A Novel Deep Unsupervised Learning Method for Sum-Rate Optimization in Device-to-Device Networks with a Quality-of-Service Constraint

BINDUBRITTA ACHARJEE

PRESENTATION OUTLINE

01 Research background

Problem statement, research objectives, and significance

02 Review of existing literature

Overview to establish the foundation and identify gaps

03 Method of research

Formulation of the optimization problem and research design

04 Results and discussions

Simulation results, major findings, and future directions

ACKNOWLEDGEMENTS

My profound gratitude goes to...

❖ **Dr. Omer Waqar**

- Ph.D., P.Eng, Senior Member IEEE
- Assistant Professor, University of the Fraser Valley, Canada
- Adjunct Faculty Member at York University, Ontario, Canada
- (Past positions include) Assistant Professor, Thompson Rivers University, Canada

❖ **Dr. Muhammad Hanif**

- Ph.D., P.Eng, Senior Member IEEE
- Assistant Professor, Thompson Rivers University, Canada

1

Research Background and Significance

Introduction to foundational studies and the importance of this topic.

Introduction to DUL in 5G D2D Networks

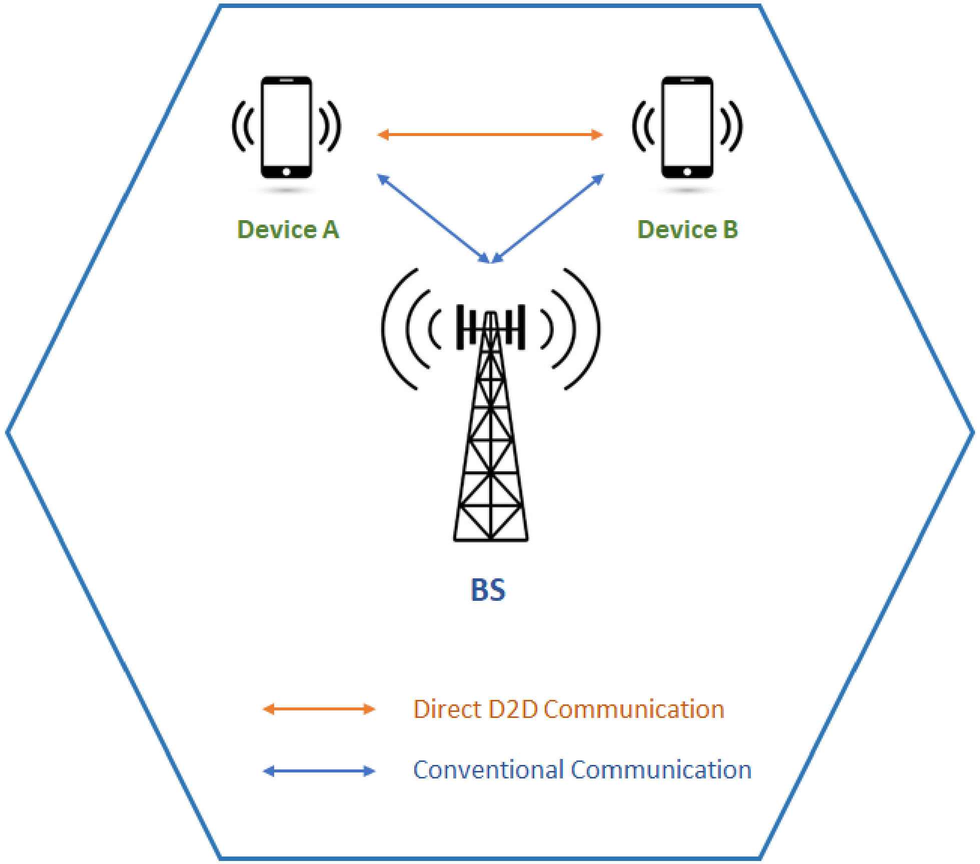


Figure 1.1: Direct Device-to-Device (D2D) communication between devices and conventional communications with a Base Station (BS)

5G The surge in mobile devices and data traffic demands solutions **beyond 5G** communication network management.

D2D Direct data exchanges among nearby mobile devices, without data relay via BSs, provide better performance.

DUL DUL uses data distribution to learn patterns without supervision, optimizing sum-rate in D2D networks.

DUL: *Deep Unsupervised Learning*

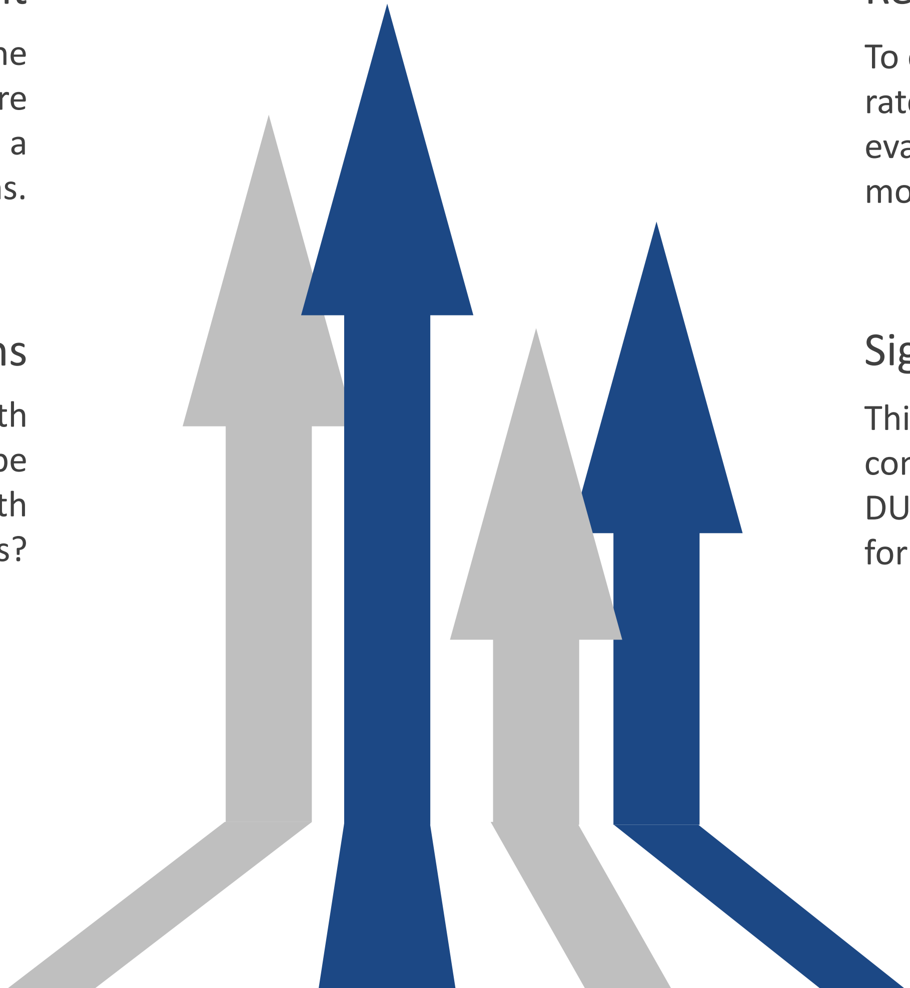
Purpose and importance of research

Problem Statement

Traditional methods for optimizing the sum-rate in D2D networks are computationally intensive. There's a need for adaptive, efficient solutions.

Research Questions

How can we improve the sum-rate with DUL? What DUL model should be proposed? How do we compare with existing DUL models?



Research Objectives

To explore a new DUL model for sum-rate optimization in D2D networks and evaluate its performance against existing models.

Significance of Study

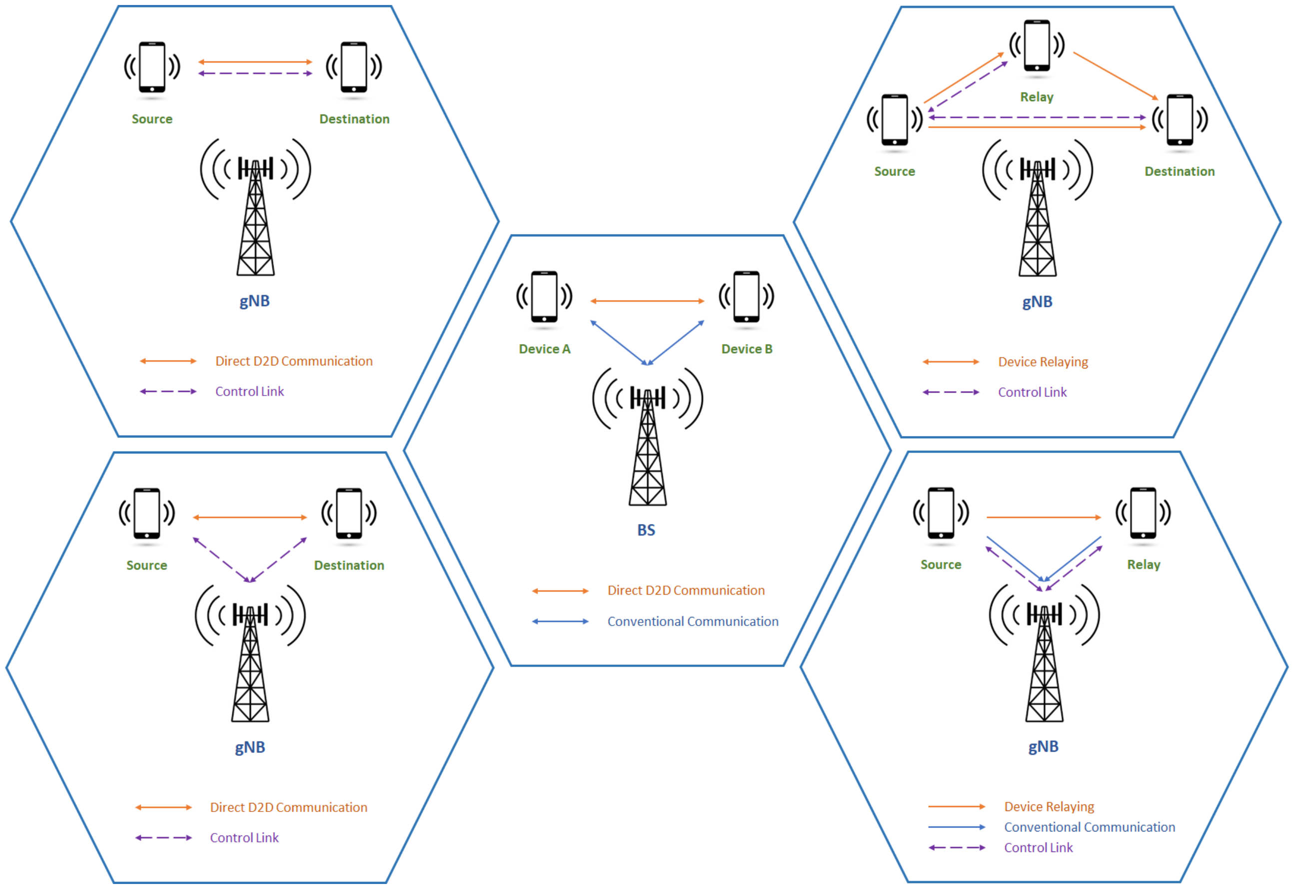
This research aims to advance wireless communication by developing adaptive DUL models for D2D networks, crucial for IoT and 5G/beyond networks.

2

Exploring the Landscape: An Overview

Evaluating current literature to build foundations and recognize gaps.

Cellular communication



Reuse time/frequency resource

leads to co-user interference, so it is important to optimize transmit power.

Cellular Telephone Specific Absorption Rate (SAR)

The FCC limit for public exposure from cellular telephones is an SAR level of 1.6 watts per kilogram

The system model

The discrete-time baseband signal received by the i-th receiver:

$$y_i = h_{i,i}x_i + \sum_{j \in K/\{i\}} h_{j,i}x_j + n_i \quad (3.1)$$

The achievable rate, i.e., Sum-Rate of the i-th receiver under Gaussian codebooks:

$$R_i(\mathbf{P}) = \log_2\left(1 + \frac{P_i|h_{i,i}|^2}{\sigma_i^2 + \sum_{j \in K/\{i\}} P_j|h_{j,i}|^2}\right) \quad (3.3)$$

The QoS, Signal-to-Interference plus Noise Ratio (SINR) for the i-th receiver:

$$SINR_i(\mathbf{P}) = \frac{P_i|h_{i,i}|^2}{\sigma_i^2 + \sum_{j \in K/\{i\}} P_j|h_{j,i}|^2} \quad (3.4)$$

The optimization problem for the i-th receiver of the system model:

$$\underset{\mathbf{P} \geq 0}{minimize} \sum_{i=1}^{K} \mathbf{P}_i$$

$$\text{subject to } SINR_i(\mathbf{P}) \geq r_{i,min}$$

Here, $r_{i,min}$ = the minimum required SINR of the i-th receiver

1964: Assignment of Transmitter Powers by Linear Programming, *F. Bock and B. Ebstein, in IEEE Transactions on Electromagnetic Compatibility, vol. 6, no. 2, pp. 36-44.* Used linear programming optimization theory.

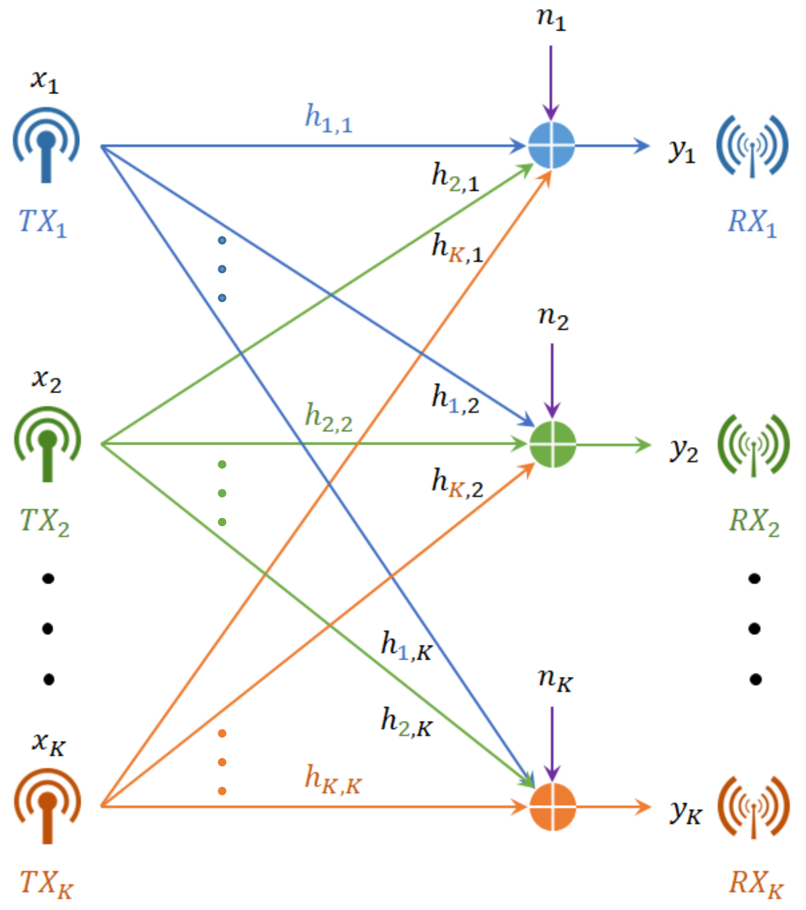


Figure 3.1: The K-user single-antenna interference channel model

What has happened since 1964?

- **Sum-Rate objective function with multiple and coupled constraints:**

The optimization problem for the i-th receiver of the system model:

$$\begin{aligned} & \underset{\mathbf{P}}{maximize} \sum_{i=1}^{K} R_i(\mathbf{P}) \\ & \text{subject to } SINR_i(\mathbf{P}) \geq r_{i,min} \\ & \text{and } (0 \leq P_i \leq P_{max}) \end{aligned} \tag{3.5}$$

Constraint elimination, from equation (3.5):

$$SINR_i(\mathbf{P}) \geq r_{i,min} \tag{3.6}$$

So, the equation in matrix form becomes:

$$\begin{pmatrix} |h_{1,1}|^2 & -r_{1,min}|h_{2,1}|^2 & \dots & -r_{1,min}|h_{K,1}|^2 \\ -r_{2,min}|h_{1,2}|^2 & |h_{2,2}|^2 & \dots & -r_{2,min}|h_{K,2}|^2 \\ \vdots & \vdots & \ddots & \vdots \\ -r_{K,min}|h_{1,K}|^2 & -r_{K,min}|h_{2,K}|^2 & \dots & |h_{K,K}|^2 \end{pmatrix}_{K\times K} \times \begin{pmatrix} P_1 \\ P_2 \\ \vdots \\ P_K \end{pmatrix}_{K\times 1} \geq \begin{pmatrix} r_{1,min}\times\sigma_1^2 \\ r_{2,min}\times\sigma_2^2 \\ \vdots \\ r_{K,min}\times\sigma_K^2 \end{pmatrix}_{K\times 1} \tag{3.9}$$

Equation (3.9) can be expressed as:

$$A_{K\times K} \times \mathbf{P}_{K\times 1} \geq \mathbf{b}_{K\times 1} \tag{3.10}$$

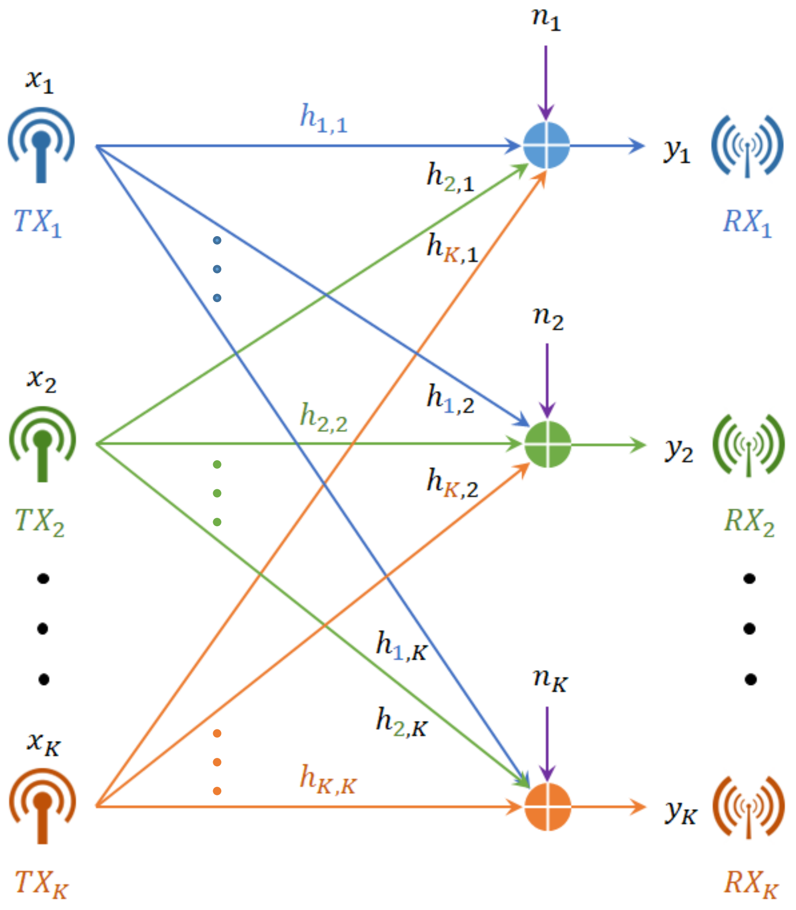


Figure 3.1: The K-user single-antenna interference channel model

What has happened since 1964?

March 2009

MAPEL: Achieving global optimality for a non-convex wireless power control problem

L. P. Qian, Y. J. Zhang and J. Huang, in IEEE Transactions on Wireless Communications, vol. 8, no. 3, pp. 1553-1563.

Introduced an algorithm that converges to an optimal solution for the WTM problem.

Identifies optimal power by refined approx. of the feasible SINR region.

The algorithm suffers from high computational complexity, which limits its practicality to small scenarios only.

October 2018

Learning to Optimize: Training Deep Neural Networks for Interference Management

H. Sun, X. Chen, Q. Shi, M. Hong, X. Fu and N. D. Sidiropoulos, in IEEE Transactions on Signal Processing, vol. 66, no. 20, pp. 5438-5453.

Applied a DNN to estimate the nonlinear input-output mapping for a interference management algorithm.

It was significantly faster than the leading algorithm used then.

This was an initial exploration into the potential of DNNs for specific problems.

March 2020

Towards Optimal Power Control via Ensembling Deep Neural Networks

F. Liang, C. Shen, W. Yu and F. Wu, in IEEE Transactions on Communications, vol. 68, no. 3, pp. 1760-1776.

A penalty term is included to ensure NN outputs meet D2D network rate constraints.

It offers flexibility in handling various types of constraints.

It provides a 'soft' boundary, requires tuning of a penalty factor, and often doesn't ensure optimality or feasibility.

December 2021

Multicell power control under rate constraints with deep learning

Y. Li, S. Han and C. Yang, in IEEE Transactions on Wireless Communications, vol. 20, no. 12, pp. 7813-7825.

A projection-based approach is employed to set up ‘hard’ boundaries.

This method meets the real-time requirements of online applications.

Using a projection necessitates an optimization solver, raising the computational complexity.

2021: DC3: A learning method for optimization with hard constraints, *P. L. Donti, D. Rolnick, and J. Z. Kolter, in International Conference Learning Representations.*

An approach different from projection, where an iterative gradient-descent-based Deep Constraint Completion and Correction (DC3) algorithm is proposed to improve solution feasibility.

While this approach guarantees linear constraint satisfaction, they employ an iterative process for satisfying the constraints during training and testing. This requirement contradicts the purpose of utilizing deep learning models as a substitute for iterative, non-data-driven optimization algorithms.

3

Proposed Methodology and Model

A comprehensive overview of the proposed strategy and DNN architecture.

Formulation for the optimization problem

For an optimization problem:

$$\begin{aligned} \text{P1}: \min_{\mathbf{x}} \quad & f(\mathbf{x}) \\ \text{subject to } & A\mathbf{x} \geq \mathbf{b} \\ \text{and } & \mathbf{0} \leq \mathbf{x} \leq \mathbf{c} \end{aligned} \tag{3.12}$$

The non-homogeneous constraint given in Equation (3.12) can be rewritten as:

$$A\mathbf{x} = \mathbf{b} + \mu \tag{3.13}$$

Here $\mu \geq \mathbf{0}$ is a K-dimensional real vector. As a consequence,

$$\mathbf{x} = \widehat{\mathbf{x}} + A^{-1}\mu \tag{3.14}$$

Where,

$$\widehat{\mathbf{x}} = A^{-1}\mathbf{b} \tag{3.15}$$

(3.12) is always satisfied implicitly, it can be expressed in terms of μ as:

$$A^{-1}\mu \leq \mathbf{c} - \widehat{\mathbf{x}} \tag{3.16}$$

Consequently, P1 can be reformulated in terms of μ as follows:

$$\begin{aligned} \text{P2}: \min_{\mu} \quad & f(\widehat{\mathbf{x}} + A^{-1}\mu) \\ \text{subject to } & \mu \geq \mathbf{0} \\ \text{and } & A^{-1}\mu \leq \mathbf{c} - \widehat{\mathbf{x}} \end{aligned} \tag{3.17}$$

$A^{-1} = \begin{bmatrix} \mathbf{a}_1 \; \mathbf{a}_2 \; \cdots \; \mathbf{a}_K \end{bmatrix}$, where $\mathbf{a}_k$ is the k-th column of A^{-1}, and $k = 1, 2, \cdots, K$. Denoting the k-th element of μ by μ_k, Equation (3.17) can be re-written as:

$$\sum_{k=1}^{K} \mu_k \mathbf{a}_k \leq \mathbf{c} - \widehat{\mathbf{x}} \tag{3.18}$$

Equivalently,

$$\mu_k \mathbf{a}_k \leq \mathbf{c} - \widehat{\mathbf{x}} - \sum_{l=1, l\neq k}^{K} \mu_l \mathbf{a}_l \tag{3.19}$$

Noting that $\mu_k \mathbf{a}_k \geq 0$ for $1 \leq k \leq K$, it comes:

$$\mu_k \mathbf{a}_k \leq \mathbf{c} - \widehat{\mathbf{x}} \tag{3.20}$$

and

$$0 \leq \mu_k \leq \beta_k \triangleq \min_{l} \frac{c_l - \widehat{x}_l}{a_{k,l}} \tag{3.21}$$

Where c_l, $\widehat{x}_l$, and $a_{k,l}$ are the l-th elements of $\mathbf{c}$, $\widehat{\mathbf{x}}$, and $\mathbf{a}_k$, respectively, and $k, l = 1, 2, \cdots, K$.

Formulation for the optimization problem

Since Equation (3.21) does not imply Equation (3.17) in all cases, it is proposed to scale μ by $\alpha \geq 0$, if Equation (3.17) is not satisfied. To this end, it is defined $\nu = \alpha\mu$, where $\alpha = 1$ if Equation (3.17) is satisfied. Otherwise, α is computed using Equation (3.17) as:

$$\alpha = \begin{cases} 1 & \text{if Equation (3.17) is satisfied;} \\ \min\limits_{l} \dfrac{c_l - \widehat{x}_l}{[A^{-1}\mu]_l} & \text{otherwise;} \end{cases} \tag{3.22}$$

and $[A^{-1}\mu]_l$ is the l-th element of $A^{-1}\mu$ with $l = 1, 2, \cdots, K$. By employing Equation (3.22) it is ensured that $A^{-1}\nu \leq \mathbf{c} - \widehat{\mathbf{x}}$, thus 100% constraint satisfaction is ensured. For noting that $\min\limits_{l} \dfrac{c_l - \widehat{x}_l}{[A^{-1}\mu]_l}$ cannot be less than unity when Equation (3.17) is satisfied, it can be re-writen α in a more compact form as:

$$\alpha = \min\left\{1, \min_{l} \frac{c_l - \widehat{x}_l}{[A^{-1}\mu]_l}\right\} \tag{3.23}$$

Regarding the Equation (3.5), α can be expressed as:

$$\alpha = \min\left\{1, \min_{1 \leq l \leq K} \frac{P_{max} - \widehat{P}_l}{[A^{-1}\mu]_l}\right\} \tag{3.24}$$

Proposed DNN model

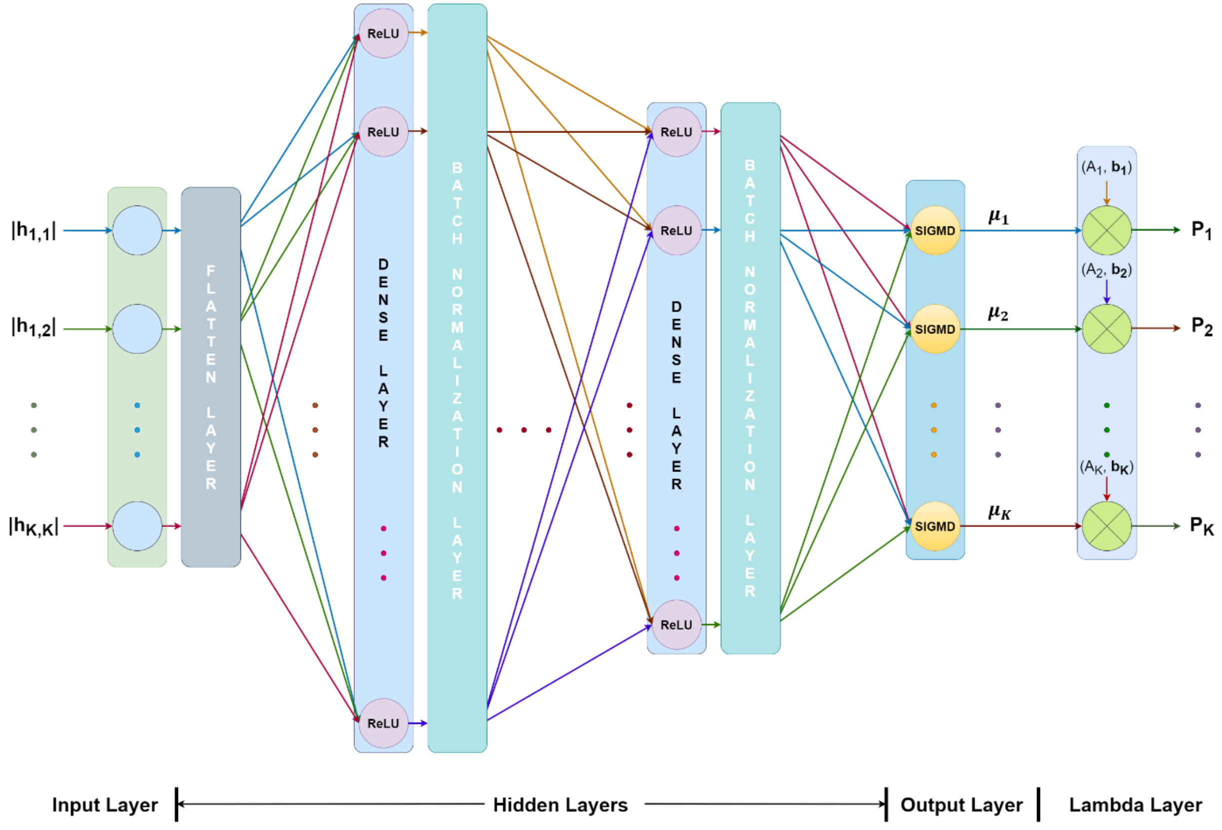


Figure 3.2: The architecture of the proposed DNN model

DNN Architecture

- Inputs are K^2 feasible transmission channel parameters in terms of $|h_{ij}|$ magnitudes.
- Then densely interconnected hidden layers with the ReLU activation function are followed by a batch normalization layer.
- K number of outputs for slack variable μ from the output layer with the sigmoid function.
- Finally, there is a lambda layer to calculate the power profile **P** that satisfies both the SINR and power constraints of the optimization problem.

The loss function is $f(\widehat{\mathbf{P}} + A^{-1}\nu)$, where $\nu = \alpha\mu$, and α is given in Equation (3.22).

$$\mathcal{L}\text{oss}_{\mathbf{DNN}} = \frac{1}{|\Psi|}\sum_{\psi\epsilon\Psi} -R_{\psi}\left(\mathbf{P}\right) \tag{3.28}$$

Here, ψ represents the mini-batch of size $|\Psi|$.

4

Results and Discussions

Key findings, interpretations, and scopes for further research.

Benchmark metrices, model parameters, and datasets

Benchmark scheme

- Power Control Network (PCNet and PCNet+)

Evaluation metrices

- Constraint Violations Probability or Hit Rate
- Average Sum-Rate in Bit/Second/Hertz

DNN model parameters

- 2 dense layers with $2K^2$ and K^2 neurons respectively
- ADAM optimizer
- Learning rate = 0.0001
- Mini-batch size = 1,000
- Epoch = 50
- P_{max} = 1.0 Watt

Datasets

- 250,000 feasible channel parameters, each 5×5 matrix for K = 5, a total of 8,750,000 (= $5 \times 7 \times 1 \times$ 250,000).
- 250,000 feasible channel parameters for each K = 5, 6, 7 & 8, a total of 5,000,000 (= $1 \times 5 \times 4 \times$ 250,000).
- An 80% : 10% : 10% split for the training, validation, and test dataset

Results for training with a given background noise power

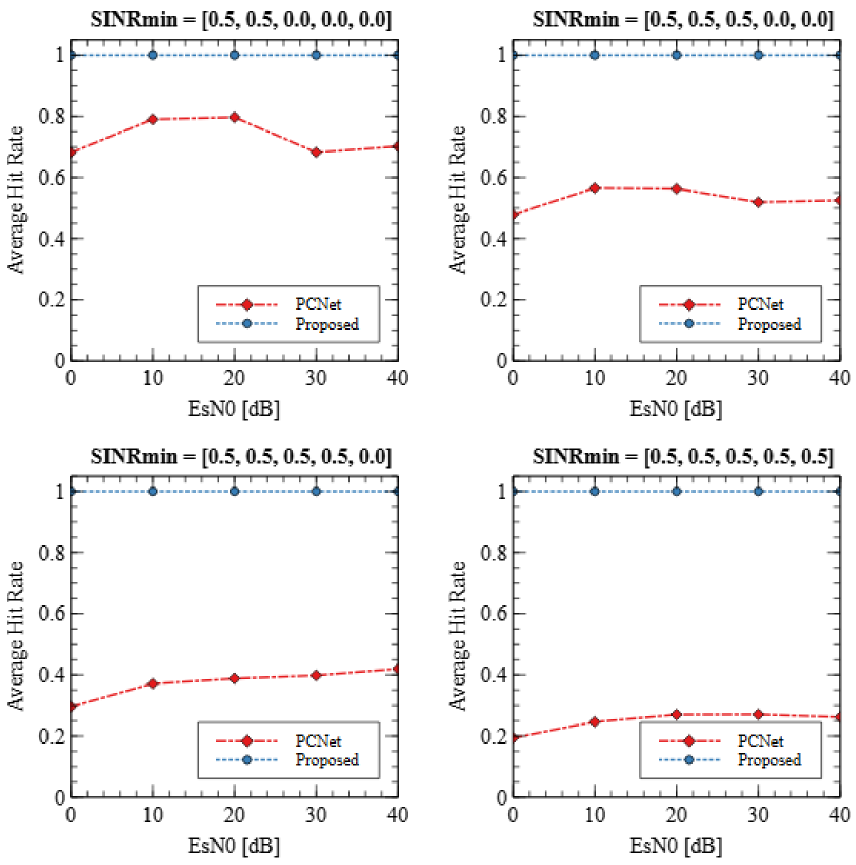


Figure A: Hit rate comparison for four QoS requirements when K = 5

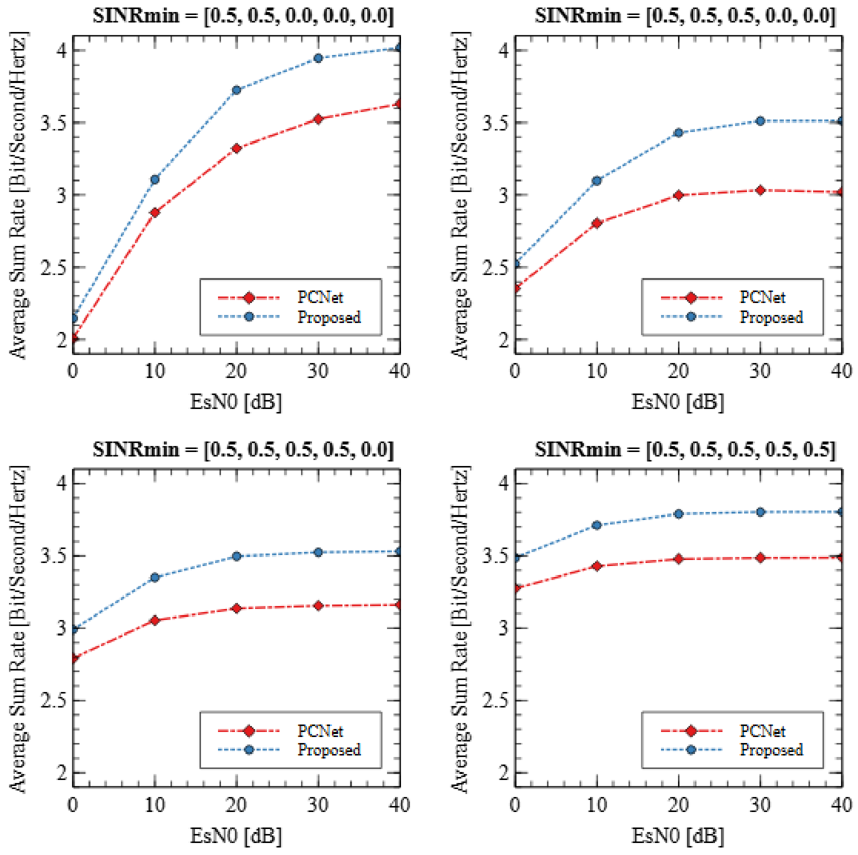


Figure B: Average sum rate comparison for four QoS requirements when K = 5

Note 1. For PCNet, the best values are at λ = 5, which are taken to be shown here.

Note 2. the noise level $EsN0$ as:

$$EsN0 = 10\log_{10}\frac{\mathbf{P}_{\max}}{\sigma^2} \qquad (4.1)$$

Results for enhanced generalization capacity

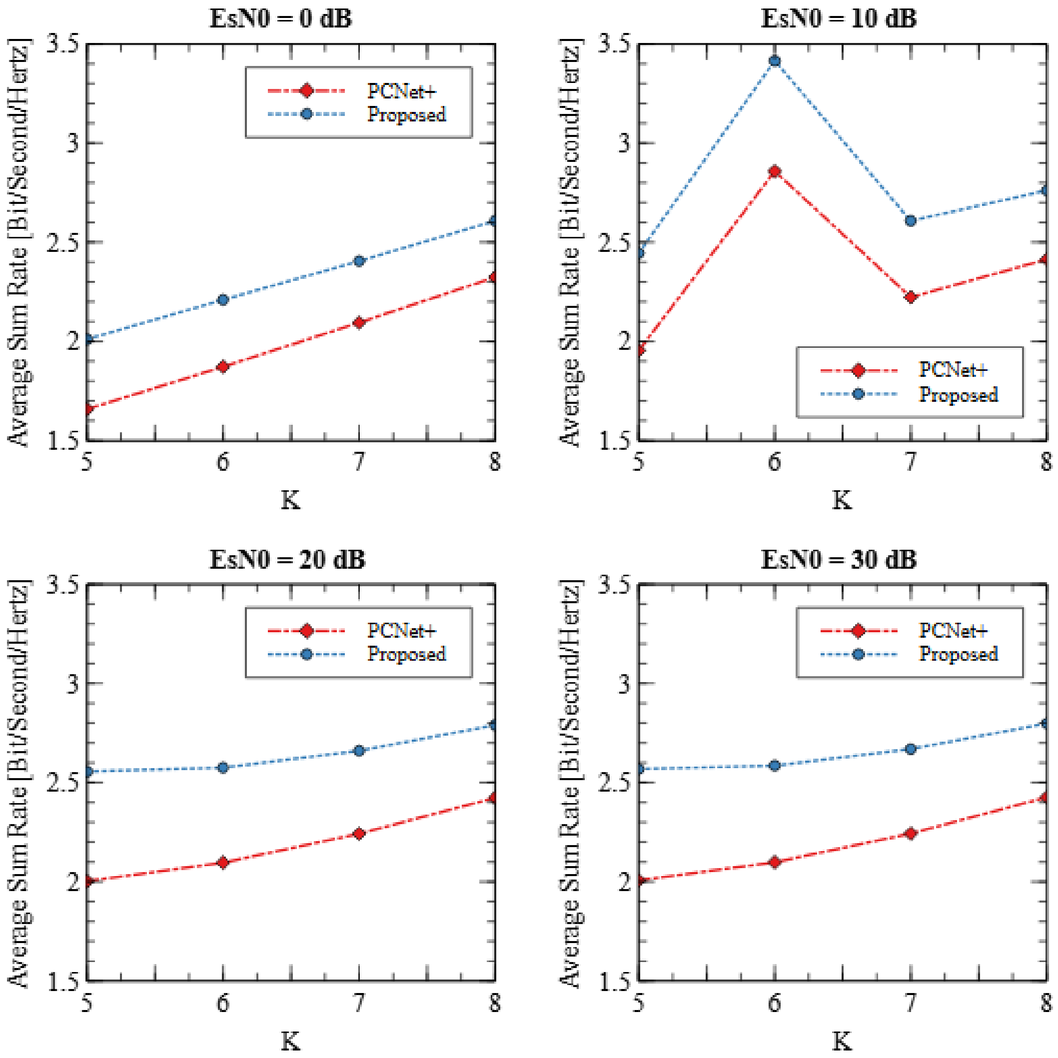


Figure C: Average sum rate comparison against the number of receivers, K, for four EsN0 values

Note 1. For PCNet+, the best values are at λ = 5, which are taken to be shown here.

Note 2. $SINR_{min}$ = 0.2 for all receivers. For example, for K = 5, $SINR_{min}$ = [0.2, 0.2, 0.2, 0.2, 0.2]

Key findings

Constraint Violations

The proposed model ensures that all the constraints are satisfied whenever the problem is feasible.

Average Sum-Rate

It outperforms the benchmark scheme for all QoS requirements and for the complete EsN0 range.

Computational Complexity

Unlike PCNet/PCNet+, the proposed model ensures solution feasibility without requiring any tuning of additional hyperparameters (λ).

Training Time

For example, the values are (46.408s vs. 49.717s) and (346.42s vs 401.94s) for PCNet/PCNet+ and the proposed scheme, respectively.

Limitations and future research

Ideal channel estimation

Should focus on enhancing the model's robustness against potential errors in channel estimation.

Centralized scheme

Developing a distributed version of this scheme could make it more adaptable to diverse network conditions and requirements.

Selecting hyperparameters

Developing strategies for hyperparameter selection that can optimize the model's performance.

Integrating other parameters

Developing a model that can generalize other system parameters like the number of users and the distribution of channel coefficients effectively and accurately is a crucial research direction.

PUBLICATION

Journals & Magazines > IEEE Wireless Communications ... > Early Access

Deep Unsupervised Learning for Optimization With Box and Monotone-Matrix Based Polytope Constraints: A Case-Study of D2D Wireless Networks

Publisher: IEEE Cite This PDF

Bindubritta Acharjee ; Muhammad Hanif ; Omer Waqar **All Authors**

Impact Factor = 6.3

Thank you for your time.