succeeds with probability $1 - p_{\text{LL}}^r$ for LL packets and with probability $1 - p_{\text{BB}}$ for BB. In the case of successful transmission the packet moves to node j if $j < k$ or leave the network if $j = k$, otherwise it remains at node i and keeps its position in the queue. Similarly, if in time step n data transmission is possible from node i to an adjacent node $j < i$ (uplink), i.e., if $R_{i,j}(n) = 1$, then the first $\lfloor hC / \sigma \rfloor$ packets of the uplink buffer in node i attempt transmission in this time step. If transmission is successful, the packet moves to the uplink buffer in node j if $j > 1$ or leaves the network if $j = 1$, otherwise it remains in node i and keeps its position in the queue.

B. System States

Under the considered assumptions, BB-type packets, being of lower priority in queues, do not affect LL traffic. Furthermore, we assume that the share of LL traffic, which is determined by the probability q_{LL} for a packet to belong to this type, is rather small, e.g. 5–10 %, and the probability of its successful transmission is close to 1. This implies that if a

link is enabled by the control policy at a certain time step, we can assume that all LL packets from the corresponding buffer are transmitted in the same time step, which is equivalent to assuming infinite link capacities for LL traffic transmission. Still, since each node is either transmitting or receiving, no packet can travel over more than one link in one time step.

The state of a queuing network is commonly described via the number of packets in the queues. However, we adopt a different approach and define the state of a buffer by the age of the “oldest” LL packet therein. By the age we understand the number of time steps elapsed since the arrival of the packet into the network. Such an approach reduces the state space of the model and permits focusing on latency optimization, which is the key performance measure in our analysis.

We assume that all packets entering the network during time step $n-1$ find themselves in the buffers receiving external arrivals at the beginning of time step n aged 1. Then, in each consecutive time step spent in the system, the age of the packet increases by 1, independently of whether the packet moves between nodes or not. If the packet leaves the system at time $n+m$ then it has experienced a latency of $m+1$ time steps.

Denote by $s^{\text{UL}}(i, n)$ the maximum LL packet age in the uplink buffer of node $i \in \mathcal{N} \setminus \{1\}$ at time n , and let $s^{\text{UL}}(i, n) = 0$ if the buffer does not contain any LL packets. For convenience of further notation, we introduce a row vector $\mathbf{s}^{\text{UL}}(n) = (s^{\text{UL}}(2, n), \dots, s^{\text{UL}}(N, n))$ and denote its components by $s_i^{\text{UL}}(n) = s^{\text{UL}}(i+1, n)$, $i = 1, \dots, N-1$. Similarly, let $s_{k,i}^{\text{DL}}(n)$ represent the maximum age of the LL packet of flow k in the downlink buffers of node $i \in \mathcal{N}$ at time n . We denote $\mathbf{s}_k^{\text{DL}}(n) = (s_{k,1}^{\text{DL}}(n), \dots, s_{k,k-1}^{\text{DL}}(n))$ and $\mathbf{s}^{\text{DL}}(n) = (\mathbf{s}_2^{\text{DL}}(n), \dots, \mathbf{s}_N^{\text{DL}}(n))$.

Now, the state of the whole system can be described by a row vector $\mathbf{s} = (\mathbf{s}^{\text{UL}}, \mathbf{s}^{\text{DL}})$ of size $N(N+1)/2-1$ having non-negative integer-valued components. However, not any vector from $\{0, 1, 2, \dots\}^{N(N+1)/2-1}$ represents a feasible system state. In particular, we note that $s_{k,i}^{\text{DL}}(n) = 0$ for all $i \in \mathcal{N} \setminus \mathcal{N}_k$. Indeed, we use zero padding to preserve the routes and \mathbf{s}^{DL} actually has only $\sum_{k \in \mathcal{N} \setminus \{1\}} (N_k - 1)$ significant entries.

Since in each time step a downlink LL packet either moves one hop to its destination or waits for the link to switch on, its age at a node is always greater than the distance from this node to the root, therefore if $s_{k,i}^{\text{DL}}(n) > 0$, then $s_{k,i}^{\text{DL}}(n) \geq N_i$. Also, because the nodes are numbered with respect to their distance from the root, if $s_{k,j}^{\text{DL}}(n) > 0$ then $i < j \Leftrightarrow s_{k,i}^{\text{DL}}(n) < s_{k,j}^{\text{DL}}(n)$.

C. System Dynamics

At the beginning of time step n we find the system in state $\mathbf{s}(n)$. Then, first, a control $\mathbf{a}(n)$ is applied and all LL packets of the corresponding buffers are transmitted over enabled links. The packets received from adjacent nodes are put into buffers for further transmission or leave the network (if arrived to node 1 in the uplink or to the destination nodes in the downlink). We denote this intermediate state of the system by $\mathbf{s}(n+0)$. After that, all external arrivals in time step n , i.e., the packets arrived to node 1 from the CN and to nodes $2, \dots, N$ from UEs, aged 0, are put into the respective buffers. Next, before the end of the time step, the age of all packets in the system

increases by 1, and then the new system state, $\mathbf{s}(n+1)$, is finalized for time step $n+1$.

Let us now derive the system equations linking $\mathbf{s}(n)$ and $\mathbf{s}(n+1)$. To describe the uplink state transitions, we will make use of a lower triangular matrix $\mathbf{R}^{\text{UL}}(n)$ of order $N-1$, whose elements are obtained from matrix $\mathbf{R}(n)$ as follows

$$R_{i,j}^{\text{UL}}(n) = \begin{cases} R_{i+1,j+1}(n) & \text{if } 1 \leq j < i < N, \\ 1 - \sum_{k=1}^i R_{i+1,k}(n) & \text{if } 1 \leq j = i < N, \\ 0 & \text{otherwise.} \end{cases} \quad (3)$$

Then,

$$s^{\text{UL}}(i, n+0) = \max_{j \in \mathcal{N} \setminus \{1\}} \{s^{\text{UL}}(j, n) R_{j-1,i-1}^{\text{UL}}(n)\}, \quad (4)$$

provides the intermediate state of the uplink buffers after applying the control and transferring the packets over enabled links. Notice that (4) resembles matrix multiplication with the summation replaced by the maximum function.

To obtain $\mathbf{s}^{\text{UL}}(n+1)$ we account for external arrivals at time step n and increase the age of all packets in the system by 1. Let $\theta_i^{\text{UL}}(n)$, $i \in \mathcal{N} \setminus \{1\}$, equal 1 or 0 depending on whether the uplink buffer of node i received external arrivals in time step n . This yields, for any $i \in \mathcal{N} \setminus \{1\}$ and $n \geq 0$,

$$s^{\text{UL}}(i, n+1) = \begin{cases} s^{\text{UL}}(i, n+0) + 1 & \text{if } s^{\text{UL}}(i, n+0) > 0, \\ \theta_i^{\text{UL}}(n) & \text{otherwise.} \end{cases} \quad (5)$$

For convenience of notation, for an arbitrary vector \mathbf{x} of length J , a $J \times J$ binary matrix \mathbf{M} and any index $i = 1, \dots, J$, we define a mapping

$$f_i(\mathbf{x}, \mathbf{M}) = \max_{j=1, \dots, J} \{x_j M_{j,i}\} + \mathbf{1}_{\max_{j=1, \dots, J} \{x_j M_{j,i}\} > 0}. \quad (6)$$

Now, the system equation for the uplink can be written as

$$s^{\text{UL}}(i, n+1) = \max \{f_{i-1}(s^{\text{UL}}(n), \mathbf{R}^{\text{UL}}(n)), \theta_i^{\text{UL}}(n)\}, \quad (7)$$

for any $i \in \mathcal{N} \setminus \{1\}$ and $n \geq 0$.

We address the downlink transmission in a similar manner. Let $\theta_k^{\text{DL}}(n) = 0$ if no LL packets of downlink flow k arrived to node 1 in time step n and $\theta_k^{\text{DL}}(n) = 1$ otherwise. For any $k \in \mathcal{N} \setminus \{1\}$ we define an upper triangular square matrix $\mathbf{R}_k^{\text{DL}}(n)$ of order $k-1$ as

$$R_{k,i,j}^{\text{DL}}(n) = \begin{cases} R_{i,j}(n) & \text{if } i, j \in \mathcal{N}_k \setminus \{k\}, i < j, \\ 1 - \sum_{l \in \mathcal{N}_k, l > i} R_{i,l}(n) & \text{if } i \in \mathcal{N}_k \setminus \{k\}, j = i, \\ 0 & \text{otherwise.} \end{cases} \quad (8)$$

For any $k \in \mathcal{N} \setminus \{1\}$, $i = 1, \dots, k-1$ and $n \geq 0$ we have

$$s_{k,i}^{\text{DL}}(n+0) = \max_{1 \leq j < k} \{s_{k,j}^{\text{DL}}(n) R_{k,j,i}^{\text{DL}}(n)\}, \quad (9)$$

and also

$$s_{k,i}^{\text{DL}}(n+1) = \begin{cases} s_{k,i}^{\text{DL}}(n+0) + 1 & \text{if } s_{k,i}^{\text{DL}}(n+0) > 0, \\ \theta_k^{\text{DL}}(n) & \text{if } i = 1 \text{ and } s_{k,1}^{\text{DL}}(n+0) = 0, \\ 0 & \text{otherwise.} \end{cases} \quad (10)$$

Using notation (6), for any $k \in \mathcal{N} \setminus \{1\}$ we can write

$$s_{k,1}^{\text{DL}}(n+1) = \max \{f_1(s_k^{\text{DL}}(n), \mathbf{R}_k^{\text{DL}}(n)), \theta_k^{\text{DL}}(n)\}, \quad (11)$$

$$s_{k,i}^{\text{DL}}(n+1) = f_i(s_k^{\text{DL}}(n), \mathbf{R}_k^{\text{DL}}(n)), \quad 1 < i < k.$$

From the control theory perspective, equations (7) and (11) relate the state of the system at time $n + 1$ to the state of the system at time n , the control used at time n , $\mathbf{a}(n)$, which determines the matrices $\mathbf{R}^{\text{UL}}(n), \mathbf{R}_2^{\text{DL}}(n), \dots, \mathbf{R}_N^{\text{DL}}(n)$, and the disturbance at time n , given by the external arrival indicator

$$\boldsymbol{\theta}(n) = \left(\theta_i^{\{\text{UL}, \text{DL}\}}(n) \right)_{i \in \mathcal{N} \setminus \{1\}} \in \{0, 1\}^{2(N-1)}. \quad (12)$$

D. Performance Measures

To estimate the quality of a control \mathbf{a} applied in a state \mathbf{s} , we evaluate the maximum LL packet age in the network after enacting the control, packet transmission and aging but before registering new external arrivals. Since the sought-for control policy is aimed at minimizing the maximum latency in the network, we can assume that by applying control \mathbf{a} in state \mathbf{s} the controller incurs a cost of

$$c_{\max}(\mathbf{s}, \mathbf{a}) = \max \left\{ \max_{i \in \mathcal{N} \setminus \{1\}} f_i(\mathbf{s}^{\text{UL}}, \mathbf{R}^{\text{UL}}(\mathbf{a})), \max_{k \in \mathcal{N} \setminus \{1\}, i \in \mathcal{N}_k \setminus \{k\}} f_i(\mathbf{s}_k^{\text{DL}}, \mathbf{R}_k^{\text{DL}}(\mathbf{a})) \right\}. \quad (13)$$

Alternatively, to additionally encourage optimization of shorter routes, we can express the incurred cost as the sum of maximum LL packet ages in all buffers, which corresponds to the unnormalized buffer-average (not packet-average) maximum packet age:

$$c_{\text{avg}}(\mathbf{s}, \mathbf{a}) = \sum_{k \in \mathcal{N} \setminus \{1\}} \left(f_k(\mathbf{s}^{\text{UL}}, \mathbf{R}^{\text{UL}}(\mathbf{a})) + \sum_{i \in \mathcal{N}_k \setminus \{k\}} f_i(\mathbf{s}_k^{\text{DL}}, \mathbf{R}_k^{\text{DL}}(\mathbf{a})) \right). \quad (14)$$

V. LINK SCHEDULING FOR LOW-LATENCY TRAFFIC

We estimate the efficiency of control for LL traffic in terms of the maximum latency experienced by packets of this type. Namely, we demand that a control policy: (i) ensures that the latency of any LL packet does not exceed some fixed value, τ , which depends on the network topology and more specifically on the length of its longest route $N_N - 1$ and (ii) minimizes the maximum latency of LL packets. For the model of the previous section, the MDP framework permits deriving a deterministic stationary optimal control policy π which maps any state \mathbf{s} to an action $\mathbf{a} = \pi(\mathbf{s})$ [43]. We provide the MDP formalization next in this section.

However, to reduce the signaling overload and considering the stringent time constraints of IAB scheduling, it is preferable to obtain some fixed policy, $\mathbf{a}^{(1)}, \mathbf{a}^{(2)}, \dots, \mathbf{a}^{(L)}$, that can be applied on a loop regardless of the system state and still provide acceptable network performance. The MDP framework can help us in this task as well, if we find an adequate way to eliminate randomness in the process model. Indeed, a sufficiently long trajectory of a process following some optimal deterministic control policy in a finite state set with deterministic transitions will necessarily loop: the process is bound to return to some state because the state set is finite, and the trajectory will repeat from this point because the transitions are deterministic. The sequence of actions of such a

loop can be taken as the fixed policy, provided that the process starts in a state of the loop or the proposed policy ensures it can reach its state. To make the process transitions deterministic we consider the worst-case, full-buffer scenario in which new LL packets join all queues receiving external arrivals at each time step. The full-buffer MDP will be considered in details further in this section and the obtained fixed policy – Tree level switching (TLS) – will be discussed in Subsection V-D.

A. MDP Formulation

An MDP formulation for the problem of optimal control for LL traffic in the IAB backhaul network with the stated above optimization criteria follows.

1) *Decision epochs (infinite horizon):*

$$\mathcal{T} = \{0, 1, 2, \dots\}. \quad (15)$$

2) *States (finite):* We demand that no LL packet experience a latency larger than a threshold τ and let the set of states,

$$\mathcal{S} \subset \{0, 1, \dots, \tau\}^{\frac{N(N+1)}{2}-1}, \quad (16)$$

consist of all such vectors $\mathbf{s} = (\mathbf{s}^{\text{UL}}, \mathbf{s}^{\text{DL}})$ that can be obtained from state $(0, \dots, 0)$ by a finite sequence of controls $\mathbf{a}_1, \dots, \mathbf{a}_J$ for some combination of disturbances $\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_J$. Such states satisfy the properties stated in Section IV-B and the number of significant entries in \mathbf{s} is $\sum_{k \in \mathcal{N} \setminus \{1\}} N_k$.

We will refer to system states that have one or more entries equal τ as terminal states and denote their set by $\mathcal{S}_\tau \subset \mathcal{S}$. The set of non-terminal states is denoted $\bar{\mathcal{S}}_\tau = \mathcal{S} \setminus \mathcal{S}_\tau$.

3) *Actions (finite):* We define the set of feasible actions in state $\mathbf{s} \in \mathcal{S}$ as

$$\mathcal{A}_\mathbf{s} = \begin{cases} \emptyset & \text{if } \mathbf{s} \in \mathcal{S}_\tau, \\ \mathcal{A} & \text{otherwise,} \end{cases} \quad (17)$$

where $\mathcal{A} = \{0, 1\}^N$. We assume that if any of the state vector entries attains the threshold latency τ , the process terminates.

4) *Expected rewards (stationary and bounded):* We define the controller reward for applying control \mathbf{a} in state \mathbf{s} as

$$r_X(\mathbf{s}, \mathbf{a}) = \tau - c_X(\mathbf{s}, \mathbf{a}), \quad X \in \{\max, \text{avg}\}, \quad (18)$$

where $c_X(\mathbf{s}, \mathbf{a})$ is given either by (13) or (14) representing, respectively, the overall or the buffer-average maximum LL packet age in all buffers after packet transmission and aging but before registering external arrivals.

5) *Transition probabilities (stationary):* It is assumed that the vectors of external arrival indicators, $\boldsymbol{\theta}(n)$, are realizations of random vectors $\Theta(n)$ that take values in $\{0, 1\}^{2(N-1)}$, are independent of $\boldsymbol{\theta}(m)$ for any $m < n$, and have probability distributions that do not depend on $\mathbf{s}(n)$ or $\mathbf{a}(n)$. Let p_i^{UL} and p_i^{DL} denote, respectively, the probabilities that $\Theta^{\text{UL}}(i, n)$ and $\Theta_i^{\text{DL}}(n)$ assume the value of 1, and thus external arrivals to different network nodes occur independently of each other.

Next, we notice that as a result of action $\mathbf{a} \in \mathcal{A}$ and external arrivals the system can move from a non-terminal state $\mathbf{s} \in \bar{\mathcal{S}}_\tau$ only to such states $\tilde{\mathbf{s}}$ that

$$1) \tilde{s}_i^{\text{UL}} = f_i(\mathbf{s}^{\text{UL}}, \mathbf{R}^{\text{UL}}(\mathbf{a})) \text{ if } f_i(\mathbf{s}^{\text{UL}}, \mathbf{R}^{\text{UL}}(\mathbf{a})) > 0 \text{ and } \tilde{s}_i^{\text{UL}} \in \{0, 1\} \text{ otherwise, } i = 1, \dots, N - 1;$$

- 2) $\tilde{s}_{k,i}^{\text{DL}} = f_i(\mathbf{s}_k^{\text{DL}}, \mathbf{R}_k^{\text{DL}}(\mathbf{a}))$ for all $i = 2, \dots, k-1$, $k \in \mathcal{N} \setminus \{1\}$.
- 3) $\tilde{s}_{k,1}^{\text{DL}} = f_1(\mathbf{s}_k^{\text{DL}}, \mathbf{R}_k^{\text{DL}}(\mathbf{a}))$ if $f_1(\mathbf{s}_k^{\text{DL}}, \mathbf{R}_k^{\text{DL}}(\mathbf{a})) > 0$ and $\tilde{s}_{k,1}^{\text{DL}} \in \{0, 1\}$ otherwise, $k \in \mathcal{N} \setminus \{1\}$.

We denote the set of all such states $\tilde{\mathbf{s}}$ for $\mathbf{a} \in \mathcal{A}$ and $\mathbf{s} \in \bar{\mathcal{S}}_\tau$ by $\mathcal{C}(\mathbf{s}, \mathbf{a}) \subset \mathcal{S}$, and let $\mathcal{C}(\mathbf{s}, \mathbf{a}) = \emptyset$ for $\mathbf{s} \in \mathcal{S}_\tau$.

Now, for any $\mathbf{s} \in \bar{\mathcal{S}}_\tau$ and $\mathbf{a} \in \mathcal{A}_\mathbf{s}$, the conditional probability to find the process in state $\tilde{\mathbf{s}} \in \mathcal{S}$ given that in the previous time step it was in state \mathbf{s} and action \mathbf{a} was taken is

$$p(\tilde{\mathbf{s}}|\mathbf{s}, \mathbf{a}) = \prod_{i=2}^N \left((p_i^{\text{UL}})^{\tilde{s}_{i-1}^{\text{UL}}} (1 - p_i^{\text{UL}})^{1 - \tilde{s}_{i-1}^{\text{UL}}} \right)^{\mathbf{1}_{f_{i-1}(\mathbf{s}^{\text{UL}}, \mathbf{R}^{\text{UL}}(\mathbf{a}))=0}} \times \left((p_i^{\text{DL}})^{\tilde{s}_{i,1}^{\text{DL}}} (1 - p_i^{\text{DL}})^{1 - \tilde{s}_{i,1}^{\text{DL}}} \right)^{\mathbf{1}_{f_1(\mathbf{s}_i^{\text{DL}}, \mathbf{R}_i^{\text{DL}}(\mathbf{a}))=0}} \quad (19)$$

if $\tilde{\mathbf{s}} \in \mathcal{C}(\mathbf{s}, \mathbf{a})$, and $p(\tilde{\mathbf{s}}|\mathbf{s}, \mathbf{a}) = 0$ for all $\tilde{\mathbf{s}} \notin \mathcal{C}(\mathbf{s}, \mathbf{a})$. Indeed, the states of the buffers not receiving external arrivals are completely determined by the original state and action, so the above expression does not consider the downlink buffers at nodes $i > 1$. Furthermore, if a buffer was not empty after applying the action in the previous state \mathbf{s} , its state is not affected by external arrivals. To reflect this we take the corresponding factor to the power of $\mathbf{1}_{f_{i-1}(\mathbf{s}^{\text{UL}}, \mathbf{R}^{\text{UL}}(\mathbf{a}))=0}$ or $\mathbf{1}_{f_1(\mathbf{s}_i^{\text{DL}}, \mathbf{R}_i^{\text{DL}}(\mathbf{a}))=0}$. Finally, after registering external arrivals the buffer state can be either 1, with probability $p_i^{\{\text{UL}, \text{DL}\}}$, or 0, with probability $1 - p_i^{\{\text{UL}, \text{DL}\}}$.

We refer to $\{\mathcal{T}, \mathcal{S}, \mathcal{A}_\mathbf{s}, r_{\{\text{max}, \text{avg}\}}(\mathbf{s}, \mathbf{a}), (\tilde{\mathbf{s}}|\mathbf{s}, \mathbf{a})\}$ as the backhaul MDP_{max} or MDP_{avg} depending on the reward r .

B. Optimal Control

To find an optimal LL-traffic-oriented IAB scheduling policy using the backhaul model MDP described above, we assume the expected total discounted reward optimality criterion with a discount factor $0 \leq \lambda < 1$ and adopt the dynamic programming method. We recall that if the system is in state \mathbf{s}_0 at the first decision epoch, the expected total discounted reward for a policy is defined as

$$Q_\lambda(\mathbf{s}_0) = \mathbb{E}_{\mathbf{s}_0} \left\{ \sum_{n=1}^{\infty} \lambda^{n-1} r(\mathbf{s}(n), \mathbf{a}(n)) \right\}. \quad (20)$$

The optimality (or Bellman) equations for this problem are then given for all $\mathbf{s} \in \mathcal{S}$ by

$$Q_\lambda(\mathbf{s}) = \max_{\mathbf{a} \in \mathcal{A}_\mathbf{s}} \left\{ r(\mathbf{s}, \mathbf{a}) + \sum_{\tilde{\mathbf{s}} \in \mathcal{C}(\mathbf{s}, \mathbf{a})} \lambda p(\tilde{\mathbf{s}}|\mathbf{s}, \mathbf{a}) Q_\lambda(\tilde{\mathbf{s}}) \right\}. \quad (21)$$

Since we have stationary rewards and transition probabilities, finite state and action sets, and rewards are bounded, there exists an optimal deterministic stationary policy given by

$$\mathbf{a}_\mathbf{s}^* \in \mathcal{A}_\mathbf{s}^* = \arg \max_{\mathbf{a} \in \mathcal{A}_\mathbf{s}} \left\{ r(\mathbf{s}, \mathbf{a}) + \sum_{\tilde{\mathbf{s}} \in \mathcal{C}(\mathbf{s}, \mathbf{a})} \lambda p(\tilde{\mathbf{s}}|\mathbf{s}, \mathbf{a}) Q_\lambda^*(\tilde{\mathbf{s}}) \right\}, \quad (22)$$

where $Q_\lambda^*(\mathbf{j})$ is the solution to (21). The solution can be obtained, e.g., by iteratively solving (21) with the criterion for stopping after iteration n given by

$$\max_{\mathbf{s} \in \mathcal{S}} |Q_\lambda^{(n)}(\mathbf{s}) - Q_\lambda^{(n-1)}(\mathbf{s})| < \epsilon(1 - \lambda)/2\lambda, \quad (23)$$

for some small $\epsilon > 0$ [43, Section 6.3.2].

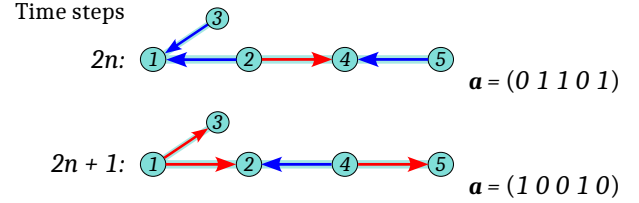


Fig. 6: TLS policy for an example IAB backhaul network.

C. Deterministic MDP for the Full-Buffer Scenario

To obtain an MDP with deterministic transitions, we adopt the worst-case full-buffer scenario and set $p_i^{\text{UL}} = p_i^{\text{DL}} = 1$ for all $i \in \mathcal{N} \setminus \{1\}$, which yields $\boldsymbol{\theta}(n) = (1, \dots, 1)$ for all n . Although this modification suffices to formalize the full-buffer backhaul MDP, we will redefine its affected components, namely, the state set and the transition probabilities.

Let the state set of the full-buffer backhaul MDP, $\mathcal{S}^{\text{FB}} \subset \mathcal{S}$, consist of all such vectors $\mathbf{s} = (\mathbf{s}^{\text{UL}}, \mathbf{s}^{\text{DL}})$ that can be obtained from state $(0, \dots, 0)$ by a finite sequence of controls $\mathbf{a}_1, \dots, \mathbf{a}_J$ with the disturbance $\boldsymbol{\theta} = (1, \dots, 1)$. Thus, for all $\mathbf{s} \in \mathcal{S}^{\text{FB}} \setminus \{0\}$ the queues receiving external arrivals cannot be empty: $s_i^{\text{UL}} > 0$, $i = 1, \dots, N-1$, and $s_{k,1}^{\text{DL}} > 0$, $k \in \mathcal{N} \setminus \{1\}$.

Transitions between states are deterministic, i.e., for any non-terminal $\mathbf{s} \in \bar{\mathcal{S}}_\tau^{\text{FB}}$ and $\mathbf{a} \in \mathcal{A}$ we have $|\mathcal{C}(\mathbf{s}, \mathbf{a})| = 1$. More specifically, upon action \mathbf{a} the process moves from state \mathbf{s} with probability 1 to state $\tilde{\mathbf{s}} = (\tilde{\mathbf{s}}^{\text{UL}}, \tilde{\mathbf{s}}^{\text{DL}})$ given by

$$\tilde{s}_i^{\text{UL}} = \max\{f_i(\mathbf{s}^{\text{UL}}, \mathbf{R}^{\text{UL}}(\mathbf{a})), 1\}, \quad i = 1, \dots, N-1, \quad (24)$$

and, for $k \in \mathcal{N} \setminus \{1\}$,

$$\tilde{s}_{k,i}^{\text{DL}} = \begin{cases} f_i(\mathbf{s}_k^{\text{DL}}, \mathbf{R}_k^{\text{DL}}(\mathbf{a})) & \text{if } i = 2, \dots, k-1, \\ \max\{f_1(\mathbf{s}_k^{\text{DL}}, \mathbf{R}_k^{\text{DL}}(\mathbf{a})), 1\} & \text{if } i = 1. \end{cases} \quad (25)$$

Thus, for $\mathbf{s} \in \mathcal{S}^{\text{FB}}$ and $\mathbf{a} \in \mathcal{A}_\mathbf{s}$, the transition probability is $p^{\text{FB}}(\tilde{\mathbf{s}}|\mathbf{s}, \mathbf{a}) = 1$ if $\tilde{\mathbf{s}}$ is given by the expressions above and zero otherwise. The collection of objects $\{\mathcal{T}, \mathcal{S}^{\text{FB}}, \mathcal{A}_\mathbf{s}, r_{\{\cdot\}}(\mathbf{s}, \mathbf{a}), p^{\text{FB}}(\tilde{\mathbf{s}}|\mathbf{s}, \mathbf{a})\}$ will be referred to as the full-buffer MDP.

D. Tree Level Switching (TLS) Fixed Control Policy

The deterministic MDP above with the reward function $r_{\text{avg}}(\mathbf{s}, \mathbf{a})$ yields a simple 2-step fixed control policy, which can be obtained for an arbitrary network as follows. Let, as previously, $N_k = |\mathcal{N}_k|$, $k \in \mathcal{N}$. Now, set for all $i \in \mathcal{N}$

$$a_i^{(1)} = \begin{cases} 1 & \text{if } N_i = 2m \text{ for some } m \in \mathbb{N}, \\ 0 & \text{if } N_i = 2m + 1 \text{ for some } m \in \mathbb{N}, \end{cases} \quad (26)$$

and $a_i^{(2)} = 1 - a_i^{(1)}$.

Then, $(\mathbf{a}^{(1)}, \mathbf{a}^{(2)})$ represents a fixed control policy that minimizes the maximum packet latency in the LL traffic transmission model under the full-buffer assumption and ensures that no packet experiences a latency exceeding

$$\tau = N_N + 1. \quad (27)$$

The resulting pattern is referred to as Tree level switching (TLS). Fig. 6 illustrates the policy for an example network. Fig. 7 shows the trajectory of the full-buffer MDP for the same

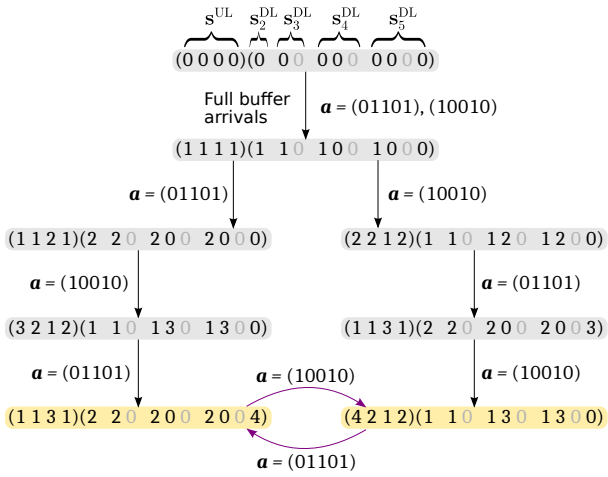


Fig. 7: The trajectory of the full-buffer backhaul MDP under the TLS fixed control policy starting from state 0. Starting with either control of the pattern, $\mathbf{a}^{(1)}$ or $\mathbf{a}^{(2)}$, the process ends up in a loop switching between two states (in yellow).

network. The latency threshold (27) stems from the following argument. Under the proposed control policy, each LL packet, once set to motion, travels non-stop because the links are enabled consecutively in either direction. Thus, the packet's travelling time along a flow k route equals the route's length $N_k - 1$. However, a packet can arrive at a "wrong" phase and wait to be set to motion. E.g., in Fig. 6 an uplink LL packet arriving to node 5 during time step 0, under the adopted assumptions, will join the buffer and not move further uplink at time 0. Then, at time 1, it cannot move either because the link is enabled in the opposite direction, so the packet remains in the buffer. At time 2, aged 2, the packet starts travelling and travels non-stop until it leaves the system. The waiting of $L = 2$ time steps is the worst-case waiting in either direction. This yields the worst-case flow k LL packet latency of $2 + N_k - 1 = N_k + 1$ and the worst-case backhaul latency of $N_N + 1$ for any LL packet, because $N_k \leq N_N$ and the pattern applies similarly to all routes.

Fig. 8 compares latency of TLS, MDP_{avg} , and MDP_{max}

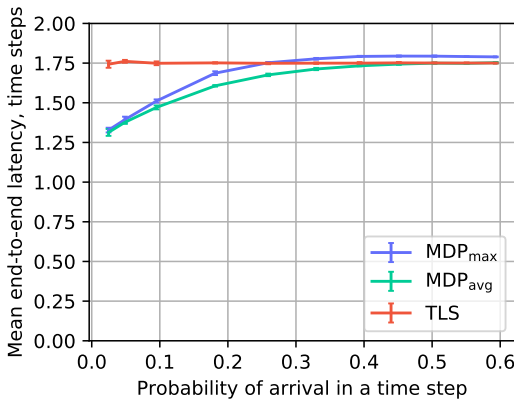


Fig. 8: Mean backhaul LL packet latency vs. the arrival probability p under MDP_{max} , MDP_{avg} and TLS control in the LL traffic transmission model.

controls. The results are presented for the topology of Fig. 10(a) with the arrival probabilities of $p_k^{\{\text{UL}, \text{DL}\}} = p$, i.e., the offered workload is evenly distributed across the routes and between uplink and downlink at each route. The figure shows the mean backhaul LL packet latencies as functions of p . The values are averages over 100 simulation runs, error bars indicate the corresponding standard deviations. The MDP-based solutions outperform the TLS fixed policy for lower workloads, where the controller can adapt the TDM pattern to suit individual arrivals. However, with the workload increasing, packets arriving to different routes uplink and downlink must be transmitted simultaneously. In these conditions, in terms of the mean latency, TLS outperforms MDP_{max} , whose focus is solely on the maximum latency, and approaches MDP_{avg} .

VI. LINK SCHEDULING FOR MIXED TRAFFIC

The TLS control policy derived in the previous section suits LL traffic, but it is not throughput-optimal. The throughput performance of a control policy is characterized by its *stability region* defined as the set of all packet arrival rates for which the network under this policy is stable, i.e., the time-average backlogs of its queues are finite. The network's *capacity region* is the set of arrival rates that the network can stably support using all possible control policies. *Throughput-optimal* policies stabilize the network whenever possible: their stability region and the capacity region of the network coincide [33].

TLS equally partitions the time resources between the uplink and downlink without adapting to the actual workload distribution. As a result, at certain workload levels TLS cannot stabilize the queues while a throughput-optimal dynamic control policy still can, although at the expense of computational and/or signaling delays. In this section we propose a customizable hybrid control policy that switches between TLS and a throughput-optimal dynamic control depending on the state of the buffers. Before formalizing our approach, we review candidate dynamic policies to be employed in such a setup. Their performance will be compared in Section VII.

A. Throughput-Oriented Control Policies

1) *Differential Backlogs*: The widely known throughput-optimal network control policy is the backpressure (BP) policy proposed in [33]. The central idea behind BP and similar capacity-maximizing control algorithms is to use differential queue backlogs. For the studied IAB network these can be expressed as follows.

Let the network links be indexed by the number of their incident nodes farther away from the donor. Denote by $Q_{l,k}^{\text{UL}}(n)$ the number of flow k packets (of both types) in the uplink queue facing link l at the beginning of time step n and let $Q_{1,k}^{\text{UL}}(n) = 0$ for all k and n . Similarly, denote by $Q_{l,k}^{\text{DL}}(n)$ the number of flow k packets in the downlink buffer facing node l . Then, the differential backlogs in the uplink are written as

$$\Delta Q_{l,k}^{\text{UL},\alpha}(n) = (Q_{l,k}^{\text{UL}}(n))^\alpha - (Q_{\max \mathcal{N}_l \setminus \{l\},k}^{\text{UL}}(n))^\alpha, \quad (28)$$

and in the downlink as

$$\Delta Q_{l,k}^{\text{DL},\alpha}(n) = \begin{cases} (Q_{l,k}^{\text{DL}}(n))^\alpha & \text{if } l = k, \\ (Q_{l,k}^{\text{DL}}(n))^\alpha - (Q_{m(l,k),k}^{\text{DL}}(n))^\alpha & \text{if } l \neq k, \end{cases} \quad (29)$$

where $m(l, k) = \min\{i \in \mathcal{N}_k : i > l\}$, $\alpha > 0$ is scaling factor.

2) *The Backpressure Policy*: The BP policy determines the action at time step n as

$$\mathbf{a}^* \in \arg \max_{\mathbf{a}} \sum_{l \in \mathcal{N} \setminus \{1\}} (W_l^{\text{UL}}(n) C_l^{\text{UL}}(n) \sum_{i < l} R_{i,l}(\mathbf{a}) + W_l^{\text{DL}}(n) C_l^{\text{DL}}(n) \sum_{j < l} R_{l,j}(\mathbf{a})), \quad (30)$$

where $C_l^X(n)$ is the capacity, in packets per time step, of link l in direction $X \in \{\text{UL}, \text{DL}\}$ and the links' weights, for $l \in \mathcal{N} \setminus \{1\}$, are given by

$$W_l^{\{\text{UL}, \text{DL}\}}(n) = \max \left\{ 0, \max_{k: l \in \mathcal{N}_k} \Delta Q_{l,k}^{\{\text{UL}, \text{DL}\}, 1}(n) \right\}. \quad (31)$$

Once an optimal action \mathbf{a}^* is determined and employed, the backpressure policy implies transmission over active links up to $C_l^{\{\text{UL}, \text{DL}\}}(n)$ packets of the flows whose differential backlogs, with $\alpha = 1$, yield the respective maxima in (31).

3) *$\alpha\beta$ Scheduling*: Modifications of the BP policy have been proposed to improve its latency performance. Thus, the $\alpha\beta$ link scheduling algorithm [35] computes, for each flow k and direction $X \in \{\text{UL}, \text{DL}\}$, the quantities

$$V_k^X(n) = \left(\sum_{l \in \mathcal{N}_k \setminus \{1\}} (Q_{l,k}^X(n))^{\alpha+1} \right)^{(\alpha+1)^{-1}} \quad (32)$$

and defines the link weights as

$$W_{l,k}^X(n) = \frac{(V_k(n))^{\beta-\alpha}}{(\lambda_k^X)^{\beta+1}} \Delta Q_{l,k}^{X,\alpha}(n), \quad (33)$$

$$W_l^X(n) = \max_{k: l \in \mathcal{N}_k} W_{l,k}^X(n), \quad (34)$$

where α and β are positive parameters. Then, the $\alpha\beta$ algorithm proceeds similarly to BP, uses (30) to determine an optimal action and serves the flows yielding the maximum weights.

4) *Per-Neighbor Queuing*: Both BP and $\alpha\beta$ are intended for networks with dynamic routing, where a packet flow can use multiple routes chosen dynamically at each time step. Both policies thus use per-flow queues in each node, which are served when the corresponding links are activated. However, it was shown in [37] that in networks with fixed routing, which is the case considered here, per-neighbor queues can be used instead of per-flow without notably degrading capacity performance. The authors of [37] stress that per-neighbor queuing, at the expense of a small loss in throughput, not only improves implementation scalability of the control policy but also substantially reduces the end-to-end packet latency.

The approach proposed in [37] suggests using the so called shadow queues – per-flow virtual packet counters – along with the per-neighbor queues storing real data packets. The shadow queues are offered traffic with rates larger than the real ones by a factor of $1 + \epsilon$ for some small $\epsilon > 0$. A throughput-optimal policy uses the shadow queues to determine optimal control, which is then applied to the real first-in first-out (FIFO) queues regardless of the flows, so shadow packets are interpreted as permits to transmit real packets.

In this work, we adopt the idea of using per-neighbor FIFO queuing while determining the control via a throughput-optimal policy based on per-flow backlogs.

B. Hybrid TLS+ Control

To combine the capacity gains provided by a throughput-oriented policy with the ultra-low-complexity and latency performance of TLS, for an IAB network with mixed traffic we propose a hybrid control method as follows. It is assumed that the IAB network controller can switch the nodes' modes in each time step and runs TLS by default. Whenever a node detects a packet age beyond a threshold δ , it switches to a dynamic throughput-oriented policy until the packet ages in all buffers are again below the threshold.

Note that the proposed approach implies the implementation in IAB of a dynamic throughput-oriented control policy such as BP along with TLS. The feasibility of BP-like control in IAB systems is due to the inherently distributed MT/DU design of the IAB-nodes and donor, where the buffer maps of the downstream nodes are explicitly signalled over the air interface. Upstream transmission requirements can be incorporated to IAB-specific BAP protocol.

The parameter δ allows to adjust the trade-off between computational efficiency and latency performance for workloads beyond the TLS stability region. In the next section, we employ simulation to compare the considered control policies, evaluate performance of hybrid TLS+BP and TLS+ $\alpha\beta$ and investigate the impact of the threshold δ value.

VII. NUMERICAL RESULTS

In this section we will provide our numerical results. After a brief description of the simulation setup and parameters, we compare performance of dynamic link scheduling policies with that of TLS. Then, we study the hybrid TLS+ approach and the impact of its threshold δ .

The example IAB topologies considered in this section are depicted in Fig. 10 and correspond to practical deployments in city centers such as Manhattan-like grids, Fig. 10(a), and along roads, Fig. 10(b). In what follows, they will be referred to as topologies A and B, respectively.

It is assumed that the number of arriving packets of each flow per time step follows Poisson distribution with mean $\lambda_k^{\text{UL}} = \lambda$ in uplink and $\lambda_k^{\text{DL}} = 4\lambda$ in downlink, so the network is downlink-dominated. The values of the other model parameters are given in Table I based on [48], [49]. The values

TABLE I: Default system parameters.

Parameter	Notation	Value
Number of layers	ν	8
Modulation order	Q_m	6
Overhead	ω	0.18
Bandwidth, MHz	B	200
Numerology	μ	3
Code rate	R_{\max}	948/1024
Resource blocks allocation	$N_{\text{PRB}}^{B,\mu}$	132
OFDM symbol duration, s	T_{symb}^{μ}	$10^{-3}/(14 \times 2^{\mu})$
Block error rate for LL	p_{LL}	0.01
Block error rate for BB	p_{BB}	0.1
Replication coding parameter for LL	r	3
Prob. for a packet to be low-latency	q_{LL}	0.05
Packet transmission size, bytes	σ	1500
Backhaul phase duration, ms	h	1
$\alpha\beta$ -scheduling policy parameters	(α, β)	(10, 25)

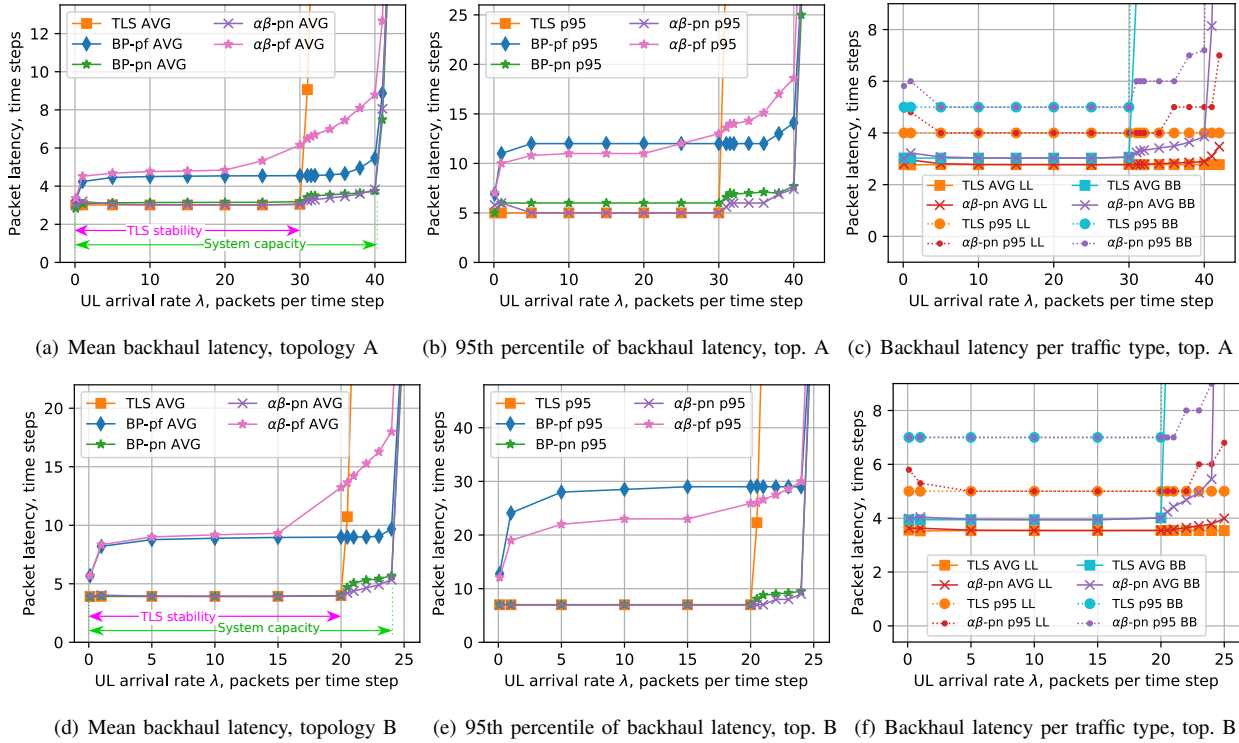


Fig. 9: Backhaul latency characteristics as functions of λ under TLS, BP and $\alpha\beta$ with per-flow (*pf*) and per-neighbor (*pn*) queuing in topologies A (the top row) and B (the bottom row). The plots to the left show the mean backhaul latency, the plots in the center the 95th percentile, and the plots to the right the mean and the 95th percentile separately for LL and BB traffic.

in Table I yield the backhaul links capacities of approximately 540 packets per time step. LL traffic constitutes 5 % of the total workload. The plots of this section show mean values averaged over 10 simulation runs.

We consider two versions of the BP and $\alpha\beta$ policies: the throughput-optimal with per-flow (*pf*) queuing and the approximate with per-neighbor (*pn*) queuing as described in Section VI-A4. The parameters of the $\alpha\beta$ scheduling algorithm

have been set to provide a better latency performance in the network of Fig. 10(a), namely, we let $\alpha = 10$ and $\beta = 25$.

A. Latency Performance of BP, $\alpha\beta$, and TLS

The plots in Fig. 9 show the mean and the 95th percentile of the backhaul latency as functions of the uplink flows' arrival rate λ . Note that in Fig. 9(a) and 9(d) we also indicate the capacity regions of the considered systems. Recall that the capacity region is defined as the range of the system load for which the queues can be stabilized.

By analyzing the results presented in Fig. 9 one can first observe that for the both considered topologies, A and B, TLS yields the smallest latency among the studied policies for a large range of λ . However, as TLS is not throughput-optimal, its latency metric soars before that of the other policies, thus indicating the boundary of the TLS stability region, which is inferior to the capacity of the system. Nevertheless, for both considered topologies the stability region of TLS, where it outperforms other policies, is rather large. Specifically, for topology A it is 75 % of the whole capacity region of the system, whereas for topology B it is approximately 83 %. Specifically, over these regions TLS outperforms BP-*pf* and $\alpha\beta$ -*pf* by approximately 70–100 % depending on the topology in terms of the mean latency, and by 150–300 % in terms of the 95th percentile. On the other hand, BP-*pn* and $\alpha\beta$ -*pn* show similar performance.

By examining the results further, we see that for the both topologies the $\alpha\beta$ algorithm with per-neighbor queuing ($\alpha\beta$ -*pn*) slightly outperforms BP-*pn*. This is why we use it as a

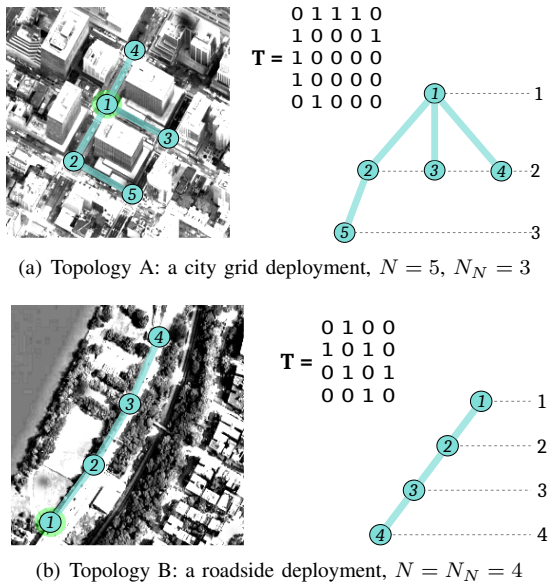


Fig. 10: Example topologies for numerical analysis.

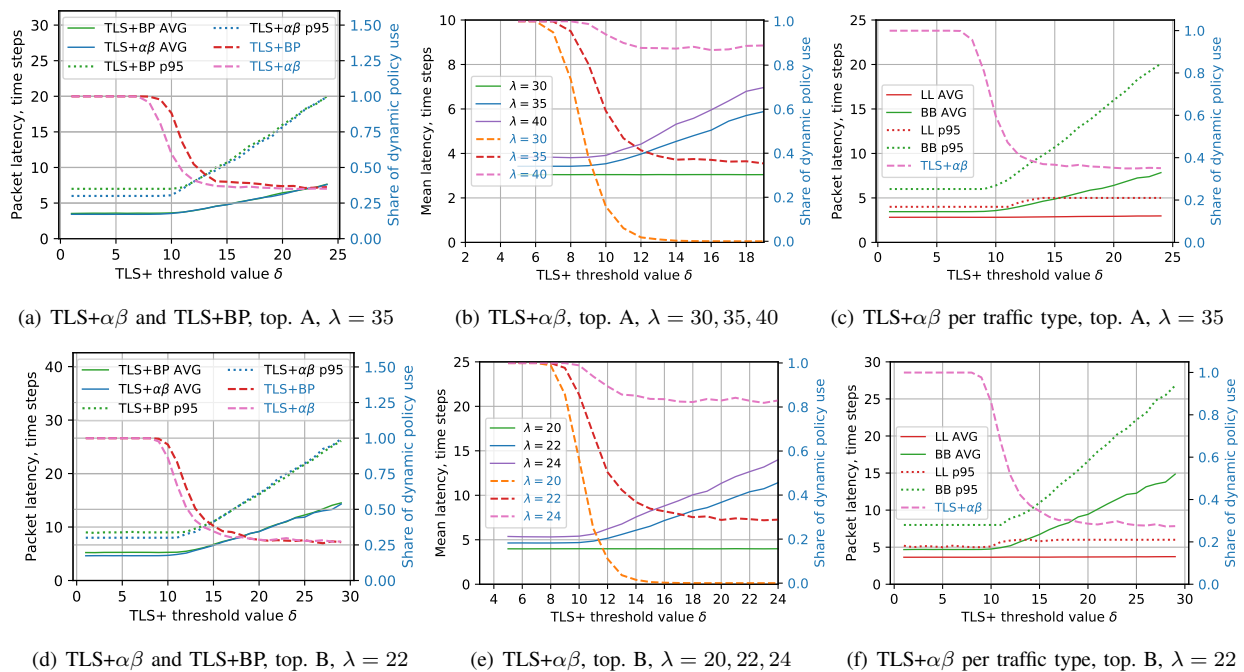


Fig. 11: Performance metrics of TLS+ as functions of the threshold δ in topologies A (the top row) and B (the bottom row). The plots to the left compare TLS+ $\alpha\beta$ and TLS+BP in terms of the mean backhaul latency (solid lines), its 95th percentile (dotted lines) and the share of time steps when the dynamic policy is used instead of TLS (dashed lines against the right-hand y-axis), for a single λ value. The plots in the center show the mean latency and the share of dynamic use under TLS+ $\alpha\beta$ for several workload levels. The plots to the right assess TLS+ $\alpha\beta$ in terms of mean latency and its 95th percentile separately for LL and BB traffic.

benchmark for TLS when analyzing latency per traffic type in Fig. 9(c) and 9(f). For the most part, the latency performance of $\alpha\beta$ -pn is comparable to that of TLS. However, the latency of LL traffic under TLS does not deteriorate as λ attains the system's capacity boundary and even beyond, which is not the case for the dynamic policies. This fact and the extent of the TLS stability region make this fixed policy a good candidate for systems prioritizing LL traffic and employing an admission control for the BB one.

B. Threshold δ of the Hybrid TLS+ Control

Having observed that the TLS control allows to minimize the latency over 75–85 % of the system capacity for both types of traffic, we now proceed investigating the dynamic hybrid TLS+ control allowing to extend the stability region of TLS. To this aim, we first assess the feasible range of the latency threshold δ serving as a trigger between the fixed TLS and the dynamic control component. Fig. 11 plots the backhaul latency characteristics (left-hand y-axis) and the proportion of time steps when the threshold is exceeded and thus the dynamic policy is used instead of TLS (right-hand y-axis) as a function of δ .

Fig. 11(a) and 11(d) detail the trade-off, for a given workload level, between the gain in computational efficiency indicated by the proportion of the dynamic policy use (dashed lines against the right-hand y-axis) and the latency performance. Note that the workload levels given by λ are fixed outside the stability region of TLS but inside the capacity region of the

system, see Fig. 9. We observe that the range, where variation of δ significantly affects the computational performance of TLS+ is rather short, e.g., if we look at TLS+ $\alpha\beta$ in Fig. 11(a), for $\delta \leq 7$ TLS is not used at all, for $8 \leq \delta \leq 10$ the use of TLS increases dramatically without latency performance degradation, and for $11 \leq \delta \leq 14$ the use of TLS still increases gradually, but the latency grows as well. An increase of δ beyond 14 does not provide any gain in TLS use but results in further degradation of latency performance.

In Fig. 11(a) and 11(d) we compare two policies, TLS+BP and TLS+ $\alpha\beta$, both with pn queuing. In our setup the action selection time for TLS averages at approximately 7×10^{-7} , whereas for BP it amounts to 2×10^{-3} and 6.2×10^{-4} , respectively, for topologies A and B, and for $\alpha\beta$ to 2.2×10^{-3} and 8.2×10^{-4} . It can be observed that TLS+ $\alpha\beta$ outperforms TLS+BP in both topologies in terms of the dynamic policy use percentage as well as latency. For this reason the remaining plots of Fig. 11 show results for TLS+ $\alpha\beta$ only.

Fig. 11(b) and 11(e) reveal that the range of δ where its variation significantly impacts performance does not change with the workload level. For λ in the stability region of TLS the use of TLS can be brought to 100 % without degrading latency performance. However, for λ outside the TLS stability region the use of the dynamic policy cannot be brought below a certain level, which increases as λ approaches the boundary of the system capacity.

Fig. 11(c) and 11(f) investigate the impact of δ on the latency metrics per traffic type. It can be observed that once δ stops affecting the proportion of TLS use, it also stops

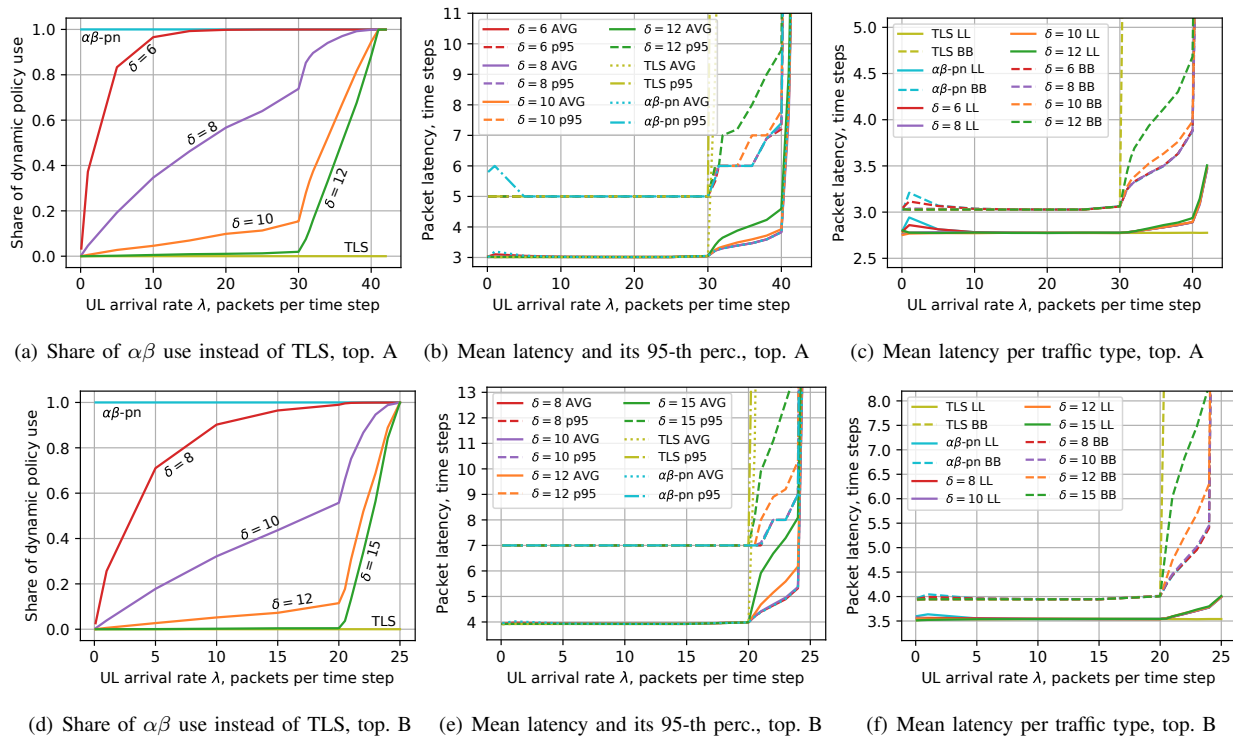


Fig. 12: Performance metrics of TLS+ $\alpha\beta$ as functions of λ for several values of δ in topologies A (the top row) and B (the bottom row). The metrics under TLS and $\alpha\beta$ -pn are provided for comparison. The plots to the left show the proportion of time steps when a dynamic policy is used instead of TLS. The plots in the center show the mean backhaul latency and its 95th percentile. The plots to the right present the mean backhaul latency per traffic type.

affecting the LL traffic performance and only the performance of BB traffic gets degraded further. Indeed, a further increase of δ for workloads beyond the stability region of TLS makes packets accumulate in some queues until their age reaches the threshold value, after which these queues get unloaded by a dynamic policy so that the packet age is again below the threshold. The accumulation of packets affects performance of BB traffic, while the control adjustment to unload the affected queues impacts LL traffic.

C. Performance of the TLS+ $\alpha\beta$ Control

We are now in position to explore the performance of the identified best control, TLS+ $\alpha\beta$. To this aim, Fig. 12 plots performance metrics of TLS+ $\alpha\beta$ as functions of λ for several values of δ . The metrics under standalone TLS and $\alpha\beta$ -pn are provided for comparison. Specifically, Fig. 12(a) and 12(d) show the proportions of dynamic policy use for topologies A and B, respectively. As we may observe for both topologies, if the workload level is in the TLS stability range, δ can be chosen such that the dynamic policy is used rarely.

By analyzing the results presented in Fig. 12(b) and 12(e), we see that, whatever the threshold δ in TLS+ $\alpha\beta$, the mean backhaul latency and its 95th percentile stay low and at the level of TLS for λ in the TLS stability region. However, for higher workloads, smaller δ yield lower latencies. Such a behaviour is, however, unavoidable as the system approaches the overloaded regime.

Finally, Fig. 12(c) and 12(f) plot the mean backhaul latency per traffic type. It can be observed that the $\alpha\beta$ algorithm and

TLS+ $\alpha\beta$ with small δ yield slightly larger latency for both traffic types under light workloads. Then, for the remainder of the TLS stability region the performance of the studied policies is similar. Finally, for workload levels beyond the TLS stability region $\alpha\beta$ and TLS+ $\alpha\beta$ yield similar latency performance for LL traffic, whereas for BB TLS+ $\alpha\beta$ with larger δ result in higher latency. Thus, we conclude that the hybrid TLS+ is needed mainly to control the latency of low-priority BB traffic.

VIII. CONCLUSIONS

Motivated by the need for a simple yet efficient link scheduling policy for half-duplex IAB systems which would reconcile the requirements of multiple traffic types, we have proposed the hybrid TLS+ control. By default, TLS+ operates in latency-oriented mode by utilizing the proposed fixed TLS policy which does not require signaling and minimizes the latency in low-to-moderate system load conditions. Whenever the packet age in the network attains a customizable threshold δ , the control switches to a more computationally challenging throughput-oriented dynamic policy capable of adapting to traffic bursts and unbalances. Depending on implementation needs, the dynamic policy in TLS+ can be chosen throughput-optimal, in order to exploit the full capacity region of the system, or reduced-complexity. The range of the threshold δ where its variation notably affects control performance is rather short and easy to expose.

The proposed TLS+ link scheduling represents a ready-to-use resource allocation solution at the level of link activation patterns. To fully exploit its benefits in practical IAB

deployments, it should be complemented with an efficient distributed yet coordinated packet-level scheduling of activated links, which is subject to future research.

REFERENCES

- [1] V. Raghavan, J. Li, A. Sampath, O. H. Koymen, and T. Luo, *Millimeter Wave Communications in 5G and Towards 6G*. CRC Press, 2024.
- [2] A. Shurakov, D. Moltchanov, A. Prikhodko, A. Khakimov, E. Mokrov, V. Begishev, I. Belikov, Y. Koucheryavy, and G. Gol'tsman, "Empirical blockage characterization and detection in indoor sub-THz communications," *Comput. Commun.*, vol. 201, pp. 48–58, 2023.
- [3] GSMA, *5G mmWave Guide: A Resource for Operators*, GSM Association, London, February 2022.
- [4] F. Agnoletto, P. Castells, E. Kolta, and D. Nichiforov-Chuang, *The economics of mmWave 5G: An assessment of total cost of ownership in the period to 2025*, GSMA Intelligence, London, 2021.
- [5] M. Cudak, A. Ghosh, A. Ghosh, and J. Andrews, "Integrated access and backhaul: A key enabler for 5G millimeter-wave deployments," *IEEE Commun. Magazine*, vol. 59, no. 4, pp. 88–94, 2021.
- [6] 3GPP, "5G; NR; NR and NG-RAN overall description; stage-2," ETSI, Technical Specification TS 138 300 V18.2.0 Release 18, Aug. 2024.
- [7] M. Polese, M. Giordani, T. Zugno, A. Roy, S. Goyal, D. Castor, and M. Zorzi, "Integrated Access and Backhaul in 5G mmWave networks: Potential and challenges," *IEEE Commun Magazine*, vol. 58, no. 3, pp. 62–68, March 2020.
- [8] 3GPP, "5G; NR; Integrated Access and Backhaul (IAB) radio transmission and reception," ETSI, Technical Specification TS 138 174 V18.4.0 Release 18, May 2024.
- [9] M. Gupta, A. Rao, E. Visotsky, A. Ghosh, and J. G. Andrews, "Learning link schedules in self-backhauled millimeter wave cellular networks," *IEEE Trans. Wireless Commun.*, vol. 19, no. 12, pp. 8024–8038, 2020.
- [10] M. Pagin, T. Zugno, M. Polese, and M. Zorzi, "Resource management for 5G NR Integrated Access and Backhaul: A semi-centralized approach," *IEEE Trans. Wireless Commun.*, vol. 21, pp. 753–767, 2022.
- [11] A. Zhivtsova, V. A. Beschastnyi, Y. Koucheryavy, and K. Samouylov, "A survey of delay-oriented dynamic link scheduling policies for 5G/6G Integrated Access and Backhaul systems," *IEEE Access*, vol. 12, 2024.
- [12] 3GPP, "5G; service requirements for the 5G system," ETSI, Technical Specification TS 122 261 V18.14.0 Release 18, July 2024.
- [13] N. Yarkina, D. Moltchanov, and Y. Koucheryavy, "Counter waves link activation policy for latency control in in-band IAB systems," *IEEE Commun. Lett.*, 2023.
- [14] C. Joo and N. B. Shroff, "Local greedy approximation for scheduling in multihop wireless networks," *IEEE Trans. Mobile Comput.*, vol. 11, no. 3, pp. 414–426, 2012.
- [15] 3GPP, "5G; NG-RAN; architecture description," ETSI, Technical Specification TS 138 401 V18.2.0 Release 18, Aug. 2024.
- [16] —, "NR; study on Integrated Access and Backhaul," 3GPP, Technical Report TR 38 874 V16.0.0 Release 16, Dec. 2018.
- [17] —, "5G; NR; Backhaul Adaptation Protocol (BAP) specification," ETSI, Technical Specification TS 138 340 V18.0.0, May 2024.
- [18] Y. Sadovaya, D. Moltchanov, W. Mao, O. Orhan, S.-p. Yeh, H. Nikopour, S. Talwar, and S. Andreev, "Integrated access and backhaul in millimeter-wave cellular: Benefits and challenges," *IEEE Commun. Magazine*, 2022.
- [19] V. F. Monteiro, F. R. M. Lima, D. C. Moreira, D. A. Sousa, T. F. Maciel, B. Makki, and H. Hannu, "Paving the way toward mobile IAB: Problems, solutions and challenges," *IEEE Open J. Commun. Soc.*, vol. 3, pp. 2347–2379, 2022.
- [20] S. Zhang, X. Xu, M. Sun, X. Tao, and C. Liu, "Joint spectrum and power allocation in 5G Integrated Access and Backhaul Networks at mmWave band," in *IEEE PIMRC 2020*, Aug. 2020, pp. 1–7.
- [21] Y. Sadovaya, D. Moltchanov, H. Nikopour, S.-p. Yeh, W. Mao, O. Orhan, S. Talwar, and S. Andreev, "Self-interference assessment and mitigation in 3GPP IAB deployments," in *IEEE ICC 2021*. IEEE, 2021, pp. 1–6.
- [22] P. Jayasinghe, A. Töllli, J. Kaleva, and M. Latva-Aho, "Traffic aware beamformer design for flexible TDD-based Integrated Access and Backhaul," *IEEE Access*, vol. 8, 2020.
- [23] B. Smida, A. Sabharwal, G. Fodor, G. C. Alexandropoulos, H. A. Suraweera, and C.-B. Chae, "Full-duplex wireless for 6G: Progress brings new opportunities and challenges," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 9, pp. 2729–2750, 2023.
- [24] C. Ghodhbane, M. Manini, P. Savelli, C. Gueguen, and X. Lagrange, "Load-efficiency-balance cell selection policy for IAB networks," in *IEEE PIMRC 2023*, 2023, pp. 1–6.
- [25] C. Madapatha, B. Makki, A. Muhammad, E. Dahlman, M.-S. Alouini, and T. Svensson, "On topology optimization and routing in Integrated Access and Backhaul networks: A genetic algorithm-based approach," *IEEE Open J. Commun. Soc.*, vol. 2, pp. 2273–2291, 2021.
- [26] H. Yin, S. Roy, and L. Cao, "Routing and resource allocation for IAB multi-hop network in 5G advanced," *IEEE Trans. Commun.*, vol. 70, no. 10, pp. 6704–6717, 2022.
- [27] N. Tafintsev, D. Moltchanov, S.-P. Yeh, N. Hosein, M. Wei, O. Orhan, S. Talwar, M. Valkama, and S. Andreev, "Joint path selection and resource allocation in multi-hop mmWave-based IAB systems," in *IEEE ICC WS 2022*, 2022, pp. 1–7.
- [28] A. H. Jazi, S. M. Razavizadeh, and T. Svensson, "Integrated Access and Backhaul (IAB) in cell-free massive MIMO systems," *IEEE Access*, vol. 11, 2023.
- [29] Y. Zhang, M. A. Kishk, and M.-S. Alouini, "Freshness-aware energy efficiency optimization for integrated access and backhaul networks," *IEEE Transactions on Wireless Communications*, 2024.
- [30] B. Zhang and I. Filippini, "Mobility-aware resource allocation for mmWave IAB networks: A multi-agent reinforcement learning approach," *IEEE/ACM Trans. Netw.*, vol. 32, no. 4, pp. 3559–3574, 2024.
- [31] A. A. Gargari, A. Ortiz, M. Pagin, W. de Sombre, M. Zorzi, and A. Asadi, "Risk-averse learning for reliable mmwave self-backhauling," *IEEE/ACM Transactions on Networking*, 2024.
- [32] M. Gupta, I. P. Roberts, and J. G. Andrews, "System-level analysis of full-duplex self-backhauled millimeter wave networks," *IEEE Trans. Wireless Commun.*, vol. 22, no. 2, pp. 1130–1144, 2023.
- [33] L. Tassiulas and A. Ephremides, "Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks," *IEEE Trans. Automatic Control*, vol. 37, no. 12, pp. 1936–1948, 1992.
- [34] —, "Dynamic server allocation to parallel queues with randomly varying connectivity," *IEEE Trans. Inf. Theory*, vol. 39, no. 2, 1993.
- [35] V. J. Venkataramanan, X. Lin, L. Ying, and S. Shakkottai, "On scheduling for minimizing end-to-end buffer usage over multihop wireless networks," in *Proc. IEEE INFOCOM 2010*, 2010.
- [36] Y. Cui, E. M. Yeh, and R. Liu, "Enhancing the delay performance of dynamic backpressure algorithms," *IEEE/ACM Trans. Netw.*, vol. 24, no. 2, pp. 954–967, 2016.
- [37] L. X. Bui, R. Srikant, and A. Stolyar, "A novel architecture for reduction of delay and queueing structure complexity in the back-pressure algorithm," *IEEE/ACM Trans. Netw.*, vol. 19, no. 6, pp. 1597–1609, 2011.
- [38] B. Ji, C. Joo, and N. Shroff, "Throughput-optimal scheduling in multihop wireless networks without per-flow information," *IEEE/ACM Trans. Netw.*, vol. 21, no. 2, pp. 634–647, 2013.
- [39] Z. Li, L. Xiang, X. Ge, G. Mao, and H.-C. Chao, "Latency and reliability of mmWave multi-hop V2V communications under relay selections," *IEEE Trans. Veh. Technol.*, vol. 69, no. 9, pp. 9807–9821, 2020.
- [40] N. Tafintsev, D. Moltchanov, S. Andreev, S.-p. Yeh, N. Himayat, Y. Koucheryavy, and M. Valkama, "Handling spontaneous traffic variations in 5G+ via offloading onto mmWave-capable UAV 'bridges'," *IEEE Trans. Veh. Technol.*, vol. 69, no. 9, pp. 10070–10084, 2020.
- [41] J. Ye and X. Ge, "Beam management optimization for V2V communications based on deep reinforcement learning," *Scientific Reports*, vol. 13, no. 1, 2023.
- [42] N. Tafintsev, D. Moltchanov, W. Mao, H. Nikopour, S.-P. Yeh, S. Talwar, M. Valkama, and S. Andreev, "Analysis of duplexing patterns in multi-hop mmWave Integrated Access and Backhaul systems," *IEEE Open J. Commun. Soc.*, vol. 5, pp. 5392–5407, 2024.
- [43] M. L. Puterman, *Markov decision processes: discrete stochastic dynamic programming*, ser. Wiley series in probability and statistics. Hoboken, NJ: John Wiley and Sons Inc., 2005.
- [44] Y. Cui, V. K. N. Lau, R. Wang, H. Huang, and S. Zhang, "A survey on delay-aware resource control for wireless systems-large deviation theory, stochastic Lyapunov drift, and distributed stochastic learning," *IEEE Trans. Inf. Theory*, vol. 58, no. 3, pp. 1677–1701, 2012.
- [45] S. Gopalam, S. V. Hanly, and P. Whiting, "Distributed resource allocation and flow control algorithms for mmWave IAB networks," *IEEE/ACM Trans. Netw.*, vol. 31, no. 6, pp. 3175–3190, 2023.
- [46] T.-K. Le, U. Salim, and F. Kaltenberger, "An overview of physical layer design for ultra-reliable low-latency communications in 3GPP releases 15, 16, and 17," *IEEE Access*, vol. 9, pp. 433–444, 2020.
- [47] S. E. Elayoubi, P. Brown, M. Deghel, and A. Galindo-Serrano, "Radio resource allocation and retransmission schemes for URLLC over 5G networks," *IEEE J. Sel. Areas Commun.*, vol. 37, pp. 896–904, 2019.
- [48] 3GPP, "5G; NR; User Equipment (UE) radio access capabilities," ETSI, Technical Specification TS 138 306 V17.0.0 Release 17, May 2022.
- [49] H. Holma, A. Toskala, and T. Nakamura, *5G technology: 3GPP New Radio*. John Wiley & Sons, 2020.