SPARQ: Efficient Entanglement Distribution and Routing in Space-Air-Ground Quantum Networks

MOHAMED SHABAN[1,2], MUHAMMAD ISMAIL[1], (Senior Member, IEEE), AND WALID SAAD[3], (Fellow, IEEE)

[1]Department of Computer Science, Tennessee Tech University, Cookeville, TN 38505 USA (e-mail: mmibrahims42@tntech.edu and mismail@tntech.edu)
[2]Department of Mathematics, Faculty of Education, Alexandria University, Egypt
[3]Bradley Department of Electrical and Computer Engineering, Virginia Tech, Blacksburg, VA 24061 USA (e-mail: walids@vt.edu)

Corresponding author: Mohamed Shaban (email: mmibrahims42@tntech.edu).

ABSTRACT In this paper, a space-air-ground quantum (SPARQ) network is developed as a means for providing a seamless on-demand entanglement distribution. The node mobility in SPARQ poses significant challenges to entanglement routing. Existing quantum routing algorithms focus on stationary ground nodes and utilize link distance as an optimality metric, which is unrealistic for dynamic systems like SPARQ. Moreover, in contrast to the prior art that assumes homogeneous nodes, SPARQ encompasses heterogeneous nodes with different functionalities further complicates the entanglement distribution. To solve the entanglement routing problem, a deep reinforcement learning (RL) framework is proposed and trained using deep Q-network (DQN) on multiple graphs of SPARQ to account for the network dynamics. Subsequently, an entanglement distribution policy, third-party entanglement distribution (TPED), is proposed to establish entanglement between communication parties. A realistic quantum network simulator is designed for performance evaluation. Simulation results show that the TPED policy improves entanglement fidelity by 3% and reduces memory consumption by 50% compared with benchmark. The results also show that the proposed DQN algorithm improves the number of resolved teleportation requests by 39% compared with shortest path baseline and the entanglement fidelity by 2% compared with an RL algorithm that is based on long short-term memory (LSTM). It also improved entanglement fidelity by 6% and 9% compared with two state-of-the-art benchmarks. Moreover, the entanglement fidelity is improved by 15% compared with DQN trained on a snapshot of SPARQ. Additionally, SPARQ enhances the average entanglement fidelity by 23.5% compared with existing networks spanning only space and ground layers.

INDEX TERMS SPARQ, quantum routing, entanglement distribution, entanglement fidelity, entanglement swapping.

I. INTRODUCTION

QUANTUM information science can potentially revolutionize various disciplines by addressing major challenges that are beyond the capabilities of classical computers. For instance, quantum computing will require a major re-evaluation of security measures because current security algorithms can be easily compromised by quantum computers. Although large quantum computers could solve certain computational problems substantially faster than classical computers [1], current quantum computers are still limited in resources to outperform their classical counterparts in practical applications. The development of large-scale quantum computers equipped with a sufficient number of qubits remains an open technological challenge that may be difficult to achieve in the foreseeable future.

However, the potential of quantum computing can still be achieved through distributed quantum computing paradigms in which numerous quantum computers collaborate to tackle a single problem. In distributed quantum computing, the division of qubits across multiple quantum computers can present challenges, particularly when multi-qubit gates, such as the CNOT gate, are involved. Since the control and target qubits may reside on different computers, they must be physically co-located to perform the operation. To overcome

this limitation, one of the techniques used is to teleport either the control or target qubit to the other system, where the gate operation can be performed, and then return it back to its original system. However, executing a large circuit in distributed quantum computing often requires frequent teleportation of qubits between systems. This process of teleporting qubits back and forth can lead to a reduction in computation speed, throughput, and increased latency if it is not carried out over an efficient network architecture. In this paradigm, quantum entanglement plays a pivotal role [2], serving as the fundamental resource for various quantum communication protocols, such as teleportation and superdense coding. In order to establish entanglement between two communication parties, entangled pairs are generated and transmitted through quantum channels. These channels suffer from losses that degrade the quality of the transmitted entangled photons. To overcome this limitation, several quantum repeaters can be used to perform entanglement swapping and directly entangle the source and destination. Subsequently, to enable distributed quantum computing, a resilient quantum network that enables seamless and on-demand entanglement distribution is required. Consequently, the following key open questions need to be answered: (a) What is the practical architecture for a near-term quantum network in terms of the type of adopted links and node functionality (sources of entanglement and repeaters)? (b) How to successfully establish an end-to-end path to create entanglement between two nodes, hence, supporting quantum teleportation and superdense coding? The establishment of such an end-to-end path can be performed with quantum routing.

A. RELATED WORKS

Network Architecture: Existing research on quantum-enabled space-air-ground networks is limited to quantum key distribution (QKD) services and does not develop a practical network architecture. The work in [3] and [4] presented quantum-enabled space-air-ground integrated networks that replace classical nodes with quantum nodes without covering the functionalities of such nodes and the specifications of the communication links. Furthermore, current space-ground projects such as China's Micius [5] and EuroQCI [6] are not sufficient to achieve a fully functional quantum Internet. For instance, Micius focuses only on QKD services. Although EuroQCI aims to provide QKD and quantum communication to member states, it heavily relies on satellites for communication. However, satellite services are degraded during bad weather conditions and may not always be available hence leading to intermittent connectivity. In addition, the considerable length of satellite links poses a potential risk of degrading link quality. Therefore, the EuroQCI architecture lacks the capability to facilitate on-demand entanglement distribution. Hence, the network architecture question is still open.

Entanglement Routing: Current research on quantum rout-

ing primarily focuses on ground-based networks. For instance, the works in [7] and [8] proposed quantum routing algorithms based on the expected end-to-end entanglement rate. Also, the authors in [9] proposed a routing algorithm for grid quantum networks that is based on link distance. The work in [10] investigated the impact of quantum routing algorithms on the key rate of QKD in grid quantum networks. In [11], entanglement was used to add quantum features to classical networks, and a shortest distance-based quantum routing algorithm was proposed. The work in [12] developed a shortest path quantum routing algorithm for handling multiple requests simultaneously. Moreover, two neural networks are trained in [13] to maximize the number of served requests within the network. The authors in [14] proposed a routing algorithm that adopts end-to-end fidelity as a routing metric. Additionally, the works in [15]–[19] used shortest path algorithms to solve the quantum routing problem. The authors in [20] developed a quantum routing algorithm for space-ground networks, but it is specifically designed for QKD services. Consequently, their algorithm may not effectively handle teleportation requests, as it prioritizes secret key rates as the optimality metric. Several factors restrict the applicability of existing quantum routing algorithms [7]–[20] in SPARQ: (a) utilizing the link distance to be the routing metric is impractical in SPARQ due to the significant variation in path lengths spanning three layers, (b) existing algorithms require intensive computations, restricting their scalability in large networks, (c) existing algorithm requires calculating the optimal path for each communication request, leading to delays that impede the on-demand entanglement distribution capabilities of SPARQ, and (d) existing algorithms are designed for static networks, limiting their applicability in SPARQ.

Entanglement Distribution: The work in [21] proposed a scheme for entanglement generation called Midpoint-source (MPS). This scheme utilizes a photon source positioned at the midpoint of an optical channel to generate entanglement between quantum dots. The receivers subsequently establish local entanglement between the communication qubit and the emitted photons with the help of Bell state measurement. The scheme is designed to be less susceptible to losses and to achieve the highest entanglement rate by repeatedly attempting to produce Bell states. However, MPS operates under the assumption of direct communication between parties, without involving intermediate repeaters. Furthermore, it primarily focuses on hardware aspects and lacks a specific policy for nodes to follow.

Several efforts have been made to design entanglement distribution policies. For instance, the work in [22] proposed a genetic algorithm to discover an optimal policy for entanglement generation and distribution. Despite its potential optimality, this policy is computationally intensive and is susceptible to delays. Alternatively, [23] employed a reinforcement learning (RL) agent to handle the task of entanglement distribution. However, this method is limited to a linear

repeater chain comprising a maximum of five nodes. Additionally, the work in [24] used dynamic programming and RL to propose an entanglement distribution policy. The works in [25] and [26] proposed policies for entanglement distribution that are based on Markov decision processes. Moreover, the work in [27] tackled the entanglement distribution challenge by formulating it as a mixed-integer nonlinear programming optimization problem, aiming to optimize the entanglement generation rate. However, it is limited to small-scale quantum networks. The work in [28] utilized exhaustive search techniques to establish entanglement between two nodes while meeting the specified quality of service requirement. This is achieved through the manipulation of various parameters, including the distance between adjacent nodes and the number of entanglement distillations conducted. However, using the exhaustive search can be computationally expensive for a large search space. Despite these research efforts, a common limitation among existing studies [22]–[27] is the assumption of uniform functionalities across network nodes which restricts their applicability in SPARQ. This limitation arises from the fact that not all nodes within SPARQ possess the capability to generate entanglement. While some existing work considers a heterogeneous network [27], it typically assumes heterogeneity in network specifications, such as varying distances, diverse applications, and various quality-of-service requirements. In contrast, a more realistic assumption of heterogeneity would be to consider the functionalities and capabilities of the nodes. Additionally, determining the optimal strategy for each communication request introduces additional delays that scale with the number of nodes, restricting their scalability in large networks like SPARQ. Thus, there is a need for an entanglement distribution policy that establishes entanglement between communication parties through a set of intermediate nodes. This policy should specify the functionality of each intermediate node as a source of entanglement and/or repeater in a way that minimizes the number of entanglement swapping operations, thereby improving the end-to-end entanglement quality and reducing memory consumption by intermediate repeaters. Also, this policy should consider network limitations such as the incapability of ground nodes to serve as sources of entanglement.

Quantum Network Simulator: Existing quantum network simulators, such as QuNetSim [29], NetSquid [30], QDNS [31], SQUANCH [32], SeQUeNCe [33], SimulaQron [34], and SimQN [35] are primarily focused on ground networks and terrestrial communications. However, to simulate large-scale space-air-ground quantum networks like SPARQ, there is a need for a quantum network simulator that considers satellites and their movements, high altitude platforms (HAPs) and their locations, and various communication channels such as fiber optical and free-space optical channels.

B. CONTRIBUTIONS

The main contribution of the paper is proposing a space-air-ground quantum (SPARQ) network, which integrates satellites, HAPs, and ground end-users to distribute entanglements and create end-to-end on-demand entangled paths. This approach introduces an air layer, shortening the links within the entanglement path and enhancing the overall quality of communication links. Hence, SPARQ has the potential to harness the global quantum Internet, enabling applications such as distributed quantum computing, quantum communications, quantum sensing, and quantum security. SPARQ has the potential to address the limitations of the prior works. To achieve this overarching goal, our key contributions include:

- We propose the SPARQ network architecture, which spans space, air, and ground layers. Furthermore, we specify node functionalities such as practical sources of entanglement, repeaters, and communication links. SPARQ facilitates quantum teleportation and superdense coding within the emerging quantum Internet, enabling the efficient realization of advanced network applications like distributed quantum computing. Simulation results demonstrate the superior performance of the proposed architecture compared to traditional satellite-terrestrial networks, exhibiting a remarkable improvement by 23.5% in the average end-to-end entanglement fidelity.

- We propose a deep RL strategy to solve the entanglement routing problem within the SPARQ network. The deep RL agent is trained following a deep Q-network (DQN) policy. Unlike the traditional DQN policies, our proposed DQN agent is trained on the dynamic network topology of SPARQ, enabling it to capture comprehensive network graph states and develop an optimal policy for real-time entanglement routing. Simulation results showed that the proposed DQN algorithm achieves a 39% improvement in the number of resolved teleportation requests compared to the shortest path baseline. Additionally, it improves the average entanglement fidelity by 15% compared with another DQN algorithm trained on a snapshot of SPARQ. Moreover, the proposed DQN model achieves up to 9% improvement in the average entanglement fidelity compared with state-of-the-art benchmark algorithms. Also, the simulation results show that the proposed DQN algorithm has minimal overhead compared to all considered baselines. The proposed DQN algorithm also significantly reduces the time required to find the routing path, demonstrating a substantial improvement in speed compared to state-of-the-art algorithms.

- We propose a new policy for on-demand entanglement distribution within SPARQ, called third-party entanglement distribution (TPED). This policy is designed to establish end-to-end entanglement between communication parties after determining the optimal routing path. The TPED policy assigns functionalities to each node

in the routing path to act as sources of entanglement and/or repeaters in a way that minimizes the number of entanglement swapping operations, while also considering network node limitations such as the incapability of ground nodes to act as a source of entanglement. The goal of TPED is to enhance the quality of established end-to-end entanglement by reducing the number of entangled pairs consumed by intermediate repeaters. Simulation results demonstrate that TPED improved the average entanglement fidelity while reducing the memory consumption by 50% compared with a benchmark. The benchmark approach involved establishing entanglement between adjacent nodes and subsequently performing entanglement swapping to achieve direct entanglement between the source and destination.

We conducted a comprehensive evaluation of the proposed SPARQ architecture and algorithms by upgrading the QuNet-Sim simulator and integrating it with Ansys Systems Tool Kit (STK). This integration accounts for realistic space-air-ground node behaviors and models. The rest of this paper is organized as follows. Section II presents the proposed SPARQ network architecture and the channel models. Section III defines the quantum routing problem within SPARQ and presents the proposed deep RL routing model for finding the optimal end-to-end entanglement path. Section IV proposes an on-demand entanglement distribution policy within SPARQ. Section V presents the SPARQ simulator which is used for performance evaluations, discusses the simulation setup, and presents the simulation results. Conclusions are drawn in Section VI.

II. SPARQ ARCHITECTURE AND CHANNEL MODEL

A. NETWORK LAYERS AND NODE FUNCTIONALITIES

As shown in the architecture of Fig. 1, the SPARQ network spans three layers as follows:

- *Space layer:* This layer consists of a satellite constellation equipped with quantum devices that can generate entangled states and manipulate qubits through measurements and processing. We consider that the satellites are positioned in a low earth orbit (LEO) at an altitude of 500 km.
- *Air layer:* This layer consists of swarms of aerial vehicles equipped with limited quantum capabilities, including a restricted number of qubits for generating entangled states and manipulating qubits through measurements and processing. The deployment of the air layer is motivated by the fact that satellite services may experience interruptions due to bad weather conditions and orbit limitations [3]. Additionally, aerial vehicles can serve as relays between ground nodes and satellites, extending communication range and enhancing link quality. Consequently, this layer is deployed to address network demands when satellite services are unavailable and to bridge the communication gap between the space and ground layers, contributing to the development of a

resilient network architecture. We use HAP drones at a height of 50 km.
- *Ground layer:* This layer is composed of end-user quantum nodes with limited quantum capabilities, allowing them to measure and process qubits.

As mentioned earlier, it is necessary to distribute entangled pairs among end users in order to enable quantum communications within SPARQ. This process involves the creation of entangled pairs and their transmission through quantum channels. However, these channels are prone to losses and transmitted photons decay rapidly as the transmission distance increases. To address this challenge, multiple quantum repeaters can be adopted, allowing for entanglement swapping and the extension of the entanglement range. Consequently, the node functionalities in SPARQ can be summarized as follows:

- *Source of entanglement:* These nodes are responsible for the generation of entangled pairs, which are the fuel for quantum networks. In SPARQ, satellites and HAPs are used as sources of entanglement.
- *Repeaters:* These nodes are responsible for performing entanglement swapping to enhance the quality of transmitted photons over long distances. In SPARQ, satellites, HAPs, and ground end-users can act as repeaters.
- *End-users:* These are the ground nodes. Entanglement routing algorithms are employed to connect two communicating parties within the ground layer. The routing path makes use of intermediate nodes to establish an end-to-end entanglement connection between any two ground nodes.

B. COMMUNICATION LINKS

Quantum communications can be realized through optical fiber channels or free space optical (FSO) channels, both technologies are adopted within SPARQ. Optical fiber offers certain advantages, such as establishing a reliable and uninterrupted connection between communicating parties. Moreover, photons transmitted over fiber links are typically unaffected by external factors such as background light or bad weather conditions. However, optical fiber technologies limit the transmission of photons to a few hundred kilometers due to issues such as polarization preservation, random scattering, and optical attenuation. Conversely, FSO technology outperforms fiber optical technology in terms of losses because of the low atmospheric absorption [36]. In addition, most of the transmission path for satellite-based FSO communication remains unaffected by either absorption or turbulence, as the atmospheric width is less than 10 km [36]. It is worth noting that FSO channels are used for real projects such as Micius [5] and are widely employed for QKD in research [3], [4], [37]–[41]. Therefore, FSO channels are adopted for communication between satellites, HAPs, and between satellites or HAPs and end-users. On the other hand, fiber optic channels are employed for communication between ground end-users.
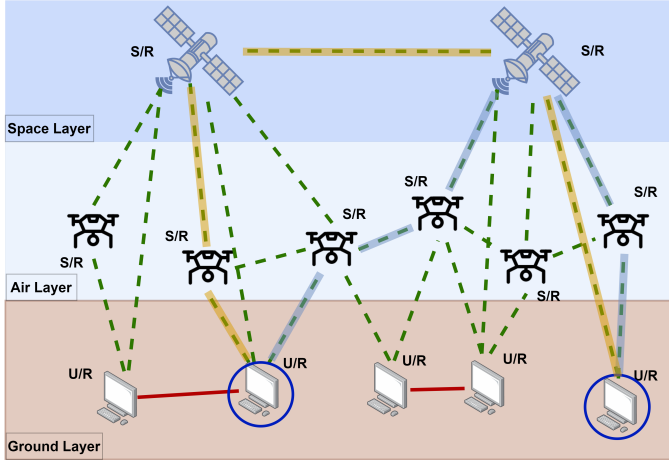
FIGURE 1: Proposed architecture of SPARQ. FSO channels are shown with green dashed lines and optical fiber channels are shown with red solid lines. S, R, and U indicate a source of entanglement, repeater, and end-user, respectively.

C. CHANNEL MODELS

We consider two channels in our SPARQ system: fiber optic and FSO channels. In each channel, we consider transmissivity to address optical losses occurring during communication. Thus, while transmitting quantum photons over a communication channel, their decay is determined by the calculated transmissivity for that specific channel.

1) Fiber optical channels

For an attenuation coefficient α and transmission distance l, the transmissivity of fiber optical channels given by [42]:

$$\eta = e^{-\alpha l}. \tag{1}$$

2) FSO channels

The transmissivity of the FSO channel considers the total optical loss that may occur in strong turbulence weather. Specifically, the transmissivity η considers three optical transmissivities as follows [43]:

$$\eta = \eta_{\mathrm{lt}}\eta_{\mathrm{eff}}\eta_{\mathrm{atm}}, \tag{2}$$

where η_{lt}, η_{eff}, and η_{atm} present the transmissivity based on turbulence, receiver efficiency, and atmospheric loss, respectively. Herein, η_{atm} is modeled by the Beer-Lambert equation [43]:

$$\eta_{\mathrm{atm}} = e^{-\alpha(\lambda,h_0)z}, \alpha(\lambda,h_0) = \alpha_0(\lambda)e^{-\frac{h_0}{6600}}, \tag{3}$$

where z is the propagation path length, λ is the carrier wavelength, h_0 represents the altitude in meters, and $\alpha_0(\lambda)$ represents the extinction factor at sea level.

Furthermore, the transmissivity based on turbulence η_{lt} is given by [43]:

$$\eta_{\mathrm{lt}} = 1 - e^{\frac{-2a_{\mathrm{R}}^2}{w_{\mathrm{lt}}^2}}, \tag{4}$$

where a_{R} is the aperture radius and w_{lt} is the long-term beam waist, which can be derived from [43]:

$$w_{\mathrm{lt}} = w_z\sqrt{1 + \frac{4}{3}q\Lambda}, \tag{5}$$

where w_z is the diffraction beam waist, while the parameters Λ and q are given by [43]:

$$\Lambda = \frac{2z}{kw_z^2}, \tag{6}$$

$$q = 0.74\sigma^2\sqrt[6]{M}, M = 35.05\frac{z}{kl_0^2}, \tag{7}$$

and l_0 is the turbulence inner scale, z is the propagation path length, $k = 2\pi/\lambda$ is the wave number, and σ^2 is the Raytov variance, which is given by [43]:

$$\sigma^2 = 2.25C_n^2 k^{\frac{7}{6}} z^{\frac{11}{6}}, \tag{8}$$

where C_n^2 is the refractive index structure that is used to characterize the atmospheric turbulence strength.

One characteristic of natural light is diffraction, a phenomenon that continuously expands the beam waist of a wave as it travels through free space. This phenomenon also leads to a constant increase in the radius of curvature of the propagating beam. We start with a Gaussian beam characterized by an initial field spot size of w_0, a carrier wavelength λ, and an initial radius of curvature R_0. As the beam propagates to a distance z where a receiver is positioned, free-space diffraction results in an enlargement of the beam's spot size to [43]:

$$w_z^2 = w_0^2\left[\left(1 - \frac{z}{R_0}\right)^2 + \left(\frac{z}{z_{\mathrm{R}}}\right)^2\right], \tag{9}$$

where $z_{\mathrm{R}} = \pi w_0^2/\lambda$ is the beam's Rayleigh length.

3) Amplitude damping channel

The amplitude damping channel is a concept that describes a type of quantum channel through which a quantum system experiences a reduction in its amplitude or damping of its state over time. This channel is used to model certain physical processes that can cause the loss of quantum information or the decay of quantum states. In quantum communication, the amplitude damping channel can be used to model the attenuation that can occur during transmission. We adopt the amplitude damping channel to degrade the entanglement quality based on the transmissivity η. Then, the amplitude damping channel can be represented by the Kraus operator as follows [44]–[46]:

$$\boldsymbol{K_0} = \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{\eta} \end{bmatrix}, \boldsymbol{K_1} = \begin{bmatrix} 0 & \sqrt{1-\eta} \\ 0 & 0 \end{bmatrix}. \tag{10}$$

The action of the amplitude damping channel on the density matrix $\boldsymbol{\rho}$ is given by:

$$\boldsymbol{\rho'} = \boldsymbol{K_0}\boldsymbol{\rho}\boldsymbol{K_0}^\dagger + \boldsymbol{K_1}\boldsymbol{\rho}\boldsymbol{K_1}^\dagger. \tag{11}$$

4) Entanglement Fidelity

Entanglement fidelity (F_e) is a metric used to quantify the similarity between an ideal entangled state and the actual state produced in a quantum system [47]. Here, we use entanglement fidelity in order to measure the quality of the final end-to-end entanglement created between the source and the destination after going through all the channel losses experienced by the intermediate nodes. Entanglement fidelity is given by [48] and [47] as follows:

$$F_e = \left(\mathrm{Tr} \left(\sqrt{ \sqrt{\boldsymbol{\rho'}} \, |\psi\rangle \langle\psi| \, \sqrt{\boldsymbol{\rho'}} } \right) \right)^2, \qquad (12)$$

where $\boldsymbol{\rho'}$ is the density matrix of the entangled state after experiencing losses, as given by (11), and $|\psi\rangle$ represents the ideal entangled state. Here, $|\psi\rangle$ is the maximally entangled Bell state $\frac{|00\rangle + |11\rangle}{\sqrt{2}}$.

In summary, considering factors such as node distances, channel types, and environmental conditions, we calculate the transmissivity, as captured by (1) and (2). This transmissivity has a direct impact on the density matrix, as illustrated in (11), which affects the quality of entanglement as captured by the entanglement fidelity in (12). The entanglement fidelity will be used to evaluate our proposed routing strategy. Leveraging entanglement fidelity as a metric to evaluate the performance of a routing algorithm provides a quantitative measure of the entanglement quality established between the source and destination.

D. MOBILITY MODEL

The mobility of satellites results in continuous changes in transmissivities, leading to dynamic connectivity and, consequently, a dynamic network topology. The STK [49] simulator is employed to model the mobility of satellites. Specifically, the utilized satellites operate in LEO at an altitude of 500 km, similar to the Micius satellite [5]. It is worth noting that the STK simulator considers the Earth's motion in simulating satellite movement, thereby providing precise satellite locations. The mobility of satellites has a significant impact on the SPARQ network's topology. For instance, Fig. 2 provides a visual representation of various SPARQ network topologies, illustrating how the network topology changes as satellites move. Fig. 2 is created using the SPARQ simulator, to be detailed in Section V-A, which uses realistic satellite movements, hence, all of the parameters of this setup are practical. As shown in Fig. 2, the mobility of satellites has a significant impact on the SPARQ network's topology. This is because of the dynamic changes in transmissivities during satellite movement. In some cases, transmissivities may become too low, leading to node disconnections, while in other scenarios, transmissivity can become high and nodes get connected. For example, in Fig. 2a, all satellites are disconnected both from each other and all network nodes due to low transmissivities. As satellites move, connectivity undergoes continuous changes based on varying transmissivities. Subsequently, in Fig. 2b, S_1 establishes connections with S_2,

A_1, G_2, and G_3 because transmissivities between S_1 and these nodes become sufficiently high. However, in Fig. 2c, S_1 is disconnected from S_2, while maintaining connections with A_1, G_2, and G_3 due to the variable transmissivities. These connections and disconnections are determined by a transmissivity threshold of 0.7 to support quantum teleportation and superdense coding, as will be discussed in Section V-B.

III. QUANTUM ROUTING WITHIN SPARQ
A. PROBLEM DEFINITION

We represent the SPARQ network as a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \{v_i^x\}_{i=1}^V$ represents a set of V nodes where $x \in \{\mathrm{s, a, g}\}$ denoting a space, air, and ground node, respectively. Herein, $\mathcal{E} = \{e_{i,j}; \, v_i^x, v_j^x \in \mathcal{V}\}$ represents the edges connecting these nodes. Each edge $e_{i,j}$ can take the form of an FSO or fiber channel based on the types of nodes v_i^x and v_j^x. Suppose node v_i^{g} wants to transmit qubits to node v_j^{g}, and there is no direct quantum channel between these two nodes, or the direct channel yields poor transmissivity. This implies that establishing an entangled pair directly between v_i^{g} and v_j^{g} is not possible. In this scenario, an intermediate node v_b^x that shares entangled pairs with both v_i^{g} and v_j^{g}, meaning that $\{e_{i,b}, e_{b,j}\} \in \mathcal{E}$, can act as a repeater performing entanglement swapping to directly establish entanglement between v_i^{g} and v_j^{g}. Therefore, the objective is to find the optimal routing path, such that, via a sequence of intermediate repeaters, a direct entanglement between communication nodes can be established with the highest possible quality. Consider the communication scenario involving the two end users represented by the black circles in Fig. 1. In this case, there are several path options, and here we highlight two potential routes in yellow and black that indirectly connect these users. Assume that entangled pairs are already distributed between adjacent nodes. For the yellow and black route, each intermediate node can act as a repeater performing entanglement swapping to directly entangle the two end users. Thus, solving the routing problem involves choosing the optimal path between these users based on a specific metric for optimality. In SPARQ, the optimal path is defined as the sequence of intermediate nodes that connects the source and destination with the highest achievable entanglement quality. This is a binary optimization problem, that is NP-hard. The dynamic network topology of SPARQ, as discussed in Section II-D, necessitates solving the routing problem in real-time. This is infeasible because it is categorized as an NP-hard problem. Thus an efficient solution is needed.

B. DEEP RL-BASED ROUTING STRATEGY

Deep RL is employed to determine the optimal route between communication nodes in SPARQ. The deep RL agent is trained using DQN. The use of deep RL and DQN for quantum routing within the SPARQ network is motivated by the following factors. Deep RL can simplify the solution of the routing problem, which is NP-hard as discussed in Section III-A, by learning through experience. In contrast, alternative approaches like stochastic optimization are impractical
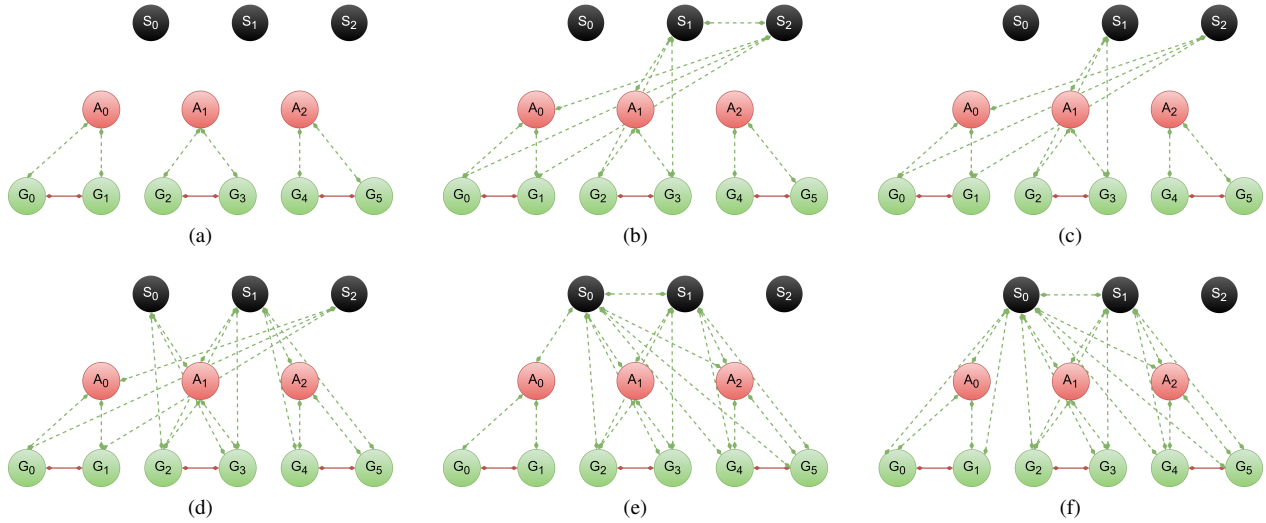
FIGURE 2: Illustration of different topologies of the same SPARQ network while satellites are moving. Ground nodes are presented in green, satellite nodes are presented in black, and HAP nodes are presented in red. FSO channels are presented in green dashed lines and optical fiber channels are presented in red solid lines. The shown node positions do not reflect the actual position, instead, this is reflected by the edges and whether they are connected or disconnected.

in SPARQ due to involving intricate computations which introduces significant delays. The requirement of computing the optimal path for each communication request further complicates the situation. Additionally, stochastic optimization approaches may converge to local minima or maxima rather than the global optimum. Furthermore, SPARQ exhibits dynamic network topology and several topologies must be captured during the training of the agent. DQN provides the adaptability required to optimize routing decisions in response to the dynamic network topology of SPARQ, ensuring efficient entanglement distribution and high-quality paths. A pivotal component of the DQN framework is the experience buffer, strategically employed to enable the agent to retain both past and recent experiences. This extensive repository of experiences equips the agent with a comprehensive understanding of the network dynamics. This comprehensive knowledge base and adaptability enable the routing decisions to be optimized in real time. The objective of the proposed DQN agent is to identify the most efficient path for entanglement distribution between nodes, considering the transmissivity between these nodes. We experimented with different hyperparameters to find the optimal settings that yield the best-performing model. We implemented the proposed DQN algorithm using a neural network with four fully connected layers. The input to the network consists of the state space $\mathcal{S}$ of the RL environment, and the output is the Q-values for each possible action. The architecture of the proposed neural network is as follows:

- Input Layer: The input layer accepts a state representation vector, which is passed through a fully connected layer. The rectified linear unit (ReLU) activation function is used to introduce non-linearity.
- Hidden Layers: Two hidden layers are implemented

with 128 neurons and 64 neurons, respectively. Also, the ReLU activation function is employed.
- Output Layer: The output layer consists of neurons equal in number to the possible actions in the environment, and its role is to produce the Q-values associated with each action.

C. ENVIRONMENT AND STATE REPRESENTATION

The environment is represented by the dynamic network graph of SPARQ. This network graph captures the current state of the network at each episode. We represent the state as a vector of transmissivities to adjacent nodes. Each node in the network is characterized by its transmissivity vector. The state vector is designed to have a fixed length, set to at least accommodate the maximum number of possible neighbors at any given node. For the cases in which a node has fewer connections, the state vector is padded with zeros to keep the fixed length. This vector serves as the state representation and is the input to the DQN agent.

D. ACTION SPACE AND REWARD FUNCTION

In the proposed model, the action space covers the collection of adjacent nodes to the current node. This signifies that at any given state s_t, the agent selects an action a_t to visit one of the neighboring nodes. However, if the agent is at a ground node, its choice is constrained to prevent the next node from also being a ground node. This restriction ensures that two consecutive ground nodes do not appear in the routing path. This is essential due to the limited capabilities of ground nodes, which are unable to generate entangled states. To determine the quality of the chosen action, we use a reward function based on the transmissivity associated with the selected action. This encourages the DQN agent to favor

actions that lead to higher transmissivity, thus optimizing the routing path for an efficient entanglement distribution process. The reward function is described as:

$$r(s,a) = \begin{cases} \min[\eta_{i,j} \forall i,j \in \mathcal{P}] & \text{if agent reach destination} \\ -1 & \text{if } \eta \leq 0.1 \\ -\frac{1}{10 \times \eta} & \text{otherwise} \end{cases}, \quad (13)$$

where $\mathcal{P}$ is the path between the source and destination nodes, which is an ordered list starting at the source and ending at the destination. As shown in (13), the agent receives a positive reward once reaching the destination. Otherwise, for each step taken, it incurs a penalty that is inversely proportional to the transmissivity of the chosen link. This means that the agent receives a smaller penalty when choosing to pass through a channel with high transmissivity, and vice versa. Additionally, the penalty is scaled and normalized within the range of 0.1 to 1. The agent incurs the minimum penalty of 0.1 when selecting a link with a transmissivity of 1, and it incurs the maximum penalty of 1 when selecting a link with a transmissivity of 0.1 or lower. The second condition of (13) ensures that the penalty does not exceed 1. The reward function in (13) is designed to prevent the agent from encountering cycles. Additionally, it encourages the agent to find an optimal path with the least number of hops. A transmissivity threshold is calculated based on acceptable entanglement fidelity value according to the application as will be shown in Section V-B.

E. TRAINING THE DQN AGENT

Q-learning is employed to train the RL agent. The agent learns from experience by selecting an action a from the action space $\mathcal{A}$, this enables the agent to acquire knowledge about the expected cumulative discounted reward $Q(s_t, a_t)$, which is known as the Q-value. In a given current state s_t, the agents select an action a_t that offers the highest Q-value from the entire set of feasible actions within its action space, such that:

$$a_t = \arg\max_{a_t \in A} Q(s_{t+1}, a_t). \quad (14)$$

We adopted the Q-learning algorithm with experience replay [50]. During training, an experience replay buffer was used to store and sample experiences (state, action, reward, next state) to break the temporal correlation in the data. The loss function used for updating the Q-network was the mean squared error loss between the predicted Q-values and the target Q-values. We applied a target network, which was a duplicate of the Q-network, to stabilize the learning process. The target network parameters were updated periodically to match those of the Q-network. The DQN was trained using mini-batch stochastic gradient descent. The Q-values are updated as follows:

$$Q(s_t, a_t) = (1-\beta)Q(s_t, a_t)$$
$$+ \beta\left(r(s_t, a_t) + \gamma \max_{a_t \in \mathcal{A}} Q(s_{t+1}, a_t)\right), \quad (15)$$

where β is the learning rate, γ is the discount factor, s_t is the current state, a_t is the action chosen by the agent, and r is the reward function. We experimented with various hyperparameters, including the discount factor (γ), exploration rate (ϵ), batch size, and replay buffer size, to find the optimal settings for our specific environment (see Appendix A).

The SPARQ network graph is dynamic as it experiences changes in the positions and connectivity of satellites over time. At each training episode, we update the environment to capture the current state of the network graph. This ensures that the DQN agent adapts to the dynamic network topology of SPARQ. Simulation results indicate the superiority of this approach compared to training the agent on a single snapshot of SPARQ, as detailed in Section V. Our DQN agent undergoes 500 training episodes, and the following steps are performed for each episode:

1) The environment (network graph) is updated to reflect the current state of the SPARQ network.
2) Edge features, represented by transmissivities, are transformed into a vector at each node, forming node features. This step allows us to represent the state space and incorporate it into our DQN model.
3) The DQN agent is trained for 100 additional episodes (mini-episodes) on the current graph.

This methodology outlines the approach used to design and train the DQN agent for routing optimization within the dynamic SPARQ network. It provides details of how the model interacts with the environment, how state representations are created, and how the model's policy is developed to determine optimal routing paths based on transmissivities. The proposed routing strategy is illustrated on the right side of Fig. 3. We employ offline training to train our agent, and once the training is complete, we deploy the agent at each node in the network. At each node, the agent relies solely on local information to determine the next node for entanglement distribution. Therefore, there is no need for centralized coordination. Specifically, the agent uses the transmissivities to adjacent nodes to determine the encountered network state and select the next node that maximizes the cumulative reward. Given that satellite trajectories are periodic, this local information effectively reflects the network state, enabling the agent to optimize routing requests based on encountered network states. Once the destination is reached, the selected routing path is shared with each node involved to follow an entanglement distribution policy. This decentralized approach eliminates the need for a central controller. Each node is equipped with its own trained agent to make real-time decisions based on local information, ensuring scalability and feasibility in a dynamic SPARQ environment. Further details on the dataset employed for training the DQN agent can be found in Appendix B.

F. CONVERGENCE

The proposed DQN algorithm's convergence is tracked by observing the loss function during the training process. The
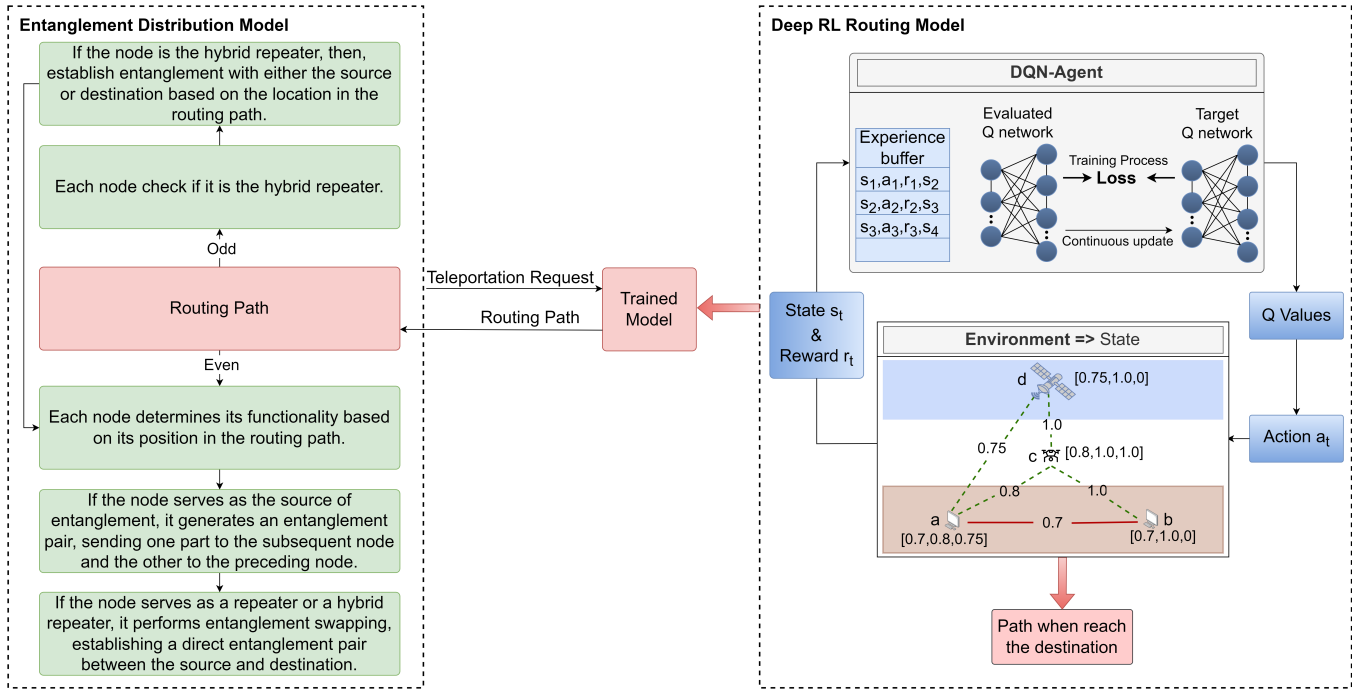
FIGURE 3: Illustration of the proposed routing model and entanglement distribution strategy. The TPED policy is employed here for entanglement distribution as it results in improved end-to-end entanglement quality, as will be detailed in Section V-D.

training is halted when the difference in loss between epochs reaches a sufficiently low value (5×10^{-4}) for 50 consecutive epochs. Observing minimal differences between successive epochs indicates stability in the learning process. When the algorithm converges, the Q-network's updates become progressively marginal, signifying that the model has learned a stable policy.

G. COMPLEXITY

The time complexity of a deep RL algorithm with a state space $\mathcal{S}$, action space $\mathcal{A}$, exploration rate ϵ, and discount factor γ can be expressed as $O(|\mathcal{S}|^2|\mathcal{A}|/(\epsilon^3(1-\gamma)^3))$ [51], where $|.|$ denotes the cardinality of a space. Here, considering a SPARQ network with V nodes, the time complexity of the proposed deeper RL algorithm is upper bounded by $O(V^3/(\epsilon^3(1-\gamma)^3))$, denoting polynomial complexity, which is considered efficient.

IV. ENTANGLEMENT DISTRIBUTION WITHIN SPARQ

Once the routing path between the source and destination is determined, the subsequent phase involves establishing entanglement between them, utilizing intermediate nodes according to a predefined policy. As mentioned earlier, existing entanglement distribution policies are not applicable in SPARQ as they assume uniform functionalities across network nodes. In SPARQ, not all network nodes possess the capability to generate entanglement pairs. The intuitive approach for entanglement distribution is to establish entanglement between adjacent nodes, considering network constraints such that satellites and HAPs can serve as sources

of entanglement, while ground nodes lack the ability to generate entangled pairs. Intermediate nodes then execute entanglement swapping operations to establish direct entanglement between the communication parties. However, this approach leads to a degradation in the end-to-end entanglement shared between the source and destination. This is because each entanglement swapping process degrades the entanglement quality due to non-maximally entangled states being shared between adjacent nodes [52]. Therefore, we propose an entanglement distribution policy, namely TPED. The TPED policy aims to minimize the number of entangled pairs consumed by the intermediate repeaters. This policy seeks to enhance the overall entanglement quality of the end-to-end entangled pair established between the source and destination. The idea of the proposed TPED policy is based on keeping the source of entanglement outside the routing path. Each source of entanglement can achieve this by generating an entangled pair and sending one part to the preceding node in the routing path, while the other part is sent to the subsequent node. This allows for the adjacent nodes of the source of entanglement to get entangled together directly. The difference between the intuitive approach and the TPED policy is shown in Fig. 4. In this figure, the entanglement distribution process can be performed in different ways. For instance, N_1 can establish entanglement with S and N_2, then, N_2 can establish entanglement with D. Now, we have a chain of entanglement ($S \longleftrightarrow N_1 \longleftrightarrow N_2 \longleftrightarrow D$) between S and D as shown in Fig. 4b. By performing entanglement swapping at nodes N_1 and N_2, S and D share a direct entangled pair that can be used for communication.

This approach requires creating three entangled pairs and performing two swapping operations to establish entanglement between S and D. In TPED, we consider generating an entangled pair at N_1, then sending half to S and the other half to N_2. Meanwhile, N_2 establishes an entanglement connection with D. Now, we have a chain of entanglement ($S \longleftrightarrow N_2 \longleftrightarrow D$) between S and D shown in Fig. 4c. By performing entanglement swapping at nodes N_2, S and D share a direct entangled pair that can be used for communication. This approach only requires two entangled pairs and a single entanglement swapping operation to establish entanglement between S and D.

The TPED assigns functionalities to each node along the routing path to establish entanglement between communication parties while minimizing the number of entanglement swaps. This approach also considers node limitations, such as the inability of ground nodes to serve as sources of entanglement. The TPED policy depends on the number of nodes involved in the communication. If the number of nodes in the routing path is odd, the policy will follow the following steps:

1) Start from the first intermediate node, which is the node after the source, to the last intermediate node, which is the node before the destination. In an alternating pattern, each node is marked as either a source of entanglement or a repeater, respectively.

2) Each source of entanglement generates an entangled pair and sends half to the preceding node and the other half to the subsequent node of the routing path. Hence, the preceding and the subsequent nodes are connected by this source of entanglement.

3) Each repeater performs entanglement swapping to directly entangle the source and the destination.

If the number of hops is even, the policy designates one intermediate node to serve a dual role, acting both as a source of entanglement and as a repeater. This node is referred to as the hybrid repeater. Meanwhile, the remaining intermediate nodes continue to serve as either sources of entanglement or repeaters. It is worth mentioning that the SPARQ network limits the functionality of the ground nodes to either end-node or repeaters. Therefore, ground nodes cannot be used as a source of entanglement. Then, the policy goes as follows:

1) The initial step is to select the hybrid repeater, where the role of the hybrid repeater is assigned only to either the first or the last intermediate node of the routing path. The selection is based on the type of the two nodes that they connect in the routing path. One of these nodes is a ground node (source or destination) and the other can be ground, satellite, or HAP. The intermediate node is selected as the hybrid repeater if it connects one ground node (source or destination) with one source of entanglement (satellite or HAP) in the routing path. In cases where both the first and last intermediate nodes are eligible for selection as the hybrid repeater, either of them can be chosen. These restrictions on the choice of

the hybrid repeater allow the network nodes to follow a deterministic policy during entanglement distribution.

2) Start from the first intermediate node to the last intermediate in the routing path. In an alternating pattern, each node except for the hybrid repeater is marked as either a source of entanglement or a repeater, respectively.

3) Each source of entanglement generates an entangled pair and sends half to the preceding node and the other half to the subsequent node of the routing path. Hence, the preceding and the subsequent nodes are connected by this source of entanglement. Also, the hybrid repeater establishes entanglement with either the source or the destination based on the selection of the hybrid repeater.

4) Each repeater and the hybrid repeater perform entanglement swapping to establish a direct entanglement between the source and the destination.

The whole process with the proposed entanglement distribution based on the TPED policy is illustrated in Fig. 3.

V. SIMULATION RESULTS AND ANALYSIS
A. SPARQ SIMULATOR

Currently, quantum network simulators, such as QuNetSim [29], NetSquid [30], QDNS [31], SQUANCH [32], SeQUeNCe [33], SimulaQron [34], and SimQN [35] primarily focus on ground networks and terrestrial communications. There is no existing quantum network simulator that can be used to carry out SPARQ's simulation. To bridge this gap, we have upgraded the QuNetSim simulator to evaluate the performance of the proposed SPARQ's architecture and the proposed routing model. This upgraded version of the QuNetSim simulator is called SPARQ simulator. We chose QuNetSim because it is easy to configure and upgrade, integrates with multiple backends, and includes several built-in protocols such as teleportation and entanglement swapping. Also, it utilizes graphs to represent the network, supports the creation of separate networks for classical and quantum communications, and offers flexibility in developing and testing routing algorithms for both classical and quantum networks. Here, we note that QuNetSim is an abstract quantum network simulator, focusing more on higher-level quantum networking protocols than on the detailed physical layer. As a result, it lacks detailed modeling of specific hardware characteristics, such as the imperfections and noise associated with quantum repeaters, detectors, or entanglement sources. Additionally, the simulator assumes perfect quantum memory, where quantum states can be stored without degradation. Our SPARQ simulator integrates QuNetSim and STK [49] to account for realistic space-air-ground node behaviors and models. While QuNetSim already features an implementation of fiber optic channels, it lacks the implementation of FSO channels. Hence, an FSO channel model was implemented as described in Section II-C2. This implementation involves creating a new class within the simulator for FSO channels. This class holds the FSO channel properties, which are configured using the constructor. Furthermore, a function is defined within the class to calculate and store the trans-
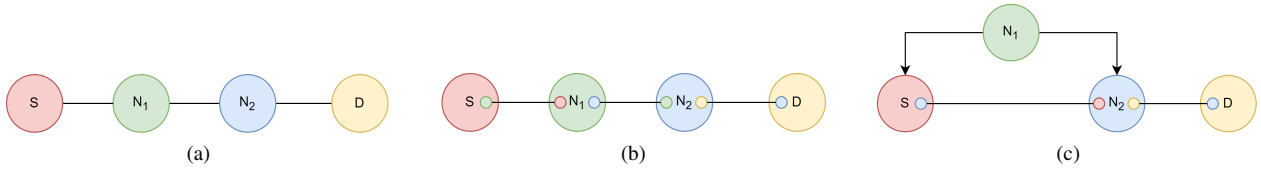
IEEE Transactions on
Quantum Engineering

FIGURE 4: Two different policies to establish entanglement between the source node S and the destination node D, the circles inside nodes represent an entanglement state. The color of the entanglement state represents the node with which an entanglement state is shared. In this figure: (a) represents the network nodes and communication channels before establishing entanglement, (b) represents the intuitive approach, and (c) represents the proposed TPED policy, where the role of N_1 is to generate an entangled pair and send one part to S and the other part to N_2.

missivity according to (2). Additionally, various functions are implemented to calculate the equations necessary for determining the transmissivity. These functions are invoked by the primary transmissivity calculation function. Notably, the setter for the distance recalculates the transmissivity for this channel. Therefore, adjusting the channel distance during satellite motion results in updating the transmissivity. This ensures the accurate computation of transmissivity for satellite links during movement. Moreover, QuNetSim defines the *Host* class for representing ground nodes within quantum networks. We extended this class to include information about node location such as latitude, longitude, and altitude. Furthermore, new classes were implemented to model HAPs and satellites within the proposed SPARQ network. These classes extend the *Host* class and implement specific properties relevant to each host type. For example, the *Satellite* class includes a movement list detailing the satellite's successive locations. Additionally, we define a thread for updating the satellite's location. This thread adjusts the satellite's current location to the next entry in the movement list. During each step of satellite movement, the satellite updates the distance to each connected node, which results in updated transmissivity. Furthermore, new functions were defined to degrade the quality of entangled states as described in (11), and measure the quality of entanglement using the entanglement fidelity as defined by (12). Furthermore, the STK simulator is used to simulate the movement of satellites. Each satellite is created in its orbit, ensuring a uniform distribution of orbits around the Earth. Subsequently, the simulation is executed to model the movement of satellites throughout one day, recording the location of each satellite at intervals of every 30 seconds. These recorded locations are used to generate movement sheets which include precise locations of each satellite while moving including latitude, longitude, and altitude. These movement sheets are then employed to simulate satellite movements within the SPARQ simulator. Specifically, when creating a satellite node within SPARQ, the corresponding sheet for that satellite is converted to a movement list. As the satellite navigates within the network, its current location is continuously updated to the subsequent entry in the movement list as described earlier.
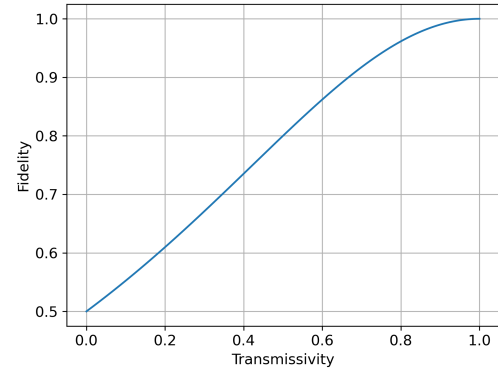

FIGURE 5: Relationship between transmissivity and entanglement fidelity.

B. IDENTIFYING THE TRANSMISSIVITY THRESHOLD

In Section II-D, we indicated that a transmissivity threshold is needed to decide the SPARQ topology, and hence, identify a routing path following our proposed RL-based strategy. Here, we conducted an experiment using the SPARQ simulator to determine this transmissivity threshold that can be used for quantum communication. This experiment aims to show the relationship between the transmissivity and the entanglement fidelity. To achieve this, we established different quantum links with different transmissivity ranging from 0 to 1 in a step of 0.01. Regardless of the node and communication channel types, the entanglement fidelity remains consistent for the same value of transmissivity, as described in (10) and (11). Therefore, for simplicity, a scenario with two ground nodes connected by a fiber optical channel is considered in this experiment. To determine the channel's parameters that should be used to achieve the desired transmissivity ranging from 0 to 1, we solve (1) setting the attenuation coefficient α to 1 for simplicity. Hence, for each channel, without loss of generality, the attenuation coefficient α is set to 1, and the distance between the two nodes is set to $-\ln(x)$, where x is the desired transmissivity. In this manner, mobility is implicitly considered. Then, an entangled pair is established across each link and the fidelity is measured for each entanglement. The relationship is shown in Fig. 5.

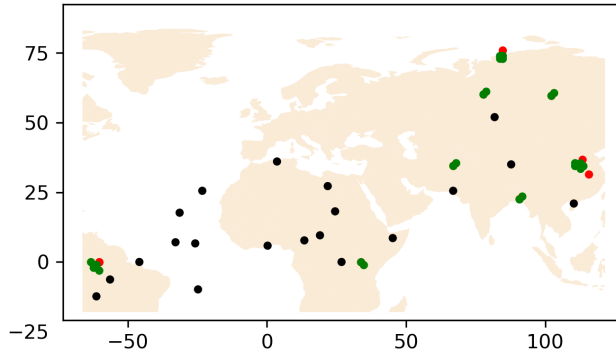As shown in Fig. 5, a transmissivity of 0.7 results in an

FIGURE 6: Setup of the SPARQ network used for simulation. Ground nodes are presented in green, satellite nodes are presented in black, and HAP nodes are presented in red.

FIGURE 7: Average training loss of the proposed DQN algorithm.

entanglement fidelity exceeding 90%, which can enable high-fidelity teleportation and hence, exchanging quantum information [53], [54]. Therefore, the transmissivity threshold is set to 0.7. The transmissivity threshold is the limit that determines whether transmissivity is sufficient for enabling teleportation and, hence, allowing quantum communications. It serves as a criterion for establishing connectivity between nodes within SPARQ, determining whether they are effectively connected or disconnected. A link is established between two nodes if its transmissivity is at least equal to the specified threshold. For instance, during satellite movements, transmissivities undergo continuous changes, leading to dynamic connections and disconnections of links based on the transmissivity threshold.

C. SIMULATION SETUP

As a proof of concept, the proposed DQN-based routing strategy is tested in a SPARQ network with 54 nodes, containing 24 ground nodes, 20 satellites, and 10 HAPs. The satellites operate in LEO at an altitude of 500 km, similar to the Micius satellite [5]. This altitude is chosen to maximize the satellites' lifetime while minimizing signal loss [55]. The HAPs are positioned at an altitude of 50 km, which is the maximum flying altitude defined by the International Telecommunication Union (ITU) for HAPs, ranging from 20 to 50 km [56]. We selected the maximum altitude for HAPs to minimize the length of satellite links, thereby enhancing link quality and coverage. The simulation parameters used in our experiments are summarized in Appendix A. The network setup is shown in Fig. 6 where ground nodes are presented in green, satellite nodes are presented in black, and HAP nodes are presented in red. The node locations are randomly selected by generating random latitude and longitude coordinates, ensuring through the Geopy [57] library that the generated locations are not over the sea. Moreover, some nodes are manually positioned to bring them into proximity with other nodes, forming local networks. Detailed information regarding the coordinates of these nodes and the configuration of satellites is provided in
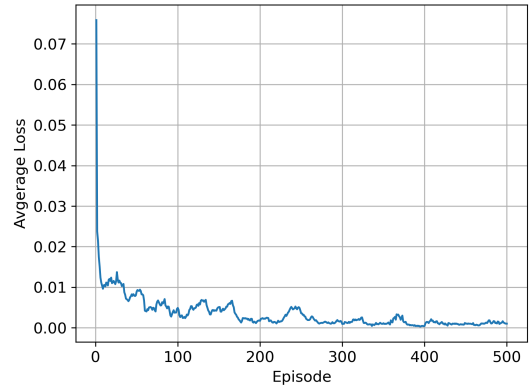
Appendix C. The network architecture is designed to have several local networks separated by large distances across several countries. Each local network contains a few ground nodes connected using fiber optical channels. Communication links are established between ground nodes if the transmissivity is at least 0.7. HAPs may exist to extend the communication distance between network nodes and fulfill teleportation requests whenever satellites are not in range. Similarly, communication links are established among HAPs and between HAPs and ground nodes if the transmissivity is at least 0.7. Each satellite follows its orbital trajectory, dynamically connecting and disconnecting from network nodes based on the transmissivity threshold. Notably, this architecture serves both classical and quantum data transmission, as quantum channels can be used for transmitting both types of data [58]. We chose this architecture to simulate the early-stage scenario of quantum networks where the quantum nodes will be distributed across vast geographical areas. For instance, IBM is working to establish quantum computers across the globe, including locations such as Germany, Japan, South Korea, and Canada in addition to existing quantum computers in the United States [59]–[61].

D. SIMULATION RESULTS

Fig. 8 compares the performance of the proposed TPED policy for on-demand entanglement distribution with the intuitive approach. The time step reflects the movements of the satellites. In this experiment, the communication parties are chosen randomly from the ground nodes, the proposed DQN algorithm is used to find the optimal routing path, and each entanglement distribution policy is used separately to establish on-demand end-to-end entanglement between the communication parties. The entanglement fidelity is measured for each entangled pair established by each policy. The experiment is repeated 100 times and the average is recorded for each time step over 100 moves of satellites. As shown in Fig. 8, the TPED policy outperforms the intuitive approach. Specifically, the TPED policy achieved an average entanglement fidelity of 0.988, while the intuitive approach achieved
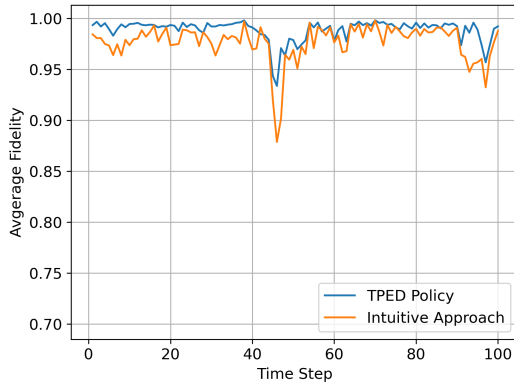
FIGURE 8: omparing the TPED policy with the intuitive approach in terms of the average entanglement fidelity.
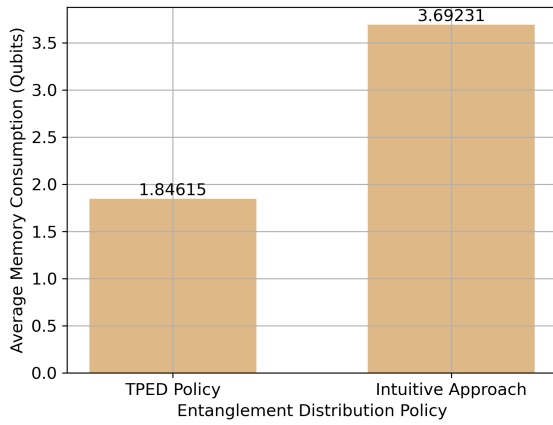

FIGURE 9: Comparing the TPED policy with the intuitive approach in terms of the average memory consumption.

an average entanglement fidelity of 0.96. This means that the TPED policy improved the average entanglement fidelity by about 3% compared with the intuitive approach. This is because TPED minimizes the number of utilized entangled pairs and minimizes the number of entanglement swapping operations performed. Furthermore, Fig. 9 compares the TPED policy and the intuitive approach in terms of average memory units consumed by intermediate repeaters. In the scenario of Fig. 9, the source and destination nodes are randomly selected from the ground nodes. Each policy is then employed to distribute entanglement between the source and destination, with the number of memory units consumed by intermediate repeaters being recorded. The experiment is repeated 100 times, and the average memory consumption is calculated. This improvement is attributed to the efficiency of the proposed TPED policy, which minimizes the number of utilized entangled pairs. As shown in Fig. 9, the TPED policy reduces memory consumption by 50%. As such, the TPED policy will be adopted in all experiments.

Fig. 10 compares the proposed DQN algorithm with the shortest path algorithm in terms of the average transmissivity

while satellites are moving. In this experiment, the shortest path algorithm is selected as a baseline for performance evaluation due to its extensive use in the literature for addressing the quantum routing problem [9]–[13], [15]–[19]. In order to conduct this experiment, the source and destination are chosen randomly from the ground nodes. Then, the proposed DQN scheme is used to find the optimal path between the source and the destination over 100 steps of satellites. Also, the Dijkstra algorithm is used to find the shortest path between the two parties involved in the communication. The experiment is repeated 100 times and the average is recorded. As shown in Fig. 10, the proposed DQN algorithm outperforms the shortest path algorithm and is able to select links with high transmissivity. This is because the shortest path algorithm only focuses on link distance and does not consider link quality. Furthermore, the proposed DQN algorithm is compared with both the shortest path algorithm and an LSTM-based RL algorithm in terms of the average fidelity of the end-to-end entanglement shared between communication parties. The implemented LSTM-based RL architecture comprises an LSTM layer with 64 neurons, followed by a fully connected layer. This LSTM layer processes a state representation vector, which is then passed into the fully connected layer to generate Q-values. These Q-values represent the expected cumulative rewards for different actions. It is worth noting that the LSTM-based RL algorithm is trained on the dynamic network topology of SPARQ, similar to the proposed DQN algorithm. The corresponding results are illustrated in Fig. 11 for the comparison with the shortest path algorithm and in Fig. 12 for the comparison with the LSTM-based RL algorithm. In order to obtain these results, the communication parties are selected randomly from ground nodes. Then, the proposed DQN, the shortest path, and the LSTM-based RL algorithms are used to find the optimal routing path over 100 satellite steps of satellites. Finally, the TPED policy is used to distribute on-demand entanglement between communication parties and the entanglement fidelity is measured. This experiment is repeated 100 times and the average entanglement fidelity is recorded in each time step. As shown in Fig. 11, the proposed DQN algorithm can distribute entangled pairs with the entanglement quality that is essential for successful teleportation (fidelity > 0.9) most of the time, while the shortest path algorithm fails to meet this requirement most of the time. Specifically, the proposed DQN algorithm is able to meet this requirement 100% of the time, while the shortest path algorithm is only able to meet this requirement 61% of the time. This indicates that the proposed DQN algorithm improved the number of resolved teleportation requests by 39% compared with the shortest path algorithm. This is because the shortest path algorithm prioritizes link distance over the link quality. Although both the proposed DQN and the LSTM-based RL algorithms follow the same pattern, the proposed DQN algorithm improves the end-to-end entanglement fidelity by about 2% as shown in Fig. 12. The comparable results can be attributed to the utilization of the same foundational RL algorithm, with both
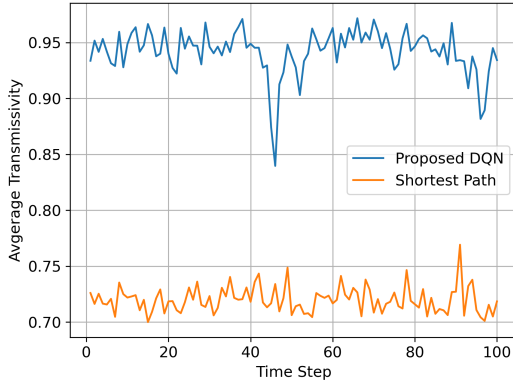
FIGURE 10: Comparing the proposed DQN algorithm with the shortest path algorithm in terms of the average transmissivity.
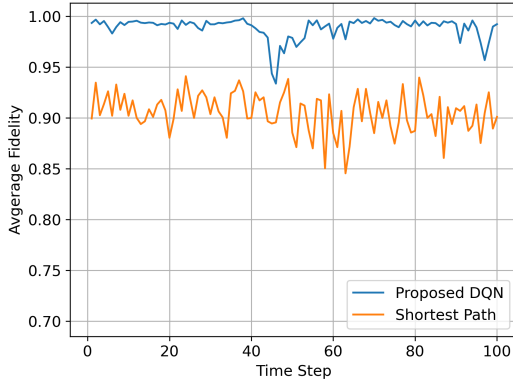

FIGURE 12: Comparing the proposed DQN algorithm with the LSTM-based RL algorithm in terms of the average entanglement fidelity.

FIGURE 11: Comparing the proposed DQN algorithm with the shortest path algorithm in terms of the average entanglement fidelity.

FIGURE 13: Comparing the proposed DQN algorithm with benchmarks [10] and [20] in terms of average entanglement fidelity.

RL algorithms trained on the dynamic network topology of SPARQ. The observed improvement can be attributed to the effective utilization of the experience buffer mechanism, which plays a pivotal role in the DQN's outperformance of the LSTM-based RL. It is worth mentioning that implementing an experience buffer in the LSTM-based RL poses a significant challenge due to the inherent nature of LSTM, which necessitates the simultaneous input of multiple related time steps. Furthermore, we have calculated the confidence interval based on the normal distribution, and with 95% confidence. We assert that the mean entanglement fidelity of the teleportation request solved by the proposed DQN algorithm falls within the range of 0.98 to 0.99.

Moreover, we compared the proposed DQN algorithm with two benchmarks proposed in [10] and [20]. In this comparison, the communication parties are randomly selected from ground nodes, and each routing algorithm is executed to find the optimal path between them over 100 time
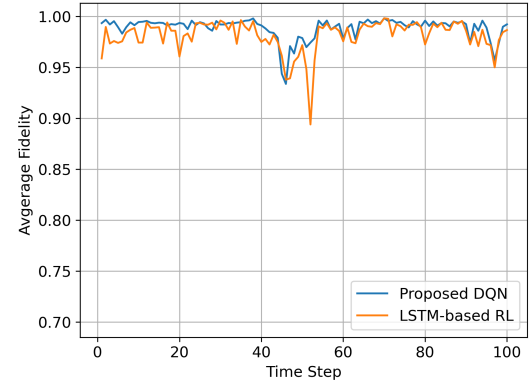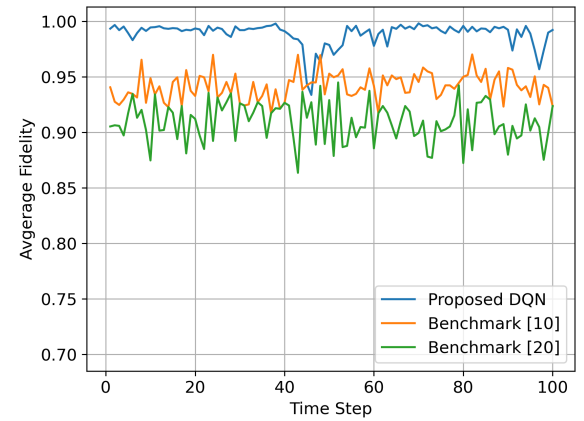
steps. Subsequently, the TPED policy is utilized to establish entanglement between the source and destination, and the resulting end-to-end entanglement fidelity is measured. This process is repeated 100 times and the average entanglement fidelity is calculated. As shown in Fig. 13, the proposed DQN model outperforms both benchmarks. In particular, the DQN algorithm improves the average entanglement fidelity by about 6% and 9% compared with [10] and [20], respectively. The enhancement over the benchmark proposed in [10] can be attributed to the consideration of both link quality and the number of hops in our DQN algorithm, whereas [10] solely focuses on the number of hops. Similarly, the improvement over the benchmark proposed in [20] can be attributed to the joint consideration of link quality and the number of hops in finding the optimal path in our DQN algorithm, while [20] prioritizes the key rate.

In order to evaluate the overhead introduced by the proposed DQN algorithm, we conducted an experiment to compare the average time required for different algorithms to find
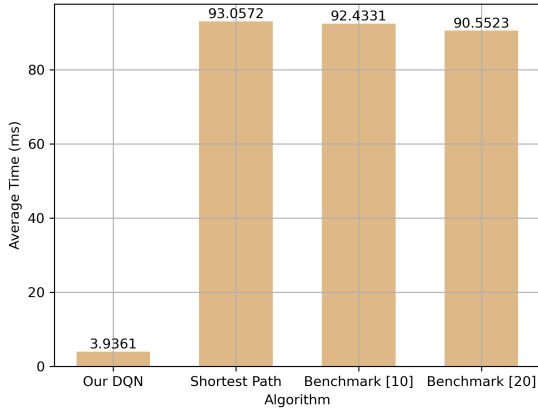
FIGURE 14: Comparing the proposed DQN algorithm with the shortest path algorithm and benchmark algorithms [10] and [20] in terms of the average time required to find the optimal path between communication parties.
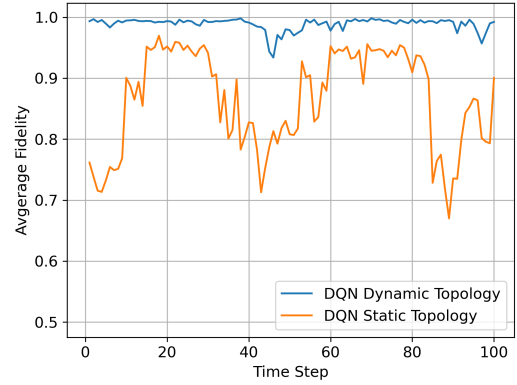
FIGURE 15: Comparing two versions of the DQN algorithm: one is trained on the dynamic network topology of SPARQ, while the other is trained on a static snapshot of SPARQ.

the optimal path between a source and destination over 100 trials. As shown in Fig. 14, the proposed DQN algorithm exhibits the lowest average time among all algorithms, with an average time of approximately 3.94 milliseconds. This highlights the efficiency of our DQN algorithm in optimizing routing requests within SPARQ. In contrast, traditional methods such as the shortest path algorithm and the benchmarks [10] and [20] require significantly more time, with their average times ranging from 90 to 93 milliseconds. This improvement can be attributed to the efficiency of RL, which simplifies the routing problem by learning through experience. As mentioned earlier, the training of the DQN agent was performed offline to ensure efficient real-time performance. Additionally, the efficient representation of the system state contributes to the agent's efficiency, as it requires only local knowledge to choose the next action. In contrast, benchmark algorithms require global knowledge, and they need to calculate the routing path for each communication request. The substantial reduction in simulation time achieved by our DQN algorithm underscores its practical viability for real-time routing optimization in dynamically changing network environments like SPARQ.

Furthermore, the proposed DQN algorithm is compared with another DQN algorithm trained in the same setup on a single snapshot of the SPARQ network. This comparison aims to highlight the significance of training the agent on the dynamic network topology of SPARQ. Similarly, in this experiment, the communication parties are chosen randomly from ground nodes, and each DQN algorithm is applied to find the optimal path between the communication parties over 100 time steps. Finally, the TPED policy is used to establish entanglement between the source and destination and the entanglement fidelity is measured and recorded. This experiment is repeated 100 times and the average entanglement fidelity is calculated. As shown in Fig. 15, the DQN algorithm that is trained on the dynamic network topology outperforms the other algorithm trained on a static snapshot

of the SPARQ network. Specifically, the proposed DQN algorithm achieves an average entanglement fidelity of 0.988, while the other algorithm achieves an average entanglement fidelity of 0.86, which indicates an improvement by about 15%. This is because training the agent on the dynamic network topology allows the agent to capture all states of the network graph and choose the optimal route based on the encountered network state.

In order to highlight the importance of the air layer within SPARQ, we trained another DQN algorithm to find the optimal routing path without the existence of the air layer. In other words, the DQN algorithm is trained on a network that spans only the space and the ground layers. Similarly, The source and the destination are selected randomly from the ground layer. To extend satellite coverage in the absence of HAPs and facilitate the establishment of paths between communication parties, we adjusted the transmissivity threshold to 0.5. Each DQN algorithm is used to find the optimal routing path within the corresponding network over 100 time steps. Then, the TPED policy is used to establish entanglement between the source and the destination, and the entanglement fidelity is measured. The average is calculated across 100 different trials and the results are presented in Fig. 16. The proposed DQN algorithm in SPARQ achieves an average entanglement fidelity of 0.988, while the architecture that spans only satellite and ground layers achieves an average entanglement fidelity of 0.8, which indicates an improvement by 23.5%. As shown in Fig. 16, the performance of the SPARQ network with the existence of the air layer is much better than the performance without the air layer. This is because of two reasons. Firstly, the air layers bridge the considerable distance between satellite nodes and ground nodes. Secondly, the air layer satisfies teleportation requests when satellites are out of range.
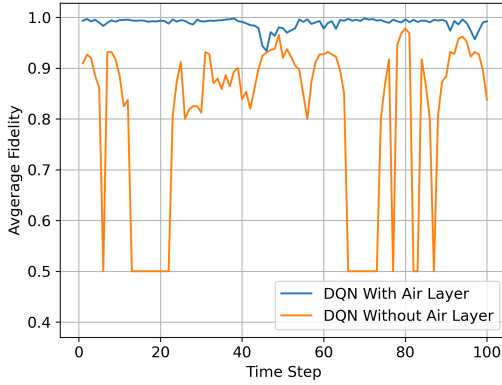
FIGURE 16: Comparing the SPARQ network performance with a network that spans only space and ground layers.

E. DISCUSSION

The proposed SPARQ network and its associated algorithms are evaluated across several key performance metrics, including average entanglement fidelity, the number of resolved teleportation requests, memory consumption, and the time required to find a routing path between communication parties. For instance, an experiment comparing SPARQ with a network that spans only space and ground layers highlights the critical role of the air layer. The inclusion of the air layer in SPARQ achieves a 23.5% improvement in average entanglement fidelity, underscoring its significant contribution to overall network performance. Within the SPARQ architecture, the TPED policy improves average entanglement fidelity by 3% compared to the intuitive approach. The proposed DQN algorithm improves the fidelity compared to state-of-the-art benchmarks while yielding significant gains in other performance metrics. For instance, the proposed TPED policy achieves a 50% reduction in memory consumption, underscoring its efficiency in managing quantum resources. Additionally, the proposed DQN algorithm enhances the number of resolved teleportation requests by 39% compared to the shortest path baseline. Furthermore, the proposed DQN algorithm also significantly reduces the time required to find the routing path, demonstrating a substantial improvement in speed over state-of-the-art algorithms, thereby optimizing network routing and resource utilization.

VI. CONCLUSION

In this paper, we have proposed a space-air-ground quantum network designed to enable quantum Internet applications in the near term. SPARQ aims to allow on-demand and seamless entanglement distribution, which enables applications such as distributed quantum computing, quantum communications, quantum sensing, and quantum security. We have then developed a DQN-based reinforcement learning algorithm to efficiently solve the routing problem between communication parties within SPARQ. Additionally, we have proposed an entanglement distribution policy in order to establish on-demand entanglement between communication parties. Fur-

thermore, we have upgraded an existing quantum network simulator to carry out SPARQ's simulation and evaluate the performance of SPARQ, the proposed DQN algorithm, and the entanglement distribution policies. Simulation results show that the proposed TPED policy improved the average end-to-end entanglement fidelity by 3% and reduced the memory consumption by 50% compared with the intuitive approach. The proposed DQN algorithm is compared with the shortest path algorithm, an LSTM-based RL algorithm, two state-of-the-art benchmark algorithms, and another DQN algorithm that is trained on a static snapshot of SPARQ. The proposed DQN algorithm demonstrated significant improvements across various algorithms. Compared with the shortest path algorithm, it enhanced the number of resolved teleportation requests by 39%. Additionally, in comparison with an LSTM-based algorithm, it achieved a 2% increase in average entanglement fidelity. Furthermore, when benchmarked against two state-of-the-art algorithms, the proposed DQN model improved the average fidelity by 6% and 9%, respectively. Moreover, it showcased an average improvement of 15% in end-to-end entanglement fidelity compared to a DQN trained on a snapshot of SPARQ. Also, to highlight the importance of the air layer, we have conducted an experiment to compare the efficiency of SPARQ and the network spanning only space and ground layers. The results also show that communication within SPARQ is more efficient when the air layer is included. SPARQ enhanced the average end-to-end entanglement fidelity by 23.5% compared with the network that spans only space and ground layers. The attained entanglement fidelity has the potential for improvement through entanglement purification but it will consume additional quantum resources. However, we leave the exploration of efficient entanglement purification to future work. Additionally, future research may consider identifying optimal satellite trajectories to improve coverage and connectivity between distant nodes. Moreover, future extensions could improve the SPARQ simulator by incorporating more realistic physical layer models that address hardware imperfections, such as quantum memory decoherence and deficiencies in quantum devices like repeaters and detectors.

.

APPENDIX A SIMULATION PARAMETERS

The attenuation coefficient is set to 0.15 dB/km for fiber optical channels as it is used in [42]. The simulation parameters of the FSO channels are set as presented in [43]. The learning rate β is set to 0.001, while the exploration rate (ϵ) is initialized to be 1 and decays by 0.9995 until reaches 0.01. The Adam optimizer is used to minimize the loss by updating the weights of the neural network. The average training loss is visualized in Fig. 7, where the average is calculated based on all losses encountered during the 100 mini-episodes used to train the model on the current network graph. The simulation parameters are summarized in Table 1.

IEEE Transactions on
Quantum Engineering

TABLE 1: Simulation parameters.

Parameter	Value
Number of satellites	20
Altitude of satellites	500 km
Number of HAPs	10
Flying height of HAPs	50 km
Number of ground nodes	24
Number of episodes	500
Number of mini-episodes	100
β	0.001
ϵ	$1 - 0.01$
decay rate of ϵ	0.9995
γ	0.99
Optimizer	Adam
Hidden layers	2
Number of neurons	64, 128
Activation function	ReLU
Transmissivity Threshold	0.7
α	0.15 dB/km
η_{eff}	1
a_R	5 m
R_0	∞
$\alpha_0(\lambda)$	$5 \times 10^{-6} \mathrm{m}^{-1}$
l_0	1 mm
w_0	20 cm
λ	800 nm

APPENDIX B TRAINING DATASET

Our dataset is dynamically generated during the training phase as the DQN agent interacts directly with the network environment to learn from experience. To facilitate the learning process, we simulate the SPARQ network before the actual system is operational, allowing the agent to encounter different network graphs and learn optimal routing policy. Transmissivities to adjacent nodes serve as a feature for training and identifying the current network state. For each graph, routing requests are randomly generated between ground nodes, and the agent's routing decisions lead to either rewards or penalties. Starting from the source node, the agent selects actions to visit one of the adjacent nodes, and it continues until reaching the destination. Initially, the agent takes random actions to explore diverse paths and strategies for finding the optimal route for entanglement distribution. As training progresses, the DQN agent uses the received rewards to update its internal Q-values. The Q-values represent the expected cumulative reward for taking a particular action in a given state. Throughout training, the exploration rate (ϵ) gradually decreases, leading to a reduced dependence on random actions. Instead, the agent increasingly prioritizes actions based on its learned policy. This transition from exploration to exploitation occurs until the agent converges toward a strategy for finding an effective path for entanglement distribution. Here, we note that the agent relies solely on local information to determine the next action at each node. This local information, represented by transmissivities to adjacent nodes, provides insight into

channel quality and connectivity. Since satellite trajectories are relative and periodic, this local information effectively represents the network state, enabling the agent to optimize routing requests based on encountered network states.

Several key factors underscore the significance of this dataset. Firstly, it enables the agent to interact with diverse network topologies and configurations, reflecting the dynamic nature of SPARQ. This interaction with a wide range of network scenarios ensures that the trained model can effectively adapt to varying environmental conditions. Secondly, it allows the agent to optimize routing requests between several communication nodes at each network state during training. Consequently, the dataset facilitates the development of a comprehensive policy for identifying the optimal routing path for entanglement distribution.

APPENDIX C NODES CONFIGURATIONS

The coordinates for the ground nodes are shown in Table 2, and the coordinates for the HAPs are listed in Table 3. Furthermore, the orbital configurations of the satellites are detailed in Table 4, including their inclination, right ascension of the ascending node (RAAN), and altitude. Notably, these orbital configurations can be used in the STK simulator to replicate the trajectories of the satellites.

TABLE 2: Coordinates of ground nodes.

Ground Nodes		
(34.315,112.732)	(34.336,112.736)	(34.357,112.740)
(34.378,112.746)	(34.399,112.750)	(34.415,112.754)
(72.823,84.759)	(72.844,84.765)	(72.865,84.771)
(72.876,84.777)	(−0.062,−60.266)	(−0.083,−60.270)
(−0.103,−60.278)	(−0.113,−60.285)	(−0.126,33.812)
(−0.105,33.816)	(59.561,102.211)	(59.568,102.251)
(22.455,90.777)	(22.434,90.774)	(34.411,66.991)
(34.426,66.972)	(60.064,77.805)	(60.043,77.813)

TABLE 3: Coordinates of HAPs.

HAPs		
(34.368,112.738)	(−0.062,−60.266)	(−0.162,−60.235)
(−0.105,33.832)	(36.644,113.444)	(34.418,66.979)
(60.044,77.848)	(59.561,102.251)	(75.834,84.763)
(22.485,90.777)		

IEEE Transactions on
QuantumEngineering

TABLE 4: Satellites orbital configurations.

Satellites		
Inclination (deg)	RAAN (deg)	Altitude (km)
36	126	500
54	126	500
105	126	500
36	233	500
72	304	500
108	15	500
125	53	500
144	91	500
0	0	500
18	158	500
73	158	500
90	160	500
127	160	500
144	160	500
163	160	500
0	200	500
18	228	500
55	300	500
90	13	500
162	160	500

REFERENCES

[1] H. P. Paudel, M. Syamlal, S. E. Crawford, Y.-L. Lee, R. A. Shugayev, P. Lu, P. R. Ohodnicki, D. Mollot, and Y. Duan, "Quantum computing and simulations for energy applications: Review and perspective," ACS Engineering Au, vol. 2, no. 3, pp. 151–196, Jun 2022. [Online]. Available: https://doi.org/10.1021/acsengineeringau.1c00033

[2] J.-Y. Wu, K. Matsui, T. Forrer, A. Soeda, P. Andrés-Martínez, D. Mills, L. Henaut, and M. Murao, "Entanglement-efficient bipartite-distributed quantum computing," Quantum, vol. 7, p. 1196, 12 2023.

[3] R. Kaewpuang, M. Xu, D. Niyato, H. Yu, and Z. Xiong, "Resource allocation in quantum key distribution (QKD) for space-air-ground integrated networks," in Proc. of IEEE 27th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD), Paris, France, Nov. 2022, pp. 71–76.

[4] M. Xu, D. Niyato, Z. Xiong, J. Kang, X. Cao, X. S. Shen, and C. Miao, "Quantum-secured space-air-ground integrated networks: Concept, framework, and case study," IEEE Wireless Communications, vol. 30, no. 6, pp. 136–143, 2023.

[5] C.-Y. Lu, Y. Cao, C.-Z. Peng, and J.-W. Pan, "Micius quantum experiments in space," Rev. Mod. Phys., vol. 94, p. 035001, Jul 2022. [Online]. Available: https://link.aps.org/doi/10.1103/RevModPhys.94.035001

[6] "The european quantum communication infrastructure (EuroQCI) initiative." [Online]. Available: https://digital-strategy.ec.europa.eu/en/policies/european-quantum-communication-infrastructure-euroqci

[7] M. Caleffi, "Optimal routing for quantum networks," IEEE Access, vol. 5, pp. 22 299–22 312, 2017.

[8] S. Shi and C. Qian, "Concurrent entanglement routing for quantum networks: Model and designs," in Proc. of the Annual Conference of the ACM Special Interest Group on Data Communication on the Applications, Technologies, Architectures, and Protocols for Computer Communication, ser. SIGCOMM '20.   New York, NY, USA: Association for Computing Machinery, Aug. 2020, p. 62–75. [Online]. Available: https://doi.org/10.1145/3387514.3405853

[9] M. Pant et al., "Routing entanglement in the quantum internet," npj Quantum Information, vol. 5, no. 1, p. 25, Mar 2019. [Online]. Available: https://doi.org/10.1038/s41534-019-0139-x

[10] O. Amer, W. Krawec, and B. Wang, "Efficient routing for quantum key distribution networks," Denver, CO, USA, Oct. 2020, pp. 137–147.

[11] J. Nötzel and S. DiAdamo, "Entanglement-enhanced communication networks," in Proc. of IEEE International Conference on Quantum Computing and Engineering (QCE), Denver, CO, USA, Oct. 2020, pp. 242–248.

[12] C. Li et al., "Effective routing design for remote entanglement generation on quantum networks," npj Quantum Information, vol. 7, no. 1, p. 10, Jan 2021. [Online]. Available: https://doi.org/10.1038/s41534-020-00344-4

[13] L. Le and T. N. Nguyen, "DQRA: Deep quantum routing agent for entanglement routing in quantum networks," IEEE Transactions on Quantum Engineering, vol. 3, pp. 1–12, 2022.

[14] Y. Zhao, G. Zhao, and C. Qiao, "E2E fidelity aware routing and purification for throughput maximization in quantum networks," in Proc. IEEE Conference on Computer Communications (INFOCOM), London, United Kingdom, May 2022, pp. 480–489.

[15] L. Yang, Y. Zhao, H. Xu, and C. Qiao, "Online entanglement routing in quantum networks," in Proc. of IEEE/ACM 30th International Symposium on Quality of Service (IWQoS), Oslo, Norway, June 2022, pp. 1–10.

[16] H. Leone, N. R. Miller, D. Singh, N. K. Langford, and P. P. Rohde, "Qunet: Cost vector analysis multi-path entanglement routing in quantum networks," 2021.

[17] Y. Zeng, J. Zhang, J. Liu, Z. Liu, and Y. Yang, "Entanglement routing design over quantum networks," IEEE/ACM Transactions on Networking, pp. 1–16, 2023.

[18] T. N. Nguyen, D. H. P. Nguyen, D. H. Pham, B.-H. Liu, and H. N. Nguyen, "Maximizing entanglement routing rate in quantum networks: Approximation algorithms," 2022.

[19] R. Van Meter, T. Satoh, T. D. Ladd, W. J. Munro, and K. Nemoto, "Path selection for quantum repeater networks," Networking Science, vol. 3, no. 1, pp. 82–95, Dec 2013. [Online]. Available: https://doi.org/10.1007/s13119-013-0026-2

[20] D. L. Bakker, Y. Jong, B. P. F. Dirks, and G. C. Amaral, "A best-path approach to the design of a hybrid space–ground quantum network with dynamic constraints," Photonics, vol. 11, no. 3, 2024. [Online]. Available: https://www.mdpi.com/2304-6732/11/3/268

[21] C. Jones, K. D. Greve, and Y. Yamamoto, "A high-speed optical link to entangle quantum dots," 2013.

[22] F. F. da Silva, A. Torres-Knoop, T. Coopmans, D. Maier, and S. Wehner, "Optimizing entanglement generation and distribution using genetic algorithms," Quantum Science and Technology, vol. 6, no. 3, p. 035007, jun 2021. [Online]. Available: https://dx.doi.org/10.1088/2058-9565/abfc93

[23] S. Haldar, P. J. Barge, S. Khatri, and H. Lee, "Fast and reliable entanglement distribution with quantum repeaters: principles for improving protocols using reinforcement learning," 2023.

[24] S. Khatri, "Policies for elementary links in a quantum network," Quantum, vol. 5, p. 537, 2021.

[25] G. Iñesta, G. Vardoyan, L. Scavuzzo, and S. Wehner, "Optimal entanglement distribution policies in homogeneous repeater chains with cutoffs," npj Quantum Information, vol. 9, no. 1, p. 46, May 2023. [Online]. Available: http://dx.doi.org/10.1038/s41534-023-00713-9

[26] S. D. Reiß and P. van Loock, "Deep reinforcement learning for key distribution based on quantum repeaters," Physical Review A, vol. 108, no. 1, p. 012406, July 2023. [Online]. Available: http://dx.doi.org/10.1103/PhysRevA.108.012406

[27] M. Chehimi and W. Saad, "Entanglement rate optimization in heterogeneous quantum communication networks," in Proc. of the 17th International Symposium on Wireless Communication Systems (ISWCS), Berlin, Germany, Sept. 2021, pp. 1–6.

[28] M. Chehimi, S. Pouryousef, N. K. Panigrahy, D. F. Towsley, and W. Saad, "Scaling limits of quantum repeater networks," in Proc. of IEEE International Conference on Quantum Computing and Engineering (QCE), vol. 01, Bellevue, Washington, USA, Sept. 2023, pp. 1205–1210. [Online]. Available: https://api.semanticscholar.org/CorpusID:258686492

[29] S. Diadamo, J. Nötzel, B. Zanger, and M. M. Beşe, "QuNetSim: A software framework for quantum networks," IEEE Transactions on Quantum Engineering, vol. 2, pp. 1–12, 2021.

[30] T. Coopmans, R. Knegjens, A. Dahlberg, D. Maier, L. Nijsten, J. de Oliveira Filho, M. Papendrecht, J. Rabbie, F. Rozpędek, M. Skrzypczyk, L. Wubben, W. de Jong, D. Podareanu, A. Torres-Knoop, D. Elkouss, and S. Wehner, "Netsquid, a network simulator for quantum information using discrete events," Communications Physics, vol. 4, no. 1, p. 164, Jul 2021. [Online]. Available: https://doi.org/10.1038/s42005-021-00647-8

[31] O. S. Ceylan and Yılmaz, "Qdns: Quantum dynamic network simulator based on event driving," in Proc. of International Conference on Information Security and Cryptology (ISCTURKEY), Ankara, Turkey, Dec. 2021, pp. 45–50.

[32] B. Bartlett, "A distributed simulation framework for quantum networks and channels," 08 2018.

[33] X. Wu, A. Kolar, J. Chung, D. Jin, T. Zhong, R. Kettimuthu, and M. Suchara, "SeQUeNCe: a customizable discrete-event simulator of quantum networks," Quantum Science and Technology, vol. 6, no. 4, p. 045027, Sep. 2021. [Online]. Available: https://dx.doi.org/10.1088/2058-9565/ac22f6

[34] A. Dahlberg and S. Wehner, "SimulaQron—a simulator for developing quantum internet software," Quantum Science and Technology, vol. 4, no. 1, p. 015001, Sep. 2018. [Online]. Available: https://dx.doi.org/10.1088/2058-9565/aad56e

[35] L. Chen, K. Xue, J. Li, N. Yu, R. Li, Q. Sun, and J. Lu, "Simqn: A network-layer simulator for the quantum network investigation," IEEE Network, vol. 37, no. 5, pp. 182–189, 2023.

[36] N. Alshaer, T. Ismail, and M. E. Nasr, "Performance evaluation and security analysis of ground-to-satellite FSO system with CV-QKD protocol," IET Communications, vol. 14, no. 10, pp. 1534–1542, 2020. [Online]. Available: https://ietresearch.onlinelibrary.wiley.com/doi/abs/10.1049/iet-com.2019.0795

[37] F. Lauterbach, M. Vaněk, M. Mehic, and M. Voznak, "A study on quantum key distribution satellite communications," in Proc. of 15th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT). Ghent, Belgium: IEEE, Oct. 2023, pp. 128–133.

[38] P. V. Trinh, A. T. Pham, A. Carrasco-Casado, and M. Toyoshima, "Quantum key distribution over fso: Current development and future perspectives," in Proc. of Progress in Electromagnetics Research Symposium (PIERS-Toyama), Toyama, Japan, Aug. 2018, pp. 1672–1679.

[39] T. V. Nguyen, H. T. Le, H. T. T. Pham, V. Mai, and N. T. Dang, "Enhancing design and performance analysis of satellite entanglement-based cv-qkd/fso systems," IEEE Access, vol. 11, pp. 112 097–112 107, 2023.

[40] M. Q. Vu, H. D. Le, T. V. Pham, and A. T. Pham, "Design of satellite-based fso/qkd systems using geo/leos for multiple wireless users," IEEE Photonics Journal, vol. 15, no. 4, pp. 1–14, 2023.

[41] M. Q. Vu, T. V. Pham, N. T. Dang, and A. T. Pham, "Design and performance of relay-assisted satellite free-space optical quantum key distribution systems," IEEE Access, vol. 8, pp. 122 498–122 510, 2020.

[42] M. Mobayenjarihani, G. Vardoyan, and D. Towsley, "Optimistic entanglement purification with few quantum memories," in Proc. of IEEE International Conference on Quantum Computing and Engineering (QCE), Broomfield, CO, USA, Oct. 2021, pp. 439–440.

[43] M. Ghalaii and S. Pirandola, "Quantum communications in a moderate-to-strong turbulent space," Communications Physics, vol. 5, no. 1, p. 38, Feb 2022. [Online]. Available: https://doi.org/10.1038/s42005-022-00814-5

[44] M. Grassl, L. Kong, Z. Wei, Z.-Q. Yin, and B. Zeng, "Quantum error-correcting codes for qudit amplitude damping," IEEE Transactions on Information Theory, vol. 64, no. 6, pp. 4674–4685, 2018.

[45] J. L. Pereira and S. Pirandola, "Bounds on amplitude-damping-channel discrimination," Phys. Rev. A, vol. 103, p. 022610, Feb 2021. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevA.103.022610

[46] A. D'Arrigo, G. Benenti, G. Falci, and C. Macchiavello, "Classical and quantum capacities of a fully correlated amplitude damping channel," Phys. Rev. A, vol. 88, p. 042337, Oct 2013. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevA.88.042337

[47] F. Tonolini, S. Chan, M. Agnew, A. Lindsay, and J. Leach, "Reconstructing high-dimensional two-photon entangled states via compressive sensing," Scientific Reports, vol. 4, no. 1, p. 6542, Oct 2014. [Online]. Available: https://doi.org/10.1038/srep06542

[48] R. Jozsa, "Fidelity for mixed quantum states," Journal of Modern Optics, vol. 41, no. 12, pp. 2315–2323, 1994. [Online]. Available: https://doi.org/10.1080/09500349414552171

[49] AGI, "Ansys STK | digital mission engineering software." [Online]. Available: https://www.ansys.com/products/missions/ansys-stk

[50] S. Adam, L. Busoniu, and R. Babuska, "Experience replay for real-time reinforcement learning control," IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), vol. 42, no. 2, pp. 201–212, 2012.

[51] S. M. Kakade, "On the sample complexity of reinforcement learning," Ph.D. dissertation, England, 2003, ph.D.

[52] S. M. Zangi, C. Shukla, A. ur Rahman, and B. Zheng, "Entanglement swapping and swapped entanglement," Entropy, vol. 25, no. 3, p. 415, 2023. [Online]. Available: https://www.mdpi.com/1099-4300/25/3/415

[53] D. Nield, "Quantum teleportation was just achieved with 90% accuracy over a 44km distance," Dec 2020. [Online]. Available: https://www.sciencealert.com/scientists-achieve-sustained-high-fidelity-quantum-teleportation-over-44-km

[54] R. Valivarthi, S. I. Davis, C. Peña, S. Xie, N. Lauk, L. Narváez, J. P. Allmaras, A. D. Beyer, Y. Gim, M. Hussein, G. Iskander, H. L. Kim, B. Korzh, A. Mueller, M. Rominsky, M. Shaw, D. Tang, E. E. Wollman, C. Simon, P. Spentzouris, D. Oblak, N. Sinclair, and M. Spiropulu, "Teleportation systems toward a quantum internet," PRX Quantum, vol. 1, p. 020317, Dec 2020. [Online]. Available: https://link.aps.org/doi/10.1103/PRXQuantum.1.020317

[55] D. King-Hele and D. M. Walker, The prediction of satellite lifetimes. Royal Aircraft Establishment, 1987.

[56] M. S. Alam, G. K. Kurt, H. Yanikomeroglu, P. Zhu, and N. D. Đào, "High altitude platform station based super macro base station constellations," IEEE Communications Magazine, vol. 59, no. 1, pp. 103–109, 2021.

[57] K. Esmukov, "Geopy." [Online]. Available: https://pypi.org/project/geopy/

[58] G. Smith, "Quantum channel capacities," in 2010 IEEE Information Theory Workshop, 2010, pp. 1–5.

[59] D. Welle, "IBM unveils first quantum computer in germany," Jun 2021. [Online]. Available: https://www.dw.com/en/ibm-unveils-first-quantum-computer-in-germany/a-57909494

[60] A. Annunziata, "IBM takes its quantum computer to japan to launch countrywide quantum initiative," Nov 2023. [Online]. Available: https://research.ibm.com/blog/ibm-japan-launch-quantum-initiative

[61] D. McCann, "First IBM quantum system in canada," Apr 2022. [Online]. Available: https://www.ibm.com/blogs/ibm-canada/2022/02/first-ibm-quantum-system-in-canada/