ADAPTIVE WIRELESS POWER ALLOCATION WITH GRAPH NEURAL NETWORKS

Navid NaderiAlizadeh[⋆] *Mark Eisen*[†] *Alejandro Ribeiro*[⋆]

[⋆] University of Pennsylvania
[†] Intel Corporation

ABSTRACT

We consider the problem of power control in wireless networks, consisting of multiple transmitter-receiver pairs communicating with each other over a single shared wireless medium. To achieve both a high total rate and a level of fairness across users, we formulate a policy optimization problem with constraints on the minimum per-user rate across network configurations with an adaptive slack parameter. To apply unsupervised learning algorithms in the dual domain, we parameterize the power control policy, slack variable, and dual parameters using graph neural networks (GNNs), which leverage the network topology to create a scalable and network-invariant processing architecture. We use a primal-dual algorithm to learn the optimal GNN parameters and demonstrate via numerical simulations the resulting GNNs' success in achieving the right balance between sum- and 5^{th} percentile rates throughout a range of network configurations.

***Index Terms*—** Wireless power allocation, graph neural networks, primal-dual learning, unsupervised learning.

1. INTRODUCTION

With the advent of 5G network deployments across the world and the discussions around beyond 5G technologies, ultra-dense network (UDN) scenarios become of critical importance. To address the challenges arising in the UDN paradigm related to the scarcity of wireless resources, a considerable amount of research has been conducted on the problem of *radio resource management* (RRM), where the goal is to efficiently and optimally allocate the limited time/frequency/spatial resources across the network. However, the fact that RRM problems are very challenging to solve in the exact form has led to the development of numerous approximate heuristic solutions [1–4].

More recently, the success of deep learning in areas such as computer vision and natural language processing led to the emergence of learning-based solutions for challenging problems in wireless communications, including for RRM [5]. In particular, for the class of power allocation problems that we consider in this work, several approaches have been proposed that leverage techniques based on supervised, unsupervised, self-supervised, and reinforcement learning, as well as graph representation learning, and meta-learning [6–14]. Most of the prior work on learning-based power control, however, has considered optimizing network-wide objective functions, such as sum-throughput, without any constraints for fair allocation of resources across the network.

In this paper, we consider the problem of power allocation in wireless networks with multiple interfering transmitter-receiver pairs, where the goal is to control the transmit power of all transmitters for maximizing the network sum-rate, while ensuring all receivers in the network are treated fairly. We show how this problem can be cast as a constrained optimization problem, where the ergodic long-term average rate of each receiver is forced to be lower-bounded by an *adaptive* minimum rate constraint, which is learned via an optimized *slack* parameter. We specifically demonstrate how we can parameterize the power control policy, as well as the minimum rates, which adapt to any given network configuration [15]. We reformulate the problem in the Lagrangian dual domain, and show that the dual coefficients can also be parameterized as policies, whose outputs vary based on the input network configuration. We show how we can use three graph neural networks (GNNs) to parameterize the aforementioned primal and dual policies, and we propose a gradient-based primal-dual framework to optimize the parameters of the GNNs. We verify, through simulation results, that the minimum rate constraints indeed adapt themselves to the network configurations under study. We further demonstrate that thanks to such adaptation, our proposed method achieves a superior tradeoff between the sum-rate and the 5^{th} percentile rate, a metric that quantifies the level of fairness in the resource allocation decisions as compared to baseline algorithms.

2. PROBLEM FORMULATION

We consider a wireless interference network with a set of m transmitters $\{\mathsf{Tx}_i\}_{i=1}^m$ and a set of m receivers $\{\mathsf{Rx}_j\}_{j=1}^m$, where each transmitter Tx_i intends to communicate to its corresponding receiver Rx_i. The channel gain between each transmitter Tx_i and each receiver Rx_j in the network is a random variable denoted by h_{ij}. We collect all the channel gains across the network in a square matrix, denoted

This work was supported in part by ARL DCIST CRA under Grant W911NF-17-2-0181.

by $\mathbf{H} \in \mathcal{H} \subseteq \mathbb{C}^{m\times m}$. The channel gain matrix $\mathbf{H}$ is assumed to consist of two components: a long-term component $\mathbf{H}^\ell \in \mathcal{H}^\ell \subseteq \mathbb{R}^{m\times m}$, resulting from path loss and shadowing, and a short-term fast fading component $\mathbf{H}^s \in \mathcal{H}^s \subseteq \mathbb{C}^{m\times m}$. Together, these two fading components form the full fading state as the channel matrix $\mathbf{H} = \mathbf{H}^\ell \mathbf{H}^s$.

Assuming that all transmissions occur at the same time and on the same frequency band, they will cause interference on each other. Therefore, each transmitter needs to set its transmit power so as to optimize a global, network-wide objective, such as sum-throughput. In particular, given a maximum transmit power of $P_{\max}$, we denote the vector of power allocation variables by $\mathbf{p} \in [0, P_{\max}]^m$, whose i^{th} component, p_i, represents the transmit power allocated to transmitter Tx_i. This implies that the signal-to-interference-plus-noise ratio (SINR) at each receiver Rx_i can be written as

$$\mathsf{SINR}_i(\mathbf{H}, \mathbf{p}) = \frac{|h_{ii}|^2 p_i}{\sigma^2 + \sum_{j\in[m]\setminus\{i\}} |h_{ji}|^2 p_j}, \ \forall i \in [m], \quad (1)$$

where σ^2 denotes the noise variance, and $[m]$ is defined as $[m] \coloneqq \{1, ..., m\}$. The Shannon capacity of the link between transmitter Tx_i and receiver Rx_i is then given by

$$f_i(\mathbf{H}, \mathbf{p}) = \log_2(1 + \mathsf{SINR}_i(\mathbf{H}, \mathbf{p})). \quad (2)$$

Due to channel variations over time, the power allocation variables also need to be modified temporally. The goal is to then determine a power allocation policy $\mathbf{p}(\mathbf{H})$ that takes as input an instantaneous channel realization $\mathbf{H}$ and determines the power levels $\mathbf{p}(\mathbf{H}) = [p_1(\mathbf{H}) \ \dots \ p_m(\mathbf{H})]^T$. Performance is then observed via the ergodic average $x_i(\mathbf{H}^\ell) = \mathbb{E}_{\mathbf{H}^s}[f_i(\mathbf{H}, \mathbf{p})]$, which denotes the throughput experienced by each receiver Rx_i over a long period of time for a given network configuration specified by large-scale fading $\mathbf{H}^\ell$.

We formulate the power allocation problem as follows. For each network configuration specified by $\mathbf{H}^\ell$, we seek to maximize a concave utility $\mathcal{U}(\mathbf{x}(\mathbf{H}^\ell))$ subject to a minimum ergodic rate constraint $f_{\min}$ for each user. These minimum capacity constraints are included so as to avoid allocating all resources to "cell-center" receivers, hence balancing the power control policy to treat "cell-center" and "cell-edge" receivers *fairly*. Since $f_{\min}$ is difficult to define a priori and may change across network configurations, we further include a constraint slack $\mathbf{z}(\mathbf{H}^\ell)$ for each configuration. This slack is itself minimized by incurring a negative utility. With the total utility averaged over large-scale fading configurations, we obtain

$$\max_{\mathbf{p},\mathbf{x},\mathbf{z}} \quad \mathbb{E}_{\mathbf{H}^\ell}\left[\mathcal{U}(\mathbf{x}(\mathbf{H}^\ell)) - \frac{\alpha}{2}\|\mathbf{z}(\mathbf{H}^\ell)\|_2^2\right], \quad (3a)$$

$$\text{s.t.} \quad \mathbf{x}(\mathbf{H}^\ell) = \mathbb{E}_{\mathbf{H}^s}\left[\mathbf{f}(\mathbf{H}^\ell\mathbf{H}^s, \mathbf{p}(\mathbf{H}^\ell\mathbf{H}^s))\right], \ \forall \mathbf{H}^\ell \quad (3b)$$

$$\mathbf{x}(\mathbf{H}^\ell) \geq f_{\min} - \mathbf{z}(\mathbf{H}^\ell), \ \forall \mathbf{H}^\ell \quad (3c)$$

$$\mathbf{p}(\mathbf{H}) \in [0, P_{\max}]^m, \ \mathbf{x}(\mathbf{H}^\ell) \geq \mathbf{0}, \ \mathbf{z}(\mathbf{H}^\ell) \geq \mathbf{0}. \quad (3d)$$

In (3), to solve for the optimal policy $\mathbf{p}(\cdot)$ we additionally optimize $\mathbf{x}(\mathbf{H}^\ell)$ and $\mathbf{z}(\mathbf{H}^\ell)$.

3. PRIMAL-DUAL LEARNING FRAMEWORK

To address the existence of constraints in (3), we reformulate it in the Lagrangian dual domain. Despite the non-convexity of the capacity function (2), it is known that under mild conditions on the channel distributions, the RRM problem nonetheless exhibits zero duality gap [16]. We can then proceed with the following reformulation without any loss in optimality.

We first introduce the Lagrangian function, with non-negative dual multiplier functions $\boldsymbol{\lambda} : \mathcal{H}^\ell \to \mathbb{R}_+^n$ and $\boldsymbol{\mu} : \mathcal{H}^\ell \to \mathbb{R}_+^n$ associated with each constraint in (3), as

$$\begin{aligned}
&\mathcal{L}(\mathbf{p}, \mathbf{x}, \mathbf{z}, \boldsymbol{\lambda}, \boldsymbol{\mu}) \\
&\coloneqq \mathbb{E}_{\mathbf{H}^\ell}\left[\mathcal{U}(\mathbf{x}(\mathbf{H}^\ell)) - \frac{\alpha}{2}\|\mathbf{z}(\mathbf{H}^\ell)\|_2^2\right] \\
&\quad - \int_{\mathcal{H}^\ell} \Big\{\boldsymbol{\lambda}(\mathbf{H}^\ell)^T\left[\mathbf{x}(\mathbf{H}^\ell) - \mathbb{E}_{\mathbf{H}^s}\left[\mathbf{f}(\mathbf{H}, \mathbf{p}(\mathbf{H}))\right]\right] \\
&\qquad + \boldsymbol{\mu}(\mathbf{H}^\ell)^T\left[f_{\min} - \mathbf{z}(\mathbf{H}^\ell) - \mathbf{x}(\mathbf{H}^\ell)\right]\Big\} d\mathbf{H}^\ell. \quad (4)
\end{aligned}$$

The Lagrangian in (4) provides a single, unconstrained objective function, which we can optimize using gradient-based methods. The dual problem involves maximizing (4) over the primal variables $\mathbf{p}, \mathbf{x}, \mathbf{z}$, while subsequently minimizing over dual variables $\boldsymbol{\lambda}, \boldsymbol{\mu}$. Due to the aforementioned zero duality gap property of the RRM problem, the solution incurs no loss in optimality relative to the original problem (3). The infinite dimensionality of the primal and dual variables in (4), however, renders the dual problem largely intractable.

Taking a standard statistical learning approach, we replace the infinite-dimensional functional optimization with optimization over a finite-dimensional parameterized form. Recall that, in the given problem, we seek to optimize not only the RRM policy $\mathbf{p}(\mathbf{H})$, but also the policies that specify long-term average rates, minimum capacity slacks, and dual weight functions. We thus replace each policy $\mathbf{y}(\cdot)$ with a respective parameterization $\mathbf{y}(\cdot; \boldsymbol{\theta}^{\mathbf{y}})$ that is fully specified by a finite-dimensional parameter vector $\boldsymbol{\theta}^{\mathbf{y}} \in \mathbb{R}^{q_{\mathbf{y}}}$. With this substitution, we obtain the parameterized Lagrangian as

$$\begin{aligned}
&\mathcal{L}_\theta\left(\boldsymbol{\theta}^{\mathbf{p}}, \boldsymbol{\theta}^{\mathbf{x}}, \boldsymbol{\theta}^{\mathbf{z}}, \boldsymbol{\theta}^{\boldsymbol{\lambda}}, \boldsymbol{\theta}^{\boldsymbol{\mu}}\right) \\
&\coloneqq \mathcal{L}\left(\mathbf{p}(\cdot; \boldsymbol{\theta}^{\mathbf{p}}), \mathbf{x}(\cdot; \boldsymbol{\theta}^{\mathbf{x}}), \mathbf{z}(\cdot; \boldsymbol{\theta}^{\mathbf{z}}), \boldsymbol{\lambda}(\cdot; \boldsymbol{\theta}^{\boldsymbol{\lambda}}), \boldsymbol{\mu}(\cdot; \boldsymbol{\theta}^{\boldsymbol{\mu}})\right). \quad (5)
\end{aligned}$$

Observe in (5) that we have used as inputs to the standard Lagrangian in (4) the parameterized policy definitions for both primal and dual policies—see [17]. The resulting parameterized Lagrangian function is now a function of the respective finite-dimensional parameter vectors. The parameterized dual problem is then defined as

$$D^*_{\boldsymbol{\theta}} \coloneqq \min_{\boldsymbol{\theta}^{\boldsymbol{\lambda}}, \boldsymbol{\theta}^{\boldsymbol{\mu}}} \ \max_{\boldsymbol{\theta}^{\mathbf{p}}, \boldsymbol{\theta}^{\mathbf{x}}, \boldsymbol{\theta}^{\mathbf{z}}} \ \mathcal{L}_\theta\left(\boldsymbol{\theta}^{\mathbf{p}}, \boldsymbol{\theta}^{\mathbf{x}}, \boldsymbol{\theta}^{\mathbf{z}}, \boldsymbol{\theta}^{\boldsymbol{\lambda}}, \boldsymbol{\theta}^{\boldsymbol{\mu}}\right). \quad (6)$$

We can now define the updates over an iteration index k for each of the primal and dual variables by either adding or

subtracting the partial gradient of $\mathcal{L}_{\boldsymbol{\theta}}\left(\boldsymbol{\theta}^{\mathbf{p}}, \boldsymbol{\theta}^{\mathbf{x}}, \boldsymbol{\theta}^{\mathbf{z}}, \boldsymbol{\theta}^{\boldsymbol{\lambda}}, \boldsymbol{\theta}^{\boldsymbol{\mu}}\right)$ with respect to that variable. For the power allocation and long-term average rate policies, this gives us the updates,

$$\boldsymbol{\theta}^{\mathbf{p}}_{k+1} = \boldsymbol{\theta}^{\mathbf{p}}_{k} + \eta_{\mathbf{p}} \nabla_{\boldsymbol{\theta}^{\mathbf{p}}} \mathcal{L}_{\boldsymbol{\theta}}\left(\boldsymbol{\theta}^{\mathbf{p}}, \boldsymbol{\theta}^{\mathbf{x}}, \boldsymbol{\theta}^{\mathbf{z}}, \boldsymbol{\theta}^{\boldsymbol{\lambda}}, \boldsymbol{\theta}^{\boldsymbol{\mu}}\right), \quad (7)$$

$$\boldsymbol{\theta}^{\mathbf{x}}_{k+1} = \boldsymbol{\theta}^{\mathbf{x}}_{k} + \eta_{\mathbf{x}} \nabla_{\boldsymbol{\theta}^{\mathbf{x}}} \mathcal{L}_{\boldsymbol{\theta}}\left(\boldsymbol{\theta}^{\mathbf{p}}, \boldsymbol{\theta}^{\mathbf{x}}, \boldsymbol{\theta}^{\mathbf{z}}, \boldsymbol{\theta}^{\boldsymbol{\lambda}}, \boldsymbol{\theta}^{\boldsymbol{\mu}}\right), \quad (8)$$

where $\eta_{\mathbf{p}}, \eta_{\mathbf{x}} > 0$ denote learning rates corresponding to the primal variables $\boldsymbol{\theta}^{\mathbf{p}}$ and $\boldsymbol{\theta}^{\mathbf{x}}$, respectively. We also include an additional step to update the slack policy parameters,

$$\boldsymbol{\theta}^{\mathbf{z}}_{k+1} = \boldsymbol{\theta}^{\mathbf{z}}_{k} + \eta_{\mathbf{z}} \nabla_{\boldsymbol{\theta}^{\mathbf{z}}} \mathcal{L}_{\boldsymbol{\theta}}\left(\boldsymbol{\theta}^{\mathbf{p}}, \boldsymbol{\theta}^{\mathbf{x}}, \boldsymbol{\theta}^{\mathbf{z}}, \boldsymbol{\theta}^{\boldsymbol{\lambda}}, \boldsymbol{\theta}^{\boldsymbol{\mu}}\right), \quad (9)$$

where $\eta_{\mathbf{z}} > 0$ denotes learning rate corresponding to the slack variables $\boldsymbol{\theta}^{\mathbf{z}}$. Likewise, we descend on the dual variables using the associated partial gradients of the Lagrangian, i.e.,

$$\boldsymbol{\theta}^{\boldsymbol{\lambda}}_{k+1} = \boldsymbol{\theta}^{\boldsymbol{\lambda}}_{k} - \eta_{\boldsymbol{\lambda}} \nabla_{\boldsymbol{\theta}^{\boldsymbol{\lambda}}} \mathcal{L}_{\boldsymbol{\theta}}\left(\boldsymbol{\theta}^{\mathbf{p}}, \boldsymbol{\theta}^{\mathbf{x}}, \boldsymbol{\theta}^{\mathbf{z}}, \boldsymbol{\theta}^{\boldsymbol{\lambda}}, \boldsymbol{\theta}^{\boldsymbol{\mu}}\right), \quad (10)$$

$$\boldsymbol{\theta}^{\boldsymbol{\mu}}_{k+1} = \boldsymbol{\theta}^{\boldsymbol{\mu}}_{k} - \eta_{\boldsymbol{\mu}} \nabla_{\boldsymbol{\theta}^{\boldsymbol{\mu}}} \mathcal{L}_{\boldsymbol{\theta}}\left(\boldsymbol{\theta}^{\mathbf{p}}, \boldsymbol{\theta}^{\mathbf{x}}, \boldsymbol{\theta}^{\mathbf{z}}, \boldsymbol{\theta}^{\boldsymbol{\lambda}}, \boldsymbol{\theta}^{\boldsymbol{\mu}}\right), \quad (11)$$

with $\eta_{\boldsymbol{\lambda}}, \eta_{\boldsymbol{\mu}} > 0$ respectively denoting learning rates corresponding to the dual variables $\boldsymbol{\theta}^{\boldsymbol{\lambda}}$ and $\boldsymbol{\theta}^{\boldsymbol{\mu}}$. Note that the proposed method is *unsupervised* as we update the primal, slack, and dual variables to optimize the objective function and constraints in (3) directly rather than with labeled solutions.

4. GNN-BASED RRM POLICIES

The choice of parameterization functions is critical in achieving optimal RRM policies with good practical performance when solving (3). Fully-connected deep neural networks (DNNs) are a proper choice here, due to their universality property [8, 18]. However, despite their theoretical properties, such a parameterization does not scale well—as the parameter dimension (particularly in the input and output layers) grows with the network size—and more critically does not generalize over varying network topologies.

Here we propose a scalable graph neural network (GNN) architecture suitable for solving the RRM problem in networks of any size. In particular, we propose to use GNNs as parameterizations for the primal and dual policies outlined in Section 3. We consider the data structure given in the form of a directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, r, w)$, where $\mathcal{V}$ denotes the set of graph nodes, connected by directed edges in $\mathcal{E}$, and $r : \mathcal{V} \to \mathbb{R}^{F_0}$ and $w : \mathcal{E} \to \mathbb{R}$ respectively denote functions that determine the initial node features and edge weights. We let $\mathcal{V} = [m]$, where each node represents a specific transmitter-receiver pair, and we define $\mathcal{E} = \{(i,j) \in \mathcal{V}^2 : i \neq j\}$.

For each node $v \in \mathcal{V}$, let $\mathbf{y}^0_v = r(v) \in \mathbb{R}^{F_0}$ denote the initial feature vector corresponding to node v. These feature vectors then go through multiple rounds, i.e., *layers*, of message passing and aggregation along the edges of the graph $\mathcal{G}$. Denoting the number of such layers by L, for every layer $l \in \{1, \ldots, L\}$, we denote the feature vector of each node v by $\mathbf{y}^l_v \in \mathbb{R}^{F_l}$, which is derived as an aggregation of the features of node v and its neighbors from the previous layer, as well as its *incoming* edge weights. More specifically, we have

$$\mathbf{y}^l_v = \Psi^l\left(\mathbf{y}^{l-1}_v, \left\{\mathbf{y}^{l-1}_u, w_{uv}\right\}_{u \in \mathcal{V}:(u,v)\in\mathcal{E}}; \boldsymbol{\theta}^l\right), \quad (12)$$

where $\Psi^l(\cdot; \boldsymbol{\theta}^l)$ denotes a potentially non-linear aggregation function, parameterized through a set of parameters $\boldsymbol{\theta}^l$. At the end of the L^{th} layer, each node will have a final feature vector $\mathbf{s}_v \coloneqq \mathbf{y}^L_v \in \mathbb{R}^{F_L}$, which we refer to as its *node embedding*.

We use the aforementioned architecture as parameterizations for the five primal and dual policies mentioned in Section 3. As discussed before, the primal power allocation policy $\mathbf{p}(\cdot)$ takes as input the instantaneous fading state, $\mathbf{H} = \mathbf{H}^{\ell}\mathbf{H}^{s}$, while the remaining four policies, namely $\mathbf{x}(\cdot)$, $\mathbf{z}(\cdot)$, $\boldsymbol{\lambda}(\cdot)$, and $\boldsymbol{\mu}(\cdot)$, take as input the large-scale fading state $\mathbf{H}^{\ell}$. We define three GNNs to implement these policies as follows:

- **Main GNN (responsible for $\mathbf{p}(\cdot)$):** For this GNN, we construct the graph $\mathcal{G}$ using the instantaneous channel gains $\mathbf{H}$. After L aggregation layers, we let $\{\mathbf{s}_v\}_{v\in\mathcal{V}}$ denote the node embeddings for all the nodes in the graph, where each node embedding lies in $\mathbb{R}^{F_L}$. Then, for each node $i \in \mathcal{V}$ corresponding to transmitter-receiver pair $(\mathsf{Tx}_i, \mathsf{Rx}_i)$, we define

$$p_i(\mathbf{H}) = P_{\max} \cdot \sigma\left(\mathbf{b}^T_{\mathbf{p}} \mathbf{s}_i\right), \quad (13)$$

  where $\sigma(\cdot)$ denotes the sigmoid function $\sigma(x) = \frac{1}{1+e^{-x}}$, and $\mathbf{b}_{\mathbf{p}} \in \mathbb{R}^{F_L}$ is a parameter vector mapping the transmitter node embeddings to its allocated transmit power. This is the only GNN used during the inference phase, and at each time step, its inference complexity for the entire network scales as $\mathcal{O}(m^2 \sum_{l=1}^{L+1} F_l F_{l-1})$ with $F_{L+1} = 1$.

- **Auxiliary GNN (responsible for $\mathbf{x}(\cdot)$ and $\mathbf{z}(\cdot)$):** The underlying graph for this GNN is defined using the large-scale fading state $\mathbf{H}^{\ell}$. We assume that the auxiliary GNN consists of L' aggregation layers and for each node $i \in \mathcal{V}$, we let $\mathbf{s}'_i \in \mathbb{R}^{F'_{L'}}$ denote its resulting node embedding. Then, we introduce two parameter vectors, $\mathbf{b}_{\mathbf{x}}$ and $\mathbf{b}_{\mathbf{z}}$, both residing in $\mathbb{R}^{F'_{L'}}$, which determine the policy outputs as follows:

$$x_i(\mathbf{H}^{\ell}) = \mathbf{b}^T_{\mathbf{x}} \mathbf{s}'_i, \quad (14)$$

$$z_i(\mathbf{H}^{\ell}) = \left[\mathbf{b}^T_{\mathbf{z}} \mathbf{s}'_i\right]_+, \quad (15)$$

  where $[\cdot]_+$ stands for $\max(\cdot, 0)$.

- **Dual GNN (responsible for $\boldsymbol{\lambda}(\cdot)$ and $\boldsymbol{\mu}(\cdot)$):** The underlying graph for this GNN is identical to the one for the auxiliary GNN. We assume that the dual GNN consists of L'' aggregation layers and for each node $i \in \mathcal{V}$, we let $\mathbf{s}''_i \in \mathbb{R}^{F''_{L''}}$ denote its resulting node embedding. Then, we introduce two parameter vectors, $\mathbf{b}_{\boldsymbol{\lambda}}$, and $\mathbf{b}_{\boldsymbol{\mu}}$, all residing in $\mathbb{R}^{F''_{L''}}$, which determine the dual policy outputs as follows:

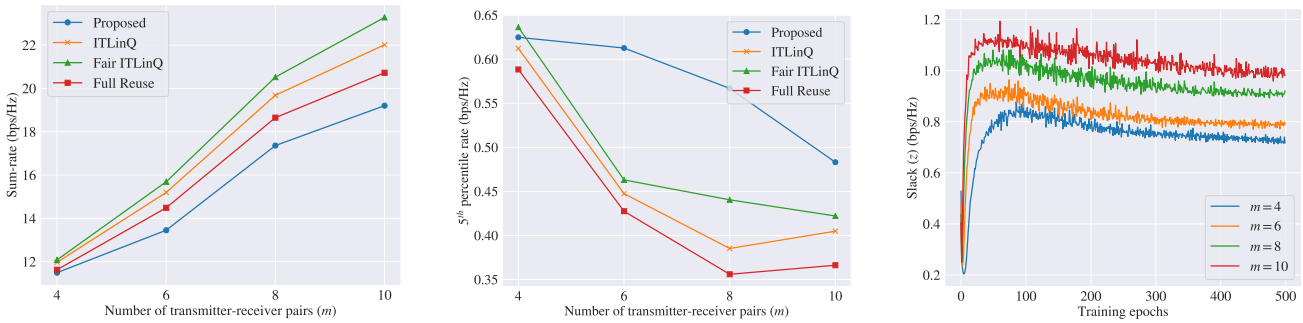

Fig. 1: Comparison of the proposed method with baseline algorithms in terms of (a) sum-rate and (b) 5^{th} percentile rate, and (c) the convergence behavior of the slack variable over the course of training.

$$\lambda_i(\mathbf{H}^\ell) = \left[\mathbf{b}_{\boldsymbol{\lambda}}^T \mathbf{s}_i''\right]_+, \tag{16}$$

$$\mu_i(\mathbf{H}^\ell) = \left[\mathbf{b}_{\boldsymbol{\mu}}^T \mathbf{s}_i''\right]_+. \tag{17}$$

5. SIMULATION RESULTS

We consider networks with $m \in \{4, 6, 8, 10\}$ transmitter-receiver pairs, dropped randomly within a $500m \times 500m$ area. We drop the transmitters and receivers uniformly at random within the network area, and ensure minimum inter-transmitter and transmitter-receiver distances of $35m$ and $10m$, respectively. The long-term fading comprises a dual-slope path-loss model [19] and log-normal shadowing with 7dB standard deviation. We also model short-term Rayleigh fading using the sum of sinusoids (SoS) [20]. The bandwidth is set to 10MHz, the noise power spectral density is assumed to be -174dBm/Hz, and the maximum transmit power is set to $P_{\max} = 10$dBm. We use a sum-rate utility function $\mathcal{U}(\mathbf{x}) = \sum_{i=1}^m x_i$, and we set $f_{\min} = 2$bps/Hz and $\alpha = 2$.

For all three GNNs, we use the GNN architecture proposed in [21] with 2 hidden layers, each containing 512 features. For the auxiliary and dual GNNs, we set $F_0' = F_0'' = 1$, where the initial feature of each node is the large-scale signal-to-noise ratio (SNR) of the corresponding transmitter-receiver pair (in dB), normalized by the Frobenius norm of the large-scale fading state (in dB). For the main GNN, we set $F_0 = 2$, where each nodes' initial features include the normalized SNR of the corresponding transmitter-receiver pair (in dB), and the normalized *exponential moving average rate* of the corresponding receiver. We set the weight of each edge to the normalized interference-to-noise ratio (INR) between the source transmitter and the destination receiver (in dB).

We use 512 training configurations, and 128 test configurations. Each configuration is run for 150 steps, where the first 50 steps is used as a warm-up period, in which all transmitters use full power to serve their receivers, and the subsequent 100 steps are used for training the policies. We set the batch size to 128. The learning rates for the main, auxiliary, and dual GNNs are initially set to 5×10^{-3}, 2×10^{-3}, and 5×10^{-4}, respectively, and they are reduced by a factor of 0.9 every 50 epochs. We compare the performance of our proposed method with the regular and Fair ITLinQ[1] methods [3], and full reuse, where all transmitters transmit with full power. For numerical stability, we clip the term $f_{\min} - \mathbf{z}(\mathbf{H}^\ell) - \mathbf{x}(\mathbf{H}^\ell)$ in the Lagrangian (4) from below at -0.05, so as to stabilize the learning process for the slack and dual policies.

Figures 1a and 1b compare the sum-rate and 5^{th} percentile rate of our proposed method and the baseline methods. As the figure shows, while slightly hurting in terms of sum-rate as compared to the baseline methods, our proposed method provides significant gains in terms of the 5^{th} percentile rate. This is thanks to the fact that, as shown in Figure 1c, our proposed algorithm learns how to adaptively elevate the slack variable so as to make the optimization problem feasible and maximize the sum-rate utility function. This leads to a much fairer resource allocation policy, where a superior trade-off between the sum-rate and the 5^{th} percentile rate can be achieved.

6. CONCLUDING REMARKS

We considered the problem of power control in wireless networks with multiple transmitter-receiver pairs. To balance fairness across users while maximizing the sum rate, we formulate a constrained optimization problem with per-user minimum rate constraints and tunable slack parameters. The resulting Lagrangian function features infinite-dimensional primal and dual variables, which are parameterized to facilitate the application of unsupervised learning algorithms. We used three graph neural networks to parameterize the primal and dual policies with node features and edge weights derived from the channel gains across the network. We then proposed a primal-dual optimization algorithm, which learns the parameters of the GNNs iteratively with stochastic gradient updates. Simulation results show the superiority of our proposed algorithm compared to baseline methods in terms of the trade-off between aggregate and 5^{th} percentile user rates.

[1] We tune the additional Fair ITLinQ hyperparameters (i.e., $\text{SNR}_{th}, \bar{M}, \bar{\eta}$) using grid search to achieve the best performance.

7. REFERENCES

[1] Qingjiang Shi, Meisam Razaviyayn, Zhi-Quan Luo, and Chen He, "An iteratively weighted MMSE approach to distributed sum-utility maximization for a MIMO interfering broadcast channel," *IEEE Transactions on Signal Processing*, vol. 59, no. 9, pp. 4331–4340, 2011.

[2] Wei Yu, Taesoo Kwon, and Changyong Shin, "Multicell coordination via joint scheduling, beamforming, and power spectrum adaptation," *IEEE Transactions on Wireless Communications*, vol. 12, no. 7, 2013.

[3] Navid Naderializadeh and Amir Salman Avestimehr, "ITLinQ: A new approach for spectrum sharing in device-to-device communication systems," *IEEE journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1139–1151, 2014.

[4] Kaiming Shen and Wei Yu, "FPLinQ: A cooperative spectrum sharing strategy for device-to-device communications," in *2017 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2017, pp. 2323–2327.

[5] Yifei Shen, Jun Zhang, SH Song, and Khaled B Letaief, "AI empowered resource management for future wireless networks," *arXiv preprint arXiv:2106.06178*, 2021.

[6] Haoran Sun, Xiangyi Chen, Qingjiang Shi, Mingyi Hong, Xiao Fu, and Nikos D Sidiropoulos, "Learning to optimize: Training deep neural networks for wireless resource management," in *2017 IEEE 18th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*. IEEE, 2017, pp. 1–6.

[7] Yifei Shen, Yuanming Shi, Jun Zhang, and Khaled B Letaief, "A graph neural network approach for scalable wireless power control," in *2019 IEEE Globecom Workshops (GC Wkshps)*. IEEE, 2019, pp. 1–6.

[8] Mark Eisen, Clark Zhang, Luiz FO Chamon, Daniel D Lee, and Alejandro Ribeiro, "Learning optimal resource allocations in wireless systems," *IEEE Transactions on Signal Processing*, vol. 67, no. 10, pp. 2775–2790, 2019.

[9] Mark Eisen and Alejandro Ribeiro, "Optimal wireless resource allocation with random edge graph neural networks," *IEEE Transactions on Signal Processing*, vol. 68, pp. 2977–2991, 2020.

[10] Mengyuan Lee, Guanding Yu, and Geoffrey Ye Li, "Graph embedding-based wireless link scheduling with few training samples," *IEEE Transactions on Wireless Communications*, vol. 20, no. 4, pp. 2282–2294, 2020.

[11] Navid Naderializadeh, Jaroslaw J Sydir, Meryem Simsek, and Hosein Nikopour, "Resource management in wireless networks via multi-agent deep reinforcement learning," *IEEE Transactions on Wireless Communications*, vol. 20, no. 6, pp. 3507–3523, 2021.

[12] Akash Doshi, Srinivas Yerramalli, Lorenzo Ferrari, Taesang Yoo, and Jeffrey G Andrews, "A deep reinforcement learning framework for contention-based spectrum sharing," *IEEE Journal on Selected Areas in Communications*, 2021.

[13] Navid Naderializadeh, "Contrastive self-supervised learning for wireless power control," in *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2021, pp. 4965–4969.

[14] Ivana Nikoloska and Osvaldo Simeone, "Fast power control adaptation via meta-learning for random edge graph neural networks," *arXiv preprint arXiv:2105.00459*, 2021.

[15] Luiz FO Chamon, Alexandre Amice, Santiago Paternain, and Alejandro Ribeiro, "Resilient control: Compromising to adapt," in *59th IEEE Conference on Decision and Control (CDC)*. IEEE, 2020, pp. 5703–5710.

[16] Alejandro Ribeiro, "Optimal resource allocation in wireless communication and networking," *EURASIP Journal on Wireless Communications and Networking*, vol. 2012, no. 1, pp. 1–19, 2012.

[17] Chengjian Sun, Changyang She, and Chenyang Yang, "Unsupervised deep learning for optimizing wireless systems with instantaneous and statistic constraints," *arXiv preprint arXiv:2006.01641*, 2020.

[18] Kurt Hornik, Maxwell Stinchcombe, and Halbert White, "Multilayer feedforward networks are universal approximators," *Neural Networks*, vol. 2, no. 5, 1989.

[19] Jeffrey G Andrews, Xinchen Zhang, Gregory D Durgin, and Abhishek K Gupta, "Are we approaching the fundamental limits of wireless network densification?," *IEEE Communications Magazine*, vol. 54, no. 10, pp. 184–190, 2016.

[20] Yunxin Li and Xiaojing Huang, "The simulation of independent Rayleigh faders," *IEEE Transactions on Communications*, vol. 50, no. 9, pp. 1503–1514, 2002.

[21] Ekagra Ranjan, Soumya Sanyal, and Partha Talukdar, "ASAP: Adaptive structure aware pooling for learning hierarchical graph representations," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, no. 04, pp. 5470–5477, 2020.