Quantum Deep Reinforcement Learning for Digital Twin-Enabled 6G Networks and Semantic Communications: Considerations for Adoption and Security

Bhaskara Narottama, *Member, IEEE,* Abrar Ul Haq, *Student Member, IEEE,* James Adu Ansere, *Member, IEEE,* Nidhi Simmons, *Senior Member, IEEE,* Berk Canberk, *Senior Member, IEEE,* Simon L. Cotton, *Fellow, IEEE,* Hyundong Shin, *Fellow, IEEE* and Trung Q. Duong, *Fellow, IEEE*

(Invited Paper)

Abstract—Recently, quantum deep reinforcement learning (Q-DRL) has started to gain attention as a potential approach for tackling complex challenges in wireless communication systems. In particular, Q-DRL, integrating quantum operations into deep learning models, can effectively handle dynamic environments and process large-scale optimizations. As future wireless networks continue to evolve, greater emphasis is being placed on context and meaning rather than raw data. New paradigms, such as semantic communications (SemComs) are essential to effectively convey meaning between transmitters and receivers. By linking SemComs with Q-DRL, future wireless networks will be capable of large-scale extractions and decoding of meaning, thereby minimizing reliance on complete context sharing between communicating parties. Together with SemComs, digital twins (DTs) have been considered as key enablers for future wireless networks. As virtual replicas of physical networks, they serve an important role in network operation, optimization, and control. In this regard, Q-DRL will also be highly beneficial for DTs in enhancing critical functions such as data management and security. This study offers fresh outlooks on how to leverage Q-DRL for SemComs in future wireless networks, augmented by the use of DTs.

Index Terms—Digital twins; quantum deep reinforcement learning; semantic communications; wireless communication networks.

B. Narottama and J. A. Ansere are with the Faculty of Engineering and Applied Science, Memorial University of Newfoundland, St. John's, NL A1C 5S7, Canada (e-mail: {bnarottama,jaansere}@mun.ca).

A. U. Haq and H. Shin are with the Department of Electronics and Information Convergence Engineering, Kyung Hee University, 1732 Deogyeong-daero, Giheung-gu, Yongin-si, Gyeonggi-do 17104, Republic of Korea (e-mail: hshin@khu.ac.kr).

S. L. Cotton and N. Simmons are with the Centre for Wireless Innovation (CWI), School of Electronics, Electrical Engineering and Computer Science, Queen's University Belfast, BT7 1NN Belfast, U.K. (e-mail: {simon.cotton, nidhi.simmons}@qub.ac.uk).

B. Canberk is with Edinburg Napier University, UK (e-mail: b.canberk@napier.ac.uk).

T. Q. Duong is with the Faculty of Engineering and Applied Science, Memorial University, St. John's, NL A1C 5S7, Canada and also with the Department of Electronic Engineering, Kyung Hee University, Yongin-si, Gyeonggi-do 17104, South Korea (e-mail: tduong@mun.ca).

Corresponding authors: Trung Q. Duong and Hyundong Shin.

DRL for SemComs in future wireless networks, augmented by the use of DTs.

I. INTRODUCTION

THE sixth generation (6G) networks aim at providing unparalleled connectivity, ultra-low latency, high data rates, and ubiquitous coverage [1]. Achieving these goals involves dealing with challenges such as rapid traffic increases, fluctuating user demands, complex spectral resource management, and varying wireless channel conditions. With the incorporation of diverse technologies such as Internet-of-Things (IoT), extended reality (XR), and massive machine-to-machine (M2M) communications, novel workflows are needed to adapt, learn, and optimize in real time [2], [3]. Semantic communications (SemComs) have gained significant interest as a key element in addressing these challenges by transmitting meaningful, task-specific information, not just raw data [4], [5], especially when combined with digital twins (DTs) [6] (i.e., high-fidelity virtual models that mirror the physical network) facilitating performance prediction, network behavior simulation, and system limitation forecasting [7]. However, integrating SemComs and DTs demands significant computational resources and advanced decision-making. To this end, quantum deep reinforcement learning (Q-DRL), which leverages quantum features such as superposition and entanglement, alongside deep reinforcement learning (DRL), will play a crucial role in optimizing complex large-scale networks in real time [8]–[11]. In the sequel, we discuss how Q-DRL addresses some of the challenges of network uncertainty and frequent context changes that will be thrust upon 6G, paving the way for resilient and intelligent wireless infrastructure of the future.

A. Background and Motivations

As we progress from 5G to 6G, the next generation of wireless networks promises more than improved data throughput and reduced latency [12], [13]. Envisioned 6G systems must deliver ultra-high reliability, massive machine-type communications, and immersive experiences ranging from XR to holographic interactions, while ensuring seamless connectivity in

remote and highly mobile environments [14]. Achieving these goals requires intelligent and autonomous decision-making tools capable of handling the highly dynamic and unpredictable nature of future wireless networks. However, conventional network optimization and machine learning (ML) techniques fall short when confronted with the sheer complexity and scale of 6G [15], [16].[1] This is where Q-DRL, SemComs, and DTs converge to reshape the wireless network ecosystem: First, quantum computing holds the promise of tackling the extraordinarily complex optimization tasks that traditional computing methods struggle to handle, while deep reinforcement learning (DRL) enables systems to refine their decision-making through continuous feedback [19].[2] Second, SemComs aim to transmit the meaning of data, improving efficiency and intelligence. Finally, DTs simulate network conditions, user behavior, and deployment strategies accurately, without disrupting real-world operations. Together, these technologies lay the foundation for 6G networks that are not only capable of higher data rates and greater reliability, but also smart, resilient, and immersive [22].

1) SemComs: Instead of transmitting every piece of information in its entirety, SemCom protocols focus on the underlying meaning or intent behind a message, sharing only the critical details [23]. By integrating semantics-aware strategies into the network, 6G systems can achieve: (i) *Minimized redundancy.* Rather than sending bulky data streams, semantic mechanisms convey condensed representations of information [24], alleviating network congestion and making better use of limited spectrum resources. (ii) *Reliability.* By recognizing context and prioritizing crucial details (for example, emergency alerts or core system updates), SemComs ensure the reliable transmissions of essential information, even under variable resource constraints [25]. (iii) *Context-driven adaptability.* As network environments, user demands, or service requirements change, semantic methods adapt how information is encoded and communicated [26], rendering wireless systems more resilient to fluctuating conditions. Nevertheless, bringing SemComs to life poses challenges, as networks must be able to infer context, understand user intent, and interpret the changing environment [27].

2) DTs: In conjunction with SemComs, DT technology emerges as a key component of 6G intelligence [28], [29].

A DT serves as a virtual counterpart to the corresponding physical network, continuously updated to mirror real-world changes in near real time. By accurately modeling network elements, traffic patterns, and user activities, operators can: (i) *Anticipate failures* by pinpointing potential bottlenecks or performance degradations in DT environments before they manifest in the physical environment, prompting solutions [30]. (ii) *Optimize resource allocation* by evaluating updated network configurations, routing approaches, or quality-of-service policies within the DT environment, minimizing risks before implementing changes to the physical network [31]. (iii) *Enhanced security and resilience* by simulating scenarios ranging from cyberattacks to natural disasters, thereby validating countermeasures and enhancing overall robustness of the wireless system [32]. However, maintaining an updated DT is not trivial, demanding massive data processing and continuous learning from real-time physical measurements. Consequently, traditional simulation frameworks and model-based strategies will face challenges in keeping up with the complexity and dynamics of wireless networks.

3) Limitations of Classical Machine Learning in 6G: While ML is valuable in current wireless networks, conventional ML algorithms, including many deep learning methods, are expected to encounter major challenges in future wireless environments [33], particularly concerning optimization scalability and changing network conditions: (i) *Enormous state and action spaces.* With billions of devices and diverse service-level agreements (SLAs), the number of possible states, e.g., channel conditions, user demands, and actions (including spectral resource allocations and scheduling decisions) grows exponentially [34], [35]. (ii) *Real-time constraints.* Key 6G use cases, ranging from ultra-responsive tactile internet applications to holographic interactions, require extremely low latency. (iii) *Rapid environment changes.* 6G infrastructure must handle transient link quality and varying spectral and time resource availability in urban, airborne, or space-based (e.g., nonterrestrial networks) scenarios [36], [37]. As these problems are rooted in the growing complexity of future wireless networks, quantum-enabled learning offers a prospective alternative for information processing information and optimization, with the key benefits as follows. On one hand, assuming N_{bits} classical bits, a classical computing system may represent one of $2^{N_{\text{bits}}}$ possible states. On the other hand, assuming $N_{\text{qubits}} = N_{\text{bits}}$ quantum bits, a quantum computing system may represent a superposition of $2^{N_{\text{qubits}}}$ possible states, thanks to quantum computing properties such as superposition and entanglement, which will be detailed in Section II. This enables quantum computing to handle large state and action spaces [38], and potentially accelerates training convergence [39].

B. Related Tutorials and Surveys

The following outlines[3] tutorials and surveys covering SemComs and DT-enabled future wireless networks, drawing at-

---

[1]Use cases such as tactile internet, augmented reality, and industrial autonomous systems demand sub-millisecond decision cycles under reliability constraints (e.g., 99.99999% or higher), as in Ultra-Reliable Low Latency Communication (URLLC) and xURLLC (Extreme URLLC). Highly parameterized neural networks within DRL workflows, particularly when processed on edge or IoT devices, may exceed available computation and energy budgets, involving challenges in satisfying ultra-low-latency requirements. To mitigate bottlenecks in delivering timely control actions for such latency-sensitive applications, well-designed quantum learning models may promote training convergence [17], [18], making reduced training iterations possible.

[2]DRL is suitable for complex and time-varying wireless scenarios, where estimating solutions via mathematical or supervised learning approaches (which typically requires labeled data) faces difficulties. For example, [20] applies DRL for managing resources in near-field internet-of-everything (IoE) for healthcare monitoring, which can also be extended to power conservation cycles, and more. Beyond its role in communications, DRL can also be used for node/user localization, as in [21]. In this study, received signal strength (RSS) data is used to locate IoT nodes without labeled data, with rewards set by exploiting *landmark points* that have reliable location and RSS data.

[3]Although this subsection primarily covers tutorials and surveys, it also features relevant technical studies, particularly those offering useful reviews and/or classifications.

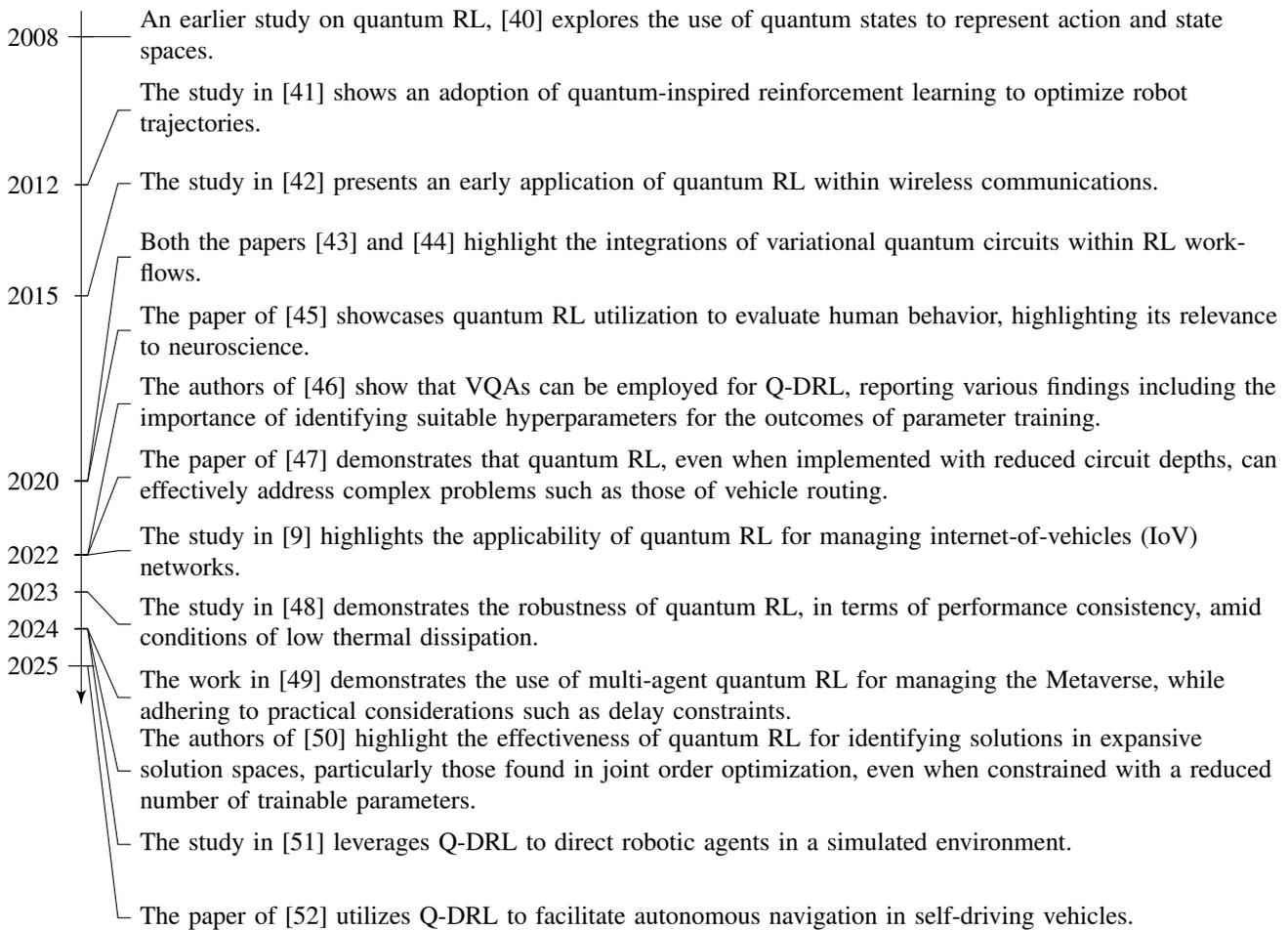|	
2008	An earlier study on quantum RL, [40] explores the use of quantum states to represent action and state spaces.
	The study in [41] shows an adoption of quantum-inspired reinforcement learning to optimize robot trajectories.
2012	The study in [42] presents an early application of quantum RL within wireless communications.
2015	Both the papers [43] and [44] highlight the integrations of variational quantum circuits within RL workflows.
	The paper of [45] showcases quantum RL utilization to evaluate human behavior, highlighting its relevance to neuroscience.
	The authors of [46] show that VQAs can be employed for Q-DRL, reporting various findings including the importance of identifying suitable hyperparameters for the outcomes of parameter training.
2020	The paper of [47] demonstrates that quantum RL, even when implemented with reduced circuit depths, can effectively address complex problems such as those of vehicle routing.
2022	The study in [9] highlights the applicability of quantum RL for managing internet-of-vehicles (IoV) networks.
2023	The study in [48] demonstrates the robustness of quantum RL, in terms of performance consistency, amid conditions of low thermal dissipation.
2024	The work in [49] demonstrates the use of multi-agent quantum RL for managing the Metaverse, while adhering to practical considerations such as delay constraints.
2025	The authors of [50] highlight the effectiveness of quantum RL for identifying solutions in expansive solution spaces, particularly those found in joint order optimization, even when constrained with a reduced number of trainable parameters.
	The study in [51] leverages Q-DRL to direct robotic agents in a simulated environment.
	The paper of [52] utilizes Q-DRL to facilitate autonomous navigation in self-driving vehicles.

Fig. 1: Timeline of quantum RL development, featuring some of the notable research contributions.

tention to the limited studies on the adoptions of Q-DRL for such networks, especially in 6G, despite the growing use of quantum RL across many fields, as outlined in Fig. 1. As presented in the figure, early QRL studies addressed challenges involving high-dimensional action and state spaces in dynamic environments, such as those in robotic locomotion [40], [41]. Subsequent studies have adopted QRL for wireless networks and/or systems [42]. Nevertheless, such studies remain limited to considering singular learning agents. More recently, efforts have been expanded to multi-agent Q-DRL, allowing the learning workflow to involve multiple nodes within wireless networks, as in [9]. The inclusion of multi-agent learning opens research efforts in (i) managing semantic-aware networks and (ii) operating and securing both physical and digital entities in DTs. These aspects have been overlooked in the related studies.

1) SemComs: The following discussion outlines a selection of representative surveys and tutorials covering SemComs. In [53], many of the different SemComs systems and techniques are surveyed. This includes cross-modal systems, semantic coded transmission, and semantic forward. The review conducted in [54], discusses various metrics relevant to SemComs, embracing those pertinent to quantum SemComs,

such as expected fidelity and Uhlmann fidelity. Human-to-human (H2H), human-to-machine (H2M), machine-to-machine (M2M), and knowledge graph-based SemCom are studied in [55]. Specifically, [55] discusses how H2M-based SemComs can enhance QoE by facilitating human-machine symbiosis via applications like interactive machine learning, whereas M2M-based SemComs improve distributed learning and split inference.

(i) *SemComs and DRL.* A number of studies have explored potential integrations of DRL and/or other AI-based approaches with SemComs. In particular, [56] discusses DRL utilizations for encoders and decoders, and highlights how SemComs' progress parallels that of DRL. Similarly, [57] explores the utilizations of representation learning to facilitate semantic reasoning and inferring information context in SemComs, whereas [58] addresses the usages of generative AI for such purposes. Furthermore, [59] reviews various types of SemComs, such as those based on weak and strong semantic information and truthlikeness. It also discusses relevant training loss functions for SemComs, and examines DRL-based approaches, e.g., collaborative DRL, utilizations for SemComs. In addition, [60] explores the use of generative difussion models

in SemComs. It provides multiple adoption cases such as the use of expert datasets, and the benefits of processes like reverse diffusion. In particular, it highlights the importance of leveraging DRL, for example the use of deep diffusion soft actor-critic (D2SAC), for SemComs. Furthermore, in [61], RL is employed for semantic bit coding, whereas [62] utilizes generative AI for the encoding and decoding of semantic information. In [63], deep learning is adopted for SemComs that leverage knowledge distillation, i.e., technique to transfer the prediction ability of a pre-trained model, while accounting for interference from multiple user terminals. Yet, these studies, while comprehensive, do not focus on leveraging quantum DRL.

(ii) *SemComs and Security.* There are growing concerns over the security of SemComs, along with its potential to improve wireless system security. In this context, [64] presents a taxonomy on resource allocation and security in SemComs, highlighting vulnerabilities, e.g., those related to knowledge bases, and addressing SemComs' role in enabling secure communications, e.g., through facilitating quantum key distributions and covert communications. Delving further into knowledge base security, studies have reviewed various approaches for data security, e.g., employing data differential and zero-knowledge proof [64], [65]. Furthermore, as deep learning, and by extension, DRL, gains relevance in SemComs, particularly for semantic extractions, security risks pertinent to it, e.g., data poisoning, shall be mitigated. In this regard, [66] reviews different strategies pertinent to training-based security, e.g., defensive distillation, as well as training-free approaches.

2) DT: Recognizing DTs' broad applicability in wireless network optimizations, management, and services, studies have introduced taxonomies to classify their benefits for future networks. The authors of [67] presented a taxonomy examining DT and wireless system integrations, identifying the ways in which DTs enhance wireless systems, and vice versa. In [68], a classification is presented that identifies the role of DTs in 6G services, including their applications in network management and blockchain.

(i) *DT and Wireless Networks.* To effectively support wireless networks, DTs shall be required to meet certain performance requirements. Such notion is particularly affirmed in [69], which calls for enhanced DT privacy (to protect user data from attackers) and scalability (to ensure DTs keep pace with future ultra-dense networks). DTs are especially apt to capture wireless network uncertainty, for example surges of streaming demand at sporting events, and dynamism, as network variables typically shift with time. In this case, [70] highlights DTs' roles in estimating the mobility of network elements, e.g., user equipment, ensuring smooth handovers in 6G. It is important to note that the capability of a wireless network can significantly impact DT operations as well. This notion has been addressed in studies, e.g., [67] and [71]; [67] emphasizes that emergent communication systems, such as machine-to-machine communications, are vital for fulfilling DT's requirements, e.g., improved coordination between the digital and the physical entities.

(ii) *DT and DRL.* Several studies highlight the interplay between DTs and AI-based approaches, among which is DRL. As such, [72] demonstrates how DTs can improve system operations, specifically in areas like fault and anomaly detection. This improvement is achieved by using AI-based approaches such as generative adversarial networks, active learning, and domain adversarial neural networks. Given its effectiveness in enhancing system operation, this interplay holds significant benefits for 6G, as highlighted by [73]. The integration of DRL and DT is projected to provide improvements in several areas, including message prediction accuracy, thereby facilitating a broad range of applications such as high-fidelity RAN modeling and network slicing. Nonetheless, while multiple studies, e.g., [67] and [68], acknowledge the benefits of integrating DRL with DTs, e.g., for twin isolation [67], they do not focus on developing quantum (or classical) DRL for such integrations, especially in the context of wireless networks. Yet, the importance of such integrations is highlighted by various studies. For example, [74] introduces the concept of reinforcement twinning (although it has not been applied to wireless networks).

The contributions of this paper, as compared to others, are summarized in Table I, highlighting a gap in the literature. The symbol "◇" indicates the paper's brief mention of (or close relation to) a topic, whereas "✓" indicates thorough discussion or even idea proposition. For instance, certain papers, while not directly mentioning DRL/Q-DRL, cover other machine learning (ML) and/or deep learning (DL) approaches, and are thus marked "◇." Regarding the discussion on Q-DRL taxonomy, the designations "general" and "wireless" indicate the paper's relevancy to a broader audience and to specialists in wireless communication networks, respectively.

C. The Primary Contributions of This Paper

The primary contributions of this paper can be outlined as follows.

- This paper systematically explores Q-DRL's potential to advance both SemComs and DT operations in intelligent and secure wireless networks. Several studies have advanced Q-DRL, by addressing its frameworks and architectures: In particular, [76] explores different approaches for quantum-enhanced policy iteration, such as quantum approximate policy iterations, and [77] presents how DRL elements can be integrated into quantum circuits, especially by representing the action and environment using operations within Grover's algorithm, which enables their interaction through quantum channels. Still, these studies do not discuss Q-DRL adoptions for future wireless

TABLE I: Comparison of Pertinent Studies

		Other Papers							This Paper
		[75] (First. 2023)	[66] (Jun. 2023)	[59] (Feb. 2023)	[56] (Fourth. 2024)	[67] (Fourth. 2022)	[73] (Aug. 2024)	[72] (Nov. 2024)	
Q-DRL realization	general wireless	-	-	-	-	-	-	-	✓
Q-DRL taxonomy		⋄ (general RL)	-	-	-	-	-	-	✓
DRL/Q-DRL for SemComs		⋄	⋄ (DL)	✓(general RL)	✓(classical DRL)	✓(classical DRL)	⋄	⋄ (using ML)	✓
DRL/Q-DRL for DTs		-	-	⋄	-	⋄	✓(classical DRL)	✓(classical DRL)	✓(quantum DRL)
security considerations		✓	✓	-	✓	✓	✓	✓	✓(with use case)

networks.[4]

- This is an early study that introduces a taxonomy of integrating Q-DRL into SemComs and DT-enabled networks. It categorizes Q-DRL architectures, learning models, agent configurations, and training strategies relevant to wireless networks. Various studies explore DRL classifications, and their adoptions for wireless networks. In particular, [75] provides classifications of DRL pertinent to its utility in facilitating communication security, while [80] offers an interesting overview of studies on DRL, classifying them based on learning principles and methods to acquire datasets, among other factors. Yet, these notable studies do not focus on integrating quantum computing and DRL.[5]

- This paper provides detailed insights on how Q-DRL can benefit both SemComs and DTs, as discussed later in Section V. Regarding SemComs, this paper highlights their prospects for 6G network operations and cross-layer design (Subsections V-A and V-B). With respect to DTs, this paper covers how Q-DRL can support autonomous digital entities, improving their operations and security (Subsection V-F). This paper also provides a security-focused Q-DRL application in DT-enabled networks, demonstrating how quantum learning models can improve resilience against eavesdroppers, which remains insufficiently explored in previous studies.

D. Paper Organization

As illustrated in Fig. 2, the rest of this paper is structured as follows. Section II discusses the fundamentals of quantum computing, establishing the comprehension necessary for the sections that follow. Section III discusses the Q-DRL realization for SemComs and DT-enabled networks. Section IV provides a Q-DRL taxonomy, classifying distinct approaches, alongside their importance for various aspects of future wireless networks. Section V discusses in detail the significance of Q-DRL for SemComs and DT–enabled networks, explaining its role in semantic extractions, to improving DT security. Section VI discusses a practical example of Q-DRL adoptions for future wireless networks, while also highlighting its security aspects, which remain underexplored in current literature. Table II illustrates the acronyms and abbreviations used across this paper.
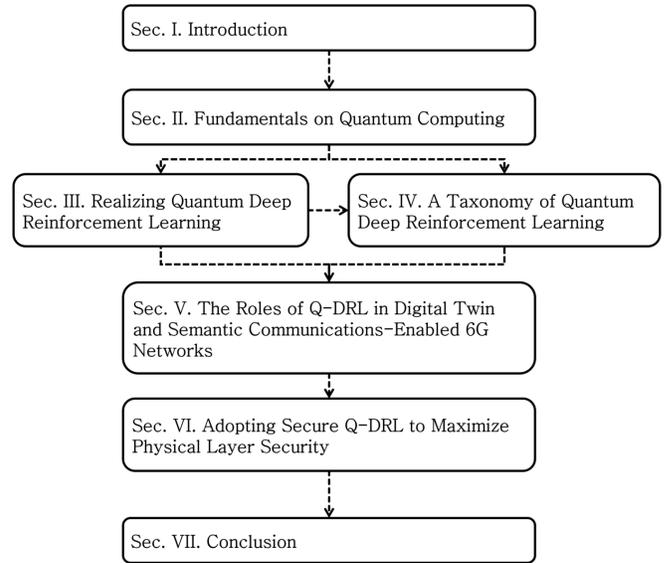
Fig. 2: The structure of this paper, with the dashed arrows (--→) indicating the logical flow between sections.

---

[4]Various studies have highlighted the effectiveness of DRL for managing wireless networks: In particular, [78] employs an Actor-Critic framework for improving computational task offloading in LEO satellite networks, and [79] leverages deep Q-network for beam tracking in ultra-massive MIMO systems, facilitating robust beamforming gain amid randomness of user mobility. Yet, these studies do not consider Q-DRL. Addressing the identified gap, this paper investigates the integration of quantum-based operations. Specifically, it examines different quantum operations (relevant to a variety of DRL frameworks) which are applicable to future wireless networks, with particular emphasis on those leveraging SemComs and DTs.

[5]Some studies do, however, address how quantum operations can be applied for DRL. The work presented in [81] addresses different approaches integrating quantum computing and reinforcement learning (RL) frameworks, and vice versa, extending from quantum computing-inspired RL to fully quantum-operated RL. Nevertheless, this study does not focus on providing guidance on how to develop such frameworks, let alone their adoptions for specific wireless network use cases, particularly those of SemComs.

II. FUNDAMENTALS ON QUANTUM COMPUTING

The following reviews quantum computing principles to support the discussion on Q-DRL, covering quantum bits, measurement, gates, circuits, and recent developments in general-purpose quantum computing.

A. Quantum Bits and Measurements

Quantum processing works on quantum bits, termed qubits, each of them can represent a superposition of 0 and 1. In this

TABLE II: Acronyms and Abbreviations

Acronym/Abbreviation	Definition
Q-DRL	quantum deep reinforcement learning
DRL	(classical) deep reinforcement learning
DTs	digital twins
SemComs	semantic communications
IoT	Internet-of-Things
qubit	quantum bit
NISQ	noisy intermediate-scale quantum
QNN(s)	quantum neural networks
NN(s)	(classical) neural networks
MIMO	multiple-input and multiple-output
UAV	unmanned aerial vehicle
VQC	variational quantum circuit
QoS	Quality-of-Service
QoE	Quality-of-Experience
FL	federated learning
MEC	mobile edge computing

Fig. 4: Co-development pathways for quantum and classical systems.

regard, the state of a particular qubit n can be expressed as: $|\psi_n\rangle = \rho_0 |0\rangle + \rho_1 |1\rangle$, where $|0\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $|1\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ are the computational basis states. In addition, ρ_0 and ρ_1 are the complex numbers pertinent to the probability of obtaining 0 and 1, upon measuring the particular qubit, where $|\rho_0|^2 + |\rho_1|^2 = 1$. This extends to a multi-qubit system, whose state is given by: $|\psi\rangle = \sum_{ꜩ=0}^{2^{N_{\text{qubit}}}-1} \rho_ꜩ |ꜩ\rangle$, where N_{qubit} signifies the number of qubits, and $\rho_ꜩ$ is related to the probability of obtaining a string of bits equivalent to the integer $ꜩ$.[6] To obtain classical values as the resulting output of a quantum circuit, which, in particular, can be used to process actions in classical-valued operations, it is necessary to perform quantum measurements. In this regard, let $\mathbf{W}$ be the measurement operator associated with observing a classical value $ꝺ$, where $ꝺ \in \{0, 1\}$. Given a qubit in the state $|y_{\text{out}}\rangle$ prior to measurement, the probability of measuring an outcome $ꝺ$ is given by $\mathsf{P}(ꝺ) = \langle y_{\text{out}}| \mathbf{W}^{\mathsf{H}}\mathbf{W} |y_{\text{out}}\rangle$, where $|y_{\text{out}}\rangle$ marks the state of the qubit (the bra notation $\langle y_{\text{out}}|$ indicates the Hermitian adjoint of $|y_{\text{out}}\rangle$).
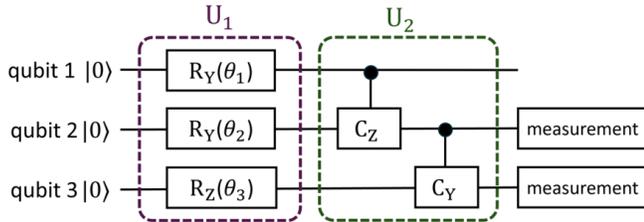

Fig. 3: A particular quantum circuit operated on three qubits, which are represented by different horizontal lines.

[6]These quantum properties, particularly quantum superposition and entanglement, facilitate quantum parallelism, which results from the interactions of $2^{N_{\text{qubits}}}$ states. In this light, quantum computing can be used to explore different optimization paths simultaneously, as highlighted in [82] (This benefit is apparent when using certain quantum algorithms, such as amplitude amplification.).
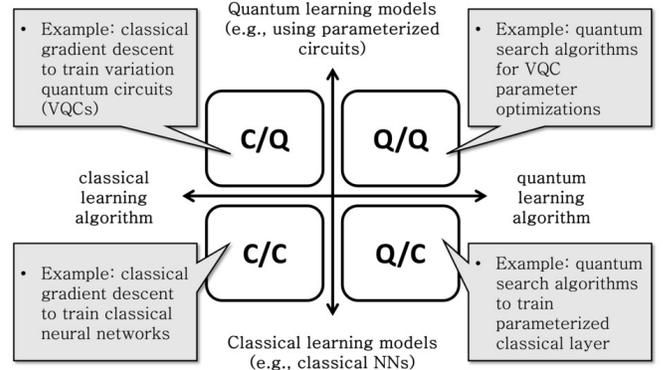
B. Quantum Gates and Circuits

An example of a quantum circuit with three qubits is depicted in Fig. 3. In this instance, the state of each qubit is initialized to $|0\rangle$. Subsequently, a series of quantum rotational gates is processed, and thus, as a unitary operation, it is representable as: $\mathsf{U}_1 = \mathsf{R}_{\mathsf{Y}}(\theta_1) \otimes \mathsf{R}_{\mathsf{Y}}(\theta_2) \otimes \mathsf{R}_{\mathsf{Z}}(\theta_3)$, where θ_1, θ_2, and θ_3 indicate their respective rotation angles. As such, the tensor product ("$\otimes$") indicates operations upon distinct qubits. These gates operate single qubit rotation along the y- and z-axes, respectively, as expressed by:

$$\mathsf{R}_{\mathsf{Y}}(\theta) = \begin{bmatrix} \cos\theta/2 & -\sin\theta/2 \\ \sin\theta/2 & \cos\theta/2 \end{bmatrix}, \quad \mathsf{R}_{\mathsf{Z}}(\theta) = \begin{bmatrix} e^{-i\theta/2} & 0 \\ 0 & e^{i\theta/2} \end{bmatrix}. \quad (1)$$

Meanwhile, a single qubit rotation on x-axis is given by: $\mathsf{R}_{\mathsf{X}}(\theta) = \begin{bmatrix} \cos\theta/2 & -i\sin\theta/2 \\ -i\sin\theta/2 & \cos\theta/2 \end{bmatrix}$. Thence, controlled-Z and controlled-Y gates respectively apply Pauli-Z and Pauli-Y operations to the corresponding target qubits when the respective controlling qubits are in the state $|1\rangle$. These two gates can be expressed as:

$$\mathsf{C}_{\mathsf{Z}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}, \quad \mathsf{C}_{\mathsf{Y}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{bmatrix}, \quad (2)$$

respectively, while the controlled-X gate is given by:

$$\mathsf{C}_{\mathsf{X}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}. \quad (3)$$

In this case, the controlled-Z and controlled-Y gates are processed in sequence, expressed as a unitary operation: $\mathsf{U}_2 = (\mathsf{C}_{\mathsf{Z}} \otimes \mathsf{I})(\mathsf{I} \otimes \mathsf{C}_{\mathsf{Y}})$, where I denotes the identity matrix. Eventually, the whole circuit can be expressed as $\mathsf{U}_2\mathsf{U}_1 |0\rangle^{\otimes 3}$, in which $|0\rangle^{\otimes 3}$ indicates that each of the three qubits is initialized to $|0\rangle$. Moreover, Pauli-X, -Y, and -Z operations can be expressed as $\mathsf{X} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, $\mathsf{Y} = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$, and $\mathsf{Z} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. Furthermore, the

Hadamard gate, expressed as $H = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$, can introduce superposition into a qubit.

C. Recent Developments

To facilitate computation on noisy intermediate-scale quantum (NISQ) computing platforms, which are prone to errors due to the presence of noise, several error mitigation, suppression, and correction approaches have been explored, particularly probabilistic error cancellation [83]. On top of that, many studies, such as [84], cover early and future fault-tolerant quantum computing and discuss their scalability. Nevertheless, studies have analyzed that sufficient computations with tolerable error can be performed on certain quantum computing platforms with acceptable noise thresholds, e.g., below 10^{-2} [85]. Furthermore, given the benefits of quantum computing and the availability of classical computing, interplay between the two is often suggested. Pertinent to Q-DRL, Fig. 4 illustrates the prospects of co-development between both types of systems, particularly in how learning models are processed and trained. For example, under the "**Q/Q**" designation, quantum search algorithms, particularly those based on Grover's algorithm, may be employed to optimize the parameters of quantum learning models.

III. REALIZING Q-DRL

In the following, we cover Q-DRL adoptions for SemComs and DT-enabled networks, covering important notions such as designing learning frameworks (see Subsection III-C), employing quantum circuits as learning models (see Subsection III-D), and training model parameters (see Subsection III-E). The design steps comprising these processes are illustrated in Fig. 5.
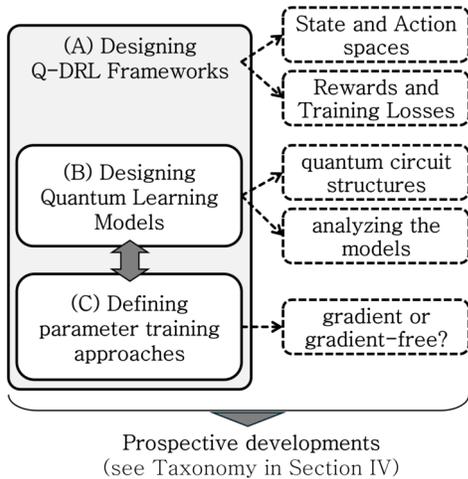

Fig. 5: The design steps in realizing Q-DRL for SemComs and DT-enabled networks, with each step explained in Section III. Further discussion and a taxonomy for the Q-DRL aspects are presented in Section IV.

A. Structural Differences Between Q-DRL and Traditional DRL

Traditional DRL relies on classical computing paradigms, where agents use classical learning models, e.g., deep neural networks, to approximate value functions and policies. By exploring and exploiting actions in a given environment, these agents iteratively learn optimal decision-making strategies. Q-DRL, on the other hand, integrates principles from both quantum computing and classical DRL to either (i) accelerate the learning process, (ii) reduce complexity via quantum parallelism, or (iii) improve performance and scalability in environments with large state-action spaces. We now highlight the key structural distinctions between Q-DRL and its antecedent, classical DRL. (i) *Quantum states and superposition.* While classical DRL uses conventional data structures, particularly vectors and tensors, to represent states, Q-DRL leverages quantum states, exploiting superposition to encode more information in a single representation. This can drastically reduce the dimensionality challenge in large state spaces. In this respect, while classical DRL typically relies on high-dimensional Euclidean spaces for representing inputs, Q-DRL extends this to Hilbert spaces, where the state vectors can represent superposed states, theoretically providing an exponential representation advantage [86]. (ii) *Quantum operators and unitary transformations.* Instead of classical matrix multiplication found in deep neural networks, Q-DRL uses quantum gates, which are unitary operators, to manipulate quantum states. This can preserve information in ways that are not possible classically (e.g., reversible computations). Accordingly, quantum learning models, especially quantum neural networks (QNNs), augment classical deep networks with quantum operations, leading to parameterized quantum circuits (PQCs) that can be trained in a manner analogous to backpropagations (Subsection III-E further elaborates on this). (iii) *Quantum measurements and observables.* Q-DRL typically involves measurements, which collapse the quantum wavefunction. The act of measurement in quantum computing yields a classical outcome from a quantum state. Designing reward functions and updating policies thus demand strategies to minimize destructive measurement effects while extracting sufficient information for training. Furthermore, quantum measurements yield probabilistic outputs governed by the Born rule, adding stochasticity. Q-DRL must incorporate this into policy updates, ensuring learning convergence despite quantum randomness. (iv) *Hybrid quantum-classical workflows.* Many practical Q-DRL architectures are hybrid, combining classical neural networks with quantum circuit blocks, as in [87]. The classical parts may handle pre-processing or certain function approximations, while quantum circuits tackle high-complexity tasks or state encoding for improved exploration. Given the constraints of NISQ devices, specialized techniques like variational quantum circuits and error mitigation strategies are required. This changes the structure of training pipelines and demands a co-design of hardware and software, as it includes, for instance, device-level executions and library compilations

[88]. By exploiting superposition, entanglement, and other quantum phenomena, Q-DRL offers the potential for more efficient exploration, rapid learning convergence, and resilience in high-dimensional or partially observable environments, crucial for DT-enabled 6G networks and semantic communications.

B. Successful Q-DRL Adoptions in Different Domains

Although still in its early stages, Q-DRL has shown promise across fields. The following adoptions demonstrate its potential to transform industry and academia, and foreshadow the advantages it may bring to DT-enabled 6G networks and semantic communications. (i) *Robotics and automated control.* In robotics, controlling arms or drones in dynamic environments often involves large state spaces (e.g., joint positions, velocities, etc). Hybrid Q-DRL has demonstrated faster convergence times when training policy networks for coordinated maneuvers compared to purely classical methods, e.g., [51]. (ii) *Portfolio optimization in finance.* Q-DRL has been employed to optimize trading strategies and portfolio allocations, e.g., [89]. Quantum parallelism in state encodings has shown potential for more efficient search within large market data sets, enabling near-real-time strategy adjustments. (iii) *Drug discovery and healthcare.* Drug discovery requires large-scale exploration of chemical compounds and molecular interactions. Q-DRL can boost the in silico search for potential drug candidates via quantum-enabled exploration of molecular conformations [90]. In healthcare, Q-DRL frameworks have been explored for personalized treatment regimens, taking patient heterogeneity into account, e.g., [91]. (iv) *Networking and resource allocation.* Preliminary research on quantum-enhanced DRL agents for adaptive routing in large-scale networks has shown improvements in latency and energy efficiency. The quantum agent efficiently explores alternative paths and traffic allocation strategies, critical in dynamic topologies. Such approaches, e.g., [92], have also demonstrated enhanced resource allocation under multiple realistic scenarios, achieving optimized usage of limited spectrum.

C. Designing Quantum DRL Frameworks

1) Defining States and Inputs: Q-DRL frameworks generally involve agents that learn to maximize rewards (and accordingly, expected rewards), through receiving feedback pertaining to past interactions and observing their respective environments. These agents command certain actions, with their policies guiding action selection. Moreover, as discussed in studies such as [93], it is possible to utilize DTs to simulate environments, so that the agent can safely interact without risking the physical environment due to sub-optimal policies. This is especially relevant in mission-critical scenarios, such as non-terrestrial or military communication networks [94]. Furthermore, Fig. 6 illustrates typical interactions between an agent and its environment: (i) *State spaces.* The state space represents a set of states reflecting the environment, that can be observed and influenced by the agent(s). For instance, for the iteration time t, the state space concerning the i-th agent

can be expressed as: $\mathcal{S}_i^{(t)} = \{\mathbf{s}_{l,i}^{(t)}\}_{l=1}^{|\mathcal{S}_i^{(t)}|}$, in which $\mathbf{s}_{l,i}^{(t)}$ indicates the l-th possible state, which can be discrete or continuous in value. For SemComs and DT-enabled networks, the state space can be the the values indicating the importance of semantic information [95], and effective semantic information [96], in addition to typical network information such as the channel state information (CSI) and the received signal-to-interference-plus-noise ratio (SINR). (ii) *Action spaces.* Furthermore, the action space consists of a set of possible actions, that can be performed by the agent(s), to influence the state of the environment. It can be expressed as: $\mathcal{A}_i^{(t)} = \{\mathbf{a}_{m,i}^{(t)}\}_{m=1}^{|\mathcal{A}_i^{(t)}|}$, where $\mathbf{a}_{m,i}^{(t)}$ denotes the m-th action vector. Concerning SemComs and DT-enabled networks, the action vectors can be variables related to the processing and information uploading pertinent to DTs, such as communication bandwidth and computation resource allocations [97], and/or node selections [98].
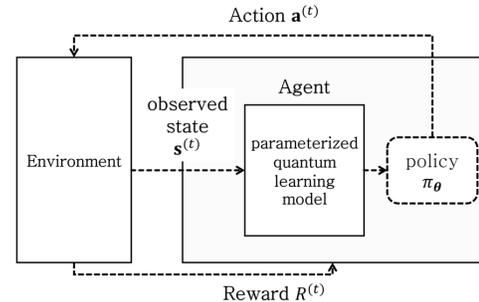

Fig. 6: Interactions between an agent and its corresponding environment in Q-DRL, where the agent employs a quantum learning model with parameter $\boldsymbol{\theta}$. At each iteration time t, it observes the environment state $\mathbf{s}^{(t)}$, and influences the environment by taking the action $\mathbf{a}^{(t)}$, guided by its policy π_θ. In turn, the environment rewards (or possibly, penalizes) the agent via $R^{(t)}$.

2) Defining Rewards and Training Losses: (i) *Rewards.* The instantaneous reward at time t represents how optimal the agent's interaction with the environment is at that particular iteration. Nevertheless, a single calculation of reward may not provide a sufficient indicator for policy performance. In this regard, cumulative rewards are often considered. For example, the cumulative reward at t can be defined as: $\hat{R}^{(t)} = \sum_{\tau=0}^{\infty}(\eta)^\tau R^{(t+\tau+1)}$, where η denotes the discount factor, and $R^{(t)}$ specifies the reward at t. In this case, $R^{(t)}$ can be defined in accordance with the performance indicator, defined in the optimization objective, reflecting how optimal the agent's/agents' interaction with the environment is. For instance, it can represent the gain in user data rate [97] and/or the reduction in offloading latency [99].

(ii) *Training losses.* Further, the loss quantifies the estimation capability of the learning model, and is vital in optimizing its parameters, as it reflects the progress of the learning process as well. For example, in the case of deep Q-network, the loss can be expressed as: $L =$

$$\mathbb{E}\left[\left(\left(R^{(t)} + \eta \max_{\mathbf{a}^{(t+1)}} \hat{Q}(\mathbf{s}^{(t+1)}, \mathbf{a}^{(t+1)} | \hat{\boldsymbol{\theta}})\right) - Q(\mathbf{s}^{(t)}, \mathbf{a}^{(t)} | \boldsymbol{\theta})\right)^2\right].$$

$Q(\cdot, \cdot | \boldsymbol{\theta})$ denotes the estimated Q-value, which represents the expected reward estimates, given the action $\mathbf{a}^{(t)}$ and the state $\mathbf{s}^{(t)}$, and can be approximated through employing quantum learning models with parameter $\boldsymbol{\theta}$. Meanwhile, $\hat{Q}(\cdot, \cdot | \hat{\boldsymbol{\theta}})$ denotes the target Q-value, approximated with a model with parameter $\hat{\boldsymbol{\theta}}$. The Q-value can be updated as: $Q(\mathbf{s}^{(t)}, \mathbf{a}^{(t)} | \boldsymbol{\theta}) \leftarrow Q(\mathbf{s}^{(t)}, \mathbf{a}^{(t)} | \boldsymbol{\theta}) + \mu \left[\left(R^{(t)} + \eta \max_{\mathbf{a}^{(t+1)}} \hat{Q}(\mathbf{s}^{(t+1)}, \mathbf{a}^{(t+1)} | \hat{\boldsymbol{\theta}})\right) - Q(\mathbf{s}^{(t)}, \mathbf{a}^{(t)} | \boldsymbol{\theta})\right]$, with μ as the defined learning rate [100].

D. Quantum Learning Models

1) Quantum Circuits for Deep Learning Models: A quantum learning model can be employed for purposes such as approximating functions in Q-DRL. In these cases, such models primarily take the form of quantum circuits, though hybrid quantum-classical models are also possible (see Section IV-A2). Each learning model typically consists of rotational and entanglement layers.

(i) *Rotational layers.* The rotational layers allow a set of parameters to influence the resulting state of a quantum learning model. Thus, in practice, these parameters function akin to weight parameters in classical neural networks. A rotational layer in particular can be in the form of a stack of parameterized operations, often realized as gates [101]. For instance, it can be defined as: $U_{rot}(\boldsymbol{\theta}) = \bigotimes_{q=1}^{N_{qubits}} U_3^{(q)}(\theta^{(q)}, \phi^{(q)}, \lambda^{(q)})$, in which $\boldsymbol{\theta}$ represents its parameter vector, N_{qubits} indicates the number of qubits, and $U_3^{(q)}(\theta^{(q)}, \phi^{(q)}, \lambda^{(q)})$ denotes the U3 gate, applied to the q-th qubit, with $\theta^{(q)}$, $\phi^{(q)}$, and $\lambda^{(q)}$ as indicating its parameters, for $\{\theta^{(q)}, \phi^{(q)}, \lambda^{(q)}\} \in \boldsymbol{\theta}$.

(ii) *Entanglement layers.* As we use multiple qubits to map and process information, we typically employ entanglement layers to connect them. Such connections are typically realized using controllable gates, such as CNOT and/or CZ gates. The structure of these layers may take different forms, as it can resemble linear, circular, parallel, and/or all-to-all connections [102]. In particular, an entanglement layer, consisting of parallel CNOT gates, can be expressed as: $U_{ent} = \bigotimes_{q=1}^{\lfloor \frac{N_{qubits}}{2} \rfloor + 1} C_X^{(2q-1)} \bigotimes_{q=1}^{\lfloor \frac{N_{qubits}}{2} \rfloor} C_X^{(2q)}$ where N_{qubits} marks the number of qubits of U_{ent}, and $C_X^{(q)}$ denotes the CNOT gate applied to the q-th qubit.

An example of the adoption of the rotational and entanglement layers for quantum learning models is provided in Fig. 7. It is worth noting that, typically, the state of the learning model circuits is initialized beforehand, particularly in the state $|0\rangle$ for each qubit. An encoding operation, such as basis or amplitude encoding, might then be executed to map classical information onto the quantum circuit, thereby preparing the input state. Several studies have investigated the effectiveness of encoding operations, particularly their impact on the learning model expressivity [103], as well as learning convergences of quantum-based reinforcement learning [104].
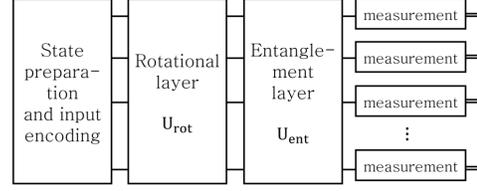

Fig. 7: An example of a quantum model circuit, with parameter ($\boldsymbol{\theta}$), primarily comprising rotational and entanglement layers, inscribed as $U_{rot}(\boldsymbol{\theta})$ and U_{ent}, respectively. In addition, the encoding operation yields the input state, while quantum measurements yield classical-valued output.

2) Analyzing Quantum Learning Models: While many studies provide empirical performance results of a quantum learning model, such as how well its training loss converges over time, these results often depend on the considered training data. Accordingly, aiming for broader applicability to generalized data, researchers have investigated additional aspects, such as dequantization, as discussed next. (i) *Quantum learning models' trainability.* In general, the trainability of a particular quantum learning model can be conveyed as the variance of the cost function, given its parameter values. It can vary depending on the structure of the employed quantum circuit. For example, given the state-efficient ansatz in [105], with θ as the parameter of the quantum model in question and C as the cost function, its trainability can be analyzed as: $\text{Var}\left[\frac{dC}{d\theta}\right] \in O\left(2^{-(N_{qubit} + N_{CNOT})}\right)$, where N_{qubit} signifies the number of qubits, while N_{CNOT} indicates the number of CNOT gates used in the circuit for qubit entanglement. Notwithstanding the preceding, there are studies that provide trainability estimates applicable to general quantum learning models. In this context, as presented in [106], the variance of the cost gradient can also be expressed as: $\text{Var}\left[\frac{dC}{d\theta}\right] = \mathbb{E}_\theta\left[\left(\frac{dC}{d\theta}\right)^2\right] - \mathbb{E}_\theta\left[\left(\frac{dC}{d\theta}\right)\right]^2$. (ii) *Dequantization.* To analyze the benefits of employing quantum learning models over their classical counterparts, [107] examines how effectively the function pertinent to a particular quantum model can be approximated through a classical learning model, a notion referred to as dequantization. The rationale for such analysis arises from the continued use of classical data for today's quantum learning models. Further, [108] addresses the relationship between dequantization and trainability in quantum learning models. Notably, it discusses how the trainable yet non-dequantizable parametrized learning models can be traced through solving the following optimization problem (among other steps): $\max_\theta \langle 0 | U^H(\theta) \mathcal{H} U(\theta) | 0 \rangle$, in which $U(\theta)$ indicates the unitary operations of the learning model with parameter θ, and $\mathcal{H}$ describes the Hamiltonian.

E. Parameter Training Approaches

Q-DRL employs parameterized learning models, requiring parameter optimization to yield desirable outputs. In general, optimizing a set of model parameters, $\theta^{(t+1)}$, for $t + 1$, can

be achieved iteratively via gradient descent, as in: $\theta_j^{(t+1)} \leftarrow \theta_j^{(t)} - \mu \nabla_{\theta_j^{(t)}} L$, for $\theta_j^{(t)} \in \boldsymbol{\theta}^{(t)}$, where μ is the learning step, L is the training loss, and $\nabla_{\theta_j^{(t)}} L$ is the loss gradient, which can be obtained through methods based on parameter-shift rule [109], [110]. An example of this approach is illustrated in Fig. 8, in which the cumulative reward values are presented with respect to the learning episodes. In this case, a mobile transmitter serves several users. Its action space includes transmit power and movement trajectory, while its state space includes user-transmitter distances. As presented in the figure, the Q-DRL approach yields higher reward values, outperforming classical QRL, demonstrating the effectiveness of its training. In addition, the reward calculation at each t, marked as $\mathcal{R}^{(t)}$, aligns with the performance indicator. In this case, $\mathcal{R}^{(t)} = \mathfrak{w} \sum_{k=1}^{K} R_k^{(t)}$, where $R_k^{(t)}$ denotes the rate of the k-th user at t, K conveys the number of served users, and $\mathfrak{w}$ is a pre-defined coefficient.

It should be noted that the need to obtain a gradient is not limited to calculating first-order gradients, but also to second-order gradients, including calculating natural gradients. Such gradients are effective, in particular, for updating policy for $t+1$, $\mathfrak{p}^{(t+1)}$, as expressed as: $\mathfrak{p}^{(t+1)} \leftarrow \mathfrak{p}^{(t)} + \zeta V$, where ζ is the learning step, and V is the value function. V can be estimated based on natural gradients, i.e., $V \leftarrow \hat{\nabla}_{\boldsymbol{\theta}} C$, where $\hat{\nabla}_{\boldsymbol{\theta}} C$ denotes the natural gradient of the cost function C [111]. Hence, there are research prospects in developing methods to obtain such second-order gradients, for optimizing quantum learning models. In what follows, we discuss the notion of developing parameter optimization methods beyond gradient-based approaches, including Rotosolve-based and search-based approaches.

1) Rotosolve-Based Approaches: Rotosolve-based approaches [112] directly return the optimized value of each parameter, instead of obtaining it via loss gradient computations. As such, studies including [112] and [113], highlight the viability of this approach, facilitating Q-DRL training through statistical error mitigation. By employing
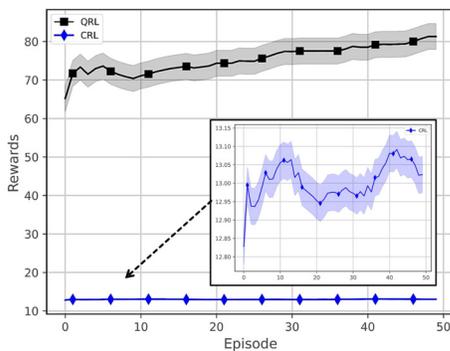

Fig. 8: An example of employing the parameter shift rule in quantum reinforcement learning, where the cumulative reward is in accordance with the performance indicators, particularly the achieved sum rate.

a Rotosolve-based approach, given the training loss L and the set of model parameters $\boldsymbol{\theta}$, the optimized i-th parameter, designated as $\hat{\theta}_i^{[k]} \in \boldsymbol{\theta}$, can be obtained as follows: $\hat{\theta}_i^{[k]} = -\frac{\pi}{2} - \arctan\left(\frac{2L_0 - L_+ - L_-}{L_+ - L_-}\right)$, in which L_0 indicates the loss where $\theta_i^{[k]} = 0$, L_+ marks the loss where $\theta_i^{[k]} = \pi/2$, and L_- signifies $\theta_i^{[k]} = -\pi/2$, towards optimized parameter $\hat{\theta}_i^{[k]}$, as defined in $\hat{\theta}_i^{[k]} = \arg\min_{\theta_i^{[k]}} L$. This approach has been a topic of emphasis in studies, as it can be applied in various settings, including reinforcement learning workflows to optimize VQE [114], and quantum-classical workflows for accelerated training convergence [115].

2) Search-Based Approaches: A number of studies have given a glimpse of promising directions for quantum search algorithms. In particular, [116] analyzes that the outputs of Grover's algorithm, a popular quantum search algorithm, might be replicable to a certain degree through the use of quantum variational circuits. In addition, [117] discusses the applicability of quantum search algorithms, such as Grover's and Dürr-Høyer algorithms, for diverse applications in wireless communications. The utilization of search algorithms, such as Grover's and Dürr-Høyer algorithms, aimed at optimizing model parameters and structures via training loss minimizations, represents a compelling research direction. Accordingly, [118] employs search methods that identifies the optimized structures of quantum embeddings.

3) Other Approaches: The following approaches can be considered for high-dimensional policy spaces and multi-agent DRL scenarios, where larger solution spaces are involved. (i) *Quantum annealing–based parameter optimization,* which involves defining an energy function over policy parameters and employs quantum annealing to identify low-energy (low-loss) configurations. This often requires the objective function to be reformulated into Ising and/or quadratic unconstrained binary optimization (QUBO) forms, as presented in studies such as [119]. (ii) *Bayesian-based parameter optimization,* involves integrating quantum kernel estimation within a Bayesian optimization algorithm, to model the policy landscape and guide parameter search. Studies such as [120] employ such an approach, iteratively refining the distribution of suitable quantum learning parameters over time, instead of deterministically selecting a parameter set.

IV. A TAXONOMY OF Q-DRL

In what follows, we discuss a Q-DRL taxonomy for Sem-Coms and DT-enabled networks, wherein we address various learning model architectures (see Subsection IV-A) and exploring distinct learning frameworks, such as employing multiple agents (see Subsection IV-B).

A. Quantum DRL Architectures

1) Variational Quantum Circuits and Quantum Kernels:
(i) *Variational quantum circuits.* Variational Quantum Circuits (VQCs) employ sets of parameterized gates, such as rotation-Y gates, allowing their output to be influenced through adjusting

gate parameters. This makes VQCs suitable for generalized learning models, as their parameters can be optimized to minimize training losses, akin to classical learning models. Different studies provide valuable insights into VQCs. In [121], different operations comprising VQC are discussed, highlighting how VQCs can be tailored based on factors, such as dimensionality of channel information, underscoring its general applicability for wireless networks, e.g., for maximizing the achieved rates in cell-free networks [122]. In addition, [101] discusses the different responses of different VQCs related to their training loss, and addresses aspects related to their trainability, such as how to derive the Hessian of their training losses. (ii) *Quantum kernels.* In certain cases, mapping input data into higher dimensions allows us to extract information that is not revealed in lower dimensions. To this end, different studies advocate the use of quantum kernels, leveraging the higher-dimensional Hilbert spaces applicable to quantum computing for important tasks such as user classifications. For instance, [123] discusses various quantum kernel-based approaches, e.g., those based on polynomial and Laplacian kernels, and their application to wireless systems, particularly in leveraging distributed quantum kernels for enhancing security.

2) Hybrid Quantum-Classical Learning Models: Depending on the sequence of model integration, the following configurations arise: (i) quantum-classical, (ii) classical-quantum, and (iii) combined architectures. (i) *Quantum-to-classical neural networks.* In quantum-to-classical architecture, input data is embedded first into the quantum NN, followed by the classical NN, as in [124]. The benefits to be had are (a) improving the precision of the QNN outputs, and (b) mitigating any quantization error that may present after quantum measurement, by adding non-linear classical layers afterward. In particular, [125] advocates quantum-classical NN adoption for beamforming optimization, a vital aspect for both present and future wireless communication systems, while [124] demonstrates their applicability to localize user terminals in a particular wireless network. (ii) *Classical-to-quantum neural networks.* In classical-to-quantum architecture, input data is first processed by classical NNs, from which the output is embedded as the input of QNNs. Using this configuration, feature extraction is conducted first through the classical NNs, thereby facilitating training convergence. Alternatively, the input dimensionality can be reduced, e.g., through classical convolutional NNs, prior to embedding the extracted features into QNNs, thereby reducing the number of qubits otherwise required to process high-dimensional data. As per [125], a classical convolutional NN is useful when extracting patterns from inputs such as images, thereby extensible to multi-dimensional inputs, as found in the wireless systems, e.g., MIMO channel information. (iii) *Combined architectures and frameworks.* Not being bound by the aforementioned configurations, both classical and quantum neural networks may be employed in any order, as follows. (a) Quantum and classical NNs as parallel learning models. In this case, the input is fed into both quantum and classical NNs, whose respective outputs jointly form the solution space, as exemplified in [126], facilitating both harmonic and non-harmonic model fitting. (b) Classical NNs to train quantum NNs (and vice versa). In such a case, exemplified by [127] the classical NN acts as an optimizer for the QNN's hyper-parameter, forming classical-quantum meta learning framework.

B. Quantum DRL Frameworks

1) Model-Free and Model-Based: (i) *Model-Free.* Model-free Q-DRL directly learns from interacting with the environment, with approaches including quantum-based adaptations of deep deterministic policy gradient (DDPG), state–action–reward–state–action (SARSA), deep Q-learning, and actor-critic. Hence, quantum-based parameterized learning models, such as variational quantum circuits, can be adopted within those model-free frameworks. For instance, quantum-based DDPG frameworks can utilize a quantum-based learning model to estimate optimal policy, while facilitating continuous action spaces, which are particularly relevant in cases involving moving transmitters, e.g., in vehicle-to-vehicle communications and unmanned aerial vehicle (UAV)-enabled systems. Furthermore, thanks to the higher effective dimension of quantum learning models [128], pioneering studies have adopted this approach to maximize wireless network performance: In particular, [129] leverages quantum-based DDPG for continuous action spaces, found in UAV-enabled wireless networks. Moreover, it can also be conveniently extended for other cases with static transmitters. Though not employing quantum-based parameterized models, [130] uses a DDPG-based approach to minimize base station transmit power, taking account of communication reliability and end-to-end latency targets as key constraints. Since model-free frameworks rely on direct interactions with the environment, there are possibilities to leverage DTs, allowing the agent to interact with digital environments instead of the real world, minimizing the risk of testing harmful solutions. In fact, such an approach to representing real-world environments aligns with DT visions, thereby affording opportunities for synergy, as elaborated in [131]. In this study, a DT, comprising long short-term memory (LSTM) and other deep neural networks, predicts states and rewards, while a DRL agent, leveraging deep-Q learning, learns and provides the policy.

(ii) *Model-based.* Unlike model-free Q-DRL, model-based Q-DRL policy learning relies less on trial-and-error interactions with real-world environments. Instead, it builds a virtual representation of the environment, i.e., a world model, to predict real-world responses without direct interaction. This potentially reduces the required training samples (data points acquired from the agent's interactions with the environment) facilitating training convergence, and curtailing consequences of testing sub-optimal policies in the real-world environment. In wireless networks, this minimizes the risk of quality-of-service degradation during training, ensuring compliance with SLA [132]. Various model-based reinforcement learning frameworks exist, with those commonly applied to wireless networks including partially observable Markov decision process (POMDP), and kernel-based reinforcement learning. Such

techniques have found use in numerous wireless network scenarios, chosen as they enable the planning of optimum solutions, e.g., routing and selected modulation schemes, as well as predicting the environment's responses, e.g., resulting delay, without direct interaction with the environment. In particular, [133] adopts two world models within a POMDP framework, representing channel occupancy and data traffic, facilitating performance maximization. Furthermore, model-based RL frameworks tend to yield higher computational complexity, as they require world models; thus, quantum-based kernels serve as alternatives to classical ones, alleviating processing demands associated with mapping data points to higher-dimensional feature spaces. The study in [134] demonstrates this by employing quantum-based unitary operations for quantum kernels. Not to mention, quantum information is processed through a higher-dimensional Hilbert space, which allows data points to be linearly separable, thus facilitating feature extraction and contributing to its utility for classification tasks.[7]

2) On- and Off-Policy: (i) *On-Policy.* The classification of on- and off-policy DRL frameworks depends on how they handle training experiences. On-policy frameworks typically compute model optimizations and evaluate policies solely using current experiences. Such an approach makes on-policy frameworks flexible and suitable for many scenarios, including their extension to meta-learning frameworks, as in [135], where an on-policy learning agent is used to optimize the hyper-parameters of other agents, leveraged to maximize packet delivery ratio. (ii) *Off-Policy.* Off-policy frameworks, unlike on-policy ones, reuse past learning experiences, for example by employing experience replay buffers, for current learning iterations. The benefits of such approaches include more efficient use of training data and (a) improved training loss convergence, though it may come with trade-offs, such as (b) higher computational space complexity and (c) heightened fluctuations in training results. Various studies have implemented off-policy frameworks, demonstrating performance benefits. In particular, [136], leverages a DRL framework with experience replay for optimizing transmit beamforming and the trajectories of moving users in a URLLC enabled network, leading to higher learning rewards. In addition, [137] employs a deep-Q network framework with an experience replay buffer in a UAV-enabled network, to optimize UAV trajectory, reconfigurable intelligent surface (RIS) phase shifting, and base station transmit precoding. The use of replay buffers is also applicable for Q-DRL frameworks. In particular, [138] employs Q-DRL with a quantum random walker and experience replay buffers to attain optimized explorations.

3) Single- and Multi-Agent: Single-agent Q-DRL assumes a solitary agent interacting with the environment, resulting in a single environment interaction for each iteration, while multi-agent Q-DRL dispatches multiple agents, allowing different agents to interact with the environment. Such multi-agent interactions may involve (i) multiple observations and states, and (ii) collaborations between different agents. (i) *Multiple observations and states.* To be specific, in a system with N_{agent} agents, at each iteration time t, we shall obtain $\{O_j^{(t)}, \mathcal{S}_j^{(t)}\}_{j=1}^{N_{\text{agent}}}$, with $O_j^{(t)}$ and $\mathcal{S}_j^{(t)}$ being the observable and state, respectively, of the j-th agent. This approach holds benefits for Q-DRL frameworks, notably facilitating richer explorations, as we can cover different action spaces, $\{\mathcal{A}_j^{(t)}\}_{j=1}^{N_{\text{agent}}}$, within a particular time frame, and effectively process varied rewards, $\{\mathcal{R}_j^{(t)}\}_{j=1}^{N_{\text{agent}}}$, concurrently. Furthermore, it is possible for agents to compete with each other in the same environments. Nonetheless, as in wireless networks we typically assume the agents are in fact controllable entities, this paper focuses on cases involving collaborating agents [80]. In such cases, possible drawbacks to be addressed include the extra computational burden, as the learning frameworks may need to estimate optimized policies associated with collaborating agents. (ii) *Agent collaborations.* Thereby, approaches facilitating agents to collaboratively interact with environments are to be explored, as highlighted in [139]. Rather than solely extending the action spaces of those in single-agent Q-DRL for different agents, this notion opens possibilities for employing action spaces that are infeasible in single-agent scenarios. For physical layer security, in particular, we may use a specific transmitter, corresponding to an agent, to send beam-focused jamming signals to a potential eavesdropper, while other transmitters are used to maintain quality-of-service for the legitimate users. Furthermore, this example highlights the need to develop individualized reward functions, as we now
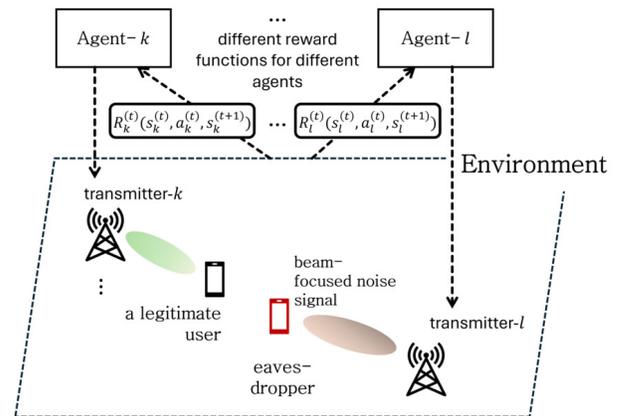

Fig. 9: An example of assigning distinct agent roles in a collaborative multi-agent approach, aimed to improve physical layer security against eavesdroppers. Since different transceivers here assume different action spaces, separate reward functions can be uniquely defined as well.

---

[7]Such quantum-based approaches are particularly useful when dealing with complex optimizations, as in network slicing [132], wherein a number of variables must be accounted for, including distinct inputs from eNB, mMTC, and URLLC slices, e.g., incoming and delivered traffic. In this regard, a quantum-based kernel can support a learning framework to optimize network slicing policies, e.g., resource block allocations under slice priority constraints, ensuring QoS and facilitating online learning, by maximizing performance with minimal prior training or data.

cannot use the same reward calculations. Further illustrated in Fig. 9, the reward calculation for transmitters sending signals to legitimate users may be based on their received data rates; it might not be the case for the one acting as a jammer, sending noise signals directed to the eavesdropper.

4) Integrating Q-DRL with Other Quantum Algorithms: It is also worthwhile to explore the interplay between Q-DRL and other algorithms, such as genetic algorithms. In particular, [140] discusses the integrations of genetic algorithms with machine learning algorithms, including those of reinforcement learning, while also exploring the use of quantum-inspired genetic algorithms. There are also possibilities to combine different operations and algorithms within the reinforcement learning frameworks, as in [141], which presents the utilization of hybrid tensor networks and variational quantum circuits within reinforcement learning frameworks.

V. THE ROLES OF Q-DRL IN SEMCOMS AND DT-ENABLED 6G NETWORKS

In the context of 6G networks, Q-DRL plays a key role in enabling advanced capabilities for SemComs and DTs. By integrating Q-DRL, 6G networks can achieve improved efficiency, adaptability, and intelligence, facilitating next-generation applications such as autonomous systems, smart cities, and immersive experiences.

A. Employing SemComs within 6G Networks

Unlike its predecessors, 6G is envisioned to transcend traditional data transmission paradigms by incorporating SemComs, an approach that focuses on transmitting meaningful information rather than raw data [142], [143]. This shift is critical to addressing the challenges posed by dynamic network conditions, time variability, and the increasing complexity of next-generation applications. The following explores the key aspects of employing SemComs within 6G networks.

1) From Data Transmission to Meaning Exchange: The paradigm shift in 6G SemComs represent a fundamental shift from conventional communication models, which prioritize the accurate delivery of bits, to a model that emphasizes the exchange of meaningful information [144]. In 6G networks, this paradigm is enabled by the following key aspects. (i) *Information context awareness.* SemComs leverage contextual information, such as user intent, environmental conditions, and application requirements, to optimize data transmission [145]. For example, in a smart city scenario, a traffic management system can prioritize transmitting critical alerts, e.g., accidents. (ii) *Knowledge graphs and ontologies.* These tools enable machines to understand and interpret data in a human-like manner. By structuring information hierarchically, SemComs reduce redundancy and ensure that only relevant data is transmitted. (iii) *Efficiency.* By focusing on meaning rather than raw data, SemComs can minimize bandwidth usage, improving resource efficiency. This is particularly crucial for 6G, which aims to support massive connectivity and ultra-low latency.

2) SemComs in Dynamic 6G Networks: 6G networks will operate in highly dynamic environments characterized by fluctuating bandwidth, latency, and user mobility [19], where SemComs are suited to address these challenges. (i) *Time-varying network conditions.* In scenarios where network resources are limited, SemComs adapt by prioritizing the transmission of certain information. For instance, during network congestion, a video streaming service might transmit only key frames or semantic descriptors instead of the entire video stream. (ii) *Real-time optimization.* Various algorithms enable semantic encoders and decoders to dynamically adjust their strategies based on real-time network feedback. This ensures consistent performance even under unpredictable conditions.

B. Cross-Layer Design and Resource Allocation in 6G Networks

By enabling collaboration across the physical (PHY), medium access control (MAC), and application layers, cross-layer optimization addresses the challenges of SemComs in dynamic 6G environments. The following discussions explore how cross-layer design can enhance SemComs, focusing on network slicing, dynamic QoS/QoE metrics, and efficient resource allocation [146].

1) Collaboration Between PHY, MAC, and Application Layers: Cross-layer optimization eliminates the traditional silos between network layers, enabling seamless coordination to prioritize semantic information. (i) *PHY layer:* At the physical layer, adaptive modulation and coding schemes can be tailored to the importance of semantic payloads. For example, critical semantic data, e.g., emergency alerts, can be transmitted using robust modulation techniques to ensure reliability, even under poor channel conditions [147]. (ii) *MAC layer:* By integrating semantic awareness, the MAC layer, playing an important role in scheduling and resource allocation, can prioritize transmissions based on the meaning and urgency of the data [148]. For instance, in a smart factory, control signals for robotic arms can be given a higher priority than routine sensor data. (iii) *Application layer:* The application layer provides context and intent, which are essential for SemComs. By sharing this information with lower layers, the network can optimize resource allocation to meet application-specific requirements [149], even under varying network conditions.

2) Network Slicing for Semantic Services vs. Conventional Data Services: Network slicing allows the creation of virtualized network segments tailored to specific services. In this context, SemComs introduce unique requirements that differ from conventional data services: (i) *Semantic service slices.* These slices are optimized for low-latency, high-reliability transmission of meaningful information [150]. For example, a network slice for autonomous vehicles would prioritize semantic data related to navigation and safety. (ii) *Conventional data slices.* These slices handle traditional data transmission, such as file downloads or video streaming, where latency and reliability requirements may be less stringent. By dynamically allocating resources to semantic and conventional slices, 6G networks can ensure that both types of services coexist efficiently [151].

3) Dynamic QoS and QoE Metrics for Semantic Payloads: SemComs require a rethinking of traditional QoS and QoE metrics to account for the importance of meaning: (i) *Dynamic QoS.* Unlike conventional data, semantic payloads may have varying levels of importance. In a healthcare application, patient vitals may require higher QoS than routine updates. Cross-layer optimization enables dynamic adjustment of QoS parameters based on the semantic context [152]. (ii) *QoE for SemComs.* QoE metrics for SemComs focus on the comprehensibility and relevance of the transmitted information. In a video conferencing application, QoE can be measured by the clarity of the speaker's intent rather than the video resolution [153].

C. Semantic Encoding and Decoding Strategies

SemComs rely on advanced encoding and decoding techniques to minimize redundancy and maximize comprehension. AI-driven approaches, particularly those leveraging Natural Language Processing (NLP) and ontology-based methods, are at the forefront of this innovation [154]. The following outlines these strategies and their role in adapting to changing network conditions.

1) NLP and Deep Learning Approaches for Semantic Extraction: AI-driven semantic encoding and decoding rely heavily on NLP and deep learning to extract and transmit meaningful information: (i) *Semantic extraction.* NLP models, such as transformers, can analyze text, audio, or video data to identify key semantic elements [155]. For example, in a voice assistant application, the system can extract the user's intent (for instance, "to play music") rather than transmitting the entire audio stream. (ii) *Deep learning for compression.* Deep learning models can compress semantic information by removing redundant data while preserving meaning [156]. For instance, a video streaming service can use AI to transmit only the most relevant frames or semantic descriptors.

2) Ontology-Based Methods to Capture Domain-Specific Meaning: Ontologies provide a structured framework for representing knowledge, enabling domain-specific SemComs. (i) *Domain-specific ontologies.* In healthcare, for example, ontologies can encode medical knowledge, allowing SemComs system to transmit patient data in a concise and meaningful way. (ii) *Interoperability.* Ontologies enable seamless communication between systems with different data formats, ensuring that semantic information is accurately interpreted across domains [157].

3) Context-Awareness and Personalization in Semantic Compression: Context-awareness and personalization are key to optimizing semantic encoding and decoding: (i) *Context-aware compression approaches.* By considering the context in which data is transmitted, systems of SemComs can prioritize relevant information. In a navigation application, the system can focus on transmitting route updates rather than static map data [146]. (ii) *Personalizations.* AI models personalize SemComs to match user preferences, allowing a news app to align summaries with interests, thus enhancing efficiency and relevance.

4) Trade-offs Between Interpretability and Efficiency: While AI-driven SemComs offer significant benefits, it also introduces trade-offs: (i) *Semantic interpretability.* Complex AI models, such as deep neural networks, may lack transparency, making it difficult to understand how semantic information is encoded and decoded. (ii) *Semantic efficiency.* Highly interpretable models, such as rule-based systems, may be less efficient in compressing and transmitting semantic information [151]. Balancing these trade-offs is essential to developing semantic systems that are both efficient and trustworthy.

D. SemComs-Assisted 6G Architecture

The integration of SemComs into 6G networks represents a transformation from typical data-centric to meaning-driven transmissions [158]. This transformation addresses the growing complexity of communication systems, the data explosion, and the demand for intelligent, context-aware networks [156].

1) Meaning-Centric Communication: Traditional communication systems focus on the accurate transmission of bits, often without considering the meaning or context of the data [147], [159]. In contrast, meaning-centric communication emphasizes the transmission of semantic information, such as intent, context, and knowledge, ensuring that the communication process aligns with the user's goals and the application's requirements. (i) *AI-driven intelligence.* Artificial Intelligence (AI) plays a central role in extracting, encoding, decoding, and interpreting semantic information. AI algorithms, such as deep learning and reinforcement learning, enable the system to understand and process complex semantic data [160]. AI extracts the intent behind a user's request, enabling the network to tailor its response accordingly. (ii) *Dynamic adaptability.* The architecture adapts to varying network conditions, user requirements, and application contexts, ensuring optimal performance in dynamic and unpredictable environments [161]. In a mobile network, the system adjusts its parameters to ensure seamless communication as users move between different coverage areas.

E. Semantic-Assisted Protocol Layers

The aforementioned SemComs-assisted 6G architecture redefines the functionality of traditional protocol layers to incorporate semantic awareness [162].

1) Semantic-Assisted Application Layer: The application layer serves as the entry point for SemComs, where meaningful information is extracted (e.g., derived from user intent), contextualized, and prepared for transmission. (i) *Semantic extraction.* AI-driven techniques, such as Natural Language Processing (NLP) and computer vision, extract semantic information from raw data [163]. In video conferencing, semantic extraction captures the speaker's intent, emotions, and key messages, rather than the full video stream, thus reducing data volume while preserving essential meaning [164]. (ii) *Context awareness.* The application layer provides contextual information, such as user intent, location, and application requirements, which is then shared with the lower layers to optimize communication [165]. In a healthcare application, the system prioritizes the

transmission of critical patient data over routine updates. (iii) *Knowledge graphs and ontologies.* These tools structure and represent semantic information in a machine-interpretable format, enabling efficient communication between domains [166]. Knowledge graphs link various data sources, while ontologies facilitate semantic understanding. In smart city applications, knowledge graphs depict the connections between traffic sensors, vehicles, and pedestrians, allowing for informed decision-making [167].

2) Semantic-Assisted Transport Layer: The transport layer ensures the reliable and efficient delivery of semantic information across the network, managing the data flow to facilitate timely and error-free transmission [168]. (i) *Semantic-aware congestion control.* Critical semantic data, such as emergency alerts, is given higher priority to ensure timely delivery, even in congested network conditions [164]. The system dynamically adjusts the transmission rate based on the importance of the data and the current network state. (ii) *Adaptive error correction.* Semantic-aware error correction techniques reconstruct lost or distorted semantic information. In scenarios where minor data losses do not compromise the overall meaning, these techniques allow for more efficient use of network resources [169]. In a video streaming application, the system reconstructs missing frames without retransmitting the entire video stream [170]. (iii) *Dynamic resource allocation.* The transport layer collaborates with the lower layers to allocate resources based on the importance of semantic information [171]. High-priority semantic data receives the necessary bandwidth and processing power. In a gaming application, the system allocates more resources to the transmission of real-time control signals, ensuring responsive experience on the users' side [168].

3) Semantic-Assisted Network Layer: The network layer routes and forwards semantic information across the network, ensuring data is delivered to the correct destination while considering its meaning and context [172]. (i) *Semantic routing.* Enhanced routing algorithms consider the meaning and context of data, making routing decisions based on importance and urgency. In a smart city application, routing decisions prioritize data related to public safety, such as traffic accident alerts, over routine updates like weather reports [173]. (ii) *Intent-based forwarding.* The network layer uses intent information provided by the application layer to make forwarding decisions, delivering semantic information to the most relevant destinations [174]. In a healthcare application, patient data is forwarded to the appropriate medical professionals based on their expertise and availability [175]. (iii) *Network slicing for semantic services.* The network layer supports the creation of dedicated slices for SemComs [176].

4) Semantic-Aware Scheduling and Resource Allocation of MAC Layer: The Medium Access Control (MAC) layer manages access to the shared communication medium, ensuring efficient and reliable transmission of semantic information. (i) *Semantic-aware scheduling.* The MAC layer prioritizes transmissions based on the meaning and urgency of semantic information [177]. By analyzing data content to determine priority [178], AI in industrial IoT scenarios may schedule

control signals for robotic arms ahead of routine sensor data to ensure real-time responsiveness. (ii) *Dynamic resource allocations.* The MAC layer dynamically allocates resources, such as time slots and frequency bands, based on the semantic context and network conditions [179]. In a smart grid application, critical control signals are allocated more resources to ensure grid stability and reliability [180]. (iii) *Collaborations with PHY layer.* The MAC layer works closely with the physical layer to optimize modulation and coding schemes for semantic payloads [181], while feedback on the importance of the data enables the physical layer to adjust its modulation and coding schemes accordingly.

5) Physical Layer: The physical layer converts data into signals for transmission over the wireless medium, ensuring they are received correctly at the destination [182]. (i) *Semantic-aware modulation.* Adaptive modulation schemes optimize the transmission of semantic information based on channel conditions, and ensure the reliability of critical semantic data transmission [183]. (ii) *Semantic-aware coding.* Various coding techniques, such as semantic error correction codes, protect semantic information from errors and distortions [184], preserving the meaning of the transmitted data, even in the presence of noise or interference. (iii) *Energy efficiency.* The physical layer minimizes energy consumption by transmitting only the most relevant semantic information, reducing the overall data volume [185], [186]. In a wearable device, the system transmits only the most critical health data, reducing energy consumption and extending battery life.

F. Leveraging DTs for 6G Networks: Towards Autonomous Digital Entities

In the following, we address various use cases of DTs, highlighting their synergy with Q-DRL, chiefly pertinent to (1) security and (2) their operations. Indeed, Q-DRL interplay with DTs are prominent in modeling various deployment scenarios of the physical entities, often referred to as level-4 DTs, or even, in achieving autonomous DTs, designated as level-5 DTs [187].

1) Digital Twins for Security: The importance of DTs for future wireless networks is particularly evident in their role for (i) covert communications and (ii) physical layer security. On the other hand, DTs take advantage of the advancements in network security, particularly (iii) data encryption. Each of these key aspects will be discussed in the following.

(i) *Physical layer security.* Physical layer security is mainly attained through exploiting, and influencing, the physical channels between transmitters and receivers, aiming to limit the adversary's decoding ability. In general, it can be defined as the difference between physical layer information received by the legitimate user and the adversary, that attempts to eavesdrop on the signal. Therefore, physical layer security maximizations, at time t, can be represented as [188]: $\max_{\mathbf{v}^{(t)}} [R_k^{(t)} - R_{\text{adv}}^{(t)}]_+$, in which $R_k^{(t)}$ denotes the rate of the k-th user under consideration, $R_{\text{adv}}^{(t)}$ marks the rate of the adversary eavesdropping the information, while $[m]_+$ stands for max $\{0, m\}$, for m is a given

scalar. In addition, $\mathbf{v}^{(t)}$ represents a particular physical layer parameter, such as transmit power, in which its optimizations are based on the channel conditions. In this matter, DTs can model physical wireless channels of the users, facilitating the optimization of the parameters relevant in maximizing the physical layer security, for example, transmit beamforming. In particular, the study in [189] employs a DT to represent the characteristics of the wireless channel, such as pathloss and delay spread. More importantly, it highlights the role of AI in extracting and learning the characteristics of real physical channels, which can be extended to Q-DRL, particularly for modeling the channels of satellite-ground communications [190].[8]

(ii) *Covert communications.* Covert communications aim to conceal the existence of the communications from adversaries. To this end, DTs can be used to model the communication channel, as a means of maximizing covert rate. For example, given the channel condition, such maximizations can be conveyed as [191]: $\max_{\mathbf{v}^{(t)}} \mathrm{R}_k^{(t)}$, such that $\chi_i^{(t)} \geq 1 - \mu$, where $\mathrm{R}_k^{(t)}$ indicates the rate of the legitimate user k, $\mathbf{v}^{(t)}$ denotes a certain physical layer parameter, $\chi_i^{(t)}$ denotes error detection probability of the attacker detecting the existence of transmissions at t, while μ is the predefined constant indicating the concealment of the communication.[9] In addition, covert communications can be extended to other network scenarios as well, such as UAV-enabled networks. In this context, UAVs can act as friendly jammers, emitting noise signals to degrade the effective rate of the adversary, as emphasized in studies such as [192]. Multiple DTs can represent multiple UAVs and, when integrated with Q-DRL, can effectively optimize jamming operations, to maximize the covert rate. It is worth highlighting that Q-DRL-based approaches have been employed for optimizing multiple UAVs in a wide range of studies, as showcased in [193], which employs multiple agents with a centralized quantum-based critic and several quantum actor networks.

(iii) *Data encryption.* Recently, there is growing effort targeted at improving intra- and inter-twin communications between distinct DTs.[10] To this end, security protocols based on post-quantum cryptography can be used: The study in [194] promotes the integration of post-quantum cryptography into blockchain frameworks, enhancing the security of IoT networks. Interestingly, it also addresses the role of post-quantum cryptography pertinent to IoT networks (and arguably, wireless networks in general) at different layers: particularly consensus and data sub-layers, employed to handle transaction integrity and records, respectively, within blockchain frameworks. Another crucial aspect, with respect to SemComs and DT-enabled networks, is the risk of semantic knowledge

poisoning, as highlighted in [195]. Nevertheless, such risks can be mitigated through adversarial training, by exploiting Q-DRL (further elaborated in Sub-subsection VI-A).

2) Quantum DRL for DT Operations: Notably, Q-DRL can also support DT-enabled wireless networks in multiple aspects, among which are optimizing (i) DT processing and (ii) operations.

(i) *Optimizing DT processing.* Q-DRL can be utilized for optimizing processing sequences in DT, aiming to minimize computational latencies and energy consumption, alongside other objectives. Such objectives involve optimizing DT deployment, in terms of computational resource allocation and synchronization between physical and digital entities, as demonstrated in studies such as [196]. When several processing units are used in executing DTs, it might be necessary to optimize the computational migrations of the DTs, chiefly when multi-tier computing is employed [197]. In such instances, it is beneficial to employ multi-tier quantum learning models, wherein higher-tier, cloud-based quantum processing units govern the computational task offloading, particularly for DT migrations, while lower-tier, edge-based units optimize their corresponding communication and computation resources. This notion can even be extended to Q-DRL, as illustrated in Fig. 10, as the computational requirements of each edge shall vary over time, depending on such factors as the arrival of clients in that particular area. In the figure, each edge unit is capable of processing DTs, and $T_{m \to n}^{(t)}$ and $T_{n \to m}^{(t)}$ respectively denote the task offloading from the m-th edge to the n-th edge, and the reverse thereof.
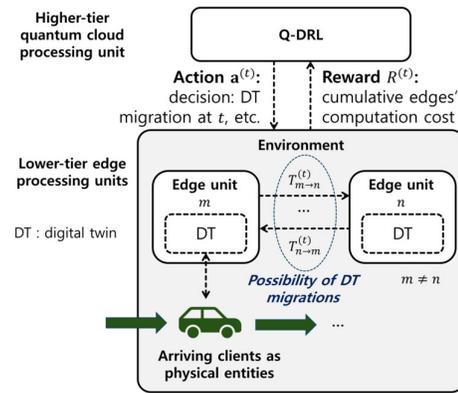

Fig. 10: One example of Q-DRL utilizations in multi-tier computing, used for task offloading, pertinent to the DT migrations.

(ii) *Optimizing DT operations.* Q-DRL is suitable for optimizing DT operations in terms of synchronization and scheduling. Though these factors are vital for DTs, for they directly pertain to how effectively DTs represent their physical counterparts, optimizing them is not trivial, owing to the challenges such as the high-mobility of user equipment, as in vehicular edge computing (VEC) networks. In this regard, Q-DRL can be used to minimize the processing latency, and/or other

---

[8]Furthermore, this opens possibilities to not only model the channel condition of the legitimate users, but also those of adversaries, at least in terms of their statistical channel information, as studied in [188].

[9]Here, it is possible to define additional constraints pertinent to the assumed system model, particularly maximum transmit power, as in [191].

[10]In his context, intra-communication addresses the information exchanges between a single twin entity and its environment, while inter-communication refers to communications between different DTs.

important performance indicators affected by it, including age-of-information. Such optimizations are demonstrated in studies such as [198], in which age-of-information minimizations are considered, although Q-DRL utilizations remain lacking. In particular, the reward pertinent to age-of-information minimizations, given M mobile edge computing (MEC) processors, at iteration time t, can be defined as: $R^{(t)} = \sum_{m=1}^{M} \xi_m^{(t)} \cdot \xi_m^{(t)}$ signifies the AoI value corresponding to the m-th MEC. If a synchronization data packet is received at a particular time slot t, it can be defined as $\xi_m^{(t)} = t - a_m^{(t)}$, where $a_m^{(t)}$ denotes the packet generation time slot, and $\xi_m^{(t)} = \xi_m^{(t-1)} + 1$ if otherwise [198]. Lastly, in cases involving large-scale deployments of interconnected MECs, Q-DRL offers benefit in processing and, thereby, facilitates extensive DT operations. This notion has been explored in studies such as [199], in which Q-DRL is used for minimizing processing latencies for task offloading.

To secure such operations, the following mitigation strategies address widely known risks. i) *Backdoor Attacks.* Adversaries may carry out backdoor attacks, particularly by incorporating malicious models/parameters into the Q-DRL workflow. These attacks can be especially harmful in a non-centralized scheme, such as federated learning, in which prediction accuracy relies on aggregating possibly tampered models [200]. ii) *Distributed Denial-of-Service (DDoS).* Adversaries may flood the computational and communication resources within Q-DRL pipelines via DDoS-based attacks, possibly extending the training duration and introduce training latency. These attacks can be especially damaging when multi-agent Q-DRL is used in time-critical scenarios (like V2X networks), as such resource congestion may greatly affect time-sensitive actions [201]. To mitigate these risks, credibility assessments of both physical and digital entities can be imposed, introducing a trust analysis layer [202]. In this regard, the credibility of the n-th entity (among the N_{ent} entities) can be computed as:

$$\vartheta_n = \frac{\sum_{m=1, m \neq n}^{N_{ent}} \epsilon_{sat, m \to n}}{N_{ent} - 1}, \tag{4}$$

where $\epsilon_{sat, m \to n}$ denotes the satisfaction score for the n-th entity as given by the m-th entity. Let us assume that $\mathbf{L}$ is an $N_{ent} \times N_{ent}$ matrix containing all such scores, so that $\epsilon_{sat, m \to n} \in \mathbf{L}$. In the case of mitigating backdoor attacks, $\epsilon_{sat, m \to n}$ may represent the contribution of the n-th entity as evaluated by the m-th entity, in which can be estimated using Shapley value.

3) Leveraging DTs for Q-DRL: Having discussed how Q-DRL can be beneficial for DTs, the following highlights practical approaches to utilizing DTs in facilitating Q-DRL training.

- *DTs to model physical environments.* By this approach, we may employ a DT comprising learning models to model a dynamic physical environment, facilitated by real-time interactions between the DT and environment. As such, we can treat the wireless environment as a physical entity, interacting with (and/or cloned by) the DT, which acts as a digital entity. The study in [203] uses such an approach, in which generative and predictive models are assembled to build a DT, and attained reduced task processing delay

in UAV-enabled wireless networks. Interestingly, when the DT is able to precisely model the environment, we can limit Q-DRL interactions with the physical environment, reducing trial and error.

- *DTs to facilitate training convergence.* By this approach, DTs can be integrated into twin-delayed deep deterministic policy gradient (TD3) algorithm (or other similar RL algorithms). The study in [204] employs this approach to improve battery energy storage systems. DTs may take action variables, such as battery charging rates, as input values and yield estimated state variables, such as the remaining battery life, as output, mimicking physical entities and making this approach applicable to energy-aware scenarios, such as low-energy wireless sensor network and ambient backscattering communications. The outcomes of DTs-assisted training can also enrich the information in experience replay, as we can obtain more data on action-state records, compared to training with physical entities alone.

VI. ADOPTING SECURE Q-DRL TO MAXIMIZE PHYSICAL LAYER SECURITY

There is a growing interest in exploring the security aspects of Q-DRL, which is being driven from two perspectives: (i) improving the security of Q-DRL itself, and (ii) using Q-DRL to enhance the security of 6G, particularly SemComs and DT -enabled networks.

A. Toward Secure Q-DRL

As discussed in studies such as [205], protecting the integrity of quantum learning models requires enhancing data security, through adding verifications to ensure that the training data is not compromised by attackers, and exploited for malicious means such as data poisoning. The risks of data poisoning and other attacks also give rise to methods such as adversarial training, as these attacks might negatively affect learning model prediction capability, whether classical or quantum training data is used, causing incorrect predictions, as highlighted in [206]. Accordingly, the study in [207] presents a secure encoding approach through randomization, and highlights that this approach is apt for reinforcement learning as well. The importance of Q-DRL security becomes even more evident when we consider its operation as machine-learning-as-a-service [208], as this may introduce risks such as model extraction attacks, making it possible for attackers to illegitimately retrieve model parameters and structures, thereby replicating the response of the trained model. In addition, distributed learning schemes, where the trained models are shared across processing units, may be assailed for model extraction attacks, as addressed in [209], this is especially pertinent in cases such as federated Q-DRL.

Further, assuming a non-centralized Q-DRL workflow, approaches such as secret key generation need to be adopted to mitigate the risks of privacy leaking and keep the integrity

of information exchange. This is particularly relevant to Q-DRL, as various types of information such as experience replay and trained parameters may be exchanged, rendering the processing nodes vulnerable to data or model poisoning [210]. In this context, universally blind quantum computing could improve the security of information exchanges between quantum-processing nodes [211].

B. A Practical Adoption

The following discussions highlight Q-DRL utility for security. First, the assumed system model and the optimization objective are addressed in Subsection VI-B1. Furthermore, drawing from the steps elaborated in Section III, we present the Q-DRL utilization in Subsection VI-B2.

1) The System Model and Optimization Objective: Let us consider a wireless network, with M access points (APs) performing downlink communications to K_{user} designated user equipments, each of which is equipped with a single antenna. Each AP is able to adjust its transmit power. The challenge is, there are K_{adv} adversary devices, each of which is equipped with a single antenna, and is assigned to eavesdrop on the message designated to a legitimate user in its proximity, where $K_{\text{adv}} = K_{\text{user}}$. To counteract, the APs strive to improve the physical layer security, by maximizing the secrecy rate of each legitimate user, under the threats of its eavesdropper (such notions is also discussed in Sub-subsection V-F1). The objective, at each iteration time t, articulates as thus:

$$\underset{\mathbf{p}^{(t)}}{\text{maximize}} \quad \zeta^{(t)} = \sum_{k=1}^{K_{\text{user}}} \left[\mathsf{R}_k^{(t)} - \mathsf{R}_{\text{adv},k}^{(t)} \right]_+ \tag{5a}$$

$$\text{subject to} \quad p_m^{(t)} P_{\text{max}} \leq P_{\text{max}}. \tag{5b}$$

$\zeta^{(t)}$ denotes the objective function, i.e., the cumulative secrecy rate. $\mathsf{R}_k^{(t)}$ and $\mathsf{R}_{\text{adv},k}^{(t)}$ denote the achieved rates of the k-th legitimate user and its associated eavesdropper, respectively, at t. $\mathbf{p}^{(t)} = [p_m^{(t)}]_{m=1}^M$ denotes the APs' transmit power coefficient. P_{max} denotes each AP's maximum transmit power. Furthermore, (5b) satisfies the transmit power budget. Without loss of generality, here we assume that $M = K_{\text{user}}$, with each m-th AP being assigned to the k-th user, in a one-to-one relationship.

2) Developing Q-DRL for the Task: Accordingly, the action space, at each time t, comprises transit power adjustments, defined as: $A^{(t)} = \mathbf{p}^{(t)}$. The state space, at each time t, comprises the channel information of the users, and can be defined as: $S^{(t)} = \mathbf{H}^{(t)}$, where $\mathbf{H}^{(t)}$ conveys the channel matrix at t. The reward, at t, can then be formulated based on (5): $R^{(t)} = \mathfrak{s}\zeta^{(t)}$, where $\mathfrak{s}$ denotes the adjustable coefficient, and is part of the hyper-parameters. In this context, increments of $R^{(t)}$ indicate improvements in secrecy rate. Then, a quantum learning model, akin to the one used in [212], is employed, and the parameter-shift rule is employed to optimize the model parameters.

The achieved cumulative secrecy rates, with respect to the training episodes are depicted in Fig. 11. In this instance,
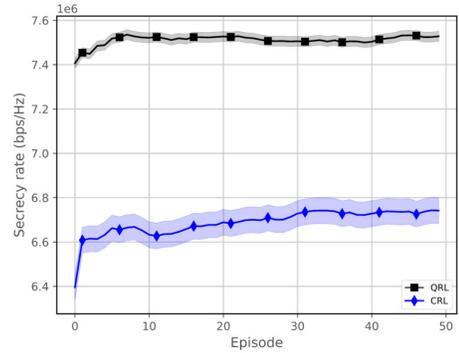

Fig. 11: The average achieved secrecy rates, with respect to the number of the episodes, of the Q-DRL utilizations detailed in Section VI-B.

$\mathbf{H}^{(t)}$ is generated for each t, requiring Q-DRL to respond to the changing environment state. As illustrated in the figure, the secrecy rate increases in accordance with the number of learning iterations, indicating Q-DRL effectiveness, even with an exceptionally high ratio of the number of eavesdroppers versus that of the legitimate users, in this case, $K_{\text{user}} = K_{\text{adv}}$. The reason for this lies in the capability of Q-DRL to learn the optimized $\mathbf{p}^{(t)}$ over time, amidst the dynamic channel information and high number of eavesdroppers.[11] Impressively, Q-DRL achieved higher secrecy rates compared to classical deep reinforcement learning (termed 'CRL' in Fig. 11). In our analysis, this can be attributed to the higher dimensionality of information processing in quantum models. These results also confirm findings from related studies: The study in [213] employs a quantum-inspired reinforcement learning (QiDRL) workflow to optimize the flight trajectory of aerial base stations, resulting in faster convergence and higher reward values compared to classical Q-learning. Furthermore, [43, Fig. 11] shows that DRL with variational quantum circuits requires fewer parameters (and thus less memory to store model information) compared to classical deep-Q learning, while attaining comparable performance in cognitive radio channel selection.

There exist opportunities to improve the security even further, particularly for SemComs and DT-enabled networks. (i) When dealing with sensitive data, it is possible to employ blockchain, and even distribute the learning process. For example, [214] adopts blockchain to improve the security of healthcare data, pertinent to federated learning scenarios, while [215] employs blockchain for DT-enabled IoT networks. (ii) Moreover, there are risks of semantic noise, wherein adversaries

---

[11]Regarding the results of Fig. 11, we assume that the eavesdroppers, potentially situated in concealed places to avoid detection, experience reduced channel gains, compared to the legitimate users. Hence, given $\mathbf{h}_k^{(t)}$ and $\mathbf{h}_{\text{adv},k}^{(t)}$ respectively as the channel of the k-th legitimate user and its eavesdropper, we assume $|\mathbf{h}_k^{(t)}|^2 \geq |\mathbf{h}_{\text{adv},k}^{(t)}|^2$, for all t and k. In addition, we set $K_{\text{user}} = 2$ and $K_{\text{adv}} = 2$. The constraint of (5b) is satisfied by normalizing each quantum circuit's outputs, each of which is responsible for yielding $p_m^{(t)}$, $\forall m, t$, respectively unto $[0, 1]$.

exploiting machine learning, to minimize semantic distances, may compromise the SemComs [216]. As a countermeasure, Q-DRL can be combined with generative AI to mitigate the semantic noise, akin to denoising approaches based on stable diffusion models in [217].

VII. CONCLUSIONS

This paper aims to promote the adoption of Q-DRL, an emerging quantum technology, for SemComs and DT-enabled wireless networks, with a focus on security. To this end, this paper offers a guide in Section III on using Q-DRL for such networks, including discussions on how to train learning model parameters. In addition thereto, this paper elaborates on the taxonomy of Q-DRL, which is surprisingly deficient in current literature, opening possibilities for further Q-DRL developments in Section IV. Thereafter, the paper discusses specific roles of Q-DRL in SemComs and DT-enabled networks, highlighting its importance and illustrating its role in strengthening physical layer security. As this topic remains in its infancy, the following reveals the prospective research directions and use cases.

A. Challenges and Future Research Directions

While the potential of Q-DRL is evident, there remain challenges pertaining to its adoption in wireless networks. It is worth noting that these challenges present compelling research directions. (i) *Security aspects of the wireless networks.* Leveraging Q-DRL might require user information to be sent as datasets, which could open vulnerabilities such as man-in-the-middle and eavesdropping attacks, especially when classical channels are being used. Employing quantum security protocols can be considered a future endeavor: For instance, [218] explores the possibility of utilizing quantum teleportation protocols to enhance the quantum neural network security, especially in federated learning settings. Interestingly, machine learning methods can themselves be used to strengthen quantum security protocols, as shown in [219]. Furthermore, network-scale employment of quantum security protocols facilitate the advancement of the quantum internet, as discussed in [220]. (ii) *Quantum-classical integrations.* Although the integration of quantum-classical processing may introduce computational bottlenecks (for instance, quantum learning operations require quantum measurements in obtaining classical valued optimization solutions) it opens up possibilities for quantum-classical integration, such as in hybrid quantum-classical DRL, as discussed earlier in Section IV-A2. Moreover, these issues may also arise from the fact that datasets are often stored as classical values. Nevertheless, such bottlenecks can indeed be turned into opportunities to streamline processing: For instance, the study in [221] employs a classical learning model to reduce the dimensionality of training datasets before embedding them into subsequent quantum learning operations, thereby minimizing logical qubit requirements.

B. Compelling Use Cases in Communication Networks

Beyond semantic communications and physical-layer security, here are further scenarios where Q-DRL could play a transformative role in the context of 6G networks and DTs.

(i) *Intelligent UAV swarm coordination.* With the surge of drones and UAVs in 6G-enabled smart cities, coordinating fleets of UAVs becomes increasingly complex: Q-DRL can leverage quantum superposition to explore multiple trajectory possibilities simultaneously, improving convergence speed in crowded airspaces. Through quantum-inspired exploration, a swarm of UAVs can avoid collisions in real time while meeting mission objectives, e.g., delivery and surveillance. A DT environment of the city's airspace can simulate weather conditions, building structures, and flight restrictions. Q-DRL agents trained in this DT can be rapidly deployed for swarm coordination.

(ii) *Quantum-enhanced federated learning in edge networks.* Federated learning (FL) becomes pivotal in 6G as it enables distributed model training across multiple edge devices. By embedding quantum techniques in local or global model updates, FL can expedite gradient search and exploration. In addition, parameter aggregation under quantum encryption schemes can enhance privacy. As 6G networks will have edge devices with varying computational capabilities, Q-DRL can manage computational resource allocation, e.g., CPU/GPU/quantum processing unit usage, handling large combinatorial possibilities in real time. Therefore, it is plausible to integrate Q-DRL into a larger federated learning (FL) workflow, where FL enables edge processors to jointly learn loss gradients, which can accelerate convergence and lighten the computational burden on central processors (e.g., cloud servers) [211]. In this wise, quantum FL, along with other non-centralized quantum learning model workflows, can be integrated with quantum invariant graph neural networks [222], useful in operating context-aware semantic wireless networks [223]. Different quantum-powered edge processors, for example, can be clustered based on the similarity of the task they perform.

(iii) *Multi-agent coordination for network slicing.* Network slicing in 6G allows virtual network slices with different Quality-of-Service (QoS) requirements (e.g., for IoT, AR/VR, and critical infrastructures). Each slice has unique latency, bandwidth, and reliability constraints. Q-DRL agents can continuously adapt slices to meet SLAs, even under dynamic network conditions (e.g., surges in user demand or hardware failures). Quantum-inspired multi-agent coordination protocols can reduce overhead among distributed network orchestrators, optimizing the system globally. In a DT context, each network slice and its performance metrics are simulated to expedite learning convergence.[12] Multiple Q-DRL agents can be complemented with quantum graph neural networks, to represent the interconnection between network nodes [226], potentially outperforming isolated classical DRL approaches.

---

[12] Classical DRL has been used for various use cases in network slicing, demonstrating the feasibility of this research direction. In particular, resource management across slices for different services [224] and varying topologies [225] can be investigated.

(iv) *Facilitating network automation.* The integration of quantum technology into 6G will substantially support operators, as it will facilitate network automation, reducing human intervention. For instance, this includes learning the network structure, enabling routing optimization, and anomaly detection [227]. Furthermore, such structural learning approaches can be handled by quantum graph neural networks (to extract interconnections between network nodes, facilitating social-aware network optimization) and by quantum learning models with knowledge distillation [206], which can improve network security by inferring the probability distribution of possible attack scenarios.

C. Beyond Communications: Emerging Use Cases

Finally, Q-DRL employments are not exclusively for advancing communication networks. They also facilitate the development of other areas, such as healthcare and the Metaverse. (i) *Smart healthcare.* Q-DRL, along with other quantum technologies, can enhance the deployment of smart healthcare, with use cases ranging from health risk detections to medical data privacy. For instance, Q-DRL and other quantum machine learning approaches can be used to detect abnormal health conditions, as in [228], in which quantum convolutional neural networks were utilized. Further, protecting medical data privacy is important, given that such data contains patients' sensitive information, such as their health records. To this end, the networks of quantum communication-enabled nodes will facilitate secure quantum internet, enhancing the privacy and security of medical data sharing [229]. (ii) *Metaverse and extended reality (XR).* Eventually, Q-DRL, and other quantum technologies, can be introduced to enhance the Metaverse and XR, since such technologies are reshaping humans interaction with other individuals and machines. In particular, quantum computing is apt to address the growing computational demands of processing the Metaverse, particularly when combined with SemComs [230]. Moreover, [231] discusses AI employments in enhancing the information extraction and processing of XR equipments, among other purposes, and thus highlighting the potential utility of Q-DRL.

D. Concluding Remarks

In conclusion, Q-DRL represents an exciting frontier for future wireless networks as it potentially accelerates learning and handles enormous state-action spaces, compared with its DRL counterpart. As 6G evolves toward extremely high data rates, ultra-low latency, and AI-native infrastructures, Q-DRL will likely play a pivotal role in empowering DTs and SemComs, and improving security. Together with the emerging 6G technologies and advanced ML techniques at the edges, Q-DRL can become a major enabler for diverse tasks across next-generation networks (including mission-critical tasks), supporting intelligent DTs, secure SemComs, and beyond.

REFERENCES

[1] B. Zhang, Z. Qin, and G. Y. Li, "Semantic communications with variable-length coding for extended reality," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 5, pp. 1038–1051, Aug. 2023.

[2] Z. Utkovski, A. Munari, G. Caire, J. Dommel, P.-H. Lin, M. Franke, A. C. Drummond, and S. Stańczak, "Semantic communication for edge intelligence: Theoretical foundations and implications on protocols," *IEEE Internet of Things Mag.*, vol. 6, no. 4, pp. 48–53, Dec. 2023.

[3] J. Xu, T.-Y. Tung, B. Ai, W. Chen, Y. Sun, and D. Gündüz, "Deep joint source-channel coding for semantic communications," *IEEE Commun. Mag.*, vol. 61, no. 11, pp. 42–48, Nov. 2023.

[4] L. Xia, Y. Sun, C. Liang, D. Feng, R. Cheng, Y. Yang, and M. A. Imran, "WiserVR: Semantic communication enabled wireless virtual reality delivery," *IEEE Wireless Commun.*, vol. 30, no. 2, pp. 32–39, Apr. 2023.

[5] Z. Q. Liew, H. Du, W. Y. B. Lim, Z. Xiong, D. Niyato, C. Miao, and D. I. Kim, "Economics of semantic communication system: An auction approach," *IEEE Trans. Veh. Technol.*, vol. 72, no. 10, pp. 13 559–13 574, May 2023.

[6] B. Bandyopadhyay, P. Kuila, M. Bey, and M. C. Govil, "Quantum-inspired differential evolution for freshness-aware caching-aided offloading in digital twin-enabled internet of vehicles," pp. 1–15, May 2024.

[7] Q. Wu, F. Liu, H. Xia, and T. Zhang, "Semantic transfer between different tasks in the semantic communication system," in *IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2022, pp. 566–571.

[8] Y. Zhao, Y. Yue, S. Hou, B. Cheng, and Y. Huang, "LaMoSC: Large language model-driven semantic communication system for visual transmission," *IEEE Trans. Cogn. Commun. Netw.*, May 2024.

[9] Y. Ren, R. Xie, F. R. Yu, T. Huang, and Y. Liu, "NFT-based intelligence networking for connected and autonomous vehicles: A quantum reinforcement learning approach," *IEEE Netw.*, vol. 36, no. 6, pp. 116–124, Nov. 2022.

[10] J. A. Ansere, S. C. Prabhashana, H. Shin, O. A. Dobre, and T. Q. Duong, "Quantum deep deterministic policy gradient for digital twin-enabled semantic IoV networks," *IEEE Trans. Network Science Eng.*, Jun. 2025.

[11] B. Narottama and S. Y. Shin, "Quantum neural networks for resource allocation in wireless communications," *IEEE Trans. Wireless Commun.*, vol. 21, no. 2, pp. 1103–1116, Feb. 2022.

[12] F. Mogyorósi, P. Revisnyei, A. Pašić, Z. Papp, I. Törös, P. Varga, and A. Pašić, "Positioning in 5G and 6G networks—a survey," *Sensors*, vol. 22, no. 13, p. 4757, Jun. 2022.

[13] Y. Fu, W. Cheng, W. Zhang, and J. Wang, "Scalable extraction based semantic communication for 6G wireless networks," *IEEE Commun. Mag.*, vol. 62, no. 7, pp. 96–102, Aug. 2023.

[14] H. Peng, Z. Zhang, Y. Liu, Z. Su, T. H. Luan, and N. Cheng, "Semantic communication in non-terrestrial networks: A future-ready paradigm," *IEEE Netw.*, Apr. 2024.

[15] I. Tomkos, D. Klonidis, E. Pikasis, and S. Theodoridis, "Toward the 6G network era: Opportunities and challenges," *IT Professional*, vol. 22, no. 1, pp. 34–38, Feb. 2020.

[16] T. Han, Q. Yang, Z. Shi, S. He, and Z. Zhang, "Semantic-preserved communication system for highly efficient speech transmission," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 1, pp. 245–259, Nov. 2022.

[17] B. Zhang, J. Liu, L. Jiang, and Q. Zhuang, "Quantum-data-driven dynamical transition in quantum learning," *npj Quantum Inf.*, vol. 11, no. 1, p. 132, Aug. 2025.

[18] S. Thanasilp, S. Wang, M. Cerezo, and Z. Holmes, "Exponential concentration in quantum kernel methods," *Nature commun.*, vol. 15, no. 1, p. 5200, Jun. 2024.

[19] Z. Zhang, Y. Xiao, Z. Ma, M. Xiao, Z. Ding, X. Lei, G. K. Karagiannidis, and P. Fan, "6G wireless networks: Vision, requirements, architecture, and key technologies," *IEEE Veh. Technol. Mag.*, vol. 14, no. 3, pp. 28–41, Jul. 2019.

[20] G. Wang, J. Yang, Y. Wei, C. Wang, K. Li, and C. Feng, "Deep-reinforcement-learning-driven patient state analysis and resource management in near-field IoE healthcare networks," *IEEE Internet of Things J.*, vol. 12, no. 13, pp. 22 647–22 657, Jul. 2025.

[21] Y. Li, X. Hu, Y. Zhuang, Z. Gao, P. Zhang, and N. El-Sheimy, "Deep reinforcement learning (DRL): Another perspective for unsupervised wireless localization," *IEEE Internet of Things J.*, vol. 7, no. 7, pp. 6279–6287, Jul. 2020.

[22] R. Chataut, M. Nankya, and R. Akl, "6G networks and the AI revolution—exploring technologies, applications, and emerging challenges," *Sensors*, vol. 24, no. 6, p. 1888, Mar. 2024.

[23] J. Dai, S. Wang, K. Tan, Z. Si, X. Qin, K. Niu, and P. Zhang, "Nonlinear transform source-channel coding for semantic communications," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 8, pp. 2300–2316, Jun. 2022.

[24] Y. Duan, Q. Du, X. Fang, Z. Xie, Z. Qin, X. Tao, C. Pan, and G. Liu, "Multimedia semantic communications: Representation, encoding and transmission," *IEEE Netw.*, vol. 37, no. 1, pp. 44–50, Apr. 2023.

[25] T. M. Getu, G. Kaddoum, and M. Bennis, "Making sense of meaning: A survey on metrics for semantic and goal-oriented communication," *IEEE Access*, vol. 11, pp. 45 456–45 492, Apr. 2023.

[26] Z. Lyu, G. Zhu, J. Xu, B. Ai, and S. Cui, "Semantic communications for image recovery and classification via deep joint source and channel coding," *IEEE Trans. Wireless Commun.*, vol. 23, no. 8, pp. 8388–8404, Jan. 2024.

[27] R. Joda, M. Elsayed, H. Abou-Zeid, R. Atawia, A. B. Sediq, G. Boudreau, M. Erol-Kantarci, and L. Hanzo, "The internet of senses: Building on semantic communications and edge intelligence," *IEEE Netw.*, vol. 37, no. 3, pp. 68–75, Dec. 2022.

[28] J. He, K. Yang, and H.-H. Chen, "6G cellular networks and connected autonomous vehicles," *IEEE Netw.*, vol. 35, no. 4, pp. 255–261, Nov. 2020.

[29] S. Mihai, M. Yaqoob, D. V. Hung, W. Davis, P. Towakel, M. Raza, M. Karamanoglu, B. Barn, D. Shetve, R. V. Prasad *et al.*, "Digital twins: A survey on enabling technologies, challenges, trends and future prospects," *IEEE Commun. Surv. Tuts.*, vol. 24, no. 4, pp. 2255–2291, Sep. 2022.

[30] R. Kumar, P. Kumar, A. Aljuhani, A. Jolfaei, A. N. Islam, and N. Mohammad, "Secure data dissemination scheme for digital twin empowered vehicular networks in Open RAN," *IEEE Trans. Veh. Technol.*, vol. 73, no. 7, pp. 9234–9247, Jul 2024.

[31] X. Xu, B. Shen, S. Ding, G. Srivastava, M. Bilal, M. R. Khosravi, V. G. Menon, M. A. Jan, and M. Wang, "Service offloading with deep Q-network for digital twinning-empowered internet of vehicles in edge computing," *IEEE Trans. Ind. Informat.*, vol. 18, no. 2, pp. 1414–1426, Feb 2022.

[32] A. Masaracchia, V. Sharma, B. Canberk, O. A. Dobre, and T. Q. Duong, "Digital twin for 6G: Taxonomy, research challenges, and the road ahead," *IEEE Open J. Commun. Soc.*, vol. 3, pp. 2137–2154, Nov 2022.

[33] B. Hazarika, K. Singh, A. Paul, and T. Q. Duong, "Hybrid machine learning approach for resource allocation of digital twin in UAV-aided Internet-of-Vehicles networks," vol. 9, no. 1, pp. 2923–2935, Jan. 2024.

[34] Z. Shao, Q. Wu, P. Fan, N. Cheng, W. Chen, J. Wang, and K. B. Letaief, "Semantic-aware spectrum sharing in internet of vehicles based on deep reinforcement learning," *IEEE Internet Things J.*, vol. 11, no. 23, pp. 38 521–38 536, Dec. 2024.

[35] Q. Pan, H. Tong, J. Lv, T. Luo, Z. Zhang, C. Yin, and J. Li, "Image segmentation semantic communication over internet of vehicles," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Mar. 2023, pp. 1–6.

[36] M. Wang, T. Zhu, T. Zhang, J. Zhang, S. Yu, and W. Zhou, "Security and privacy in 6G networks: New areas and new challenges," *Digital Commun. Netw.*, vol. 6, no. 3, pp. 281–291, Aug. 2020.

[37] W. Wu, C. Zhou, M. Li, H. Wu, H. Zhou, N. Zhang, X. S. Shen, and W. Zhuang, "AI-native network slicing for 6G networks," *IEEE Wireless Commun.*, vol. 29, no. 1, pp. 96–103, Feb. 2022.

[38] H. W. Fentaw, S. Campbell, and S. Caton, "Exploring quantum control landscape and solution space complexity through optimization algorithms and dimensionality reduction," *Sci. Reports*, vol. 15, no. 1, p. 14605, Apr. 2025.

[39] N. Schetakis, P. Bonfini, N. Alisoltani, K. Blazakis, S. I. Tsintzos, A. Askitopoulos, D. Aghamalyan, P. Fafoutellis, and E. I. Vlahogianni, "Quantum neural networks with data re-uploading for urban traffic time series forecasting," *Sci. Reports*, vol. 15, no. 1, p. 19400, Jun. 2025.

[40] D. Dong, C. Chen, H. Li, and T.-J. Tarn, "Quantum reinforcement learning," *IEEE Trans. Sys., Man, Cybernetics, Part B (Cybernetics)*, vol. 38, no. 5, pp. 1207–1220, Oct. 2008.

[41] D. Dong, C. Chen, J. Chu, and T.-J. Tarn, "Robust quantum-inspired reinforcement learning for robot navigation," *IEEE/ASME Trans. Mechatron.*, vol. 17, no. 1, pp. 86–97, Feb. 2012.

[42] S. Nuuman, D. Grace, and T. Clarke, "A quantum inspired reinforcement learning technique for beyond next generation wireless networks," in *Proc. IEEE Wireless Commun. Netw. Conf. Workshops (WCNCW)*, 2015, pp. 271–275.

[43] 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, Aug. 2020.

[44] O. Lockwood and M. Si, "Reinforcement learning with quantum variational circuit," in *Proc. AAAI Conf. Artificial Intell., Interact. Digit. Entertain.*, vol. 16, no. 1, 2020, pp. 245–251.

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

[46] A. Skolik, S. Jerbi, and V. Dunjko, "Quantum agents in the gym: a variational quantum algorithm for deep Q-learning," *Quantum*, vol. 6, p. 720, May 2022.

[47] F. Sanches, S. Weinberg, T. Ide, and K. Kamiya, "Short quantum circuits in reinforcement learning policies for the vehicle routing problem," *Phys. Rev. A*, vol. 105, no. 6, p. 062403, Jun. 2022.

[48] M. L. Olivera-Atencio, L. Lamata, M. Morillo, and J. Casado-Pascual, "Quantum reinforcement learning in the presence of thermal dissipation," *Phys. Rev. E*, vol. 108, no. 1, p. 014128, Jul. 2023.

[49] S. Park, J. Chung, C. Park, S. Jung, M. Choi, S. Cho, and J. Kim, "Joint quantum reinforcement learning and stabilized control for spatio-temporal coordination in metaverse," *IEEE Trans. Mobile Comput.*, vol. 23, no. 12, pp. 12 410–12 427, Dec. 2024.

[50] M. Franz, T. Winker, S. Groppe, and W. Mauerer, "Hype or heuristic? quantum reinforcement learning for join order optimisation," in *Proc. IEEE Int. Conf. Quantum Comput., Eng. (QCE)*, vol. 01, 2024, pp. 409–420.

[51] H. Hohenfeld, D. Heimann, F. Wiebe, and F. Kirchner, "Quantum deep reinforcement learning for robot navigation tasks," *IEEE Access*, vol. 12, pp. 87 217–87 236, Jun. 2024.

[52] A. Sinha, A. Macaluso, and M. Klusch, "Nav-Q: quantum deep reinforcement learning for collision-free navigation of self-driving cars," *Quantum Machine Intell.*, vol. 7, no. 1, pp. 1–20, Feb. 2025.

[53] T. Getu, G. Kaddoum, and M. Bennis, "Semantic communication: A survey on research landscape, challenges, and future directions," *Authorea Preprints*, Nov. 2023.

[54] T. M. Getu, G. Kaddoum, and M. Bennis, "Making sense of meaning: A survey on metrics for semantic and goal-oriented communication," *IEEE Access*, vol. 11, pp. 45 456–45 492, Apr. 2023.

[55] Q. Lan, D. Wen, Z. Zhang, Q. Zeng, X. Chen, P. Popovski, and K. Huang, "What is semantic communication? a view on conveying meaning in the era of machine intelligence," *J. Commun., Info. Netw.*, vol. 6, no. 4, pp. 336–371, Dec. 2021.

[56] Z. Lu, R. Li, K. Lu, X. Chen, E. Hossain, Z. Zhao, and H. Zhang, "Semantics-empowered communications: A tutorial-cum-survey," *IEEE Commun. Surv., Tutor.*, vol. 26, no. 1, pp. 41–79, Firstquarter 2024.

[57] C. Chaccour, W. Saad, M. Debbah, Z. Han, and H. V. Poor, "Less data, more knowledge: Building next generation semantic communication networks," *IEEE Commun. Surv., Tutor.*, pp. 1–41, Jun. 2024.

[58] C. Liang, H. Du, Y. Sun, D. Niyato, J. Kang, D. Zhao, and M. A. Imran, "Generative AI-driven semantic communication networks: Architecture, technologies and applications," *IEEE Trans. Cogn. Commun. Netw.*, Jul. 2024.

[59] D. Wheeler and B. Natarajan, "Engineering semantic communication: A survey," *IEEE Access*, vol. 11, pp. 13 965–13 995, Feb. 2023.

[60] H. Du, R. Zhang, Y. Liu, J. Wang, Y. Lin, Z. Li, D. Niyato, J. Kang, Z. Xiong, S. Cui, B. Ai, H. Zhou, and D. I. Kim, "Enhancing deep reinforcement learning: A tutorial on generative diffusion models in network optimization," *IEEE Commun. Surv., Tutor.*, vol. 26, no. 4, pp. 2611–2646, Fourthquarter 2024.

[61] X. Li, J. Shi, and Z. Chen, "Task-driven semantic coding via reinforcement learning," *IEEE Trans. Image Process.*, vol. 30, pp. 6307–6320, Jul. 2021.

[62] W. Yang, Z. Xiong, H. Du, Y. Yuan, and T. Q. Quek, "Semantic change driven generative semantic communication framework," in *Proc. IEEE Wireless Commun., Netw. Conf.*, 2024, pp. 01–06.

[63] C. Liu, Y. Zhou, Y. Chen, and S.-H. Yang, "Knowledge distillation-based semantic communications for multiple users," *IEEE Trans. Wireless Commun.*, vol. 23, no. 7, pp. 7000–7012, Jul. 2024.

[64] D. Won, G. Woraphonbenjakul, A. B. Wondmagegn, A. T. Tran, D. Lee, D. S. Lakew, and S. Cho, "Resource management, security, and privacy issues in semantic communications: A survey," *IEEE Commun. Surv., Tutor.*, pp. 1–41, Oct. 2024.

[65] Y. Lin, H. Du, D. Niyato, J. Nie, J. Zhang, Y. Cheng, and Z. Yang, "Blockchain-aided secure semantic communication for AI-generated content in metaverse," *IEEE Open J. Comput. Soc.*, vol. 4, pp. 72–83, Mar. 2023.

[66] H. Du, J. Wang, D. Niyato, J. Kang, Z. Xiong, M. Guizani, and D. I. Kim, "Rethinking wireless communication security in semantic internet of things," *IEEE Wireless Commun.*, vol. 30, no. 3, pp. 36–43, Jun. 2023.

[67] L. U. Khan, Z. Han, W. Saad, E. Hossain, M. Guizani, and C. S. Hong, "Digital twin of wireless systems: Overview, taxonomy, challenges, and opportunities," *IEEE Commun. Surv., Tutor.*, vol. 24, no. 4, pp. 2230–2254, Fourthquarter 2022.

[68] A. Masaracchia, V. Sharma, B. Canberk, O. A. Dobre, and T. Q. Duong, "Digital twin for 6G: Taxonomy, research challenges, and the road ahead," *IEEE Open J. Commun. Soc.*, vol. 3, pp. 2137–2150, Nov. 2022.

[69] N. P. Kuruvatti, M. A. Habibi, S. Partani, B. Han, A. Fellan, and H. D. Schotten, "Empowering 6G communication systems with digital twin technology: A comprehensive survey," *IEEE Access*, vol. 10, pp. 112 158–112 186, Oct. 2022.

[70] A. Jolly, B. Akgun, B. Sachdev, D. Ravichandran, M. Jayabalan, R. Amiri, V. Jain, V. Chande, Y. Chen, R. Acharya, C. Tiwari, C. Woodahl, S. Yerramalli, A. Kannan, X. Zhang, A. Kumar, H. Hong, J. Boyd, R. Prakash, S. B. Mummana, S. Govindappa, and J. Wilson, "Digital twin powered next generation wireless networks: Construction, validation, and applications," in *Proc. IEEE Vehi. Technol. Conf.*, 2024, pp. 1–7.

[71] L. Bariah, H. Sari, and M. Debbah, "Digital twin-empowered communications: A new frontier of wireless networks," *IEEE Commun. Mag.*, vol. 61, no. 12, pp. 24–36, Dec. 2023.

[72] B. Qin, H. Pan, Y. Dai, X. Si, X. Huang, C. Yuen, and Y. Zhang, "Machine and deep learning for digital twin networks: A survey," *IEEE Internet Things J.*, vol. 11, no. 21, pp. 34 694–34 716, Nov. 2024.

[73] Z. Tao, W. Xu, Y. Huang, X. Wang, and X. You, "Wireless network digital twin for 6G: Generative AI as a key enabler," *IEEE Wireless Commun.*, vol. 31, no. 4, pp. 24–31, Aug. 2024.

[74] L. Schena, P. A. Marques, R. Poletti, S. Ahizi, J. Van den Berghe, and M. A. Mendez, "Reinforcement twinning: From digital twins to model-based reinforcement learning," *J. Comput. Sci.*, vol. 82, p. 102421, Oct. 2024.

[75] X. Lu, L. Xiao, P. Li, X. Ji, C. Xu, S. Yu, and W. Zhuang, "Reinforcement learning-based physical cross-layer security and privacy in 6G," *IEEE Commun. Surv., Tutor.*, vol. 25, no. 1, pp. 425–466, Firstquarter 2023.

[76] E. A. Cherrat, I. Kerenidis, and A. Prakash, "Quantum reinforcement learning via policy iteration," *Quantum Machine Intell.*, vol. 5, no. 2, p. 30, Jul. 2023.

[77] 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.

[78] H. Zhang, H. Zhao, R. Liu, A. Kaushik, X. Gao, and S. Xu, "Collaborative task offloading optimization for satellite mobile edge computing using multi-agent deep reinforcement learning," *IEEE Trans. Veh. Technol.*, Oct. 2024.

[79] H. Park, H. Chung, A. Conti, M. Z. Win, and S. Kim, "Robust near-field beam tracking via deep Q-network for THz communications," in *Proc. Int. Conf. Info. Fusion*, 2024, pp. 1–5.

[80] A. Feriani and E. Hossain, "Single and multi-agent deep reinforcement learning for AI-enabled wireless networks: A tutorial," *IEEE Commun. Surv., Tutor.*, vol. 23, no. 2, pp. 1226–1252, Secondquarter 2021.

[81] T. H. Su, S. Shresthamali, and M. Kondo, "Quantum framework for reinforcement learning: integrating Markov Decision Process, quantum arithmetic, and trajectory search," *arXiv preprint arXiv:2412.18208*, Dec. 2024.

[82] S. Markidis, "What is quantum parallelism, anyhow?" in *Proc. ISC High Perform.*, 2024, pp. 1–12.

[83] E. Van Den Berg, Z. K. Minev, A. Kandala, and K. Temme, "Probabilistic error cancellation with sparse pauli–Lindblad models on noisy quantum processors," *Nat. phy.*, vol. 19, no. 8, pp. 1116–1121, May 2023.

[84] A. Katabarwa, K. Gratsea, A. Caesura, and P. D. Johnson, "Early fault-tolerant quantum computing," *PRX quantum*, vol. 5, no. 2, p. 020101, Jun. 2024.

[85] A. M. Childs and W. Van Dam, "Quantum algorithms for algebraic problems," *Revi. Modern Phy.*, vol. 82, no. 1, pp. 1–52, Jan. 2010.

[86] E. Peters, J. Caldeira, A. Ho, S. Leichenauer, M. Mohseni, H. Neven, P. Spentzouris, D. Strain, and G. N. Perdue, "Machine learning of high dimensional data on a noisy quantum processor," *npj Quantum Info.*, vol. 7, no. 1, p. 161, Nov. 2021.

[87] J. Park, J. Cha, S. Y.-C. Chen, S. Yoo, and H.-H. Tseng, "Over the quantum rainbow: Explaining hybrid quantum reinforcement learning," in *Proc. IEEE Inter. Conf. Quantum Comput., Engineer.*, vol. 01, 2024, pp. 1583–1594.

[88] F. Chen, L. Jiang, H. Müller, P. Richerme, C. Chu, Z. Fu, and M. Yang, "NISQ computing: A security-centric tutorial and survey," *IEEE Circuits Syst. Mag.*, vol. 24, no. 1, pp. 14–32, Firstquarter 2024.

[89] Y. Qiu, R. Liu, and R. S. Lee, "The design and implementation of a deep reinforcement learning and quantum finance theory-inspired portfolio investment management system," *Expert Sys. with App.*, vol. 238, p. 122243, Mar. 2024.

[90] G. Kumar, S. Yadav, A. Mukherjee, V. Hassija, and M. Guizani, "Recent advances in quantum computing for drug discovery and development," *IEEE Access*, vol. 12, pp. 64 491–64 509, May 2024.

[91] D. Niraula, J. Jamaluddin, M. M. Matuszak, R. K. T. Haken, and I. E. Naqa, "Quantum deep reinforcement learning for clinical decision support in oncology: Application to adaptive radiotherapy," *Sci. Reports*, vol. 11, no. 1, p. 23545, Dec. 2021.

[92] X. Wei, X. Gao, K. Ye, C.-Z. Xu, and Y. Wang, "A quantum reinforcement learning approach for joint resource allocation and task offloading in mobile edge computing," *IEEE Trans. Mobile Comput.*, vol. 24, no. 4, pp. 2580–2593, Apr. 2025.

[93] Y. Dai, K. Zhang, S. Maharjan, and Y. Zhang, "Deep reinforcement learning for stochastic computation offloading in digital twin networks," *IEEE Trans. Industrial Info.*, vol. 17, no. 7, pp. 4968–4977, Jul. 2021.

[94] H. El Alami and D. B. Rawat, "Reinforcement learning-enabled satellite constellation reconfiguration and retasking for mission-critical applications," in *Proc. IEEE Military Commun. Conf.*, 2024, pp. 938–943.

[95] Y. Wang, M. Chen, T. Luo, W. Saad, D. Niyato, H. V. Poor, and S. Cui, "Performance optimization for semantic communications: An attention-based reinforcement learning approach," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 9, pp. 2598–2613, Sep. 2022.

[96] Z. Shao, Q. Wu, P. Fan, N. Cheng, W. Chen, J. Wang, and K. Ben Letaief, "Semantic-aware spectrum sharing in internet of vehicles based on deep reinforcement learning," *IEEE Internet Things J.*, vol. 11, no. 23, pp. 38 521–38 536, Dec. 2024.

[97] Y. Dai, K. Zhang, S. Maharjan, and Y. Zhang, "Deep reinforcement learning for stochastic computation offloading in digital twin networks," *IEEE Trans. Industrial Info.*, vol. 17, no. 7, pp. 4968–4977, Jul. 2020.

[98] X. Zhou, X. Zheng, X. Cui, J. Shi, W. Liang, Z. Yan, L. T. Yang, S. Shimizu, I. Kevin, and K. Wang, "Digital twin enhanced federated reinforcement learning with lightweight knowledge distillation in mobile networks," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 10, pp. 3191–3211, Oct. 2023.

[99] W. Sun, H. Zhang, R. Wang, and Y. Zhang, "Reducing offloading latency for digital twin edge networks in 6G," *IEEE Trans. Veh. Technol.*, vol. 69, no. 10, pp. 12 240–12 251, Oct. 2020.

[100] X. Tang, X. Li, R. Yu, Y. Wu, J. Ye, F. Tang, and Q. Chen, "Digital-twin-assisted task assignment in multi-UAV systems: A deep reinforcement learning approach," *IEEE Internet Things J.*, vol. 10, no. 17, pp. 15 362–15 375, Sep. 2023.

[101] P. Huembeli and A. Dauphin, "Characterizing the loss landscape of variational quantum circuits," *Quantum Sci., Technol.*, vol. 6, no. 2, p. 025011, Feb. 2021.

[102] Y. Gujju, A. Matsuo, and R. Raymond, "Quantum machine learning on near-term quantum devices: Current state of supervised and unsupervised techniques for real-world applications," *Phy. Rev. Applied*, vol. 21, no. 6, p. 067001, Jun. 2024.

[103] M. Schuld, R. Sweke, and J. J. Meyer, "Effect of data encoding on the expressive power of variational quantum-machine-learning models," *Phy. Rev. A*, vol. 103, no. 3, p. 032430, Mar. 2021.

[104] E. Andrés, M. P. Cuéllar, and G. Navarro, "Efficient dimensionality reduction strategies for quantum reinforcement learning," *IEEE Access*, vol. 11, pp. 104 534–104 553, Sep. 2023.

[105] X. Liu, G. Liu, H.-K. Zhang, J. Huang, and X. Wang, "Mitigating barren plateaus of variational quantum eigensolvers," *IEEE Trans. Quantum Eng.*, vol. 5, pp. 1–19, Mar. 2024.

[106] V. Heyraud, Z. Li, K. Donatella, A. Le Boité, and C. Ciuti, "Efficient estimation of trainability for variational quantum circuits," *PRX Quantum*, vol. 4, no. 4, p. 040335, Dec. 2023.

[107] S. Shin, Y. S. Teo, and H. Jeong, "Dequantizing quantum machine learning models using tensor networks," *Phy. Rev. Res.*, vol. 6, no. 2, p. 023218, May 2024.

[108] E. Gil-Fuster, C. Gyurik, A. Pérez-Salinas, and V. Dunjko, "On the relation between trainability and dequantization of variational quantum learning models," *arXiv preprint arXiv:2406.07072*, Jun. 2024.

[109] M. Schuld, V. Bergholm, C. Gogolin, J. Izaac, and N. Killoran, "Evaluating analytic gradients on quantum hardware," *Phy. Rev. A*, vol. 99, no. 3, p. 032331, Mar. 2019.

[110] K. Mitarai, M. Negoro, M. Kitagawa, and K. Fujii, "Quantum circuit learning," *Phy. Rev. A*, vol. 98, no. 3, p. 032309, Sep. 2018.

[111] I. Grondman, L. Busoniu, G. A. Lopes, and R. Babuska, "A survey of actor-critic reinforcement learning: Standard and natural policy gradients," *IEEE Trans. Sys., Man, Cybernetics*, vol. 42, no. 6, pp. 1291–1307, Nov. 2012.

[112] M. Ostaszewski, E. Grant, and M. Benedetti, "Structure optimization for parameterized quantum circuits," *Quantum*, vol. 5, p. 391, Jan. 2021.

[113] H. C. Watanabe, R. Raymond, Y.-Y. Ohnishi, E. Kaminishi, and M. Sugawara, "Optimizing parameterized quantum circuits with free-axis single-qubit gates," *IEEE Trans. Quantum Eng.*, vol. 4, pp. 1–16, Jun. 2023.

[114] M. Ostaszewski, L. M. Trenkwalder, W. Masarczyk, E. Scerri, and V. Dunjko, "Reinforcement learning for optimization of variational quantum circuit architectures," *Proc. Adv. Neural Inf. Process. Syst.*, vol. 34, pp. 18182–18194, 2021.

[115] J. Tilly, G. Jones, H. Chen, L. Wossnig, and E. Grant, "Computation of molecular excited states on IBM quantum computers using a discriminative variational quantum eigensolver," *Phys. Rev. A*, vol. 102, no. 6, p. 062425, Dec. 2020.

[116] M. E. Morales, T. Tlyachev, and J. Biamonte, "Variational learning of Grover's quantum search algorithm," *Phy. Rev. A*, vol. 98, no. 6, p. 062333, Dec. 2018.

[117] P. Botsinis, D. Alanis, Z. Babar, H. V. Nguyen, D. Chandra, S. X. Ng, and L. Hanzo, "Quantum search algorithms for wireless communications," *IEEE Commun. Surv., Tutor.*, vol. 21, no. 2, pp. 1209–1242, Secondquarter 2019.

[118] N. Nguyen and K.-C. Chen, "Quantum embedding search for quantum machine learning," *IEEE Access*, vol. 10, pp. 41444–41456, Apr. 2022.

[119] Z. Ye, X. Qian, and W. Pan, "Quantum topology optimization via quantum annealing," *IEEE Trans. Quantum Eng.*, vol. 4, pp. 1–15, Apr. 2023.

[120] N. Nguyen and K.-C. Chen, "Bayesian quantum neural networks," *IEEE Access*, vol. 10, pp. 54110–54122, Apr. 2022.

[121] 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. 141007–141024, Jul. 2020.

[122] B. Narottama and T. Q. Duong, "Quantum neural networks for optimal resource allocation in cell-free MIMO systems," in *Proc. IEEE Global Commun. Conf.*, 2022, pp. 2444–2449.

[123] A. Swaminathan and M. Akgün, "Distributed and secure kernel-based quantum machine learning," *arXiv preprint arXiv:2408.10265*, Aug. 2024.

[124] S. Mittal, Y. Chand, and N. K. Kundu, "Hybrid quantum neural network based indoor user localization using cloud quantum computing," in *Proc. IEEE Reg. 10 Symp.*, 2024, pp. 1–8.

[125] J. Zhang, G. Zheng, T. Koike-Akino, K.-K. Wong, and F. Burton, "Hybrid quantum-classical neural networks for downlink beamforming optimization," *IEEE Trans. Wireless Commun.*, Aug. 2024.

[126] M. Kordzanganeh, D. Kosichkina, and A. Melnikov, "Parallel hybrid networks: An interplay between quantum and classical neural networks," *Intell. Comput.*, vol. 2, p. 0028, Oct. 2023.

[127] G. Verdon, M. Broughton, J. R. McClean, K. J. Sung, R. Babbush, Z. Jiang, H. Neven, and M. Mohseni, "Learning to learn with quantum neural networks via classical neural networks," *arXiv preprint arXiv:1907.05415*, Jul. 2019.

[128] A. Abbas, D. Sutter, C. Zoufal, A. Lucchi, A. Figalli, and S. Woerner, "The power of quantum neural networks," *Nature Comput. Sci.*, vol. 1, no. 6, pp. 403–409, Jun. 2021.

[129] Silvirianti, B. Narottama, and S. Y. Shin, "UAV coverage path planning with quantum-based recurrent deep deterministic policy gradient," *IEEE Trans. Veh. Technol.*, vol. 73, no. 5, pp. 7424–7429, May. 2024.

[130] H. Hu, M. Hernandez, Y. G. Kim, K. J. Ahmed, K. Tsukamoto, and M. J. Lee, "DDPG-based wireless resource allocation for time-constrained applications," in *Proc. IEEE Wireless Commun., Netw. Conf.*, 2024, pp. 1–6.

[131] Z. Zhang, Y. Huang, C. Zhang, Q. Zheng, L. Yang, and X. You, "Digital twin-enhanced deep reinforcement learning for resource management in networks slicing," *IEEE Trans. Commun.*, vol. 72, no. 10, pp. 6209–6224, Oct. 2024.

[132] J. J. Alcaraz, F. Losilla, A. Zanella, and M. Zorzi, "Model-based reinforcement learning with kernels for resource allocation in RAN slices," *IEEE Trans. Wireless Commun.*, vol. 22, no. 1, pp. 486–501, Jan. 2023.

[133] J. I. Park, J. B. Chae, and K. W. Choi, "Model-based deep reinforcement learning framework for channel access in wireless networks," *IEEE Internet Things J.*, vol. 11, no. 6, pp. 10150–10167, Mar. 2023.

[134] J. Gilvarry and V. C. Emeakaroha, "Achieving reinforcement learning speedup via quantum kernel estimation," in *Proc. IEEE Int. Symp. Technol. Soc.*, 2023, pp. 1–8.

[135] J. Wan, S. Lin, Z. Zhang, J. Zhang, and T. Zhang, "Scheduling real-time wireless traffic: A network-aided offline reinforcement learning approach," *IEEE Internet Things J.*, vol. 10, no. 24, pp. 22331–22340, Dec. 2023.

[136] R. Luo, H. Tian, W. Ni, J. Cheng, and K.-C. Chen, "Deep reinforcement learning enables joint trajectory and communication in Internet of Robotic Things," *IEEE Trans. Wireless Commun.*, vol. 23, no. 12, pp. 18154–18168, Dec. 2024.

[137] M. Sherman, S. Shao, X. Sun, and J. Zheng, "Optimizing AoI in UAV-RIS-assisted IoT networks: Off policy versus on policy," *IEEE Internet Things J.*, vol. 10, no. 14, pp. 12401–12415, Jul. 2023.

[138] C. Huang, Y. Guo, Z. Zhu, M. Si, D. Blankenberg, and R. Liu, "Quantum exploration-based reinforcement learning for efficient robot path planning in sparse-reward environment," in *Proc. IEEE Int. Conf. Robot, Hum. Interact. Commun.*, 2024, pp. 516–521.

[139] J. Hao, T. Yang, H. Tang, C. Bai, J. Liu, Z. Meng, P. Liu, and Z. Wang, "Exploration in deep reinforcement learning: From single-agent to multiagent domain," *IEEE Trans. Neural Netw. Learning Sys.*, vol. 35, no. 7, pp. 8762–8782, Jul. 2023.

[140] R. V. Anand, R. Magesh, I. Alagiri, M. G. Brahmam, B. Balusamy, and F. Benedetto, "Quantum channel optimization: Integrating quantum-inspired machine learning with genetic adaptive strategies," *IEEE Access*, vol. 12, pp. 80397–80417, Jun. 2024.

[141] S. Y.-C. Chen, C.-M. Huang, C.-W. Hsing, H.-S. Goan, and Y.-J. Kao, "Variational quantum reinforcement learning via evolutionary optimization," *Machine Learning: Sci., Technol.*, vol. 3, no. 1, p. 015025, Feb. 2022.

[142] H. Yang, A. Alphones, Z. Xiong, D. Niyato, J. Zhao, and K. Wu, "Artificial-intelligence-enabled intelligent 6G networks," *IEEE Netw.*, vol. 34, no. 6, pp. 272–280, Oct. 2020.

[143] K. Lu, Q. Zhou, R. Li, Z. Zhao, X. Chen, J. Wu, and H. Zhang, "Rethinking modern communication from semantic coding to semantic communication," *IEEE Wireless Commun.*, vol. 30, no. 1, pp. 158–164, May 2022.

[144] M. Giordani, M. Polese, M. Mezzavilla, S. Rangan, and M. Zorzi, "Toward 6G networks: Use cases and technologies," *IEEE Commun. Mag.*, vol. 58, no. 3, pp. 55–61, Mar. 2020.

[145] X. Luo, H.-H. Chen, and Q. Guo, "Semantic communications: Overview, open issues, and future research directions," *IEEE Wireless Commun.*, vol. 29, no. 1, pp. 210–219, Jan. 2022.

[146] C. Chaccour, W. Saad, M. Debbah, Z. Han, and H. V. Poor, "Less data, more knowledge: Building next generation semantic communication networks," *IEEE Commun. Surv. Tuts.*, Jun. 2024.

[147] W. Yang, H. Du, Z. Q. Liew, W. Y. B. Lim, Z. Xiong, D. Niyato, X. Chi, X. Shen, and C. Miao, "Semantic communications for future internet: Fundamentals, applications, and challenges," *IEEE Commun. Surv. Tuts.*, vol. 25, no. 1, pp. 213–250, Nov. 2022.

[148] D. Wheeler and B. Natarajan, "Engineering semantic communication: A survey," *IEEE Access*, vol. 11, pp. 13965–13995, Feb. 2023.

[149] H. Xie, Z. Qin, G. Y. Li, and B.-H. Juang, "Deep learning enabled semantic communication systems," *IEEE Trans. Signal Process.*, vol. 69, pp. 2663–2675, Apr. 2021.

[150] B. Güler, A. Yener, and A. Swami, "The semantic communication game," *IEEE Trans. Cogn. Commun. Netw.*, vol. 4, no. 4, pp. 787–802, Feb. 2018.

[151] G. Shi, Y. Xiao, Y. Li, and X. Xie, "From semantic communication to semantic-aware networking: Model, architecture, and open problems," *IEEE Commun. Mag.*, vol. 59, no. 8, pp. 44–50, Sep. 2021.

[152] H. Xie, Z. Qin, and G. Y. Li, "Semantic communication with memory," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 8, pp. 2658–2669, Jun. 2023.

[153] Z. Weng and Z. Qin, "Semantic communication systems for speech transmission," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 8, pp. 2434–2444, Jun. 2021.

[154] Z. Yang, M. Chen, G. Li, Y. Yang, and Z. Zhang, "Secure semantic communications: Fundamentals and challenges," *IEEE Netw.*, vol. 38, no. 6, pp. 513–520, Jun. 2024.

[155] C. Dong, H. Liang, X. Xu, S. Han, B. Wang, and P. Zhang, "Semantic communication system based on semantic slice models propagation," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 1, pp. 202–213, Nov. 2022.

[156] Q. Lan, D. Wen, Z. Zhang, Q. Zeng, X. Chen, P. Popovski, and K. Huang, "What is semantic communication? a view on conveying meaning in the era of machine intelligence," *J. Commun. Informa. Netw.*, vol. 6, no. 4, pp. 336–371, Dec. 2021.

[157] A. Maedche and S. Staab, "Ontology learning for the semantic web," *IEEE Intell. Sys.*, vol. 16, no. 2, pp. 72–79, Mar. 2001.

[158] D. Huang, F. Gao, X. Tao, Q. Du, and J. Lu, "Toward semantic communications: Deep learning-based image semantic coding," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 1, pp. 55–71, Nov. 2022.

[159] Z. Qin, F. Gao, B. Lin, X. Tao, G. Liu, and C. Pan, "A generalized semantic communication system: From sources to channels," *IEEE Wireless Commun.*, vol. 30, no. 3, pp. 18–26, Jun. 2023.

[160] H. Xie, Z. Qin, X. Tao, and Z. Han, "Toward intelligent communications: Large model empowered semantic communications," *IEEE Commun. Mag.*, vol. 63, no. 1, pp. 69–75, Jul. 2024.

[161] J. Wu, C. Wu, Y. Lin, T. Yoshinaga, L. Zhong, X. Chen, and Y. Ji, "Semantic segmentation-based semantic communication system for image transmission," *Digital Commun. Netw.*, vol. 10, no. 3, pp. 519–527, Jun. 2024.

[162] F. Jiang, Y. Peng, L. Dong, K. Wang, K. Yang, C. Pan, and X. You, "Large AI model-based semantic communications," *IEEE Wireless Commun.*, vol. 31, no. 3, pp. 68–75, Jun. 2024.

[163] P. Jiang, C.-K. Wen, S. Jin, and G. Y. Li, "Wireless semantic communications for video conferencing," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 1, pp. 230–244, Nov. 2022.

[164] W. Zhang, Y. Wang, M. Chen, T. Luo, and D. Niyato, "Optimization of image transmission in cooperative semantic communication networks," *IEEE Trans. Wireless Commun.*, vol. 23, no. 2, pp. 861–873, Jun. 2023.

[165] H. Hu, X. Zhu, F. Zhou, W. Wu, R. Q. Hu, and H. Zhu, "One-to-many semantic communication systems: Design, implementation, performance evaluation," *IEEE Commun. Lett.*, vol. 26, no. 12, pp. 2959–2963, Sep. 2022.

[166] X. Jiang and A.-H. Tan, "Crctol: A semantic-based domain ontology learning system," *J. American Soc. Infom. Sci. Technol.*, vol. 61, no. 1, pp. 150–168, Jan. 2010.

[167] R. Navigli, P. Velardi, and A. Gangemi, "Ontology learning and its application to automated terminology translation," *IEEE Intell. Sys.*, vol. 18, no. 1, pp. 22–31, Jan. 2003.

[168] R. Meng, S. Gao, D. Fan, H. Gao, Y. Wang, X. Xu, B. Wang, S. Lv, Z. Zhang, M. Sun *et al.*, "A survey of secure semantic communications," *J. Netw. Comput. Appl.*, p. 104181, Apr. 2025.

[169] Z. Zhao, Z. Yang, M. Chen, Z. Zhang, and H. V. Poor, "A joint communication and computation design for probabilistic semantic communications," *Entropy*, vol. 26, no. 5, p. 394, Apr. 2024.

[170] S. Barbarossa, D. Comminiello, E. Grassucci, F. Pezone, S. Sardellitti, and P. Di Lorenzo, "Semantic communications based on adaptive generative models and information bottleneck," *IEEE Commun. Mag.*, vol. 61, no. 11, pp. 36–41, Nov. 2023.

[171] W. F. Lo, N. Mital, H. Wu, and D. Gündüz, "Collaborative semantic communication for edge inference," *IEEE Wireless Commun. Lett.*, vol. 12, no. 7, pp. 1125–1129, Mar. 2023.

[172] Y. Liu, X. Wang, Z. Ning, M. Zhou, L. Guo, and B. Jedari, "A survey on semantic communications: technologies, solutions, applications and challenges," *Digital Commun. Netw.*, Jun. 2023.

[173] S. Tang, Q. Yang, L. Fan, X. Lei, A. Nallanathan, and G. K. Karagiannidis, "Contrastive learning based semantic communications," *IEEE Trans. Commun.*, vol. 72, no. 10, pp. 6328–6343, May 2024.

[174] H. Xie, Z. Qin, and G. Y. Li, "Task-oriented multi-user semantic communications for VQA," *IEEE Wireless Commun. Lett.*, vol. 11, no. 3, pp. 553–557, Dec. 2021.

[175] B. Wang, R. Li, J. Zhu, Z. Zhao, and H. Zhang, "Knowledge enhanced semantic communication receiver," *IEEE Commun. Lett.*, vol. 27, no. 7, pp. 1794–1798, May 2023.

[176] L. Xia, Y. Sun, C. Liang, L. Zhang, M. A. Imran, and D. Niyato, "Generative AI for semantic communication: Architecture, challenges, and outlook," *IEEE Wireless Commun.*, vol. 32, no. 1, pp. 132–140, Feb. 2025.

[177] B. Tang, L. Huang, Q. Li, A. Pandharipande, and X. Ge, "Cooperative semantic communication with on-demand semantic forwarding," *IEEE Open J. Commun. Soc.*, vol. 5, pp. 349–363, Dec. 2023.

[178] Y. Wang, M. Chen, T. Luo, W. Saad, D. Niyato, H. V. Poor, and S. Cui, "Performance optimization for semantic communications: An attention-based reinforcement learning approach," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 9, pp. 2598–2613, Jul. 2022.

[179] J. Ren, Z. Zhang, J. Xu, G. Chen, Y. Sun, P. Zhang, and S. Cui, "Knowledge base enabled semantic communication: A generative perspective," *IEEE Wireless Commun.*, vol. 31, no. 4, pp. 14–22, Aug. 2024.

[180] P. Zhang, W. Xu, Y. Liu, X. Qin, K. Niu, S. Cui, G. Shi, Z. Qin, X. Xu, F. Wang *et al.*, "Intellicise wireless networks from semantic communications: A survey, research issues, and challenges," *IEEE Commun. Surv. Tuts.*, vol. 27, no. 3, pp. 2051–2084, Aug. 2024.

[181] H. Du, J. Wang, D. Niyato, J. Kang, Z. Xiong, M. Guizani, and D. I. Kim, "Rethinking wireless communication security in semantic internet of things," *IEEE Wireless Commun.*, vol. 30, no. 3, pp. 36–43, Jun. 2023.

[182] H. Xie, Z. Qin, G. Y. Li, and B.-H. Juang, "Deep learning based semantic communications: An initial investigation," in *Proc. IEEE Global Commun. Conf.*, Dec. 2020, pp. 1–6.

[183] T. M. Getu, G. Kaddoum, and M. Bennis, "A survey on goal-oriented semantic communication: Techniques, challenges, and future directions," *IEEE Access*, Mar. 2024.

[184] Q. Fu, H. Xie, Z. Qin, G. Slabaugh, and X. Tao, "Vector quantized semantic communication system," *IEEE Wireless Commun. Lett.*, vol. 12, no. 6, pp. 982–986, Mar. 2023.

[185] Z. Qin, J. Ying, D. Yang, H. Wang, and X. Tao, "Computing networks enabled semantic communications," *IEEE Netw.*, vol. 38, no. 2, pp. 122–131, Jan. 2024.

[186] A. Li, X. Liu, G. Wang, and P. Zhang, "Domain knowledge driven semantic communication for image transmission over wireless channels," *IEEE Wireless Commun. Lett.*, vol. 12, no. 1, pp. 55–59, Oct. 2022.

[187] D.-Y. Jeong, M.-S. Baek, T.-B. Lim, Y.-W. Kim, S.-H. Kim, Y.-T. Lee, W.-S. Jung, and I.-B. Lee, "Digital twin: Technology evolution stages and implementation layers with technology elements," *IEEE Access*, vol. 10, pp. 52 609–52 620, May 2022.

[188] L. You, J. Wang, W. Wang, and X. Gao, "Secure multicast transmission for massive MIMO with statistical channel state information," *IEEE Sig. Process. Lett.*, vol. 26, no. 6, pp. 803–807, Jun. 2019.

[189] Y. Zhang, R. He, B. Ai, M. Yang, R. Chen, C. Wang, Z. Zhang, and Z. Zhong, "Generative adversarial networks based digital twin channel modeling for intelligent communication networks," *China Commun.*, vol. 20, no. 8, pp. 32–43, Aug. 2023.

[190] S. Park, S. Jung, and J. Kim, "Dynamic quantum federated learning for satellite-ground integrated systems using slimmable quantum neural networks," *IEEE Access*, vol. 12, pp. 58 239–58 247, Apr. 2024.

[191] C. Wang, X. Chen, J. An, Z. Xiong, C. Xing, N. Zhao, and D. Niyato, "Covert communication assisted by UAV-IRS," *IEEE Trans. Commun.*, vol. 71, no. 1, pp. 357–369, Jan. 2023.

[192] X. Jiang, X. Chen, J. Tang, N. Zhao, X. Y. Zhang, D. Niyato, and K.-K. Wong, "Covert communication in UAV-assisted air-ground networks," *IEEE Wireless Commun.*, vol. 28, no. 4, pp. 190–197, Aug. 2021.

[193] C. Park, W. J. Yun, J. P. Kim, T. K. Rodrigues, S. Park, S. Jung, and J. Kim, "Quantum multiagent actor–critic networks for cooperative mobile access in multi-UAV systems," *IEEE Internet Things J.*, vol. 10, no. 22, pp. 20 033–20 048, Nov. 2023.

[194] H. Gharavi, J. Granjal, and E. Monteiro, "Post-quantum blockchain security for the Internet of Things: Survey and research directions," *IEEE Commun. Surv., Tutor.*, vol. 26, no. 3, pp. 1748–1774, Third quarter 2024.

[195] Y. Wang, Z. Su, S. Guo, M. Dai, T. H. Luan, and Y. Liu, "A survey on digital twins: Architecture, enabling technologies, security and privacy, and future prospects," *IEEE Int. Things J.*, vol. 10, no. 17, pp. 14 965–14 987, Sep. 2023.

[196] W. Liu, Y. Fu, Y. Guo, F. Lee Wang, W. Sun, and Y. Zhang, "Two-timescale synchronization and migration for digital twin networks:

A multi-agent deep reinforcement learning approach," *IEEE Trans. Wireless Commun.*, vol. 23, no. 11, pp. 17 294–17 309, Nov. 2024.

[197] Z. Chen, W. Yi, A. Nallanathan, and J. A. Chambers, "Distributed digital twin migration in multi-tier computing systems," *IEEE J. Sel. Topics Sig. Process.*, vol. 18, no. 1, pp. 109–123, Jan. 2024.

[198] Q. Zhang, Y. Wang, and Z. Li, "Scheduling of digital twin synchronization in industrial internet of things: A hybrid inverse reinforcement learning approach," *IEEE Internet Things J.*, vol. 12, no. 5, pp. 5137–5147, Mar. 2025.

[199] A. Paul, K. Singh, C.-P. Li, O. A. Dobre, and T. Q. Duong, "Digital twin-aided vehicular edge network: A large-scale model optimization by quantum-DRL," *IEEE Trans. Vehi. Technol.*, vol. 74, no. 2, pp. 2156–2173, Feb. 2025.

[200] K. Walter, M. Mohammady, S. Nepal, and S. S. Kanhere, "Optimally mitigating backdoor attacks in federated learning," *IEEE Trans. Dependable and Secure Comput.*, vol. 21, no. 4, pp. 2949–2963, Jul. 2024.

[201] K. A. Simpson, S. Rogers, and D. P. Pezaros, "Per-host DDoS mitigation by direct-control reinforcement learning," *IEEE Trans. Netw. Service Mgt.*, vol. 17, no. 1, pp. 103–117, Mar. 2020.

[202] Z. Lv, D. Chen, B. Cao, H. Song, and H. Lv, "Secure deep learning in defense in Deep-Learning-as-a-Service computing systems in digital twins," *IEEE Trans. Comput.*, vol. 73, no. 3, pp. 656–668, Jul. 2024.

[203] K. Liang, Y. Wang, Z. Li, G. Zheng, K.-K. Wong, and C.-B. Chae, "Digital twin assisted deep reinforcement learning for computation offloading in AAV systems," *IEEE Trans. Vehi. Technol.*, vol. 74, no. 5, pp. 8466–8471, May 2025.

[204] I. Saba, A. H. Alobaidi, S. Alghamdi, and M. Tariq, "Digital twin and TD3-enabled optimization of xEV energy management in vehicle-to-grid networks," *IEEE Access*, vol. 13, pp. 92 495–92 506, May 2025.

[205] N. Franco, A. Sakhnenko, L. Stolpmann, D. Thuerck, F. Petsch, A. Rüll, and J. M. Lorenz, "Predominant aspects on security for quantum machine learning: Literature review," in *Proc. IEEE Int. Conf. Quantum Computing, Eng.*, vol. 01, 2024, pp. 1467–1477.

[206] S. Lu, L.-M. Duan, and D.-L. Deng, "Quantum adversarial machine learning," *Phy. Rev. Res.*, vol. 2, no. 3, p. 033212, Aug. 2020.

[207] W. Gong, D. Yuan, W. Li, and D.-L. Deng, "Enhancing quantum adversarial robustness by randomized encodings," *Phy. Rev. Res.*, vol. 6, no. 2, p. 023020, Apr. 2024.

[208] M. Xue, C. Yuan, H. Wu, Y. Zhang, and W. Liu, "Machine learning security: Threats, countermeasures, and evaluations," *IEEE Access*, vol. 8, pp. 74 720–74 742, Apr. 2020.

[209] Y. Sun, N. S. T. Chong, and H. Ochiai, "Information stealing in federated learning systems based on generative adversarial networks," in *Proc. IEEE Int. Conf. Sys., Man, Cybernetics*, 2021, pp. 2749–2754.

[210] A. Yazdinejad, A. Dehghantanha, H. Karimipour, G. Srivastava, and R. M. Parizi, "A robust privacy-preserving federated learning model against model poisoning attacks," *IEEE Trans., Info. Forensics, Security*, vol. 19, pp. 6693–6708, Jun. 2024.

[211] C. Qiao, M. Li, Y. Liu, and Z. Tian, "Transitioning from federated learning to quantum federated learning in internet of things: A comprehensive survey," *IEEE Commun. Surv., Tutor.*, vol. 27, no. 1, pp. 509–545, Feb. 2025.

[212] S. Lloyd, M. Schuld, A. Ijaz, J. Izaac, and N. Killoran, "Quantum embeddings for machine learning," *arXiv:2001.03622*, Oct. 2020.

[213] Y. Li, A. H. Aghvami, and D. Dong, "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, Sep. 2021.

[214] Z. A. E. Houda, A. S. Hafid, L. Khoukhi, and B. Brik, "When collaborative federated learning meets blockchain to preserve privacy in healthcare," *IEEE Trans. Network Sci. Engineer.*, vol. 10, no. 5, pp. 2455–2465, Sep. 2023.

[215] P. Kumar, R. Kumar, A. Kumar, A. A. Franklin, S. Garg, and S. Singh, "Blockchain and deep learning for secure communication in digital twin empowered industrial IoT network," *IEEE Trans. Network Sci. Engineer.*, vol. 10, no. 5, pp. 2802–2813, Sep. 2023.

[216] J. Kang, J. He, H. Du, Z. Xiong, Z. Yang, X. Huang, and S. Xie, "Adversarial attacks and defenses for semantic communication in vehicular metaverses," *IEEE Wireless Commun.*, vol. 30, no. 4, pp. 48–55, Aug. 2023.

[217] N. Li, Y. Deng, and D. Niyato, "Goal-oriented semantic communication for wireless video transmission via generative AI," *arXiv:2502.20927*, Feb. 2025.

[218] B. Narottama and S. Y. Shin, "Federated quantum neural network with quantum teleportation for resource optimization in future wireless communication," *IEEE Trans. Veh. Technol.*, vol. 72, no. 11, pp. 14 717–14 733, Nov. 2023.

[219] C. Biswas, M. M. Haque, and U. Das Gupta, "A modified key sifting scheme with artificial neural network based key reconciliation analysis in quantum cryptography," *IEEE Access*, vol. 10, pp. 72 743–72 757, Jul. 2022.

[220] Y. Cao, Y. Zhao, Q. Wang, J. Zhang, S. X. Ng, and L. Hanzo, "The evolution of quantum key distribution networks: On the road to the Qinternet," *IEEE Commun. Surv., Tutor.*, vol. 24, no. 2, pp. 839–894, Secondquarter 2022.

[221] F. Fan, Y. Shi, T. Guggemos, and X. X. Zhu, "Hybrid quantum-classical convolutional neural network model for image classification," *IEEE Trans. Neural Netw., Learning Sys.*, vol. 35, no. 12, pp. 18 145–18 159, Dec. 2024.

[222] M. A. Jahin, M. A. Masud, M. W. Suva, M. F. Mridha, and N. Dey, "Lorentz-equivariant quantum graph neural network for high-energy physics," *IEEE Trans. Artificial Intelligence*, Mar. 2025, early Access.

[223] M. Chehimi, C. Chaccour, C. K. Thomas, and W. Saad, "Quantum semantic communications for resource-efficient quantum networking," *IEEE Commun. Lett.*, vol. 28, no. 4, pp. 803–807, Apr. 2024.

[224] A. Filali, Z. Mlika, S. Cherkaoui, and A. Kobbane, "Dynamic SDN-based radio access network slicing with deep reinforcement learning for URLLC and eMBB services," *IEEE Trans. Netw. Sci., Engineer.*, vol. 9, no. 4, pp. 2174–2187, Jul. 2022.

[225] Y. Huang, W. Xu, Y. Dai, S. Maharjan, and Y. Zhang, "Graph deep reinforcement learning for multi-cycle queuing and scheduling in deterministic networking," *IEEE Trans. Netw. Sci., Engineer.*, vol. 12, no. 2, pp. 1297–1310, Mar. 2025.

[226] J. Zheng, Q. Gao, M. Ogorzałek, J. Lü, and Y. Deng, "A quantum spatial graph convolutional neural network model on quantum circuits," *IEEE Trans. Neural Netw. Learning Sys.*, vol. 36, no. 3, pp. 5706–5720, Mar. 2025.

[227] Y. Wu, G. Lin, and J. Ge, "Knowledge-powered explainable artificial intelligence for network automation toward 6G," *IEEE Netw.*, vol. 36, no. 3, pp. 16–23, May 2022.

[228] Z. Qu, W. Shi, B. Liu, D. Gupta, and P. Tiwari, "IoMT-based smart healthcare detection system driven by quantum blockchain and quantum neural network," *IEEE J. Biomedical, Health Info.*, vol. 28, no. 6, pp. 3317–3328, Jun. 2024.

[229] K. Sutradhar, R. Venkatesh, and P. Venkatesh, "A review on smart healthcare employing quantum internet of things," *IEEE Eng. Mgt. Rev.*, pp. 1–9, Aug. 2024.

[230] U. Khalid, M. S. Ulum, A. Farooq, T. Q. Duong, O. A. Dobre, and H. Shin, "Quantum semantic communications for metaverse: Principles and challenges," *IEEE Wireless Commun.*, vol. 30, no. 4, pp. 26–36, Aug. 2023.

[231] E. A. Tuli, J.-M. Lee, and D.-S. Kim, "Integration of quantum technologies into metaverse: Applications, potentials, and challenges," *IEEE Access*, vol. 12, pp. 29 995–30 019, Feb. 2024.