

# Learning Based User Scheduling in Reconfigurable Intelligent Surface Assisted Multiuser Downlink

Zhongze Zhang<sup>✉</sup>, *Graduate Student Member, IEEE*, Tao Jiang<sup>✉</sup>, *Graduate Student Member, IEEE*,  
and Wei Yu<sup>✉</sup>, *Fellow, IEEE*

**Abstract**—Reconfigurable intelligent surface (RIS) is capable of intelligently manipulating the phases of the incident electromagnetic wave to improve the wireless propagation environment between the base-station (BS) and the users. This paper addresses the joint user scheduling, RIS configuration, and BS beamforming problem in an RIS-assisted downlink network with limited pilot overhead. We show that graph neural networks (GNN) with permutation invariant and equivariant properties can be used to appropriately schedule users and to design RIS configurations to achieve high overall throughput while accounting for fairness among the users. As compared to the conventional methodology of first estimating the channels then optimizing the user schedule, RIS configuration and the beamformers, this paper shows that an optimized user schedule can be obtained directly from a very short set of pilots using a GNN, then the RIS configuration can be optimized using a second GNN, and finally the BS beamformers can be designed based on the overall effective channel. Numerical results show that the proposed approach can utilize the received pilots more efficiently than the conventional channel estimation based approach, and can generalize to systems with an arbitrary number of users.

**Index Terms**—Deep learning, graph neural network, proportional fairness, reconfigurable intelligent surface, scheduling.

## I. INTRODUCTION

RECONFIGURABLE intelligent surface (RIS) is envisioned as a key enabling technology for a smarter radio environment [2]–[6], due to its capability to manipulate the phases of wireless signals to enhance the transmission environment and to improve the network utility (e.g., sum rate [7] or minimum rate [8]–[10]). This paper addresses a key problem of user scheduling for the RIS-assisted wireless cellular network. In a downlink transmission environment with a base-station (BS) equipped with  $M$  antennas, which can serve at most  $M$  users simultaneously over the same resource block, if the total number of users in the network  $K$  is greater than  $M$ , how

should the BS optimally schedule a subset of users at each timeslot in conjunction with the optimal RIS configuration and the BS beamforming to achieve high network throughput while ensuring fairness across the users?

The conventional approach to user scheduling typically follows a two-step approach of aiming to accurately recover the channel state information (CSI) in the first step, then based on the estimated CSI, optimizing downlink scheduling in the second step. However, the two-step approach is not necessarily optimal as it suffers from three major shortcomings. First, accurate recovery of a large number of unknown channel coefficients is costly in terms of pilot training overhead. The number of channel coefficients that need to be estimated scales with the size of RIS, which typically consists of hundreds of passive elements. Pilot training overhead occupies resources within the channel coherence period that could otherwise be used for data transmission. Second, the channel estimation process typically aims to recover the channel according to some arbitrary metric, e.g., mean squared error, which may not necessarily align with the ultimate communication objective. Finally, even if the CSI is known perfectly, the user achievable rates are nonconvex functions of the RIS configuration and BS beamformers; further, the scheduling problem is discrete. So, the overall problem is a mixed discrete and continuous non-convex optimization problem, for which finding an optimal solution is computationally intensive and difficult to achieve in a real-time application.

To address these challenges, recent works have advocated learning based approaches to directly map the received pilot sequences to the RIS configuration and the beamformers that maximize the network objective [11], [12], in effect bypassing explicit channel estimation. However, the discrete optimization problem of user scheduling, an important but challenging part of the overall network optimization, is not considered in these works.

This paper focuses on the scheduling aspect of an RIS-assisted multiuser downlink network. Toward this end, we first make an observation that the overall optimization problem has the following permutation invariant and equivariant properties: if the ordering of users is permuted, the same subset of users should be scheduled and the same reflection coefficients should be used (i.e., permutation invariance), while the beamforming vectors should be permuted in the same way (i.e., permutation equivariance). This motivates us to make use of permutation invariant and equivariant graph neural networks (GNN) to learn the mapping from the received pilots and the user weights to the

Manuscript received 24 November 2021; revised 31 March 2022; accepted 30 April 2022. Date of publication 27 May 2022; date of current version 29 August 2022. This work was supported by Huawei Technologies Canada Ltd. Co. Part of this work was presented at the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2022 [DOI: 10.1109/ICASSP43922.2022.9746441]. The guest editor coordinating the review of this manuscript and approving it for publication was Dr. Marco Di Renzo. (Corresponding author: Zhongze Zhang.)

The authors are with the Edward S. Rogers Sr. Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON M5S3G4, Canada (e-mail: ufo.zhang@mail.utoronto.ca; taoca.jiang@mail.utoronto.ca; weiyu@ece.utoronto.ca).

Digital Object Identifier 10.1109/JSTSP.2022.3178213

optimized RIS configuration and beamforming vectors. This is a generalization of the neural network architecture proposed in the previous work [11], which deals only with beamforming and reflective coefficients design, to a system-level optimization that also accounts for the user priorities in the scheduling process. Scheduling in a multiuser network is in general challenging to learn, because the scheduling choice can be highly sensitive to the user priority weights. In this paper, we advocate the following three-stage system-level optimization approach. Specifically, we show that in the first stage, a GNN applied to all potential users but with very short pilots can already produce an optimized schedule while accounting for the user priorities. This is achieved by adopting an *implicit scheduling* scheme and by inferring the schedule from the set of output beamforming vectors. In the second stage, a second GNN applied only to the scheduled users but with longer pilots is used to design the RIS configurations. In the final stage, we insert additional pilots to estimate the overall effective channel (or if the pilot budget is limited, reuse the existing pilots), then re-optimize the beamformers. Numerical results show that the proposed algorithm can learn to maximize the network utility with significantly reduced pilot overhead as compared to the conventional channel estimation based approach, and can generalize to a network with an arbitrary number of users.

#### A. Related Work

User scheduling in RIS-assisted networks has not been thoroughly studied. Many existing proportional fairness based scheduling algorithms that have been successful in conventional networks without RIS [13]–[16] are difficult to directly migrate to an RIS-assisted network, because the RIS can work in collaboration with BS beamformers to improve channel gain or suppress interference for some users, thus influencing scheduling decisions. Moreover, most of the published works involving RIS only focus on joint beamforming and reflection coefficients design, where the BS optimizes beamforming matrix and RIS phase shift vector based on perfect CSI to improve energy efficiency [17]–[19], to maximize weighted sum rate [20]–[22], to maximize the minimum rate [8]–[10], to maximize the system throughput [23] or to maximize the received signal-to-interference-and-noise ratio (SINR) [24]. None of these works address the problem of user scheduling in dense RIS-assisted networks. Also, most of these works are based on the perfect CSI assumption, which is an unrealistic assumption as the RIS, due to its passive structure, is not equipped with radio frequency components to perform active pilot symbol transmission, or signal processing units as receivers to estimate the channel [3].

Recently, data-driven approaches have been proposed to address the challenge in CSI acquisition and transmission optimization. This paper is inspired by the successes in [11], [12], [25], which show that it is possible to bypass explicit channel estimation and to directly use the received pilot symbols to optimize the downlink beamforming configuration in their respective network setting. In an RIS-assisted network, [12] uses a deep neural network with supervised learning to approximate the mapping from the received pilot sequences to the optimized RIS

phase matrix and downlink beamforming vector for a single-user setting. In a multiuser setting, [11] uses a GNN to exploit the permutation invariant and permutation equivariant properties of multiuser sum-rate maximization problem, and to learn the mapping from the received pilot sequences to the RIS configuration and a set of beamforming vectors in an unsupervised fashion. In the hybrid beamforming network, [25] utilizes a deep neural network to design downlink analog beamformer based on the received pilot frames, thus bypassing channel estimation. Those works have shown that the rich information in the received pilot symbols is sufficient to directly design beamforming configuration. However, their proposed solutions do not account for user scheduling, which is a crucial yet challenging component of network design.

In [26], the authors perform link scheduling using a novel deep neural network architecture based solely on the geographic locations of the devices. Similarly, [27] makes use of location information and tackles link scheduling problem using a deep learning based graph embedding process. In [28], the link scheduling problem is solved using a graph convolutional network based solution with user interference relationship as the primary input. However, these existing works on scheduling are for device-to-device networks without the RIS. In this paper, we investigate the mapping from the received pilots to a user schedule and the associated RIS and beamforming design, in recognition of the rich channel information contained in the pilot symbols. Specifically, we adopt a data-driven approach to learn such a mapping, and propose a deep learning based multi-stage framework that uses the received pilot symbols as input to perform joint scheduling and beamforming in an RIS-assisted network with reduced pilot overhead.

#### B. Main Contribution

We propose to learn the mapping from the received pilot sequences to user scheduling, RIS configuration and beamformer using a GNN and train the model to directly maximize the network objective. A key motivation for adopting GNN as the scheduler is that it naturally captures the permutation invariant and equivariant properties of the optimization problem, i.e., regardless of the ordering of the users, it should result in the same scheduling with the same RIS reflection coefficients and the same beamforming vectors with permuted indices. GNN also allows generalizability across the users. For instance, the number of users in a dense network is constantly changing depending on the user traffic. If we adopt a fully connected neural network as the scheduler, the scheduler would not be able to generalize to different sizes of user pool except by re-training. In contrast, a GNN can generalize to different numbers of users by simply adding and removing components in its feature extraction and information exchange stages [29].

We propose a permutation invariant and equivariant GNN architecture for mapping the received pilots to the optimized variables. In particular, the RIS and the users are modelled as nodes in the GNN and each node is associated with a nodal feature, also known as the representation vector. The input features, i.e., received pilots, are encoded into the representation

vectors. Subsequently, the representation vectors are updated layer-by-layer by exchanging information with the neighboring nodes. After updating through multiple layers, the updated representation vectors contain the right representation of information to design the RIS configuration and beamformers, from which the schedule can be inferred.

However, a large fully connected GNN is difficult to train with high-dimensional features, because the functional landscape over a large number of users with a nonconvex objective is highly non-trivial. To address this scalability issue, we propose to use two GNNs. The first GNN takes very short received pilots and the user weights from all potential users as inputs in order to make the scheduling decisions in a first stage. We then use a second GNN to design the RIS configuration for the scheduled users in a second stage. The second GNN works on a smaller set of users, so it can take longer pilot sequences as nodal features.

Using short pilot sequences to determine the schedule in the first stage is a reasonable approach, because unlike beamforming and RIS configuration, scheduling is not a strong function of the channel realizations. Thus, coarse knowledge about the channel is already sufficient to enable the scheduler to choose the (near) optimal set of users. After the schedule is determined, a second GNN that takes input from the scheduled users only, but with longer pilots, are then used in the second stage to determine the RIS configurations.

Further, we propose a beamformer fine-tuning stage as the third stage to refine the beamformers using the weighted minimum mean square error (WMMSE) algorithm [30] based on the explicitly estimated channels for additional performance gain. Depending on the pilot overhead constraint, we adopt two choices of channel estimation schemes. The first scheme involves an additional pilot training with a fixed user schedule and fixed RIS coefficients. In this case, the low-dimensional effective channels of the scheduled users can be accurately estimated using additional short pilots in each scheduling instance. But if the total pilot overhead budget is limited, we can also employ a second channel estimation scheme which reuses the pilots in the previous stages to estimate the high-dimensional channels, then subsequently design the beamformers based on the combined channels. The conference version of this paper [1] considers only the first scheme. This journal version of the paper considers both channel estimation schemes which enables the proposed framework to adapt to both fast and slow-fading environments.

In summary, the main contributions of this paper are as follows:

- 1) Two separate GNNs are used first to schedule a subset of users, then to design the RIS reflection coefficients, so that together with the optimized BS beamformers, the users can achieve high throughput while ensuring fairness. Both GNNs use the received pilot symbols and the user priority weights as inputs. The model is permutation invariant with respect to the schedule and the RIS reflection coefficients, and permutation equivariant with respect to the beamforming vectors. It allows generalizability to different sizes of the user pool.
- 2) A hybrid data-driven and CSI-based framework is used, where the user schedule and the RIS reflection coefficients are designed using GNNs, and the beamformers

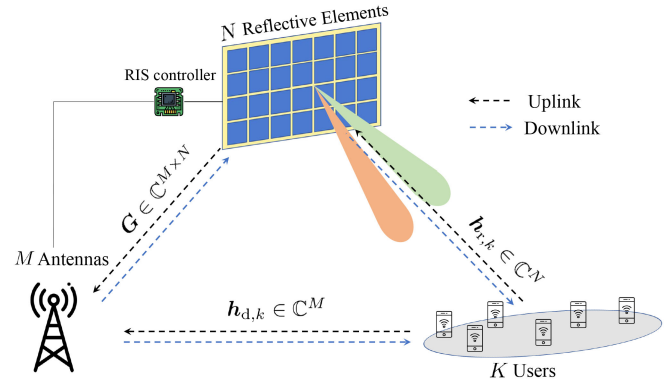


Fig. 1. RIS-assisted multiuser network.

are designed based on the estimated CSI. The framework contains three stages, where scheduling and RIS design are separated into the first and the second stages in order to reduce the neural network training complexity. A third stage is used to fine-tune the beamformers.

- 3) An uplink pilot placement strategy is proposed, in which the pilot symbols received over random setting of the uplink RIS phases are used to design the user schedule and the RIS reflection coefficients, with possibly additional pilot symbols (received over the optimized uplink RIS configuration) used to design the BS beamformers.
- 4) Numerical simulations show that the proposed framework achieves better network utility compared to the conventional two-step approach, and can generalize to scenarios with an arbitrary number of users.

### C. Organization of the Paper and Notations

The remaining paper is organized as follows. Section II introduces the system model, proportional fairness scheduling, and problem formulation. Section III describes the proposed pilot placement structure, deep learning framework and GNN architecture. Numerical results are provided in Section IV. The paper concludes with Section V.

*Notations:* We use  $a$ ,  $\mathbf{a}$ , and  $\mathbf{A}$  to denote scalar, vector, and matrix respectively;  $\mathbf{A}^\top$ ,  $\mathbf{A}^H$  and  $\mathbf{A}^{-1}$  to denote transpose, Hermitian and inverse;  $|\cdot|$  and  $(\cdot)^*$  to denote the modulus and conjugate;  $[a]_j$  to denote the  $j$ -th element of vector  $\mathbf{a}$ ;  $\text{vec}(\mathbf{A})$  to denote matrix  $\mathbf{A}$  in vector form; and  $\text{diag}(\mathbf{a})$  to denote the diagonal matrix with the entries of  $\mathbf{a}$  on the diagonal. We use  $\mathcal{R}(\cdot)$  and  $\mathcal{I}(\cdot)$  to denote the real and imaginary component of a complex value;  $\mathcal{CN}(\cdot, \cdot)$  to denote a complex Gaussian distribution;  $\mathbb{E}(\cdot)$  to denote the expectation of a random variable. Finally, we use  $|\mathcal{S}|$  to denote the cardinality of the set  $\mathcal{S}$ , and  $\mathbf{I}_b$  to denote a  $b \times b$  identity matrix.

## II. SYSTEM MODEL AND PROBLEM FORMULATION

### A. System Model

Consider a downlink RIS-assisted multiuser MISO network, in which a BS equipped with  $M$  antennas serves  $K$  single-antenna users as shown in Fig. 1. An RIS with  $N$  passive elements is placed between the BS and the users to enhance the SINR of the received signal at the users. The RIS reflection



coefficients are denoted as  $\boldsymbol{\theta} = [e^{j\delta_1}, e^{j\delta_2}, \dots, e^{j\delta_N}]^T \in \mathbb{C}^N$  with  $\delta_n \in [0, 2\pi)$  as the phase shift of the  $n$ -th element. These reflective coefficients can be controlled independently, from an RIS controller, typically located at the BS.

We consider the scenario in which the number of users in a cell is much larger than the number of antennas at the BS i.e.,  $K > M$ , whereas the BS can only serve at most  $M$  users simultaneously by employing spatial multiplexing. Thus, the users need to take turns to get served. The BS serves each set of scheduled users in what is called a scheduling timeslot. The duration of the scheduling slots is determined by the system latency requirement, and is typically in the order of milliseconds.

To maximize the network utility function while ensuring fairness amongst users, the scheduler needs to strategically choose a subset of users in each scheduling timeslot. Let  $\beta_k^t \in \{0, 1\}$  be the indicator variable that indicates the scheduling decision of the user  $k$  at the  $t$ -th scheduling timeslot

$$\beta_k^t = \begin{cases} 0 & \text{if not scheduled,} \\ 1 & \text{otherwise.} \end{cases} \quad (1)$$

Let  $\boldsymbol{\beta}^t = [\beta_1^t, \beta_2^t, \dots, \beta_K^t]^T \in \mathbb{C}^K$ . Then, we have  $\|\boldsymbol{\beta}^t\|_1 \leq M$  since at most  $M$  users are scheduled in a scheduling timeslot.

We adopt a *block-fading* model in which the channels are assumed to be constant across multiple scheduling timeslots within a coherence period, then change independently in subsequent coherent periods. The number of scheduling timeslots within the channel coherence period depends on the channel fading characteristics. If the channel is fast-fading with a relatively short channel coherence period, there would be only a few scheduling timeslots within the coherence period. Conversely, if the channel is slow fading, many more scheduling timeslots would be available within a channel coherence period.

The channel models are as shown in Fig. 1, where  $\mathbf{h}_{d,k} \in \mathbb{C}^M$  denotes the direct link channel from the BS to the user  $k$ ,  $\mathbf{h}_{r,k} \in \mathbb{C}^N$  denotes the reflected link channel from the RIS to the user  $k$ , and  $\mathbf{G} \in \mathbb{C}^{M \times N}$  denotes the common channel from the BS to the RIS.

Let  $s_k^t \in \mathbb{C}$  be the information symbol to be transmitted from the BS to the user  $k$  and let  $\mathbf{w}_k^t \in \mathbb{C}^M$  denote the associated beamforming vector of the user  $k$  at the  $t$ -th scheduling timeslot. We can denote the beamforming matrix at the BS as  $\mathbf{W}^t = [\mathbf{w}_1^t, \dots, \mathbf{w}_K^t] \in \mathbb{C}^{M \times K}$ . In each scheduling timeslot, we have  $\|\mathbf{w}_k^t\|^2 = 0$  if  $\beta_k^t = 0$ , otherwise  $\|\mathbf{w}_k^t\|^2 > 0$ . The transmitted signal at the BS at the  $t$ -th scheduling timeslot can be denoted as:

$$\mathbf{s}^t = \sum_{k=1}^K \mathbf{w}_k^t \beta_k^t s_k^t. \quad (2)$$

Consider a single  $t$ -th scheduling timeslot. The received signal at the user  $k$  is a combination of the signals from the direct path and the reflection path through the RIS, which is given by

$$\begin{aligned} r_k^t &= \mathbf{h}_{d,k}^\top \mathbf{s}^t + (\mathbf{G} \text{diag}(\boldsymbol{\theta}^t) \mathbf{h}_{r,k})^\top \mathbf{s}^t + n_k \\ &= (\mathbf{h}_{d,k} + \mathbf{A}_k \boldsymbol{\theta}^t)^\top \sum_{i=1}^K \mathbf{w}_i^t \beta_i^t s_i + n_k, \end{aligned} \quad (3)$$

where  $\mathbf{A}_k = \mathbf{G} \text{diag}(\mathbf{h}_{r,k}) \in \mathbb{C}^{M \times N}$  is the cascade channel between the BS and the user  $k$  through the reflection at the RIS,  $\boldsymbol{\theta}^t$  is the RIS reflection coefficients at the  $t$ -th scheduling timeslot, and  $n_k \sim \mathcal{CN}(0, \sigma_d^2)$  is the downlink additive white Gaussian noise.

The achievable rate of the user  $k$  at the  $t$ -th scheduling timeslot can be expressed as:

$$R_k^t = \log \left( 1 + \frac{|(\mathbf{h}_{d,k} + \mathbf{A}_k \boldsymbol{\theta}^t)^\top \mathbf{w}_k^t|^2 \beta_k^t}{\sum_{i \neq k} |(\mathbf{h}_{d,k} + \mathbf{A}_k \boldsymbol{\theta}^t)^\top \mathbf{w}_i^t|^2 \beta_i^t + \sigma_d^2} \right). \quad (4)$$

Here, we assume that the effective channel gains can be accurately estimated at the user side using a few additional downlink pilots prior to data transmission, so the rate in (4) is achievable [31].

### B. Proportional Fairness Scheduling

To ensure fairness in user scheduling, we define a network utility, which is a function of the long-term average rate achieved by each user, computed according to an exponentially weighted moving average:

$$\bar{R}_k^t = (1 - \gamma) \bar{R}_k^{t-1} + \gamma R_k^{t-1}, \quad (5)$$

where  $0 \leq \gamma \leq 1$  is the forgetting factor and  $\bar{R}_k^t$  is average rate of the  $k$ -th user prior to the  $t$ -th scheduling timeslot. Defining a non-decreasing and concave utility function  $U(\cdot)$  for each user, the network-utility maximization problem is that of maximizing

$$\sum_{k=1}^K U(\bar{R}_k^t). \quad (6)$$

In the proportional fairness scheduling problem, the utility function is chosen as  $U(\cdot) = \log(\cdot)$ . Assuming a small  $\gamma R_k^{t-1}$ , the maximization of log-utility can be approximated by the maximization of weighted sum-rate [32], so that problem (6) can be approximately solved by maximizing the following:

$$\sum_{k=1}^K \alpha_k^t R_k^t, \quad (7)$$

where the weight  $\alpha_k^t$  is the inverse of  $\bar{R}_k^t$ :

$$\alpha_k^t = \frac{\partial U(\bar{R}_k^t)}{\partial R} \bigg|_{R=\bar{R}_k^t} = \frac{\partial \log(\bar{R}_k^t)}{\partial R} \bigg|_{R=\bar{R}_k^t} = \frac{1}{\bar{R}_k^t}. \quad (8)$$

### C. Problem Formulation

To maximize the weighted sum rate in (7), it is necessary to acquire the knowledge about the channels. Assuming channel reciprocity and that the system operates in the time division duplex (TDD) mode, we rely on an uplink pilot transmission phase to gain information about the channel. The channel coherence period is therefore divided into two phases, i.e., Phase-I for uplink pilot training and Phase-II for downlink data transmission. Downlink data transmission phase consists of multiple scheduling timeslots as discussed in Section II-A.

In the pilot training phase, all  $K$  users transmit pilot sequences at the same time. Let the pilot sequence of the  $k$ -th user be  $x_k(\ell)$ ,

$\ell = 1, \dots, L$ . The BS receives a combination of the signal from the direct path and the signal reflected off the RIS, so the received pilots at the BS can be expressed as

$$\mathbf{y}(\ell) = \sum_{k=1}^K (\mathbf{h}_{d,k} + \mathbf{A}_k \boldsymbol{\theta}(\ell)) x_k(\ell) + \mathbf{n}(\ell), \ell = 1, \dots, L, \quad (9)$$

where  $\boldsymbol{\theta}(\ell)$  is the uplink RIS configuration at the  $\ell$ -th instance and  $\mathbf{n}(\ell) \sim \mathcal{CN}(0, \sigma_u^2 \mathbf{I})$  is the uplink additive noise vector.

In this paper, we focus on optimizing the system variables at each timeslot based on instantaneous CSI (in contrast to works that use statistical CSI, e.g., [33]). The conventional system design would first estimate the channels then optimize the system variables. A key observation here is that since the ultimate goal is to maximize the network utility (6), instead of explicitly estimating the CSI, we can pursue a data-driven approach to the system-level optimization problem by directly mapping the received pilots and the user weights to the optimized user schedule, RIS configuration, and the beamforming variables.

Conceptually, the optimization problem in each scheduling slot can be thought of as:

$$\begin{aligned} & \underset{(\boldsymbol{\beta}^t, \mathbf{W}^t, \boldsymbol{\theta}^t)}{\text{maximize}} \quad \sum_k \alpha_k^t R_k^t \\ & f(\{\mathbf{y}(\ell)\}_{\ell=1}^L, \boldsymbol{\alpha}^t) \\ & \text{subject to} \quad \sum_k \|\mathbf{w}_k^t\|^2 \leq P_d, |\boldsymbol{\theta}^t_n| = 1, \forall n, \\ & \quad \sum_k \beta_k^t \leq M, \beta_k^t \in \{0, 1\}, \forall k, \end{aligned} \quad (10)$$

where  $\boldsymbol{\alpha}_t = [\alpha_1^t, \dots, \alpha_K^t]^\top$  and  $P_d$  denotes the downlink transmission power constraint.

Finding the optimal functional mapping  $f(\cdot)$  for problem (10) is computationally challenging as it is a mixed discrete (scheduling) and continuous (RIS reflection coefficients, beamforming matrix) optimization problem with nonconvex objective and nonconvex constraints. Moreover, it is difficult to decouple the optimization variables as they are closely interrelated. For example, to minimize interference, we should schedule users whose channels are orthogonal, i.e., users who are geographically far from one another, but in order to leverage the full benefit of the RIS, the scheduled users should be in close proximity in order to take the advantage of the more directional beams from the RIS. This type of trade-off calls for a solution that is capable of modelling the dependencies between the optimization variables.

To address the challenges in solving problem (10), we propose to utilize deep neural network as a powerful function approximator [34] to model the mapping function  $f(\cdot)$  and to learn the transmission strategy from data. However, designing a single neural network to learn such a mapping is not trivial, due to its complicated functional landscape. In the next section, we describe a multi-stage approach to learn such a mapping.

### III. MULTI-STAGE LEARNING FRAMEWORK

Weighted sum rate can be a challenging objective to learn. It has been shown in [26] that the addition of user weights as

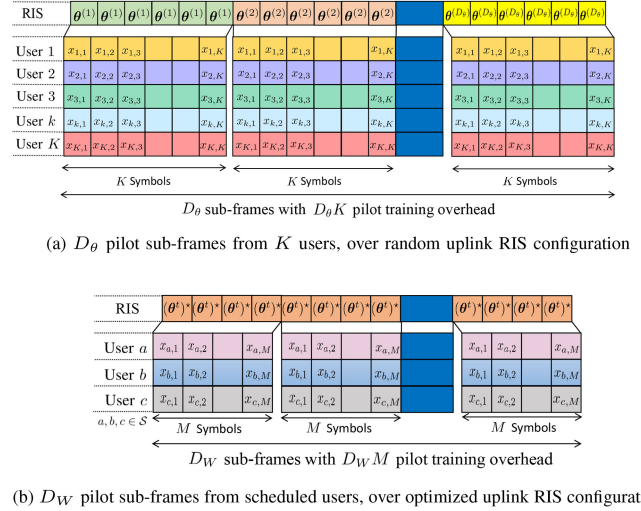


Fig. 2. Uplink pilot frame structure for: (a) scheduling and RIS configuration; and (b) beamforming.

input to the neural network imposes new learning challenge. This is due to the fact that a small change in user weights can drastically change the scheduling and the associated RIS design and beamforming vectors. Thus, to learn the function mapping, the neural network would need to sample a large amount of data, including the entire space of user weights, but the distributions of user weights are highly non-uniform and difficult to sample.

In this paper, we propose a multi-stage approach to solving the weighted sum-rate maximization problem (10). The idea is that user scheduling and RIS configuration can be designed in the first two stages in a data-driven fashion with one set of uplink pilots. Once the scheduled users and the RIS reflection coefficients are fixed, the optimal beamforming matrix is now a function of the effective channel, and can be designed in a third stage based on the estimated CSI, by either re-using the existing pilots or with possibly additional pilots.

Below we first describe the proposed uplink pilot placement strategy, then the GNN architecture for learning the user schedule in stage one and the RIS configuration in stage two, and finally the design of beamformers in stage three.

#### A. Pilot Placement for Scheduling and RIS Design

For scheduling and RIS design in the first two stages, we adopt the pilot transmission protocol proposed in [35] for pilots design and the uplink reflection coefficients at the RIS in the uplink pilot phase. The user pilots are set to be orthogonal, denoted as  $\mathbf{x}_k$ 's, where  $\mathbf{x}_k^H = [x_{k,1}, x_{k,2}, \dots, x_{k,K}]$ . Note that since there are a total of  $K$  potential users in the scheduling pool, the user pilot length is set to be  $K$  to ensure pilot orthogonality. The total training phase is equally partitioned to  $D_\theta$  sub-frames, where each sub-frame is of symbol duration  $K$ , as shown in Fig. 2(a). The reflection coefficients at the RIS are randomly chosen from sub-frame to sub-frame but remain fixed within each sub-frame, while the users repeatedly transmit the same mutually orthogonal pilot sequences over the  $D_\theta$  sub-frames. Let  $\boldsymbol{\theta}^{(d)}$  be the uplink RIS configuration in the  $d$ -th sub-frame.

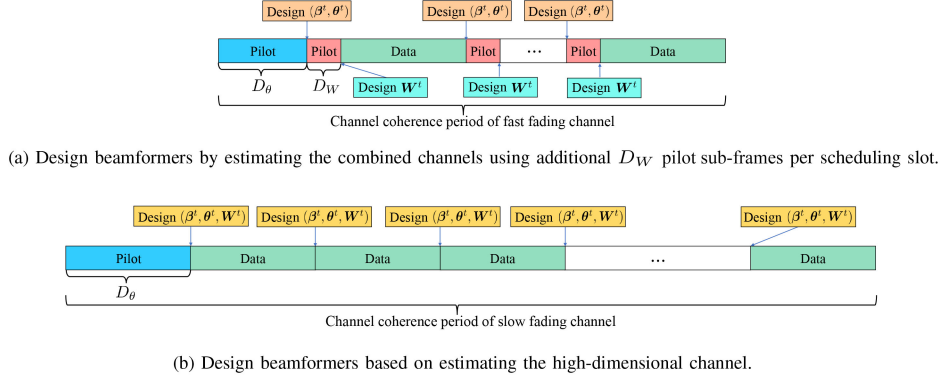


Fig. 3. Pilot placement structures for scheduling, RIS configuration and beamformer design over a coherence period.

The overall received pilots in the  $d$ -th sub-frame  $\mathbf{Y}^{(d)} \in \mathbb{C}^{M \times K}$  are given by

$$\mathbf{Y}^{(d)} = \sum_{k=1}^K \left( \mathbf{h}_{d,k} + \mathbf{A}_k \boldsymbol{\theta}^{(d)} \right) \mathbf{x}_k^H + \mathbf{N}^{(d)}, d = 1, \dots, D_\theta, \quad (11)$$

where  $\mathbf{N}^{(d)}$  is the noise matrix whose columns are independently and identically distributed following the distribution  $\mathcal{CN}(0, \sigma_u^2 \mathbf{I})$ . We can leverage the orthogonality of the received pilots sequence in each sub-frame and decorrelate  $\mathbf{Y}^{(d)}$  at the BS. Specifically, since  $\mathbf{x}_k^H \mathbf{x}_i = 0$  if  $i \neq k$  and  $\mathbf{x}_k^H \mathbf{x}_k = K P_u$ , with  $P_u$  as the uplink pilot transmission power, the contribution from the user  $k$  in the  $d$ -th sub-frame is given by:

$$\mathbf{y}_k^{(d)} = \frac{1}{K} \mathbf{Y}^{(d)} \mathbf{x}_k. \quad (12)$$

The collection of received pilot sequences from the user  $k$  in overall  $D_\theta$  sub-frames can be denoted as:

$$\mathbf{Y}_k^{D_\theta} = [\mathbf{y}_k^{(1)}, \mathbf{y}_k^{(2)}, \dots, \mathbf{y}_k^{(D_\theta)}]. \quad (13)$$

Note that the overall pilot training overhead over a total of  $D_\theta$  sub-frames is  $D_\theta K$ . This same set of received pilots  $\mathbf{Y}_k^{D_\theta}$  are used to optimize the user scheduling and the RIS configuration in the first two stages.

### B. Pilot Placement for Beamforming Design

For the beamforming design for the scheduled users in the third stage, two pilot strategies are possible. Observe that once the scheduling and the RIS configuration are fixed, the combined channel is now of lower dimension, and is given by

$$\mathbf{h}_{c,k} = \mathbf{h}_{d,k} + \mathbf{A}_k (\boldsymbol{\theta}^t)^*, \quad (14)$$

where  $(\boldsymbol{\theta}^t)^*$  denotes the optimized RIS reflection coefficients at the  $t$ -th scheduling timeslot. To obtain the combined channel  $\mathbf{h}_{c,k}$ , we have the option of either: (i) estimating the low-dimensional channel with additional pilots training; or (ii) estimating the high-dimensional channel  $\mathbf{h}_{d,k}$  and  $\mathbf{A}_k$  without additional pilots training, then computing  $\mathbf{h}_{c,k}$ . These two channel estimation strategies each have their advantages depending

on the total pilot overhead budget, and whether the channel is fast-fading or slow-fading.

1) *Estimating Low-Dimensional Combined Channel:* With fixed scheduling and RIS configuration, estimating the combined channel needs only relatively short pilots [36]. In particular, since at most  $M$  users are scheduled, we can assign orthogonal pilots of length  $M$  to the scheduled users, denoted as  $\tilde{\mathbf{x}}_k$ , over the optimized RIS reflection coefficient, in a separate uplink training phase, over  $D_W$  sub-frames, as shown in Fig. 2(b).

The received pilots at the BS in the  $d$ -th sub-frame are given by

$$\tilde{\mathbf{Y}}^{(d)} = \sum_{k \in \mathcal{S}} \mathbf{h}_{c,k} \tilde{\mathbf{x}}_k^H + \tilde{\mathbf{N}}^{(d)}, d = 1, \dots, D_W, \quad (15)$$

where  $\mathcal{S}$  denotes a set of scheduled user. Based on the orthogonality of the transmitted pilots, we decorrelate the received pilots to obtain the contribution from the  $k$ -th user:

$$\begin{aligned} \tilde{\mathbf{y}}_k^{(d)} &= \frac{1}{M} \tilde{\mathbf{Y}}^{(d)} \tilde{\mathbf{x}}_k \\ &\triangleq \mathbf{h}_{c,k} + \tilde{\mathbf{n}}_k^{(d)}, \end{aligned} \quad (16)$$

where  $\tilde{\mathbf{n}}_k^{(d)} \triangleq \tilde{\mathbf{N}}^{(d)} \tilde{\mathbf{x}}_k / M$ . The collection of received pilot sequences from the user  $k$  in overall  $D_W$  sub-frames can be denoted as:

$$\tilde{\mathbf{Y}}_k^{D_W} = [\tilde{\mathbf{y}}_k^{(1)}, \tilde{\mathbf{y}}_k^{(2)}, \dots, \tilde{\mathbf{y}}_k^{(D_W)}]. \quad (17)$$

We can then use linear minimum mean-squared error (LMMSE) estimation to estimate the combined channel of the  $k$ -th user:

$$\begin{aligned} \hat{\mathbf{h}}_{c,k} &= \mathbb{E}[\mathbf{h}_{c,k}] + (\tilde{\mathbf{Y}}_k^{D_W} - \mathbb{E}[\tilde{\mathbf{Y}}_k^{D_W}]) \\ &\quad \left( \mathbb{E}[(\tilde{\mathbf{Y}}_k^{D_W} - \mathbb{E}[\tilde{\mathbf{Y}}_k^{D_W}])^H (\tilde{\mathbf{Y}}_k^{D_W} - \mathbb{E}[\tilde{\mathbf{Y}}_k^{D_W}])] \right)^{-1} \\ &\quad \mathbb{E}[(\tilde{\mathbf{Y}}_k^{D_W} - \mathbb{E}[\tilde{\mathbf{Y}}_k^{D_W}])^H (\mathbf{h}_{c,k} - \mathbb{E}[\mathbf{h}_{c,k}])]. \end{aligned} \quad (18)$$

2) *Estimating High-Dimensional Channels:* The pilot overhead of directly estimating a low-dimensional channel is  $M D_W$ . This is to be performed at every scheduling timeslot, as shown in Fig. 3(a), because the effective low-dimensional channel is a function of the scheduled users and the optimized RIS configuration, which are different in each scheduling timeslot.

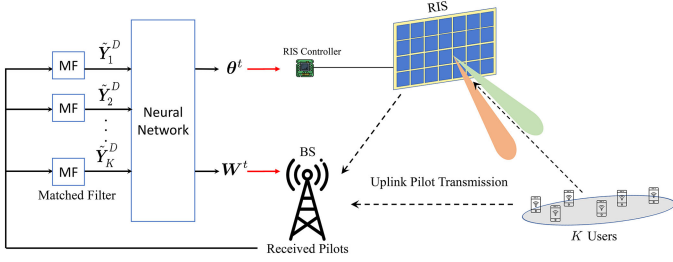


Fig. 4. GNN framework.

Although the above pilot overhead is small, when the channel is slow-fading so there are many scheduling timeslots within a channel coherence period, the total pilot overhead may still be considerable. Thus, when the total pilot budget is limited, it may be more advantageous to re-use the received pilots  $\{\mathbf{Y}^{(d)}\}_{d=1}^{D_\theta}$  to estimate the high-dimensional channel  $\mathbf{h}_{d,k}$  and  $\mathbf{A}_k$ . In this way, the combined channel  $\mathbf{h}_{c,k}$  can be obtained from  $(\mathbf{h}_{d,k}, \mathbf{A}_k, (\theta^t)^*)$  thus avoiding the extra pilot overhead, as shown in Fig. 3(b).

Recall that the collection of the received pilots over  $D_\theta$  from the  $k$ -th user for scheduling is  $\mathbf{Y}_k^{D_\theta}$ , whose column can be written as

$$\begin{aligned} \mathbf{y}_k^{(d)} &= \mathbf{h}_{d,k} + \mathbf{A}_k \theta^{(d)} + \mathbf{n}_k^{(d)} \\ &\triangleq \mathbf{H}_k \mathbf{q}^{(d)} + \mathbf{n}_k^{(d)}, \end{aligned} \quad (19)$$

where  $\mathbf{n}_k^{(d)} \triangleq \mathbf{N}^{(d)} \mathbf{x}_k / K$ . We denote the high-dimensional channel matrix as  $\mathbf{H}_k \triangleq [\mathbf{h}_{d,k}, \mathbf{A}_k] \in \mathbb{C}^{M \times (N+1)}$  and denote the combined phase shifts as  $\mathbf{q}^{(d)} \triangleq [1, \theta^{(d)\top}]^\top \in \mathbb{C}^{(N+1)}$ . We can estimate the channel matrix  $\mathbf{H}_k$  using LMMSE estimation as follows:

$$\begin{aligned} \hat{\mathbf{H}}_k &= \mathbb{E}[\mathbf{H}_k] + (\mathbf{Y}_k^{D_\theta} - \mathbb{E}[\mathbf{Y}_k^{D_\theta}]) \\ &\quad \left( \mathbb{E}[(\mathbf{Y}_k^{D_\theta} - \mathbb{E}[\mathbf{Y}_k^{D_\theta}])^H (\mathbf{Y}_k^{D_\theta} - \mathbb{E}[\mathbf{Y}_k^{D_\theta}])] \right)^{-1} \\ &\quad \mathbb{E}[(\mathbf{Y}_k^{D_\theta} - \mathbb{E}[\mathbf{Y}_k^{D_\theta}])^H (\mathbf{H}_k - \mathbb{E}[\mathbf{H}_k])]. \end{aligned} \quad (20)$$

We can adjust  $D_\theta$  to ensure the estimation accuracy of  $\mathbf{h}_{d,k}$  and  $\mathbf{A}_k$ . Importantly, no additional pilot training is needed for designing the beamformers, as shown in Fig. 3(b).

### C. GNN for User Scheduling and RIS Configuration

We now introduce the deep learning model for user scheduling and RIS configuration in the first two stages of the overall framework. A key building block of the proposed data-driven approach is a GNN architecture that takes the received pilots and the user weights as inputs and produces the optimized RIS configuration and BS beamformers as output, as shown in Fig. 4.

The use of GNN to model the inter-relationship between the RIS and the users is crucial, because the proposed GNN architecture allows certain permutation invariant and equivariant properties to be observed. That is, if the ordering of the users is permuted, the neural network should permute the set of beamforming vectors accordingly, while keeping the same scheduling

set and the same reflection coefficients. This property is difficult to learn by a conventional fully connected neural network, but is embedded in the architecture of a GNN [29]. The earlier work [11] shows that such a GNN can be trained to generate interpretable RIS configurations and BS beamformers, but only for the setting in which the user schedule is fixed a priori. In this paper, we treat the more challenging setting in which the user schedule also needs to be optimized.

Toward this end, we adopt a GNN architecture, as shown in Fig. 5, which consists of  $K+1$  fully connected nodes. Node 0 corresponds to the RIS, and node 1 to  $K$  corresponds to beamforming vectors of user 1 to user  $K$ . Each node is associated with a representation vector denoted as  $\mathbf{v}_k, k \in \{0, \dots, K\}$ . The idea is to encode useful information about each node in the representation vectors. The initial representation vectors  $\mathbf{v}_k^0$  are functions of the input features  $\pi_k$ 's, which include the weight and the received pilots of the user  $k$  over  $D$  sub-frames

$$\mathbf{v}_0^0 = g_\theta(\phi_{\text{mean}}(\pi_1, \dots, \pi_K)), \quad (21a)$$

$$\mathbf{v}_k^0 = g_w(\pi_k), k = 1, \dots, K, \quad (21b)$$

where  $\pi_k = [\alpha_k; \text{vec}(\mathcal{R}(\mathbf{Y}_k^{D_\theta})); \text{vec}(\mathcal{I}(\mathbf{Y}_k^{D_\theta}))]$  and  $\phi_{\text{mean}}$  is the element-wise mean function. Here,  $g_w(\cdot)$  and  $g_\theta(\cdot)$  are fully connected neural networks.

Then the nodal representation vectors are updated layer by layer, taking the representation vectors in the previous layer as input. Thus, the update of the RIS node is a function of itself and all user nodes; similarly, the update of a user node is a function of itself, the RIS node and all other user nodes. This updating rule allows the GNN to learn the interference amongst users. Specifically, the update rule in the  $(z+1)$ -th layer is given as [11]

$$\mathbf{v}_0^{z+1} = f_1(f_2(\mathbf{v}_0^z), \phi_{\text{mean}}(f_3(\mathbf{v}_1^z), \dots, f_3(\mathbf{v}_K^z))), \quad (22a)$$

$$\mathbf{v}_k^{z+1} = f_4(\mathbf{v}_k^z, f_2(\mathbf{v}_0^z), \phi_{\text{max}}(\{f_5(\mathbf{v}_j^z)\}_{\forall j \neq 0, j \neq k})), \quad (22b)$$

where  $f_1(\cdot), f_2(\cdot), f_3(\cdot), f_4(\cdot), f_5(\cdot)$  are fully connected neural networks. Here,  $\phi_{\text{max}}$  is the element-wise max functions.

After  $Z$  iterations, the final representation vectors  $\mathbf{v}_k^Z$ 's would contain the right representation of information to design the RIS configuration and the beamformers. The  $\mathbf{v}_k^Z$ 's pass through linear layer with  $2N$  or  $2M$  fully connected units

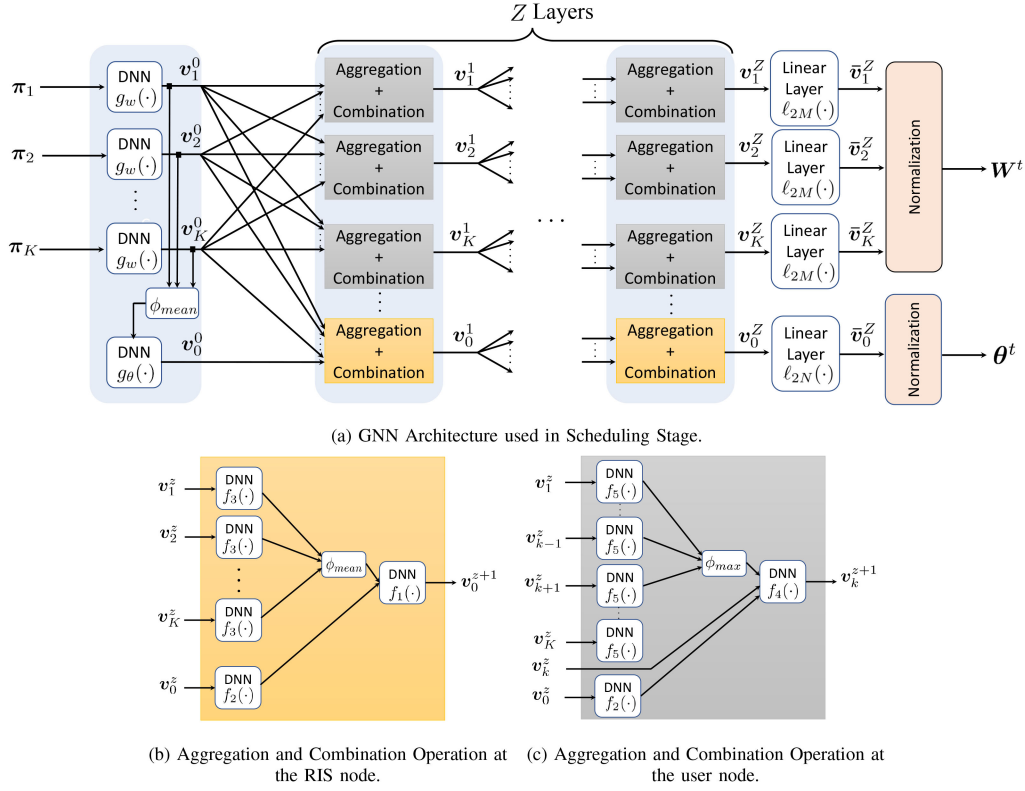
$$\bar{\mathbf{v}}_0^Z = \ell_{2N}(\mathbf{v}_0^Z) \in \mathbb{R}^{2N}, \quad (23a)$$

$$\bar{\mathbf{v}}_k^Z = \ell_{2M}(\mathbf{v}_k^Z) \in \mathbb{R}^{2M}, k = 1, \dots, K. \quad (23b)$$

Subsequently,  $\bar{\mathbf{v}}_k^Z$ 's are normalized so that the RIS reflection coefficients can be deduced from  $\bar{\mathbf{v}}_0^Z$  and the beamforming vector associated with the user  $k$  can be deduced from  $\bar{\mathbf{v}}_k^Z, k \in \{1, \dots, K\}$  as in [11].

The specific GNN architecture adopted here is similar to the one in [11] in which the constituent components all obey permutation invariant and equivariant properties, but with a key difference that the proportional fairness weights for all the users are also used as input to the GNN. Incorporating priority weights in the neural network for scheduling is in general highly non-trivial, because the optimized system parameters can be very sensitive to small perturbations in the weights [26]. To




 Fig. 5. GNN Architecture for  $K$  users.

tackle this issue, this paper proposes to use two GNNs, one for scheduling and one for RIS configuration in order to optimize the overall objective.

More specifically, training a single GNN both to perform scheduling and to find the optimal RIS configuration is challenging, because such a GNN would need to have  $K+1$  nodes, where  $K$  can be large in a dense network, and each node would be associated with high-dimensional input features over the  $D_\theta$  pilot sub-frames. Instead, we propose to use a GNN with  $(K+1)$  nodes over  $D_\beta$  pilot sub-frames in the first stage just to produce the scheduling of up to  $M$  users, where  $D_\beta$  can be as small as 1 to make the training manageable. Then, another GNN with  $(M+1)$  nodes over  $D_\theta$  pilot sub-frames is adopted in the second stage to produce the optimized RIS configuration, where  $D_\theta$  can be considerably larger than  $D_\beta$ , but the pilots are re-used so that the overall pilot length is  $D_\theta$ . We describe the two GNNs in more detail below.

**1) Scheduling:** In the first stage, a GNN with  $K+1$  nodes is used to learn the optimized schedule from a set of very short pilots of  $D_\beta$  sub-frames. As shown in Fig. 6, the inputs to the neural network are the user weights  $(\alpha_1^t, \alpha_2^t, \dots, \alpha_K^t)$  and the decorrelated received pilot sub-frames  $(Y_1^{D_\beta}, Y_2^{D_\beta}, \dots, Y_K^{D_\beta})$ . The output is the RIS configuration and beamformers designed for all  $K$  users. The idea is that unlike BS beamforming or RIS configuration, which are strong functions of the channel, scheduling can be done with only a coarse estimation of the overall channel strength or directions. In practical implementation, the  $D_\beta$  sub-frames can be part of the overall  $D_\theta$  sub-frames of the received pilots, since the channels are constant within the coherence time. In the training phase, the loss function is set to

be the negative of the weighted sum rate of  $K$  users

$$-\mathbb{E} \left[ \sum_{k=1}^K \alpha_k^t R_k^t \right]. \quad (24)$$

Since at this stage, we only use very short received pilots to train the GNN, the reflection coefficients and the beamforming vectors learned by the GNN are quite suboptimal, but it is good enough to produce the scheduling decision. To do so, we adopt the implicit scheduling scheme and infer the schedule from the powers of the beamforming vectors  $w_k^t$ 's learned by the GNN. Specifically, a user with larger beamforming power is more likely to be scheduled than those with smaller beamforming power. This is reasonable, because the GNN accounts for multiuser interference and designs the beamforming vectors together with the RIS reflection coefficients to maximize the objective [11]. Thus, the users to whom the GNN allocates more power are the users who make the most significant contributions to the weighted sum rate objective, and they should be included in the schedule. Therefore, if we select  $M$  users with the highest powers from the set  $\{\|w_1^t\|^2, \|w_2^t\|^2, \dots, \|w_K^t\|^2\}$ , this should be a good set of users to schedule.

Note that the output  $\theta^t$  is discarded as in this stage our goal is to learn a scheduling decision using only  $D_\beta$  received pilot sub-frames. Regardless, we still keep  $\theta^t$  in the training process to account for the contribution of RIS in the maximization of the objective. The reflection coefficients are designed in the next stage using another GNN that can map high-dimensional received pilots to a better  $\theta^t$ .



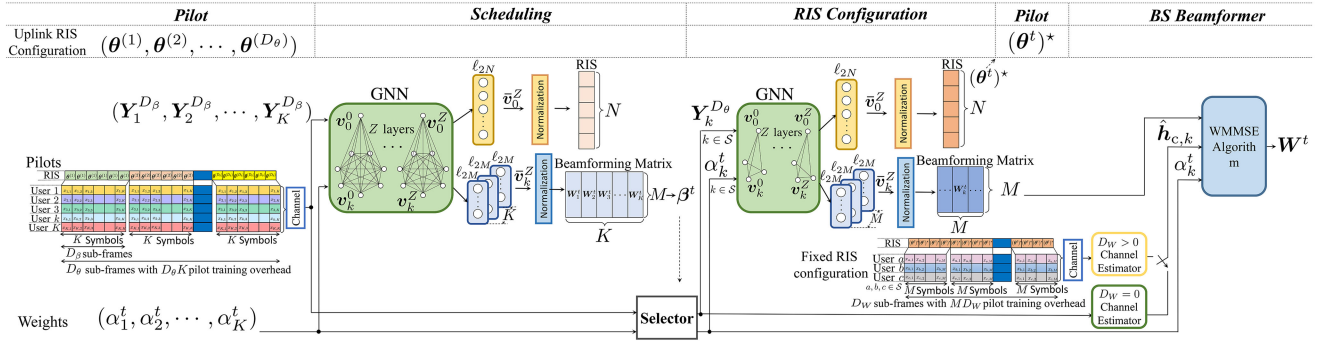


Fig. 6. The overall three-stage framework for scheduling, RIS configuration, and beamforming in multiuser downlink.

2) *RIS Configuration*: In the second stage, we design better reflection coefficients using a second GNN over the  $M$  scheduled users but with longer  $D_\theta$  pilot sub-frames. Such a GNN has a reduced number of nodes, i.e.,  $M + 1$ , so it can take much longer received pilot sequences as input without experiencing training difficulty. In particular, the inputs to the GNN for RIS design are the user weights and  $D_\theta$  decorrelated received pilot sub-frames of the scheduled users, i.e.,  $\alpha_k^t$  and  $\mathbf{Y}_k^{D_\theta}$  with  $k \in \mathcal{S}$ , as shown in Fig. 6. The output is the RIS configuration and beamformers designed for the scheduled users. The GNN is trained to maximize the weighted sum rate of the  $M$  scheduled users to produce the optimized RIS configuration  $(\theta^t)^*$ . The loss function is expressed as

$$-\mathbb{E} \left[ \sum_{k \in \mathcal{S}} \alpha_k^t R_k^t \right]. \quad (25)$$

The designed RIS reflection coefficients  $(\theta^t)^*$  are employed as the RIS configuration for the scheduled users, but the beamformers are discarded and re-optimized in the next stage.

The overall three-stage framework is shown in Fig. 6. Note that in the first two stages, the GNN for scheduling has larger dimension than the GNN for RIS configuration. But the GNN for RIS configuration takes considerably longer pilots than the GNN for scheduling.

#### D. Beamforming Stage

Although the BS beamforming vectors are already produced in the previous stages, there is still large room to fine-tune the beamforming vectors. For example in the high SNR regime, an optimal beamforming design should closely resemble a zero-forcing (ZF) solution that can eliminate all the interference. But a neural network may experience difficulties in learning a ZF beamformer in the high SNR regime as compared to learning a matched-filtering solution as shown in [37].

For this reason, we propose to fine-tune the beamforming matrix using the WMMSE algorithm based on the estimated channel  $\hat{\mathbf{h}}_{c,k}$  in a third stage. Specifically, we use the following iterative updating rule to obtain an stationary solution [30]:

$$\xi_k = \left( \sum_{j \in \mathcal{S}} |\hat{\mathbf{h}}_{c,k}^H \mathbf{w}_j^t|^2 + \sigma_d^2 \right)^{-1} \hat{\mathbf{h}}_{c,k}^H \mathbf{w}_k^t, \quad (26a)$$

$$\nu_k = \left( 1 - \xi_k^* \hat{\mathbf{h}}_{c,k}^H \mathbf{w}_k^t \right)^{-1}, \quad (26b)$$

$$\mathbf{w}_k^t = \alpha_k^t \xi_k \nu_k \left( \lambda \mathbf{I}_M + \sum_{j \in \mathcal{S}} \alpha_j^t |\xi_j|^2 \nu_j \hat{\mathbf{h}}_{c,j} \hat{\mathbf{h}}_{c,j}^H \right)^{-1} \hat{\mathbf{h}}_{c,k}, \quad (26c)$$

where  $\lambda$  denotes the dual variable for the downlink transmission power constraint.

To estimate the channel of the  $M$  scheduled users, we can use one of the two channel estimation schemes discussed in Section III-B. We could either employ additional pilots to estimate the effective low-dimensional channel, or estimate the high-dimensional channel without additional pilot overhead.

#### E. Pilot Overhead

The overall pilot overhead of the three-stage framework is as follows. The scheduling and RIS configuration stages take  $D_\theta$  pilot sub-frames, each of length  $K$ . Then, we have two cases depending on the two strategies in the beamforming stage.

1) *With Additional Pilots* ( $D_W > 0$ ): In this case, the scheduled users transmit some additional orthogonal pilots over the channel with fixed uplink RIS configuration  $(\theta^t)^*$ . Depending on the SNR, we can use  $M \times D_W$  pilot sequences. This CSI acquisition strategy allows us to obtain an accurate estimation of the low-dimensional channel with a small pilot overhead for each scheduling timeslot. In this case, the overall pilot overhead over all three stages is given by

$$L = K \times D_\theta + M \times D_W \times \Upsilon, \quad (27)$$

where  $K \times D_\theta$  accounts for the pilots used to design the schedule and the RIS configuration in the first two stages, and  $\Upsilon$  is the number of scheduling slots.

2) *Without Additional Pilots* ( $D_W = 0$ ): Alternatively, we can re-use the received  $D_\theta$  pilot sub-frames from all  $K$  users to estimate the direct channel  $\mathbf{h}_{d,k}$  and cascade channel  $\mathbf{A}_k$  for the scheduled users. In this case, the total pilot overhead is

$$L = K \times D_\theta. \quad (28)$$

In the next section, we provide a comparison of the two cases in terms of performance versus the overall pilot overhead in various scenarios.

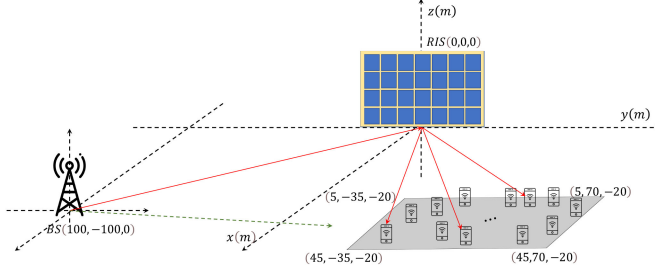


Fig. 7. Simulation setting for RIS-assisted multiuser network.

#### IV. NUMERICAL RESULTS

##### A. Simulation Environment

We consider an RIS-assisted multiuser MISO network with  $M = 8$  BS antennas,  $N = 128$  RIS reflective elements, and  $K = 32$  users. In the  $(x, y, z)$  coordinates, the BS and the RIS are located at  $(100\text{ m}, -100\text{ m}, 0\text{ m})$  and  $(0\text{ m}, 0\text{ m}, 0\text{ m})$  respectively. The user locations are uniformly generated within a rectangular area on the  $x$ - $y$  plane  $(25 \pm 20\text{ m}, 17.5 \pm 52.5\text{ m}, -20\text{ m})$ , as shown in Fig. 7. We assume that the direct link channel follows Rayleigh fading

$$\mathbf{h}_{d,k} = \rho_{0,k} \tilde{\mathbf{h}}_{d,k}, \quad (29)$$

where  $\tilde{\mathbf{h}}_{d,k} \sim \mathcal{CN}(0, \mathbf{I})$  and  $\rho_{0,k}$  denotes the pathloss between BS and user  $k$ . The reflection channels  $\mathbf{h}_{r,k}$ ,  $\mathbf{G}$  are assumed to follow Rician fading model:

$$\mathbf{h}_{r,k} = \rho_{1,k} \left( \sqrt{\frac{\epsilon}{1+\epsilon}} \tilde{\mathbf{h}}_{r,k}^{\text{LOS}} + \sqrt{\frac{1}{1+\epsilon}} \tilde{\mathbf{h}}_{r,k}^{\text{NLOS}} \right), \quad (30a)$$

$$\mathbf{G} = \rho_2 \left( \sqrt{\frac{\epsilon}{1+\epsilon}} \tilde{\mathbf{G}}^{\text{LOS}} + \sqrt{\frac{1}{1+\epsilon}} \tilde{\mathbf{G}}^{\text{NLOS}} \right), \quad (30b)$$

where  $\rho_{1,k}$  and  $\rho_2$  denote the path losses between the RIS and the  $k$ -th user/BS. The path-loss models of the direct and reflected paths are  $32.6 + 36.7 \log(d_1)$  and  $30 + 22 \log(d_2)$ , respectively, where  $d_1$  and  $d_2$  denote the corresponding link distance. Here,  $\tilde{\mathbf{h}}_{r,k}^{\text{NLOS}}$  and  $\tilde{\mathbf{G}}^{\text{NLOS}}$  denote the non-line-of-sight components and their entries are generated independently according to  $\mathcal{CN}(0, 1)$ . We assume that there are  $\Upsilon = 50$  scheduling time-slots in slow-fading channel and  $\Upsilon = 5$  scheduling time-slots in fast-fading channel. The transmission powers for both uplink and downlink are 15 dBm. The bandwidth is 10 MHz with a background noise of  $-170$  dBm/Hz. The Rician factor  $\epsilon$  is set to 10. The forgetting factor  $\gamma$  in exponentially moving averaging of user rates is set to 0.01.

##### B. Baseline Scheduling Strategy

As baseline comparisons, the following greedy scheduling (GS) algorithm and exhaustive search (ES) scheduling algorithm for solving problem (10) are implemented.

The greedy scheduling algorithm consists of two phases: an uplink channel estimation phase and a downlink proportional fairness scheduling phase. Uplink channel estimation is performed using the method in [35], and the downlink weighted

---

##### Algorithm 1: Baseline GS with BCD.

---

- 1: Initialize the set of scheduled users  $\mathcal{S} = \{\emptyset\}$ ;
- 2: Initialize random RIS phase shift vector  $\boldsymbol{\theta}^t$ ;
- 3: Set  $\mathbf{w}_k^t = (\mathbf{h}_{d,k} + \mathbf{A}_k \boldsymbol{\theta}^t)$ ,  $k = 1, \dots, K$ ;
- 4: Compute and sort weighted single-user rate  $\alpha_k^t \hat{R}_k^t$  where

$$\hat{R}_k^t = \log \left( 1 + \frac{|(\mathbf{h}_{d,k} + \mathbf{A}_k \boldsymbol{\theta}^t)^H \mathbf{w}_k^t|^2}{\sigma_d^2} \right);$$

- 5: Select the user with the largest  $\alpha_k^t \hat{R}_k^t$  to add to  $\mathcal{S}$ ;
  - 6: For each unscheduled user, test whether adding that user to  $\mathcal{S}$  improves the objective (10). Select the user with the largest improvement to add to  $\mathcal{S}$ . Repeat until adding another user no longer improves the objective.
  - 7: Fix  $\mathcal{S}$  and  $\mathbf{W}^t$ , update  $\boldsymbol{\theta}^t$  using the Riemannian conjugate gradient (RCG) algorithm [38];
  - 8: Fix  $\mathcal{S}$  and  $\boldsymbol{\theta}^t$ , update  $\mathbf{W}^t$  using WMMSE [30];
  - 9: Fix  $\boldsymbol{\theta}^t$  and  $\mathbf{W}^t$ , update  $\mathcal{S}$  by adding the unscheduled user which improves the objective (10) the most, if any;
  - 10: Repeat step 7–9 until convergence.
- 

sum-rate maximization problem is solved with a greedy scheduler and a block coordinate descent (BCD) approach for optimizing the beamformers and the RIS configuration as in [20].

1) *Uplink Channel Estimation*: Channel estimation is performed at the beginning of channel coherence period, in which the BS collects all of the pilot symbols from  $K$  users. Thus, the total pilot overhead for a coherence interval is calculated as

$$L = K \times D_H, \quad (31)$$

where  $D_H$  denotes the number of pilot sub-frames used to estimate the high-dimensional channels  $\mathbf{A}_k$ 's and  $\mathbf{h}_{d,k}$ 's via the LMMSE estimator (20).

2) *Downlink Proportional Fairness Scheduling*: The joint scheduling, RIS configuration, and beamforming problem can be solved iteratively as described in Algorithm 1. The user scheduling is based on greedily choosing the user that maximizes the overall optimization objective of weighted sum rate.

In the simulations, we evaluate the performance of the proposed data-driven approach against the baseline approach for the cases both with perfect CSI and with estimated CSI.

As another baseline comparison, the ES scheduling algorithm tries every possible combination of users to identify the optimal scheduling. For each combination of users, the RIS configuration is designed using the GNN in Section III-C2, and beamforming matrix is designed using WMMSE algorithm as in Section III-D. The set of users that achieves the highest objective in (10) is the scheduled user set. The ES scheduling algorithm is not scalable as the number of users becomes large. It serves as a performance upper bound for the proposed GNN-based scheduler.

##### C. Simulation Result

The GNN based three-stage joint scheduling, RIS configuration, and beamforming framework as shown in Fig. 6 is

TABLE I  
HYPERPARAMETERS OF DEEP NEURAL NETWORKS

Label	Activation Function	Size
$\ell_{2N}(\cdot)$	Linear	$2N$
$\ell_{2M}(\cdot)$	Linear	$2M$
$g_w(\cdot)$	Relu	$2MD_\theta \times 1024 \times 512$
$g_\theta(\cdot)$	Relu	$512 \times 1024 \times 512$
$f_1(\cdot), f_2(\cdot), f_3(\cdot), f_4(\cdot), f_5(\cdot)$	Relu	$512 \times 512 \times 512$

implemented using parameters in Table I. For both the GNN for scheduling and the GNN for RIS configuration, we implement two-layer ( $Z = 2$ ) models using Tensorflow [39]. In the scheduling stage, we set  $D_\beta = 1$ . In the training phase, the GNNs sample 102,000 training data including the channel vectors, pilot sequences, and uniformly generated weights in each epoch. The Adam optimizer [40] is used for the neural network to adjust its neuron weights in an unsupervised fashion to minimize the loss function. We examine its performance in both slow-fading and fast-fading channel environments below.

1) *Slow-Fading Channel*: We first examine the performance of GNN based algorithms vs. the baseline in terms of the *cumulative distribution function* (CDF) of user average rate in Fig. 8(a). Each point on the CDF curve represents an individual user's average instantaneous rate averaged across  $10^2$  i.i.d channel realizations, which is equivalent to  $10^2\Upsilon$  scheduling timeslots. From Fig. 8(a), the proposed data-driven algorithm with 592 pilot symbols (including  $D_W = 1$  pilot per scheduling slot in the third beamforming stage) demonstrates performance close to the baseline approach with perfect CSI, and significantly outperforms the baseline approach with estimated CSI using 640 or 1280 pilot symbols. This implies that the proposed algorithm, which is based on directly maximizing system objective based on the received pilots and the user priority weights using a neural network without explicit channel reconstruction, can significantly reduce the amount of the pilot overhead as compared to the conventional approach.

It is also interesting to see that the proposed algorithm demonstrates a strong CDF performance in the low-rate regime (below 40-percentile range). This is because the scheduled users tend to have either strong channels or high weights (if not both). A strong CDF performance in the low-rate range implies that the GNN scheduler is sensitive to user weights and often schedules users with high weights and weak channels.

Further, from Fig. 8(c) and Fig. 8(d), we observe that the proposed algorithm but without additional pilot in the beamforming stage also outperforms the baseline approach consistently, and in the short pilot region, can be superior to the approach of using additional pilot for beamforming in the third stage. The performance gain against the baseline is again due to the more direct and efficient use of the received pilot sequences by the GNN models, without explicit channel estimation as the intermediary.

Finally, we remark that a slight performance improvement can be obtained by the ES scheduling algorithm. However, exhaustive search is computationally complex and not scalable as the number of users becomes large.

We note that there is a performance gap between both proposed algorithms and the perfect CSI case. For the case without the additional pilot in the beamforming stage, the gap is due to the inaccurate estimation of the high-dimensional channel  $\mathbf{A}_k$ ,  $\mathbf{h}_{d,k}$ ,  $k = 1, \dots, K$ . The performance gap is reduced when additional pilots are used to estimate the low-dimensional effective channel  $\mathbf{h}_{c,k}$ ,  $k \in \mathcal{S}$ . But the gap is still not zero. This is because of the inherent limitation in the neural network architecture and training.

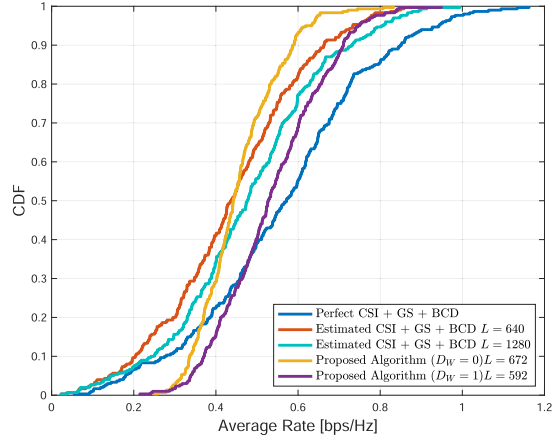
2) *Fast-Fading Channel*: In a fast-fading environment, the channel coherence period is shorter, thus there are fewer number of scheduling timeslots available. From Fig. 8(d) and Fig. 8(f), we observe that using additional pilots in the beamforming stage is overall the best approach. Note that both proposed approaches always outperform the baseline.

We also note that the log utility and the sum rate are higher in a fast-fading setting as compared to a slow-fading setting. This is due to multiuser diversity. In a slow-fading scenario, the channels remain fixed for a long period of time. At first, the scheduler tends not to schedule users with weak channels. But after a large number of scheduling timeslots, the scheduler has no choice but to schedule those users, which can result in a lower sum rate. This phenomenon is not as prevalent in a fast-fading channel, where the user channels change more rapidly and the scheduler is less likely to be forced to schedule a user with a poor channel condition.

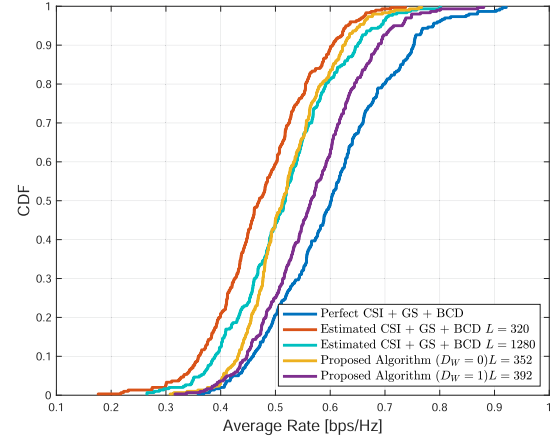
#### D. Complexity Analysis

The complexity of the proposed learning based approach consists of two neural network inference stages and an iterative beamformer optimization stage. The complexity of the GNN in the scheduling stage is  $O(MK\tau + (K+1)^2\epsilon Z + (MK + N)\delta)$ , where  $\tau$ ,  $\delta$ , and  $\epsilon$  denote the dimension of the input layer, dimension of the output layer, and the computational complexity of the fully connected aggregation and combination operations respectively. Recall that  $Z$  denotes the number of layers of the GNN. Similarly, the complexity of GNN for the RIS design is  $O(M^2\tau + (M+1)^2\epsilon Z + (M^2 + N)\delta)$ . The complexity of the beamformer optimization stage is  $O(\ell_\lambda \ell_w (2NM^2 + M^3))$  [20], where  $\ell_\lambda$  and  $\ell_w$  denote the iteration number of inner loops. The total complexity is therefore  $O((K^2 + M^2)Z + MK + M^2 + N + \ell_\lambda \ell_w (NM^2 + M^3))$  after discarding lower order terms. It is important to note that both the training and inference processes are highly parallelizable using modern graphic process units (GPUs), so that in practice the proposed approach can be executed very efficiently.

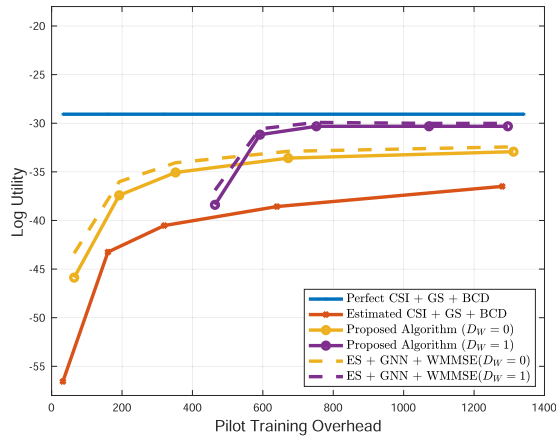
In contrast, the baseline GS strategy in Algorithm 1 adopts an iterative structure that solves scheduling, RIS reflection coefficients and beamformer sub-problems sequentially until convergence. In particular, the complexities associated with calculating the schedule, the RIS reflection coefficients, and the beamforming vectors are respectively  $O(MNK \log(K))$ ,  $O(\ell_r M^2 N^2)$ , and  $O(\ell_\lambda \ell_w (2NM^2 + M^3))$  [20], where  $\ell_r$  denotes the iteration number of inner loops. The total complexity is  $O(\ell_O(MNK \log(K) + \ell_r M^2 N^2 + \ell_\lambda \ell_w (NM^2 + M^3)))$ , where  $\ell_O$  denotes the iteration number of the outer loop. In



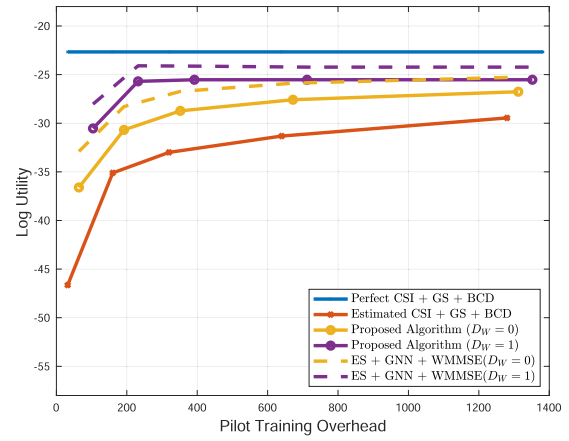
(a) Cumulative distribution function (CDF) of user rates in slow-fading channel.



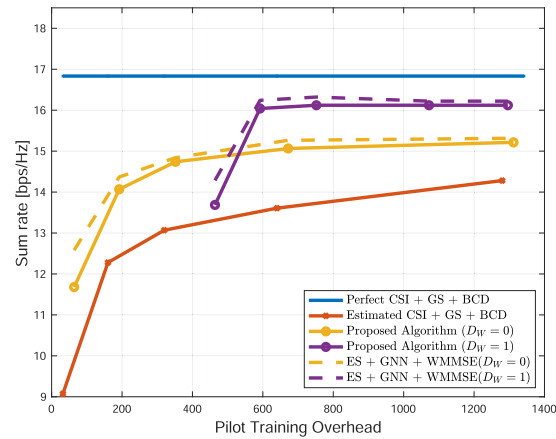
(b) Cumulative distribution function (CDF) of user rates in fast-fading channel.



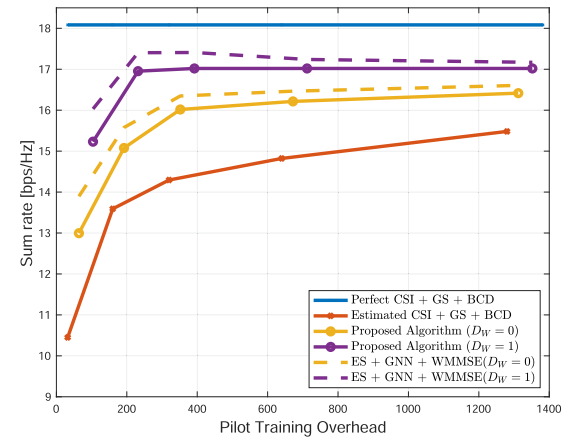
(c) Network-wide log utility of slow-fading channel.



(d) Network-wide log utility of fast-fading channel.



(e) Sum rate of slow-fading channel.



(f) Sum rate of fast-fading channel.

 Fig. 8. Performance of GNN based scheduling in an RIS-assisted downlink system with  $N = 128$ ,  $M = 8$ ,  $K = 32$ ,  $P_d = 15$  dBm.

practice, both the GS and ES algorithms are orders of magnitude slower than the proposed neural network based approach.

### E. Generalization to Different Sizes of User Pool

In an RIS-assisted network, parameters such as the number of antennas at the BS or size of the RIS may be static but the number of users is constantly changing depending on the coverage of the

BS and the user traffic. Thus, it is important for the proposed algorithm to demonstrate generalizability to scenarios with a different number of users.

Generalizability in the first two stages is achieved by adopting the same aggregation and combination operation for all user nodes throughout the GNN architectures, i.e.,  $g_w(\cdot)$ ,  $f_1(\cdot)$ ,  $f_2(\cdot)$ ,  $f_3(\cdot)$ ,  $f_4(\cdot)$ ,  $f_5(\cdot)$  and  $\ell_{2M}(\cdot)$ , such that the updating rule of the user nodes is independent of the number of users. Consequently,



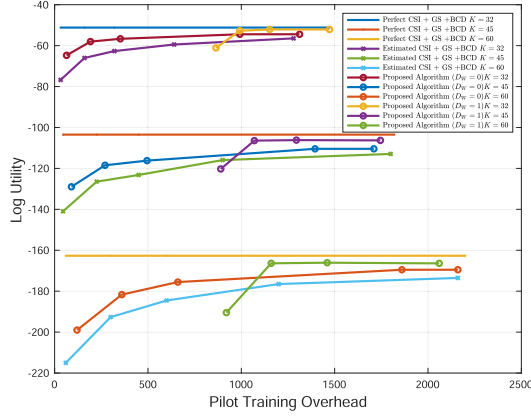


Fig. 9. Network-wide log utility of slow-fading channel with different number of users,  $N = 128$ ,  $M = 16$ ,  $P_d = 5$  dBm.

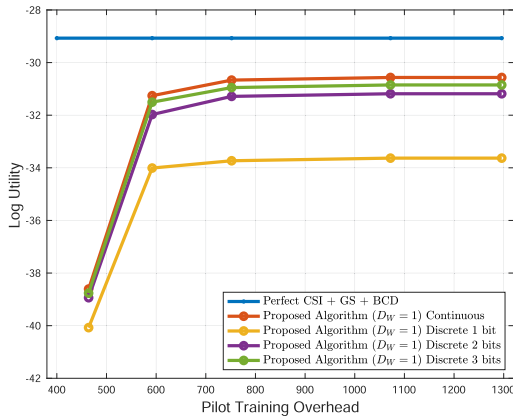


Fig. 10. Network-wide log utility of slow-fading channel with different number of control bits,  $N = 128$ ,  $M = 8$ ,  $K = 32$ ,  $P_d = 15$  dBm.

when the total population of the users changes from  $K$  to  $K'$ , the learned aggregation and combination operation can still be applied to the new graph with  $K'$  user nodes. In the final BS beamforming stage, generalizability is ensured as the BS beamforming optimization is not a function of the total number of users, but a function of the number of BS antennas. In this way, the overall three-stage framework can generalize to systems with a different number of users.

We test the generalizability of the proposed three-stage framework as follows. We train the GNN on a user pool of size 32, and test the scheduler's performance in networks where the number of users is 45 and 60. As seen in Fig. 9, the proposed algorithms can generalize well. It can converge close to the perfect CSI baseline, and can significantly reduce the pilot overhead as compared to the conventional channel estimation based approach.

#### F. Discrete vs. Continuous Phase Shifter

Finally, we test the proposed algorithm on RIS with discrete phase shifts. In practical implementation, manufacturing RIS elements with infinite precision phase shifts may be costly or even infeasible due to hardware limitations [41]. Thus, the optimization of the RIS system with discrete phase shifters is of interest. To this end, we quantize the continuous reflecting

coefficients by a finite number of levels. For an RIS with  $b$  control bits, we uniformly divide the range  $[0, 2\pi)$  to  $2^b$  partitions. The mid-value of each partition is the representation value of each partition. For example, 2 bits correspond to  $2^2$  representation values:  $\pi/4$ ,  $3\pi/4$ ,  $5\pi/4$  and  $7\pi/4$ .

Fig. 10 shows the performance of the proposed algorithm in log utility with discrete phases at the RIS. It can be observed that 2-bit quantization can already achieve 98% performance of the continuous phase shift case. This implies that with only 2 control bits, per element an RIS can already deliver a network utility performance similar to an RIS with continuous phase shifts.

#### V. CONCLUSION

This paper considers a proportional fairness scheduling problem in a multiuser RIS-assisted MISO network. Specifically, we show that a GNN can simultaneously tackle the challenges in CSI acquisition, and in designing the proportionally fair scheduling, the optimal RIS phase shifts, and the beamformers, under limited pilot overhead. Numerical results show significant gain over the conventional approach of first explicitly estimating the channel, then performing network optimization. The proposed approach can be generalized to scenarios with an arbitrary number of users. It shows considerable promise of using a machine learning approach for *discrete optimization*, but also points to the importance of judiciously designing the overall framework, the pilot placement structure, and the appropriate neural network architecture.

#### REFERENCES

- [1] Z. Zhang, T. Jiang, and W. Yu, "User scheduling using graph neural networks for reconfigurable intelligent surface assisted multiuser downlink communications," in *Proc. IEEE Int. Conf. Acoust. Speech, Signal Process. (ICASSP)*, 2022, pp. 8892–8896.
- [2] E. Basar, M. Di Renzo, J. de Rosny, M. Debbah, M.-S. Alouini, and R. Zhang, "Wireless communications through reconfigurable intelligent surfaces," *IEEE Access*, vol. 7, pp. 116753–116773, 2019.
- [3] M. Di Renzo *et al.*, "Smart radio environments empowered by reconfigurable intelligent surfaces: How it works, state of research, and the road ahead," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2450–2525, Nov. 2020.
- [4] F. Liu *et al.*, "Intelligent metasurfaces with continuously tunable local surface impedance for multiple reconfigurable functions," *Phys. Rev. Appl. Gen. Phys.*, vol. 11, no. 4, pp. 2331–7019, Apr. 2019.
- [5] C. Liaskos, S. Nie, A. Tsioliaridou, A. Pitsillides, S. Ioannidis, and I. Akyildiz, "A new wireless communication paradigm through software-controlled metasurfaces," *IEEE Commun. Mag.*, vol. 56, no. 9, pp. 162–169, Sep. 2018.
- [6] L. Li *et al.*, "Electromagnetic reprogrammable coding-metasurface holograms," *Nature Commun.*, vol. 8, no. 1, Aug. 2017, Art. no. 197.
- [7] J. Lyu and R. Zhang, "Hybrid active/passive wireless network aided by intelligent reflecting surface: System modeling and performance analysis," *IEEE Trans. Wireless Commun.*, vol. 20, no. 11, pp. 7196–7212, Nov. 2021.
- [8] Q.-U.-A. Nadeem, H. Alwazani, A. Kammoun, A. Chaaban, M. Debbah, and M. S. Alouini, "Intelligent reflecting surface-assisted multi-user MISO communication: Channel estimation and beamforming design," *IEEE Open J. Commun. Soc.*, vol. 1, pp. 661–680, May 2020.
- [9] Q.-U.-A. Nadeem, A. Kammoun, A. Chaaban, M. Debbah, and M.-S. Alouini, "Asymptotic max-min SINR analysis of reconfigurable intelligent surface assisted MISO systems," *IEEE Trans. Wireless Commun.*, vol. 19, no. 12, pp. 7748–7764, Dec. 2020.
- [10] G. Yang, X. Xu, and Y.-C. Liang, "Intelligent reflecting surface assisted non-orthogonal multiple access," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, 2020, pp. 1–6.

- [11] T. Jiang, H. V. Cheng, and W. Yu, "Learning to reflect and to beamform for intelligent reflecting surface with implicit channel estimation," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 7, pp. 1931–1945, Jul. 2021.
- [12] Ö. Özdoğan and E. Björnson, "Deep learning-based phase reconfiguration for intelligent reflecting surfaces," in *Proc. Asilomar Conf. Signals, Syst. Comput.*, 2020, pp. 707–711.
- [13] X. Ge, H. Jin, and V. C. M. Leung, "CDF-based scheduling algorithm for proportional throughput fairness," *IEEE Commun. Lett.*, vol. 20, no. 5, pp. 1034–1037, May 2016.
- [14] Y. Shimbo, N. Hiruma, and F. Maehara, "Performance evaluation of MU-MIMO THP with user scheduling in terms of system capacity and fairness," in *Proc. IEEE Int. Symp. Antennas Propag.*, 2017, pp. 1–2.
- [15] G. Liu, J. Zhang, F. Jiang, and W. Wang, "Joint spatial and frequency proportional fairness scheduling for MIMO OFDMA downlink," in *Proc. Int. Conf. Wireless Commun. Netw. Mobile Comput.*, 2007, pp. 491–494.
- [16] G. Femenias and F. Riera-Palou, "Scheduling and resource allocation in downlink multiuser MIMO-OFDMA systems," *IEEE Trans. Commun.*, vol. 64, no. 5, pp. 2019–2034, May 2016.
- [17] Q. Wu and R. Zhang, "Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming," *IEEE Trans. Wireless Commun.*, vol. 18, no. 11, pp. 5394–5409, Nov. 2019.
- [18] C. Huang, A. Zappone, G. C. Alexandropoulos, M. Debbah, and C. Yuen, "Reconfigurable intelligent surfaces for energy efficiency in wireless communication," *IEEE Trans. Wireless Commun.*, vol. 18, no. 8, pp. 4157–4170, Aug. 2019.
- [19] C. Huang, A. Zappone, M. Debbah, and C. Yuen, "Achievable rate maximization by passive intelligent mirrors," in *Proc. IEEE Int. Conf. Acoust. Speech, Signal Process. (ICASSP)*, 2018, pp. 3714–3718.
- [20] H. Guo, Y. Liang, J. Chen, and E. G. Larsson, "Weighted sum-rate maximization for reconfigurable intelligent surface aided wireless networks," *IEEE Trans. Wireless Commun.*, vol. 19, no. 5, pp. 3064–3076, May 2020.
- [21] C. Pan *et al.*, "Multicell MIMO communications relying on intelligent reflecting surfaces," *IEEE Trans. Wireless Commun.*, vol. 19, no. 8, pp. 5218–5233, Aug. 2020.
- [22] Z. Li, M. Hua, Q. Wang, and Q. Song, "Weighted sum-rate maximization for multi-IRS aided cooperative transmission," *IEEE Wireless Commun. Lett.*, vol. 9, no. 10, pp. 1620–1624, Oct. 2020.
- [23] X. Cao *et al.*, "Reconfigurable intelligent surface-assisted aerial-terrestrial communications via multi-task learning," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 10, pp. 3035–3050, Jun. 2021.
- [24] B. Yang, X. Cao, C. Huang, C. Yuen, L. Qian, and M. D. Renzo, "Intelligent spectrum learning for wireless networks with reconfigurable intelligent surfaces," *IEEE Trans. Veh. Technol.*, vol. 70, no. 4, pp. 3920–3925, Apr. 2021.
- [25] K. M. Attiah, F. Sahrabi, and W. Yu, "Deep learning approach to channel sensing and hybrid precoding for TDD massive MIMO systems," in *Proc. IEEE Glob. Commun. Conf. (GLOBECOM) Workshops*, 2020, pp. 1–6.
- [26] W. Cui, K. Shen, and W. Yu, "Spatial deep learning for wireless scheduling," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 6, pp. 1248–1261, Jun. 2019.
- [27] M. Lee, G. Yu, and G. Y. Li, "Graph embedding-based wireless link scheduling with few training samples," *IEEE Trans. Wireless Commun.*, vol. 20, no. 4, pp. 2282–2294, Apr. 2021.
- [28] Z. Zhao, G. Verma, C. Rao, A. Swami, and S. Segarra, "Link scheduling using graph neural networks," 2021. [Online]. Available: <https://arxiv.org/abs/2109.05536>
- [29] S. Ravanbakhsh, J. Schneider, and B. Póczos, "Equivariance through parameter-sharing," in *Proc. Inter. Conf. Mach. Learn.*, 2017, pp. 2892–2901.
- [30] Q. Shi, M. Razaviyayn, Z.-Q. Luo, and C. He, "An iteratively weighted MMSE approach to distributed sum-utility maximization for a MIMO interfering broadcast channel," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4331–4340, Sep. 2011.
- [31] A. Ghazanfari, T. Van Chien, E. Björnson, and E. G. Larsson, "Model-based and data-driven approaches for downlink massive MIMO channel estimation," *IEEE Trans. Commun.*, vol. 70, no. 3, pp. 2085–2101, Mar. 2022.
- [32] J. M. H. Etienne F. Chaponniere, P. J. Black, and D. N. C. Tse, "Transmitter directed code division multiple access system using path diversity to equitably maximize throughput," U.S. Patent 3, 45, 700, Jun. 30, 1999.
- [33] T. V. Chien *et al.*, "Controlling smart propagation environments: Long-term versus short-term phase shift optimization," in *Proc. IEEE Int. Conf. Acoust. Speech, Signal Process. (ICASSP)*, 2022, pp. 5348–5352.
- [34] S. Liang and R. Srikant, "Why deep neural networks for function approximation?," in *Proc. Int. Conf. Learn. Representations*, 2017.
- [35] J. Chen, Y.-C. Liang, H. V. Cheng, and W. Yu, "Channel estimation for reconfigurable intelligent surface aided multi-user MIMO systems," 2019. [Online]. Available: <https://arxiv.org/pdf/1912.03619.pdf>
- [36] T. Van Chien, H. Q. Ngo, S. Chatzinotas, M. Di Renzo, and B. Ottersten, "Reconfigurable intelligent surface-assisted cell-free massive MIMO systems over spatially-correlated channels," *IEEE Trans. Wireless Commun.*, early access, Dec. 29, 2021, doi: [10.1109/TWC.2021.3136925](https://doi.org/10.1109/TWC.2021.3136925).
- [37] Y. Shen, J. Zhang, S. H. Song, and K. B. Letaief, "AI empowered resource management for future wireless networks," in *Proc. IEEE Int. Mediterranean Conf. Commun. Netw.*, 2021, pp. 252–257.
- [38] N. Boumal, B. Mishra, P.-A. Absil, and R. Sepulchre, "Manopt, a matlab toolbox for optimization on manifolds," *J. Mach. Learn. Res.*, vol. 15, no. 1, pp. 1455–1459, 2014.
- [39] M. Abadi *et al.*, "Tensorflow: A system for large-scale machine learning," in *Proc. USENIX Conf. Operating Syst. Des. Implementation*, 2016, pp. 265–283.
- [40] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. Int. Conf. Learn. Representations*, 2015.
- [41] X. Tan, Z. Sun, D. Koutsonikolas, and J. M. Jornet, "Enabling indoor mobile millimeter-Wave networks based on smart reflect-arrays," in *Proc. IEEE Conf. Comput. Commun. (INFOCOM)*, 2018, pp. 270–278.



**Zhongze Zhang** (Graduate Student Member, IEEE) received the B.A.Sc. degree in computer engineering in 2020 from the University of Toronto, Toronto, ON, Canada, where he is currently working toward the M.A.Sc. degree with the Department of Electrical and Computer Engineering. His research interests include wireless communications, digital signal processing, machine learning, and optimization.



**Tao Jiang** (Graduate Student Member, IEEE) received the B.E. degree in telecommunications engineering from Xidian University, Xi'an, China, in 2017, and the M.E. degree in computer science from ShanghaiTech University, Shanghai, China, in 2020. He is currently working toward the Ph.D. degree with the University of Toronto, Toronto, ON, Canada. His main research interests include optimization theory, machine learning, and wireless communications.



**Wei Yu** (Fellow, IEEE) received the B.A.Sc. degree in computer engineering and mathematics from the University of Waterloo, Waterloo, ON, Canada, in 1997, and the M.S. and Ph.D. degrees in electrical engineering from Stanford University, Stanford, CA, USA, in 1998 and 2002, respectively. Since 2002, he has been with the Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON, Canada, where he is currently a Professor and holds the Canada Research Chair (Tier I) in Information Theory and Wireless Communications. He is a

Fellow of the Canadian Academy of Engineering and a member of the College of New Scholars, Artists, and Scientists of the Royal Society of Canada. Prof. Wei Yu was the President of the IEEE Information Theory Society in 2021, and has been on its Board of Governors since 2015. From 2017 to 2018, he was the Chair of the Signal Processing for Communications and Networking Technical Committee of the IEEE Signal Processing Society. From 2015 to 2016, he was an IEEE Communications Society Distinguished Lecturer. He was the recipient of the Steacie Memorial Fellowship in 2015, IEEE Marconi Prize Paper Award in Wireless Communications in 2019, IEEE Communications Society Award for Advances in Communication in 2019, IEEE Signal Processing Society Best Paper Award in 2008, 2017 and 2021, Journal of Communications and Networks Best Paper Award in 2017, and IEEE Communications Society Best Tutorial Paper Award in 2015. He is currently the Area Editor of IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, and was an Associate Editor for IEEE TRANSACTIONS ON INFORMATION THEORY, IEEE TRANSACTIONS ON COMMUNICATIONS, and IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS.