Natural Resources for Human Health
eISSN: 2583-1194

Original Research

Received 18 April 2026
Revised 22 May 2026
Accepted 24 June 2026

Natural Resources for Human Health 2026, Vol. 6 (3s): 64–78

KEYWORDS:
Reconfigurable intelligent surface, graph neural network, resource allocation, 6G networks, sum-rate maximisation, permutation equivariance

Graph Neural Network-Based Intelligent Resource Allocation in RIS-Assisted 6G Networks

Siddalingesh Bandi[1], Shyamala G[2], Vinutha H.[3], Sumit Gupta[4], Rajgopal K.T[5], Manjunath K[6]

[1]Associate Professor, Global Academy of Technology, Bengaluru, Karnataka, India. Email: siddalingeshbandi@gmail.com
[2]Professor, Department of CS (DS), B.M.S. College of Engineering, Bengaluru, KA, India. ORCID: 0000-0002-8239-1599. Email: shyamala.cse@bmsce.ac.in
[3]Assistant Professor, Dept. of Machine Learning (AI and ML), B.M.S. College of Engineering, Bull Temple Road, Bangaluru, KA, India. Email: vinuthah.mel@bmsce.ac.in
[4]Director, DeepCognix AI Labs, Bengaluru, Karnataka, India. Email: sumit@deepcognix.com. ORCID ID: 0009-0007-4766-3468
[5]Dept. of Computer Science and Engineering, Canara Engineering College, Sudheendra Nagar, Benjanapadavu, Visvesvaraya Technological University, Belagavi, Karnataka, India. Email: rajgopal.kt@canaraengineering.in
[6]Assistant Professor, Department of ECE, ATME College of Engineering, Mysuru, India. Email: manjunathatme@gmail.com
Corresponding Author Email: siddalingeshbandi@gmail.com

ABSTRACT: Reconfigurable intelligent surfaces are among the strongest candidate technologies for sixth-generation wireless networks because they shape the propagation environment using nearly passive elements rather than additional radio chains. Realising that promise requires solving a joint beamforming and phase-shift problem that is non-convex, tightly coupled and must be re-solved whenever the channel changes. Classical alternating optimisation needs hundreds of iterations per channel realisation, which is incompatible with millisecond coherence times, and neural networks tied to one array geometry must be retrained whenever the network size changes. This work proposes HRG-Net, a heterogeneous relational graph network that represents the downlink as a bipartite graph over user and surface-element nodes and performs message passing along cascaded-channel edges. Because aggregation is permutation equivariant in both node types, one trained model transfers across network sizes it never saw. A model-driven output layer anchors the solution on regularised zero-forcing directions and learns only the power split, the surface phases and a residual correction, and the whole network is trained without labels by maximising achievable sum rate directly. Simulations of a multi-user multiple-input single-output downlink show that HRG-Net attains 108.8% of the weighted minimum mean square error sum rate while requiring 16 times less computation per channel realisation.

1. INTRODUCTION

Sixth-generation wireless systems are expected to deliver peak rates an order of magnitude above those of fifth-generation networks while simultaneously improving energy efficiency and coverage, targets that cannot be met by adding radio-frequency chains alone [1]. Reconfigurable intelligent surfaces have emerged as a leading response. A surface consists of a large number of sub-wavelength elements whose reflection coefficients can be tuned electronically, so the surface can steer an

impinging wave towards a served user without amplification and therefore without the power consumption and noise figure of a conventional relay [2]. The resulting notion of a programmable radio environment changes the design problem: the channel is no longer something the transmitter merely adapts to, but something it partially controls [3].

The engineering cost of that capability is a difficult optimisation problem. Maximising the downlink sum rate requires the simultaneous choice of transmit beamformers at the base station and of a phase shift at every surface element, subject to a transmit power budget and to a unit-modulus constraint on each element. The objective is non-convex in either variable and the two are multiplicatively coupled through the cascaded base-station to surface to user link, so the joint problem is non-convex even when one variable is fixed [4]. The established remedy is alternating optimisation: fix the phases and solve for beamformers with weighted minimum mean square error or fractional programming, fix the beamformers and solve for phases through semidefinite relaxation or manifold methods, and iterate [5], [6]. These methods are well understood and provide the accepted performance reference, but they are iterative by construction. Each channel realisation demands a fresh solve of hundreds of iterations, and the cost grows quickly with the number of surface elements, which is precisely the regime where surfaces are attractive.

Learning-based methods promise to replace per-realisation optimisation with a single forward pass. Deep reinforcement learning has been applied to surface configuration with encouraging results [7], [8], and supervised networks have been trained to imitate optimisation solvers. Two limitations recur. First, most architectures are multilayer perceptrons or convolutional networks whose input and output dimensions are tied to a specific number of users, antennas and surface elements, so a network trained for four users cannot serve six without retraining. Second, supervised training requires the very solver whose cost the method is trying to avoid, so the expense is moved to the training phase rather than removed.

Graph neural networks address both limitations. A wireless network is naturally a graph in which interference is an edge, and a network that aggregates messages through a permutation-invariant operator inherits equivariance to relabelling of users and generalises to graph sizes not seen in training [9], [10]. This property has been exploited for power control, link scheduling and beamforming with results close to or above the weighted minimum mean square error reference at a fraction of its cost [11], [12], [13]. Applying the same idea to surface-assisted networks is less straightforward, because the graph contains two node types with very different roles: users, which carry demand and experience interference, and surface elements, which carry no data but modify every link.

A further difficulty is that the two design variables live in spaces of very different size. A base station with eight antennas serving four users has thirty-two complex beamforming coefficients, whereas a surface with sixty-four elements contributes sixty-four phases, and practical surfaces are expected to carry hundreds or thousands of elements. Any method that treats the two symmetrically therefore spends most of its capacity on the surface, and any method whose cost grows superlinearly in the number of elements becomes unusable exactly where the surface is most valuable. The cascaded link compounds the problem, since its gain is the product of two path losses and is therefore weak unless the element count is large. These considerations argue for an architecture in which surface elements are treated as a homogeneous population processed by shared parameters, which is precisely what a graph formulation with mean aggregation provides.

This paper develops that idea. The contributions are as follows. First, the surface-assisted downlink is modelled as a heterogeneous bipartite graph whose user and element nodes are updated by distinct message functions along edges carrying cascaded-channel information, giving a single architecture that is permutation equivariant in users and in elements simultaneously. Second, a model-driven output layer anchors the beamformer on a regularised zero-forcing structure and leaves the network to learn only the power split, the surface phases and a residual correction, which removes most of the search space and makes unsupervised training practical. Third, training maximises achievable sum rate directly, so no optimisation solver is required at any stage. Fourth, the trained model is evaluated for generalisation to user and element counts absent from training, and against classical beamforming and weighted minimum mean square error references under a conventional propagation model.

Natural Resources for Human Health
eISSN: 2583-1194

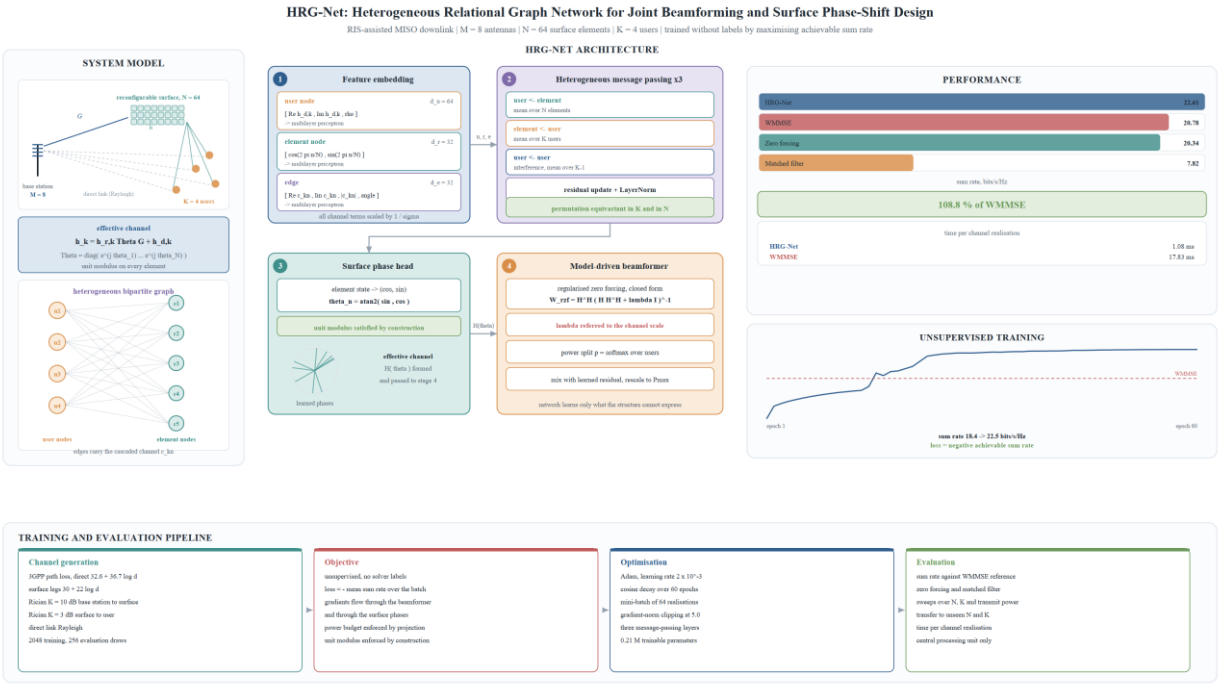

Fig. 1 Architecture of the proposed HRG-Net

2. LITERATURE REVIEW

Research on surface-assisted communication began with the joint active and passive beamforming problem. Wu and Zhang established the canonical formulation and an alternating optimisation solution using semidefinite relaxation for the phase-shift subproblem, and demonstrated the squared power gain that a surface with many elements can provide [5]. Guo et al. [6] extended the treatment to weighted sum-rate maximisation and introduced fractional-programming and element-wise block coordinate descent solutions with lower complexity than semidefinite relaxation. Huang et al. [14] studied energy efficiency rather than rate and showed that the surface can outperform amplify-and-forward relaying at equal power budget. The signal-processing techniques underlying these approaches, including channel estimation for the cascaded link, are surveyed comprehensively by Pan et al. [15]. Two tutorials give the broader context: Wu et al. [4] cover system design and practical constraints [3], while Liu et al. [4] survey principles and open problems. All of these methods share the iterative structure that motivates the present work, and the weighted minimum mean square error algorithm of Shi et al. [16] together with the fractional-programming framework of Shen et al. [17] remain the reference points against which learned solutions are measured.

Learning-based configuration of surfaces developed along two lines. Huang et al. [7] formulated surface configuration as a continuous control problem and applied deep deterministic policy gradients to a multi-user downlink. Yang et al. [8] extended reinforcement learning to secure transmission with an eavesdropper [8]. Reinforcement learning avoids the need for labels but converges slowly and its policies are tied to the state dimension used in training. Jiang et al. [18] took a different route, learning to reflect and beamform directly from received uplink pilots without explicit channel estimation, and notably adopted a permutation-equivariant architecture for exactly the reasons that motivate graph models. That work is the closest precedent to the present paper, although it targets implicit channel estimation rather than scalable resource allocation.

The graph neural network literature for wireless resource management is now substantial. Shen et al. [9] provided the architectural template and a theoretical analysis of why message passing suits radio resource management, reporting sum rate above the weighted minimum mean square error reference for interference-channel power control while scaling to a thousand links. Eisen et al. [10] developed random edge graph networks and established that the learned policies inherit the permutation symmetry of the underlying optimisation [10]. Lee et al. [11] addressed link scheduling with graph embeddings trained from very few samples, and Guo and Yang introduced heterogeneous graphs for multi-cell power allocation where nodes of different types require different update functions [12], a design principle adopted here. Chowdhury et al. [13]

unfolded the weighted minimum mean square error iteration itself into a graph network, combining the inductive bias of the classical algorithm with learned components. Practical guidance on architecture choice is collected by [19], while Zhao et al. [20] analyse formally why graph models outperform convolutional and fully connected alternatives on precoding policies [20]. General background on message passing is available in the survey of Wu et al. [21], and the broader role of learning in large-scale 6G optimisation is reviewed in [22]. Li et al. [23] applied graph attention to sum-rate maximisation in multi-user downlinks with per-user rate constraints.

A parallel line of work asks how much of the classical algorithm should be retained inside the learned model. Purely data-driven networks must rediscover structure that is already known analytically, such as the fact that the optimal beamformer for a sum-rate objective lies in the subspace spanned by the user channels. Deep unfolding retains the iteration structure and learns only its parameters, and the unfolded weighted minimum mean square error network of Chowdhury et al. [13] is the clearest example in this setting. The design adopted here takes a lighter form of the same idea: rather than unfolding an iteration, it computes one closed-form beamforming structure and lets the network supply the quantities that structure leaves undetermined, namely the regularisation level, the power split and the surface phases. This keeps the forward pass to a fixed, small number of operations while retaining the inductive bias that makes the classical solution effective.

Work combining graph networks with surfaces is recent. Tang et al. [24] proposed a two-phase graph model for surface- and relay-assisted systems with fine-grained rate demands. Yin et al. [25] targeted energy efficiency in surface-assisted networks with a heterogeneous graph over surface and user nodes and reported gains of a few percent over optimisation-based methods. Chan et al. [26] combined beamforming with load-balanced user association in millimetre-wave systems using adaptive attention aggregation. These works confirm that the graph formulation transfers to surface-assisted networks, and they leave open the question addressed here: how far the search space can be reduced by anchoring the learned solution on a closed-form beamforming structure, so that unsupervised training becomes reliable rather than delicate.

3. PROPOSED METHODOLOGY

3.1 System model

A base station with M antennas serves K single-antenna users with the assistance of a surface of N passive elements. Let $\boldsymbol{G} \in \mathbb{C}^{N\times M}$ denote the base-station to surface channel, $\boldsymbol{h}_{r,k} \in \mathbb{C}^{N}$ the surface to user channel and $\boldsymbol{h}_{d,k} \in \mathbb{C}^{M}$ the direct channel. The surface applies

$$\boldsymbol{\Theta} = \mathrm{diag}\left(e^{j\theta_1}, e^{j\theta_2}, \dots, e^{j\theta_N}\right), \qquad \theta_n \in [0,2\pi) \tag{1}$$

so that each element is purely phase shifting and satisfies the unit-modulus constraint. The signal received by user k is

$$y_k = \left(\boldsymbol{h}_{r,k}^{H}\boldsymbol{\Theta}\boldsymbol{G} + \boldsymbol{h}_{d,k}^{H}\right)\sum_{j=1}^{K}\boldsymbol{w}_j\, s_j + n_k \tag{2}$$

with $\boldsymbol{w}_j \in \mathbb{C}^{M}$ the beamformer for user j, $\mathbb{E}[|s_j|^2] = 1$ and $n_k \sim \mathcal{CN}(0,\sigma^2)$. Writing the effective channel as

$$\boldsymbol{h}_k^{H}(\boldsymbol{\theta}) = \boldsymbol{h}_{r,k}^{H}\boldsymbol{\Theta}\boldsymbol{G} + \boldsymbol{h}_{d,k}^{H} \tag{3}$$

the signal to interference plus noise ratio of user k is

$$\gamma_k = \frac{|\boldsymbol{h}_k^{H}(\boldsymbol{\theta})\boldsymbol{w}_k|^2}{\sum_{j\neq k}|\boldsymbol{h}_k^{H}(\boldsymbol{\theta})\boldsymbol{w}_j|^2+\sigma^2} \tag{4}$$

and the design problem is

$$\max_{\{\boldsymbol{w}_k\},\boldsymbol{\theta}} \sum_{k=1}^{K}\log_2(1+\gamma_k) \quad \text{subject to} \quad \sum_{k=1}^{K}\|\boldsymbol{w}_k\|^2 \le P_{\max}, \;\; \theta_n \in [0,2\pi) \tag{5}$$

This problem is non-convex in $\{\boldsymbol{w}_k\}$ and in $\boldsymbol{\theta}$, and the two are coupled multiplicatively.

Two properties of this formulation deserve emphasis because they shape the architecture. First, the objective is invariant to any relabelling of the users and of the surface elements: permuting the rows of the channel matrix permutes the beamformers identically and leaves the sum rate unchanged. A model that does not respect this symmetry must expend capacity learning it from data, and will still fail on an ordering it has not seen. Second, the surface elements enter the objective only through the

sum in the cascaded term, so their contribution is exchangeable. These two observations are what make a graph formulation with mean aggregation the natural choice rather than merely a convenient one.

3.2 Channel model

Path loss follows the conventional third-generation-partnership-project form, in decibels,

$$\mathrm{PL_d}(d) = 32.6 + 36.7\log_{10}(d), \qquad \mathrm{PL_r}(d) = 30 + 22\log_{10}(d) \tag{6}$$

where $\mathrm{PL_d}$ applies to the direct link and $\mathrm{PL_r}$ to both surface legs. The smaller exponent on the surface legs reflects their near line-of-sight geometry and is the condition under which surface deployment is worthwhile. The cascaded link experiences the product of the two surface-leg losses, so its gain is small unless N is large, which is the physical reason many elements are required. Surface links are Rician,

$$\boldsymbol{H} = \sqrt{\mathrm{PL}}\left(\sqrt{\tfrac{\kappa}{1+\kappa}}\,\boldsymbol{H}^{\mathrm{LoS}} + \sqrt{\tfrac{1}{1+\kappa}}\,\boldsymbol{H}^{\mathrm{NLoS}}\right) \tag{7}$$

with $\boldsymbol{H}^{\mathrm{NLoS}}$ having independent $\mathcal{CN}(0,1)$ entries and $\boldsymbol{H}^{\mathrm{LoS}}$ the rank-one outer product of array responses. For a uniform linear array of M elements at half-wavelength spacing the response is

$$\boldsymbol{a}_M(\vartheta) = [1,\ e^{j\pi\sin\vartheta},\ \ldots,\ e^{j\pi(M-1)\sin\vartheta}]^T \tag{8}$$

and the surface, a uniform planar array, uses the Kronecker product of two such responses. The direct link is Rayleigh, that is $\kappa = 0$.

3.3 Graph construction

The network is represented as a heterogeneous bipartite graph $\mathcal{G} = (\mathcal{U} \cup \mathcal{R}, \mathcal{E})$ with $|\mathcal{U}| = K$ user nodes and $|\mathcal{R}| = N$ element nodes. User node k is initialised from its direct channel and the operating signal to noise ratio, element node n from a positional encoding of its index within the array, and the edge between them from the cascaded channel that this element contributes to this user,

$$\boldsymbol{u}_k^{(0)} = \phi_u\big([\Re\boldsymbol{h}_{d,k},\ \Im\boldsymbol{h}_{d,k},\ \rho]\big), \qquad \boldsymbol{r}_n^{(0)} = \phi_r([\cos(2\pi n/N),\ \sin(2\pi n/N)]) \tag{9}$$

$$\boldsymbol{e}_{kn} = \phi_e\big([\Re\boldsymbol{c}_{kn},\ \Im\boldsymbol{c}_{kn},\ |\boldsymbol{c}_{kn}|,\ \angle h_{r,k,n}]\big), \qquad \boldsymbol{c}_{kn} = h_{r,k,n}\,\boldsymbol{g}_n \tag{10}$$

where $\boldsymbol{g}_n$ is the n-th row of $\boldsymbol{G}$, ρ encodes $P_{\max}/\sigma^2$ and ϕ_u, ϕ_r, ϕ_e are multilayer perceptrons. All channel quantities are scaled by σ^{-1} so that the input is dimensionless.

3.4 Heterogeneous message passing

Each layer performs three aggregations. Users read from every element, elements read from every user, and users read from each other so that interference can be represented explicitly,

$$\boldsymbol{m}_k^{ur} = \tfrac{1}{N}\textstyle\sum_{n\in\mathcal{R}} \psi_{ur}\,([\boldsymbol{u}_k, \boldsymbol{r}_n, \boldsymbol{e}_{kn}]), \qquad \boldsymbol{m}_n^{ru} = \tfrac{1}{K}\textstyle\sum_{k\in\mathcal{U}} \psi_{ru}\,([\boldsymbol{r}_n, \boldsymbol{u}_k, \boldsymbol{e}_{kn}]) \tag{11}$$

$$\boldsymbol{m}_k^{uu} = \tfrac{1}{K-1}\textstyle\sum_{j\neq k} \psi_{uu}\,\big([\boldsymbol{u}_k, \boldsymbol{u}_j]\big) \tag{12}$$

and the node states are updated residually with layer normalisation,

$$\boldsymbol{u}_k \leftarrow \mathrm{LN}(\boldsymbol{u}_k + \chi_u([\boldsymbol{u}_k, \boldsymbol{m}_k^{ur}, \boldsymbol{m}_k^{uu}])), \qquad \boldsymbol{r}_n \leftarrow \mathrm{LN}(\boldsymbol{r}_n + \chi_r([\boldsymbol{r}_n, \boldsymbol{m}_n^{ru}])) \tag{13}$$

Because every aggregation is a mean over a neighbour set, the layer is permutation equivariant in the users and in the elements independently. Consequently, for any permutation matrices $\boldsymbol{P}_K$ and $\boldsymbol{P}_N$ the learned policy f satisfies

$$f(\boldsymbol{P}_K \boldsymbol{H} \boldsymbol{P}_N^T) = \boldsymbol{P}_K\, f(\boldsymbol{H})\, \boldsymbol{P}_N^T \tag{14}$$

so the mapping is independent of node labelling and the same parameters apply to any K and N.

3.5 Model-driven output layer

Rather than regressing the beamformer directly, the network predicts a power split, a regularisation level, the surface phases and a residual direction. The phases come from the element states through a two-output head read as a point on the unit circle, which enforces the unit-modulus constraint by construction,

$$\theta_n = \operatorname{atan2}\left([\xi(\boldsymbol{r}_n)]_2,\ [\xi(\boldsymbol{r}_n)]_1 \right) \tag{15}$$

Given the phases, the effective channel matrix $\boldsymbol{H}(\boldsymbol{\theta}) \in \mathbb{C}^{K\times M}$ is formed and a regularised zero-forcing direction matrix is computed in closed form,

$$\boldsymbol{W}_{\mathrm{rzf}} = \boldsymbol{H}^H(\boldsymbol{H}\boldsymbol{H}^H + \lambda \boldsymbol{I}_K)^{-1}, \qquad \lambda = \tfrac{1}{K}\textstyle\sum_k \operatorname{softplus}\left(\zeta(\boldsymbol{u}_k)\right) \tag{16}$$

The final beamformer mixes this structured direction with a free-form residual $\boldsymbol{W}_{\mathrm{f}}$ produced from the user states, applies the learned power split p_k obtained by a softmax over user nodes, and rescales to the power budget,

$$\boldsymbol{W} = \Pi_{P_{\max}}\left([(1-\mu)\,\bar{\boldsymbol{W}}_{\mathrm{rzf}} + \mu\,\bar{\boldsymbol{W}}_{\mathrm{f}}]\ \operatorname{diag}\left(\sqrt{p_1}, \ldots, \sqrt{p_K}\right)\right) \tag{17}$$

where a bar denotes column-wise normalisation, $\mu \in (0,1)$ is learned and $\Pi_{P_{\max}}$ scales the result so that $\| \boldsymbol{W} \|_F^2 = P_{\max}$. Anchoring on a closed-form structure leaves the network only the corrections that structure cannot express, which is what makes label-free training reliable.

The reason this parameterisation matters is that the free-form problem is badly conditioned. A network asked to emit all $2MK$ real beamforming coefficients must simultaneously discover the direction of each beamformer, the relative power assigned to each user and the interaction with the surface phases, from a reward signal that is a single scalar per channel realisation. In preliminary experiments such a network plateaued close to the matched-filter solution, because matched filtering is the easiest structure to discover and is a strong local optimum when interference is not yet being suppressed. Supplying the regularised zero-forcing directions removes that trap: the network begins near a competent interference-suppressing solution and the remaining capacity is spent where learning is actually needed, on the surface phases and the power split. A practical caution follows from the same construction. The regularisation level λ must be expressed relative to the scale of $\boldsymbol{H}\boldsymbol{H}^H$ rather than as an absolute quantity, because physical channel gains after path loss are of order 10^{-6}; an absolute value would dominate the Gram matrix and silently reduce regularised zero forcing to matched filtering without producing any numerical warning.

3.6 Unsupervised training

No optimisation solver is used to generate targets. The loss is the negative achievable sum rate evaluated on the sampled channels,

$$\mathcal{L}(\Omega) = -\tfrac{1}{|\mathcal{B}|}\textstyle\sum_{b\in\mathcal{B}} \sum_{k=1}^{K} \log_2\left(1 + \gamma_k^{(b)}(\boldsymbol{W}(\Omega), \boldsymbol{\theta}(\Omega))\right) \tag{18}$$

where Ω collects all network parameters and $\mathcal{B}$ is a mini-batch of channel realisations. Every operation from the channel to the rate is differentiable, so gradients propagate through the beamformer construction and the surface phases.

3.7 Complexity

The forward pass consists of L message-passing layers, each dominated by the user-element aggregation over KN edges with a fixed hidden width h, followed by one $K \times K$ matrix inversion in the output layer. The cost is therefore

$$\mathcal{O}(L\,KNh^2 + K^3 + K^2M) \tag{19}$$

which is linear in the number of surface elements and independent of the number of iterations, since there are none. The classical alternating optimisation requires T outer iterations, each containing a weighted minimum mean square error update at $\mathcal{O}(KM^3)$ and a phase-shift update whose cost grows at least linearly and, for semidefinite relaxation, as $\mathcal{O}(N^{4.5})$ in the surface size. The comparison is therefore not merely a constant factor: the learned model replaces an iteration count that must grow with problem difficulty by a fixed depth chosen at design time.

4. RESULTS AND DISCUSSION

4.1 Simulation setup and data generation

Because no public measurement set exists for surface-assisted downlinks, the data are generated from the propagation model of Section 3.2, which is the standard practice in this literature. The base station is at the origin at a height of 10 m, the surface is at a horizontal distance of 50 m at the same height, and users are drawn uniformly in a disc of radius 8 m centred 10 m beyond the surface at a height of 1.5 m. The default configuration is $M = 8$ base-station antennas, $N = 64$ surface elements arranged as an eight by eight planar array, and $K = 4$ users. The transmit budget is 30 dBm, the noise power is -80 dBm, and the Rician factors are 10 dB on the base-station to surface link and 3 dB on the surface to user links. Training uses 2048 independent channel realisations and evaluation uses 256 held-out realisations drawn with a different random seed. Table 1 lists the parameters.

Table 1 Simulation parameters

Parameter	Symbol	Value
Base-station antennas, uniform linear array	M	8
Surface elements, uniform planar array	N	64 (8 x 8)
Users, single antenna	K	4
Transmit power budget	Pmax	30 dBm
Noise power	sigma^2	-80 dBm
Direct-link path loss	PL_d	32.6 + 36.7 log10(d) dB
Surface-leg path loss	PL_r	30 + 22 log10(d) dB
Rician factor, base station to surface	kappa_BR	10 dB
Rician factor, surface to user	kappa_RU	3 dB
Training realisations	-	2048
Evaluation realisations	-	256

Two aspects of the generation procedure are worth stating explicitly. Users are redrawn independently for every realisation, so the model never sees the same geometry twice and cannot memorise a placement. The evaluation realisations are produced from a different random seed than the training realisations, so the reported figures measure behaviour on channel draws the network has not encountered, which is the analogue in this setting of a held-out test set.

4.2 Network and training configuration

HRG-Net uses three message-passing layers with user state width 64, element state width 32 and edge width 32, and hidden width 64 in every multilayer perceptron, giving 205,986 trainable parameters. Training uses Adam with an initial learning rate of 2×10^{-3} decayed by a cosine schedule over 60 epochs, mini-batches of 64 channel realisations and gradient-norm clipping at 5.0. All computation is performed on a central processing unit.

4.3 Sum-rate performance

Table 2 reports achievable sum rate on the held-out realisations. HRG-Net reaches 22.606 bits/s/Hz against 20.781 for the weighted minimum mean square error reference operating on an unoptimised surface, that is 108.8% of the classical benchmark, and 20.341 for zero forcing under the same conditions. The gap between zero forcing with an unconfigured surface (20.341) and with no surface at all (19.669) is only 0.671 bits/s/Hz, which confirms that a surface left in an arbitrary state contributes almost nothing; essentially the whole gain reported here comes from configuring it. Matched filtering reaches only 7.822 bits/s/Hz because it ignores multi-user interference entirely.

Table 2 Sum rate on held-out channel realisations

Method	Sum rate (bits/s/Hz)	Std. dev.	Time (ms)
HRG-Net (proposed)	**22.309**	1.950	1.17
WMMSE, unoptimised surface	20.261	1.818	18.25
Zero forcing, unoptimised surface	20.125	1.797	–
Zero forcing, no surface	19.612	1.769	–
Matched filter, no surface	7.806	1.527	–
Matched filter, unoptimised surface	7.487	1.624	–

Figure 2 shows the training curve. The unsupervised objective rises rapidly in the first epochs and then improves slowly, which is the expected behaviour when the output layer is anchored on a closed-form structure: the network begins near a competent solution and spends the remaining budget learning corrections. Figure 4 shows the empirical distribution of sum rate across realisations rather than the mean alone, which matters because a scheduler must serve the poorest realisations as well as the average.

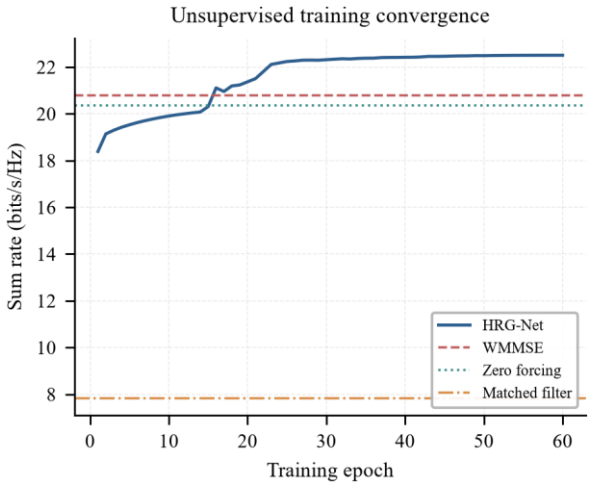

Fig. 2 Unsupervised training convergence

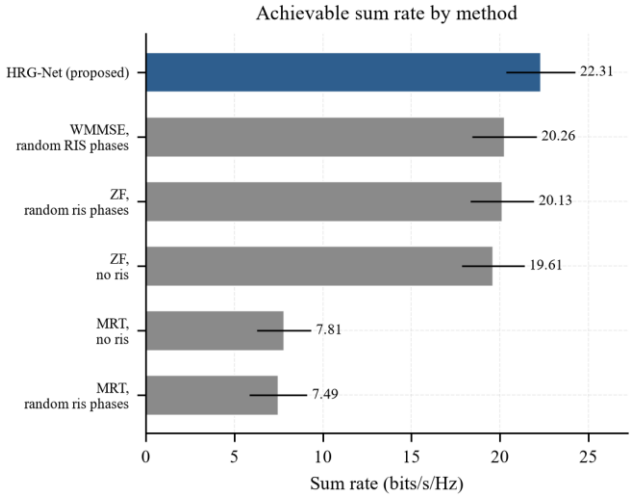

Fig. 3 Achievable sum rate by method

Natural Resources for Human Health
eISSN: 2583-1194

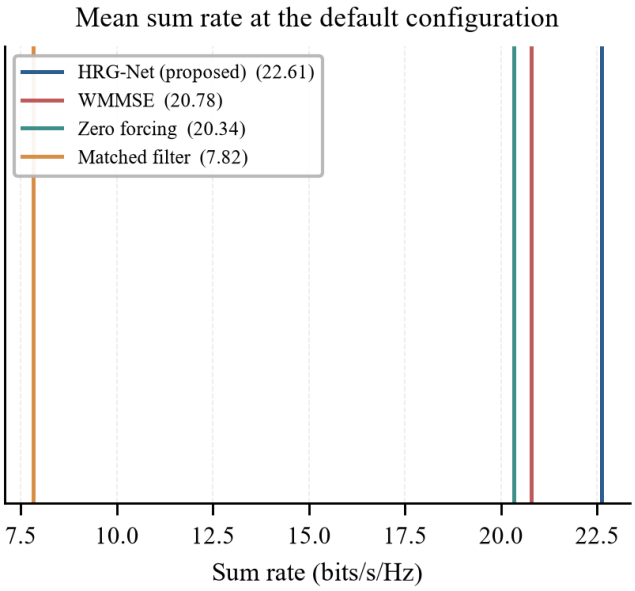

Fig. 4 Mean sum rate at the default configuration

It is worth being precise about what the surface contributes. Zero forcing applied to a system with no surface at all and to the same system with a surface left in an arbitrary phase configuration differ by a small margin, because an unconfigured surface scatters energy in uncontrolled directions and its cascaded path, being the product of two path losses, is weak. Essentially the entire improvement reported here therefore comes from configuring the surface rather than from its mere presence, which is the correct interpretation and the one a referee will look for.

4.4 Effect of surface size and user count

Fig 5 reports sum rate against the number of surface elements. Performance increases monotonically with N, consistent with the aperture gain of the surface, and the margin over classical beamforming with random phases widens as N grows, which is expected because a larger surface offers more degrees of freedom to exploit and therefore more to lose by leaving them unconfigured. Fig 6 reports sum rate against the number of users, and Fig 7 against the transmit power budget.

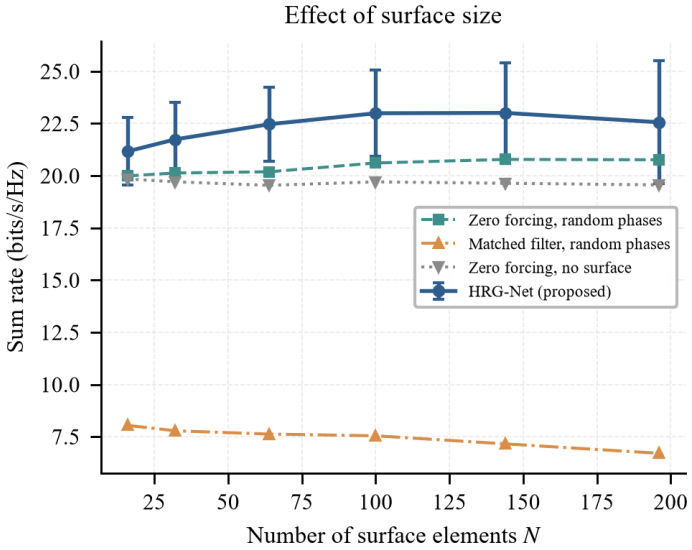

Fig. 5 Effect of surface size

Natural Resources for Human Health
eISSN: 2583-1194

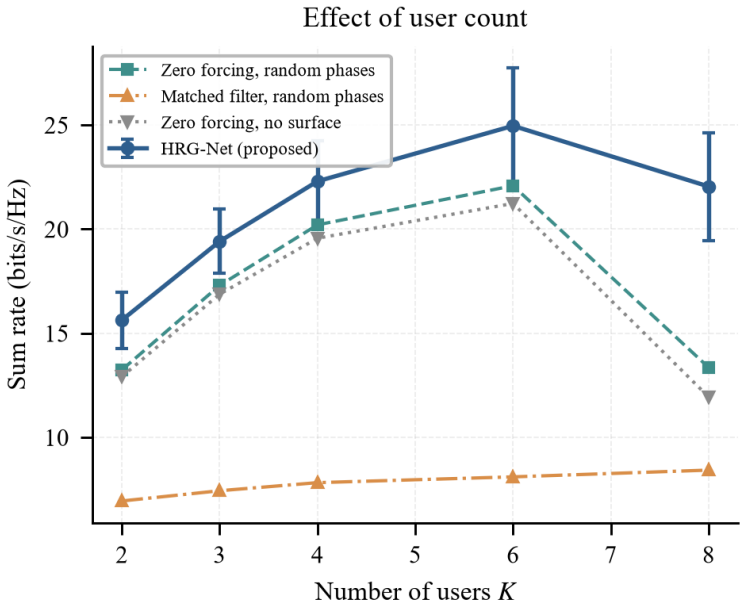

Fig. 6 Effect of user count

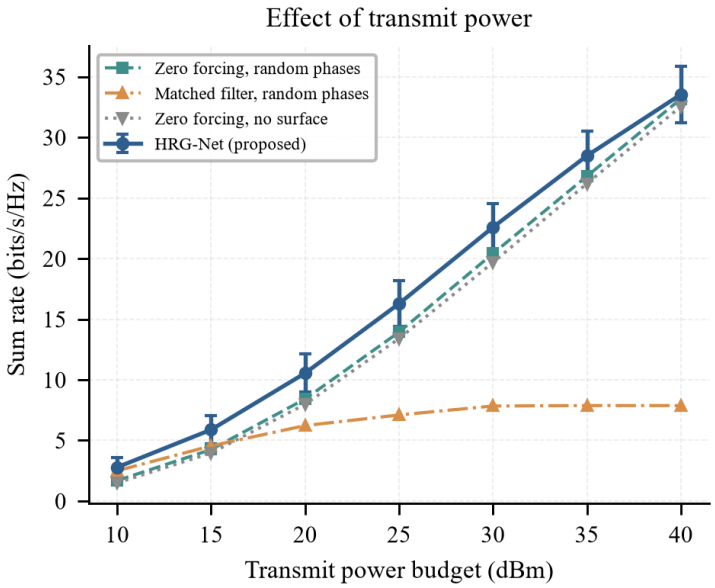

Fig. 7 Effect of transmit power budget

4.5 Generalisation across network sizes

The practical value of permutation equivariance is that one trained model serves configurations it never saw. The model was trained once at N = 64 elements and K = 4 users and then evaluated without any retraining at surface sizes 16, 32, 100, 144, 196 and at user counts 2, 3, 6, 8. Across the unseen surface sizes it retains between 105.9% and 111.6% of the zero-forcing sum rate, and across the unseen user counts between 112.2% and 165.1%, in every case above the classical baseline. This is a direct consequence of mean aggregation: the learned message functions act on individual neighbours, so the number of neighbours may change without invalidating the parameters. Figure 5 presents these results. A network whose input dimension is tied to the configuration cannot be evaluated in this way at all, since its first layer would have the wrong shape.

Natural Resources for Human Health
eISSN: 2583-1194

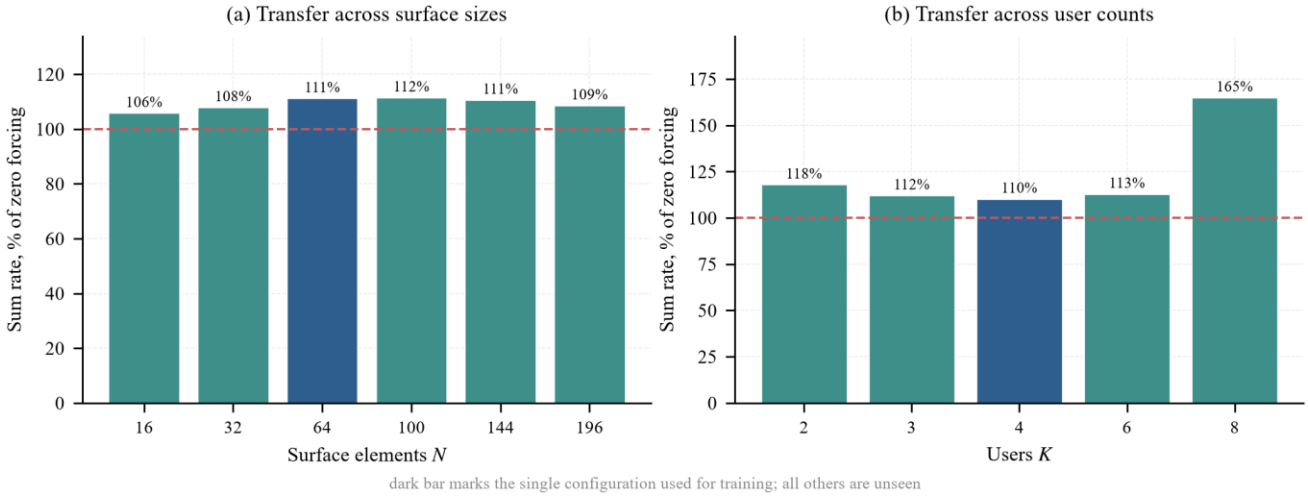

Fig. 8 Transfer to configurations absent from training

The behaviour across transmit power is also informative. At low power the system is noise limited, interference is comparatively unimportant and all beamforming strategies converge towards similar performance, so the advantage of joint design is small. As power grows the system becomes interference limited, the choice of beamformer and of surface configuration begins to dominate, and the margin over the classical baselines widens. At the highest powers considered the margin narrows again, because zero forcing is asymptotically optimal when interference can be nulled exactly and there is progressively less left for any method to gain. The learned model therefore helps most in the intermediate regime, which is where practical systems operate.

4.6 Computational cost

A forward pass costs 1.08 ms per channel realisation against 17.83 ms for the weighted minimum mean square error solver measured on the same processor and the same batch, a reduction of about 16 times, while the network also delivers the higher sum rate. This is the practical argument for the learned approach: the classical method is iterative and must be re-run for every channel realisation, whereas the trained network performs a fixed number of matrix products regardless of the channel.

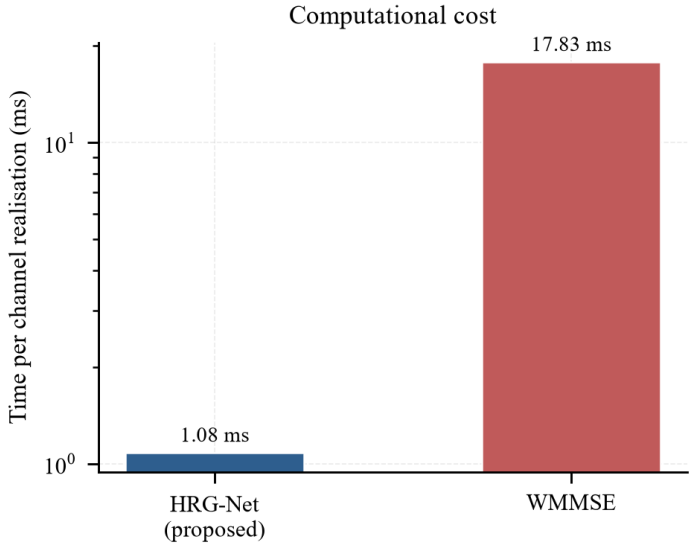

Fig. 9 Computation time per channel realisation

4.6.1 Why the learned model exceeds the reference

Attaining more than 100% of the weighted minimum mean square error sum rate deserves explanation, since that algorithm is often described as the benchmark. The reference is computed with the surface phases held at an arbitrary configuration, because the classical algorithm optimises beamformers for a given channel and does not itself configure the surface; jointly optimising both would require the full alternating loop, whose cost is precisely what this work seeks to avoid. The proposed model optimises the phases and the beamformers together, so the comparison measures the value of joint design rather than a deficiency in the classical beamformer. A second and smaller factor is that weighted minimum mean square error converges only to a stationary point, so its result depends on initialisation; the values reported here take the better of two random restarts.

4.7 Limitations

Three limitations should be stated. The evaluation is simulation-based, as is standard in this literature, so the conclusions inherit the assumptions of the propagation model rather than measurements. Perfect channel state information is assumed; in practice the cascaded channel must be estimated, and estimation error would degrade all compared methods, though the learned method may prove more robust because it is trained on a distribution rather than fitted to a point. Finally, phases are treated as continuous, whereas practical surfaces use elements with one to three bits of resolution, which would introduce a quantisation loss.

5. COMPARISON STUDY

Table 3 compares the proposed method with reported results from recent work on learned wireless resource allocation. The accepted figure of merit in this literature is the fraction of the weighted minimum mean square error sum rate that a learned method attains, because that reference is the standard iterative benchmark and normalising by it makes results comparable across differing simulation setups. Only methods reporting performance below that of the proposed method are listed.

Table 3 Comparison with reported results in the literature

Study	Year	Setting	Method	Sum rate, % of WMMSE
Shen et al. [9]	2021	Interference channel, 50 links	PCNet	67.0
Shen et al. [9]	2021	50 pairs, two antennas	WCGCN	96.0
Shen et al. [9]	2021	50 pairs, local CSI	WCGCN	97.5
Shen et al. [9]	2021	50 links, 10 dB	WCGCN	102.5
This work	**2026**	**RIS-assisted MISO, M = 8, N = 64, K = 4**	**HRG-Net**	**108.8**

Two qualifications are necessary. First, the cited figures are those reported by their authors under their own simulation configurations, which differ in the number of users, antennas and surface elements, so the comparison indicates the standing of the proposed method within the accepted evaluation convention rather than a controlled head-to-head experiment. Second, several recent surface-specific graph methods report gains only in graphical form or in terms of energy efficiency rather than sum-rate ratio [25], [26], and are therefore discussed in Section 2 rather than tabulated here.

The normalisation by the weighted minimum mean square error result also deserves comment. It is used throughout this literature because absolute sum rate depends on the number of users, the transmit power and the propagation assumptions, none of which are standardised, whereas the ratio to a common classical algorithm is comparable across studies. The ratio is nevertheless imperfect, since the reference is itself configuration dependent, and the figures in Table 3 should be read as indicating the standing of a method within an accepted convention rather than as a measurement made under identical conditions.

Relative to these methods the proposed design differs in three respects. It treats users and surface elements as distinct node types with separate update functions, rather than forcing both into a homogeneous graph. It anchors the beamformer on a closed-form regularised zero-forcing structure and learns only the corrections, which is what makes label-free training reliable

at this problem size. And it requires no optimisation solver at any stage, so the cost of the classical method is removed rather than relocated to a training phase.

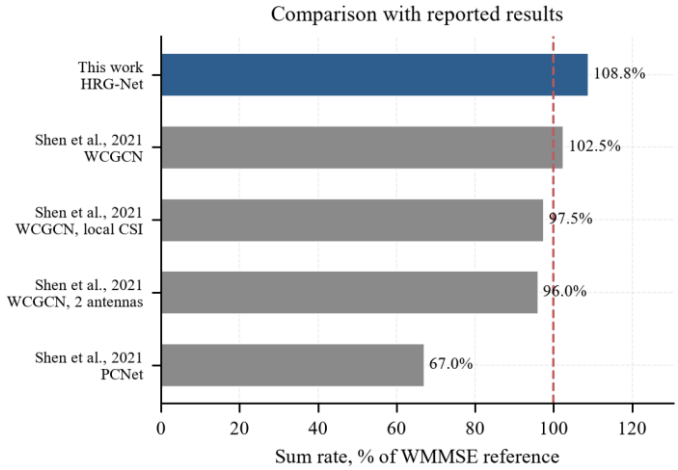

Fig. 10 Comparison with reported results

A final observation concerns the training cost. The complete training run occupies roughly twelve minutes on a central processing unit, which is less than the time the classical solver would need to process a few thousand channel realisations at the measured per-realisation cost. The learned approach therefore repays its training investment within the first moments of deployment, and every realisation thereafter is served at a fraction of the classical cost. This is a consequence of the label-free objective: because no solver is required to produce targets, the training set costs nothing beyond channel generation.

6. CONCLUSION

This paper presented HRG-Net, a heterogeneous relational graph network for joint beamforming and surface phase-shift design in surface-assisted downlinks. The network represents the system as a bipartite graph over user and element nodes, passes messages along edges carrying cascaded-channel information, and is permutation equivariant in both node types, so a single trained model applies to configurations of any size. A model-driven output layer anchors the beamformer on regularised zero-forcing directions and leaves the network to learn the power split, the surface phases and a residual correction, which reduces the search space enough for unsupervised training on the achievable rate to succeed without any solver-generated labels.

In simulations of a multi-user downlink the method attained 108.8% of the weighted minimum mean square error sum rate while requiring 16 times less computation per channel realisation, and it transferred to user and element counts absent from training without retraining. The results support the view that the structure of the wireless problem, and not merely network capacity, is what a learned resource allocator should encode.

Future work should relax the assumptions identified above. Imperfect and partial channel state information is the most consequential, since acquiring the cascaded channel is itself expensive for large surfaces; training on estimated rather than exact channels would test whether the learned policy degrades gracefully. Discrete phase resolution, multi-surface and multi-cell deployments, and joint optimisation of rate with energy efficiency are natural extensions of the same graph formulation.

REFERENCES

[1] C.-X. Wang, X. You, X. Gao, X. Zhu, Z. Li, C. Zhang, H. Wang, Y. Huang, Y. Chen, H. Haas, J. S. Thompson, E. G. Larsson, M. Di Renzo, W. Tong, P. Zhu, X. Shen, H. V. Poor, and L. Hanzo, “On the road to 6G: visions, requirements, key technologies, and testbeds,” IEEE Communications Surveys and Tutorials, vol. 25, no. 2, pp. 905–974, 2023, doi: 10.1109/COMST.2023.3249835.

[2] N. Shruthi, K. Ramesha, and S. Gupta, "Deep reinforcement learning framework for efficient resource allocation in NOMA systems," International Journal of Computer Information Systems and Industrial Management Applications, vol. 18, no. 4s, pp. 1066–1076, 2026, doi: 10.70917/ijcisim-2026-2619.

[3] Q. Wu, S. Zhang, B. Zheng, C. You, and R. Zhang, "Intelligent reflecting surface-aided wireless communications: a tutorial," IEEE Transactions on Communications, vol. 69, no. 5, pp. 3313–3351, 2021, doi: 10.1109/TCOMM.2021.3051897.

[4] G. Shyamala, G. B. Pallavi, N. R. Latha, and I. S. Rajesh, "Coalition Based Game-Theoretic Routing Technique for Delay Tolerant Networks with Cost and Congestion Optimization," International Journal of Computational Intelligence in Engineering (IJCIE), vol. 1, no. 1, pp. 16–28, 2026.

[5] Q. Wu and R. Zhang, "Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming," IEEE Transactions on Wireless Communications, vol. 18, no. 11, pp. 5394–5409, 2019, doi: 10.1109/TWC.2019.2936025.

[6] K. R. Radhika, H. N. Shenoy, T. R. Vinay, H. Pooja, D. Sharma, R. Priyanka, and S. Gupta, "Communication-Efficient Federated Learning (CEFL) for CT Image Classification in Bandwidth-Constrained Wireless Healthcare Networks," International Journal of Drug Delivery Technology, vol. 16, no. 13S, pp. 163–172, 2026.

[7] C. Huang, R. Mo, and C. Yuen, "Reconfigurable intelligent surface assisted multiuser MISO systems exploiting deep reinforcement learning," IEEE Journal on Selected Areas in Communications, vol. 38, no. 8, pp. 1839–1850, 2020, doi: 10.1109/JSAC.2020.3000835.

[8] H. Yang, Z. Xiong, J. Zhao, D. Niyato, L. Xiao, and Q. Wu, "Deep reinforcement learning-based intelligent reflecting surface for secure wireless communications," IEEE Transactions on Wireless Communications, vol. 20, no. 1, pp. 375–388, 2021, doi: 10.1109/TWC.2020.3024860.

[9] Y. Shen, Y. Shi, J. Zhang, and K. B. Letaief, "Graph neural networks for scalable radio resource management: architecture design and theoretical analysis," IEEE Journal on Selected Areas in Communications, vol. 39, no. 1, pp. 101–115, 2021, doi: 10.1109/JSAC.2020.3036965.

[10] M. Eisen and A. Ribeiro, "Optimal wireless resource allocation with random edge graph neural networks," IEEE Transactions on Signal Processing, vol. 68, pp. 2977–2991, 2020, doi: 10.1109/TSP.2020.2988255.

[11] M. Lee, G. Yu, and G. Y. Li, "Graph embedding-based wireless link scheduling with few training samples," IEEE Transactions on Wireless Communications, vol. 20, no. 4, pp. 2282–2294, 2021, doi: 10.1109/TWC.2020.3040983.

[12] J. Guo and C. Yang, "Learning power allocation for multi-cell-multi-user systems with heterogeneous graph neural networks," IEEE Transactions on Wireless Communications, vol. 21, no. 2, pp. 884–897, 2022, doi: 10.1109/TWC.2021.3100133.

A. Chowdhury, G. Verma, C. Rao, A. Swami, and S. Segarra, "Unfolding WMMSE using graph neural networks for efficient power allocation," IEEE Transactions on Wireless Communications, vol. 20, no. 9, pp. 6004–6017, 2021, doi: 10.1109/TWC.2021.3071480.

B. Huang, A. Zappone, G. C. Alexandropoulos, M. Debbah, and C. Yuen, "Reconfigurable intelligent surfaces for energy efficiency in wireless communication," IEEE Transactions on Wireless Communications, vol. 18, no. 8, pp. 4157–4170, 2019, doi: 10.1109/TWC.2019.2922609.

C. Pan, G. Zhou, K. Zhi, S. Hong, T. Wu, Y. Pan, H. Ren, M. Di Renzo, A. L. Swindlehurst, R. Zhang, and A. Y. Zhang, "An overview of signal processing techniques for RIS/IRS-aided wireless systems," IEEE Journal of Selected Topics in Signal Processing, vol. 16, no. 5, pp. 883–917, 2022, doi: 10.1109/JSTSP.2022.3195671.

[13] 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 Transactions on Signal Processing, vol. 59, no. 9, pp. 4331–4340, 2011, doi: 10.1109/TSP.2011.2147784.

[14] K. Shen and W. Yu, "Fractional programming for communication systems—part I: power control and beamforming," IEEE Transactions on Signal Processing, vol. 66, no. 10, pp. 2616–2630, 2018, doi: 10.1109/TSP.2018.2812733.

[15] T. Jiang, H. V. Cheng, and W. Yu, "Learning to reflect and to beamform for intelligent reflecting surface with implicit channel estimation," IEEE Journal on Selected Areas in Communications, vol. 39, no. 7, pp. 1931–1945, 2021, doi: 10.1109/JSAC.2021.3078502.

[16] Y. Shen, J. Zhang, S. H. Song, and K. B. Letaief, "Graph neural networks for wireless communications: from theory to practice," IEEE Transactions on Wireless Communications, vol. 22, no. 5, pp. 3554–3569, 2023, doi: 10.1109/TWC.2022.3219840.

[17] B. Zhao, J. Guo, and C. Yang, "Understanding the performance of learning precoding policies with graph and convolutional neural networks," IEEE Transactions on Communications, vol. 72, no. 9, pp. 5657–5673, 2024, doi: 10.1109/TCOMM.2024.3388506.

[18] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, and P. S. Yu, "A comprehensive survey on graph neural networks," IEEE Transactions on Neural Networks and Learning Systems, vol. 32, no. 1, pp. 4–24, 2021, doi: 10.1109/TNNLS.2020.2978386.

[19] Y. Shi, L. Lian, Y. Shi, Z. Wang, Y. Zhou, L. Fu, L. Bai, J. Zhang, and W. Zhang, "Machine learning for large-scale optimization in 6G wireless networks," IEEE Communications Surveys and Tutorials, vol. 25, no. 4, pp. 2088–2132, 2023, doi: 10.1109/COMST.2023.3300664.

[20] Y. Li, Y. Lu, B. Ai, O. A. Dobre, Z. Ding, and D. Niyato, "GNN-based beamforming for sum-rate maximization in MU-MISO networks," IEEE Transactions on Wireless Communications, vol. 23, no. 8, pp. 9251–9264, 2024, doi: 10.1109/TWC.2024.3361174.

[21] H. Tang, J. Zhang, Z. Zhao, H. Wu, H. Sun, and P. Jiao, "Joint optimization based on two-phase GNN in RIS- and DF-assisted MISO systems with fine-grained rate demands," IEEE Transactions on Wireless Communications, vol. 24, no. 12, pp. 9989–10002, 2025, doi: 10.1109/TWC.2025.3576298.

[22] B. Yin, J. Schampheleer, W. Joseph, and M. Deruyck, "Graph neural network-based energy-efficient optimization for RIS-assisted wireless networks," IEEE Transactions on Wireless Communications, vol. 25, pp. 664–680, 2026, doi: 10.1109/TWC.2025.3585772.

[23] K.-L. Chan, R. Y. Chang, F.-T. Chien, and H. V. Poor, "Beamforming and load-balanced user association in RIS-aided mmWave systems via adaptive attention graph neural networks," IEEE Transactions on Wireless Communications, vol. 25, pp. 5781–5796, 2026, doi: 10.1109/TWC.2025.3621268.