关于提案的声明

该提案最初由李元健博士发起和撰写，旨在申请 Virgina Tech’s Presidential Postdoctoral Fellowship。该提案已于 2022 年底提交至委员会审批，尽管未被评审团选中，我保留对此提案所含内容的全部版权和知识产权。任何使用、复制或传播此提案或其任何部分的行为，须注明李元健博士为原作者。李元健博士保留所有权利。

Author’s Note on Proposal

This proposal was originally initiated and authored by Dr. Yuanjian Li for the Royal Commission for Virgina Tech’s Presidential Postdoctoral Fellowship. It was submitted to the committee in late 2022 as part of my application for the fellowship. While the proposal was not selected by the panel, I retain full copyright and intellectual ownership over the content presented herein. Any use, reproduction, or distribution of this proposal or any part thereof must acknowledge Dr. Yuanjian Li as the original author. All rights are reserved.

Distributed, Quantum Multi-Agent Deep Reinforcement Learning for Wireless 6G Systems

Objectives. The sixth-generation (6G) wireless cellular system will be an *artificial intelligence (AI)-native network* in which most of the protocol stack is designed using data-machine learning (ML) techniques as envisioned by industry, academia, and standardization bodies. For example, deep reinforcement learning (DRL) and multi-agent DRL (MADRL) solutions are being considered to address various 6G problems [1–4] ranging from transceiver design to resource management. Distributed MADRL solutions particularly have several benefits for designing AI-native 6G systems that range from the possibility of deploying them at the edge of the wireless network to their inherent ability to accommodate agents with heterogeneous capabilities. However, existing DRL and MADRL schemes [1–4] have limited performance in terms of robustness, latency, and efficiency. Moreover, they cannot handle non-stationary and partially observable environments, and they face major complexity, overhead, and computational challenges due to their limited scalability to scenarios with large action spaces or large number of devices and their reliance on updating a large number of deep neural network (DNN) parameters. To overcome these challenges, we propose a *novel framework of quantum MADRL (QMADRL)* that combines the benefits of quantum mechanics with those of distributed MADRL, and that will be *designed to address major wireless 6G challenges*. Recent works on quantum machine learning (QML) [5, 6] showed its benefits for improving efficiency and enhancing ML generalization. Several works [7–12] merged quantum mechanics and DRL to boost the learning efficiency and robustness of DRL. For example, comparable or better learning performance with much lighter parameter update of the quantum DRL algorithm was shown in [7], compared to DNN-based DRL. A *fundamental question* is whether one can build a novel QMADRL framework that can be used to design AI-native 6G systems, with low latency and high reliability. Although there are some recent works on QRL in [7, 11–14], these works do not leverage quantum DRL to address wireless 6G problems, and they mostly rely on single agent or they make assumptions that are impractical for real 6G systems, e.g., quantum errors were not considered [14] and full observability of environment was naively adopted [15].

The *goal of this project* is thus to lay the theoretical foundations of *distributed QMADRL* for radio resource coordination in *AI-native wireless 6G* systems. First, we will develop new distributed QMADRL algorithms tailored towards solving complex optimization problems in 6G networks, such as problems of radio resource management, e.g., beamforming design, spectrum access, and resource (energy, bandwidth) allocation. Decentralized actors will be redesigned with variational quantum circuit (VQC) including state encoding as well as parametric quantum circuit (PQC) and measurements, while VQC-based critic can remain centralized for dealing with non-stationarity. Here, we will investigate the impact of various PQC ansatzes. For QMADRL, we will study the signaling overhead and information sharing among distributed agents, and we will theoretically and empirically show how and when quantum designs can reduce such overhead, while identifying its impact on the learning in several 6G use cases. Our designs will incorporate practical consideration on partial observability of complex 6G wireless environments. The application of VQC-aided training architecture for QMADRL can significantly reduce the number of training parameters, which could help overcome the scalability and latency challenges of DNN-based MADRL. VQC can be implemented on conventional computers, with the help of related libraries, e.g., PennyLane, IBM Qiskit, Google Cirq, and Torch Quantum. Next, we will leverage promising techniques such as space compression [16], hypernetwork [17], and meta-learning [18] to enhance the robustness and scalability of the proposed QMADRL for large-scale and heterogeneous 6G networks. We will investigate practical implementation considerations for our QMADRL. For instance, in the current noise intermediate-scale quantum (NISQ) era, quantum errors caused by quantum decoherence and imprecision of quantum gates will inevitably jeopardize the learning performance of VQC-based schemes. Hence, we will explicitly investigate the *fundamental effects of quantum noise on our QMADRL* solutions, and we will propose efficient strategies (e.g., hybrid optimization [19] or ensemble learning [20]) to counter this shortcoming. Finally, we will develop quantum-inspired MADRL schemes in which principles from quantum physics are used to aid classical MADRL action selection and experience replay, building on Dr. Li's prior works [21–23]. In short, this research will *contribute simultaneously to AI, quantum computing, and wireless 6G systems* thus providing scientific foundations for the promising interdisciplinary area of QMADRL applied to 6G systems.

Fellow Qualifications. Dr. Li has a unique expertise that spans DRL, wireless systems, and quantum-aided ML with an impressive record of over 10 publications in these areas [21–33], most of which appeared in top-tier venues such as the IEEE Transactions on Wireless Communications (the most prestigious wireless journal). Beyond designing novel DRL solutions for wireless systems in [21–25], during his PhD at the prestigious King's College London (UK), Dr. Li developed some of the first quantum-aided DRL algorithms [21–23] for drone-assisted wireless networks. These results will be a key building block for this research. Dr. Li is also an expert in designing wireless network protocols as done in [26–33]. Hence, the candidate has a very distinguished blend of expertise, across AI, wireless networks, and quantum, that is rarely found in postdoc applicants and which uniquely qualifies him for this award.

Contribution to Fellow Career Goals. Dr. Li's career goal is to become a faculty member at a world-leading university, e.g., VT or a peer-like institution. This fellowship is an optimal m[redacted]r boosting chances of achieving his long-term aspiration. During this fellowship, Dr. Li will work with both Dr. [redacted]'s group, with expertise on quantum computing, and Dr. [redacted]'s group with expertise in AI and wireless networ s. T s is a very unique opportunity for allowing Dr. Li to gain extensive knowledge and forge multi-disciplinary expertise in quantum computing, distributed AI, and QMADRL solutions for empowering future 6G wireless systems. He will be encouraged to publish in top-tier venues for AI, quantum, and wireless, and to write grant proposals, which will be an invaluable skill. This training will uniquely qualify him for multiple faculty positions in different areas. This fellowship will also offer Dr. Li invaluable real-world expertise in leading research proposals, and it will help him expand his scholarship, track record and academic visibility by interacting with the large network of international and national collaborators (including those at Virginia Tech) of the two mentors. The mentors will hone Dr. Li's leadership skills and boost his visibility by encouraging him to lead collaborative projects and intiatives (e.g., workshops etc.).

Impact on Mentor and VT

[redacted]

[redacted]

[redacted]

[redacted]

Mentoring Plan

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

References

[1] J. Hu, H. Zhang, L. Song, R. Schober, and H. V. Poor, "Cooperative internet of UAVs: Distributed trajectory design by multi-agent deep reinforcement learning," *IEEE Trans. Commun.*, vol. 68, no. 11, pp. 6807–6821, Aug. 2020.

[2] M. Chen, U. Challita, W. Saad, C. Yin, and M. Debbah, "Artificial neural networks-based machine learning for wireless networks: A tutorial," *IEEE Commun. Surv. Tutor.*, vol. 21, no. 4, pp. 3039–3071, Jul. 2019.

[3] S. Talwar, N. Himayat, H. Nikopour, F. Xue, G. Wu, and V. Ilderem, "6G: Connectivity in the era of distributed intelligence," *IEEE Commun. Mag.*, vol. 59, no. 11, pp. 45–50, Dec. 2021.

[4] K. Zhang, Z. Yang, and T. Başar, "Multi-agent reinforcement learning: A selective overview of theories and algorithms," *Handbook of Reinforcement Learning and Control*, pp. 321–384, 2021.

[5] Y. Liu, S. Arunachalam, and K. Temme, "A rigorous and robust quantum speed-up in supervised machine learning," *Nat. Phys.*, vol. 17, no. 9, pp. 1013–1017, Sep. 2021.

[6] M. C. Caro, H.-Y. Huang, M. Cerezo, K. Sharma, A. Sornborger, L. Cincio, and P. J. Coles, "Generalization in quantum machine learning from few training data," *Nat. commun.*, vol. 13, no. 1, pp. 1–11, Aug. 2022.

[7] S. Y.-C. Chen, C.-H. H. Yang, J. Qi, P.-Y. Chen, X. Ma, and H.-S. Goan, "Variational quantum circuits for deep reinforcement learning," *IEEE Access*, vol. 8, pp. 141 007–141 024, Jul. 2020.

[8] D. Dong, C. Chen, H. Li, and T.-J. Tarn, "Quantum reinforcement learning," *IEEE Trans. Syst. Man. Cybern. B*, vol. 38, no. 5, pp. 1207–1220, Jul. 2008.

[9] D. Dong, C. Chen, J. Chu, and T.-J. Tarn, "Robust quantum-inspired reinforcement learning for robot navigation," *IEEE/ASME Trans. Mechatronics*, vol. 17, no. 1, pp. 86–97, Dec. 2010.

[10] P. Fakhari, K. Rajagopal, S. Balakrishnan, and J. Busemeyer, "Quantum inspired reinforcement learning in changing environment," *New Math. Nat. Comput.*, vol. 9, no. 03, pp. 273–294, Nov. 2013.

[11] V. Saggio, B. E. Asenbeck, A. Hamann, T. Strömberg, P. Schiansky, V. Dunjko, N. Friis, N. C. Harris, M. Hochberg, D. Englund *et al.*, "Experimental quantum speed-up in reinforcement learning agents," *Nature*, vol. 591, no. 7849, pp. 229–233, Mar. 2021.

[12] J.-A. Li, D. Dong, Z. Wei, Y. Liu, Y. Pan, F. Nori, and X. Zhang, "Quantum reinforcement learning during human decision-making," *Nat. Hum. Behav.*, vol. 4, no. 3, pp. 294–307, Mar. 2020.

[13] D. Heimann, H. Hohenfeld, F. Wiebe, and F. Kirchner, "Quantum deep reinforcement learning for robot navigation tasks," *arXiv preprint arXiv:2202.12180*, 2022.

[14] Y. Xu, R. Yan, Y. Wang, and D. Jiahong, "A multi-agent quantum deep reinforcement learning method for distributed frequency control of islanded microgrids," *IEEE Trans. Control. Netw. Syst.*, Jan. 2022.

[15] W. J. Yun, Y. Kwak, J. P. Kim, H. Cho, S. Jung, J. Park, and J. Kim, "Quantum multi-agent reinforcement learning via variational quantum circuit design," *arXiv preprint arXiv:2203.10443*, 2022.

[16] D. Yin, Z. Zhao, Y. Dai, and H. Long, "A novel multi-agent deep reinforcement learning approach," in *J. Phys. Conf. Ser.*, vol. 1757, no. 1, 2021, p. 012097.

[17] M. Goutay, F. A. Aoudia, and J. Hoydis, "Deep hypernetwork-based MIMO detection," in *Proc. IEEE Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, May 2020, pp. 1–5.

[18] T. Hospedales, A. Antoniou, P. Micaelli, and A. Storkey, "Meta-learning in neural networks: A survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 9, pp. 5149–5169, May 2021.

[19] L. Gentini, A. Cuccoli, S. Pirandola, P. Verrucchi, and L. Banchi, "Noise-resilient variational hybrid quantum-classical optimization," *Phys. Rev. A*, vol. 102, no. 5, p. 052414, Nov. 2020.

[20] D. Silver, T. Patel, and D. Tiwari, "Quilt: Effective multi-class classification on quantum computers using an ensemble of diverse quantum classifiers," in *Pro. AAAI Conf. Artif. Intell.*, vol. 36, no. 8, Jun. 2022, pp. 8324–8332.

[21] Y. Li, A. H. Aghvami, and D. Dong, "Path planning for cellular-connected UAV: A DRL solution with quantum-inspired experience replay," *IEEE Trans. Wireless Commun.*, vol. 21, no. 10, pp. 7897–7912, Apr. 2022.

[22] ——, "Intelligent trajectory planning in UAV-mounted wireless networks: A quantum-inspired reinforcement learning perspective," *IEEE Wireless Commun. Lett.*, vol. 10, no. 9, pp. 1994–1998, Jun. 2021.

[23] Y. Li and A. H. Aghvami, "Intelligent UAV navigation: A DRL-QiER solution," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Seoul, South Korea, May 2022, pp. 1–6.

[24] ——, "Covertness-aware trajectory design for UAV: A multi-step TD3-PER solution," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Seoul, South Korea, May 2022, pp. 1–6.

[25] D. Ouyang, R. Zhao, Y. Li, R. Guo, and Y. Wang, "Antenna selection in energy harvesting relaying networks using Q-learning algorithm," *China Commun.*, vol. 18, no. 4, pp. 64–75, Apr. 2021.

[26] Y. Li, R. Zhao, Y. Deng, F. Shu, Z. Nie, and A. H. Aghvami, "Harvest-and-opportunistically-relay: Analyses on transmission outage and covertness," *IEEE Trans. Wireless Commun.*, vol. 19, no. 12, pp. 7779–7795, Aug. 2020.

[27] Y. Li, R. Zhao, Y. Wang, G. Pan, and C. Li, "Artificial noise aided precoding with imperfect CSI in full-duplex relaying secure communications," *IEEE Access*, vol. 6, pp. 44 107–44 119, Jul. 2018.

[28] Y. Li, R. Zhao, L. Fan, and A. Liu, "Antenna mode switching for full-duplex destination-based jamming secure transmission," *IEEE Access*, vol. 6, pp. 9442–9453, Jan. 2018.

[29] Y. Li, R. Zhao, X. Tan, and Z. Nie, "Secrecy performance analysis of artificial noise aided precoding in full-duplex relay systems," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Singapore, Dec. 2017, pp. 1–6.

[30] D. Ouyang, R. Zhao, and Y. Li, "Analysis and optimization of wireless powered untrusted relay system with multiple destinations," *Phys. Commun.*, vol. 42, p. 101161, Oct. 2020.

[31] Z. Nie, R. Zhao, and Y. Li, "A full-duplex SWIPT relaying protocol based on discrete energy state," in *Proc. IEEE Int. Symp. Wireless Personal Multimedia Commun. (WPMC)*, Bali, Dec. 2017, pp. 500–505.

[32] X. Tan, R. Zhao, and Y. Li, "Large-scale antennas analysis of untrusted relay system with cooperative jamming," in *Proc. IEEE Int. Conf. Netw. Serv. Manag. (CNSM)*, Tokyo, Japan, Nov. 2017, pp. 1–5.

[33] D. Ouyang, R. Zhao, and Y. Li, "Wireless energy harvesting relaying networks combined with antenna selection," in *Proc. IEEE Int. Symp. Wireless Personal Multimedia Commun. (WPMC)*, Lisbon, Portugal, Nov. 2019, pp. 1–5.

[34] [redacted]

[35] [redacted]

[36] [redacted]

[37] [redacted]

[38] [redacted]

[39] [redacted]

[40] [redacted]

[41] [redacted]

[42] [redacted]

[43] [redacted]

[44] [redacted]

[45] [redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

李元健，伦敦国王学院博士

分布式量子多智能体深度强化学习用于无线 6G 系统

背景 第六代 (6G) 无线蜂窝系统将成为人工智能 (AI) 原生网络，其大部分协议栈将使用数据-机器学习 (ML) 技术进行设计，这是行业、学术界和标准化机构所设想的。例如，深度强化学习 (DRL) 和多智能体 DRL (MADRL) 解决方案正在被考虑用于解决从收发器设计到资源管理的各种 6G 问题 [1-4]。分布式 MADRL 解决方案在设计 AI 原生 6G 系统方面具有多种优势，这些优势包括在无线网络边缘部署它们的可能性以及它们能够容纳异构智能体的固有能力。然而，现有的 DRL 和 MADRL 方案 [1-4] 在鲁棒性、延迟和效率方面的表现有限。此外，它们无法处理非平稳和部分可观察的环境，并且由于它们依赖于更新大量深度神经网络 (DNN) 参数，面对大动作空间或大量设备时，还面临复杂性、开销和计算挑战。为了克服这些挑战，我们提出了一种新的量子 MADRL (QMADRL) 框架，该框架结合了量子力学和分布式 MADRL 的优势，并将设计用于解决主要的无线 6G 挑战。最近的量子机器学习 (QML) 研究 [5, 6] 显示了其在提高效率和增强 ML 泛化能力方面的优势。若干研究 [7-12] 将量子力学和 DRL 结合起来，以提升 DRL 的学习效率和鲁棒性。例如，在 [7] 中，展示了量子 DRL 算法相比基于 DNN 的 DRL，参数更新量更少且学习表现相当或更好。一个根本性的问题是，是否能够构建一个新的 QMADRL 框架，该框架可以用于设计具有低延迟和高可靠性的 AI 原生 6G 系统。虽然最近在 [7, 11-14] 中有一些关于量子强化学习 (QRL) 的研究，但这些工作没有利用量子 DRL 来解决无线 6G 问题，且大多依赖于单智能体，或者做出了在真实 6G 系统中不切实际的假设，例如，量子误差未被考虑 [14]，且天真地采用了环境的完全可观察性 [15]。

本项目的目标是为 AI 原生无线 6G 系统中的无线资源协调奠定分布式 QMADRL 的理论基础。首先，我们将开发新型分布式 QMADRL 算法，专门用于解决 6G 网络中的复杂优化问题，例如无线资源管理问题，如波束成形设计、频谱访问和资源（能量、带宽）分配。分布式智能体将使用变分量子电路 (VQC) 进行设计，包括状态编码以及参数化量子电路 (PQC) 和测量，而 VQC 基于的评论者可以保持集中以应对非平稳性。在这里，我们将研究各种 PQC 插入符的影响。对于 QMADRL，我们将研究分布式智能体之间的信号开销和信息共享，并从理论上和实验证明量子设计如何以及何时能够减少此类开销，同时识别其对多个 6G 用例中学习的影响。我们的设计将结合复杂 6G 无线环境部分可观察性的实际考虑。 VQC 辅助的训练架构应用于 QMADRL 可以显著减少训练参数的数量，这有助于克服基于 DNN 的 MADRL 所面临的可扩展性和延迟挑战。 VQC 可以在传统计算机上实现，并在相关库的帮助下实现，例如 PennyLane、IBM Qiskit、Google Cirq 和 Torch Quantum。接下来，我们将利用有前景的技术，例如空间压缩 [16]、超网络 [17] 和元学习 [18]，以增强所提出的 QMADRL 在大规模和异构 6G 网络中的鲁棒性和可扩展性。我们还将研究 QMADRL 的实际实施考虑。例如，在当前的噪声中等规模量子 (NISQ) 时代，由于量子退相干和量子门不准确引发的量子误差将不可避免地影响基于 VQC 的方案的学习性能。因此，我们将明确研究量子噪声对我们 QMADRL 解决方案的基本影响，并提出有效的策略（例如，混合优化 [19] 或集成学习 [20]）来应对这一缺点。最后，我们将开发量子启发的 MADRL 方案，在这些方案中，量子物理学的原理被用于辅助经典 MADRL 行动选择和经验回放，建立在李博士的早期工作 [21-23] 的基础上。简而言之，本研究将同时为 AI、量子计算和无线 6G 系统做出贡献，从而为 QMADRL 应用于 6G 系统这一有前途的跨学科领域提供科学基础。