Path Selection and Rate Allocation in Self-Backhauled mmWave Networks

Trung Kien Vu*, Chen-Feng Liu*, Mehdi Bennis*, Mérouane Debbah†, and Matti Latva-aho*
*Centre for Wireless Communications, University of Oulu, Finland
†Mathematical and Algorithmic Sciences Lab, Huawei France R&D, Paris, France, and
†CentraleSupelec, Universite Paris-Saclay, Gif-sur-Yvette, France
E-mail: {trungkien.vu, chen-feng.liu, mehdi.bennis, matti.latva-aho}@oulu.fi, merouane.debbah@huawei.com

***Abstract*—We investigate the problem of multihop scheduling in self-backhaul mmWave networks in which self-backhauled small cells are considered. Owing to the high path loss of mmWave, multihop routes between the macro base station and the intended users via small cells need to be carefully selected. This paper addresses the fundamental question: "*how to select the best paths and to allocate rates over paths subject to latency constraints with a guaranteed probability*?". To answer this question, we propose a new system design, which factors in channel variations and network dynamics. The problem is cast as a network utility maximization subject to a bounded delay constraint with a guaranteed probability and network stability. By leveraging the stochastic optimization, the studied problem is then decoupled into path/route selection and rate allocation, whereby learning the best paths is done by means of a *regret learning* algorithm which builds the probability distribution of the best paths. Via numerical results, our approach ensures reliable communication with a guaranteed probability of 99.9999%, and reduces latency by 50.64% and 92.9% as compared to *baselines* with and without learning, respectively.**

I. Introduction

In order to provide reliable communication with an over-the-air latency of few milliseconds and extreme throughput, a number of candidate solutions are currently investigated for 5G: 1) higher frequency spectrum, e.g., centimeter and millimeter waves (mmWaves); 2) advanced spectral-efficient techniques, e.g., massive multiple-input multiple-output (MIMO); and 3) ultra-dense self-backhauled small cell deployments [1].

In this paper, we are motivated by the combination of the above techniques, which holds the promise of providing great enhancements of the overall system performance [1], [2]. To do so, an in-band wireless backhauling solution allows to deploy such ultra-dense small cells [2], [3], in which massive MIMO and mmWave are combined to provide the wireless backhaul since mmWave frequency bands offer huge bandwidth to meet the exponentially growing traffic demands [3], [4]. However, operating at higher frequency bands experiences high propagation attenuation [4], which requires packing more antennas to achieve highly directional gain [5]. Hence, mmWave and massive MIMO are combined together [4]. However, mmWave communication requires higher transmit power and is very sensitive to blockage, when transmitting over a long distance [3], [4]. Hence, instead of using a single hop [3], [6], a multihop self-backhauling architecture is a promising solution [7], [8]. The authors in [9] studied the multihop routing for device-to-device communication, focusing on maximizing the quality for multimedia applications. More importantly, allowing multihop transmissions raises a problem of increased delay, which has been ignored. Hence, there is a need for fast and efficient multihop scheduling with respect to traffic dynamics and channel variance in self-backhauled mmWave networks [7].

Main contributions: Considering a multihop self-backhauled mmWave network, we propose a system design to ensure ultra-reliable and low latency communication (URLLC). In particular, our goal is to maximize a general network utility subject to a probabilistic delay constraint and network stability. Leveraging the stochastic optimization, the problem is decoupled into multihop path/route selection and rate allocation. The proposed approach answers the following questions: (i) *over which paths the traffic flow should be forwarded*? and (ii) *what is the data rate per flow/subflow* while ensuring low-latency and ultra-reliability constraints?

We study a scenario where statistical system information is *not* available, that makes the problem of path selection very difficult. Instead, based on the historical system information, we leverage *regret learning* techniques to build an empirical distribution of the system dynamics to aid in selecting the best paths. The *regret* is defined as the difference between the average utility when choosing the same paths in previous times, and its average utility obtained by constantly selecting different paths. The idea is that the regret is minimized over time so as to choose the best paths [10].

Related work: Path selection and multipath congestion control are well-studied in [11] in which the aggregate utility is increased as more paths are provided. However, splitting data into too many paths leads to increased signaling overhead and makes the traffic congested. Moreover, [11] did *not* consider the problem of providing URLLC. In this paper, we determine the best paths to maximize the network throughput subject to the delay bound violation constraint with a tolerable probability (reliability). A recent work in [12] has studied the multi-hop relaying transmission challenges for mmWave systems, aiming at maximizing the network throughput, while taking account into traffic dynamics and link qualities. In our work, we also model the network utility maximization, while considering the channel variations and network dynamics. More importantly, we further address two fundamental questions: (i) how to select the best paths while taking account into traffic dynamics and link qualities and (ii) how to capture the URLLC while maximizing the network utility in multi-hop self-backhauled mmWave networks. While our previous work [6] studied URLLC-centric mmWave networks for single hop transmission, in this work we extend it to the multihop wireless backhaul scenario and study a joint path selection and rate allocation optimization.

II. System Model

Let us consider a multi-hop heterogeneous cellular network (HCN) which consists of a macro base station (MBS), a set of

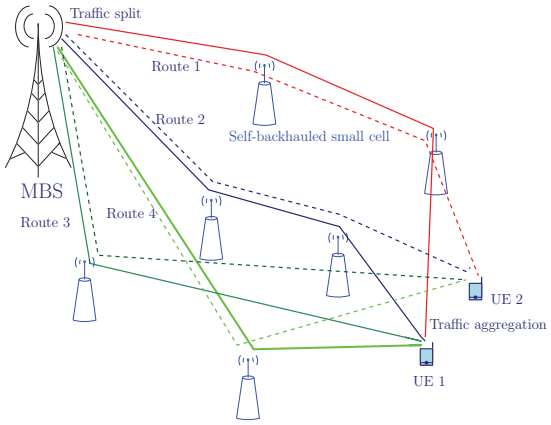

Fig. 1: 5G multihop self-backhauled mmWave networks.

B self-backhauled small cell base stations (SCBSs), and a set $\mathcal{K}$ of UEs K single-antenna user equipments (UEs) as shown in Fig 1. Let $\mathcal{B} = \{0, 1, \cdots, B\}$ denote the set of all BSs in which index 0 refers to the MBS. We consider the downlink transmission in which the traffic is generated from the MBS to UEs via self-backhauled SCBSs. The in-band wireless backhaul is used to provide backhaul among base stations (BSs). Each BS is equipped with N_b transmitting antennas and we denote the propagation channel between BS b and UE k as $\mathbf{h}_{(b,k)} = \sqrt{N_b}\mathbf{\Theta}_{(b,k)}^{1/2}\tilde{\mathbf{h}}_{(b,k)}$ [3], where $\mathbf{\Theta}_{(b,k)} \in \mathbb{C}^{N_b \times N_b}$ depicts the antenna spatial correlation, and the elements of $\tilde{\mathbf{h}}_{(b,k)} \in \mathbb{C}^{N_b \times 1}$ are independent and identically distributed (i.i.d.) with zero mean and variance $1/N_b$.

The network topology is modeled as a directed graph $\mathcal{G} = (\mathcal{N}, \mathcal{L})$, where $\mathcal{N} = \mathcal{B} \cup \mathcal{K}$ represents the set of nodes including BSs and UEs. $\mathcal{L} = \{(i,j) | i \in \mathcal{B}, j \in \mathcal{N}\}$ denotes the set of all directional edges (i,j) in which nodes i and j are the transmitter and the receiver, respectively.

We consider a (stochastic) queueing network operating in discrete time $t \in \mathbb{Z}^+$ [13]. There are F independent data at the MBS. Each data traffic is destined for one UE, whereas one UE receives multiple data streams, i.e., $F \geq K$. Hereafter, we refer to data traffic as data flow. We use $\mathcal{F}$ to represent the set of F data flows. The MBS splits each flow f into multiple sub-flows which are sent through a set of disjoint paths. The traffic aggregation capability is assumed at the UEs [14].

We assume that there exits Z_f number of disjoint routes from the MBS to the UE for flow f. For any disjoint route $m = \{1, \cdots, Z_f\}$, we denote $\mathcal{Z}_f^m$ as the route state, which contains all route information such as topology and queue states for every hop. Let $\mathcal{Z}_f = \{\mathcal{Z}_f^1, \cdots, \mathcal{Z}_f^m, \cdots \mathcal{Z}_f^{Z_f}\}$ denote route states observed by flow f. We use the flow-split indicator vector $\mathbf{z}_f = (z_f^1, \cdots, z_f^{Z_f})$ to denote how the MBS splits flow f, where $z_f^m = 1$ means path m is used to send data for flow f. Otherwise, $z_f^m = 0$. Let $\mathcal{N}_i^{(\mathrm{o})}$ denote the set of the next hops from node i via a directional edge. We denote the next hop of flow f from BS b as $b_f^{(\mathrm{o})}$.

In addition, $\mathbf{h} = \left(\mathbf{h}_{(i,j)} | (i,j) \in \mathcal{L}\right)$ consists of the channel propagations and we denote $p_{(i,j)}^f$ as the transmit power of node i assigned to node j for flow f, such that $\sum_{f \in F} \sum_{j \in \mathcal{N}_i^{(\mathrm{o})}} p_{(i,j)}^f \leq P_i^{\max}$, where $P_i^{\max}$ is the maximum transmit power of node i. We have the power constraint as

$$\mathcal{P} = \left\{ p_{(i,j)}^f \geq 0, i, j \in \mathcal{N}, \Big| \sum_{f \in \mathcal{F}} \sum_{j \in \mathcal{N}_i^{(\mathrm{o})}} p_{(i,j)}^f \leq P_i^{\max} \right\}. \quad (1)$$

Vector $\mathbf{p} = (p_{(i,j)}^f | \forall i, j \in \mathcal{N}, \forall f \in \mathcal{F})$ denotes the transmit power over all flows. Therefore, for a given channel state and transmit power, the data rate in edge (i,j) over flow f can be posted as a function of channel state and transmit power, i.e., $R_{(i,j)}^f(\mathbf{h}, \mathbf{p}) = \log_2\left(1 + \frac{p_{(i,j)}^f |\mathbf{h}_{(i,j)}|^2}{\sum_{f' \in \mathcal{F}} \sum_{i' \neq i} p_{(i',j)}^{f'} |\mathbf{h}_{(i',j)}|^2 + 1}\right)$, such that $\sum_{f \in \mathcal{F}} R_{(i,j)}^f = R_{(i,j)}$. We denote $\mathbf{R} = (R_{(i,j)}^f | \forall i, j \in \mathcal{N}, \forall f \in \mathcal{F})$ as a vector of data rates over all flows.

Let $Q_b^f(t)$ denote the queue length[1] at BS b at time slot t for flow f. The evolution of the queue length at the MBS is

$$Q_0^f(t+1) = \left[Q_0^f(t) - \sum_{m=1, 0_f^{(\mathrm{o})} \in \mathcal{Z}_f^m}^{Z_f} z_f^m R_{(0,0_f^{(\mathrm{o})})}^f(t), 0 \right]^+ + \mu^f(t). \quad (2)$$

where $[Q]^+ \triangleq \max\{Q, 0\}$ and $\mu^f(t)$ is the data arrival at the MBS during slot t which is independent and identical distributed (i.i.d.) over time with the mean value $\bar{\mu}^f$ [13]. Due to the disjoint paths, the incoming rate at the SCBS is either from one SCBS or the MBS, which is denoted as $b^{(\mathrm{I})}$. The evolution of the queue length at the SCBS b is given by

$$Q_b^f(t+1) \leq \left[Q_b^f(t) - R_{(b,b_f^{(\mathrm{o})})}^f(t), 0 \right]^+ + R_{(b^{(\mathrm{I})},b)}^f(t). \quad (3)$$

III. Problem Formulation

Assume that the MBS determines which routes to split a given flow f with a given probability distribution, i.e., $\boldsymbol{\pi}_f = \left(\pi_f^1, \cdots, \pi_f^{Z_f} | \pi_f^m = \Pr(z_f = z_f^m)\right)$. Here, $\boldsymbol{\pi}_f$ is the probability mass function (PMF) of the flow-split vector, i.e., $\sum_{m=1}^{Z_f} \Pr(z_f^m) = 1$. We denote $\boldsymbol{\pi} = \{\boldsymbol{\pi}_1, \cdots, \boldsymbol{\pi}_f, \cdots, \boldsymbol{\pi}_F\} \in \Pi$ as the global probability distribution of all flow-split vectors in which Π is the set of all possible global PMFs. Let $\bar{x}_0^f$ denote the achievable average rate of flow f, where $\bar{x}_0^f \triangleq \lim_{t \to \infty} \frac{1}{t} \sum_{\tau=1}^{t} \sum_{m=1, 0_f^{(\mathrm{o})} \in \mathcal{Z}_f^m}^{Z_f} \mathbb{E}\left[z_f^m R_{(0,0_f^{(\mathrm{o})})}^f(\tau)\right]$. Vector $\bar{\mathbf{x}} = (\bar{x}_0^1, \cdots, \bar{x}_0^F)$ denotes the time average of rates over all flows. Let $\mathcal{R}$ denote the rate region, which is defined as the convex hull of the average rates, i.e., $\bar{\mathbf{x}} \in \mathcal{R}$.

We define U_0 as a network utility function, i.e., $U_0(\bar{\mathbf{x}}) = \sum_{f \in \mathcal{F}} U\left(\bar{x}_0^f\right)$ [11]. Here, $U(\cdot)$ is assumed to be a twice differentiable, concave, and increasing L-Lipschitz function for all $\bar{\mathbf{x}} \geq 0$. According to Little's law [15], the queuing delay is defined as the ratio of the queue length to the average arrival rate. By taking into account the probabilistic bounded delay constraint for each flow/subflow, the following network utility maximization (NUM) is formulated:

$$\text{OP:} \max_{\boldsymbol{\pi}, \mathbf{x}, \mathbf{p}} \quad U_0(\bar{\mathbf{x}}) \quad (4a)$$

$$\text{subject to} \quad \Pr\left(\frac{Q_b^f(t)}{\bar{\mu}_f} \geq \beta\right) \leq \epsilon, \forall t, f \in \mathcal{F}, b \in \mathcal{B}, \quad (4b)$$

$$\lim_{t \to \infty} \frac{\mathbb{E}\left[|Q_b^f|\right]}{t} = 0, \forall f \in \mathcal{F}, \forall b \in \mathcal{B}, \quad (4c)$$

$$\mathbf{x}(t) \in \mathcal{R}, \quad (4d)$$

$$\boldsymbol{\pi} \in \Pi, \quad (4e)$$

$$\text{and (1)},$$

where $\Pr(\cdot)$ denotes the probability operator, β reflects the maximum allowed delay requirement for UEs, and $\epsilon \ll 1$ is the target probability for reliable communication. The

[1]The queues at the UEs are assumed to be empty and to be removed completely from network.

probabilistic delay constraint (4b) implies the probability that the delay for each flow at node b is greater than β is very small, which captures the constraints of ultra-low latency and reliable communication. It is also used to avoid congestion for each flow f at any point (BS) in the network, if the queue length is greater than $\beta\bar{\mu}^f$. More importantly, (4b) forces the transmission of all BSs, and (4c) maintains network stability.

Constraints relaxation: The above problem has a non-linear probabilistic constraint (4b), which cannot be solved directly. Hence, we replace the non-linear constraint (4b) with a linear deterministic equivalent by applying Markov's inequality [6] in which we have $\Pr(X \geq a) \leq \mathbb{E}[X]/a$ for a non-negative random variable X and $a > 0$, where $\mathbb{E}[\cdot]$ denotes the expectation operator. Thus, we relax (4b) as

$$\mathbb{E}\big[Q_b^f(t)\big] \leq \bar{\mu}^f \epsilon\beta. \tag{5}$$

Assuming that $\mu^f(t)$ follows a Poisson arrival process, we derive the expected queue length in (2) for the MBS, i.e.,

$$\mathbb{E}[Q_0^f(t)] = t\bar{\mu}^f - \sum_{\tau=1}^{t} \sum_{m=1,0_f^{(\mathrm{o})} \in \mathcal{Z}_f^m} \pi_f^m z_f^m R^f_{(0,0_f^{(\mathrm{o})})}(\tau), \tag{6}$$

and the expected queue length in (3), for each SCBS, i.e.,

$$\mathbb{E}[Q_b^f(t)] = \sum_{\tau=1}^{t} \sum_{m} \pi_f^m z_f^m \left(R^f_{(b^{(\mathrm{I})},b)}(\tau) - R^f_{(b,b_f^{(\mathrm{o})})}(\tau) \right). \tag{7}$$

Subsequently, combining the URLLC constraint (5) and (6), we obtain, for the MBS,

$$\begin{aligned} &\bar{\mu}^f(t-\epsilon\beta) - \sum_{\tau=1}^{t-1} \sum_{m=1,0_f^{(\mathrm{o})} \in \mathcal{Z}_f^m} \pi_f^m z_f^m R^f_{(0,0_f^{(\mathrm{o})})}(\tau) \\ &\leq \sum_{m=1,0_f^{(\mathrm{o})} \in \mathcal{Z}_f^m} \pi_f^m z_f^m R^f_{(0,0_f^{(\mathrm{o})})}(t). \end{aligned} \tag{8}$$

Similarly, for each SCBS $b = \{1, \cdots, B\}$, we have

$$\begin{aligned} &-\bar{\mu}^f \epsilon\beta + \sum_{\tau=1}^{t-1} \sum_{m} \pi_f^m z_f^m \left(R^f_{(b^{(\mathrm{I})},b)}(\tau) - R^f_{(b,b_f^{(\mathrm{o})})}(\tau) \right) \\ &\leq \sum_{m} \pi_f^m z_f^m \left(R^f_{(b,b_f^{(\mathrm{o})})}(t) - R^f_{(b^{(\mathrm{I})},b)}(t) \right), \end{aligned} \tag{9}$$

by combining (5) and (7). With the aid of the above derivations, we consider (8) and (9) instead of (4b) in the original problem (4).

In practice, since the statistical information of all candidate paths to decide $\boldsymbol{\pi}_f, \forall\, f \in \mathcal{F}$, is not available beforehand, that would be very challenging when solving (4). One solution is that paths are randomly assigned to each flow which does not guarantee optimality, whereas applying the high-complexity exhaustive search is not practical. Therefore, we propose a low-complexity approach by invoking Lyapunov stochastic optimization in which the local optimal is obtained [13] .

IV. Learning-Aided Path Selection and Rate Allocation

Let us start by rewriting (4) equivalently as [13]

$$\begin{aligned} \text{RP:} \max_{\bar{\boldsymbol{\varphi}},\boldsymbol{\pi},\mathbf{p}} \quad & U_0(\bar{\boldsymbol{\varphi}}) && \text{(10a)} \\ \text{subject to} \quad & \bar{\varphi}_0^f - \bar{x}_0^f \leq 0,\ \forall f \in \mathcal{F}, && \text{(10b)} \\ & \text{(1), (4c), (4e), (8), (9),} \end{aligned}$$

where the new constraint (10b) is introduced to replace the rate constraint (4d) with new auxiliary variables $\boldsymbol{\varphi} = \big(\varphi_0^1, \cdots, \varphi_0^F\big)$. In (10b), $\bar{\boldsymbol{\varphi}} \triangleq \lim_{t\to\infty} \frac{1}{t} \sum_{\tau=1}^{t} \mathbb{E}\left[|\boldsymbol{\varphi}(\tau)|\right]$. In order to ensure the inequality constraint (10b), we introduce a virtual queue vector $Y_0^f(t)$, which is given by

$$Y_0^f(t+1) = \left[Y_0^f(t) + \varphi_0^f(t) - x_0^f(t) \right]^+, \forall f \in \mathcal{F}. \tag{11}$$

Then, we write the conditional Lyapunov drift-plus-penalty for slot t as

$$\mathbb{E}\left[L\left(\Sigma(t+1)\right) - L\left(\Sigma(t)\right) | \Sigma(t) \right] - \nu U_0\left(\bar{\boldsymbol{\varphi}}\right), \tag{12}$$

where $L(\Sigma(t)) \triangleq \frac{1}{2}\left[\sum_{f=1}^{F} \sum_{b=0}^{B} Q_b^f(t)^2 + \sum_{f=1}^{F} Y_0^f(t)^2 \right]$ is the quadratic Lyapunov function of the queue backlogs $\Sigma(t) = (\mathbf{Q}(t), \mathbf{Y}(t))$ [13]. Here, ν is a control parameter, which is chosen to trade off utility optimality and queue length. Note that the stability of $\Sigma(t)$ assures that the constraints of problem (4c) and (10b) are held. Subsequently, following the straightforward calculations of the Lyapunov optimization which are omitted here for space, we obtain

$$\begin{aligned} (12) \quad \leq \quad & \sum_{f=1}^{F} \sum_{b=1}^{B} \sum_{m} \pi_f^m z_f^m Q_b^f \left(R^f_{(b^{(\mathrm{I})},b)} - R^f_{(b,b_f^{(\mathrm{o})})} \right) \\ & - \sum_{f=1}^{F} \sum_{m=1,0_f^{(\mathrm{o})} \in \mathcal{Z}_f^m} \pi_f^m z_f^m Q_0^f R^f_{(0,0_f^{(\mathrm{o})})} \qquad (13) \\ & + \sum_{f=1}^{F} \left[Y_0^f \varphi_0^f - \nu U\left(\varphi_0^f\right) - Y_0^f x_0^f \right] + C. \end{aligned}$$

Due to space limitation, we omit the details of the constant C, which does not influence the system performance [13].

Decoupling problems: The solution to (10) can be obtained by minimizing the upper bound in (13) in which we have three decoupled subproblems as follows: The flow-split vector and the probability distribution are determined by

$$\begin{aligned} \text{SP1:} \min_{\boldsymbol{\pi},\mathbf{z}} \quad & \sum_{f=1}^{F} \Xi_f \\ \text{subject to} \quad & \text{(4e),} \end{aligned}$$

where

$$\begin{aligned} \Xi_f \quad = \quad & \sum_{b=1}^{B} \sum_{m} \pi_f^m z_f^m Q_b^f \left(R^f_{(b^{(\mathrm{I})},b)} - R^f_{(b,b_f^{(\mathrm{o})})} \right) \\ & - \sum_{m=1,0_f^{(\mathrm{o})} \in \mathcal{Z}_f^m} \pi_f^m z_f^m Q_0^f R^f_{(0,0_f^{(\mathrm{o})})}. \end{aligned}$$

Then, we select the optimal auxiliary variables by solving

$$\begin{aligned} \text{SP2:} \min_{\boldsymbol{\varphi}|\mathbf{z}} \quad & \sum_{f=1}^{F} \left[Y_0^f \varphi_0^f - \nu U\left(\varphi_0^f\right) \right] \\ \text{subject to} \quad & \varphi_0^f(t) \geq 0,\ \forall f \in \mathcal{F}. \end{aligned}$$

Finally, the transmit power is decided by

$$\begin{aligned} \text{SP3:} \min_{\mathbf{x},\mathbf{p}|\mathbf{z}} \quad & \sum_{f=1}^{F} -Y_0^f x_0^f \\ \text{subject to} \quad & \text{(1), (8), (9).} \end{aligned}$$

A. Learning-Aided Route Selection

To select the optimal routes in SP1, we leverage regret learning which exploits the historical system information such as queue state and channel state [10]. The intuition behind this approach is that the regret learning method results in maximizing the long-term utility for each flow, by leveraging stochastic optimization [13], which allows to design the optimal utility-delay tradeoff.

Recall that $\mathbf{z}_f$ represents the flow-split vector given to flow f and $z_f^m = 1$ means path m is used to send data for flow f. The MBS determines routes for each flow with a given probability (mixed strategy). The optimal strategies mean that the MBS does not wish to change its strategy for any flow where any deviation does not offer better utility gain for all flows. We denote $u_f^m = u_f\left(z_f^m, \mathbf{z}_f^{-m}\right)$ as an utility function of flow f when using path m. The vector $\mathbf{z}_f^{-m}$ denotes the flow-split vector excluding path m. The MBS probably chooses more than one path to deliver data, from SP1, the utility gain of flow f is

$$u_f = \sum_m u_f^m = -\Xi_f.$$

To exploit the historical information, the MBS determines a flow-split vector for each flow f from $\mathcal{Z}_f$ based on the PMF from the previous stage $t-1$, i.e.,

$$\boldsymbol{\pi}_f\left(t-1\right) = \left(\pi_f^1\left(t-1\right), \cdots, \pi_f^m\left(t-1\right) \cdots, \pi_f^{Z_f}\left(t-1\right)\right). \tag{14}$$

Here, we define $\mathbf{r}_f(t) = \left(r_f^1\left(t\right), \cdots, r_f^m\left(t\right) \cdots, r_f^{Z_f}\left(t\right)\right)$ as a regret vector of determining flow-split vector for flow f. For a given flow-split vector associated to each flow, a positive regret means each flow could obtain a higher payoff with this flow-split vector during the previous time instant. Otherwise, it implies that the MBS has no regret to determine that flow-split vector for the flow. Hence, the MBS tends to determine the flow-split vector with highest regret in which the mixed-strategy probability is given as

$$\pi_f^m\left(t\right) = \frac{\left[r_f^m\left(t\right)\right]^+}{\sum_{m' \in \mathcal{Z}_f}\left[r_f^{m'}\left(t\right)\right]^+}. \tag{15}$$

In addition, the MBS is allowed to change its decision (flow-split vectors) for all flows over time yielding the highest regrets and exploring other decisions with lower regrets with a non-zero probability. To capture such behavior, we introduce the Boltzmann-Gibbs (BG) distribution, $\boldsymbol{\beta}_f^m\left(\tilde{\mathbf{r}}_f(t)\right)$, which is the solution of the following optimization problem:

$$\begin{aligned}\boldsymbol{\beta}_f^m\left(\tilde{\mathbf{r}}_f(t)\right) = \underset{\boldsymbol{\pi}_f \in \Pi}{\operatorname{argmax}} \sum_{m \in \mathcal{Z}_f} & \left[\pi_f^m\left(t\right)\tilde{r}_f^m\left(t\right)\right. \\ & \left. -\kappa_f \pi_f^m\left(t\right)\ln(\pi_f^m\left(t\right))\right],\end{aligned} \tag{16}$$

where $\tilde{\mathbf{r}}_f(t) = \left(\tilde{r}_f^1\left(t\right), \cdots, \tilde{r}_f^m\left(t\right) \cdots, \tilde{r}_f^{Z_f}\left(t\right)\right)$ is the estimated regret vector of flow f, and the tradeoff factor κ_f is used to balance between exploration and exploitation. Choosing a small value of κ_f, the MBS tends to decide routes for each flow with highest regret and does not change its routes in the future. While the tradeoff factor goes to infinity, all routes having a uniform distribution are equally chosen over time. Hence, by judiciously considering the mixed-strategy regret and the tradeoff factor, we obtain a mixed strategy probability that exploits certain routes to maximize the expected utility while exploring the rest of routes. Hence, the MBS tends to maximize the long-term utility of each flow $\bar{u}_f(t) = \frac{1}{t}\sum_{\tau=1}^{t} u_f(\tau)$.

For given set of $\tilde{\mathbf{r}}_f(t)$ and κ_f, we solve (16) to find the probability distribution in which the solution determining the disjoint routes for each flow f is given as

$$\beta_f^m(\tilde{\mathbf{r}}_f(t)) = \frac{\exp\left(\frac{1}{\kappa_f}\left[\tilde{r}_f^m\left(t\right)\right]^+\right)}{\sum\limits_{m' \in \mathcal{Z}_f}\exp\left(\frac{1}{\kappa_f}\left[\tilde{r}_f^{m'}\left(t\right)\right]^+\right)}. \tag{17}$$

We denote $\tilde{u}\left(t\right)$ as the estimated utility of flow f at time instant t with action $\mathbf{z}_f$, i.e, $\tilde{\mathbf{u}}_f(t) = \left(\tilde{u}_f^1\left(t\right), \cdots, \tilde{u}_f^m\left(t\right) \cdots, \tilde{u}_f^{Z_f}\left(t\right)\right)$. In addition, $\hat{u}_f(t)$ denotes the utility observed by flow f, i.e., $\hat{u}_f(t) = u_f(t-1)$. Finally, we propose the learning mechanism at each time instant t as follows.

Learning procedure: The estimates of the utility, regret, and probability distribution functions are performed, and are updated for all actions as follows:

$$\begin{cases}\tilde{u}_f^m\left(t\right) = & \tilde{u}_f^m\left(t-1\right) + \xi_f(t)\mathbb{I}_{\{\mathbf{z}_f = \mathbf{z}_f^m\}}\left(\hat{u}_f(t) - \tilde{u}_f^m\left(t-1\right)\right), \\ \tilde{r}_f^m\left(t\right) = & \tilde{r}_f^m\left(t-1\right) + \gamma_f(t)\left(\tilde{u}_f^m(t) - \hat{u}_f(t) - \tilde{r}_f^m\left(t-1\right)\right), \\ \pi_f^m(t) = & \pi_f^m(t-1) + \iota_f(t)\left(\beta_f^m(\tilde{\mathbf{r}}_f(t)) - \pi_f^m(t-1)\right),\end{cases} \tag{18}$$

Here, $\xi_f(t)$, $\gamma_f(t)$, and $\iota_f(t)$ are the learning rates which are chosen to satisfy the convergence properties (please see [10] for more details and convergence proof). Based on the probability distribution as per (18), the MBS determines the flow-split vector for each flow f as defined in Section III. Note that the learning-aided route selection[2] is performed in a long-term period to ensure that the routes do not suddenly change such that the SCBSs have enough time to release traffic from the queues.

B. Selection of Auxiliary Variables

The optimal values of auxiliary variables are found by solving SP2 which is a convex optimization problem. Let φ_0^{f*} be the optimal solution obtained by the first order derivative of the objective function of SP2. Assuming a logarithmic utility function [3], [11], we have

$$\varphi_0^{f*}(t) = \max\left\{\frac{\nu}{Y_0^f},\, 0\right\}.$$

C. Rate Allocation

Consider $R_{(i,j)}^f = \log(1 + p_{(i,j)}^f |g_{(i,j)}(\mathbf{h})|^2)$ as the transmission rate, where the effective channel gain[3] for mmWave channels can be modeled as $g_{(i,j)}(\mathbf{h}) = \frac{\tilde{g}_{(i,j)}(\mathbf{h})}{1+I^{\max}}$ [5]. Here, under the noise-limited regime of mmWave channels $\tilde{g}_{(i,j)}(\mathbf{h})$ and $I^{\max}$ denote the normalized channel gain and the maximum interference, respectively [3]. Denoting the left hand side (LHS) of (8) and (9) as D_b^f for simplicity, the optimal values of flow control $\mathbf{x}$ and transmit power $\mathbf{p}$ are found by minimizing

[2]Moreover, the routes are decided before transmission, which helps designing smart beamforming and beam-allignment to achieve higher directional gain, while mitigating interference.

[3]The effective channel gain captures the path loss, channel variations, and interference penalty (Here, the impact of interference is considered small due to highly directional beamforming and high pathloss for interfered signals at mmWave frequency band.).

$$\min_{\mathbf{x},\mathbf{p}|\mathbf{z}} \quad \sum_{f=1}^{F} -Y_0^f x_0^f \tag{19a}$$

$$\text{subject to} \quad 1+p^f_{(0,0^{(\mathrm{o})}_f)}|g_{(0,0^{(\mathrm{o})}_f)}|^2 \geq e^{x_0^f}, \forall f \in \mathcal{F}, \tag{19b}$$

$$\frac{1+p^f_{(b,b^{(\mathrm{o})}_f)}|g_{(b,b^{(\mathrm{o})}_f)}|^2}{1+p^f_{(b^{(\mathrm{I})},b)}|g_{(b^{(\mathrm{I})},b)}|^2} \geq e^{D_b^f}, \forall b \in \mathcal{S}^f, f \in \mathcal{F}, \tag{19c}$$

$$\sum_{f\in F} p^f_{(b,b^{(\mathrm{o})}_f)} \leq P_b^{\max}, \forall b \in \mathcal{B}, \forall f \in \mathcal{F}. \tag{19d}$$

The constraint (19c) is non-convex, but the LHS of (19c) is an affine-over-affine function, which is jointly convex w.r.t the corresponding variables [16], [17]. In this regard, we introduce the slack variable y to (19c) and rewrite it as

$$\frac{2+p^f_{(b,b^{(\mathrm{o})}_f)}|g_{(b,b^{(\mathrm{o})}_f)}|^2}{2} \geq \sqrt{y^2+\left(\frac{p^f_{(b,b^{(\mathrm{o})}_f)}|g_{(b,b^{(\mathrm{o})}_f)}|^2}{2}\right)^2}, \tag{20}$$

$$\frac{y^2}{1+p^f_{(b^{(\mathrm{I})},b)}|g_{(b^{(\mathrm{I})},b)}|^2} \geq e^{D_b^f}. \tag{21}$$

Here, the constraint (20) holds a form of the second-order cone inequalities [17], while the LHS of constraint (21) is a quadratic-over-affine function which is iteratively replaced by the first order to achieve a convex approximation as follow

$$\frac{2yy^{(l)}}{1+p^{f(l)}_{(b^{(\mathrm{I})},b)}|g_{(b^{(\mathrm{I})},b)}|^2} - \frac{y^{(l)2}\left(1+p^f_{(b^{(\mathrm{I})},b)}|g_{(b^{(\mathrm{I})},b)}|^2\right)}{\left(1+p^{f(l)}_{(b^{(\mathrm{I})},b)}|g_{(b^{(\mathrm{I})},b)}|^2\right)^2}. \tag{22}$$

Here, the superscript l denotes the lth iteration. Hence, we iteratively solve the approximated convex problem of (19) as **Algorithm** 1 in which the approximated problem is given as

$$\min_{\mathbf{x},\mathbf{p}|\mathbf{z}} \quad \sum_{f=1}^{F} -Y_0^f x_0^f \tag{23}$$
$$\text{subject to} \quad \text{(19b), (19d), (20), (22).}$$

Finally, the information flow diagram of the learning-aided path selection and rate allocation approach is shown in Fig. 2, where the rate allocation is executed in a short-term period.

Algorithm 1 Iterative rate allocation

Initialization: set $l=0$ and generate initial points $y^{(l)}$.
repeat
  Solve (23) with $y^{(l)}$ to get the optimal value $y^{(l)\star}$.
  Update $y^{(l+1)} := y^{(l)\star}$; $l := l+1$.
until Convergence

V. Numerical Results

In this section, we provide numerical results by assuming two flows from the MBS to two UEs, while the number of available paths for each flow is four [11]. The MBS determines two best routes from four most popular routes[4]. Each route contains one or two relays, the one-hop distance is

[4]As studied in [11], it suffices for a flow to maintain at least two paths provided that it repeatedly selects new paths at random and replaces if the latter provides higher throughput.

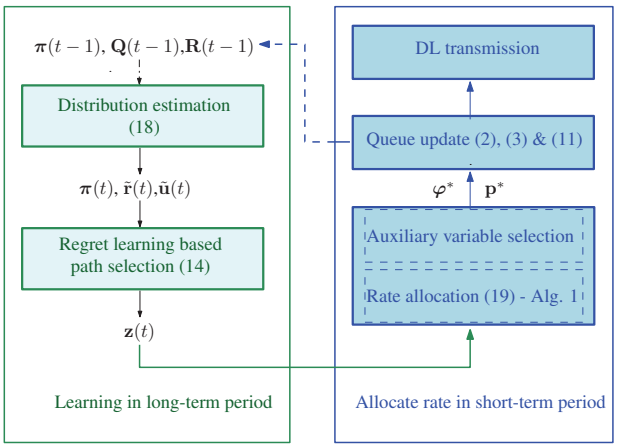

Fig. 2: Information flow diagram of the learning-aided path selection and rate allocation.

varying from 50 to 100 meters. The maximum transmit power of MBS and each SC are 43 dBm and 30 dBm, respectively. The SC antenna gain is 5 dBi and the number of antennas at each BS is $N_b = 8$. We assume that the traffic flow is divided equally into two subflows, the arrival rate for each subflow is varying from 2 to 5 Gbps. The path loss is modeled as a distance-based path loss with the line-of-sight (LOS) model[5] for urban environments at 28 GHz with 1 GHz of bandwidth [18]. The maximum delay requirement β and the target reliability probability ϵ are set to be 10 ms and 5%, respectively [6]. For the learning algorithm, the Boltzmann temperature (tradeoff factor) κ_f is set to 5, while the learning rates $\xi_f(t)$, $\gamma_f(t)$, and $\iota_f(t)$ are set to $\frac{1}{(t+1)^{0.5}}$, $\frac{1}{(t+1)^{0.55}}$, and $\frac{1}{(t+1)^{0.6}}$, respectively [10].

Furthermore, we compare our proposed scheme with the following baselines:

- ***Baseline* 1** refers to the leaning-based multihop framework in which the probabilistic latency constraint (4b) is not considered.
- ***Baseline* 2** considers the probabilistic latency constraint (4b), but randomly chooses paths to delivery data (without learning function).
- ***Baseline* 3** does not consider the probabilistic latency constraint (4b), and randomly chooses paths to delivery data (without learning function).
- ***Single hop*** scheme: The MBS delivers data to UEs over one single hop at long distance in which the probability of LOS communication is low, and then the blockage needs to be taken into account [18].

In Fig. 3, we report the average one-hop delay[6] versus the mean arrival rates $\bar{\mu}$. As we increase $\bar{\mu}$, ***baseline* 3** violates the latency constraints, whereas our proposed algorithm outperforms the other ***baselines***. The average one-hop delay of ***baseline* 1** with learning outperforms ***baselines* 2** and **3**, whereas our proposed scheme reduces latency by 50.64%, 81.32% and 92.9% as compared to ***baselines* 1**, **2**, and **3**, respectively, when $\lambda = 4.5$ Gbps. When $\lambda = 5$ Gbps, the average delay of all ***baselines*** increases, violating the delay requirement of 10 ms, while our proposed scheme is robust to the latency requirement. Moreover, for throughput

[5]We assume that the probability of LOS communication is high, while the impact of other channel models is left for future work.

[6]The average end-to-end delay can be defined as the sum of the average one-hop delay of all hops.

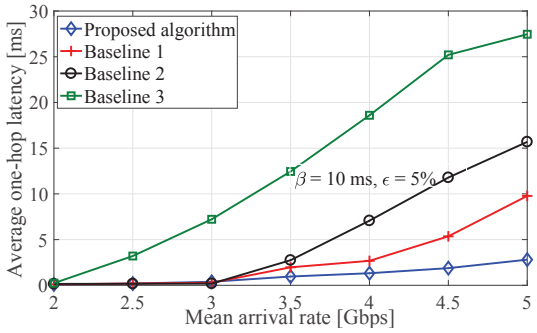

Fig. 3: Average one-hop delay versus mean arrival rates, $\epsilon = 5\%$, $\beta = 10$ ms, $\kappa = 5$.

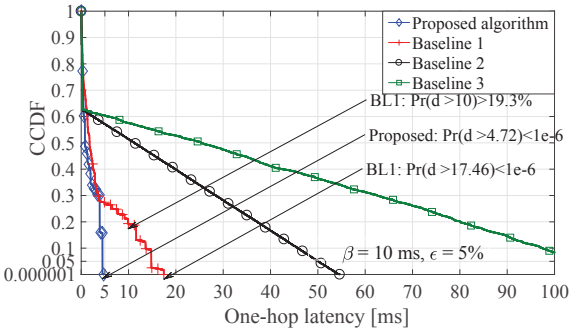

Fig. 4: CCDF of the one-hop latency, $\bar{\mu} = 4.5$ Gbps, $\epsilon = 5\%$, $\beta = 10$ ms, $\kappa = 5$.

comparison, we observe that as $\lambda = 4.5$ Gbps, our proposed algorithm is able to deliver 4.4874 Gbps of average network throughput per each subflow, while the ***baselines* 1, 2,** and **3** deliver 4.4759, 4.4682, and 4.3866 Gbps, respectively. Here, the ***single hop*** scheme only delivers 3.55 Gbps due to the blockage, which resulting in large delay.

In Fig. 4, we report the tail distribution (complementary cumulative distribution function (CCDF)) of latency to showcase how often the system achieves a delay greater than the target delay levels. In contrast to the average delay, the tail distribution is an important metric to reflect the URLLC characteristic. For instance, at $\lambda = 4.5$ Gbps, by imposing the probabilistic latency constraint, our proposed approach ensures reliable communication with better guaranteed probability, i.e, $\Pr(\text{Delay} > 4.72\text{ms}) < 1e-6$. In contrast, ***baseline* 1** with learning violates the latency constraint with high probability, where $\Pr(\text{Delay} > 10\text{ms}) = 19.3\%$ and $\Pr(\text{Delay} > 17.46\text{ms}) < 1e-6$, while the performance of ***baselines* 2** and **3** gets worse.

We further plot the tail distribution of one-hop latency versus the guaranteed probability ϵ as shown in Fig. 5, . By varying ϵ from 5% to 15%, the system achieves a delay greater than the target delay levels with higher probability. As can be seen in Fig. 5, the probability that the system achieves a delay greater than 4 ms increases when increasing ϵ.

VI. Conclusion

In this paper, we have proposed a multihop scheduling to ensure URLLC by incorporating the probabilistic latency constraint in 5G self-backhauled mmWave networks. In particular, the problem is modeled as a network utility maximization subject to the probabilistic latency/reliability constraints and queue stability. We have proposed a dynamic approach, which adapts to channel variations and system dynamics. We leverage the stochastic optimization to decouple the studied problem into path selection and rate allocation. While the learning algorithm is applied to select the best paths by using historical system information. Numerical results show that our proposed framework reduces latency by 50.64% and 92.9% as compared to *baselines* with and without learning, respectively.

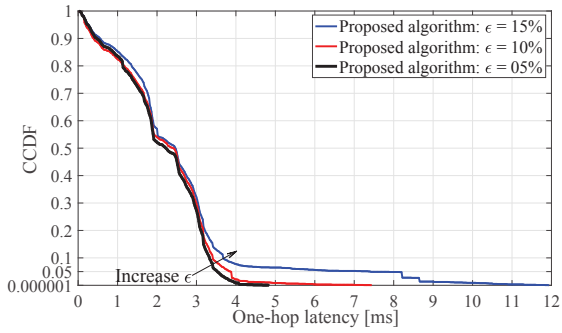

Fig. 5: CCDF of the one-hop latency versus the guaranteed probability ϵ, $\beta = 10$ ms, $\kappa = 5$, and $\bar{\mu} = 4.5$ Gbps.

Acknowledgment

The authors would like to thank the Finnish Funding Agency for Technology and Innovation (Tekes), Nokia, Huawei, and Anite for project funding. The Academy of Finland funding through the grant 284704 is also acknowledged.

References

[1] J. G. Andrews et al., "What Will 5G Be?" *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1065–1082, June 2014.

[2] A. Anpalagan, M. Bennis, and R. Vannithamby, *Design and Deployment of Small Cell Networks*. Cambridge University Press, 2015.

[3] T. K. Vu et al., "Joint load balancing and interference mitigation in 5G heterogeneous networks," *IEEE Transactions on Wireless Communications*, vol. 16, no. 9, pp. 6032–6046, Sep. 2017.

[4] T. S. Rappaport et al., "Millimeter wave mobile communications for 5G cellular: It will work!" *IEEE Access*, vol. 1, pp. 335–349, 2013.

[5] S. Hur et al., "Millimeter wave beamforming for wireless backhaul and access in small cell networks," *IEEE Trans. Commun.*, vol. 61, no. 10, pp. 4391–4403, Oct. 2013.

[6] T. K. Vu et al., "Ultra-reliable and low latency communication in mmWave-enabled massive MIMO networks," *IEEE Communications Letters*, vol. 21, no. 9, pp. 2041–2044, Sep. 2017.

[7] Qualcomm Technologies, Inc, "Leading the path towards 5G with LTE Advanced Pro," White Paper, 2016.

[8] S. Singh, F. Ziliotto, U. Madhow, E. M. Belding, and M. J. Rodwell, "Millimeter wave WPAN: Cross-layer modeling and multi-hop architecture," in *Proc. 26th IEEE Int. Conf. Comput. Commun.*, Barcelona, Spain, May 2007, pp. 2336–2340.

[9] N. Eshraghi et al., "Millimeter-wave device-to-device multi-hop routing for multimedia applications," in *Proc. IEEE Int. Conf. Commun.*, Kuala Lumpur, Malaysia, May 2016, pp. 1–6.

[10] M. Bennis, S. M. Perlaza, and M. Debbah, "Learning coarse correlated equilibria in two-tier wireless networks," in *Proc. IEEE Int. Conf. Commun.*, Ottawa, ON, Canada, Jun. 2012, pp. 1592–1596.

[11] P. Key et al., "Path selection and multipath congestion control," *Commun. ACM*, vol. 54, no. 1, pp. 109–116, Jan. 2011.

[12] B. Sahoo, C.-H. Yao, and H.-Y. Wei, "Millimeter-wave multi-hop wireless backhauling for 5g cellular networks," in *Proc. IEEE 85th Vehicular Technology Conf.*, Sydney, Australia, June 2017, pp. 1–6.

[13] M. J. Neely, "Stochastic network optimization with application to communication and queueing systems," *Synthesis Lectures on Communication Networks*, vol. 3, no. 1, pp. 1–211, 2010.

[14] A. Zakrzewska et al., "Dual connectivity in LTE HetNets with split control-and user-plane," in *Proc. IEEE Global Commun. Conf. Workshops*, Atlanta, GA, USA, Dec. 2013, pp. 391–396.

[15] J. D. Little and S. C. Graves, "Little's law," in *Building intuition*. Springer, 2008, pp. 81–100.

[16] K. G. Nguyen et al., "Achieving energy efficiency fairness in multicell MISO downlink," *IEEE Communications Letters*, vol. 19, no. 8, pp. 1426–1429, 2015.

[17] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[18] M. R. Akdeniz, Y. Liu, M. K. Samimi, S. Sun, S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter wave channel modeling and cellular capacity evaluation," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1164–1179, Jun. 2014.