Stochastic Arrow-Hurwicz Algorithm for Path Selection and Rate Allocation in Self-Backhauled mmWave Networks

Abhinav Sharma, K. Lakshmanan, Ruchir Gupta, *Senior Member, IEEE*, and Atul Gupta

***Abstract*— One of the key promoters of 5G deployment using millimetre waves is multi-hop self-backhauling. However, the rate allocation and path selection for these networks are challenging tasks. For this purpose, we consider a network utility maximization problem with probabilistic stability constraints. We propose a novel and simple method based on the stochastic Arrow-Hurcwicz algorithm to the optimization problem with gradient estimation using a smoothed functional technique. Numerical results show an improved utility and faster performance compared to the benchmark.**

***Index Terms*— Smoothed functional algorithms, Arrow-Hurwicz algorithm, simulation, mmWave networks, stochastic optimization, ultra-dense small cells.**

I. Introduction

SELF-BACKHAUL millimetre Wave (mmWave) [1] deployed with Ultra-Dense Network (UDN) using small cell base station (SCBS) [2] is one of the most promising technology for 5G transmission. Here, a large number of SCBSs are placed for each macro base station (MBS), which is essential in mmWave transmission. The major hurdle in this setup is capacity and resource allocation [3]. To improve the throughput over longer distances [4] massive MIMO technology has been employed, which allows a high degree of spatial multiplexing [5]. MBS uses a massive number of antennas to transmit mmWave to SCBS and vice-versa. This architecture uses multi-hop and multi-path data transfers where the MBS selects paths and allocates power to SCBS. However, hopping results in increased delay for the data packet transfers. Additionally, the shared bandwidth of the self-backhaul architecture needs to be optimized for optimal data rate allocation of different data flows. Thus, optimizing path selection and rate allocation for efficient communication are crucial components of this architecture.

One of the most notable works in this regard is done by Vu *et al.* [3] who formulated the problem of joint path selection and rate allocation into a network utility optimization problem with network stability and bounded latency constraints. They divided the problem into sub-problems and solved them with reinforcement learning (RL) and successive convex approximation techniques. A major issue in handling this non-convex optimization is the non-linear probability constraint on network queues. Their model [3] approximates the probability constraint with constraint in expectation using Markov inequality. However, the approach is rather crude as the probability constraint need not be satisfied even if the condition in expectation is satisfied. It then uses SOCP based approximation method, which is complicated. Thus, we believe there is room for significant improvement.

We propose a novel technique to solve the path selection and rate allocation problem of self-backhaul mmWave networks by using smoothed functional approach in the Arrow-Hurwicz algorithm. First, we approximate the probabilistic stability constraint with an indicator function using the smoothed stochastic technique. Then, we utilize the framework of the Arrow-Hurwicz algorithm [6] and smoothed functional gradient estimates [7] for the optimization. This resultant algorithm is faster, shows improved network utility and has a lower average queue length compared to the work of Vu *et al.* [3].

The rest of the letter is organized as follows: Section II describes the problem definition. Section III introduces the proposed algorithm. Section IV provides the numerical results against benchmark. Conclusions are drawn in Section V.

II. Problem Definition

Consider a downlink transmission with one macro base station (MBS) of a multi-hop heterogeneous cellular network [3]. Assume a set $\mathcal{B}$ of self-backhauled SCBS and set $\mathcal{K}$ of UEs. Let $\{0, 1, \ldots, B\}$ be the indices for the base stations with MBS getting index 0. Let, the total SCBS and UEs be $\mathcal{N}$. At the MBS, let there is a set $\mathcal{F}$ of F independent data flows. We consider discrete time intervals. Each flow $f \in \mathcal{F}$ is split by MBS into two or more sub-flows which are joined together at UEs. The paths of these sub-flows are disjoined and decided by the MBS from among Z_f number of disjoint paths. For a disjoint path m, $\mathcal{Z}_f^m$ denotes the path state, which has queue state information along with the topology for every hop. Let $\mathcal{Z}_f$ denote the state observed by flow f and $\mathbf{z}_f$ denote the flow-split indicator vector. Here, the vector $\mathcal{Z}_f = \{\mathcal{Z}_f^1, \ldots, \mathcal{Z}_f^{Z_f}\}$ and $\mathbf{z}_f = (z_f^1, \ldots, z_f^{Z_f})$ where, z_f^m sets to 1 whenever the path m is selected to send data for flow f.

Let the set of next hops from node i via a directional edge be $\mathcal{N}_i^{(o)}$. Let the upper limit of the transmit power for any node i be $P_i^{\max}$. For any flow f, transmit power is assigned from transmitter node i to receiver node j and is denoted by $p_{(i,j)}^f$. Next, as the transmit power cannot be negative for any network, we have:

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

Manuscript received September 15, 2021; revised October 23, 2021; accepted November 21, 2021. Date of publication December 3, 2021; date of current version March 10, 2022. This work was supported by Nokia Solutions and Networks, Bengaluru. The associate editor coordinating the review of this letter and approving it for publication was C. Jiang. *(Corresponding author: Ruchir Gupta.)*

Abhinav Sharma and Atul Gupta are with the Indian Institute of Information Technology Design and Manufacturing, Jabalpur, Jabalpur 482005, India.

K. Lakshmanan and Ruchir Gupta are with the IIT BHU, Varanasi 221005, India (e-mail: rgupta.cse@iitbhu.ac.in).

Digital Object Identifier 10.1109/LCOMM.2021.3131328

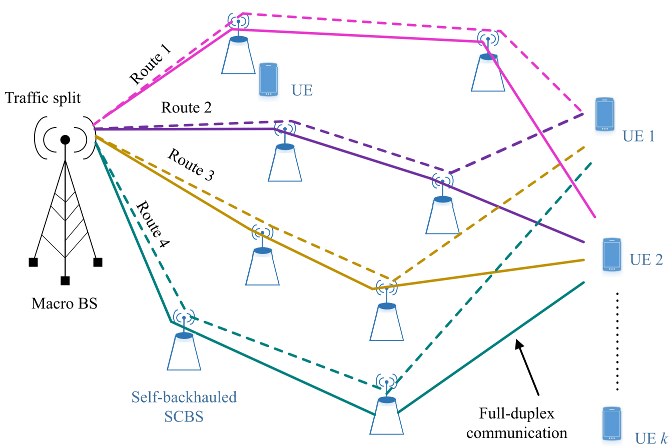

Fig. 1. 5G multi-hop self-backhauled mm-wave network.

as the power constraint. We assume the queue length at BS i to be $Q_f^i(t)$ for flow f and time slot t. For flow f, let $r_f^{(i,j)}(\boldsymbol{H},\boldsymbol{p})$ denote the data rate in edge (i,j) with channel state $\boldsymbol{H}$ and transmit power $\boldsymbol{p}$ as arguments. Let, $\boldsymbol{r}$ be used to refer data rates vector over all the flows. Figure 1 shows MBS communicating with UE-1 and UE-2, where the communication links are shown as a dashed and solid line, respectively.

A. Optimization Problem

The probability that a particular path m is selected for a sub-flow of flow f is denoted by $\Pr\left(z_f = z_f^m\right) = \pi_f^m$. Thus for all the sub-flows of f we can define a probability mass function $\boldsymbol{\pi}_f = (\pi_f^1,\dots,\pi_f^{Z_f})$ where $\sum_{m=1}^{Z_f} Pr(z_f^m) = 1$. Let the probability distribution and the set containing all possible global probability mass functions of all flow-split vectors be denoted by Π and $\boldsymbol{\pi}$ respectively where $\boldsymbol{\pi} \in \Pi$. Note that, $\boldsymbol{\pi} = \{\pi_1,\dots,\pi_f,\dots,\pi_F\}$.

Further, let the achievable average rate of flow $\bar{x}_f$ for the flow f be:

$$\bar{x}_f \triangleq \lim_{t\leftarrow\infty} \frac{1}{t}\sum_{\tau=0}^{t-1} x_f(\tau),$$

with $x_f(\tau)$ be given as:

$$x_f(\tau) = \sum_{m=1,i_f^{(o)}\in\mathcal{Z}_f^m}^{Z_f} \mathbb{E}_{\boldsymbol{H},\boldsymbol{p}}\left[\pi_f^m r_f^{(i,i_f^{(o)})}(\tau)\right]\Big|_{i=0}.$$

Now, assume that $x_f(t)$ is non negative, and is bounded from above with a maximum achievable rate $a_f^{\max}$ at all times for flow f. The constraint on this rate is:

$$0 \le x_f(t) \le a_f^{\max} \tag{2}$$

The vector $\bar{\boldsymbol{x}} = (\bar{x}_1,\dots,\bar{x}_F)$ is the rate averaged over a period of time over all flows. We can now define the objective function as a network utility function $U_0(\bar{x})$, which is a twice differentiable, concave and increasing L-Lipschitz for all $\bar{x} \ge 0$. Formally, $U_0(\bar{\boldsymbol{x}}) = \sum_{f\in F} U(\bar{x}_f)$ [8], [9].

The objective is to allocate resources across the network, such that the overall utility is maximized, while confirming to latency, power and reliability constraints. Thus, the optimization problem is modelled as a network utility maximization problem, which is given by:

$$\max_{\pi,\boldsymbol{x},\boldsymbol{p}} \; U_0(\bar{\boldsymbol{x}}) \text{ s.t.} \tag{3a}$$

$$\Pr\left(\frac{Q_f^i(t)}{\bar{a}_f} \ge d^{\max}\right) \le \epsilon, \text{ for all } t, f\in F, i\in B, \tag{3b}$$

$$\lim_{t\to\infty} \frac{\mathbb{E}\left[|Q_f^i|\right]}{t} = 0, \text{ for all } f\in F, i\in B \tag{3c}$$

$$x(t) \in \mathcal{R}, \tag{3d}$$

$$\pi \in \Pi, \tag{3e}$$

and constraints (1), (2).

In equation (3b), the ratio of queue length ($Q_f^i(t)$) and average arrival rate ($\bar{a}_f$) represents average queuing delay, which is considered in accordance to Little's law [10]. Hence, the constraint dictates that the probability of average queuing delay being greater than $d^{\max}$ is less than ϵ for each flow f at node i. This constraint is essential to ensure reliable and ultra-low latency communication [11]. This avoids congestion in the network by limiting the queue build-up at any BS. The following constraint (3c) is to ensure the stability of the network. Other constraints are boundary conditions on rates (3d, 2), flow split probabilities (3e), and power (1).

III. Smoothed Functional Stochastic Arrow-Hurwicz Algorithm

The basic idea of the proposed model is to consider the probability constraint as a constraint on the expected value of an indicator function. Then use a standard stochastic gradient-based technique to solve the problem. Finally, since the indicator function is itself discontinuous, we consider a smoothed approximation of it using approximation by convolution (AC) [6].

Algorithm 1 Arrow-Hurwicz Algorithm

1: Set NUM-ITERATIONS to a large number.
2: **for** $k = 1 \to$ NUM-ITERATIONS **do**
3: Draw an independent sample ε^{k+1}.
4: Compute the stochastic gradient $\nabla_u J(u^k,\varepsilon^{k+1})$ and $\nabla_u G(u^k,\varepsilon^{k+1})$.
5: Update u^{k+1} and Lagrangian multiplier λ^{k+1} as follows:

$$u^{k+1} = \mathcal{J}_U\Big(u^k - \omega^k\big(-\nabla_u J(u^k,\varepsilon^{k+1}) + \nabla_u G(u^k,\varepsilon^{k+1})\lambda^k\big)\Big), \tag{4}$$

$$\lambda^{k+1} = \mathcal{J}_+\left(\lambda^k + \rho^k\left(G(u^{k+1},\varepsilon^{k+1}) - \alpha\right)\right). \tag{5}$$

6: **end for**

A. Stochastic Arrow-Hurwicz Algorithm

Before moving on to the proposed solution, we first discuss the Arrow-Hurwicz algorithm.

Consider the following stochastic optimization problem:

$$\max_{u \in U} -J(u) \text{ such that } \boldsymbol{G}(u) \leq \alpha \tag{6}$$

where u be the decision variable in a Hilbert space $\mathcal{U}$, $J : \mathcal{U} \rightarrow \mathbb{R}^d$ is the cost function, and $\boldsymbol{G} : \mathcal{U} \rightarrow \mathbb{R}$ be the constraint function. Then the Lagrangian is given by:

$$L(u, \lambda) = J(u) + \lambda(\boldsymbol{G}(u) - \alpha) \tag{7}$$

where $\lambda \in \mathbb{R}^d$ is a multiplier.

If $\boldsymbol{G} = \mathbb{E}(G(u, \varepsilon))$ in the above equation, i.e., constraint in expectation, then equation (6) can be written as:

$$\max_{u \in U} -J(u, \varepsilon) \text{ s.t. } \mathbb{E}(G(u, \varepsilon)) \leq \alpha \tag{8}$$

Now, if equation (3b) can be written in form of expectation, then the equation (8) can correspond to equation (3a).

The Arrow-Hurwicz algorithm (in 1), finds the unbiased estimate of the gradient of the Lagrangian L. $\mathcal{J}$ denotes a projection operator. As in any stochastic approximation algorithm the step-size sequences ω^k should satisfy the following two conditions.

$$\sum_{k \in \mathbb{N}} \omega^k = \infty, \quad \sum_{k \in \mathbb{N}} (\omega^k)^2 < \infty, \quad \text{same goes for } \rho^k \tag{9}$$

B. Smoothed Stochastic Gradient

Suppose we have the probability constraint $\mathbb{P}\left(G(u, \varepsilon) \leq \alpha\right) \geq \pi$. This can written as expectation constraint using an indicator function as:

$$S(u) \geq \pi, \text{ where } S(u) = \mathbb{E}\bigg(\mathbb{I}_{\mathbb{R}^+}\left(\alpha - G(u, \varepsilon)\right)\bigg). \tag{10}$$

where $\mathbb{I}_{\mathbb{R}^+}$ is the indicator function and $S(u) = \mathbb{P}(G(u, \varepsilon) \leq \alpha)$. This gradient is difficult to compute because of the discontinuity in the indicator function [6]. The standard way is to derive a smooth approximation of this function and then find the gradient.

Consider a smooth function $h : \mathbb{R} \rightarrow \mathbb{R}$ such that: h has a unique maximum at $x = 0$,

$$\forall x, \quad h(x) = h(-x); \quad h(x) \geq 0; \quad \int_{-\infty}^{+\infty} h(x) dx = 1. \tag{11}$$

Then, convolute this with the indicator function $\mathbb{I}_{\mathbb{R}^+}$ and get

$$S_k(u) = \frac{1}{k}\mathbb{E}\left(\int_0^{+\infty} h\left(\frac{y - \alpha + G(u, \varepsilon)}{k}\right) dy\right).$$

The expectation can be denoted by $\mathbb{E}\left(s_k(u, \varepsilon)\right)$. We can then derive the following:

$$(s_k)'_u(u, \varepsilon) = -\frac{1}{k} h\left(\frac{G(u, \varepsilon) - \alpha}{k}\right) G'_u(u, \varepsilon). \tag{12}$$

Hence,

$$S'_k(u) = \mathbb{E}\left((s_k)'_u(u, \varepsilon)\right). \tag{13}$$

Thus the previous equation (12) can be used as an unbiased estimate of $S'(u)$. We note that this bias vanishes when r approaches 0 [6].

C. Changed Objective Function

Let $Q_f^i(t)$ denote the queue length for flow f at BS i and time slot t. At the MBS $(i = 0)$, the update equation for the queue length is

$$Q_f^i(t+1) = \left(Q_f^i(t) - \sum_{m=1, i_f^{(o)} \in \mathcal{Z}_f^m}^{Z_f} r_f^{(i, i_f^{(o)})}(t), 0\right)^+ + a_f(t), \tag{14}$$

where $a_f(t)$ is the data arrival at the MBS for flow f at time slot t, which is independent and identically distributed (i.i.d.) over time with a mean value as $\bar{a}_f$.

Let $\mathcal{A}(\boldsymbol{\pi}, \boldsymbol{p}, \varepsilon) = \min_{i,f}\left[\frac{Q_f^i(t)}{\bar{a}_f}\right]$, then the probability constraint is given by

$$\Pr\left(\mathcal{A}(\boldsymbol{\pi}, \boldsymbol{p}, \varepsilon) \geq d^{\max}\right) \leq \epsilon$$

in equation (3b) corresponds to $\mathbb{P}\left(G(u, \varepsilon) \leq \alpha\right) \geq \pi$ in equation (10). Note that, by introducing a minima of $Q_f^i(t)/\bar{a}_f$, the inequality of (3b) still holds. Thus, from equation (3b) and equation (10), the modified constraint in place of (3b) is given by:

$$S\left(\boldsymbol{\pi}, \boldsymbol{p}\right) \leq \epsilon, \tag{15}$$

where $S(\boldsymbol{\pi}, \boldsymbol{p})$ denotes the expected value of the indicator function and is given by, $S(\boldsymbol{\pi}, \boldsymbol{p}) = \mathbb{E}\left(\mathbb{I}_{\mathbb{R}^+}\left(-d^{\max} + \mathcal{A}\left(\boldsymbol{\pi}, \boldsymbol{p}\right)\right)\right) := \mathbb{E}(g_1(\boldsymbol{\pi}, \boldsymbol{p}))$.

Now, to find the gradient estimate, assume that $h(\cdot) : \mathbb{R} \implies \mathbb{R}$ is a smooth function that follows equation (11), the gradient for $S(\boldsymbol{\pi}, \boldsymbol{p})$ is calculated according to equation (12), i.e., :

$$S'_k(\boldsymbol{\pi}, \boldsymbol{p}) = \mathbb{E}\left((\nabla g_1)_{\boldsymbol{\pi}, \boldsymbol{p}}(\boldsymbol{\pi}, \boldsymbol{p}, \varepsilon)\right). \tag{16}$$

where

$$\nabla_{\boldsymbol{\pi}} g_1(\boldsymbol{\pi}, \boldsymbol{p}, \varepsilon) = \frac{1}{k} h\left(\frac{\mathcal{A}(\boldsymbol{\pi}, \boldsymbol{p}, \varepsilon) - d^{\max}}{k}\right) \nabla_{\boldsymbol{\pi}} \mathcal{A}(\boldsymbol{\pi}, \boldsymbol{p}, \varepsilon) \tag{17}$$

$$\nabla_{\boldsymbol{p}} g_1(\boldsymbol{\pi}, \boldsymbol{p}, \varepsilon) = \frac{1}{k} h\left(\frac{\mathcal{A}(\boldsymbol{\pi}, \boldsymbol{p}, \varepsilon) - d^{\max}}{k}\right) \nabla_{\boldsymbol{p}} \mathcal{A}(\boldsymbol{\pi}, \boldsymbol{p}, \varepsilon) \tag{18}$$

To apply Arrow-Hurwicz, we approximate the limiting constraint in equation (3c) in the following way:

$$\frac{\mathbb{E}\left(|Q_f^i|\right)}{t^c} \leq 1, \quad \forall t, f \in F, i \in B \tag{19}$$

where $0 < c < 1$ is a constant. This constraint is to ensure the stability of the network. Next, we denote $g_2(\boldsymbol{\pi}, \boldsymbol{p}) = \max_{i,f}\left[\mathbb{E}\left(-|Q_f^i|\right)/t^c\right]$ and to be consistent with the algorithm, the inequality is denoted as $g_2(\boldsymbol{\pi}, \boldsymbol{p}) \geq -1, \forall t, f \in F, i \in B$.

It is unknown whether $g_2(\boldsymbol{\pi}, \boldsymbol{p})$ is differentiable or not. Therefore, for gradient estimation, we perform a smoothed functional gradient estimation which involves convolution with a Z_f-dimensional multivariate $(0, \beta_2^2)$-distributed Gaussian [7]. For some small scalar $\beta_2 > 0$ and a large positive integer M_2, the gradient estimator with respect to $\boldsymbol{\pi}$ is given by:

$$\nabla_{\boldsymbol{\pi}} g_2(\boldsymbol{\pi}, \boldsymbol{p}) \approx \frac{1}{M_2}\frac{1}{\beta_2}\sum_{t=1}^{M_2}\eta(t)(g_2(\boldsymbol{\pi}+\beta_2\eta(t),\boldsymbol{p}) - g_2(\boldsymbol{\pi},\boldsymbol{p})) \tag{20}$$

Here $\eta(t) \triangleq (\eta_1(t),\ldots,\eta_N(t))^T$, with $\eta_i(t), i = 1,\ldots,N, n \geq 0$, being independent $Z_f(0,1)$-distributed random variables. Similarly, the gradient estimator with respect to $\boldsymbol{p}$ is given by:

$$\nabla_{\boldsymbol{p}} g_2(\boldsymbol{\pi}, \boldsymbol{p}) \approx \frac{1}{M_2}\frac{1}{\beta_2}\sum_{t=1}^{M_2}\eta(t)(g_2(\boldsymbol{\pi},\boldsymbol{p}+\beta_2\eta(t)) - g_2(\boldsymbol{\pi},\boldsymbol{p})) \tag{21}$$

Similar treatment is required for the estimation of $\nabla g_1(\cdot)$ in equation (17), (18), as it is difficult to determine whether $\mathcal{A}$ is differentiable or not. Thus, the gradient estimate for $\nabla_{\boldsymbol{\pi}}\mathcal{A}(\boldsymbol{\pi},\boldsymbol{p},\varepsilon)$ and $\nabla_{\boldsymbol{p}}\mathcal{A}(\boldsymbol{\pi},\boldsymbol{p},\varepsilon)$ is given by:

$$\nabla_{\boldsymbol{\pi}}\mathcal{A}(\boldsymbol{\pi},\boldsymbol{p},\varepsilon) = \frac{1}{M_1\beta_1}\sum_{t=1}^{M_1}\zeta(t)\Big(\mathcal{A}(\boldsymbol{\pi}+\beta_1\zeta(t),\boldsymbol{p},\varepsilon) - \mathcal{A}(\boldsymbol{\pi},\boldsymbol{p},\varepsilon)\Big) \tag{22}$$

$$\nabla_{\boldsymbol{p}}\mathcal{A}(\boldsymbol{\pi},\boldsymbol{p},\varepsilon) = \frac{1}{M_1\beta_1}\sum_{t=1}^{M_1}\zeta(t)(\mathcal{A}(\boldsymbol{\pi},\boldsymbol{p}+\beta_1\zeta(t),\varepsilon) - \mathcal{A}(\boldsymbol{\pi},\boldsymbol{p},\varepsilon)) \tag{23}$$

where, M_1 is a large positive integer, β_1 belongs to Z_f dimensional multivariate $(0,\beta_1^2)$ distributed Gaussian and ζ is a $Z_f(0,1)$ distributed random variable similar to η. Thus equation (17), (18) can be expressed as:

$$\nabla_{\boldsymbol{\pi}} g_1(\boldsymbol{\pi},\boldsymbol{p},\varepsilon) = \frac{1}{kM_1\beta_1}h\bigg(\frac{\mathcal{A}(\boldsymbol{\pi},\varepsilon)-d^{\max}}{k}\bigg)\sum_{t=1}^{M_1}\zeta(t) \times(\mathcal{A}(\boldsymbol{\pi}+\beta_1\zeta(t),\boldsymbol{p},\varepsilon)-\mathcal{A}(\boldsymbol{\pi},\boldsymbol{p},\varepsilon)) \tag{24}$$

Similarly:

$$\nabla_{\boldsymbol{p}} g_1(\boldsymbol{\pi},\boldsymbol{p},\varepsilon) = \frac{1}{kM_1\beta_1}h\bigg(\frac{\mathcal{A}(\boldsymbol{\pi},\boldsymbol{p},\varepsilon)-d^{\max}}{k}\bigg)\sum_{t=1}^{M_1}\zeta(t) \times(\mathcal{A}(\boldsymbol{\pi},\boldsymbol{p}+\beta_1\zeta(t),\varepsilon)-\mathcal{A}(\boldsymbol{\pi},\boldsymbol{p},\varepsilon)) \tag{25}$$

To estimate the gradient of utility function, we use an i.i.d. random variable η_3 and generate two simultaneous simulations of $\bar{x}$ with respect to each $\boldsymbol{\pi}$ and $\boldsymbol{p}$.

$$\frac{\partial U(\bar{x})}{\partial \boldsymbol{\pi}} \approx \frac{1}{\beta_3}\left(U(\bar{x}_{\boldsymbol{\pi}+\beta_3\eta_3}) - U(\bar{x}_{\boldsymbol{\pi}})\right) \tag{26}$$

$$\frac{\partial U(\bar{x})}{\partial \boldsymbol{p}} \approx \frac{1}{\beta_4}\left(U(\bar{x}_{\boldsymbol{p}+\beta_4\eta_3}) - U(\bar{x}_{\boldsymbol{p}})\right) \tag{27}$$

As we have the gradient estimates, the update parameters for the Arrow-Hurwicz algorithm can be formulated as:

$$\boldsymbol{\pi}^{k+1} = \mathcal{J}_{\Pi}\bigg[\boldsymbol{\pi}^k - \omega^k\bigg(-\nabla_{\boldsymbol{\pi}}U_0(\boldsymbol{\pi}^k,\varepsilon^{k+1}) + \nabla_{\boldsymbol{\pi}}g_1(\boldsymbol{\pi}^k,\boldsymbol{p}^k,\varepsilon^{k+1})\lambda_1^k + \nabla_{\boldsymbol{\pi}}g_2(\boldsymbol{\pi}^k,\boldsymbol{p}^k,\varepsilon^{k+1})\lambda_2^k\bigg)\bigg] \tag{28}$$

$$\boldsymbol{p}^{k+1} = \mathcal{J}_{\mathcal{P}}\bigg[\boldsymbol{p}^k - \omega^k\bigg(-\nabla_{\boldsymbol{p}}U_0(\boldsymbol{p}^k,\varepsilon^{k+1}) + \nabla_{\boldsymbol{p}}g_1(\boldsymbol{\pi}^k,\boldsymbol{p}^k,\varepsilon^{k+1})\lambda_1^k + \nabla_{\boldsymbol{p}}g_2(\boldsymbol{\pi}^k,\boldsymbol{p}^k,\varepsilon^{k+1})\lambda_2^k\bigg)\bigg] \tag{29}$$

$$\lambda_1^{k+1} = \mathcal{J}_{+}\Big[\lambda_1^k + \rho^k\left(g_1(\boldsymbol{\pi}^{k+1},\boldsymbol{p}^{k+1},\varepsilon^{k+1}) - \epsilon\right)\Big]. \tag{30}$$

$$\lambda_2^{k+1} = \mathcal{J}_{+}\Big[\lambda_2^k + \rho^k\left(g_2(\boldsymbol{\pi}^{k+1},\boldsymbol{p}^{k+1},\varepsilon^{k+1}) + 1\right)\Big]. \tag{31}$$

Algorithm 2 Arrow-Hurwicz Smoothed Functional Algorithm

1: Set NUM-ITERATIONS to a large number.
2: **for** $t = 1 \to$ NUM-ITERATIONS **do**
3: Compute the stochastic gradient for the objective (equation (3a)) w.r.t. $\boldsymbol{\pi}$ and $\boldsymbol{p}$, i.e., $\nabla_{\boldsymbol{\pi}}U_0(\bar{x})$ and $\nabla_{\boldsymbol{p}}U_0(\bar{x})$.
4: Compute the gradient estimates for constraints from equation (20), (21), (24), (25).
5: Update $\boldsymbol{\pi}^{(t+1)}$, $\boldsymbol{p}^{(t+1)}$ and Lagrangian multiplier $\lambda^{(t+1)}$ as per equation (28)-(31).
6: **end for**

IV. Numerical Analysis

All the experiments are performed in MATLAB 2017b on a Intel® Core™ i5 CPU@1.80GHz and 8-GB RAM running on Windows 10 OS.

We consider one MBS and eight SCBS with a one-hop distance ranging from $50m$ to $100m$. For each SCBS and MBS, the maximum transmits power is capped at 30 dBm and 43 dBm, respectively. The small antenna arrays for BS is 8 and for UE is 2. Here, the SC antenna gain is set to 5 dBi. We use the general channel model of arbitrary antenna arrays. The number of RF chains assigned at BS is eight, and UE is two, the same as the flows from MBS to each UE. For each flow, MBS selects two paths from a total of four available routes [9].

In experiments, we have taken the utility function as $U(\bar{x}) = \log(\bar{x}+1)$ [3], [12], but the results will be valid for other utility functions too, if they satisfy the necessary requirements mentioned in Section-II. Each traffic flow is divided equally into two sub-flows for which the arrival rate varies from 2 to 5 Gbps. The maximum delay requirement is set to 10 ms whereas the target reliability probability is 5% [11].

We model distance-based path loss for blockage, line-of-sight (LOS), or non-LOS (NLOS) states as $PL(i,j)$ at 28 GHz with 1 GHz system bandwidth for urban environments [13]. For a distance d, the LOS probability is set to $\exp(-0.006d)$ and NLOS probability is $1-\exp(-0.006d)$. The side lobe gain

TABLE I
COMPUTATIONAL TIME (IN 10^3 sec) OF ALGORITHM

Algorithms	Mean Arrival Rates						
	2	2.5	2.86	3.5	4	4.5	5
AH-SF	1.44 ± 0.063	1.64 ± 0.026	2.56 ± 0.012	2.95 ± 0.034	3.13 ± 0.049	3.47 ± 0.022	3.79 ± 0.060
Baseline	1.84 ± 0.042	2.26 ± 0.033	3.28 ± 0.041	3.74 ± 0.068	3.86 ± 0.052	3.98 ± 0.044	4.25 ± 0.064

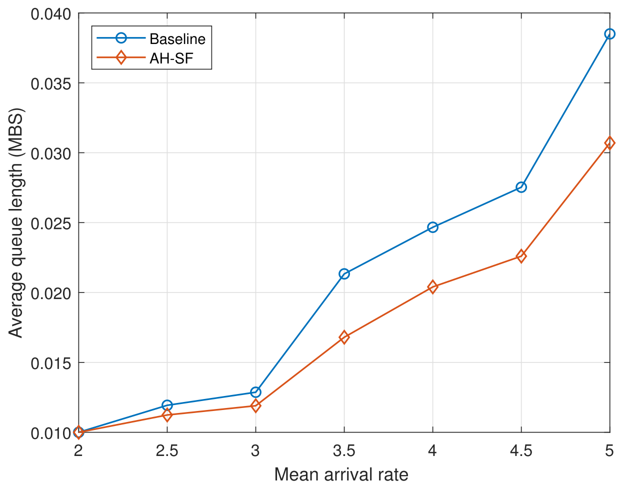

Fig. 2. Average queue length with different mean arrival rates.

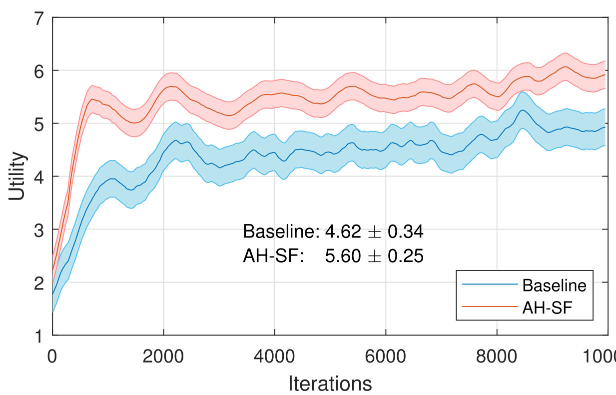

Fig. 3. The utility versus the number of iterations.

for analog beamforming is taken as 0.25. The beamwidth is set to $\pi/4$ radians at transmitter and $\pi/3$ radians at receiver.

We have compared our model with the current state-of-art [3], which is considered as the baseline. For the baseline, the boltzman temperature is set to 10 and the learning rates considered are $1/(t+1)^{0.51}, 1/(t+1)^{0.55}$, and $1/(t+1)^{0.6}$.

The proposed approach results in a shorter queue length as conveyed in Figure 2. Table I shows the computational time of two models for different mean arrival rates. It shows that the proposed model is computationally more efficient than the baseline. The reason might be that Vu el al.'s work uses an RL model for path selection, requiring more training time. This additional overhead is reflected in the total computational time of the two models.

When comparing the network utility, it is observed that the utility of the proposed algorithm is 5.6 ± 0.25. In contrast, the baseline is 4.62 ± 0.34, and thus the proposed algorithm scores $21.2\% \pm 0.36\%$ better utility score than the baseline (Figure 3).

V. CONCLUSION

This letter presents an algorithm for maximization under uncertainty constraints using the stochastic Arrow-Hurwicz algorithm with smoothed functional gradient estimates. First, the probability constraints are approximated with an indicator function using the smoothed stochastic approach. Then Arrow-Hurwicz algorithm is used for optimization with the help of smoothed stochastic gradient estimates. Finally, we provided numerical results on the problem of path selection and rate allocation in self-backhauled mmWave networks where our proposed algorithm outperformed the state-of-the-art technique. For future work, we will provide the convergence results of the proposed algorithm.

ACKNOWLEDGMENT

The authors would like to thank Nokia Solutions and Networks, Bengaluru, for supporting this work.

REFERENCES

[1] A. N. Uwaechia and N. M. Mahyuddin, "A comprehensive survey on millimeter wave communications for fifth-generation wireless networks: Feasibility and challenges," *IEEE Access*, vol. 8, pp. 62367–62414, 2020.
[2] I. A. Alablani and M. A. Arafah, "An adaptive cell selection scheme for 5G heterogeneous ultra-dense networks," *IEEE Access*, vol. 9, pp. 64224–64240, 2021.
[3] T. K. Vu, M. Bennis, M. Debbah, and M. Latva-Aho, "Joint path selection and rate allocation framework for 5G self-backhauled mm-wave networks," *IEEE Trans. Wireless Commun.*, vol. 18, no. 4, pp. 2431–2445, Apr. 2019.
[4] A. Bonfante *et al.*, "5G massive MIMO architectures: Self-backhauled small cells versus direct access," *IEEE Trans. Veh. Technol.*, vol. 68, no. 10, pp. 10003–10017, Oct. 2019.
[5] M. H. Alsharif and N. Nordin, "Evolution towards fifth generation (5G) wireless networks: Current trends and challenges in the deployment of millimetre wave, massive MIMO, and small cells," *Telecommun. Syst.*, vol. 64, no. 4, pp. 617–637, Apr. 2017.
[6] L. Andrieu, G. Cohen, and F. Vázquez-Abad, "Stochastic programming with probability constraints," HAL, Bengaluru, India, Tech. Rep. Hal-00166149, 2007.
[7] S. Bhatnagar, H. Prasad, and L. Prashanth, *Stochastic Recursive Algorithms for Optimization*, vol. 434. London, U.K.: Springer-Verlag, 2013.
[8] T. K. Vu *et al.*, "Joint load balancing and interference mitigation in 5G heterogeneous networks," *IEEE Trans. Wireless Commun.*, vol. 16, no. 9, pp. 6032–6046, Sep. 2017.
[9] P. Key, L. Massoulie, and D. Towsley, "Path selection and multipath congestion control," in *Proc. 26th IEEE Int. Conf. Comput. Commun. (INFOCOM)*, May 2007, pp. 143–151.
[10] J. D. C. Little and S. C. Graves, *Little's Law*. Cambridge, MA, USA: Massachusetts Institute Technology," 2008.
[11] T. K. Vu *et al.*, "Ultra-reliable and low latency communication in mmWave-enabled massive MIMO networks," *IEEE Commun. Lett.*, vol. 21, no. 9, pp. 2041–2044, Sep. 2017.
[12] A. Sinha and E. Modiano, "Network utility maximization with heterogeneous traffic flows," in *Proc. 16th Int. Symp. Modeling Optim. Mobile, Ad Hoc, Wireless Netw. (WiOpt)*, May 2018, pp. 1–8.
[13] T. Bai, V. Desai, and R. W. Heath, Jr., "Millimeter wave cellular channel models for system evaluation," in *Proc. Int. Conf. Comput., Netw. Commun. (ICNC)*, Feb. 2014, pp. 178–182.