Efficient and Flexible Multi-Qubit Entanglement Transmission in Quantum Networks

Yanan Gao, Song Yang, *Senior Member, IEEE,* Fan Li, *Member, IEEE,* Youqi Li, *Member, IEEE,* Liehuang Zhu, *Senior Member, IEEE,* Stojan Trajanovski, *Member, IEEE,* and Xiaoming Fu, *Fellow, IEEE*

***Abstract*—The unprecedented advancements in quantum technology have opened new prospects for the widespread adoption of quantum applications, placing new demands on the information transmission capabilities of large-scale quantum networks. Long-distance and stable entanglements are deemed as the lifeline in quantum network communication. However, some weaknesses, e.g., quantum decoherence, scarce quantum memory, and uneven-quality entanglement, of the quantum entanglement hinder the development. In this paper, we propose *Sophon*, an online transmission framework for quantum networks, which utilizes high-dimensional entanglements to concurrently transmit multi-qubit data to satisfy the transmission requirements of the real-time request set. We first model the quantum network with multi-qubit entanglement represented by *W* quantum state and then formulate the Entanglement Routing and Qubit Provisioning (*ERQP*) problem as a global-local optimization process. To solve the *ERQP* problem, we distributedly regard each network node as an RL agent for resource provisioning and extend the step-updating of the Markov Decision Process by introducing a centralized controller for entanglement route selection to optimize local and global objectives, respectively. Extensive simulations demonstrate, on the self-made simulation platform, *Sophon* achieves a $21.89\% - 66.52\%$ decrease in the communication cost, and is more robust on different scales of the network topology and the request set than the baselines.**

***Index Terms*—Quantum networks, multi-qubit entanglement, multi-agent reinforcement learning, branch-and-bound**

The work of Song Yang is partially supported by the National Natural Science Foundation of China under Grant 62472028, 62172038, the Beijing Natural Science Foundation under Grant 4232033, and the National Key R&D Program of China under Grant 2023YFB3107300. The work of Fan Li is partially supported by the National Natural Science Foundation of China under Grant 62372045. The work of Youqi Li is partially supported by the Beijing Natural Science Foundation under Grant L254058 and the Beijing Institute of Technology Research Fund Program for Young Scholars. The work of Liehuang Zhu is partially supported by Yunnan Provincial Major Science and Technology Special Plan Projects under Grant 202302AD080003. The work of Xiaoming Fu is partially supported by the Horizon Europe CODECO project under Grant 101092696, and in part by the Horizon Europe COVER project under Grant 101086228. *(Corresponding author: Song Yang)*

Y. Gao, S. Yang, F. Li, and Y. Li are with the School of Computer Science and Technology, Beijing Institute of Technology, Beijing 100081, China. e-mail: {yanangao, S.Yang, fli, liyouqi}@bit.edu.cn.

L. Zhu is with the School of Cyberspace Science and Technology, Beijing Institute of Technology, Beijing 100081, China. e-mail: liehuangz@bit.edu.cn.

S. Trajanovski is with Microsoft, W2 6BD London, United Kingdom. e-mail: sttrajan@microsoft.com.

X. Fu is with the Institute of Computer Science, University of Göttingen, 37077 Göttingen, Germany. e-mail: Fu@cs.uni-goettingen.de.

I. Introduction

Benefiting from the inherent privacy protection and high-speed transmission properties, the quantum mechanism is revolutionizing the conventional communication paradigm [1]. In light of the support around the world, the quantum application has witnessed unprecedented advancements [2], [3], such as Quantum Key Distribution [4], distributed quantum computing [5], quantum nuclear magnetic resonance [6], etc. As the carrier for quantum information transmission, the quantum network promises to provide reliable transmission for massive data to ensure seamless communication, thereby facilitating the widespread adoption of these quantum applications. With the strong data-carrying and noise-resistance abilities, the high-dimensional entanglement state is deemed as a potential approach [7], [8].

In quantum networks, qubit transmission relies on quantum entanglement, a medium-free communication channel established between two quantum nodes. Long-distance End-to-End (E2E) entanglements for responding to the quantum request are crucial components in large-scale quantum networks. However, the quantum decoherence, caused by the interaction noise between the quantum state and its surroundings [9], [10], brings more uncertainty to the entanglement establishment. It gradually deteriorates the entanglement quality and breaks the transmission connection, especially in long-distance E2E entanglements. Moreover, the scarce qubit memory of quantum nodes limits the establishment of more entanglement links, which hinders the quantum network from completing the transmission requirements of the request set. Therefore, reasonably routing E2E entanglement routes and efficiently provisioning qubit resources to relieve the pressure of quantum decoherence to satisfy massive data transmission requirements is an urgent issue in large-scale quantum networks. In this paper, we call it an Entanglement Routing and Qubit Provisioning (*ERQP*) problem.

Quantum network modeling is the first issue to be considered in solving the *ERQP* problem. Most of the existing model maximizes the network throughput by establishing E2E one-qubit entanglements as many as possible [11]–[15]. This kind of entanglement state merely transmits one qubit of data in a one-time entanglement transmission. The one-qubit entanglement state is susceptible to interference from noise induced by quantum decoherence. Recently, a type of multi-qubit entanglement brings a noticeable improvement in the resistance to the noise [16]–[18]. It is generated by the weighted summation of multiple qubit states between two

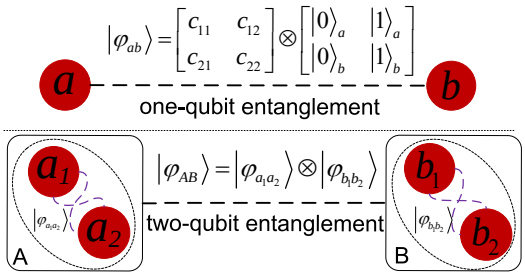


Fig. 1: One-qubit and two-qubit entanglements.

quantum nodes, and proven to be more noise-resistance [16]. We provide an example to demonstrate this entanglement scheme and its stronger correlations. As shown in Fig. 1, one-qubit entanglement $|\varphi_{ab}\rangle$ has two pairs of Hilbert coordinate bases $((|0\rangle_a, |1\rangle_a), (|0\rangle_b, |1\rangle_b))$ and is represented by a matrix with four coefficients. In contrast, the two-qubit entanglement $|\varphi_{AB}\rangle = |\varphi_{a_1a_2}\rangle \otimes |\varphi_{b_1b_2}\rangle$ has 4 pairs of Hilbert coordinate bases and its coefficient matrix has 2^4 elements. Part of the representing coefficients is modified when the entanglement state interferes with the decoherence noise. A larger representing matrix is more noise-resistant because more coefficients have better fault tolerance [16]. This high-dimensional multi-qubit entanglement scheme can implement more efficient (quantity and reliability) data transmission than the one-qubit entanglement. Therefore, establishing E2E multi-qubit entanglements to satisfy transmission requirements of massive requests is a valuable issue in large-scale quantum networks.

The second issue is the optimization methodology. Most of the previous studies optimize network utility by the centralized methods, such as extended Dijkstra [11], randomized rounding [12], opportunistic method [19], and self-made iteration [20]. Running these algorithms, with concurrent iterations and one-shot solving in large-scale networks, brings excessive communication expenditure caused by frequent information aggregation. Some works adopt entirely distributed methods to relieve the communication burden and achieve local privacy guarantee, such as graph-based distance routing [21], neighbor-based controlling [22]. These local-view algorithms damage utility performance because of the low input of valuable network information. There exists an objective difference between centralized and distributed control. The global objective serves quantum users in such a way that the requests can be successfully and quickly satisfied with the required data transmission, while the local objective focuses on the node benefits, such as the resource utilization ratio and energy consumption. Hence, designing a flexible transmission framework to balance the communication expenditure and the network performance with the joint optimization of global and local objectives is also an imperative issue in large-scale quantum networks.

In this paper, we propose an effective and flexible transmission framework (*Sophon*) in quantum networks with high-dimensional multi-qubit entanglement states. We resort to a high-dimensional quantum state (*W* state) to first formulate the multi-qubit quantum state and then quantum network components containing E2E entanglements, state decoherence, and data transmission process. Next, we define and model the *ERQP* problem as a joint optimization problem to globally maximize the network throughput and locally maximize the memory usage while constrained by manifold characteristics on the entanglement quality, the response delay, and the node capacity. Furthermore, we propose *Sophon*, which runs a centralized control platform in the quantum cloud and equips the Reinforcement Learning (RL) agent on each network node. The distributed RL agent takes the local information to learn the flexible qubit resource provisioning policies while maximizing memory usage. The centralized controller aims to optimize the global throughput by selecting reasonable E2E entanglement routes for each request using an improved link-level branch-and-bound algorithm. During the interaction process, the global network implements one-time entanglement transmission after obtaining the distributed resource provisioning decisions and the centralized route selection results. After that, the whole network environment completes a one-time global state transition. To the best of our knowledge, this is the first work studying qubit transmission based on the multi-qubit E2E entanglement. Our contributions are as follows:

- We use *W* quantum state to formulate multi-qubit entanglement and quantum network components, based on which we formulate the *ERQP* problem as a global-local optimization problem with manifold constraints.
- We propose a data transmission framework, *Sophon*, which adopts distributed, RL-based resource provision agents in each network node, cooperating with a centralized E2E entanglement routing solving by a link-level branch-and-bound algorithm.
- Extensive simulations reveal that (i) *Sophon* achieves a $41.1\%-69.5\%$ decrease in the communication cost and a $87.6\%-98.8\%$ decrease in the algorithm running time compared to the baselines and (ii) *Sophon* is robust and flexible on the heterogeneous requests, various topology scales, and different constraint thresholds.

The rest of this paper is organized as follows. Section II demonstrates the existing works, and Section III introduces some preliminaries of quantum networks. We present the network model and problem definition in Section IV, and elaborate *Sophon* in Section V. The evaluation results are shown and analyzed in Section VI. We conclude this work in Section VII.

II. Related Work

A. One-Qubit Entanglement Transmission

The entanglement distribution quantum network is first introduced with one-qubit entanglements that can transmit only one qubit of quantum information. Gyongyosi *et al.* [21] embed the entangled overlay network with the multi-level (different hop number) and probabilistic entangled links onto a simple graph-based network and then reduce the decentralized routing of finding the shortest route to a

statistical estimation problem. It faces a high complexity in responding to multiple SD pairs due to the only-one-request limitation. Shi *et al.* [11] respectively propose a pre-computed route selection and recovery assisted by the cost-ascending queue with the greedy-resource scheme, and a runtime extended Dijkstra's algorithm to greedily find the best one among all SD pairs' routes according to the cost metric. It doesn't consider the fairness of SD pairs and lacks the feasibility of link recovery. To remedy the weakness of [11], Zhao *et al.* [12], [15] merge the route recovery and E2E entanglement establishment to maximize the throughput of multiple SD pairs by firstly determining the optimal candidate entanglement links, link-based selecting the optimal route, and then, repairing failed routes using the remaining qualified links. Different from the fixed number of SD pairs in [12], [15], Zeng *et al.* [23] sequentially maximize the number of user pairs and the network throughput to make full use of network resources. The above works all regard the uncertainty of entanglement establishment caused by quantum decoherence as the homogeneous [21] or heterogeneous [11], [12], [15], [23] probabilities of entanglement links. Zhao *et al.* [24] select the primary links to maximize the throughput for multiple SD pairs, which provides a progressive fidelity qualification. Different from [24] that quantifies fidelity by using the bit flip error, Gu *et al.* [25] utilize the isotropic error model to formulate the entanglement quality. And, they consider maximizing the throughput with buffered entanglement reserve, the same as the overlay network applied by Pouryousef *et al.* [26]. Inspired by these works, our quantum network responds to multiple and flexible SD pairs in fairness, quantifies the entanglement quality with fidelity, and introduces the link-level buffering capability of network nodes.

B. Multi-Qubit Entanglement Transmission

The multi-qubit entanglement can concurrently transmit multiple qubit information between two network nodes, including two schemes of the multiple nodes connected as a line and as a graph topology. Zhong *et al.* [27] experimentally report a three-qubit entanglement in a quantum network comprising two superconducting nodes with the deterministic entanglement fidelity, which verifies the feasibility of multi-qubit entanglement between two network nodes. Mannalath *et al.* [28] extract maximally entangled GHZ states from a repeater line generated by finding the shortest route with the weights of the smallest combined neighborhood in the quantum network, where each network node contains only one qubit. The above two works belong to the node-line scheme, where each multi-qubit entanglement transmits multiple qubit information. BEGHELLI *et al.* [29] try to establish a long-distance high-dimensional entanglement state among multiple network nodes through entanglement fusion and BSM operations, which is utilized in the collaborative computing and encryption of multiple quantum parties. Similar to [29] that establishes GHZ state among several network nodes, Zeng *et al.* [30] model a multi-user entanglement as a graph-based degree-constrained minimum spanning tree problem to implement a more powerful quantum processor. These two works are the graph scheme, which transmits only one qubit via an entanglement route due to being inherently based on the one-qubit entanglement between any two nodes. Here, we explore the node-line scheme where the multi-qubit entanglement is established by the weighted summation of multiple one-qubit entanglements and formulated as the high-dimensional W state between two network nodes [16], [17].

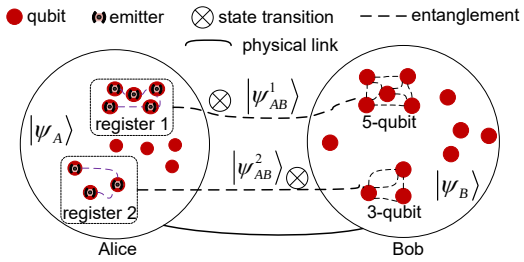


Fig. 2: Two multi-qubit entanglements established between two network nodes.

III. Preliminaries

A. Multi-Qubit Entanglement

The d-qubit entanglement can be constructed by d-dimensional entanglement gate (Clifford gate) group, formulating the entanglement into a weighted summation of one-qubit entanglement [18]. Based on [18], [17] introduces the quantum register to emit an array of qubits once, and establishes a multi-qubit entanglement formulated in the W state. For example, 3-dimensional W state is denoted as $|W\rangle = \frac{1}{\sqrt{3}}(|001\rangle + |010\rangle + |100\rangle)$.

The quantum register is a container accommodating an array of qubits and emitters in a quantum node. Each emitter emits a qubit particle once a time. The quantum state of the d-qubit system in a quantum register is described as:

$$|\varphi\rangle = \bigotimes_{j=0}^{d-1} (\sqrt{1-p}|0\rangle_j + \sqrt{p}|1\rangle_j) \, |\emptyset\rangle \,, \tag{1}$$

where p is the probability of each qubit measured in state $|1\rangle$. $\otimes$ denotes the state transition of dimension expansion. $|\emptyset\rangle$ denotes the vacuum of the electromagnetic field. When d emitters in a register concurrently emit d entangled particles, a high-dimensional entanglement state is attempted to be successfully established in the receiver. This step is repeated for f times until successful. The established state is:

$$|\psi\rangle_d^f = \frac{1}{\sqrt{d}} \sum_{j=0}^{d-1} |e_j\rangle \, |j\rangle^{\otimes f} \tag{2}$$

where f is the number of attempts, i.e., the emission times of a group of laser emitters.

As shown in Fig. 2, Alice has registers 1 and 2, respectively, accommodating 5 and 3 qubit emitters. After each register completes the particle emission, two multi-qubit entanglement states $|\psi_{AB}^1\rangle$ and $|\psi_{AB}^2\rangle$ are established. Bob receives 5 entangled particles in $|\psi_{AB}^1\rangle$ and 3 ones

in $\left|\psi_{AB}^2\right\rangle$. As shown in Fig. 3, Alice's register 2 emits entangled particles with $f = 4$-time attempts, during each of which 3 entangled particles are attempted to be emitted to Bob. There establishes a 3-qubit entanglement represented by the W state as $\left|\psi_{AB}^2\right\rangle_3^4 = \frac{1}{\sqrt{3}} \sum_{j=0}^{2} |e_j\rangle\,|j\rangle^{\otimes 4}$. This link-level entanglement $\left|\psi_{AB}^2\right\rangle_3^4$ can successfully transmit 3-qubit information from Alice to Bob once time.

Using the W state with the weighted overlapping of d-qubit entanglements to write the d-dimensional state transition $\otimes$ in Eq. (2):

$$|\psi\rangle_d^f = \sqrt{fd} \sum_{i=0}^{f-1} \sum_{j=0}^{d-1} |e_i\rangle\,|j\rangle. \tag{3}$$

and introducing Eq. (1), i.e., $|\psi\rangle = |\varphi\rangle$, we can obtain:

$$|\psi\rangle_d^f = \sqrt{fd}p(1-p)^{f\cdot d-1} \sum_{i=0}^{f-1} \sum_{j=0}^{d-1} |e_{ij}\rangle, \tag{4}$$

where $|e_{ij}\rangle$ is the unit state, the $(i \cdot d + j)$th qubit is in state $|1\rangle$ and all others are in state $|0\rangle$.

B. Quantum Decoherence

We focus on the collapsed decoherence model [9], [10], which sheds light on the behavior of quantum systems in realistic scenarios. It occurs when the masses of qubit particles in registers significantly exceed those of scattered surrounding particles. Considering a quantum entangled state $|\psi\rangle$, subjected to collapsed decoherence, it evolves with an exponential item as:

$$|\psi\rangle = |\psi\rangle\, e^{-\tau(\Delta x)^2}, \tag{5}$$

where the measured scattering constant τ can characterize the physical properties of the system-surrounding interaction. Δx is the physical distance between two entangled particles, within which the system maintains coherence. Jointly considering Eq. (4) and Eq. (5), we obtain the collapsed decoherence of d-dimensional qubit entanglement:

$$|\psi\rangle_d^f = \sum_{i=0}^{f-1} \sum_{j=0}^{d-1} \sqrt{fd}p(1-p)^{f\cdot d-1} e^{-\tau(\Delta x)^2} |e_{ij}\rangle. \tag{6}$$

Additionally, fidelity, valued in $(0, 1)$, is an index to evaluate the entanglement quality. We utilize the Schumacher fidelity [31], which extends the state similarity to the maximal transition probability. Decoherence weakens the entanglement quality while decreasing the fidelity. The Schumacher fidelity of multi-qubit entanglement $|\psi\rangle_d^f$ under the collapse decoherence [9], [10] is denoted as:

$$F(|\psi\rangle_d^f) = \sqrt{fd}p(1-p)^{f\cdot d-1} e^{-\tau(\Delta x)^2}. \tag{7}$$

which indicates that the emission times f and the transmission dimension d contribute equally to $F(|\psi\rangle_d^f)$.

When the entanglement state cannot be quantified in the practical scenario, the physical system guarantees the successful entanglement establishment through f-times emission attempts. Li *et al.* [32] design a hardware system for entanglement quantification, which can successfully establish the entanglement state only through 2-times emission attempts. Here, we employ a fidelity metric derived from the collapse decoherence model to assess high-dimensional entanglement quality, rather than relying on system-level statistical methods (emission attempts). We set $f = 1$, i.e., a single emission event per emitter. This simplification allows the analysis to focus exclusively on how the entanglement quality is determined by its dimensionality.

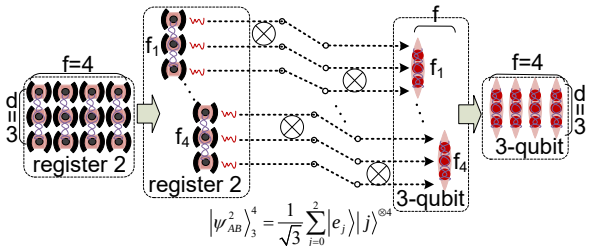


Fig. 3: A multi-qubit entanglement established by the laser emitter group and represented in *W* state.

C. Quantum Network Components

Quantum Requests: The quantum request originates from either classical or quantum applications, requiring a certain amount of qubit information transmitted via a quantum network. It contains a source node to send data and a destination node to receive data.

Quantum Repeaters: The quantum repeater is required to forward the quantum information along the E2E entanglement route. Each quantum repeater should provide two types of resources: storage resources and computing resources.

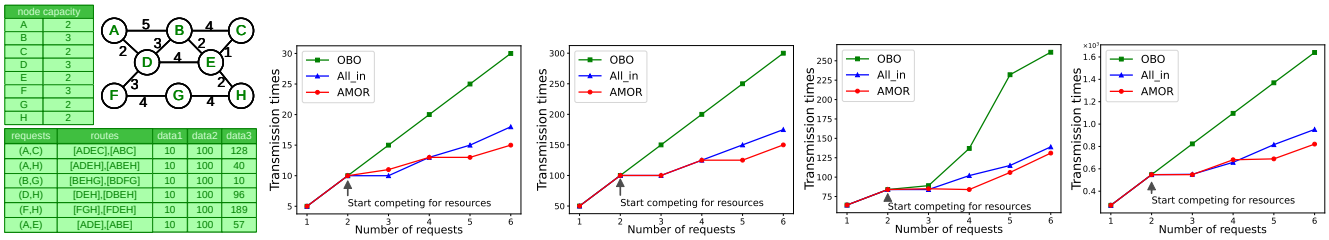
Quantum Memory: Each network node's memory is required to temporarily maintain the multi-qubit information for a time-bin interval in one-time transmission. The occupied memory is released until the request is satisfied with its transmission requirement.

Quantum Cloud Platform: It is a cluster of many high-performance quantum computing processors. A centralized controller runs on a common quantum server employed in the quantum cloud platform. It has a global view of the real-time network environment with the comprehensive authority to manage network resources.

IV. Quantum Network Model

A. Motivating Example

Before formulating the problem, we develop an example of the *ERQP* problem to demonstrate the innovation of our transmission framework. Fig. 4(a) shows the example network topology consisting of 8 quantum nodes and 10 edges with their physical length. The node memories and the request set are also provided in the chart. Each request has the top two shortest candidate routes generated between its source and destination nodes. The request set has four types of data requirements: homogeneous requirements (10 and 100 qubits), and non-homogeneous requirements (a given requirement set and a sampled requirement set).



(a) The given problem (b) Transmission times with data requirement $data1$ (c) Transmission times with data requirement $data2$ (d) Transmission times with data requirement $data3$ (e) Transmission times with data requirement $data4$

Fig. 4: A motivation example on three transmission schemes (OBO, All_in, AMOR) with different patterns of data requirements $data1, data2, data3$, whose values are taken from Fig. 4(a). $data4$ is sampled from a Gaussian distribution $(55, 1)$ and clipped to the range $[10, 100]$. The result in Fig. 4(e) is the average of 10 runs.

TABLE I: Example problem models

Framework	Variables	Problem models
OBO [11], [13], [33]	$\chi_k^{r,t}$	$obj: max \sum_{k\in\mathcal{P}_r} \chi_k^{r,t}, \forall r \in R$ $s.t.: \sum_{r\in R}\sum_{k\in\mathcal{P}_r} H_{k,n}^r \chi_k^{r,t} \le \mathcal{C}_n$
All_in [12], [24]	$\chi_k^{r,t}$	$obj: max \sum_{r\in R}\sum_{k\in\mathcal{P}_r} \chi_k^{r,t}$ $s.t.: \sum_{r\in R}\sum_{k\in\mathcal{P}_r} H_{k,n}^r \chi_k^{r,t} \le \mathcal{C}_n$
AMOR [14], [23]	$\mathbb{X}_k^{r,t}$, $Y_k^{r,t}$	$obj: max \sum_{r\in R}\sum_{k\in\mathcal{P}_r} \mathbb{X}_k^{r,t}$ $s.t.: \sum_{r\in R}\sum_{k\in\mathcal{P}_r} H_{k,n}^r \mathbb{X}_k^{r,t} \le \mathcal{C}_n$ $\sum_{k\in\mathcal{P}_r} Y_k^{r,t} \le 1, \forall r \in R$

We regard one-time entanglement transmission as a network time slot where all network resources are considered in the provisioning decision for a request set. Based on this, we explore three representative transmission schemes: One-By-One (*OBO*) where the overall network resource serves only one request in one-time transmission, *All_in*, where all requests can be simultaneously responded to in the network, and each request is served by multiple entanglement routes with the multi-qubit transmission, At Most One Route (*AMOR*) where all requests can be simultaneously responded to in the network, and each request is guaranteed to be responded through at most one entanglement route with the multi-qubit transmission. In these schemes, when one-time transmission cannot complete the data transmission requirement of the request set, the residual transmission tasks are executed after the previously used resources are released. *All_in* is applied in [12], [24], allowing one entanglement route to transmit merely two rather than multiple qubit information. *AMOR* is proposed in [14], [23] regarding an entanglement route as a whole that transmits the same number of qubits in all entanglement links through this route. Compared to *All_in*, *AMOR* guarantees the transmission fairness for each request.

Table I lists the problem models of these three transmission frameworks. $\chi_k^{r,t}$ are integer variables denoting how many two-qubit E2E entanglements are established along the k-th route of request r in the t-th transmission. $H_{k,n}^r$ are binary constants to represent whether request r's k-th route traverses node n (1) or not (0). $\mathbb{X}_k^{r,t}$ are integer variables depicting the number of qubits that request r's k-th route can transmit in the t-th transmission. $Y_k^{r,t}$ are binary variables, denoting whether request r selects the k-th route to establish E2E entanglement in the t-th transmission. $Y_k^{r,t}$ can be deduced from $\mathbb{X}_k^{r,t}$. All models aim to maximize the total transmission data volume while being limited by node memories. The optimal solutions to these problems are obtained by using the global traversal. We set $\chi_k^{r,t} \in \{0,1,2\}, \mathbb{X}_k^{r,t} \in \{0,1,2\}$. $Y_k^{r,t} = 1$ if $\mathbb{X}_k^{r,t} > 0$, otherwise $Y_k^{r,t} = 0$.

Shown in Fig. 4, three transmission schemes respond to $data1, data2, data3$ transmission requirements (shown in Fig. 4(a)), respectively. Each scheme is completed with a different number of transmissions. It is evident that *OBO* requires an increasing number of transmission times as requests increase, especially with exponential increases when the data requirements are non-homogeneous. This is because the data requirements of partial requests cannot be completed in one-time transmission, while others must wait. Network resources cannot be fully utilized in one-time transmission, even as the number of requests or the data volume requirement continues to increase. *All_in* and *AMOR* both perform better than *OBO* when the network resources are not enough to support all transmission tasks. With the network resources becoming severely scarce, *AMOR* maintains its performance advantage under homogeneous data demand conditions in Fig. 4(b) and Fig. 4(c). With non-homogeneous data requirements, *AMOR* still performs better with fewer transmission times when the number of requests increases, as shown in Fig. 4(d). It is because *AMOR* is more sensitive to resource contention than *All_in* and *AMOR* ensures the possibility that each request can be responded to fairly in one-time transmission, which brings a long-term benefit.

B. Problem Definition

In the given example, we assume the k-th candidate route of request r can transmit $\mathbb{X}_k^{r,t}$ qubits in the t-th transmission, i.e., each node in the route $\mathcal{P}_r^k$ allocates the same $\mathbb{X}_k^{r,t}$ qubits for r. In the *ERQP* problem, local quantum node n possesses the decision-making ability to allocate $X_n^{r,t}$ qubit resources

TABLE II: Notations

Parameters	Descriptions
$\mathcal{G}(\mathcal{N},\mathcal{E})$	The quantum network topology
R	The given request set
$r(s_r,d_r,V_r)$	Request r with its node pair $< s_r,d_r >$ and the data transmission requirement V_r
$\mathcal{P}_r$	The given physical routes of request r
$\mathcal{F}_{thr}$	The fidelity threshold for all requests
$\mathcal{D}_{thr}$	The delay threshold for each E2E entanglement
$\mathcal{C}_m$	The capacity of node n
$H^r_{k,n}$	A boolean array to denote whether request r's the kth route traverses node n (1) or not (0)
$M^{r,t}_n$	The number of resources that stores request r's qubit information on node n in the tth transmission
$\Phi^{r,t}_{n,n+1,k}$	The transmitted data volume of entanglement link $l_{n,n+1}$ on the kth route of request r
Variables	Descriptions
$Y^{r,t}_k$	Boolean variable and it is 1 if the physical route P_k is selected to generate the tth E2E entanglement for request r
$X^{r,t}_n$	Integer variable, which indicates how many qubits node n allocates for request r in the tth transmission

for r and considers the local optimization objective. The transmission data volumes of request r on each entanglement link should be flexible and vary with the node capacity. Therefore, we extend the *AMOR* scheme on the link-level entanglement link by depicting the qubit transmission on each entanglement link. The link-level *AMOR* is more fine-grained in the routing and resource provisioning decisions compared to the route-level *AMOR* [14], [23]. We first formulate some link-level transmission components related to the variables. Table II provides the notations.

1) Link-Level Entanglement Transmission: We extend *AMOR* by using a link-level data transmission to adapt each network node's local decision-making ability. A network node can temporarily maintain qubit resources to wait to be transmitted to the next node. In the t-th transmission, node n maintains $M^{r,t}_n$ qubit resources for request r and pre-allocates $X^{r,t}_n$ qubit resources for the k-th route of r to be prepared for transmitting. The transmitting data volume of the entanglement link $l_{n,n+1}$ on route $\mathcal{P}^k_r$ is denoted by:

$$\Phi^{r,t}_{n,n+1,k} = min(X^{r,t}_n, M^{r,t}_n + \Phi^{r,t}_{n-1,n,k}). \tag{8}$$

where $n-1$ and $n+1$ are the predecessor and successor of node n. The accumulated data volume of request r on node d_r in the t-th transmission is:

$$\Phi^{r,t}_{*} = M^{r,t}_{d_r} + \Phi^{r,t}_{d_r-1,d_r,k}. \tag{9}$$

Once the data requirement of request r is successfully satisfied, i.e., $\Phi^{r,t}_{*} \geq V_r$, the occupied qubit resources in all network nodes are released to prepare for the data transmission of other requests.

2) E2E Entanglement Fidelity: The E2E entanglement fidelity can be obtained by the product of multi-qubit entanglement links on the E2E entanglement route [24]. The E2E fidelity of route $P^{k,t}_r$ is computed by:

$$\begin{aligned} F_{P^{k,t}_r} &= \prod_{n\in P^{k,t}_r} F(\left|\psi_{n(n+1)}\right\rangle^f_d) \\ &= \prod_{n\in P^{k,t}_r} \sqrt{\Phi^{r,t}_{n,n+1,k}}p(1-p)^{\Phi^{r,t}_{n,n+1,k}-1}. \end{aligned} \tag{10}$$

Combined with the property of quantum decoherence in Eq. (7), Eq. (10) is transformed as:

$$F_{P^{r,t}_k} = \prod_{n\in P^{r,t}_k} \sqrt{\Phi^{r,t}_{n,n+1,k}}p(1-p)^{\Phi^{r,t}_{n,n+1,k}-1}P^k_r(e), \tag{11}$$

where $P^k_r(e) = e^{-\tau|P^{r,t}_k|^2}$. We deduce from Eq. (11) that the transmitting data volume on entanglement links and the length of the E2E entanglement route are critical factors to quantify E2E entanglements indicated by E2E fidelity.

3) Transmission Delay: The transmission delay of request r in the t-th transmission is denoted as:

$$D^t_r = \sum_{k\in\mathcal{P}_r}\sum_{i,j\in V} Y^{r,t}_k L^r_k + D^{t-1}_r, \tag{12}$$

where D^{t-1}_r is the previously accumulated delay until the $t-1$-th transmission, recorded during the transmission process.

4) Problem Definition: Suppose there is a given undirected graph to represent a quantum network denoted by $\mathcal{G}(\mathcal{N},\mathcal{E})$ and a set of requests R. $\mathcal{N}$ and $\mathcal{E}$ are respectively node set and edge set. Quantum node $n \in \mathcal{N}$ has its memory capacity of $\mathcal{C}_n$. Request $r \in R$ is characterized by $r =< s_r,d_r,V_r >. < s_r,d_r >, s_r,d_r \in \mathcal{N}$ denotes the source-destination node pair. V_r is the data transmission requirement that necessitates being served by the quantum network. Here, we assume the route set $\mathcal{P}_r$ between s_r and d_r node pair is known [1] in advance. $\mathcal{D}_{thr}$ is the delay threshold for each request, and $\mathcal{F}_{thr}$ is the fidelity threshold for E2E entanglements, both of which are determined by the properties of quantum applications. We define the Entanglement Routing and Qubit Provisioning (*ERQP*) problem with high-dimensional entanglement transmission:

Definition 1. *Given a network $\mathcal{G}(\mathcal{N},\mathcal{E})$ and a set of requests R, the* ERQP *problem*[2]*, involving routing E2E multi-qubit entanglements and allocating qubit resources for a request set R, with the cooperative objectives of globally maximizing the transmitting data volume while locally maximizing the network node's resource utilization, such that:*

- *The fidelity and delay of each E2E entanglement are no less than the given thresholds $\mathcal{F}_{thr}$ and $\mathcal{D}_{thr}$ in one-time entanglement transmission.*
- *The total resource usage for all requests must not exceed the memory capacity in any quantum node.*

[1] In the given route set $\mathcal{P}$, some routes connecting the node pair $< s,d >$ are sorted in order by physical length.

[2] The ERQP problem is essentially a network resource allocation problem, while it has the specific characteristic of quantum mechanism (fidelity and the classical communication delay after being measured). Also, its objectives can be changed according to the requirements of quantum applications due to the joint optimization setup in our proposed framework.

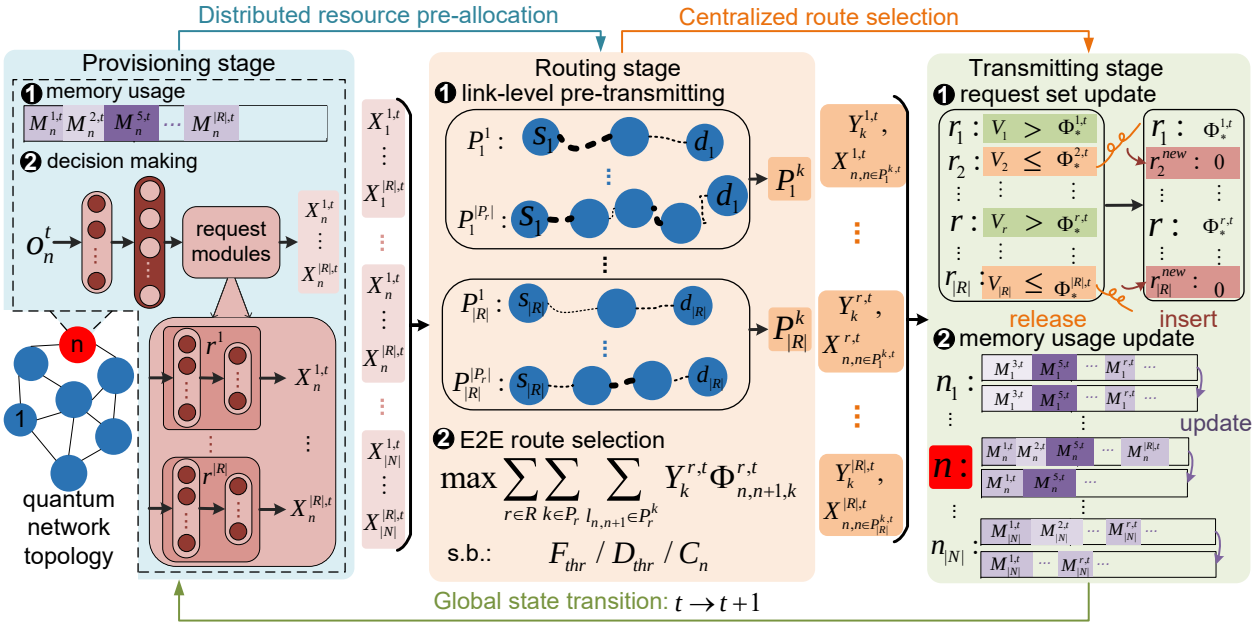


Fig. 5: Sophon's working framework.

V. Sophon Design

A. Overview

Sophon is an online link-level transmission framework for quantum networks, which uses distributed RL agents to dynamically adjust qubit resource provisioning based on the real-time data transmission requirements computed by the centralized controller while balancing the global-local optimization objectives. It is highly challenging to directly solve the defined *ERQP* problem in Def. 1 because (1) treating joint objectives as an integrated entity neglects their individual characteristics, (2) the large number of variables and constraints leads to an exponential increase in computing complexity. *Sophon* decomposes the solving process into three stages: provisioning, routing, and transmitting.

As shown in Fig. 5, each network node is equipped with an RL agent. An RL agent's one-time state transition consists of three stages. In the provisioning stage, each RL agent pre-allocates qubit resources for all requests based on the local-view network information o_n^t. In the routing stage, *Sophon* uses the resource provisioning policies $X_n^{r,t}$ output by all distributed RL agents to compute the link-level data transmission according to Eq. (8). According to the edge weight parameters calculated based on the link-level transmitting data volume, *Sophon* selects the optimal candidate route to establish E2E entanglement to maximize the total transmitting data volume on the centralized control platform. The centralized controller outputs the selected routes $Y_k^{r,t}$ and the determined provisioning policies $X_n^{r,t}$ for each request. In the transmitting stage, data transmission occurs based on the known route and resource decisions. The request set R is updated to release the satisfied requests and insert new ones by comparing the required data V_r and the real-time completed data volume $\chi_*^{r,t}$ in r's destination node d_r. Each node's memory is updated. The occupied qubit resources of the satisfied requests are released to make resources available for new requests. The quantum network environment completes a global state transition from the t-th to the $(t+1)$-th transmission. Each network node goes into the provisioning stage of the next time slot. *Sophon* extends the classical Markov interaction between RL agents and the network environment by introducing the global E2E route selection.

B. Provisioning Stage

In the provisioning stage, each network node aims to cooperatively complete the transmission tasks with its neighbors while sufficiently utilizing individual resources. Each RL agent in the node inputs the local-observed network information into the policy network to obtain actions and compute the local reward. The policy network is trained by the step-to-step accumulated interaction experience with the network environment. The quantum network with all trained agents self-adaptively transmits the qubit data to update the request set. We take the Deep Q-learning Network (DQN) as an example of RL algorithms [3] to learn the distributed pre-provision policies. DQN contains the policy network, which learns the mapping between the observations and the Q-value, and the target network, which is used to stabilize the learning process with a periodic parameter copy from the policy network. DQN agent n is updated by:

$$Q(o_n^t, a_n^t) \xleftarrow{\nabla} \alpha(rd_n^t + \gamma max_a Q(o_n^{t+1}, a_n^{t+1}) - Q(o_n^t, a_n^t)), \quad (13)$$

where α is the step size. γ is the discounted reward factor, rd^t is the local reward of agent n in the t-th interaction. $max_a Q$ is the maximum Q-value selected from the output of the target network based on the observations o_n^{t+1}.

[3] Sophon integrates the centralized route selection into the Markov Decision Process, so each type of RL algorithm can be adapted here.

1) RL Details: We design the RL details as follows:

- *Episode and Step:* An episode refers to a human-made data transmission interval of the request set $\mathcal{R}$ starting from the initial network state during the training process. R is continuously updated with completed requests, so the episode length is limited by the given T, i.e., there are a total of T-time transmissions in an episode. One-time transmission is an RL interaction between the agent and the network environment consisting of three stages.
- *Local Observations:* The local observations o_n of agent n contain the candidate route information, memory usage, memory remaining, and the neighbor information. The candidate route information is denoted by $H^r_{k,n}$, which is 1 if the k-th candidate route of request r traverses node n, otherwise 0. The memory usage is $M^{r,t}_n, r \in R$ and the memory remaining is $\mathcal{C}_n - \sum_{r \in R} M^{r,t}_n$. The neighbor information is denoted by $I_{n,n'}$, which is 1 if there is a physical link between node n and node n'. The length of agent n's local observations is $|R|(|\mathcal{P}_r|+1)+|\mathcal{N}|+1$.
- *Action:* The action a_n of agent n is $X^{r,t}_n$, which denotes how many qubit resources agent n pre-allocates for request r. We set X_{thr} as the upper bound of variables $X^{r,t}_n$ because the transmission ability of the entanglement link is limited by its physical properties. To transmit data for multiple requests concurrently, we design "request modules" to individually determine the action $X^{r,t}_n$ for each request. The action space size of each agent is $|R|X_{thr}$.
- *Reward:* The local reward rd^t_n of agent n contains the memory usage, the contribution to the requests, and the payment given by the global controller:

$$rd^t_n = \sum_{r \in \mathcal{R}} Y^{r,t}_k H^r_{k,n}(M^{r,t}_n + X^{r,t}_n) + py^t_n, \tag{14}$$

where py^t_n is the global payment, which denotes the number of completed requests. $Y^{r,t}_k$ and $M^{r,t}_n$ are limited by the constraints in the centralized routing stage and impact the process of memory pre-provisioning in the t-th entanglement transmission of the distributed agents.

2) An Improved Policy Network: To concurrently provision resources for multiple requests, we design an improved policy network architecture that divides the output layer of the policy network into $|R|$ request modules. Each request module has two full connection layers and individually determines its provisioning policies $X^{r,t}_n$. The improved policy network is trained using the following process.

Given is a mini-batch of tuples (O, A, RD, O') with the size of bs sampled from the experience pool, where $O = \{o_i\}^{\mathcal{N}}_{i=1}$ and $A = \{a_i\}^{\mathcal{N}}_{i=1}$. For agent n, the policy network contains two parts: $\mathcal{F}^{bf,n}_{eval}$ with the parameter $\theta^{bf,n}_{eval}$ which are the common neural layers and the "request modules" $\mathcal{F}^{r,n}_{eval}$. The output of $\mathcal{F}^{bf,n}_{eval}$ is $\mathcal{F}^{bf,n}_{eval,t}(o^t_n; \theta^{bf,n}_{eval,t})$. The output of the request module r is:

$$\mathcal{F}^{r,n}_{eval,t}(\frac{1}{|R|}\mathcal{F}^{bf,n}_{eval,t}(o^t_n; \theta^{bf,n}_{eval,t}); \theta^{r,n}_{eval,t}), r \in R. \tag{15}$$

Therefore, the outputs of the policy (evaluation) network in the t-th transmission are denoted as:

$$eval\_outs_t = concat(\mathcal{F}^{r,n}_{eval,t}(o^t_n; \theta^n_{eval,t})), r \in R. \tag{16}$$

$concat$ is the concatenation operation. Q-values of policy networks are deduced as:

$$Q^t_{eval} = concat(\mathcal{F}^{r,n}_{eval,t}(o^t_n, a^t_n; \theta^{r,n}_{eval,t}))^{\mathcal{N}}_{n=1} \tag{17}$$

where $a^t_n \in A$ is the historical action for the observations o^t_n, which is used to choose the Q^t_{eval}. Considering Eq. (17) and Eq. (13), all DQN agents are trained concurrently with the mini-batch experience.

C. Routing Stage

The routing stage aims to select the optimal route from the candidate route set for E2E entanglement establishment based on the pre-allocated resource policies $X^{r,t}_n$ obtained in the provisioning stage. With the given $X^{r,t}_n$, the link-level data transmission $\chi^{r,t}_{n,n+1,k}$ is computed by Eq. (8). With the different transmission abilities of entanglement links, we formulate an exact Integer Linear Programming (ILP) to maximize the total transmission data volume to select the optimal E2E entanglement routes for the request set R, constrained by the E2E entanglement quality, the transmission delay, and the node capacity.

1) Problem Formulation: The E2E route selection is formulated as follows:

$$obj: \max \sum_{r \in R} \sum_{k \in P_r} \sum_{n \in P^k_r} Y^{r,t}_k H^{r,t}_{k,n} \Phi^{r,t}_{n,n+1,k}, \forall t \in T \tag{18}$$

$$s.t.: Y^{r,t}_k \prod_{n \in P^k_r} \sqrt{\Phi^{r,t}_{n,n+1,k}} H^{r,t}_{k,n} p (1-p)^{\Phi^{r,t}_{n,n+1,k}-1} P^k_r(e) \geq \mathcal{F}_{thr}, \forall r \in R, \forall t \in T \tag{19}$$

$$\sum_{k \in \mathcal{P}_r} Y^{r,t}_k L^r_k + D^{t-1}_r \leq \mathcal{D}_{thr}, \forall r \in R, \forall t \in T \tag{20}$$

$$\sum_{r \in R} \sum_{k \in P_r} Y^{r,t}_k H^{r,t}_{k,n}(M^{r,t}_n + \Phi^{r,t}_{n-1,n} - \Phi^{r,t}_{n,n+1}) \leq \mathcal{C}_n, \forall n \in P^k_r, \forall t \in T \tag{21}$$

$$\sum_{k \in \mathcal{P}_r} Y^{r,t}_k \leq 1, \forall r \in R, \forall t \in T \tag{22}$$

The objective is to maximize the total transmitting data volume in the t-th transmission. Eq. (19) ensures the fidelity of each E2E entanglement is larger than the threshold $\mathcal{F}_{thr}$. Eq. (20) imposes that the transmission delay of each request r must be less than the threshold $\mathcal{D}_{thr}$. Eq. (21) states the node capacity limitation. Eq. (22) ensures the t-th E2E entanglement of each request to be established by at most one physical route.

We solve the above ILP with polynomial time complexity by relaxing the integer variables $Y^{r,t}_k$ to the fractional variables $\widetilde{Y}^{r,t}_k \in [0,1]$. By using the Linear Programming (LP) solver (an interior-point method [34]), we obtain fractional solutions $\widetilde{Y}^{r,t}_k$ that can satisfy the fidelity, delay, and memory constraints while maximizing the global transmission volume of data. Then, we improve the branch-and-bound

algorithm, which is adopted in [14], [35] to select the approximately optimal E2E entanglement routes by introducing *the local-objective filter* and *the multiple-constraint filter* in each branch.

Algorithm 1 Obtain the Integer Solution

Input: Current request r, $X_n^{r,t}$, sig;
Output: Integer solution $Y_k^{\dagger,r,t}$, sig;
1: Compute the transmission parameters $\Phi_{n,n+1,k}^{r,t}$;
2: Initialize the queue $Y = \emptyset$;
3: $\widetilde{Y}_k^{r,t}$, obj = LP_solver($\Phi_{n,n+1,k}^{r,t}$, sig);
4: **if** $obj_{glocal} > obj + \sum_{r\in sig[r]>0} \sum_{k\in P_r} \sum_{n\in P_r^k} \Phi_{n,n+1,k}^{r,t}$ **then**
5: $sig[r] = -1$;
6: $return\ Y_k^{\dagger,r,t}, sig$;
7: **end if**
8: **for** $\widetilde{Y}_k^{r,t}, r \in R$ **do**
9: **if** $\widetilde{Y}_k^{r,t} = 1$ **then**
10: $Y_k^{\dagger,r,t} = 1$, $sig[r] = k$;
11: **else**
12: $\widetilde{Y}_k^{\dagger,r,t} \to Y$;
13: **end if**
14: **end for**
15: **if** $len(Y) = 0$ **then**
16: $return\ Y_k^{\dagger,r,t}, sig$;
17: **else**
18: Sort Y in descending order;
19: Find the maximum $\widetilde{Y}_k^{r,t} \in Y, \forall r \in R$ that $sig[r] = 0$;
20: *Link-Level Branch-and-Bound* $(r, None, sig)$;
21: **end if**

2) Link-Level Branch-and-Bound Algorithm: Algorithm 1 is the entrance of the E2E route selection algorithm. Lines 3-4 are the process of *the local objective filter* where the new branch is cut if its optimal objective $obj + \sum_{r\in sig[r]>0} \sum_{k\in P_r} \sum_{n\in P_r^k} \Phi_{n,n+1,k}^{r,t}$ is smaller than the recorded global objective obj_{global}. Then, part of integer solutions $Y_k^{\dagger,r,t}$ and its signals $sig[r]$ are determined for the cases where $\widetilde{Y}_n^{r,t} = 1$, while the remaining $\widetilde{Y}_n^{r,t}$ are appended into the queue Y in Lines 5-9. The Link-Level Branch-and-Bound Algorithm executes with the maximum $\widetilde{Y}_n^{r,t}$ selected from the queue Y sorted in descending order. Initially, the selected route signals $sig[r]$ are set to be $-1, \forall r \in R$, where $sig[r] = 0$ denotes no route selection and $sig[r] = k$ denotes the selection of the k-th route. obj_{global} is set to be 0.

In Algorithm 2, all candidate routes containing the sub-route P are first filtered by *the multiple constraints* in Lines 2-4. It indicates that if the entanglement quality, the transmission delay, and the traversed node memory of the sub-route P meet all the requirements, the candidate routes enter into the next branch. Lines 5-9 describe the determination of $Y_k^{\dagger,r,t}$ under the condition where the number of qualified candidate routes $\widetilde{P}_r$ is smaller than 1. Then, lines 10-12 update the globally recorded objective obj_{global} based on the given $X_n^{r,t}$ and the determined sig. This branching node is a new request that goes into Line 13.

Algorithm 2 Link-Level Branch-and-Bound Algorithm

Input: Current request r, sub-route P, $X_n^{r,t}$, sig;
1: Initialize feasible routes $\widetilde{P}_r$;
2: **for** $P' \in P_r$ **do**
3: **if** $P \subseteq P'$ and P is feasible on Constraints Eq. (19), Eq. (20), Eq. (21) **then**
4: $\widetilde{P}$.append(P');
5: **end if**
6: **end for**
7: **if** $|\widetilde{P}_r| \leq 1$ **then**
8: **if** $|\widetilde{P}_r| = 1$ **then**
9: $sig[r] = \widetilde{P}_r[0]$;
10: **else**
11: $sig[r] = 0$;
12: **end if**
13: **if** $sig[r] > -1, \forall r \in R$ **then**
14: Compute $\widetilde{\Phi}'$ based on sig and $X_n^{r,t}$;
15: $obj_{global} = \widetilde{\Phi}'$ if $obj_{global} < \widetilde{\Phi}'$;
16: **end if**
17: *Obtain the Integer Solution* $(r, X_n^{r,t}, sig)$;
18: **else**
19: Find the minimum i s.t. exist two routes $P'(n_i), P''(n_i') \in \widetilde{P}_r$, satisfies $n_i \neq n_i'$;
20: **for** $P' \in \widetilde{P}_r$ **do**
21: $\widetilde{P} = P$.append(n_{i+1}), $n_i, n_{i+1} \in P'$;
22: *Link-Level Branch-and-Bound* $(r, \widetilde{P}, X_n^{r,t}, sig)$;
23: **end for**
24: **end if**

Lines 15-18 demonstrate the branching process with at least two qualified candidate routes. After finding the minimum overlap node n_i among all feasible routes, the current sub-route P is extended by individually adding the next node $n_{i+1}, P'(n), P'(n_{i+1}) \in \widetilde{P}_r$ to recursively go into Link-Level Branch-and-Bound.

3) Performance Analysis: Algorithm 1 is not only the entry point of solving the problem but also manages the inter-dependent, concurrent calls to Alg. 2. We regard Algorithms 1 and 2 as an entity for analysis.

Theorem 1. *The link-level branch-and-bound algorithm is an approximation algorithm to solve* ERQP*'s routing problem based on the known $X_n^{r,t}$ and the candidate routes $\mathcal{P}$. It achieves the performance approximation of 1.5.*

Proof. Let $\hat{\Phi}_{n,n+1,k}^{r,t} = -\Phi_{n,n+1,k}^{r,t}$, the routing problem is transformed into a Variant of the Integer Multi-commodity Flow problem (VIMF), formulated as the following:

$$obj: \min \sum_{r\in R} \sum_{k\in P_r} \sum_{n\in P_r^k} Y_k^{r,t} H_{k,n}^{r,t} \hat{\Phi}_{n,n+1,k}^{r,t} \tag{23}$$

$$s.t.: Y_k^{r,t} \prod_{n\in P_r^k} \sqrt{\hat{\Phi}_{n,n+1,k}^{r,t}} \frac{H_{k,n}^{r,t} p P_r^k(e)}{(1-p)^{\hat{\Phi}_{n,n+1,k}^{r,t}+1}} \leq \mathcal{F}_{thr}, \tag{24}$$

$$\sum_{k\in\mathcal{P}_r} Y_k^{r,t} L_k^r \leq \mathcal{D}_{thr} - D_r^{t-1} \tag{25}$$

$$\sum_{r\in R}\sum_{k\in P_r} Y_k^{r,t} H_{k,n}^{r,t}(M_n^{r,t}+\hat{\Phi}_{n,n+1}^{r,t}-\hat{\Phi}_{n-1,n}^{r,t}) \le \mathcal{C}_n \tag{26}$$

$$\sum_{k\in\mathcal{P}_r} Y_k^{r,t} \le 1 \tag{27}$$

By using the column-generation approach to solve VIMF, only a small subset of all columns is in the optimal solution, which indicates that all columns with positive reduced cost can be ignored, i,e., the problem is solved if the reduced costs of all columns are positive.

Let $\vartheta_1^{r,t}$, $\vartheta_2^{r,t}$, $\vartheta_3^{r,t}$ be respectively the unrestricted dual variables associated with constraints (24), (25), (27), and $-\pi_{n,n+1}^{r,t}$ be the non-negative dual variable associated with constraints (26). Then, the reduced cost of the column k for request r in the t-th transmission is:

$$\begin{aligned}\hat{\Lambda}_k^{r,t} = &\sum_{n\in P_r^k}(\hat{\Phi}_{n,n+1,k}^{r,t}+C_{n,n+1}^{r,t}\pi_{n,n+1}^{r,t})H_{k,n}^{r,t} \\ &-\sum_{i=1,2,3}\vartheta_i^{r,t}, \forall r\in R, t\in T.\end{aligned} \tag{28}$$

where $C_{n,n+1}^{r,t}=(M_n^{r,t}+\hat{\Phi}_{n,n+1}^{r,t}-\hat{\Phi}_{n-1,n}^{r,t})$. Each column is identified by solving the shortest path problem for each request with the edge weight $(\hat{\Phi}_{n,n+1,k}^{r,t}+\hat{C}_n^t\pi_{n,n+1}^t)$. The reduced cost of the shortest path is $\hat{\Lambda}_{k^*}^{r,t}$. If for all requests, $\hat{\Lambda}_{k^*}^{r,t}\ge 0$, the VIMF problem is solved. Otherwise, the shortest path k^* is added to VIMF's master problem. Hence, our VIMF problem conforms to the branching rules in [35].

For the integer solution, the column must be generated on the node of the branch-and-bound tree. When an optimal solution to an LP relaxation is infeasible to the ILP, the lifted cover inequalities are considered to cut off the infeasible solution. VIMF has a simple 0-1 knapsack inequalities (26) to deduce the link-level lifted cover inequality:

$$\sum_{r\in R^{cov}}\sum_{k\in P_r} Y_k^{r,t}H_{k,n}^{r,t}+\sum_{r\in\bar{R}^{cov}}\delta_r\sum_{k\in P_r}Y_k^{r,t}H_{k,n}^{r,t}\le |R^{cov}|-1, \tag{29}$$

which cuts off the infeasible solutions for the link-level adjustment and conforms to the cutting rules in [35]. R^{cov} is the minimal cover of requests for the link $(n,n+1)$, and $\bar{R}^{cov}=R\backslash R^{cov}$. Moreover, *the local-objective filter* introduces the link-level inequalities $obj_{global}\le obj+obj_{determined}$ and *the multiple-constraint filter* judges the feasibility of sub-route, both to cut off infeasible branches.

Consequently, our link-level branch-and-bound algorithm can achieve $f^*(Y^*)\le f(\widetilde{Y})=f(Y^\dagger)+\epsilon$, where ϵ is the remaining difference between the integer solution and the optimal LP solution. We assume $\epsilon=50\%f(Y^\dagger)$, which is a strong relaxation in almost all cases, except for the worst performance of $\epsilon=3.63\%f(Y^\dagger)$ due to the termination caused by a time out. Thus, we obtain $f^*(Y^*)=1.5f(Y^\dagger)$. □

D. Transmitting stage

Actual link-level data transmission and memory updating occur in the transmitting stage. Distributed agents individually maximize local rewards to pre-allocate the qubit resources $X_n^{r,t}$ for all requests. In the routing stage, the centralized controller maximizes the total data transmission volume to select the E2E entanglement routes $Y_k^{r,t}$ for all requests based on the transmission parameters $\Phi_{n,n+1,k}^{r,t}$. With the determined $Y_k^{r,t}$ and $X_n^{r,t}$, the t-th transmission executes to update memory usages to complete the global state transition. The memory updating of each node is:

$$M_n^{r,t+1}=\begin{cases}\widetilde{M}_n^{r,t+1}, & M_{d_r}^{r,t}<V_r\\ 0, & M_{d_r}^{r,t}\ge V_r\end{cases}, \tag{30}$$

where

$$\widetilde{M}_n^{r,t+1}=M_n^{r,t}+\sum_{k\in\mathcal{P}_r}\sum_{n\in P_r^k}Y_k^{r,t}H_{k,n}^{r,t}(\Phi_{n-1,n,k}^{r,t}-\Phi_{n,n+1}^{r,t}), \tag{31}$$

where $M_{d_r}^{r,t}\ge V_r$ indicates the request r is successfully served by the network until the t-th transmission. The memory usage $M_n^{r,t}$ for r is released after a new request is inserted into the request set $\mathcal{R}$ to fill up the position.

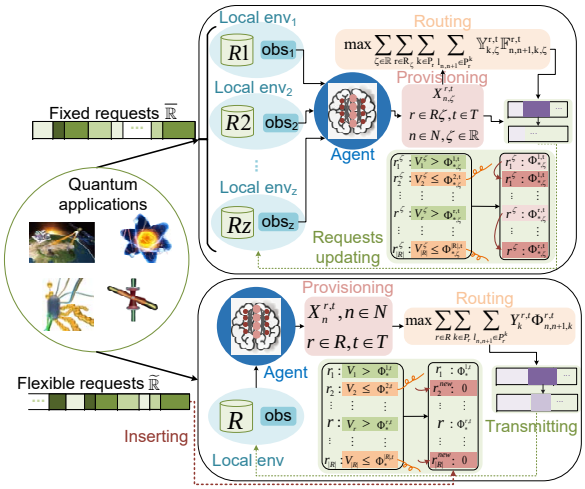


Fig. 6: Sophon's workflow after being deployed in the network.

E. Deployment

Each agent is trained to cooperatively transmit qubit data with constraints to satisfy the requirements of quantum application requests while individually optimizing its resource utilization with local observations for selfish consideration. The size of the inputs of each agent is fixed as $|R|(|\mathcal{P}_r|+1)+|\mathcal{N}|+1$, which is related to the topology scale and the request pool capacity. We provide two deployment schemes for the trained RL agents to apply *Sophon*:

1) Fixed Requests: Given is a fixed number of requests $\overline{\mathbb{R}}$ with different data transmission requirements generated by various quantum applications. $\overline{\mathbb{R}}$ is divided into z request pools, where $z=\left\lceil\frac{\overline{\mathbb{R}}}{|R|}\right\rceil$. The Rz request pool can be masked by 0 if it is not filled up by the divided requests. Each agent maintains z local environments for each request pool, but all local environments share all node memory resources.

After the provisioning process, we obtain the resource pre-allocation policy $\mathbb{X}_{n,\zeta}^{r,t}$. The global routing is to select the E2E routes by solving the extensive ILP that maximizes the transmission data volumes of $\overline{\mathbb{R}}$, where $\mathbb{F}$ is the extensive

Φ for the ζ-th request pool. In the transmitting process, the node memory usage is updated in terms of the determined $\mathbb{X}_{n,\zeta}^{r,t}$ and $\mathbb{Y}_{k,\zeta}^{r,t}$ by releasing the occupied resources of the completed requests, while each request pool deletes the completed requests and randomly copies the uncompleted ones to fill itself up. Agent terminates the interaction with the request pool $R\zeta, \zeta \in \overline{\mathbb{R}}$ when the required data transmission of all requests inside is satisfied. In Fig. 6, the ζ-th request pool $R\zeta$ copies requests r_1^ζ and r^ζ to fill up the request pool.

2) Flexible Requests: Give is an endless request queue $\widetilde{\mathbb{R}}$. An initial request pool R, selecting $|R|$ requests in turn from $\widetilde{\mathbb{R}}$, builds an initial local environment to start up the agent. With the determined resource provisioning and routing policies, the node updates its memory usage and judges whether the requests in the request pool R were successfully responded to. The completed requests are deleted from R, and their occupied resources are released. The request pool R asks new requests from $\widetilde{\mathbb{R}}$ to fill the pool. As shown in Fig. 6, the new requests r_2^{new} and $r_{|R|}^{new}$ initially has no data memory usage.

VI. Performance Evaluation

A. Simulation Setup

We evaluate the performance of *Sophon* on a self-made simulator, where the network topology is generated by using the *NetworkX* library[4] and *Waxman* model [36]. *Sophon* is implemented using Python 3.6 for RL and Branch-and-Bound algorithms while using Gurobi 9.1 as an LP solver. We train and run *Sophon* on an Intel (R) Core (TM) i7 Windows 10 64-bit system with NVIDIA GeForce GTX 1660 SUPER. We provide public access to the data, code, and results of the problem example in Section IV-A and performance evaluation in Section VI at the GitHub [5]. The performance setup details are as follows:

Network Topology: The network topology is randomly generated as an area of 100K units by 100K units square, which contains 18 quantum nodes with an average of 3 neighbors of each node. We randomly generate requests with different source and destination nodes, where the required data volume of each request is uniformly picked up from the interval $[10, 20]$. The node capacities of all quantum nodes are generated by a Gaussian distribution with the expectation $\mu = 17$, the variance $\sigma = 2$, and the clipping by the bound $[9, 26]$. We adopt Yen's algorithm [37] to obtain the first $|P_r| = 3$ candidate routes for each request.

Quantum State: We set the probability of a quantum state to be $|1\rangle$ as 0.5 and the scattering constant τ of decoherence can be set as 0.1 in a specific environment [9]. X_{thr}, the upper bound of variables $X_n^{r,t}$, is 4. The size of a request pool $\mathcal{R}$ is initially 5. The thresholds of entanglement fidelity [6] and transmission delay are initially $10e-8$ and 150.

RL Agent: The size of experience pool is $10e3$. The discounted reward factor is 0.9. The learning rate α is $10e-3$. The batch size of the inputs is 32. We save all distributed agents after the training process of *Sophon* with $10e4$ episodes. The step limitation of an episode is $T = 10$.

Baselines: *Sophon* deploys each trained agent for resource provision on the network node and the route selection algorithm on the centralized platform. We first verify the effectiveness of Sophon by comparing it with OBO and All_in transmission frameworks and then introduce two kinds of All_in frameworks (EFiRAP and Multi_R) to further demonstrate the flexibility of Sophon. The baselines are as follows:

- *EFiRAP:* An *All_in* [12], [24] scheme where it maximizes the network throughput by selecting the qualified entanglement route limited by the node memory from all candidate routes. Because of the computing complexity caused by the traversal of all possible candidate routes, we use a simple rounding to obtain the results when the problem scale becomes larger.
- *Multi_R:* A route-level *AMOR* [14], [23] scheme where it maximizes the network throughput by selecting at-most-one qualified route of a request limited by the node memory. We obtain the solution by using the branch-and-bound algorithm.

Performance Indicators: We provide the following indicators to evaluate the effectiveness, flexibility, robustness, and scalability of Sophon:

- *The communication cost:* It refers to t_{done} when all transmission requirements of a request set are completed for t_{done}-times entanglement transmission, denoted as:

$$obj : \min t_{done}, \forall r \in R$$
$$s.t. : \sum_{t \in t_{done}} \sum_{k \in P_r} \sum_{n \in P_r^k} Y_k^{r,t} H_{k,n}^{r,t} \Phi_{n,n+1,k}^{r,t} \geq V_r. \quad (32)$$

- *The algorithm running time:* It refers to the time during which the decision-making algorithms running on our experimental platform output the routing and provisioning decisions for one-time entanglement transmission.
- *The convergence of Sophon:* We adopt the number of the completed requests, the total transmission data volume $\sum_t \widetilde{\Phi}$ in an episode, the transmitting data of all completed requests $\sum_r \Phi^*$, and the average transmitting ratio to evaluate the convergence performance of *Sophon*. The average transmitting ration is $(\sum_r \Phi^* - \sum_t \widetilde{\Phi})/|R|$.
- *Quantum Bit Error Rate (QBER):* It is related to the bit-flip error rate caused by the collapse decoherence here. Assuming the bit-flip error rate of a qubit in the multi-qubit entanglement is p'', we can derive the correlation among QBER, fidelity, and entanglement dimension as [7]:

$$F_{p''}^d = C_d^d (1-p'')^d + \sum_{k=2}^{d} \left[C_d^{d-k} \left(k! \sum_{kk=0}^{k} \frac{(-1)^{kk}}{kk!} \right) (1-p'')^{d-k} (p''/2)^k \right]. \quad (33)$$

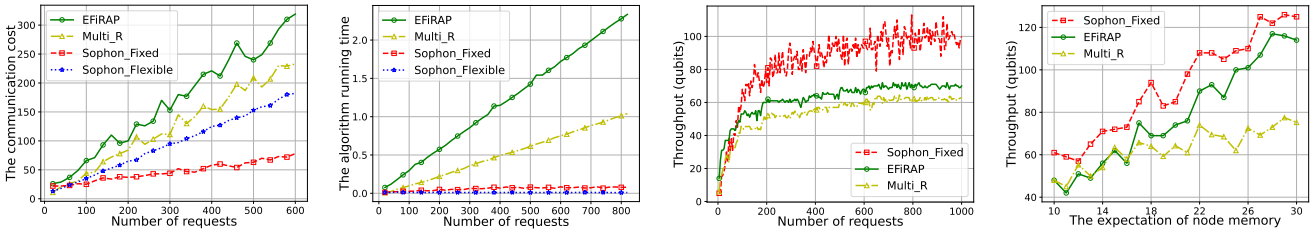
With the given fidelity F^d and entanglement dimension d of a link-level entanglement, its QBER p'' can be obtained by Eq. (33).

[4] https://networkx.org/

[5] https://github.com/yanangao1709/Sophon.git

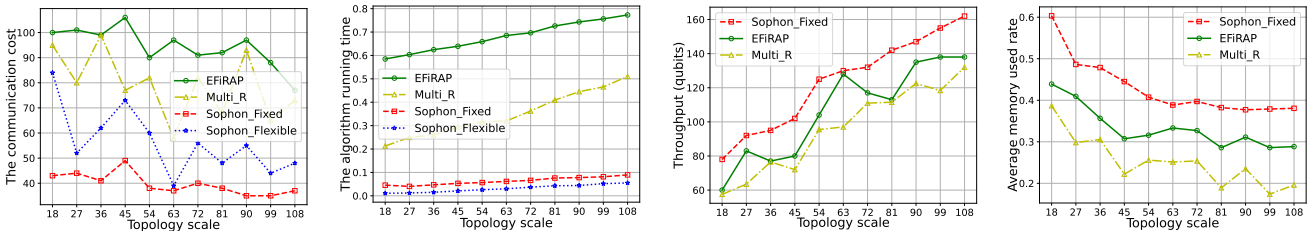
[6] In the practical quantum system, the fidelity varies from the scatter constant determined by the physical property.

[7] Sophon.git: Sophon/Document/Appendices.pdf.



(a) c.c. & request scale (b) a.r.t. & request scale (c) thr. & request scale (d) thr. & n.m.

Fig. 7: The comparison performance on a fixed topology.



(a) c.c. & top.s (b) a.r.t. & top.s. (c) thr. & top.s. (d) a.m.u.r. & top.s.

Fig. 8: The comparison performance on various scales of topology.

- *Signal-to-Noise Ratio (SNR)*: It refers to the ratio between the pure state and the noised state, which is denoted as $\frac{|\varphi\rangle F(|\varphi\rangle)}{|\varphi\rangle(1-F(|\varphi\rangle))}$. $|\psi\rangle$ is the initial state of a multi-qubit E2E entanglement, and $F(|\varphi\rangle)$ is the fidelity probability in the collapse decoherence model, obtained by Eq. (7).
- *Quantum Phased Error Rate (QPER):* It refers to the phased-flip error caused by the collapse decoherence here, which has no equivalent in classical computing, like the QBER. Assuming the phased-flip error rate is p', we derive the correlation among QPER, fidelity, and entanglement dimension as:

$$F_{p'}^{d} = p'^{d} + (1-p')^{d}. \tag{34}$$

With the given fidelity F^d and entanglement dimension d of a link-level entanglement, its QPER p' can be obtained by Eq. (34).

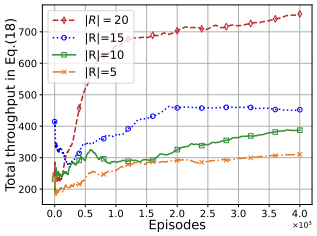
B. Performance Analysis

The Comparison Performance on a Fixed Topology: After deploying the agent trained on a fixed topology with 18 nodes, we obtain the comparison performance with baselines on a fixed topology in Fig. 7. In the implementation of EFiRAP and Multi_R, the node memory is occupied until the request is completely satisfied with its transmission requirement. Fig. 7(a) shows the increased tendency of the communication cost when the scale of the request set gradually becomes larger from 20 to 600. More requests require more entanglement transmission times. *AMOR* scheme, containing Multi_R and Sophon, performs better than *All_in* (EFiRAP) in the long-term communication process due to the fairness consideration on each request. Sophon, as a link-level *AMOR* scheme, requires less communication cost than Multi_R, a route-level *AMOR* scheme, because of the more fine-grained provisioning on the qubit resources.

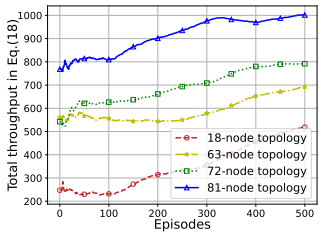
Fig. 7(b) and Fig. 7(c) respectively display the algorithm running time and the throughput obtained in one-time entanglement transmission. EFiRAP needs the most time to complete the request set because it tries all possible entanglement routes and iteratively runs the LP solver to find the integer solution through traversing methods like DP. Sophon, both fixed and flexible schemes, generates the routing and resource provisioning policies faster than Multi_R, even though it is more fine-grained. This is because the distributed agent shares the solving complexity of qubit provisioning, which decreases the size of the solution space. However, from the short-term view, Multi_R performs worse on the one-time throughput than EFiRAP due to the selection limitation of the candidate route, as shown in Fig. 7(c). Benefiting from the more fine-grained resource provisioning, Sophon compensates for this deficiency with the best throughput performance. Because of the limitation of the network resource, the one-time throughput is gradually convergent as the scale of the request set becomes larger.

Fig. 7(d) shows the impacts of enlarged node memory on the one-time throughput. We respectively generate the node memory scales by adjusting the expectation of the Gaussian Distribution, from which each node memory is selected randomly. It is obvious that larger node memory brings more throughput with 200 requests on a fixed topology. Sophon is more sensitive to the network node memory than the baselines due to the better fine-grained provisioning.

The Comparison Performance on Various Scales of Topology. Fig. 8 displays the comparison performance with baselines on the topology of (18,27,36,45,54,63,72,81,90,99,108) nodes. With a fixed (200) number of requests, the communication cost gradually decreases as the topology scale enlarges. This is because a larger topology brings more resources to satisfy the

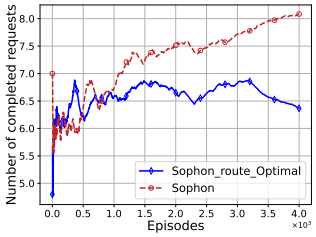


(a) Total thr. & $|R|$

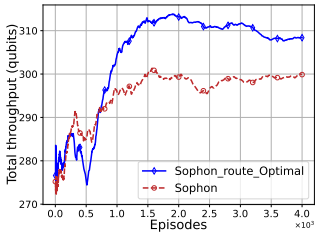


(b) Total thr. & top.s.

Fig. 9: The convergence performance of Sophon.

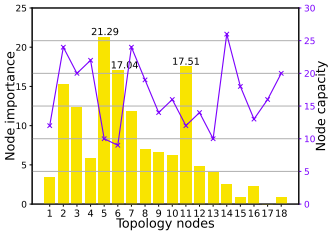


(a) The completed requests

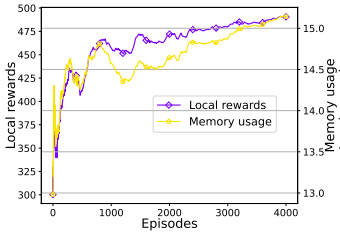


(b) The total throughput

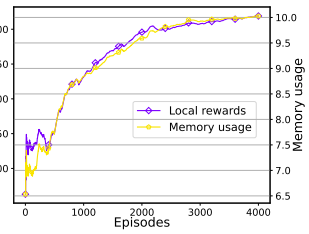
Fig. 10: Approximate-Optimal route selection



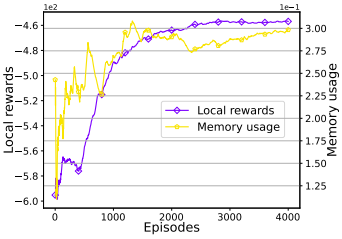
(a) Node importance



(b) Node agent 5



(c) Node agent 6



(d) Node agent 11

Fig. 11: The convergence performance of distributed agents.

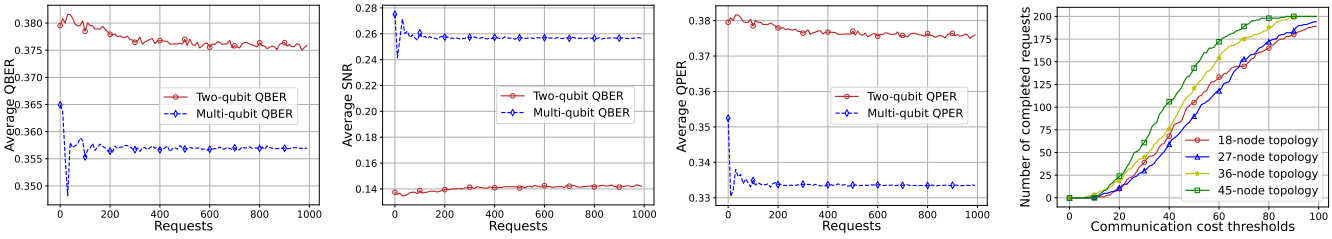
transmission requirement of the request set. The resource competition in each network node is relieved due to fewer requests traveling from them. For the same reason, each request can be allocated more qubit resources in the network node, which brings more throughput in one-time entanglement transmission. As shown in Fig. 8(c), all methods obtain more short-term throughput as the topology scale enlarges. Therefore, Sophon works well under the circumstances, no matter whether the resource competition is relaxing or tense in network nodes.

Fig. 8(b) provides the algorithm running time for responding to the fixed request set on various topology scales. We can see that the a.r.t of all methods becomes longer because a larger topology enlarges the solution space, which increases the problem's complexity. Sophon runs faster due to decomposing the solution space into the routing and resource provisioning stages. Fig. 8(d) displays the average memory used rate of the network nodes as the topology scale. A larger topology makes the network resource redundant for the fixed request set, so the a.m.u.r. becomes smaller.

The Convergence Performance of the Global Objective: Each RL agent's training episode contains $T = 10$ times of interaction with the network environment, i.e., 10 times of E2E entanglement transmission. The total throughput is the summation of all times in an episode. As shown in Fig. 9 and Fig. 10, the total throughput gradually converges as the episode becomes larger. Fig. 9(a) illustrates the convergence process of the total throughput with various sizes $(5, 10, 15, 20)$ of the request pool. Because the agent with the request pool of $|R| = 20$ can respond to 20 requests, its total throughput is larger than others. $R = 20$ agents converge much more quickly since they have more interaction attempts with the network environment. Fig. 9(b) displays the convergence process of the total throughput on various scales of topology. We can see that a larger topology conducts a better total throughput after converging.

Approximate-Optimal Route Selection: We train the distributed agent with the $|R| = 5$ request pool and select the optimal policy for route selection from $4^5 = 1024$ (the number of candidate routes is 3) solutions in the routing stage. Obtaining the optimal entanglement route from the candidate route set for each request is hard because the solution space is exceedingly larger with the increasing complexity of $(|P| + 1)^{|R|}$. Sophon adopts the approximate algorithm to reduce computational complexity and globally aims to maximize the total throughput. Therefore, in Fig. 10(a), the total throughput of Sophon is smaller than Sophon_route_optimal's after the convergence. Fig. 10(b) shows the convergence of the number of completed requests between Sophon and Sophon_route_optimal, where Sophon unexpectedly performs better. We take the average value of these results to make the curve smooth. We infer that this is because (1) the number of completed requests is considered in the local reward of each distributed agent in Eq. (14), (2) the total throughput of Sophon_route_optimal is redundant for the requirement of the completed request, while RL can eliminate these redundancies through the interaction with the approximate route selection. Distributed RL agents bridge the gap between the optimal solution and the approximate solution in route selection while simultaneously maximizing the global-local joint optimization objectives.

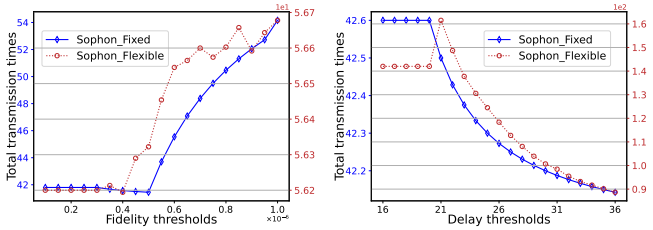
Convergence Performance of Distributed Agents: We randomly generate 50 requests with their data transmission requirements. The top 3 shortest candidate routes of each request are obtained using Yen's algorithm. We record the used times of each node in all candidate routes of the randomly



(a) Average QBER (b) Average SNR (c) Average QPER (a) The c.c. thresholds

Fig. 12: The noise-resistance of multi-qubit transmission

Fig. 13: The c.c. thresholds



(a) Fidelity threshold (b) Delay threshold

Fig. 14: Sophon's robustness on different thresholds

selected request set as node importance" and repeat this process for 100 times to obtain Fig. 11(a). The node importance of nodes $5, 6, 11$ are respectively $21.29, 17.51, 17.04$, and their memory capacities are smaller than others, which leads to more complicated resource provisioning. Therefore, we provide the performance of their agents. We exhibit the local reward (in Eq.(14)) convergence of these agents in Fig. 11(b), Fig. 11(c) and Fig. 11(d). 3 agents converge approximately at 3700, 3500 and 2600 episodes. We display the changes in memory usage of each agent, which converges with the local rewards. It implies that more and more node resources are effectively utilized in the training process. The maximized memory usage of node 5 and node 6 is larger than their capacities because of the accumulated memory usage in an episode.

The noise-resistance ability: We first obtain the fidelity by using Equation $F^d = F(|\psi\rangle_d^f) = \frac{1}{\sqrt{fd}}p(1-p)^{f\cdot d-1}e^{-\tau(\Delta x)^2}$ of a link-level entanglement between two network nodes. Then, with a known fidelity F^d and entanglement dimension d, we obtain the link-level QBER p'' by using Eq. (33). The average QBER is obtained by $\sum_{r\in R} Y_k^r \frac{\sum_{h=0}^{|P_k^r|} p_h''}{|P_k^r|}$, while the average SNR is obtained by $\sum_{r\in R} Y_k^r \frac{\sum_{h=0}^{|P_k^r|} F_h^d}{|P_k^r|}$, where h denotes the h-th link of the route P_k^r. We provide the performance analysis on various scales of the request set to verify the advantages of multi-qubit transmission. As shown in Figure 12(a), the average QBER in the two-qubit scheme is much higher than the ones in the multi-qubit scheme. It indicates that multi-qubit transmission in Sophon is more robust than two-qubit transmission because the multi-qubit scheme has a better noise-resistance ability and higher entanglement quality. Moreover, as shown in Figure 12(b), the multi-qubit transmission scheme has a larger SNR compared to the two-qubit transmission scheme. It results from the stability of the high-dimensional entanglement state in terms of the Equations $F(|\psi\rangle_d^f) = \frac{1}{\sqrt{fd}}p(1-p)^{f\cdot d-1}e^{-\tau(\Delta x)^2}$ and $\frac{|\varphi\rangle F(|\varphi\rangle)}{|\varphi\rangle(1-F(|\varphi\rangle))}$.

Similar to QBER and SNR, we obtain the average QPER of a request set by $\sum_{r\in R} Y_k^r \frac{\sum_0^{|P_k^r|} p'}{|P_k^r|}$. Especially, QBER is the same as QPER with the formulation $F^d = p^2 + (1-p)^2$ in the two-qubit entanglement transmission scheme. In the multi-qubit transmission scheme, QPER is formulated as a structured form to obtain the link-level fidelity, through which QPER of a link-level entanglement can be obtained by the inverse function $F_{p'}^{-1}$ with the known fidelity F^d. We experiment with Sophon on different scales of the request set, respectively, with the two-qubit and multi-qubit transmission scheme to obtain the average QPER. As shown in Figure 12(c), the multi-qubit transmission scheme has a smaller average QPER than the two-qubit transmission. Moreover, with the increasing scale of the request set, the average QPER maintains an approximately stable trend.

Robustness on Thresholds: We run the deployed *Sophon* to respond 20 requests with $|R| = 5$ agent for 50-times to obtain the average performance of different thresholds. As shown in Fig. 14, *Sophon* is significantly robust and feasible for different scales of constraints. The increase of the fidelity threshold brings a transmission burden to *Sophon* with more total transmitting times. The increase in the delay threshold accelerates the transmission process. That's because the fidelity threshold is the upper bound of entanglement quality, while the delay threshold is the lower bound of route delay. A higher fidelity threshold and a smaller delay threshold mean that more candidate routes of the request set are considered, which can obtain larger $\widetilde{\Phi}$ and further affect the overall transmission performance.

We set various communication cost thresholds to respond to 200 requests on different scales of network topology $(1827, 36, 45$ nodes). Fig. 13 illustrates that fewer times of entanglement transmission can only complete the transmission tasks of fewer requests. Moreover, a larger topology is more sensitive to the limitation of the communication cost thresholds. This is because a larger topology possesses more network resources to support the transmission requirements of the request set, which can complete the transmission task by consuming fewer times of entanglement transmission.

VII. Conclusion

In this paper, we study the problem of transmission optimization in quantum networks. Focusing on the weak transmission ability of the one-qubit quantum entanglement and the limitation of network resources, we introduce high-dimensional multi-qubit quantum entanglements with W state representation, which is more noise-resistant for quantum decoherence. We formulate the quantum network model with multi-qubit entanglement and define the Entanglement Routing and Qubit Provisioning (*ERQP*) problem to jointly optimize the global-local objectives with manifold constraints. We propose *Sophon*, an online transmission framework that decomposes the solution space into quantum resource pre-provision and entanglement route selection. We use multi-agent RL to pre-allocate resources and extend the Markov process by introducing an approximately optimal branch-and-bound algorithm to select E2E entanglement routes. The extensive simulations on the self-made network topology demonstrate that *Sophon* is efficient and flexible to respond quickly to massive and heterogeneous transmission requirements while optimizing the joint objectives.

References

[1] S. Wehner, D. Elkouss, and R. Hanson, "Quantum internet: A vision for the road ahead," *Science*, vol. 362, no. 6412, p. eaam9288, 2018.

[2] H. Omaar. (2023) The u.s. approach to quantum policy. [Online]. Available: https://itif.org/publications/2023/10/10/the-us-approach-to-quantum-policy/

[3] J. P. (2023) Chinese quantum companies and national strategy 2023. [Online]. Available: https://thequantuminsider.com/2023/04/13/chinese-quantum-companies-and-national-strategy-2023/

[4] V. M. Acosta, D. Dequal, M. Schiavon, A. Montmerle-Bonnefois, C. B. Lim, J.-M. Conan, and E. Diamanti, "Analysis of satellite-to-ground quantum key distribution with adaptive optics," *New Journal of Physics*, vol. 26, no. 2, p. 023039, 2024.

[5] D. Cuomo, M. Caleffi, K. Krsulich, F. Tramonto, G. Agliardi, E. Prati, and A. S. Cacciapuoti, "Optimized compiler for distributed quantum computing," *ACM Transactions on Quantum Computing*, vol. 4, no. 2, pp. 1–29, 2023.

[6] N. Aslam, H. Zhou, E. K. Urbach, M. J. Turner, R. L. Walsworth, M. D. Lukin, and H. Park, "Quantum sensors for biomedical applications," *Nature Reviews Physics*, vol. 5, no. 3, pp. 157–169, 2023.

[7] M. Erhard, M. Krenn, and A. Zeilinger, "Advances in high-dimensional quantum entanglement," *Nature Reviews Physics*, vol. 2, no. 7, pp. 365–381, 2020.

[8] L. Bulla, K. Hjorth, O. Kohout, J. Lang, S. Ecker, S. P. Neumann, J. Bittermann, R. Kindler, M. Huber, M. Bohmann *et al.*, "Distribution of genuine high-dimensional entanglement over 10.2 km of noisy metropolitan atmosphere," *Physical Review A*, vol. 107, no. 5, p. L050402, 2023.

[9] M. A. Schlosshauer, *Decoherence and the quantum-to-classical transition*. Springer Science & Business Media, 2007.

[10] M. Schlosshauer, "Quantum decoherence," *Physics Reports*, vol. 831, pp. 1–57, 2019.

[11] S. Shi and C. Qian, "Concurrent entanglement routing for quantum networks: Model and designs," in *ACM SIGCOMM*, 2020, pp. 62–75.

[12] Y. Zhao and C. Qiao, "Redundant entanglement provisioning and selection for throughput maximization in quantum networks," in *IEEE INFOCOM*, 2021.

[13] J. Li, M. Wang, K. Xue, R. Li, N. Yu, Q. Sun, and J. Lu, "Fidelity-guaranteed entanglement routing in quantum networks," *IEEE Transactions on Communications*, vol. 70, no. 10, pp. 6748–6763, 2022.

[14] Y. Zeng, J. Zhang, J. Liu, Z. Liu, and Y. Yang, "Multi-entanglement routing design over quantum networks," in *IEEE INFOCOM 2022*, 2022, pp. 510–519.

[15] G. Zhao, J. Wang, Y. Zhao, H. Xu, L. Huang, and C. Qiao, "Segmented entanglement establishment with all-optical switching in quantum networks," *IEEE/ACM Transactions on Networking*, 2023.

[16] V. Srivastav, N. H. Valencia, W. McCutcheon, S. Leedumrongwatthanakun, S. Designolle, R. Uola, N. Brunner, and M. Malik, "Quick quantum steering: Overcoming loss and noise with qudits," *Physical Review X*, vol. 12, no. 4, p. 041023, 2022.

[17] T. J. Bell, J. F. Bulmer, A. E. Jones, S. Paesani, D. P. McCutcheon, and A. Laing, "Protocol for generation of high-dimensional entanglement from an array of non-interacting photon emitters," *New Journal of Physics*, vol. 24, no. 1, p. 013032, 2022.

[18] M. Ringbauer, M. Meth, L. Postler, R. Stricker, R. Blatt, P. Schindler, and T. Monz, "A universal qudit quantum processor with trapped ions," *Nature Physics*, vol. 18, no. 9, pp. 1053–1057, 2022.

[19] A. Farahbakhsh and C. Feng, "Opportunistic routing in quantum networks," *arXiv preprint arXiv:2205.08479*, 2022.

[20] 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.

[21] L. Gyongyosi and S. Imre, "Decentralized base-graph routing for the quantum internet," *Physical Review A*, vol. 98, no. 2, p. 022310, 2018.

[22] Y. Zhao and C. Qiao, "Distributed transport protocols for quantum data networks," *IEEE/ACM Transactions on Networking*, vol. 31, no. 6, pp. 2777–2792, 2023.

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

[24] Y. Zhao, G. Zhao, and C. Qiao, "E2e fidelity aware routing and purification for throughput maximization in quantum networks," in *Proceedings of the IEEE INFOCOM*, 2022.

[25] H. Gu, Z. Li, R. Yu, X. Wang, F. Zhou, and J. Liu, "Fendi: High-fidelity entanglement distribution in the quantum internet," *arXiv preprint arXiv:2301.08269*, 2023.

[26] S. Pouryousef, N. K. Panigrahy, and D. Towsley, "A quantum overlay network for efficient entanglement distribution," in *IEEE INFOCOM 2023-IEEE Conference on Computer Communications*. IEEE, 2023, pp. 1–10.

[27] Y. Zhong, H. S. Chang, A. Bienfait, and et al., "Deterministic multi-qubit entanglement in a quantum network," *Nature*, vol. 590, pp. 571–575, 2021.

[28] V. Mannalath and A. Pathak, "Multiparty entanglement routing in quantum networks," *Physical Review A*, vol. 108, no. 6, p. 062614, 2023.

[29] E. Sutcliffe and A. Beghelli, "Multi-user entanglement distribution in quantum networks using multipath routing," *IEEE Transactions on Quantum Engineering*, 2023.

[30] Y. Zeng, J. Zhang, X. Shang, J. Liu, Z. Liu, and Y. Yang, "Multi-user entanglement routing design over quantum internets," in *2024 IEEE 44th International Conference on Distributed Computing Systems (ICDCS)*. IEEE, 2024, pp. 266–276.

[31] Y.-C. Liang, Y.-H. Yeh, P. E. Mendonça, R. Y. Teh, M. D. Reid, and P. D. Drummond, "Quantum fidelity measures for mixed states," *Reports on Progress in Physics*, vol. 82, no. 7, p. 076001, 2019.

[32] Y. Li, S.-Y. Huang, M. Wang, C. Tu, X.-L. Wang, Y. Li, and H.-T. Wang, "Two-measurement tomography of high-dimensional orbital angular momentum entanglement," *Physical Review Letters*, vol. 130, no. 5, p. 050805, 2023.

[33] C. Li, T. Li, Y.-X. Liu, and P. Cappellaro, "Effective routing design for remote entanglement generation on quantum networks," *npj Quantum Information*, vol. 7, no. 1, pp. 1–12, 2021.

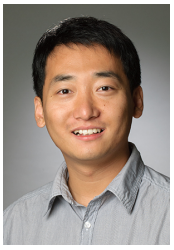
[34] M. Mitzenmacher and E. Upfal, *Probability and computing: Randomization and probabilistic techniques in algorithms and data analysis*. Cambridge university press, 2017.

[35] C. Barnhart, C. A. Hane, and P. H. Vance, "Using branch-and-price-and-cut to solve origin-destination integer multicommodity flow problems," *Operations Research*, vol. 48, no. 2, pp. 318–326, 2000.

[36] B. Waxman, "Routing of multipoint connections," *IEEE Journal on Selected Areas in Communications*, vol. 6, no. 9, pp. 1617–1622, 1988.

[37] J. Y. Yen, "Finding the k shortest loopless paths in a network," *management Science*, vol. 17, no. 11, pp. 712–716, 1971.

Yanan Gao received the B.S. degree from the School of Computer Science and Technology, Shandong University of Science and Technology, Qingdao, China, in 2015, and the M.S. degree in software engineering from Beijing Forestry University, Beijing, China, in 2020. She is currently a Ph.D. student at the School of Computer Science and Technology, Beijing Institute of Technology. Her research interests include reinforcement learning algorithms and quantum networks.



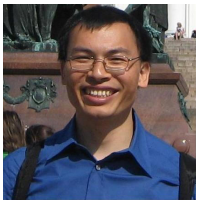
Song Yang (Senior Member, IEEE) is currently an associate professor at the School of Computer Science at Beijing Institute of Technology, China. Song Yang received the B.S. degree in software engineering and the M.S. degree in computer science from Dalian University of Technology, Dalian, Liaoning, China, in 2008 and 2010, respectively, and the Ph.D. degree from Delft University of Technology, The Netherlands, in 2015. From August 2015 to August 2017, he worked as postdoc researcher for the EU FP7 Marie Curie Actions CleanSky Project in Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen (GWDG), Göttingen, Germany. His research interests focus on data communication networks, cloud/edge computing, and network function virtualization.

Fan Li (Member, IEEE) received the Ph.D. degree in computer science from the University of North Carolina at Charlotte in 2008, MEng degree in electrical engineering from the University of Delaware in 2004, MEng and BEng degrees in communications and information system from Huazhong University of Science and Technology, China in 2001 and 1998, respectively. She is currently a professor at the School of Computer Science in Beijing Institute of Technology, China. Her current research focuses on wireless networks, ad hoc and sensor networks, and mobile computing. Her papers won Best Paper Awards from IEEE MASS (2013), IEEE IPCCC (2013), ACM MobiHoc (2014), and Tsinghua Science and Technology (2015). She is a member of IEEE and ACM.

Youqi Li (Member, IEEE) is an assistant professor in the School of Computer Science, Beijing Institute of Technology, Beijing, China. He received the Ph.D. degree and the BS degrees in Computer science and technology from Beijing Institute of Technology, Beijing, China, in 2020 and 2015, respectively. His research interests include IoT, mobile crowd sensing, privacy, game theory, and online machine learning.

Liehuang Zhu (Senior Member, IEEE) received the BE and ME degrees from Wuhan University, Wuhan, China, in 2001 and 1998, respectively, and the PhD degree in computer science from the Beijing Institute of Technology, Beijing, China, in 2004. He is currently a professor at the School of Cyberspace Security and Technology, Beijing Institute of Technology, Beijing, China. He has published more than 150 peer-reviewed journal or conference papers. He has been granted a number of IEEE Best Paper awards, including IWQoS 17', TrustCom 18', and ICA3PP 20'. His research interests include security protocol analysis and design, blockchain, wireless sensor networks, and cloud computing.

Stojan Trajanovski (Member, IEEE) received his Ph.D. degree (cum laude, 2014) from Delft University of Technology, The Netherlands, and his master degree in Advanced Computer Science (with distinction, 2011) from the University of Cambridge, United Kingdom. He is currently an applied scientist at Microsoft, working in London, UK, and Bellevue, WA, USA. He was in a similar role with Philips Research in Eindhoven, Netherlands from 2016 to 2019. Before that, he spent some time as a postdoctoral researcher at the University of Amsterdam and at Delft University of Technology. He successfully participated at international science olympiads, winning a bronze medal at the International Mathematical Olympiad (IMO) in 2003. His main research interests include network science & complex networks, machine learning, game theory, and optimization algorithms.



Xiaoming Fu (Fellow, IEEE) received his Ph.D. in computer science from Tsinghua University, Beijing, China in 2000. He was then a research staff at the Technical University Berlin until joining the University of Göttingen, Germany in 2002, where he has been a professor in computer science and heading the Computer Networks Group since 2007. He has spent research visits at the universities of Cambridge, Uppsala, UPMC, Columbia, UCLA, Tsinghua, Nanjing, Fudan, and PolyU of Hong Kong. Prof. Fu's research interests include network architectures, protocols, and applications. He is currently an editorial board member of IEEE Network, IEEE Transactions on Network and Service Management, IEEE Transactions on Network Science and Engineering, IEEE Networking Letters, and Elsevier Computer Communications, and has served on the organization or program committees of leading conferences such as INFOCOM, ICNP, ICDCS, MOBICOM, MOBIHOC, CoNEXT, ICN and Networking. He is an IEEE Fellow, an IEEE Communications Society Distinguished Lecturer, an ACM Distinguished Member, a fellow of IET, and a member of the Academia Europaea.