Sum rate maximization of D2D networks with energy constrained UAVs through deep unsupervised learning

Benjamin Lea
Department of Engineering
Thompson Rivers University (TRU)
Kamloops, BC, Canada
leab17@tru.ca

Debaditya Shome
School of Electronics Engineering
KIIT University
Odisha, India
debadityashome@ieee.org

Omer Waqar
Department of Engineering
Thompson Rivers University (TRU)
Kamloops, BC, Canada
owaqar@tru.ca

Jabed Tomal
Department of Mathematics and Statistics
Thompson Rivers University (TRU)
Kamloops, BC, Canada
jtomal@tru.ca

***Abstract*—We consider a system model in which several energy harvesting (EH) unmanned aerial vehicles (UAVs), often known as drones, are deployed with device-to-device (D2D) communication networks. For the considered system model, we formulate an optimization problem that aims to find an optimal transmit power vector which maximizes the sum rate of the D2D network while also meets the minimum energy requirements of the UAVs. Because of the nature of the system model, it is necessary to deliver solutions in real time i.e., within a channel coherence time. As a result, conventional non-data-driven optimization methods are inapplicable, as either their run-time overheads are prohibitively expensive or their solutions are significantly suboptimal. In this paper, we address this problem by proposing a deep unsupervised learning (DUL) based hybrid scheme in which a deep neural network (DNN) is complemented by the full power scheme. It is shown through simulations that our proposed hybrid scheme provides up to 91% higher sum rate than an existing fully non-data driven scheme and our scheme is able to obtain solutions quite efficiently, i.e., within a channel coherence time.**

***Index Terms*—Device-to-device (D2D) communications, energy harvesting (EH), non-convex optimization, unmanned aerial vehicles (UAVs) and deep unsupervised learning (DUL).**

I. Introduction

It is envisioned that the device-to-device (D2D) networks will be integral component of the beyond fifth generation (B5G) and/or sixth generation (6G) communication systems owing to their ability to enhance the overall performance significantly. In fact, it is reported in [1] that D2D communications enable an increase in spectrum utilization up to 50% when compared to the conventional cellular communications. However, to realize promising gains of such networks, D2D-generated interference has to be managed which is a challenging task. To this end, power control [2] is considered as the most appropriate method for management of the D2D-generated interference. In a power control method, transmit powers of the D2D transmitters are adjusted in a manner either to maximize any performance metric (e.g., sum rate) or to meet any quality-of-service (QoS) requirement.

Moreover, unmanned aerial vehicles (UAVs) have been gaining popularity due to their ability to support plethora of daily life applications. Leveraging their ability to dynamically configure their altitudes and speeds, UAVs cannot only fill in the coverage holes of any communication system, but also can provide high quality line-of-sight (LoS) links. Motivated by this fact, researchers have been proposing to integrate UAVs into the next-generation communication systems [3]. Furthermore, UAVs also facilitate to provide on-demand communication such as during rescue services, emergency scenarios etc. However, there exists unique challenges for widespread deployment of UAVs due to the stringent constraints imposed by their size, weight, and power (SWAP) limitations. Nevertheless, some advancements have been made recently in reducing the size and weight of the UAVs, limited power supply of

This work is supported by funding from the Internal Research Fund (IRF) and the Undergraduate Research Experience Award Program (UREAP) of the Thompson Rivers University (TRU), BC, Canada.

978-1-5386-5541-2/18/$31.00 ©2021 IEEE

the UAVs still poses a formidable challenge. To this end, wireless energy harvesting (EH) is considered as a breakthrough technology as it is capable to provide an uninterruptible power supply to the UAVs [4]. It is worth mentioning here that wireless EH also facilitates to either completely replace the batteries or to decrease the number/size of required batteries, thus enables reduction in the size and weight of the UAVs that ultimately increases their flight times.

Keeping in mind the aforementioned discussion and significance of the D2D communications and the wireless EH assisted UAVs, we integrate these two paradigms in this paper. More specifically, we consider the power control problem for the D2D networks in presence of the energy constrained UAVs that can harvest energy wirelessly. It is well known that power control for the D2D networks is a non-deterministic polynomial-time (NP) hard problem [5], thus it is difficult to obtain high-quality solution for this problem within a channel coherence time (typically on the order of few milliseconds). Moreover, the minimum energy requirement constraint for the UAVs further exacerbates this problem. This motivates us to address this problem by leveraging deep learning (DL) theory that has shown remarkable success in various fields such as computer vision, natural language processing (NLP) etc. Our main contributions are as follows:

(i) We provide an analytical framework that characterizes the sum rate for the D2D network and the amount of EH for the UAVs.

(ii) By leveraging the above-mentioned framework, we develop an analytical optimization problem with an aim to maximize the sum rate while meeting the minimum energy requirement of each UAV.

(iii) We propose a deep unsupervised learning (DUL) based new scheme that provides a high-quality solution for the aforementioned NP-hard problem. Moreover, it is shown that the sum rate achievable through our proposed scheme is significantly larger than the existing fully non-data driven scheme, i.e., full power scheme.

The rest of this paper is organized as follows: Section II provides an overview of the related works. System model is given in Section III. Moreover, section IV provides details about our proposed scheme. Simulation results are provided in section V. Finally, we conclude the paper in section VI.

II. Overview of Related Works

In [6], joint optimization for user association, trajectory and power control for multi-UAVs enabled wireless network was carried out. In particular, the authors proposed to use block coordinate descent in conjunction with the successive convex optimization techniques. In [7], a UAV-aided secure communication network was considered. The minimum secrecy rate was maximized by optimizing the user scheduling, power and trajectory control for UAVs using the block successive upper bound minimization techniques. It is worth mentioning here that the non data-driven optimization techniques as considered in [6], [7] and references therein. Generally, non data-driven optimization techniques have either prohibitive time-complexity or these techniques provide highly non-optimal results (e.g., heuristic-based algorithms). Considering this, recently, the researchers started to employ machine learning (ML)-based solutions for optimizing the parameters of the UAV-enabled wireless networks.

Researchers in [8] proposed a ML based solution, specifically a reinforcement learning (RL) approach (Q-learning algorithm) was utilized that optimizes the trajectory and placement of the UAV base stations based on real time user movement. For this, the researchers used data that was randomly generated and suggested to use a more practical data set for future research (e.g., data from social networks such as Twitter and Facebook can be used for collecting position information of mobile users) [6]. Moreover, in [9], researchers proposed a method to optimize deployment and trajectory of UAV for providing wireless communications to high-speed trains. Specifically, the authors proposed Soft Actor-Critic algorithm to optimize trajectory for communicating with the train and Support Vector Machine (SVM) algorithm to find optimal initial deployment. Furthermore, in [10] it was proposed to use a deep reinforcement learning algorithm for optimizing the trajectory of the UAV for communication with wireless sensor networks. In [11] researchers used big data and DL to offer a solution that determine the optimal UAV battery size, and determine parameters for defining the relationship between quality of service and UAV efficiency. In contrast to the aforementioned works, [12] considered a fixed wing UAV because fixed wings facilitate longer flight times and higher payload. In this paper, the authors proposed a circular flight trajectory as solution for how fixed wing UAV should fly to provide coverage to an area. Furthermore, in [13] the researchers proposed a deep reinforcement learning (DRL) method, The DRL method is for controlling multiple UAVs to provide base station coverage for a large area. The proposed method used two DL NNs to learn about the UAVs environments and make decisions to optimize service. It is evident that most of the prior works considered RL for training the NNs. RL provides good performance, however there exists some convergence issues owing to the presence of exploitation-exploration tradeoff. Therefore, in order to avoid this issue, in some works, the authors considered supervised learning (SL) approach for training the deep neural networks (DNNs) (e.g., see [14] and references therein). However, SL requires labeled data-sets which is typically impractical to generate (due to prohibitively high time complexity), particularly for the NP-hard problems.

In order to overcome aforementioned issues of the SL and RL based schemes, some researchers recently proposed unsupervised (UL) based schemes. For instance, in [15] the authors proposed a method for solving NP-hard generalized assignment problems using DNN and showed that DUL is very efficient compared to traditional optimization methods. Similarly, in [5] the researchers proposed a DUL method for optimizing the sum rate, however the constraint for the minimum energy requirements are not considered. Recently, a DUL based scheme is proposed in [16] for maximizing the throughput of an intelligent reflecting surface (IRS) based wireless powered networks. Considering a huge advantage of UL of being an efficient learning mechanism, in this paper, we propose DUL based scheme that maximizes the sum rate of the D2D network while meeting the minimum energy requirements of each UAV.

III. System Model

We consider a D2D network with energy constrained UAVs, as shown in Fig. 1. The system has K transmitters, K receivers, and N energy harvesting UAVs. The received signal at the ith receiver is modeled as:

$$y_i = \sqrt{P_i} h_{i,i}, x_i + \sum_{j \in K/\{i\}} \sqrt{P_j} h_{j,i} x_j + z_i, \quad \forall i, j = 1, 2, ..., K, P_j \,\varepsilon\, (0, P_{\text{max}}], \tag{1}$$

where x_j is the transmit symbol of jth transmitter with average transmit of P_j , h_{ji} is the channel coefficient between the jth transmitter and ith receiver and z_i is the additive noise at the ith receiver. Moreover, h_{ji} and z_i follow the circularly symmetric complex Gaussian (CSCG) distributions with zero means, unit and σ^2 variances, respectively. P_max denotes the maximum allowable transmit power.

Similarly, the signal received at the nth energy constrained UAV is represented as:

$$y_n = \sum_{j=1}^{K} \sqrt{P_j} g_{n,j} x_j + w_n \qquad \forall n = 1, 2, ..., N, \tag{2}$$

where g_{nj} is the channel coefficient between nth UAV and jth transmitter, and w_n is the additive noise at nth UAV. The channel coefficient g_{nj} and the noise w_n follow zero mean CSCG distributions with unit and σ^2 variances, respectively. The energy harvested at the nth UAV given the EH efficiency (η), is

$$\begin{aligned} E_n &= \eta |\sum_{j=1}^{K} \sqrt{P_j} g_{n,j}|^2 \\ &= \eta P_{total} |g_n|^2, \end{aligned} \tag{3}$$

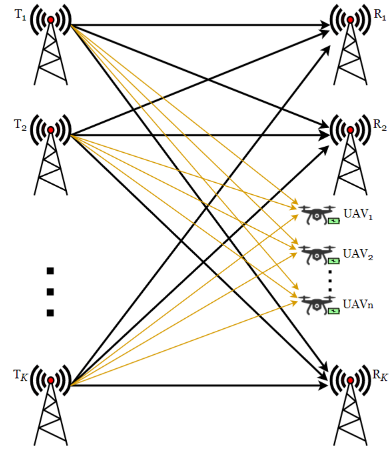

Fig. 1: System Model

where $P_{total} = \sum_{j=1}^{K} P_j$ and g_n is the equivalent channel coefficient for the nth UAV that is a zero mean and unit variance CSCG random variable.

The sum rate for the ith receiver is given as:

$$R_i(\mathbf{P}) = \log_2 \left(1 + \frac{P_i |h_{i,i}|^2}{\sigma^2 + \sum_{j \in K/\{i\}} P_j |h_{j,i}|^2} \right), \tag{4}$$

where $\mathbf{P} = [P_1, P_2, ..., P_K]$ represents a transmit power vector.

In this paper, our aim is to maximize the data rates between all pairs of the transmitters and receivers of the D2D network with two constraints; (C1) P_j must not exceed the maximum transmit power, $P_{\max}$ and (C2) amount of energy harvested at each nth UAV should meet its minimum energy requirement ($e_{\min,n}$). Formally, we can write

$$\begin{aligned} &\max_{\mathbf{P}} R(\mathbf{P}) = \sum_{i=1}^{K} R_i(\mathbf{P}) \quad \forall i = 1, 2, ..., K. \\ &\text{s.t.} \\ &\text{C1}: \quad 0 \leq P_j \leq P_{\text{max}}, \forall j = 1, 2, ..., K. \\ &\text{C2}: \quad E_n \geq e_{\min,n}, \ \forall n = 1, 2, ..., N. \end{aligned} \tag{5}$$

IV. Our Proposed Scheme

In this section, we propose a new DUL based scheme that provides a high-quality solution for the optimization problem given in eq. (5). The details of our proposed scheme are given in the subsequent subsections.

A. Architecture of DNN

We utilize a deep feedforward NN as shown in Fig. 2. The input layer of the neural network takes the feature vector ($\mathbf{f}$). Our feature vector is the concatenation of the magnitude of the channel coefficients, i.e., $\mathbf{f} = \{|h_{1,1}|, |h_{0,1}|, ..., |h_{k,k}|, |g_1|, ..., |g_n|\}$. Following

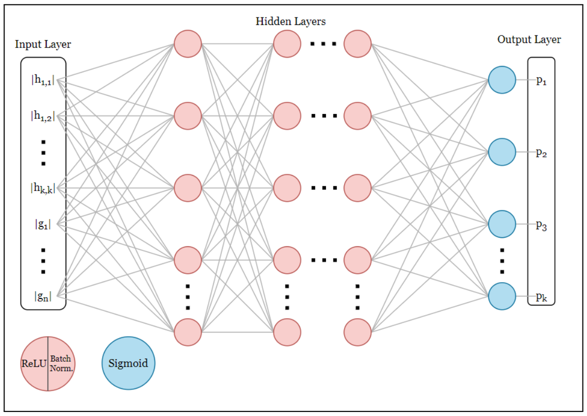

Fig. 2: Architecture of the DNN

this are the densely connected hidden layers. The output of each node in the hidden layer is calculated using a ReLU activation function which is then followed by a batch normalization layer. Finally, the output layer uses the sigmoid activation function. The sigmoid activation is very important as the sigmoid function will cause the output value to be between 0 and 1. This allows us to multiply this output with a given $P_{\max}$, thus easily meeting the C1 constraint of the optimization problem.

B. Training mechanism

Next step is to define a custom loss function which is a function of the sum rate and the C2 constraint of the optimization problem. This custom loss function is then minimized by the DNN. Following the penalty-based optimization approach, we define the following loss function (L)

$$L = \frac{1}{|\mathcal{F}|}\sum_{f\in\mathcal{F}}\Big[-R(\mathbf{P}) + \lambda\sum_{n=1}^{N}\left(\mathsf{ReLU}^2\left(e_{\min,n} - E_n\right)\right) \tag{6}$$

where λ is the penalty parameter and $\mathcal{F}$ denotes the mini-batch of size $|\mathcal{F}|$. The DNN aims to minimize the loss function, therefore maximizes the sum rate $R(\mathbf{P})$. The term $\lambda\sum_{n=1}^{N}\left(\mathsf{ReLU}^2\left(e_{min,n} - E_n\right)\right)$ is know as the penalty term that causes the loss function to increase if the constraint C2 is not met. In this way, the NN is penalized and is pushed to meet this energy constraint. The term λ acts as a multiplier to the penalty term and enforces the constraint when λ is large and relaxes the constraint when λ is small. Therefore, λ is considered as a hyperparameter the value of which should be chosen wisely in order to achieve a better sum rate-constraint violation tradeoff. It is important mentioning here that our customized loss function enables the DNN to train in an unsupervised manner and efficiently because labeled datasets are not required.

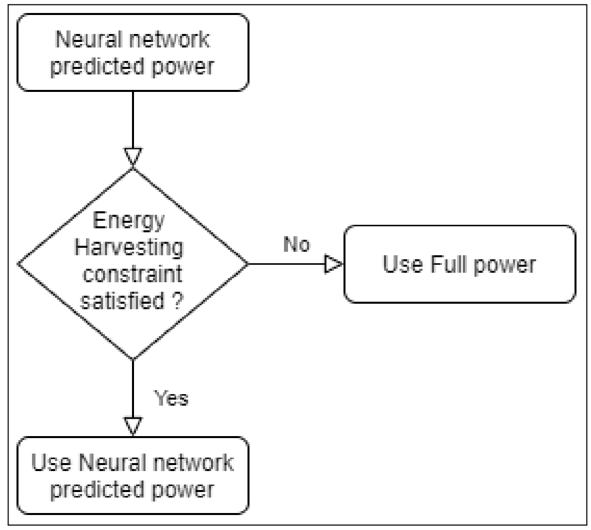

Fig. 3: Illustration of the proposed hybrid scheme

C. Hybrid scheme

Although the parameter λ allows us to achieve a tradeoff between the sum rate and the constraint violation probability (e.g., see Fig. 8 and Fig. 9), DNN does not guarantee to meet the constraint C2. In order to make sure to meet constraint C2 of the optimization problem, we further propose a hybrid scheme in which the full power and DNN based schemes are complemented as shown in Fig. 3, given on top of this page. It is shown in the next section that our proposed hybrid scheme guarantees to meet both constraints of the optimization problem while achieving much higher sum rate as compared to the full power scheme (which is a fully non-data driven scheme) within a typical channel coherence time.

V. Simulation Results

In this section, we provide simulation results of our proposed scheme and compare these results with that of the non-data driven scheme i.e., full power scheme. In full power scheme, all the transmitters transmit with $P_{\max}$. The specifications of machine that is used to train and test the DNN are given in Table 1. Furthermore, Table 2 and Table 3 provide values of the hyperparameters and system parameters of the DNN and the system model, respectively. The DNN was created and run using Python with the package Tensorflow V2 with Keras.The datasets comprise of all feasible examples. Furthermore, the test dataset is utilized to take an average for the sum rate and to evaluate the constraint violation probability. Specifically, constraint violation probability represents a ratio of the number

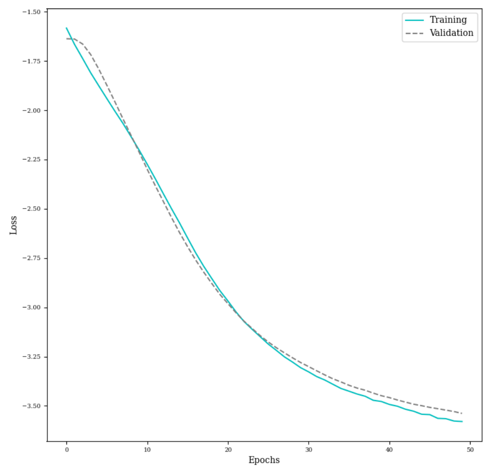

Fig. 4: Training with learning rate = 0.0001

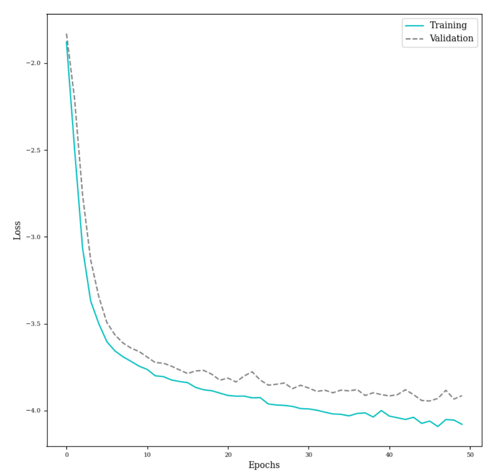

Fig. 5: Training with learning rate = 0.001

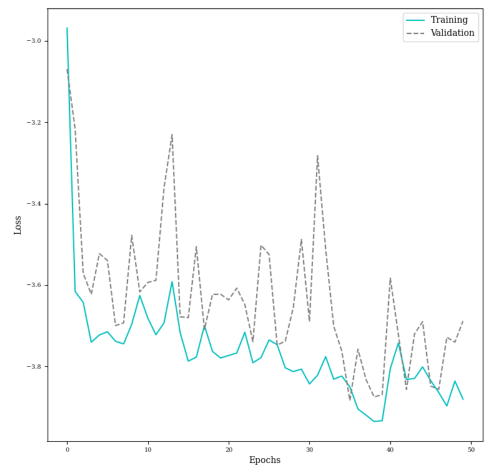

Fig. 6: Training with learning rate = 0.01

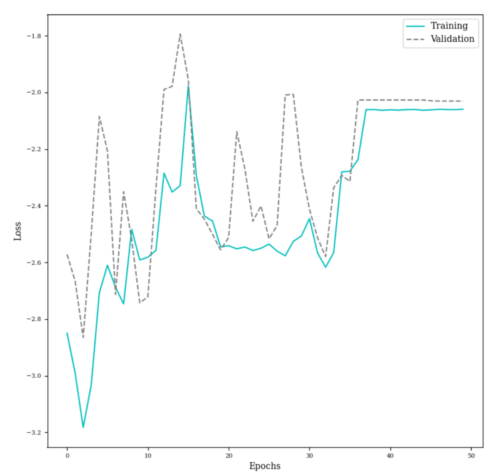

Fig. 7: Training with learning rate = 0.1

of examples that do not meet constraint C2 to the total number of examples in the test dataset.

As learning rate is one of the most important hyperparameters for the DNN, we tried four different learning rates as shown in the Fig. 4 to Fig. 7. It is clear from these figures that learning rate = 0.0001 is the most appropriate for our DNN as both the training and validation losses decrease consistently and does not result in overfitting. Therefore, learning rate = 0.0001 is taken in the remaining simulation results.

Fig. 8 and Fig. 9 show tradeoff between the sum rate and the constraint violation as a function of the penalty parameter λ. As expected, both constraint violation probability and the sum rate decrease with increasing λ. Thus, for high sum rate we need a small value of λ, whereas for small constraint violation probability, a high value of λ is required. It is evident from Fig. 8 that it is highly unlikely that constraint violation probability reaches to zero even for very large value of λ, therefore we propose a hybrid scheme and show its achievable sum rate in Fig. 10. It is clear from Fig. 10 that our proposed hybrid scheme achieves up to 91% higher sum rate (for $\lambda = 50$) as compared to the full power scheme and also guarantees to meet the energy harvesting constraint i.e., the constraint violation probability is zero for the proposed hybrid scheme.

TABLE I: Machine specifications

Component	**Specification**
CPU	Intel(R) Core(TM) i7@ 2.60GHz
GPU	NVIDIA GeForce GTX 1650
GPU memory	8113 MB
CUDA cores	896

TABLE II: Hyperparameters

Hyperparameter	Value
Batch size	100
Learning rate	0.0001
Epochs	50
Optimizer	Adam
Training data size	10000
Validation data size	2000
Test data size	1000

TABLE III: System parameters

System parameter	Value
K	4
N	3
P_{max}	1
$e_{\min,n}$	0.5
η	1
σ^2	0.05

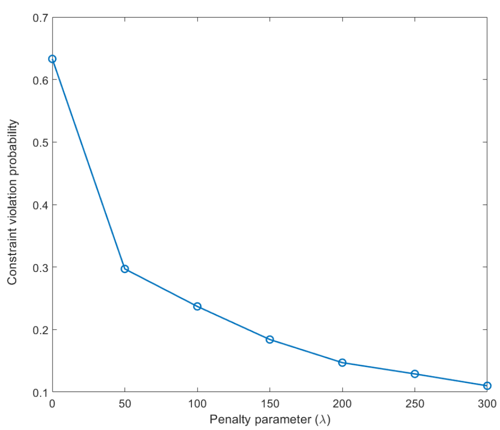

Fig. 8: Constraint violation probability plot

It is important mentioning here that we found out through simulations that the average time taken by the hybrid scheme to obtain the sum rate is approximately 27.45 milliseconds. This shows that our proposed hybrid scheme can obtain a high-quality solution for the considered NP-hard problem within a channel coherence time, thus can serve as an excellent scheme for the practical wireless communication systems.

VI. Conclusion

In this paper, we formulate an optimization problem for maximizing the sum rate of the D2D networks while

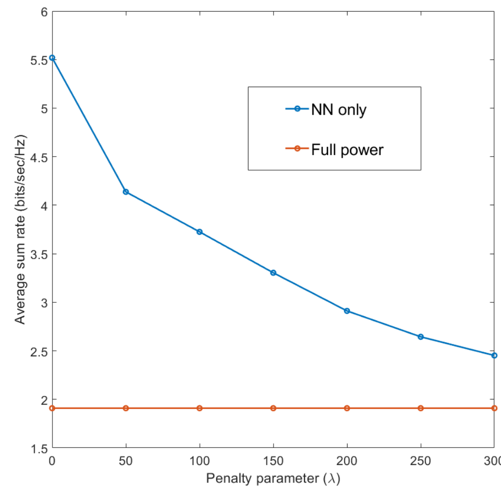

Fig. 9: Sum rates with NN only and full power schemes

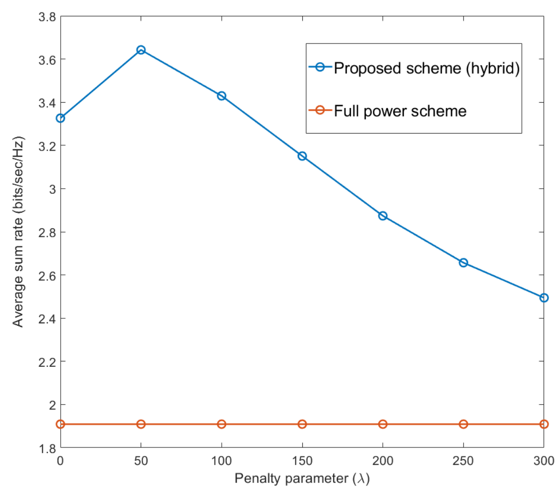

Fig. 10: Sum rate with proposed hybrid and full power schemes

considering the constraints of the maximum allowable transmit power and the minimum energy requirements of the UAVs. As the optimization problem is an NP-hard, we propose a new DUL based hybrid scheme that obtains a high-quality solution quite efficiently. In particular, it is shown through simulations that our proposed scheme achieves a significant higher sum rate as compared to the full power scheme. As a future work, we intend to incorporate a path-loss model and compare our results with other state-of-the-art non-data driven schemes.

References

[1] S.-Y. Lien, C.-C. Chien, F.-M. Tseng, and T.-C. Ho, "3GPP device-to-device communications for beyond 4G cellular net-

works," *IEEE Communications Magazine*, vol. 54, no. 3, pp. 29–35, 2016.

[2] C. Liu and B. Natarajan, "Power-aware maximization of ergodic capacity in D2D underlay networks," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 3, pp. 2727–2739, 2017.

[3] Y. Zeng, Q. Wu, and R. Zhang, "Accessing from the sky: A tutorial on UAV communications for 5G and beyond," *Proceedings of the IEEE*, vol. 107, no. 12, pp. 2327–2375, 2019.

[4] A. M. Le, L. H. Truong, T. V. Quyen, C. V. Nguyen, and M. T. Nguyen, "Wireless power transfer near-field technologies for unmanned aerial vehicles (UAVs): A review," *EAI Endorsed Transactions on Industrial Networks and Intelligent Systems*, vol. 7, 1 2020.

[5] F. Liang, C. Shen, W. Yu, and F. Wu, "Towards optimal power control via ensembling deep neural networks," *IEEE Transactions on Communications*, vol. 68, no. 3, pp. 1760–1776, 2019.

[6] Q. Wu, Y. Zeng, and R. Zhang, "Joint trajectory and communication design for multi-UAV enabled wireless networks," *IEEE Transactions on Wireless Communications*, vol. 17, no. 3, pp. 2109–2121, 2018.

[7] H. Lee, S. Eom, J. Park, and I. Lee, "UAV-aided secure communications with cooperative jamming," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 10, pp. 9385–9392, 2018.

[8] Y. Liu, Z. Qin, Y. Cai, Y. Gao, G. Y. Li, and A. Nallanathan, "UAV communications based on non-orthogonal multiple access," *IEEE Wireless Communications*, vol. 26, no. 1, pp. 52–57, 2019.

[9] Y. M. Park, Y. K. Tun, and C. S. Hong, "Optimized deployment of multi-UAV based on machine learning in UAV-hst networking," in *2020 21st Asia-Pacific Network Operations and Management Symposium (APNOMS)*, pp. 102–107, IEEE, 2020.

[10] Y. Nie, J. Zhao, J. Liu, J. Jiang, and R. Ding, "Energy-efficient UAV trajectory design for backscatter communication: A deep reinforcement learning approach," *China communications*, vol. 17, no. 10, pp. 129–141, 2020.

[11] W. Guo, "Partially explainable big data driven deep reinforcement learning for green 5G UAV," in *ICC 2020-2020 IEEE International Conference on Communications (ICC)*, pp. 1–7, IEEE, 2020.

[12] Y. Zeng and R. Zhang, "Energy-efficient UAV communication with trajectory optimization," *IEEE Transactions on Wireless Communications*, vol. 16, no. 6, pp. 3747–3760, 2017.

[13] C. H. Liu, Z. Chen, J. Tang, J. Xu, and C. Piao, "Energy-efficient UAV control for effective and fair communication coverage: A deep reinforcement learning approach," *IEEE Journal on Selected Areas in Communications*, vol. 36, no. 9, pp. 2059–2070, 2018.

[14] A. Carrio, C. Sampedro, A. Rodriguez-Ramos, and P. Campoy, "A review of deep learning methods and applications for unmanned aerial vehicles," *Journal of Sensors*, vol. 2017, 2017.

[15] A. Kaushik, M. Alizadeh, O. Waqar, and H. Tabassum, "Deep unsupervised learning for generalized assignment problems: A case-study of user-association in wireless networks," in *2021 IEEE International Conference on Communications Workshops (ICC Workshops)*, pp. 1–6, 2021.

[16] A. Mehmood, O. Waqar, and M. ur Rahman, "Throughput maximization of an IRS-assisted wireless powered network with interference: A deep unsupervised learning approach," 2021 (Available on: ArXiv).