



# Optimizing Power Allocation for D2D Communication with URLLC under Rician Fading Channel: A Learning-to-Optimize Approach

Owais Muhammad<sup>1</sup>, Hong Jiang<sup>1,\*</sup>, Mushtaq Muhammad Umer<sup>1</sup>, Bilal Muhammad<sup>2</sup> and Naeem Muhammad Ahtsam<sup>3</sup>

<sup>1</sup>School of Information Engineering, Southwest University of Science and Technology, Mianyang, 621010, China

<sup>2</sup>School of Software Engineering, Northeastern University, Shenyang, 110167, China

<sup>3</sup>School of Information and Software Engineering, University of Electronic Sciences and Technology, Chengdu, 610000, China

\*Corresponding Author: Hong Jiang. Email: jianghong@swust.edu.cn

Received: 15 April 2023; Accepted: 21 June 2023; Published: 11 September 2023

**Abstract:** To meet the high-performance requirements of fifth-generation (5G) and sixth-generation (6G) wireless networks, in particular, ultra-reliable and low-latency communication (URLLC) is considered to be one of the most important communication scenarios in a wireless network. In this paper, we consider the effects of the Rician fading channel on the performance of cooperative device-to-device (D2D) communication with URLLC. For better performance, we maximize and examine the system's minimal rate of D2D communication. Due to the interference in D2D communication, the problem of maximizing the minimum rate becomes non-convex and difficult to solve. To solve this problem, a learning-to-optimize-based algorithm is proposed to find the optimal power allocation. The conventional branch and bound (BB) algorithm are used to learn the optimal pruning policy with supervised learning. Ensemble learning is used to train the multiple classifiers. To address the imbalanced problem, we used the supervised undersampling technique. Comparisons are made with the conventional BB algorithm and the heuristic algorithm. The outcome of the simulation demonstrates a notable performance improvement in power consumption. The proposed algorithm has significantly low computational complexity and runs faster as compared to the conventional BB algorithm and a heuristic algorithm.

**Keywords:** D2D; URLLC; rician fading; supervised learning

## 1 Introduction

URLLC is one of the most important scenarios in 5G, whose goal is to make it possible for new services and applications to have high reliability, availability, and minimal latency [1]. D2D communication is a promising solution with URLLC, adopted as a vital communication scenario in 5G mobile communication networks, and it is becoming increasingly important to offer end-to-



This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

end services with low latency and high reliability [2]. URLLC has very stricter requirements, such as 99.999% reliability (i.e.,  $10^{-5}$  packet error probability) under the end-to-end latency of 1 ms. URLLC has become a primary goal for several applications, including unmanned aerial vehicles, intelligent transportation systems, industrial automation, vehicle-to-vehicle communication, and tactile internet [3,4].

To enhance power allocation, traditional wireless networks have been developed with long-packet transmission scenarios, where ensuring high reliability and low latency at the same time is generally difficult [3]. D2D communication is utilized in various scenarios to enhance the Quality of Service (QoS) of the network by reducing power consumption, lowering latency, and improving reliability, resulting in a significant improvement in performance [5,6]. Moreover, resource allocation and power allocation are major issues for D2D communication underlying cellular networks. D2D communication can decrease latency and enhance network capacity, thereby optimizing resource allocation. D2D communications reduce overall power usage due to users' proximity, which is impossible in conventional cellular communications [7]. D2D communications are emerging as a potential technique for meeting the URLLC's strict standards [8,9].

In wireless communication, resource allocation management affects the performance of the optimization. Machine learning has demonstrated a high degree of efficacy in resolving non-convex optimization issues related to D2D communication, such as power allocation [10] and interference management [11], that affect communication performance. Global optimization algorithms, such as the BB algorithm, have exponential complexity, and the majority of recent studies have concentrated on heuristic or suboptimal algorithms. Machine learning is a relatively new technique for balancing performance gaps and computational complexity that has proven effective in addressing difficult optimization problems, particularly those that are non-convex and NP-hard [12,13]. This field of research is referred to as learning-to-optimize for resource allocation in wireless communication in order to address wireless network optimization problems [14]. Learning-to-optimization aims to minimize the computational complexity and resources needed to obtain solutions that are nearly optimal [12].

## 2 Related Work

Some studies examine the D2D architecture with URLLC, where they attempt to preserve URLLC's strict QoS requirements. Chang et al. [15] propose an autonomous probability-based D2D transmission method under the Rayleigh fading channel in URLLC to minimize the transmission power. A frequency-division duplex system is used in the uplink and downlink spectrums of D2D communication under cellular networks with URLLC requirements for optimized resource allocation. An unbalanced distribution of data traffic between both frequency bands could lead to less effective use of network resources [8]. For improving the transmission power for a finite block-length rate with the discrete-time block-fading channel, a resource allocation algorithm is proposed to maximize the achievable rate and optimize the power allocation for D2D communication with URLLC requirements [3]. Similarly, in [16], the proposed real-time wireless control systems for D2D communication were primarily focused on transmission power optimization within the confines of URLLC, and a Rayleigh fading channel is used, which decreases power consumption but is not as effective as Rician fading. The authors in [17] formulated an optimal power allocation problem in an uplink D2D communication scenario under a cellular network with Rayleigh fading, maximized the overall system throughput of the communication system, and ensured the URLLC requirements.

Numerous studies have been done on optimization-based algorithms for wireless resource allocation problems [18]. The majority of wireless resource allocation problems are non-convex or NP-hard problems, and optimization algorithms are known for being either excessively time-consuming or exhibiting significant performance gaps [19]. In order to achieve efficient and nearly optimal resource allocation, learning-to-optimize is a disruptive technique for solving resource allocation problems. Learning-to-optimize combines knowledge and techniques from wireless communications, mathematical optimization, and machine learning. It aims to develop algorithms that can dynamically and efficiently solve optimization problems while effectively balancing the computational complexity and the optimality gap in optimization problems. Policy learning is one of the major subcategory of learning-to-optimization; training an agent to discover the best solution to a problem within a specified algorithm is known as policy learning [20]. In order to solve the mixed integer nonlinear programming (MINLP) problem for resource allocation in D2D communication, the authors in [21] propose a policy learning method for node pruning in the BB algorithm.

This paper investigates uplink D2D communication underlying the cellular system in a single-cell environment with a Rician fading channel. URLLC is utilized in D2D communication to deal with the extremely high QoS requirements for D2D communication. In order to maintain satisfactory performance for all users, we focus on maximizing and analyzing the minimum rate of D2D users. However, the theoretical analysis that corresponds to this idea is complex because of the complicated interference patterns of the system under consideration. Finding a solution to the problem of maximizing the minimal rate and analyzing the interference effects on the maximized minimal rate is challenging. In this paper, we use a policy learning approach to solve the power allocation problem with low computational complexity in D2D communication. The following are the contributions of this research:

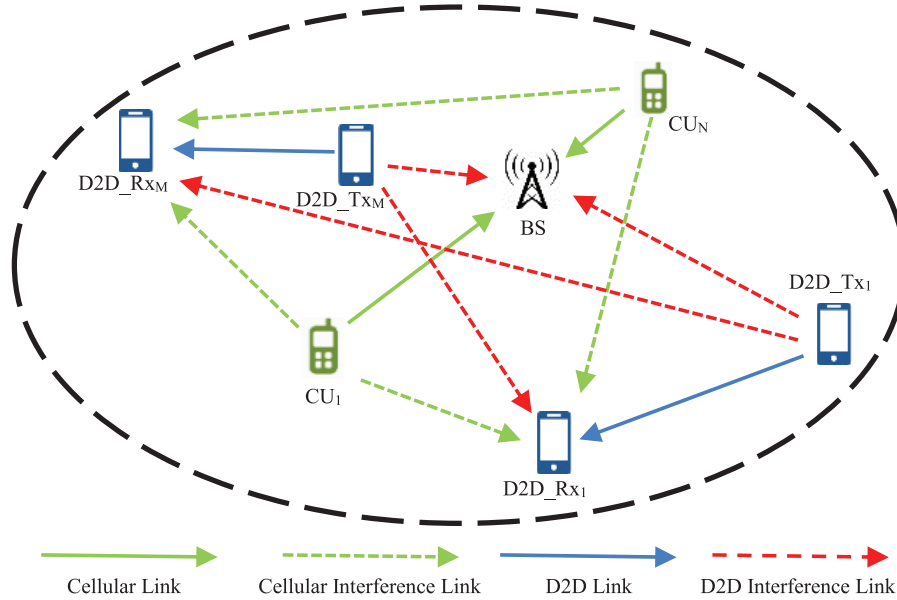
- A framework has been developed to analyze D2D communication in a cellular network. This framework utilizes the Rician fading channel to enable D2D communication that meets URLLC QoS requirements. We maximize the minimal rate of D2D users. The formulated problem is a non-convex problem with a complex expression of the achievable rate, which is solved by using the proposed scheme.
- We propose a fast iterative learning-to-optimize-based algorithm to maximize the minimal achievable rate  $R_m^*$ , by searching for optimal power allocation with significantly low computational complexity.
- The conventional BB algorithm is used to generate the training sample sets and learn the optimal pruning policy for power allocation in D2D communication. To address the imbalanced problem, we used an undersampling technique, and ensemble learning is used with supervised learning, which involves training multiple classifiers and combining their outputs to enhance overall performance. The computational complexity of the proposed algorithm is significantly less than that of the BB algorithm and heuristic algorithm.

The rest of the paper is structured in the following manner: [Section 3](#) provides a brief introduction to the system model. [Section 4](#) formulates the problem, and [Section 5](#) presents the learning-to-optimize-based algorithm to solve the problem. [Section 6](#) of the paper contains the simulation and numerical results, while the conclusion can be found in [Section 7](#).

*Notation:* The following notations will be used throughout the paper:  $\Pr \{ \cdot \}$  denotes the probability and  $E [ \cdot ]$  denotes the expectation of a random variable.  $I_0 (\cdot)$  is the zeroth-order modified Bessel function of the first kind.

### 3 System Model

In this section, we consider uplink D2D communication within a cellular system in a single-cell network, as depicted in Fig. 1. We consider  $N$  cellular users (CUs) and  $M$  D2D pairs where  $M \leq N$ . The cellular user set is  $U = \{1, 2, \dots, N\}$  and the D2D pairs set is  $L = \{1, 2, \dots, M\}$ . We assume that D2D pairs reuse the uplink channels of cellular users to transmit data. The resource reuse indicator is  $p_{n,m}$ , if D2D pair  $m$  reuses cellular user  $n$  channel, then  $p_{n,m} = 1$  otherwise  $p_{n,m} = 0$  [22,23].



**Figure 1:** System model

We can suppose that the slow fading is comprised of known constants and the fast fading is comprised of random variables where  $h_m^D$  and  $h_{n,m}^{CD}$  are fast-fading coefficients,  $h_{n,m}^{CD}$  is interference from other D2D links are distributed exponentially with a zero mean and a unit variance. The coefficients of slow fading from the D2D transmitter (D2D-Tx)  $m$  to the receiver is  $g_m^D$  and interference channel power gain between the CUE  $n$  to D2D receiver (D2D-Rx)  $m$  is  $g_{n,m}^{CD}$  [24–26]. D2D users get interference from the cellular users and the D2D pairs because they use the same spectrum resources. The signal-to-interference-plus-noise ratios (SINR) of the  $m^{\text{th}}$  D2D link can be formulated as

$$s_m^D = \frac{P_m^D g_m^D h_m^D}{\sum_{n \in U} p_{n,m} P_n^C g_{n,m}^{CD} h_{n,m}^{CD} + \sigma^2} \quad (1)$$

where  $P_n^C$  and  $P_m^D$  is the transmit power of cellular link  $n$  and transmit power of D2D link  $m$ .  $\sigma^2$  is the additive noise power.

#### Outage Probability

The outage probability of D2D links is analytically expressed as

$$P_{out}^{D2D} = Pr \left[ s_m^D \leq s_0 = \frac{s_D}{s_I} \leq s_0 \text{ or } s_D \leq s_{th} \right] \quad (2)$$

where  $s_0$  is the threshold of minimum SINR and  $s_{th}$  is the threshold of minimum instantaneous signal power.  $s_D$  is the signal power, i.e.,  $s_D = P_m^D g_m^D h_m^D$  and the total interference power is  $s_I$ , i.e.,

$s_I = \sum_{n \in U} p_{n,m} P_n^C g_{n,m}^{CD} h_{n,m}^{CD}$ .  $s_I$  represent the cross-tire and co-tire interference because of interference from cellular users and interference from the D2D transmitter to other D2D pairs used. The sum of the instantaneous powers of all active sources of interference, represented by  $W_I$ , equals the total instantaneous interference power, denoted as  $s_I$ , i.e.,  $s_I = \sum_{i=1}^{W_I} s_i$  [27,28]. According to Eq. (2), the D2D outage probability can be expressed as

$$P_{out}^{D2D} = \{s_m^D \leq s_0\} = Pr \left( \frac{P_m^D g_m^D h_m^D}{\sum_{n \in U} p_{n,m} P_n^C g_{n,m}^{CD} h_{n,m}^{CD}} \leq s_0 \right) \quad (3)$$

From the CUE to the D2D receiver, the interference power is usually substantially greater than the noise power [25]. We make the assumption that D2D links are limited by interference and that the impact of noise power on the outage probability can be ignored [7]. Eq. (3) can be expressed in the probability density function (PDF) as

$$P_{out}^{D2D} = 1 - \int_{s_{th}}^{\infty} \left( \int_0^{\frac{s_D}{s_0}} f_{s_I}(s_I) ds_I \right) f_{s_D}(s_D) ds_D \quad (4)$$

The instantaneous signal power PDF is  $f_{s_D}(s_D)$  and  $f_{s_I}(s_I)$  is the total interference power [27]. The signal received from the desired user is Rician distributed, and there are  $W_I$  independent identically distributed (i.i.d.) Rayleigh's interference in the system. The PDF of total instantaneous interference power  $s_I$  express as

$$f_{s_I}(s_I) = \frac{s_I^{W_I-1}}{\bar{s}_I^{W_I} (W_I - 1)!} \exp \left( -\frac{s_I}{\bar{s}_I} \right) \quad (5)$$

$\bar{s}_I$  is the statistical average of the  $s_I$  and the PDF of instantaneous signal power  $s_D$  express as

$$f_{s_D}(s_D) = \frac{(K+1)}{\bar{s}_D} e^{\left[ -K - \frac{(K+1)s_D}{\bar{s}_D} \right]} I_0 \left( 2\sqrt{\frac{K(K+1)s_D}{\bar{s}_D}} \right) \quad (6)$$

where  $K$  is the Rician fading parameter,  $I_0(\cdot)$  is the first kind of modified Bessel function with zeroth-order and  $\bar{s}_D$  is statistically average of the instantaneous signal power [25]. The outage probability of D2D can be rewritten by using Eqs. (5) and (6) in Eq. (4).

$$P_{out}^{D2D} = 1 - \int_{s_{th}}^{\infty} \left( \int_0^{\frac{s_D}{s_0}} \frac{s_I^{W_I-1}}{\bar{s}_I^{W_I} (W_I - 1)!} \exp \left( -\frac{s_I}{\bar{s}_I} \right) ds_I \right) \left( \frac{(K+1)}{\bar{s}_D} e^{\left[ -K - \frac{(K+1)s_D}{\bar{s}_D} \right]} I_0 \left( 2\sqrt{\frac{K(K+1)s_D}{\bar{s}_D}} \right) \right) ds_D \quad (7)$$

Solving the inner integral of Eq. (7) and getting the outage probability as

$$P_{out}^{D2D} = \frac{\exp \left[ -K + \frac{K}{\left( 1 + \frac{\bar{s}_D}{(K+1)\bar{s}_I} \right)} \right]}{\left( 1 + \frac{\bar{s}_D}{(K+1)s_0\bar{s}_I} \right)} \quad (8)$$

*Proof:* See [Appendix A](#).

In URLLC scenarios, we anticipate users to send short packets to achieve low latency, and the successful packet error probability is used as a measure of high reliability. The uplink channel capacity expression for D2D communication is as follows:

$$R_m = \frac{B_m}{\ln 2} \left[ C_m - \sqrt{\frac{V_m}{T_m B_m}} fQ^{-1}(\varepsilon_m) \right] \quad (9)$$

where  $C_m$  denotes the Shannon capacity and  $V_m$  is the channel dispersion.  $T_m$  is the transmission time delay,  $B_m$  is the bandwidth,  $\varepsilon_m$  transmission error probability (i.e., packet error probability), and  $fQ^{-1}(\cdot)$  is the inverse of the  $Q$ -function [14,29]. The Shannon capacity can be expressed as the following based on the received SINR:

$$C_m = \log(1 + s_m^D) \quad (10)$$

The capacity loss resulting from transmission errors is represented by channel dispersion  $V_m$  [30], which can be represented as

$$V_m = \left( 1 - \frac{1}{(1 + s_m^D)^2} \right) \approx 1 \quad (11)$$

When SINR is higher than 5 dB then  $V_m \approx 1$  [31]. We analyze the ergodic capacity of  $C_m$ , ergodic capacity is obtained by experiencing all the channel fading states, which are expressed as

$$E[\log(1 + s_m^D)] = \int_0^\infty \log(1 + z) f_{s_m^D}(z) dz \quad (12)$$

where the expectation  $E[\cdot]$  is taken over the fast-fading distribution. The following theorem presents the expression to compute the ergodic capacity.

**Theorem 1:** In D2D communication, the ergodic capacity is given by

$$C_m^* = \int_0^\infty \frac{1 - F_{s_m^D}(z)}{1 + z} dz \quad (13)$$

where  $F_{s_m^D}(z) = Pr(s_m^D \leq s_0)$  is given in [Eq. \(8\)](#)

*Proof:* See [Appendix B](#).

Then [Eq. \(9\)](#) can be rewritten as

$$R_m = \frac{B_m}{\ln 2} \left[ C_m^* - \sqrt{\frac{V_m}{T_m B_m}} fQ^{-1}(\varepsilon_m) \right] \quad (14)$$

The probability of a packet error during the transmission of  $b_m$  bits from the sender to receiver  $m$  within the transmission duration  $T_m$  can be formally expressed as

$$\varepsilon_m^* = fQ \left\{ \sqrt{\frac{T_m B_m}{V_m}} \left[ C_m^* - \frac{b_m \ln 2}{T_m B_m} \right] \right\} \quad (15)$$

where  $b_m = R_m^* T_m$ , is the number of bits to be transmitted in each transmission. In order to ensure the reliability requirement of URLLC, the following constraint must be satisfied.

$$\varepsilon_m^* \leq \varepsilon_{max} \quad (16)$$

$\varepsilon_{max}$  is the maximum packet error probability bounded by the URLLC QoS requirements. The successful transmission can be expressed as  $1 - \varepsilon_m^* \geq 1 - \varepsilon_{max}$ , i.e., each packet should be successfully delivered with a probability greater than  $1 - \varepsilon_{max}$ . The primary communication constraint  $\sum_{m=1}^L P_m^D \leq P_{max}^D$  is restricted wireless resources, and the URLLC QoS requirement includes a limitation on the time delay for communication as  $T_m \leq T_{max}$ .

#### 4 Problem Formulation

In this section, we formulate the optimization problem to maximize the minimal achievable rate among the D2D users where the power allocation is optimized, which is described as:

$$P1 : \max_{P_m^D} \min_m \{R_m^*\} \quad (17a)$$

$$\text{s.t. } \varepsilon_m^* \leq \varepsilon_{max}, m = 1, 2, \dots, L \quad (17a)$$

$$T_m \leq T_{max}, m = 1, 2, \dots, L \quad (17b)$$

$$\sum_{m=1}^L P_m^D \leq P_{max}^D, m = 1, 2, \dots, L \quad (17c)$$

$$Pr \{s_m^D \leq s_0\} \leq \gamma_0, m = 1, 2, \dots, L \quad (17d)$$

The objective of this optimization problem is to maximize the minimal achievable rate  $R_m^*$ , by searching for overall optimal transmission power  $P_m^D$ . The first constraint in Eq. (17a) is used to ensure the reliability of D2D users. The constraint in Eq. (17b) is used to ensure the transmission time delay (the delay cannot exceed the maximum transmission time delay  $T_{max}$ ). Eq. (17c) is the transmission power constraint, and there are limitations on the total transmission power,  $P_{max}^D$  is the maximum transmission power, and in Eq. (17d)  $\gamma_0$  is the maximum outage probability constraint.

The aforementioned problem  $P1$  is difficult because, as demonstrated by Eqs. (1) and (14), the achievable rate  $R_m^*$  is not convex due to interference, which is challenging to solve for power allocation. To overcome this challenge, we use the bisection-based method, and to handle the non-convex objective function in  $P1$ , an auxiliary variable  $t_0$  is introduced to simplify the objective function and formulate a new problem  $\hat{P}1$  from power allocation problem  $P1$ :

$$\hat{P}1 : \max_{P_m^D} t_0 \quad (18a)$$

$$\text{s.t. } \varepsilon_m^* \leq \varepsilon_{max}, m = 1, 2, \dots, L \quad (18a)$$

$$T_m \leq T_{max}, m = 1, 2, \dots, L \quad (18b)$$

$$P_m^D \geq 0, m = 1, 2, \dots, L \quad (18c)$$

$$\sum_{m=1}^L P_m^D \leq P_{max}^D, m = 1, 2, \dots, L \quad (18d)$$

$$R_m^* \geq t_0, m = 1, 2, \dots, L \quad (18e)$$



$$Pr \{s_m^D \leq s_0\} \leq \gamma_0, m = 1, 2, \dots, L \quad (18f)$$

To solve the problem  $\hat{P}1$  and  $t_0$ , can be accomplished by the max-min rate. Where (18c) assures that non-negative power is assigned to each user, and (18d) establishes the maximum transmission power  $P_{max}^D$ . We can assign the value of  $t_0$  to a certain  $t_1$  and determine whether the max-min rate can accomplish  $t_1$ . We find the optimal auxiliary variable  $t_1$  from  $\hat{P}1$  and use in  $P1'$ . All D2D users' achievable rates are greater than  $t_1$ , and the total power consumption  $\sum_{m=1}^L P_m^D$  is examined to determine whether it is less than the maximum transmit power  $P_{max}^D$ .  $t_1$  can only be achieved if the minimized power consumption is less than  $P_{max}^D$ ; otherwise, it cannot be achieved and formulate a subsequent problem  $P1'$ :

$$P1' : \min_{P_m^D} \sum_{m=1}^L P_m^D$$

$$\text{s.t. } \varepsilon_m^* \leq \varepsilon_{max}, m = 1, 2, \dots, L \quad (19a)$$

$$T_m \leq T_{max}, m = 1, 2, \dots, L \quad (19b)$$

$$P_m^D \geq 0, m = 1, 2, \dots, L \quad (19c)$$

$$\sum_{m=1}^L P_m^D \leq P_{max}^D, m = 1, 2, \dots, L \quad (19d)$$

$$R_m^* \geq t_1, m = 1, 2, \dots, L \quad (19e)$$

$$Pr \{s_m^D \leq s_0\} \leq \gamma_0, m = 1, 2, \dots, L \quad (19f)$$

According to this significant finding, by solving the following set of equations, the optimal solution of  $P1'$  can be achieved:

$$R_m = t_1$$

$$\Leftrightarrow \frac{B_m}{\ln 2} \left[ C_m - \sqrt{\frac{V_m}{T_m B_m}} fQ^{-1}(\varepsilon_m^*) \right] = t_1$$

$$\Leftrightarrow P_m^D = P_{out}^{D2D} \left\{ \exp \left[ \frac{T_m \ln 2 (t_1)}{T_m B_m} + \sqrt{\frac{1}{T_m B_m}} fQ^{-1}(\varepsilon_m^*) \right] - 1 \right\} \quad (20)$$

*Proof:* See [Appendix C](#).

### **Conventional Branch and Bound Algorithm**

The traditional BB algorithm often addresses non-convex optimization problems by repeatedly exploring a tree with high computational complexity [21]. The BB algorithm is a binary tree search problem where the original problem is represented at the root node and each leaf node represents a subproblem within its corresponding subregion [32]. All of the tree's branches have been explored to estimate the objective function's upper bound and optimal power allocation [33]. If the optimal transmit power exceeds the maximum transmit power, that particular branch is eliminated and removed from the area being searched. Once all the branches have been investigated, the solution that possesses the minimum value is regarded as the ultimate optimal solution [12,34]. All unexplored branches are pruned, if the solution is found on the last branch, the majority of the branches will not



be pruned; in this situation, the computational complexity will increase. Because of the BB algorithm's limited performance and high computational complexity, we proposed a learning-to-optimize-based algorithm with low computational complexity that outperforms the conventional BB algorithm.

## 5 Learning-to-Optimize Approach

The learning-to-optimize approach is used to find a near-optimal solution for the power allocation problem in D2D communication with URLLC QoS constraints and provide low computational complexity. The conventional BB algorithm is used to learn the optimal pruning policy, which is the process of learning a complicated step in a particular algorithm and using that policy to produce optimal results with low computational complexity.

### 5.1 Proposed Learning-to-Optimize-Based Algorithm

The search process of the BB algorithm can be seen as a problem of making sequential decisions in a binary tree. We generate the unbalanced training sample sets from the BB algorithm because the number of preserve nodes is fewer than the number of prune nodes. An undersampling technique is utilized to address the unbalanced problem. We sample multiple disjoint subsets of the majority set with the same size as the minority set and then train a separate classifier for each dataset. Along with the minority set, a balanced training set can be achieved for each classifier. Several classifiers are being trained via ensemble learning, and their performance is subsequently improved by combining them. A node in the tree is either pruned or not in each decision based on the pruning policy. We are considering using supervised learning to develop a pruning policy that acts like a binary classifier. The feature node serves as the input, and the decision to either preserve or prune is the output. A learning-to-optimize-based algorithm learns the pruning policy from the BB algorithm search process and uses the features and labels to optimize the solution and ensure that it effectively identifies nodes to determine whether a node should be pruned or preserved with low computational complexity. Pruning policy explores every branch of the tree to determine the upper bound, lower bound, and optimal transmit power for the objective function. The related branch is deleted if the optimal transmit power is greater than the maximum transmit power. The ultimate optimal solution is the one that among all explored branches, has the lowest value. Fig. 2 shows the complete process of the training and testing phases. The major components of the learning-to-optimize-based algorithm are the classifier, feature design, flow of the algorithm, and computation complexity, which are defined in the following steps:

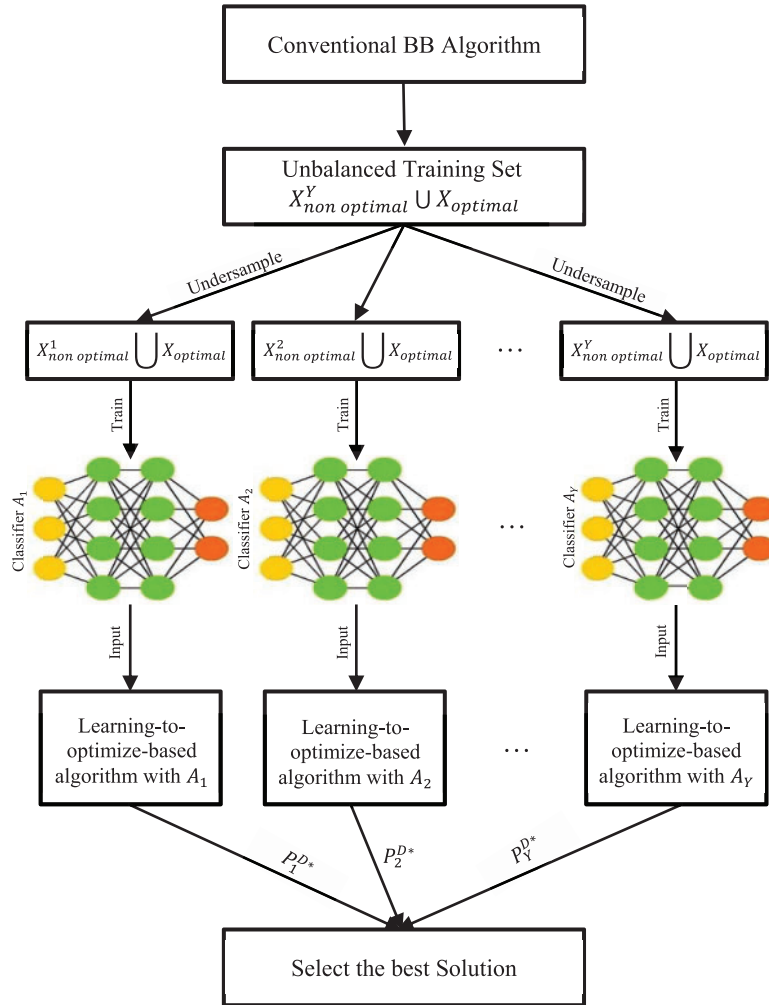
### 5.2 Classifier

Ensemble learning is used in conjunction with supervised learning to train multiple classifiers and combine their outputs for improved performance [35]. Using neural networks as a classifier in pruning policy involves training an  $L$ -layer binary classifier model to identify and remove unimportant parameters to reduce the model's complexity while maintaining accuracy. Rectified linear unit function, i.e.,  $Reul = \max(0, \cdot)$  used for input and hidden layers as an active function, and softmax function used as an active function for output layer. The probability of each classification being preserved or pruned is defined as

$$o_L [y] = \frac{\exp(i_L [y])}{\sum_{y'=1}^2 \exp(i_L [y'])}, y = 1, 2 \quad (21)$$

where  $i_L [y]$  is the input layer and  $o_L [y]$  is the output layer of the  $L$ -layer. The label  $u$  for the optimal solution is denoted as  $(1, 0)$ , and for the non-optimal solution, it is denoted as  $(0, 1)$ . The cross-entropy provided by the loss function default value is defined as

$$Loos = -u[1] \log(o_L[1]) - u[2] \log(o_L[2]) \tag{22}$$



**Figure 2:** The complete process of the training and testing phases

### 5.3 Feature Design

To effectively train classifiers, feature design is very important. An optimal solution to the problem  $P1'$  that maps the features,  $\phi(\alpha)$  is the feature node, where  $\phi(\cdot)$  denote the feature map. The nodes that contain the optimal solution within their feasible regions should be designated as “preserve,” while all other nodes should be categorized as “prune”. The input feature has a significant impact on classification accuracy and computational complexity. To improve accuracy, we need to find effective features that are closely related to the problem. We distinguish between two types of features: independent features and dependent features.

### 5.3.1 Independent Features

Independent features contain information about the conventional BB algorithm search processes, such as node features, branching features, and tree features; they contain the upper bound, lower bound, optimal value, and argument set  $\alpha$ .

### 5.3.2 Dependent Features

The dependent features are strongly associated with specific problems such as URLLC QoS constraints, channel state information (CSI), SINR, and power allocation constraints for D2D communication.

## 5.4 Flow of Algorithm

The proposed algorithm consists of a training dataset generated by the BB algorithm. An undersampling method is utilized for an imbalanced training dataset that integrates ensemble learning to train multiple classifiers and combine them to improve performance. The proposed learning-to-optimize-based algorithm explains in the following key steps:

In steps 1-2, the conventional BB algorithm is used to produce the training dataset. The features of every node that was explored are recorded by the algorithm. Nodes that have a feasible region that includes the best possible solution are identified as “preserve,” whereas the other nodes are designated as “prune” and store all the features and optimal solution into  $X$ .

In steps 3-4, the training set generated by the BB algorithm is unbalanced because there are fewer nodes that are preserved compared to the nodes that are pruned. Using this training set with traditional supervised learning, the classifier may not obtain the optimal solution for optimal nodes and bias for non-optimal nodes. We used the undersampling supervised learning technique to avoid the unbalanced training set problem. Undersampling creates multiple training sets that include both the minority class and a random subset of the majority class and then trains a classifier on each of these sets [36]. We divide the training set  $X$  into two subsets as  $X_{optimal}$  and  $X_{non-optimal}$  denotes optimal set and non-optimal set. Randomly split the  $X_{non-optimal}$  according to  $X_{optimal}$ , where  $|X_{non-optimal}| = |X_{optimal}|$  and  $X_{non-optimal}^1, X_{non-optimal}^2, \dots, X_{non-optimal}^Y$  from  $X_{non-optimal}$ ,  $\forall y \in \{1, 2, \dots, Y\}$ . The size of the  $X_{non-optimal}^y$  subset  $Y$  is based on the size of  $X_{optimal}$  and  $X_{non-optimal}$  where  $Y \in \{1, 2, \dots, \lceil |X_{non-optimal}| / |X_{optimal}| \rceil\}$ . The balanced training subset is denoted as  $X_{non-optimal}^y \cup X_{optimal}$  where  $y = 1, 2, \dots, Y$ , now we have  $Y$  training sets that can train for  $Y$  classifiers.

In step 5, we used an ensemble-based test that is similar to ensemble learning to test the multiple classifiers. Specifically, the proposed learning-to-optimize-based algorithm is run for  $Y$  times, each time for the different classifiers.

In steps 6–7, the upper bound is set to  $P^{ub} = \log \left( 1 + \frac{P_{max}^D g_{max}^D}{\sigma^2} \right)$ , where  $g_{max}^D$  the maximal value among all  $g_m^D$  and lower bound set to  $P^{lb} = 0$ .

In step 8, instead of using a counter, we establish a tolerance threshold  $\Delta$  to determine when to terminate the loop.

In step 9, set the auxiliary variable  $t_1$  and obtain the  $P_m^D$  by solving Eq. (20) with  $t_1$ .

In steps 10–16, ensemble learning is used for better performance,  $A_y(\cdot)$  is the output of the  $y$  – th classifier and  $\phi(\alpha)$  is the features. If  $A_y(\phi(\alpha))$  is greater than or equal to the threshold  $\omega$  then it is an optimal solution; otherwise, non-optimal. The threshold  $\omega \in [1, 0]$  is employed to restrict the search space. If  $\sum_{m=1}^L P_m^D \leq P_{max}^D$  and  $A_y(\phi(\alpha)) \geq \omega$  is optimal, then  $P_y^{D*} = P_m^D$  and lower bound set to  $t_1$ . If

$\sum_{m=1}^L P_m^D > P_{max}^D$  and  $A_y(\phi(\alpha)) < \omega$  is not optimal, then the upper bound is set to  $t_1$ . By comparing the  $P_m^D$  with  $P_{max}^D$ , we can examine whether  $t_1$  can be achieved as mentioned.

In step 20, we get the most optimal solution among all classifiers. The proposed learning-to-optimize-based algorithm is shown in Algorithm 1.

---

**Algorithm 1:** Learning-to-optimize-based Algorithm

---

```

1   Apply the conventional BB algorithm to find a solution to the problem and generate the set
    of data that is used for training.
2   All features and optimal solutions save into  $X$ .
3   Derived two sets  $X_{optimal}$  and  $X_{non-optimal}$  from set  $X$ .
4   Train the  $y - th$  classifier using  $X_{non-optimal}^y \cup X_{optimal}^y$ , where  $X_{non-optimal}^1, X_{non-optimal}^2, \dots, X_{non-optimal}^Y, y = 1, 2, \dots, Y$ .
5   for  $y = 1, 2, \dots, Y$  do
6       Set  $P^{ub} = \log\left(1 + \frac{P_{max}^D g_{max}^D}{\sigma^2}\right)$ ;
7       Set  $P^{lb} = 0$ ;
8       while  $P^{ub} - P^{lb} > \Delta$  do
9           Set  $t_1 = \frac{(P^{ub} + P^{lb})}{2}$ , solve the optimization problem  $P1'$  via Eq. (20) to obtain  $P_m^D$ ;
10          if  $\sum_{m=1}^L P_m^D \leq P_{max}^D$  and  $A_y(\phi(\alpha)) \geq \omega$  then
11              Set  $P^{lb} = t_1$ ;
12              if  $P_m^D \leq P_y^{D*}$  then
13                   $P_y^{D*} = P_m^D$ ;
14              end if
15          else
16              Set  $P^{ub} = t_1$ ;
17          end if
18      end while
19  end for
20  Select the best solution from the different classifiers  $P^{D*} = \arg\min_{y=1,2,\dots,Y} P_y^{D*}$ 

```

---

### 5.5 Complexity Analysis

The complexity of the learning-to-optimize-based algorithm can be evaluated by measuring the expected number of nodes explored. The anticipated number of examined nodes and the number of relaxed problems solved is  $O(M^2)$ , where  $M$  is the depth of the learning-to-optimize-based algorithm corresponding to the search procedure [12].

## 6 Simulation and Numerical Results

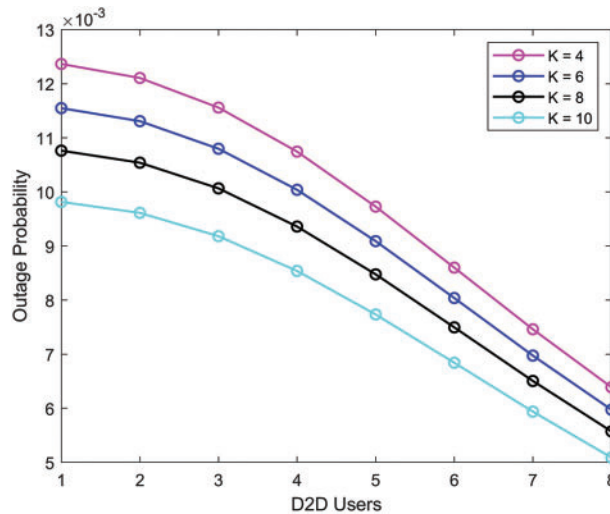
The performance of the proposed method is illustrated in this section through simulation results. We evaluate the performance achieved by the learning-to-optimize-based algorithm and compare it with the BB algorithm and heuristic algorithm. The optimization problem is to maximize the minimal achievable rate where the power allocation is optimized under the Rician fading channel. It is further obtained that the proposed learning-to-optimize-based algorithm minimizes the power consumption under the URLLC requirements. The dataset encompasses various data elements, including node

features, branching features, tree features, upper bound, lower bound, optimal value, the argument set  $\alpha$ , CSI, URLLC QoS constraints, SINR, and power allocation constraints for D2D communication. The error tolerance threshold  $\Delta$  is set to 0.001, and the learning rate for each classifier is set to 0.0005. Table 1 summarizes the simulation parameters we used.

**Table 1:** Simulation parameters

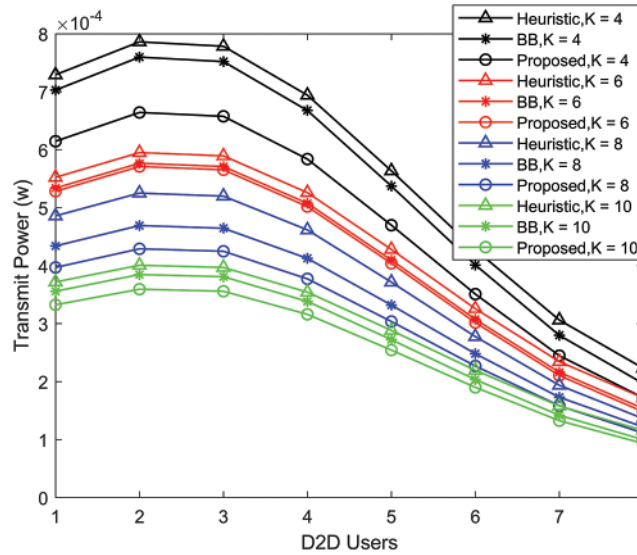
Parameters	Value
Number of cellular users $U_N$	10
Number of D2D users $L_M$	8
Rician $K$ -factor $K$	1, 2, 4, 6, 8, 10
Noise spectral density	-174 dBm/Hz
Cellular D2D links path loss model	$148 + 40\log(d)$
Shadowing standard deviation	10 dB
Maximum transmit power $P_{max}^D$	20 dBm
Maximum bandwidth $B_{max}$	20 MHz
Maximum transmission error probability $\varepsilon_{max}$	$1 \times 10^{-5}$
Maximum transmission time delay $T_{max}$	0.1 ms
Maximum tolerance $\Delta$	$1 \times 10^{-3}$
Outage probability threshold $\gamma_0$	$1 \times 10^{-2}$
Classification threshold $\omega$	[0, 1]

The outage probability of D2D communication is demonstrated in Fig. 3 with different Rician  $K$ -factors. The Rician parameter  $K$  affects the outage probability significantly in D2D communication. The outage probability decreases as the Rician  $K$ -factor increases because there is a stronger line of sight (LoS) component and a weaker propagation loss as a result of multiple paths. We observe that communication quality improves significantly when there is a LoS between communicators.



**Figure 3:** Outage probability of D2D communication with different Rician  $K$ -factor

Fig. 4 illustrates the transmit power performance of the learning-to-optimize-based algorithm compared with the BB algorithm and heuristic algorithm. We used different Rician  $K$ -factors, and we can see that the learning-to-optimize-based algorithm provides the optimal solution and minimum power consumption in D2D communication.



**Figure 4:** Optimal transmit power of D2D communication under URLLC with different Rician  $K$ -factor

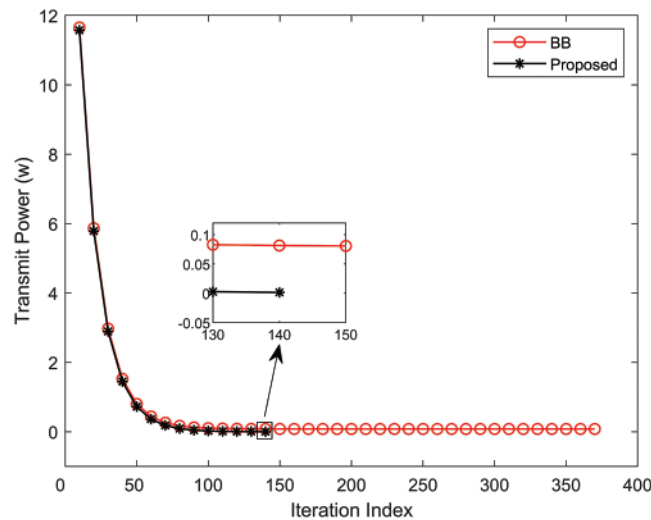
At the same time, there is a considerable performance gap between the heuristic algorithm, BB algorithm, and the learning-to-optimize-based algorithm, which shows the optimality of our proposed algorithm. When the Rician  $K$ -factor increases, the transmit power is optimized because a higher  $K$ -factor means a stronger LoS component and a weaker propagation loss due to multiple paths. We notice that when there is a LoS between communicators, the quality of communication significantly improves, leading to higher data rates for users.

Table 2 presents the computation time and average transmit power for different Rician  $K$ -factors. The learning-to-optimize-based algorithm provides optimal results with lower computational complexity. It can be seen that the proposed learning-to-optimize-based algorithm runs faster than the BB algorithm and heuristic algorithm and also provides optimal power allocation. Heuristic algorithms have high computational complexity but faster than the BB algorithm.

Fig. 5 plots the gap between the upper bound ( $P^{ub}$ ) and lower bound ( $P^{lb}$ ) of the transmit power returned by the proposed algorithm and BB algorithm. We can observe that the gap  $P^{ub} - P^{lb}$  of the learning-to-optimize-based algorithm reduces much faster than the BB algorithm.

**Table 2:** Average transmit power (w) and average computation time (s)

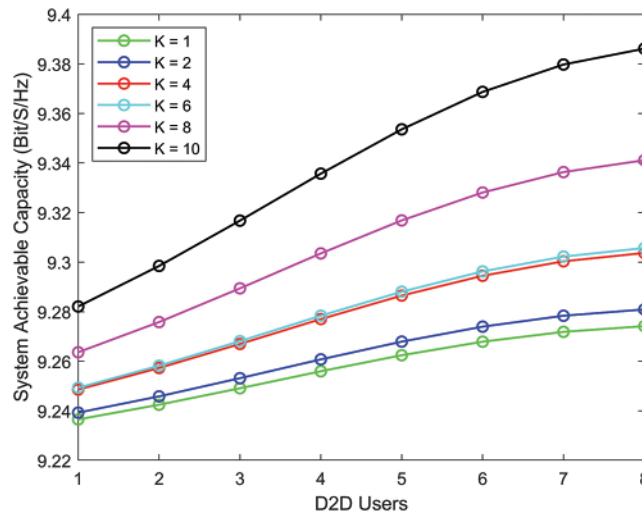
K	BB algorithm		Learning-to-optimize-based algorithm		Heuristic algorithm	
	Average transmit power	Average computation time	Average transmit power	Average computation time	Average transmit power	Average computation time
1	$7.21 \times 10^{-4}$	19.93	$5.67 \times 10^{-4}$	4.41	$8.70 \times 10^{-4}$	9.35
2	$6.68 \times 10^{-4}$	15.09	$5.19 \times 10^{-4}$	4.32	$8.23 \times 10^{-4}$	9.18
4	$5.10 \times 10^{-4}$	27.02	$4.46 \times 10^{-4}$	4.33	$5.36 \times 10^{-4}$	9.16
6	$3.89 \times 10^{-4}$	18.85	$3.83 \times 10^{-4}$	4.37	$4.07 \times 10^{-4}$	9.23
8	$3.15 \times 10^{-4}$	23.85	$2.88 \times 10^{-4}$	4.45	$3.52 \times 10^{-4}$	9.04
10	$2.58 \times 10^{-4}$	26.36	$2.41 \times 10^{-4}$	4.36	$2.74 \times 10^{-4}$	9.15



**Figure 5:** Convergence behavior of learning-to-optimize-based algorithm and BB algorithm with  $K = 8$

In Fig. 6, by comparing the achievable capacity of D2D communication, we can notice that the Rician  $K$ -factor increases and the system capacity also improves. This is due to the strong LoS communication and less loss of multi-path propagation. The learning-to-optimize-based algorithm performed better than the conventional BB algorithm and heuristic algorithm. The overall performance of the learning-to-optimize-based algorithm is better than the conventional BB algorithm and a heuristic algorithm. Our proposed algorithm provides the minimum power consumption in URLLC D2D communication under the Rician fading channel, which shows the optimality of our proposed algorithm.





**Figure 6:** Achievable capacity of D2D communication with different  $K$  values

## 7 Conclusion

This paper focused on finding the optimal power allocation in D2D communication. We considered an uplink D2D communication underlying the cellular system in a single-cell environment, and the Rician fading channel was investigated. The impact of the Rician  $K$ -factor was examined in the simulation and numerical results. We formulated the optimization problem of power allocation and maximized the minimal rate of D2D communication. The learning-to-optimize-based algorithm is proposed for optimizing the power allocation under the constraints of URLLC with Rician fading in D2D communication. The pruning policy learns from the BB algorithm, and the unbalanced dataset problem is handled by the undersampling method. Ensemble learning is used with supervised learning to train and combine multiple classifiers for better performance. The learning-to-optimize-based algorithm iteratively achieves an optimal solution and is compared with the BB algorithm and heuristic algorithm. The study has found that the learning-to-optimize-based algorithm has considerably lower computational complexity than the BB algorithm and heuristic algorithm and achieves better performance than the conventional BB algorithm.

**Funding Statement:** This work was supported in part by the National Natural Science Foundation of China under Grant 61771410, in part by the Sichuan Science and Technology Program 2023NSFSC1373, and in part by Postgraduate Innovation Fund Project of SWUST 23zx7101.

**Conflicts of Interest:** The authors declare that they have no conflicts of interest to report regarding the present study.

## References

- [1] D. Zhai, R. Zhang, Y. Wang, H. Sun, L. Cai *et al.*, "Joint user pairing, mode selection, and power control for D2D-capable cellular networks enhanced by nonorthogonal multiple access," *IEEE Internet of Things Journal*, vol. 6, no. 5, pp. 8919–8932, 2019.
- [2] C. She and C. Yang, "Available range of different transmission modes for ultra-reliable and low-latency communications," in *Proc. IEEE 85th Vehicular Technology Conf. (VTC Spring)*, Sydney, NSW, Australia, pp. 1–5, 2017.

- [3] Z. Chu, W. Yu, P. Xiao, F. Zhou, N. Al-Dhahir *et al.*, "Opportunistic spectrum sharing for D2D-based URLLC," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 9, pp. 8995–9006, 2019.
- [4] P. Gandotra and R. K. Jha, "Device-to-device communication in cellular networks: A survey," *Journal of Network and Computer Applications*, vol. 71, pp. 99–117, 2016.
- [5] R. I. Ansari, C. Chrysostomou, S. A. Hassan, M. Guizani, S. Mumtaz *et al.*, "5G D2D networks: Techniques, challenges, and future prospects," *IEEE Systems Journal*, vol. 12, no. 4, pp. 3970–3984, 2017.
- [6] F. Jameel, Z. Hamid, F. Jabeen, S. Zeadally and M. A. Javed, "A survey of device-to-device communications: Research issues and challenges," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 3, pp. 2133–2168, 2018.
- [7] R. Yin, C. Zhong, G. Yu, Z. Zhang, K. K. Wong *et al.*, "Joint spectrum and power allocation for D2D communications underlying cellular networks," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 4, pp. 2182–2195, 2015.
- [8] B. Singh, Z. Li and M. A. Uusitalo, "Flexible resource allocation for device-to-device communication in FDD system for ultra-reliable and low latency communications," in *Proc. of Advances in Wireless and Optical Communications (RTUWO)*, Riga, Latvia, pp. 186–191, 2017.
- [9] Y. Wu, D. Wu, L. Ao, L. Yang and Q. Fu, "Contention-based radio resource management for URLLC-oriented D2D communications," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 9, pp. 9960–9971, 2020.
- [10] H. Sun, X. Chen, Q. Shi, M. Hong, X. Fu *et al.*, "Learning to optimize: Training deep neural networks for interference management," *IEEE Transactions on Signal Processing*, vol. 66, no. 20, pp. 5438–5453, 2018.
- [11] W. Cui, K. Shen and W. Yu, "Spatial deep learning for wireless scheduling," *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 6, pp. 1248–1261, 2019.
- [12] Y. Shen, Y. Shi, J. Zhang and K. B. Letaief, "LORM: Learning to optimize for resource management in wireless networks with few training samples," *IEEE Transactions on Wireless Communications*, vol. 19, no. 1, pp. 665–679, 2019.
- [13] Z. Zhang and M. Tao, "A learning based branch-and-bound algorithm for single-group multicast beamforming," in *Proc. of IEEE Global Communications Conf. (GLOBECOM)*, Madrid, Spain, pp. 1–6, 2021.
- [14] A. Zappone, M. Di Renzo, M. Debbah, T. T. Lam and X. Qian, "Model-aided wireless artificial intelligence: Embedding expert knowledge in deep neural networks for wireless system optimization," *IEEE Vehicular Technology Magazine*, vol. 14, no. 3, pp. 60–69, 2019.
- [15] B. Chang, G. Zhao, Z. Chen, P. Li and L. Li, "D2D transmission scheme in URLLC enabled real-time wireless control systems for tactile internet," in *Proc. of 2019 IEEE Global Communications Conf. (GLOBECOM)*, Waikoloa, HI, USA, pp. 1–6, 2019.
- [16] B. Chang, L. Li, G. Zhao, Z. Chen and M. A. Imran, "Autonomous D2D transmission scheme in URLLC for real-time wireless control systems," *IEEE Transactions on Communications*, vol. 69, no. 8, pp. 5546–5558, 2021.
- [17] I. O. Sanusi, K. M. Nasr and K. Moessner, "Resource allocation for a reliable D2D enabled cellular network in factories of the future," in *Proc. of European Conf. on Networks and Communications (EuCNC)*, Dubrovnik, Croatia, pp. 89–93, 2020.
- [18] M. Hong and Z. Q. Luo, "Signal processing and optimal resource allocation for the interference channel," *Academic Press Library in Signal Processing*, vol. 2, pp. 409–469, 2014.
- [19] Z. Zhang and M. Tao, "Learning-based branch-and-bound for non-convex complex modulus constrained problems with applications in wireless communications," *IEEE Transactions on Wireless Communications*, vol. 21, no. 6, pp. 3752–3763, 2021.
- [20] Y. Bengio, A. Lodi and A. Prouvost, "Machine learning for combinatorial optimization: A methodological tour d'horizon," *European Journal of Operational Research*, vol. 290, no. 2, pp. 405–421, 2021.
- [21] M. Lee, G. Yu and G. Y. Li, "Learning to branch: Accelerating resource allocation in wireless networks," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 1, pp. 958–970, 2019.
- [22] D. Feng, L. Lu, Y. Yuan-Wu, G. Y. Li, G. Feng *et al.*, "Device-to-device communications underlying cellular networks," *IEEE Transactions on Communications*, vol. 61, no. 8, pp. 3541–3551, 2013.

- [23] L. Liang, G. Y. Li and W. Xu, "Resource allocation for D2D-enabled vehicular communications," *IEEE Transactions on Communications*, vol. 65, no. 7, pp. 3186–3197, 2017.
- [24] A. Goldsmith, *Wireless Communications*. UK: Cambridge University Press, 2005.
- [25] M. Peng, Y. Li, T. Q. Quek and C. Wang, "Device-to-device underlaid cellular networks under Rician fading channels," *IEEE Transactions on Wireless Communications*, vol. 13, no. 8, pp. 4247–4259, 2014.
- [26] Y. Wang, M. Chen, N. Huang, Z. Yang and Y. Pan, "Joint power and channel allocation for D2D underlaying cellular networks with Rician fading," *IEEE Communications Letters*, vol. 22, no. 12, pp. 2615–2618, 2018.
- [27] K. M. S. Huq, S. Mumtaz and J. Rodriguez, "Outage probability analysis for device-to-device system," in *Proc. of IEEE Int. Conf. on Communications (ICC)*, Kuala Lumpur, Malaysia, pp. 1–5, 2016.
- [28] H. C. Yang and M. S. Alouini, "Closed-form formulas for the outage probability of wireless communication systems with a minimum signal power constraint," *IEEE Transactions on Vehicular Technology*, vol. 51, no. 6, pp. 1689–1698, 2002.
- [29] H. He, H. Daume III and J. M. Eisner, "Learning to search in branch and bound algorithms," *Advances in Neural Information Processing Systems*, vol. 27, no. 1, pp. 1–11, 2014.
- [30] G. Durisi, T. Koch and P. Popovski, "Toward massive, ultrareliable, and low-latency wireless communication with short packets," in *Proc. of the IEEE*, vol. 104, no. 9, pp. 1711–1726, 2016.
- [31] C. Sun, C. She, C. Yang, T. Q. Quek, Y. Li *et al.*, "Optimizing resource allocation in the short blocklength regime for ultra-reliable and low-latency communications," *IEEE Transactions on Wireless Communications*, vol. 18, no. 1, pp. 402–415, 2018.
- [32] C. Lu and Y. F. Liu, "An efficient global algorithm for single-group multicast beamforming," *IEEE Transactions on Signal Processing*, vol. 65, no. 14, pp. 3761–3774, 2017.
- [33] Y. Zhang, Y. Yang and L. Dai, "Energy efficiency maximization for device-to-device communication underlaying cellular networks on multiple bands," *IEEE Access*, vol. 4, no. 1, pp. 7682–7691, 2016.
- [34] T. H. Cormen, C. E. Leiserson, R. L. Rivest and C. Stein, *Introduction to Algorithms*. USA: MIT press, 2022.
- [35] M. A. Arbib, *The Handbook of Brain Theory and Neural Networks*. USA: MIT Press, 2002.
- [36] X. Y. Liu, J. Wu and Z. H. Zhou, "Exploratory undersampling for class-imbalance learning," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 39, no. 2, pp. 539–550, 2008.
- [37] A. Nuttall, "Some integrals involving the Q<sub>M</sub> function (Corresp.)," *IEEE Transactions on Information Theory*, vol. 21, no. 1, pp. 95–96, 1975.
- [38] Y. D. Yao and A. Sheikh, "Outage probability analysis for microcell mobile radio systems with cochannel interferers in Rician/Rayleigh fading environment," *Electronics Letters*, vol. 13, no. 26, pp. 864–866, 1990.

## Appendix A

Regarding the outage probability of the D2D link, we have

$$P_{out}^{D2D} = \{s_m^D \leq s_0\} = Pr \left( \frac{P_m^D g_m^D h_m^D}{\sum_{n \in U} p_{n,m} P_n^C g_{n,m}^{CD} h_{n,m}^{CD}} \leq s_0 \right) \quad (23)$$

$$P_{out}^{D2D} = 1 - \int_{s_{th}}^{\infty} \left( \int_0^{\frac{s_D}{s_0}} f_{s_I}(s_I) ds_I \right) f_{s_D}(s_D) ds_D \quad (24)$$

The outage probability of D2D can be rewritten by using Eqs. (5) and (6) in Eq. (24)

$$P_{out}^{D2D} = 1 - \int_{s_{th}}^{\infty} \left( \int_0^{\frac{s_D}{s_0}} \frac{s_I^{W_I-1}}{\bar{s}_I^{W_I} (W_I - 1)!} \exp\left(-\frac{s_I}{\bar{s}_I}\right) \right) \left( \frac{(K+1)}{\bar{s}_D} e^{-K - \frac{(K+1)s_D}{\bar{s}_D}} I_0 \left( 2\sqrt{\frac{K(K+1)s_D}{\bar{s}_D}} \right) \right) ds_D \quad (25)$$

Solving the inner integral of Eq. (25) and getting the outage probability as

$$P_{out}^{D2D} = 1 - Q_1 \left( \sqrt{2K}, \sqrt{\frac{2(K+1)s_{th}}{\bar{s}_D}} \right) + \frac{K+1}{\bar{s}_D} \sum_{j=0}^{W_I-1} \frac{1}{j!} \int_{s_{th}}^{\infty} \left( \frac{s_D}{s_0 \bar{s}_I} \right)^j \exp \left[ -K - \left( \frac{1}{s_0 \bar{s}_I} + \frac{K+1}{\bar{s}_D} \right) s_D \right] I_0 \left( 2 \sqrt{\frac{K(K+1)s_D}{\bar{s}_D}} \right) ds_D \quad (26)$$

where  $Q_1(\cdot, \cdot)$  is the first-order Marcum  $Q$ -function. Solving the integral and using  $Q$ -function to simplify Eq. (26) [37], the outage probability can be expressed as

$$P_{out}^{D2D} = 1 - Q_1 \left( \sqrt{2K}, \sqrt{\frac{2(K+1)s_{th}}{\bar{s}_D}} \right) + \frac{a^2}{2K} \sum_{j=0}^{W_I-1} \frac{\beta_j}{j!} Q_{2j+1,0}(a, b) \quad (27)$$

$$\text{where } \beta_j = \frac{\exp \left[ -K + \frac{K}{\left( 1 + \frac{\bar{s}_D}{(K+1)s_0 \bar{s}_I} \right)} \right]}{\left( 2 + \frac{2(K+1)s_0 \bar{s}_I}{\bar{s}_D} \right)^j}, a = \sqrt{\frac{2K}{\left( 1 + \frac{\bar{s}_D}{(K+1)s_0 \bar{s}_I} \right)}} \text{ and } b = \sqrt{2 \left( 1 + K + \frac{\bar{s}_D}{s_0 \bar{s}_I} \right) \frac{s_0}{\bar{s}_D}}$$

where  $W_I = 1$  [28,38] the Eq. (27) can be expressed as

$$P_{out}^{D2D} = 1 - Q_1 \left( \sqrt{2K}, \sqrt{\frac{2(K+1)s_{th}}{\bar{s}_D}} \right) + \frac{a^2}{2K} \exp \left[ -K + \frac{a^2}{2} \right] Q_1(a, b) \quad (28)$$

The interference-limited  $s_{th} = 0$ , Eq. (28) can be

$$P_{out}^{D2D} = \frac{\exp \left[ -K + \frac{K}{\left( 1 + \frac{\bar{s}_D}{(K+1)\bar{s}_D} \right)} \right]}{\left( 1 + \frac{\bar{s}_D}{(K+1)s_0 \bar{s}_I} \right)} \quad (29)$$

### Appendix B. Proof of Theorem 1

Applying integration-by-parts

$$C_m = E \left[ \log \left( 1 + s_m^D \right) \right] = \int_0^{\infty} \log(1+z) f_{s_m^D}(z) dz \quad (30)$$

$$= \int_{z=0}^{\infty} \int_{y=0}^z \frac{1}{1+y} f_{s_m^D}(z) dy dz \quad (31)$$

$$= \int_{y=0}^{\infty} \frac{1}{1+y} dy \int_{z=y}^{\infty} f_{s_m^D}(z) dz \quad (32)$$

$$= \int_0^{\infty} \frac{1 - F_m^D(z)}{1+z} dz \quad (33)$$

The Eq. (33) is derived by substituting  $F_m^D(z) = Pr\{s_m^D \leq s_0\}$ , where  $Pr\{s_m^D \leq s_0\}$  is given in Eq. (8).

### Appendix C

The minimum transmit power can be derived from Eq. (9) with an auxiliary variable  $t_1$

$$R_m = t_1 \quad (34)$$

$$\frac{T_m B_m}{T_m \ln 2} \left[ C_m - \sqrt{\frac{1}{T_m B_m}} fQ^{-1}(\varepsilon_m^*) \right] = t_1 \quad (35)$$

$$\frac{T_m \ln 2(t_1)}{T_m B_m} = C_m - \sqrt{\frac{1}{T_m B_m}} fQ^{-1}(\varepsilon_m^*) \quad (36)$$

$$C_m = \frac{T_m \ln 2(t_1)}{T_m B_m} + \sqrt{\frac{1}{T_m B_m}} fQ^{-1}(\varepsilon_m^*) \quad (37)$$

where  $C_m = \log(1 + s_m^D)$

$$\log \left( 1 + \frac{P_m^D g_m^D h_m^D}{\sum_{n \in U} p_{n,m} P_n^C g_{n,m}^{CD} h_{n,m}^{CD} + \sigma^2} \right) = \frac{T_m \ln 2(t_1)}{T_m B_m} + \sqrt{\frac{1}{T_m B_m}} fQ^{-1}(\varepsilon_m^*) \quad (38)$$

$$1 + \frac{P_m^D g_m^D h_m^D}{\sum_{n \in U} p_{n,m} P_n^C g_{n,m}^{CD} h_{n,m}^{CD} + \sigma^2} = \exp \left[ \frac{T_m \ln 2(t_1)}{T_m B_m} + \sqrt{\frac{1}{T_m B_m}} fQ^{-1}(\varepsilon_m^*) \right] \quad (39)$$

$$P_m^D = \frac{\sum_{n \in U} p_{n,m} P_n^C g_{n,m}^{CD} h_{n,m}^{CD} + \sigma^2}{g_m^D h_m^D} \left\{ \exp \left[ \frac{T_m \ln 2(t_1)}{T_m B_m} + \sqrt{\frac{1}{T_m B_m}} fQ^{-1}(\varepsilon_m^*) \right] - 1 \right\} \quad (40)$$

where  $\frac{\sum_{n \in U} p_{n,m} P_n^C g_{n,m}^{CD} h_{n,m}^{CD} + \sigma^2}{g_m^D h_m^D}$  solved by outage probability Eq. (3)

$$P_m^D = P_{out}^{D2D} \left\{ \exp \left[ \frac{T_m \ln 2(t_1)}{T_m B_m} + \sqrt{\frac{1}{T_m B_m}} fQ^{-1}(\varepsilon_m^*) \right] - 1 \right\} \quad (41)$$