Quantum Reinforcement Learning: Concepts, Models, and Applications

Joongheon Kim

Korea University, Seoul, Republic of Korea
joongheon@korea.ac.kr
https://joongheon.github.io

Abstract. This tutorial presents the fundamental theory of quantum reinforcement learning (QRL) and its emerging applications. Thanks to the recent evolution in quantum computing, numerous research results have been proposed for the development of theories and applications of QRL. In this tutorial paper, the advantages and benefits of QRL-based models and algorithms are also discussed, i.e., fast learning convergence, high action-dimension scalability, and efficient training parameter utilization. Based on the design concept and advantages, QRL-based models and algorithms are used for various emerging applications. Among them, this paper introduces the applications of QRL-based models and algorithms in terms of future multimedia systems, autonomous mobility services, and distributed computing platforms.

Keywords: Quantum Computing, Quantum Reinforcement Learning, Reinforcement Learning Applications

1 Introduction

In modern computing and communication research trends, that there have been a lot of attentions and interests in quantum algorithms and computation related technologies, such as quantum communications [1,2], quantum optimization [3–5], quantum visual learning via quantum convolutional neural networks [6–8], quantum recurrent neural networks [9], quantum distributed learning [10–13], and neural network design with quantum computing concepts [14–17]. Among the various technical research topics in quantum algorithms and computation, the research on quantum reinforcement learning has been also widely and extensively conducted due to its possibility for the applications to various mobile and networking systems [18–26].

To enable the design and implementation of QRL algorithms and models, the advantages of the use of QRL-based algorithms can be fully utilized, i.e., fast learning convergence, high action-dimension scalability, and efficient training parameter utilization. These advantages can be definitely one of major reasons for facilitating QRL algorithms and models in various emerging engineering applications.

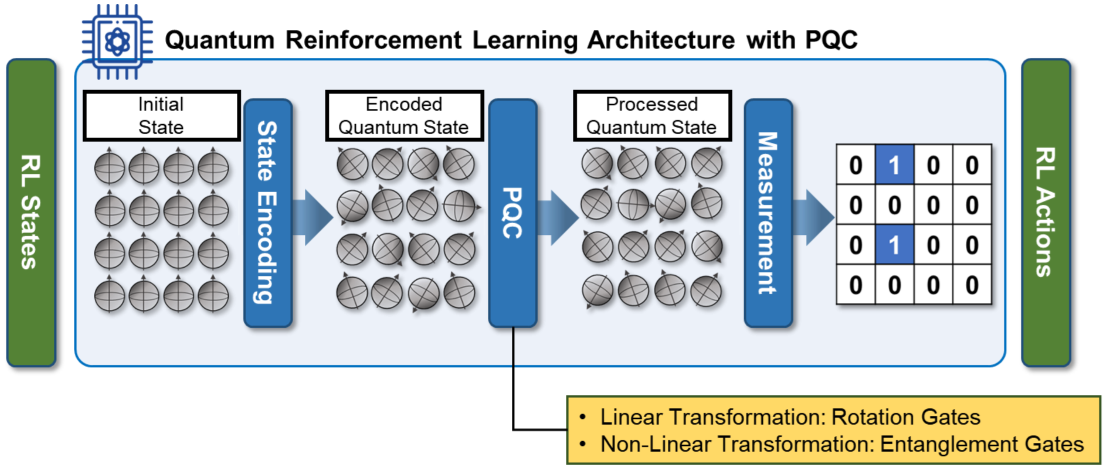

Fig. 1: Reference model for QRL-based algorithms.

Moreover, the extensions of QRL models to multi-agent cooperation are also essentially required for modern distributed multi-device autonomous control. Therefore, it is also worthy to consider multi-agent extensions of QRL-based algorithms and models. These algorithms and models are for realizing multi-agent reinforcement learning theory and applications while taking care of the limitations to modern quantum computing platforms under modern noisy intermediate-scale quantum (NISQ) era [26].

Lastly, this paper presents the major emerging mobility and networking applications of QRL-based algorithms and models, such as future multimedia systems (e.g., metaverse), autonomous mobility services with unmanned aerial vehicles (UAVs), and distributed computing platforms (e.g., federated learning and split learning).

The rest of this paper is organized as follows. Firstly, Sec. 2 introduces the basic theory of QRL-based algorithms. In addition, Sec. 3 presents the major potential applications of QRL-based algorithms. Lastly, Sec. 4 concludes this paper.

2 Quantum Reinforcement Learning Theory

One of the majorly considered QRL models, which is called variational quantum circuit (VQC), consists of state encoder, parameterized quantum circuit (PQC), and measurement, as presented in Fig. 1 [20, 27, 31].

- **State Encoder.** The state encoder transforms classical input state data into encoded quantum states because classical input state data cannot be optimally utilized with quantum gates and circuits. In our considering QRL-based algorithms and models, this state encoder is organized by quantum rotation gates.

- **Parameterised Quantum Circuit (PQC).** After completing the state encoder as the first part of QRL-based algorithms and models, the quantum state information produced by the state encoder reaches PQC which is the core part for imitating the hidden layer functionalities and characteristics of classical deep neural network (DNN)-based models. In this PQC, training is performed on top of the quantum data by rotation quantum gates which conduct phase rotations on the quantum state over x-, y-, and z-axes in 3D Bloch Hilbert sphere in order to represent linear function approximation. Moreover, entanglement gates can be also used in order to represent non-linear function approximation by causing quantum entanglement between two or more qubits using controlled NOT (CNOT) gates. This concept is equivalent to the multiple hidden layer and activation function computation in classical DNN-based reinforcement learning models.
- **Measurement.** After completing state encoder and PQC, the measurement procedure should be conducted for quantum computing algorithm computation. It serves as the counterpart to the output layer in classical DNN-based reinforcement learning models. Before the measurement, quantum states probabilistically exist according to superposition-stated qubits. As a result, quantum data cannot be used for computation unless the qubits are deterministically stabilized which can be achieved by measuring the quantum states which are equivalent to the projection computation on quantum states.

The potential novelties and advantages of QRL-based models and algorithms comparing to classical DNN-based reinforcement learning models are as follows [20].

- **Fast Learning Convergence.** In QRL-based models and algorithms, the *parameter shift rule* is used for loss function optimization by conducting parameter optimization during training. This *parameter shift rule* offers simplistic and direct approaches under the characteristics of qubits and quantum computing. Therefore, the training/learning can be accelerated in QRL-based models and algorithms.
- **High Action-Dimension Scalability.** In the major problems of modern reinforcement learning theory and applications in real-worlds, there are many issues where they have to deal with the huge dimensions of sophistic actions in each agent. However, the efficiency of reinforcement learning models and algorithms can be significantly compromised by the high-dimensionality dilemma. Thus, the convergence cannot be achieved when the dimensions of reinforcement learning actions are huge. In order to tackle the issue, the method which is called *basis measurement* is proposed in order to provide promising solutions [27]. The basis measurement has the potential to simplify the decision-making procedure of QRL-based agents by reducing its action dimensions into a logarithmic-scale [27]. Therefore, it effectively relaxes the related computational overheads and enhances efficiency by avoiding potential high qubit error rates due to large-scale qubit utilization.

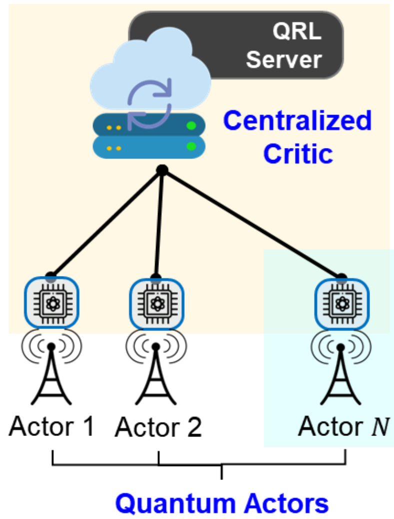

Fig. 2: Reference model for QRL-based actor-critic algorithms.

– **Efficient Training Parameter Utilization.** QRL-based models and algorithms are capable for achieving similar performance comparing to DNN-based reinforcement learning models and algorithms with fewer training parameters, according to the nature of superposition in quantum computing and algorithms. In the quantum computing and algorithms, the superposition allows the QRL-based models and algorithms exist in multiple states at once, and thus, it expands the representational capacity of QRL-based models and algorithms.

In addition to PQC-based QRL models and algorithms, it is required to extend for multi-agent reinforcement learning computation and cooperation. As illustrated in Fig. 2, a QRL-based centralized-critic and multiple-actors model can be one of potential models for realizing multi-agent reinforcement learning computation [25, 28]. The fundamental learning architecture of our proposed QRL-based centralized-critic and multiple-actors model is originated from conventional actor-critic network [29], where the actor and critic models are designed based on QRL-based models. Here, one of the major advantages of QNN-based design and implementation in actor/critic models is fundamentally based on the fact that the iterative training/learning for the models is obviously faster than conventional DNN-based models and algorithms. Hence, training multiple actor models can be done in a given time. Moreover, this QRL-based centralized-critic and multiple-actors model allows to design multiple QRL-based models where the input size is for single agent. Therefore, it avoids the situation to design large-scale QRL-based models and algorithms. Furthermore, if only one actor network is utilized in our considering QRL-based centralized-critic and multiple-

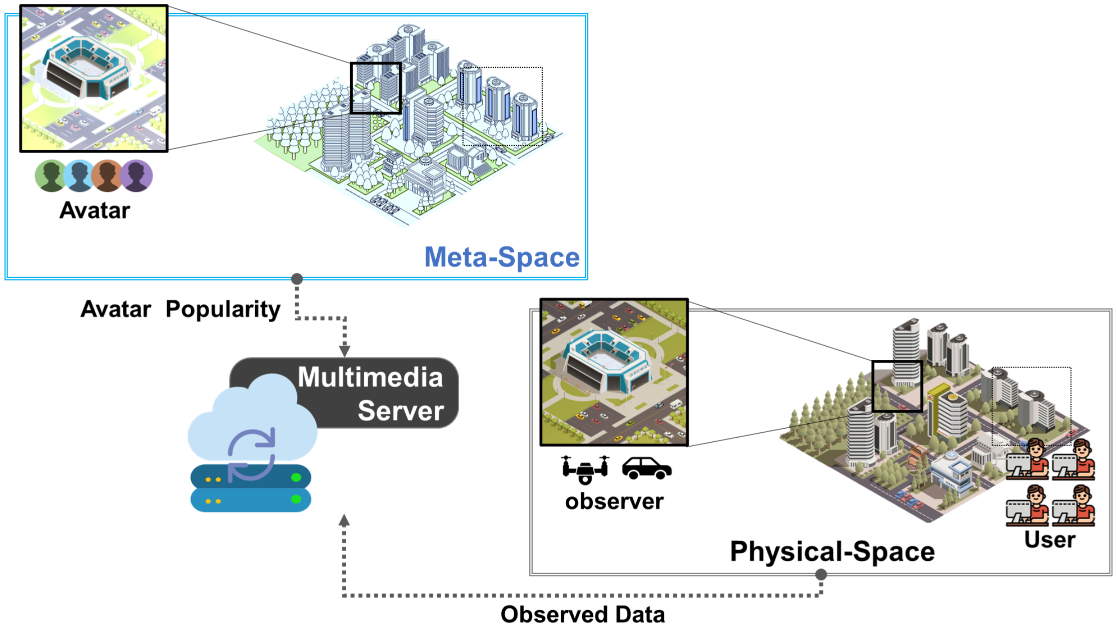

Fig. 3: Reference model for the separation between meta-space and physical-space [19].

actors model, it can directly be the model of single-agent reinforcement learning computation.

Lastly, this QRL-based centralized-critic and multiple-actors model has been used for emerging reinforcement learning applications such as multi-robot coordination in smart factory applications [23] and multi-UAV cooperation in mobile network access applications [22].

3 Applications

Our considering QRL-based models and QRL-based centralized-critic and multiple-actors models have been widely and frequently used for many emerging mobility and networking applications, such as future multimedia systems (refer to Sec. 3.1), autonomous mobility services (refer to Sec. 3.2), and distributed computing platforms (refer to Sec. 3.3).

3.1 Future Multimedia Systems

In future multimedia systems such as metaverse [19, 21] and digital-twin [30], the interests of users are in their meta-space regions whereas the users are physically located in their physical-spaces. Therefore, the main objective should be the meta-space utility maximization subject to physical-space constraints, as illustrated in Fig. 3. Furthermore, massive virtual identities (named to avatars in Fig. 3) can be located in meta-space which can introduce massive action-dimension (i.e., scheduling dimension) computation. Therefore, quantum-based control/scheduling algorithms such as our proposed QRL-based models and QRL-based centralized-critic and multiple-actors models are essentially required.

3.2 Autonomous Mobility Services

In autonomous mobility services, as mentioned in the previous section, our proposed QRL-based centralized-critic and multiple-actors model has been used for multi-robot coordination in smart factory applications [23] and multi-UAV cooperation in mobile network access applications [22]. The used of these our proposed algorithms is beneficial because it can reduce action-dimensions which is essentially required for high-resolution action control. Furthermore, fewer learning parameter utilization is also beneficial for light-weight UAV/drone and embedded robot platforms.

3.3 Distributed Computing Platforms

For distributed computing platforms and applications, our proposed QRL-based centralized-critic and multiple-actors model has been widely used for many applications. This can be also used for quantum federated learning [31] and quantum split learning [10]. Especially, as one of major distributed computing applications, satellite networks can utilize quantum federated learning based on the advantages in QRL-based models and algorithms, as illustrated in Fig. 4.

4 Concluding Remarks and Future work

This paper introduces the theory of quantum reinforcement learning (QRL) and its emerging mobility and networking applications. In addition, the advantages of QRL-based models are also discussed, i.e., fast learning convergence, high action-dimension scalability, and efficient training parameter utilization. Based on the theory and advantages, QRL-based models and algorithms can be utilized for various emerging applications suc has future multimedia systems, autonomous mobility services, and distributed computing platforms.

As future work, it should be worthy to consider data-intensive performance evaluation with various scenarios and settings.

Acknowledgement

This work was supported in part by the National Research Foundation of Korea under Grant 2022R1A2C2004869, and in part by the Institute of Information and Communications Technology Planning and Evaluation (IITP) grant funded

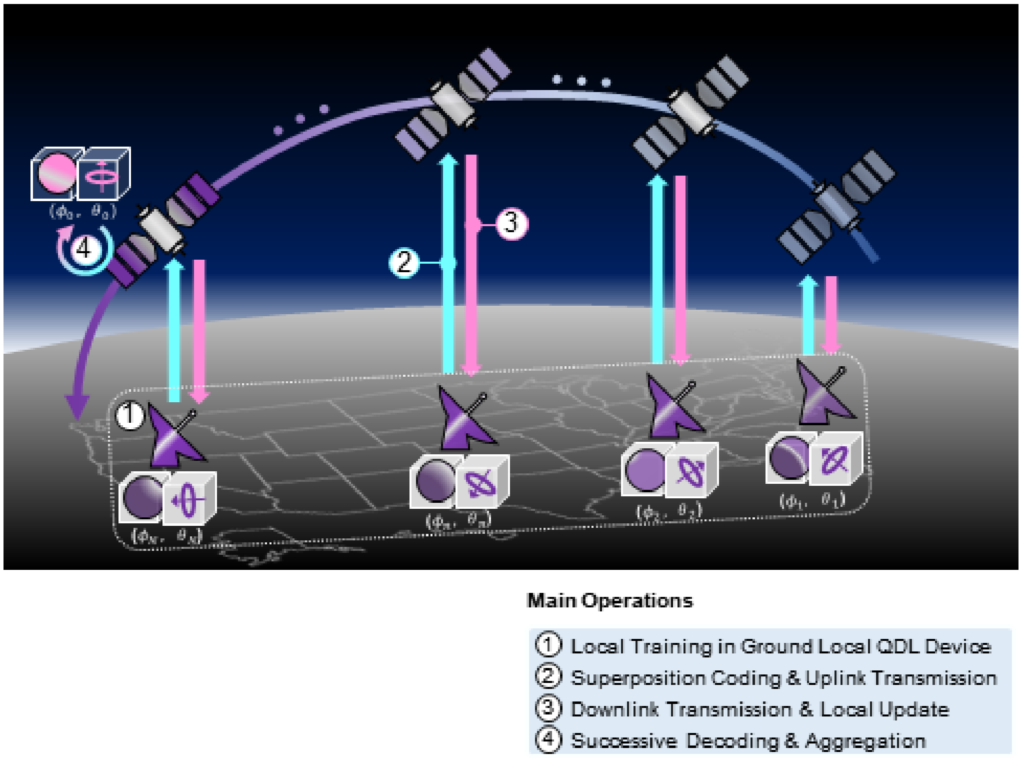

Fig. 4: Reference model for distributed quantum federated learning in satellite networks [31].

by Korean Government [Ministry of Science and Information and Communications Technology (MSIT)] (Intelligent 6G Wireless Access System) under Grant 2021-0-00467. Joongheon Kim is a corresponding author of this paper (e-mail: `joongheon@korea.ac.kr`, ORCID: `https://orcid.org/0000-0003-2126-768X`).

References

1. Jiang, X., Itzler, M., O'Donnell, K., Entwistle, M., Owens, M., Slomkowski, K., Rangwala, S.: InP-Based Single-Photon Detectors and Geiger-Mode APD Arrays for Quantum Communications Applications. IEEE Journal of Selected Topics in Quantum Electronics. 21, 5-16 (2015). https://doi.org/10.1109/JSTQE.2014.2358685
2. Fernandez, V., Gómez-García, J., Ocampos-Guillén A., Carrasco-Casado, A. Correction of Wavefront Tilt Caused by Atmospheric Turbulence Using Quadrant Detectors for Enabling Fast Free-Space Quantum Communications in Daylight. IEEE Access. 6, 3336-3345 (2018). https://doi.org/10.1109/ACCESS.2018.2791099
3. Kim, J., Kwak, Y., Jung, S., Kim, J.-H.: Quantum Scheduling for Millimeter-Wave Observation Satellite Constellation. In: Proceedings of the IEEE VTS Asia-

Pacific Wireless Communications Symposium (APWCS), pp. 1-5 (2021). https://doi.org/10.1109/APWCS50173.2021.9548770
4. Choi, J., Oh, S., Kim, J.: The Useful Quantum Computing Techniques for Artificial Intelligence Engineers. In: Proceedings of the IEEE International Conference on Information Networking (ICOIN), pp. 1-3 (2020). https://doi.org/10.1109/ICOIN48656.2020.9016555
5. Choi, J., Kim, J.: A Tutorial on Quantum Approximate Optimization Algorithm (QAOA): Fundamentals and Applications. In: Proceedings of the IEEE International Conference on Information and Communication Technology Convergence (ICTC), pp. 138-142 (2019). https://doi.org/10.1109/ICTC46691.2019.8939749
6. Baek, H., Yun, W.J., Park, S., Kim, J.: Stereoscopic Scalable Quantum Convolutional Neural Networks. Neural Networks. 165, 860-867 (2023). https://doi.org/10.1016/j.neunet.2023.06.027
7. Oh, S., Choi, J., Kim, J.: A Tutorial on Quantum Convolutional Neural Networks (QCNN). In: Proceedings of the IEEE International Conference on Information and Communication Technology Convergence (ICTC), pp. 236-239 (2020). https://doi.org/10.1109/ICTC49870.2020.9289439
8. Oh, S., Choi, J., Kim, J.-K., Kim, J.: Quantum Convolutional Neural Network for Resource-Efficient Image Classification: A Quantum Random Access Memory (QRAM) Approach. In: Proceedings of the IEEE International Conference on Information Networking (ICOIN), pp. 50-52 (2021). https://doi.org/10.1109/ICOIN50884.2021.9333906
9. Choi, J., Oh, S., Kim, J.: A Tutorial on Quantum Graph Recurrent Neural Network (QGRNN). In: Proceedings of the IEEE International Conference on Information Networking (ICOIN), pp. 46-49 (2021). https://doi.org/10.1109/ICOIN50884.2021.9333917
10. Park, S., Baek, H., Kim, J.: Quantum Split Learning for Privacy-Preserving Information Management. In: Proceedings of the ACM Conference on Information and Knowledge Management (CIKM), pp. 4239-4243 (2023). https://doi.org/10.1145/3583780.3615144
11. Kwak, Y., Yun, W.J., Kim, J.P., Cho, H., Park, J., Choi, M., Jung, S., Kim, J.: Quantum Distributed Deep Learning Architectures: Models, Discussions, and Applications. ICT Express. 9, 486-491 (2023). https://doi.org/10.1016/j.icte.2022.08.004
12. Kim, J.: Quantum Federated Learning for Vehicular Computing Scenarios. In: Proceedings of the IEEE International Conference on Information and Communication Technology Convergence (ICTC), pp. 168-172 (2023). https://doi.org/10.1109/ICTC58733.2023.10392753
13. Kim, J.P., Yun, W.J., Baek, H., Kim, J.: Modern Trends in Quantum AI: Distributed and High-Definition Computation. In: Proceedings of the IEEE International Conference on Information Networking (ICOIN), pp. 750-754 (2023). https://doi.org/10.1109/ICOIN56518.2023.10048995
14. Kwak, Y., Yun, W.J., Jung, S., Kim, J: Quantum Neural Networks: Concepts, Applications, and Challenges. In: Proceedings of the IEEE International Conference on Ubiquitous Future Networks (ICUFN), pp. 413-416 (2021). https://doi.org/10.1109/ICUFN49451.2021.9528698
15. Park, S., Baek, H., Yoon, J.W., Lee, Y.K., Kim, J.: AQUA: Analytics-driven Quantum Neural Network (QNN) User Assistance for Software Validation. Future Generation Computer Systems. 159, 545-556 (2024). https://doi.org/10.1016/j.future.2024.05.047

16. Park, S., Feng, H., Park, C., Lee, Y.K., Jung, S., Kim, J.: EQuaTE: Efficient Quantum Train Engine for Run-Time Dynamic Analysis and Visual Feedback in Autonomous Driving. IEEE Internet Computing. 27, 24-31 (2023). https://doi.org/10.1109/MIC.2023.3307395
17. Park, S., Feng, H., Yun, W.J., Park, C., Lee, Y.K., Jung, S., Kim, J.: Demo: EQuaTE: Efficient Quantum Train Engine Design and Demonstration for Dynamic Software Analysis. In: Proceedings of the IEEE International Conference on Distributed Computing Systems (ICDCS), pp. 1009-1012 (2023). https://doi.org/10.1109/ICDCS57875.2023.00116
18. Kwak, Y., Yun, W.J., Jung, S., Kim, J.-K., Kim, J.: Introduction to Quantum Reinforcement Learning: Theory and PennyLane-based Implementation. In: Proceedings of the IEEE International Conference on Information and Communication Technology Convergence (ICTC), pp. 416-420 (2021). https://doi.org/10.1109/ICTC52510.2021.9620885
19. Park, S., Chung, J., Park, C., Jung, S., Choi, M., Cho, S., Kim, J.: Joint Quantum Reinforcement Learning and Stabilized Control for Spatio-Temporal Coordination in Metaverse. IEEE Transactions on Mobile Computing (Early Access). https://doi.org/10.1109/TMC.2024.3407883
20. Park, S., Kim, J.P., Park, C., Jung, S., Kim, J.: Quantum Multi-Agent Reinforcement Learning for Autonomous Mobility Cooperation. IEEE Communications Magazine. 62, 106-112 (2024). https://doi.org/10.1109/MCOM.020.2300199
21. Park, S., Baek, H., Kim, J.: Quantum Reinforcement Learning for Spatio-Temporal Prioritization in Metaverse. IEEE Access. 12, 54732-54744 (2024). https://doi.org/10.1109/ACCESS.2024.3390042
22. Park, C., Yun, W.J., Kim, J.P., Park, S., Rodrigues, T.K., Jung, S., Kim, J.: Quantum Multi-Agent Actor-Critic Networks for Cooperative Mobile Access in Multi-UAV Systems. IEEE Internet of Things Journal. 10, 20033-20048 (2023). https://doi.org/10.1109/JIOT.2023.3282908
23. Yun, W.J., Kim, J.P., Jung, S., Kim, J.-H., Kim, J.: Quantum Multiagent Actor-Critic Neural Networks for Internet-Connected Multirobot Coordination in Smart Factory Management. IEEE Internet of Things Journal. 10, 9942-9952 (2023). https://doi.org/10.1109/JIOT.2023.3234911
24. Yun, W.J., Park, J., Kim, J.: Quantum Multi-Agent Meta Reinforcement Learning. In: Proceedings of the AAAI Conference on Artificial Intelligence, pp. 11087–11095 (2023). https://doi.org/10.1609/aaai.v37i9.26313
25. Yun, W.J., Kwak, Y., Kim, J.P., Cho, H., Jung, S., Park, J., Kim, J.: Quantum Multi-Agent Reinforcement Learning via Variational Quantum Circuit Design. In: Proceedings of the IEEE International Conference on Distributed Computing Systems (ICDCS), pp. 1332-1335 (2022). https://doi.org/10.1109/ICDCS54860.2022.00151
26. Park, S., Kim, J.: Quantum Reinforcement Learning for Large-Scale Multi-Agent Decision-Making in Autonomous Aerial Networks. In: Proceedings of the IEEE VTS Asia-Pacific Wireless Communications Symposium (APWCS), pp. 1-4 (2023). https://doi.org/10.1109/APWCS60142.2023.10233966
27. Baek, H., Park, S., Kim, J.: Logarithmic Dimension Reduction for Quantum Neural Networks. In: Proceedings of the ACM Conference on Information and Knowledge Management (CIKM), pp. 3738–3742 (2023). https://doi.org/10.1145/3583780.3615240
28. Yun, W.J., Park, S., Kim, J., Shin, M., Jung, S., Mohaisen, D.A., Kim, J.-H.: Cooperative Multiagent Deep Reinforcement Learning for Reliable Surveillance via

Autonomous Multi-UAV Control. IEEE Transactions on Industrial Informatics. 18, 7086-7096 (2022). https://doi.org/10.1109/TII.2022.3143175
29. Barto, A.G., Sutton, R.S., Anderson, C.W.: Looking Back on the Actor-Critic Architecture. IEEE Transactions on Systems, Man, and Cybernetics: Systems. 51, 40-50 (2021). https://doi.org/10.1109/TSMC.2020.3041775
30. Jung, S., Baek, H., Kim, J.: Neural Myerson Auction for Truthful and Distributed Mobile Charging in UAV-Assisted Digital-Twin Networks. IEEE Transactions on Intelligent Vehicles (Early Access). https://doi.org/10.1109/TIV.2024.3396556
31. Park, S., Jung, S., Kim, J.: Dynamic Quantum Federated Learning for Satellite-Ground Integrated Systems using Slimmable Quantum Neural Networks. IEEE Access. 12, 58239-58247 (2024). https://doi.org/10.1109/ACCESS.2024.3392429