Multi-user Entanglement Routing Design over Quantum Internets

Yiming Zeng*‡, Jiarui Zhang*‡, Xiaojun Shang § Ji Liu*, Zhenhua Liu*, Yuanyuan Yang*
*Stony Brook University, Stony Brook, NY 11794, USA
§University of Texas at Arlington, Arlington, TX 76019, USA
{yiming.zeng, jiarui.zhang.2, ji.liu, zhenhua.liu, yuanyuan.yang}@stonybook.edu, xiaojun.shang@uta.edu

***Abstract*—Quantum Internet has potential capabilities far beyond the traditional Internet and is thus a promising future platform for communication and computation. Entanglement is a cornerstone of quantum mechanics and forms the basis of numerous quantum applications in the quantum Internet. While existing studies primarily focus on two-user entanglement, a plethora of applications necessitates the leap to multi-user entanglement. This paper tackles the fundamental problem of multi-user entanglement routing in the quantum Internet, aiming to entangle multiple quantum users with a high entanglement rate. We abstract the problem as a novel graph routing problem, which is not readily addressed by existing graph problem solutions due to the unique characteristics of the quantum Internet. To address this problem, we first consider a sufficient condition ensuring a feasible solution's existence and design an algorithm with the optimal solution. Given the NP-Completeness and NP-Hardness of determining a feasible solution's existence and deriving an optimal solution in general cases, respectively, we propose two heuristic algorithms to offer efficient solutions, which are shown, via extensive simulations, to outperform the existing algorithms in terms of entanglement rates.**

***Index Terms*—Quantum Internet; Multi-user Entanglement Routing; Entanglement-Swapping under Bell State Measurements**

I. Introduction

The quantum Internet, a future vision of global computing and communication, is poised to revolutionize how we compute and transmit secure information by exploiting the unique properties of quantum mechanics. Notably, quantum computing offers the potential to dramatically reduce the computational complexity for certain types of tasks, such as those involving Shor's algorithm for factoring integers [1], and the quantum linear system algorithm for solving linear equations [2]. Furthermore, quantum communication enables the generation, storage, and processing of information at levels of performance that significantly surpass those achievable by traditional means [3]. Some trail small-scale quantum networks, as the prototype of the quantum Internet, have been designed to handle the creation, transmission, and detection of entangled qubits [4], [5].

The essential feature of the quantum Internet is *entanglement* – a phenomenon in which quantum bits (qubits) become interlinked and the state of one instantly influences the other, regardless of the distance separating them [6]. This unique attribute serves as the foundational underpinning of potential quantum applications for the quantum Internet. However, the quantum entanglement process is probabilistic and unstable as qubits are inherently fragile and extremely susceptible to noise and disturbances in the environment [7]. The successful entanglement rate decreases exponentially with the distance between quantum users [7], making the establishment of reliable entanglement with a high entanglement rate a particularly daunting task. Meanwhile, the quantum users in the quantum Internet trying to be entangled through qubits could be distant from others [8].

Generating reliable entanglement with a high entanglement rate is crucial for the quantum Internet. To address that, current *entanglement-swapping* technology based on Bell State Measurements (BSMs) through quantum switches is an effective and widely-tested solution for enabling high entanglement rate entanglement [9]. These devices are quantum processors that work as relays aiming to extend the range of entanglement by entangling qubits across smaller, manageable distances. Fig. 1 illustrates an example where Alice and Bob each share independent Bell states with a switch. Following this, the switch executes BSMs to perform the entanglement-swapping. As a result, Alice and Bob become entangled through a Bell state pair, while qubits within the switch are freed. However,

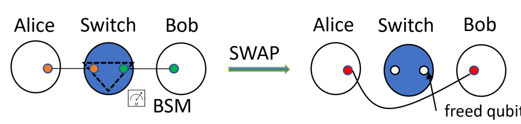


Fig. 1. An example of swapping based on BSMs.

quantum switches are restricted by their limited capacity of quantum memories (i.e., qubits), which are essential for entanglement-swapping as the entanglement demands from quantum users far beyond the switches' capacity. Quantum switches typically have a very limited number of qubits, often around 10 in real quantum experiments [10]. Meanwhile, it is still very difficult, if not impossible, to build a quantum switch with a large number of qubits embedded in the near future. Therefore, the efficient utilization of qubits within the switch is crucial for facilitating entanglement.

Existing studies mainly focus on the design of two-user entanglement routing via entanglement-swapping based on BSMs to support quantum communication between pairs

‡ Both authors contributed equally to this research.

of quantum users [11]–[19]. This is because, according to No-Cloning theorem [6], unmeasured qubits can only be (re)transmitted between quantum user pairs instead of broadcasting among more than two quantum users.

However, many applications go beyond two-user entanglement and necessitate multi-user entanglement, such as quantum error correction [20], quantum secret sharing [21], quantum cryptography [22], and distributed quantum computing [8], [23], [24]. Some real small-scale experiments for multi-user entanglement (e.g., three nodes) have been executed [10], [25], [26]. Taking distributed quantum computing as an example, the state-of-the-art quantum computing processor currently can only support up to 127 qubits [27]. However, many quantum computing tasks (e.g., quantum annealing [28], quantum machine learning [29], quantum chemistry [30]) might require a computing capacity up to tens of thousands of qubits, far exceeding the capabilities of a single monolithic quantum computing processor. To counter this limitation, distributed quantum computing is widely considered to augment computational power by multi-user entanglement. Specifically, a collection of monolithic quantum computing processors are entangled over switches and optical fibers within the quantum Internet to augment computational power [8], [23], [24]. Thus, efficiently establishing multi-user entanglement will be an important stride toward a fully functional quantum Internet. However, approaches for two-user entanglement cannot be directly extended to multi-user entanglement due to the more stringent entanglement requirements and a significant increase in complexity associated with a larger number of entangled users.

In this paper, we focus on a fundamental issue that the quantum Internet must address to support a variety of quantum applications: *how can we generate multi-user entanglement efficiently at a high entanglement rate within the quantum Internet?*

Most existing studies about multi-user entanglement (routing) [31]–[35] utilize an entanglement-swapping technique called $n-$fusion [36], [37] to form/distribute Greenberger–Horne–Zeilinger (GHZ) states among multiple users. An example of $n-$fusion ($n = 3$) is shown in Fig. 2, where the switch takes GHZ projective measurements and thus forms a 3-GHZ state among three quantum users. However, under current and near-term technical conditions, $n-$fusion has two main limitations compared to swapping under BSM. First, $n-$fusion has a lower successful swapping rate [38]–[40]. Performing GHZ measurements of $n-$fusion in the real operation is much harder than BSMs as GHZ measurements involve manipulating multiple qubits that are inherently fragile. A measurement failure at a switch that is to entangle multiple users may disrupt entanglement among all users, especially considering $n-$fusion's lower successful swapping rate. Second, it produces GHZ states that are less stable compared to Bell states. Maintaining entanglement among multiple users is more challenging with GHZ states, as GHZ states are more susceptible to noise, which leads to decoherence [7]. As such, $n-$fusion results in a lower successful entanglement rate and makes the entanglement less reliable.

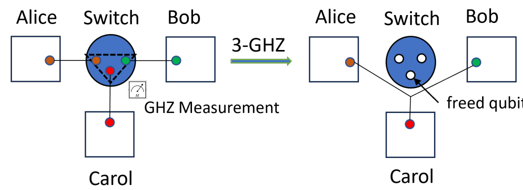


Fig. 2. 3-fusion: it fuses three quantum links and entangles three qubits.

Entanglement-swapping under BSMs is a well-established and more reliable method, as it operates on only two qubits at a time. This approach has been widely tested and applied in real-world experiments [4], [5], [9]. For quantum applications where the stability of entanglement is important, entangling quantum users through entanglement-swapping under BSMs is much more prevalent than $n-$fusion when possible [39]. Moreover, considering that BSM can be viewed as a specific case of n-fusion (i.e., $n = 2$) in theory [36], [39], a comprehensive understanding of this fundamental model, which closely aligns with real-world scenarios, is essential to navigate the complexities of multi-user entanglement design effectively.

A. Contribution

Noting that existing research primarily focuses on two-user entanglement, which significantly constrains the quantum applications that can be supported, or often oversimplifies network constraints such as the capacity of switches, or considers a less reliable entanglement-swapping method (i.e., $n-$fusion), this paper sets out to address these open questions. The key aim of this paper is to investigate how to enable efficient multi-user entanglement, which is a fundamental capability for the quantum Internet.

Problem Description. We model the quantum Internet as a network with arbitrary topology, where quantum switches have limited capacity and perform entanglement swapping via BSMs. Based on this model, we consider an entanglement routing problem for establishing multi-user entanglement for a set of users, with the goal of maximizing the entanglement rate. This fundamental model captures the core and essential properties of the quantum Internet. It is readily extendable to more complex situations, such as those accounting for fidelity decay or concurrent routing of multiple independent entanglement groups. Moreover, the model can serve as the basis for other related quantum research such as quantum mapping and the architectural design of the quantum Internet [8], and various applications mentioned before.

Contribution 1. We formulate the multi-user entanglement problem as a novel graph routing problem. To the best of our knowledge, this is an uncharted graph routing problem that does not fit directly into any existing classic graph routing problems, nor can it draw directly upon their existing solutions. This is due to the unique characteristics of the quantum Internet. Specifically, the 'connectivity' (entanglement) of vertex (quantum users), and non-additive optimization objectives present distinct and significant technical challenges compared to classic graph routing problems (Detailed analyses are provided in Sec. III-A). This new graph problem is not only crucial to multi-user entanglement in the quantum Internet but also adds a novel intersection between classic graph routing and quantum networking.

Contribution 2. We prove that determining the existence of a feasible solution to the problem is NP-complete, and further demonstrate that obtaining the optimal solution to this problem is NP-hard. This indicates that there are no known polynomial-time algorithms for finding *feasible* or *optimal* solutions unless P=NP.

Contribution 3. We develop three algorithms to address this problem. Specifically, we first design an algorithm with the optimal solution under the condition that the switch has sufficient capacity to support the entanglement. Given that the determination of a solution's existence and solving the problem are respectively NP-complete and NP-hard in general cases, we propose two heuristic algorithms to find solutions when switch capacity is limited.

Contribution 4. We evaluate the performance of our algorithms through a series of simulations across different parameters and topologies. Our evaluations demonstrate that the proposed algorithms are adaptable to various network topologies and can outperform existing methods. By adjusting the network topology, size, number of qubits, and exchange rate, we show how each variable affects the performance of the algorithm, thus providing a perspective on how each variable affects the MUERP problem. We design a simulation to uniformly and randomly remove edges in the network, which shows that the performance of our algorithm is mainly affected by some critical edges in the network structure.

II. Quantum Network Model

In this section, we introduce the quantum Internet model and present the problem statement. Figure 3 shows an example of the considered model.

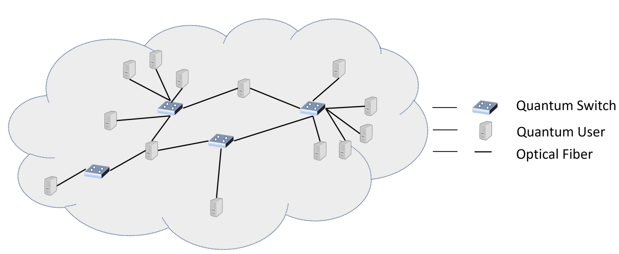


Fig. 3. An Example of the Quantum Internet. The terminologies are introduced in Sec. II-A.

A. Terminology

Quantum Users: A quantum user is a quantum processor or a quantum computing node that consists of a set of quantum processors interconnected. We assume that a quantum user has enough quantum memory to participate in entanglement with others. The set of quantum users is represented as $\mathcal{U} = \{u_1, u_2, \cdots, u_{|\mathcal{U}|}\}$, attempting to achieve entanglement among themselves, where $|\mathcal{U}|$ indicates the total number of users within the set $\mathcal{U}$.

Quantum Switches: The set of quantum switches is denoted as $\mathcal{R}$ with a total of $|\mathcal{R}|$ switches. The switch $r \in \mathcal{R}$ has Q_r qubits. The switches employ entanglement-swapping under BSMs to execute qubit swapping. The successful swapping rate for any qubits is uniform and represented as $q \in [0, 1]$. The node set, denoted as $\mathcal{V}$, is formed by the union of $\mathcal{U}$ and $\mathcal{R}$, i.e., $\mathcal{V} = \mathcal{U} \cup \mathcal{R}$.

Optical Fibers: Edge e_{v_i,v_j} represents an optical fiber cable connecting nodes $v_i \in \mathcal{V}$ and $v_j \in \mathcal{V}$ ($i \neq j$). A node could either be a user or a switch. A quantum link over the optical fiber entangles two neighboring nodes through a Bell state pair $\frac{|00\rangle+|11\rangle}{\sqrt{2}}$. The successful entanglement rate over the quantum link is determined by the link length and the physical material of the optical fiber, represented by $p = \exp(-\alpha L)$, where $\exp(x) = \mathbf{e}^x$ and $\mathbf{e}$ is the Euler's number approximately equal to 2.71828. α is a constant dependent on the physical material and L is the length the optical fiber. Each optical fiber comprises several independent cores, each of which can function as a quantum link for the entanglement. We assume that optical fibers, due to their lower cost and multiple-core design, have adequate capacity to support entanglement in the quantum Internet [41].

Network Topology: The quantum network consists of quantum users, quantum switches, and optical fibers. Users and switches are connected through optical fibers. The network is abstracted as an undirected graph $G = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \mathcal{U} \cup \mathcal{R}$, and $\mathcal{E} = \{e_{v_i,v_j} | v_i, v_j \in \mathcal{V}\}$.

B. Entanglement Process

The entanglement process in the quantum Internet for multiple quantum users, while similar to that for pairs of users [12], has increased complexity due to the greater number of entangled users and the entanglement topology's complexity. This process can be summarized as follows:

In brief, a central node (i.e., a server for traditional communication and computing in the network) collects entanglement requests from users and, using all available network information like topology and switches' capacity, formulates entanglement routes in an offline process. This plan is then sent to all switches and users via traditional means. Finally, with synchronized internal clocks across switches, the network executes the entanglement process to generate quantum links based on the pre-designed routes, with switches performing entanglement-swapping to establish entanglement between users.

C. Routing Metric

We define *entanglement rate*, i.e., the successful entanglement rate of multiple quantum users trying to be entangled, as the routing metric to evaluate the performance.

Entanglement Rate of a Quantum Channel: A *quantum channel* is a path consisting of quantum links and switches that connect two quantum users. For a quantum user pair $\langle u_i, u_j \rangle$, fix a quantum channel $A = \{v_0, v_1, v_2, \cdots, v_{l-1}, v_l\}$, where $v_0 = u_i$, $v_l = u_j$, $v_1, v_2, \cdots, v_{l-1}$ are quantum switches listed as the order in the path from the source u_i to the destination u_j, and l denotes the distance of A, i.e., the number of its quantum links. The adjacent nodes are connected by one quantum link. Each quantum switch along A will assign 2 qubits for the quantum channel.

Building a quantum channel for a pair of quantum users requires all quantum links between adjacent nodes and all switches in the channel to generate successful entanglement and swapping simultaneously during the fixed time period. Therefore, the entanglement rate for quantum channel A is:

$$P_A = q^{l-1}\Pi_{i=0}^{l} p_{i,i+1} = q^{l-1}\exp(-\alpha \sum_{i=0}^{l-1} L_{i,i+1}), \qquad (1)$$

where $L_{i,i+1}$ denotes the length of optical fibers between two neighboring nodes v_i and v_{i+1}. In this paper, we only consider the case there is at most one quantum channel between a quantum user pair. For example, in Fig. 4a, Alice and Bob are connected through one quantum channel consisting of two quantum links (represented by brown dashed lines) and a switch. If we assume the entanglement rate of any quantum link to be p and the swapping rate of the switch to be q, then the entanglement rate between Alice and Bob equates to $p^2 q$.
Entanglement Rate of Multi-users: We assume that for any subset of quantum users within $\mathcal{U}$ that includes more than one user, there are no quantum channels that create loops.

Definition 1. ***Entanglement Tree.*** *Given a quantum user set $\mathcal{U}$, achieving entanglement among all users within the set necessitates that these users form a tree where users are vertices and quantum channels are edges. We name this tree as an* ***entanglement tree****.*

An example is shown in Figure 4a. Alice, Bob, and Carol are entangled via two independent quantum channels, as represented by the brown dashed line and green dot-dashed line, respectively. These two channels specifically connect Alice with Bob and Alice with Carol, respectively. However, there is no channel connecting Bob and Carol due to the switch capacity limitation.

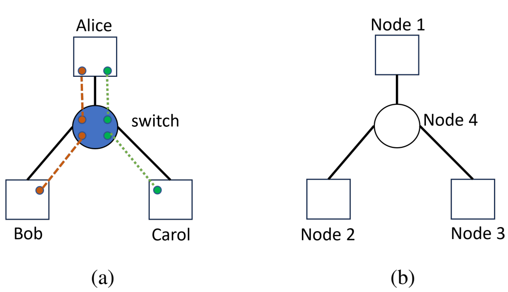


Fig. 4. Examples of connectivity in the quantum Internet and a classic graph model. (a) An entanglement example of three quantum users. Black lines indicate the optical fiber. The brown dashed line and the green dot-dashed line are quantum links over optical fibers for two independent quantum channels. (b) An example of connectivity of Node 1, 2, and 3 in the classic graph model, where Node 1, 2, and 3 form a Steiner tree in the graph.

Let $\mathcal{A}$ denote the set of quantum channels forming an entanglement tree that connects all quantum users in $\mathcal{U}$. Since there are no loops among users, to achieve multi-user entanglement among all users, all quantum channels need to be successfully entangled.

Therefore, the entanglement rate for the entanglement of the set $\mathcal{U}$ is the product of entanglement rates of all quantum channels of set $\mathcal{A}$, which is expressed as:

$$P = \Pi_{A \in \mathcal{A}} P_A. \qquad (2)$$

D. Problem Statement

In this subsection, we introduce a basic model of the multi-user entanglement routing problem in the quantum Internet with a general network topology, where switches use BSM entanglement-swapping. We assume that there is at most one quantum channel between one pair of quantum users, as our main focus is to address how to efficiently establish entanglement among multi-users. The objective is to maximize the entanglement rate of multi-users.

We model this problem as a graph routing problem, thereby laying the groundwork for potential applications in other graph-based contexts or extensions to more sophisticated entanglement routing scenarios, e.g., fidelity-aware entanglement, entanglement for multiple entanglement sets, and more.

For clarity, we revisit some notations mentioned before. In an undirected connected graph $G = (\mathcal{V}, \mathcal{E})$, $\mathcal{V} = \mathcal{U} \cup \mathcal{R}$ represents the vertex set, $\mathcal{E}$ indicates the optical fiber set where its elements are known as edges. e_{v_i,v_j} represents the edge between vertices $v_i \in \mathcal{V}$ and $v_j \in \mathcal{V}$ $(i \neq j)$. The vertex set $\mathcal{V}$ consists of two disjoint subsets: $\mathcal{U}$ for users and $\mathcal{R}$ for switches. We assume that G has no self-loops.

Definition 2. *A* ***channel*** *is a path with width 1 through vertices in $\mathcal{R}$ and edges in $\mathcal{E}$ that connects a pair of nodes in $\mathcal{U}$.*

Definition 3. *The* ***capacity*** *of a vertex in set $\mathcal{R}$ is the maximum number of channels that can support, i.e., $\lfloor Q_r/2 \rfloor, \forall r \in \mathcal{R}$. We assume that each vertex in set $\mathcal{U}$ has sufficient capacity to meet the channel requirements.*

Definition 4. *The* ***value*** *of a channel is defined as Eq. (1). The* ***value*** *of an entanglement tree is defined as Eq. (2).*

Problem. ***Multi-user Entanglement Routing Problem (MUERP)****: The MUERP is to route channels to form a spanning entanglement tree that can span $\mathcal{U}$, with the objective of maximizing the value of the tree while ensuring that the capacities of vertices in $\mathcal{R}$ are not exceeded.*

III. Problem Analyses

In this section, we present the technical challenges of the *MUERP*. In particular, we discuss unique characters brought by the quantum Internet that discriminate this problem from existing classic graph problems. Then, we prove that determining the existence of a feasible solution to this problem is NP-Complete, and driving an optimal solution is NP-Hard.

A. Unique Features of the Quantum Internet

The quantum Internet has some unique properties from the quantum aspect that make it different from classic graph models (e.g., spanning tree, Steiner tree [42]) in graph theory, which makes the *MUERP* significantly challenging.

The *main difference* between the model in the quantum Internet and the classic graph theory is the *connectivity of vertices. In graph theory, the connectivity is based on vertices and edges.* A set of vertices is connected if there exists a path between every pair of vertices in the set. More importantly, an edge in the graph can be a part of multiple paths simultaneously. Generally, once an edge connects two vertices, the connectivity will be guaranteed for any path over this edge.

However, the case is different in the model of quantum Internet. *The 'connectivity' (entanglement) of 'vertices' (quantum users) is based on 'vertices' and independent 'paths' (quantum channels).* While the quantum Internet is built on the topology of switches and optical fibers, achieving entanglement requires that quantum users are interconnected via quantum channels, consisting of quantum links between pairs of qubits over optical fibers. Analogous to a path in graph theory, a quantum channel connects a pair of quantum users. However, the qubits used in a quantum link of a channel for a specific user pair cannot be utilized by other quantum channels simultaneously. To ensure 'connectivity' (entanglement), multiple quantum users should be connected in pairs through independent quantum channels that do not share quantum links (qubits), as opposed to simple edge (optical fiber) connections. Furthermore, the number of channels is limited by the capacity of switches. Once a switch's qubits are exhausted, no additional channels can be established over it.

Fig. 4 presents two examples demonstrating the connectivity of both the quantum Internet and a classic graph under identical graph topology. In Fig. 4a, three quantum users are entangled through two independent quantum channels, utilizing four qubits within the switch to generate four quantum links. As such, the switch exhausts its capacity and cannot generate more quantum links. In Fig. 4b, Nodes 1, 2, and 3 are directly connected through Node 4 and their respective edges. If the switch only contains two qubits, thus enabling the service of only one quantum channel, then the entanglement of three quantum users is not feasible. However, this does not disrupt the connectivity in its original graph model.

In the classic graph theory, the description of the Steiner minimal tree problem [43] is similar to *MUERP* problem. Given a graph, the graphical Steiner minimal tree problem requires a minimum weight tree subgraph of the given graph that spans a given set of nodes on the graph. The key distinction between this problem and the MUERP problem lies in whether the number of quantum channels traversed by each node can surpass a specified limit. In the graphical Steiner minimum tree problem, there is no such restriction, allowing nodes to have any number of degrees. In the MUERP problem, the degree needs careful consideration, particularly when an edge may be shared by multiple channels. For instance, consider a graph featuring a central node and several leaf nodes. While the graphical Steiner minimum tree can be directly connected through the central point, MUERP necessitates determining the feasibility of a solution based on the number of qubits at the central node.

Another minor difference is the *optimization objective*. In classic graph models, the objective is usually the linear summation of cost from individual edges [42]. In the quantum Internet, as shown in Eq. (1) and Eq. (2), the objective is the product of quantum links (edges) or channels (paths).

B. Challenges

These two unique features bring significant research challenges. Firstly, due to the novel requirements for the *connectivity of vertices*, multi-user entanglement presents significantly greater challenges compared to two-user entanglement. While the latter only involves single-path routing [11], [12], [18], the former demands strict 'connectivity' among more than three vertices. As a result, this problem cannot be classified under existing classic graph problems, nor can it directly utilize their existing solutions. Secondly, the *optimization objective* (referenced by Eq. (1) and Eq. (2)) is not the linear summation of edge cost, presenting considerable hurdles for algorithmic design.

In Theorem 1, we prove that determining the existence of a feasible solution to this problem is NP-complete. Additionally, in Theorem 2, we demonstrate that finding an optimal solution is NP-hard.

Theorem 1. *Determining the existence of a feasible solution to the MUERP problem is NP-complete.*

Proof. We name the problem of whether there exists a feasible solution to the MUERP as *E-MUERP*. The proof relies on the NP-completeness of another problem, known as the Degree-constrained Spanning Tree Problem (*DCSTP*), which is NP-Complete [44]. Given a graph, the *DCSTP* is to determine if there exists a spanning tree that the degree of each vertex in the spanning tree does not exceed an upper bound k. We can reduce the *DCSTP* to the *E-MUERP*. In other words, if we can solve the *E-MUERP*, then we can solve the *DCSTP*.

Given a graph, we assume that all vertices are quantum users, and the limit number of qubits of a vertex is k. If we can find a tree that spans all the quantum users by quantum channels and satisfies the vertices' capacity limitation, then this tree is also a solution to the spanning tree in the graph with the degree limit k. Therefore, the *E-MUERP* is more difficult than the *DCSTP*, which means it is NP-complete. □

Theorem 2. *The MUERP is NP-Hard.*

Proof. We prove the NP-hardness of the *MUERP* by reducing from the Degree-constrained Minimum Spanning Tree (*DCMST*), which is known to be NP-hard [45]. The *DCMST* takes a graph as input and seeks the minimum spanning tree such that the degree of each node does not exceed a given bound k. To reduce the *DCMST* to the *MUERP*, we construct an instance of the *MUERP* from the *DCMST* input graph by setting each vertex as a quantum user and the degree bound k as the capacity limit on quantum channels for each user. A feasible solution to this *MUERP* instance that spans all users under the capacity constraints maps to a valid degree-bounded minimum spanning tree for the original graph. Since the *DCMST* is NP-hard, the *MUERP* must also be NP-hard. □

As demonstrated in Theorem 1, there are no known efficient algorithms to find *feasible* solutions in polynomial time unless P=NP. Theorem 2 indicates the same conclusion for *optimal* solutions. It underscores the significant computational challenge in efficiently solving the *MUERP*.

IV. Entanglement Routing Algorithm Design

In this section, we first present an algorithm to address the challenge of *the product objective*. This algorithm seeks to identify quantum channels with the maximum entanglement rate between a fixed pair of quantum users. Subsequently, this algorithm will serve as a basic function in the design of entanglement routing algorithms.

Next, we introduce three entanglement routing algorithms. In the first algorithm, we explore a special case of the original problem, assuming that the number of qubits in a switch is equal to or greater than twice the number of quantum users, i.e., $Q_{v_i} \geq 2|\mathcal{U}|$. This is a sufficient condition ensuring that switches have enough capacities to serve quantum users. We demonstrate that the proposed algorithm under this condition can produce an optimal solution. Since the *MUERP* is NP-hard, and determining its feasible solution existence is NP-complete without this condition, we further propose two heuristic algorithms to find solutions of the *MUERP*.

A. Find a Quantum Channel with Maximum Entanglement Rate

In this subsection, we propose an algorithm to find a quantum channel with the maximal entanglement rate between a pair of quantum users. This algorithm will be a basic function utilized in the subsequent design of entanglement routing algorithms. The challenge is that Eq. (1) is not the linear summation of the values of single quantum links. Therefore, existing classic algorithms (e.g., Dijkstra's algorithm [42]) cannot be applied directly.

To address this, we use the logarithm of Eq. (1), allowing us to add the logarithms of product terms consecutively. It is because $\ln(\prod_{i=1}^{T} t_i) = \sum_{i=1}^{T} \ln t_i$. As a result, we transfer each term $t_i \in [0, 1]$ to $-\ln(t_i) \in [0, +\infty]$. This transformation allows us to use shortest-path-related methods to find the channels with maximum entanglement rates.

We design Algorithm 1 to compute the channel with maximum entanglement rate $RATE_{u_i,u_j}$ between two quantum users, u_i and u_j. The algorithm consists of multiple rounds. At the beginning of each round, it selects an unvisited node u_k with the maximum entanglement rate from u_i. Then, the algorithm attempts to improve the entanglement rate (see Line 12 to Line 14). The algorithm repeats the process until no further improvements are possible. Through the $Prev$ array, the algorithm constructs a quantum channel by tracing the previous nodes.

Algorithm 1 Maximum Entanglement Rate of a Channel

Input: $G = (\mathcal{V} = (\mathcal{U} \cup \mathcal{R}), \mathcal{E}), < u_i, u_j >$
Output: $A_{u_i,u_j}, RATE_{u_i,u_j}$
1: $\mathcal{A} \leftarrow \emptyset, RATE_{u_i,u_j} \leftarrow 0$
2: $Dist_{u_k} \leftarrow \infty, \forall u_k \in \mathcal{V}$
3: $Dist_{u_i} \leftarrow 0$
4: $Visit_{u_k} \leftarrow 0, \forall u_k \in \mathcal{V}$
5: **while** $\exists u_k, s.t. Visit_{u_k} = 0$ **do**
6: Select u_k with minimum $Dist_{u_k}$ and $Visit_{u_k} = 0$
7: **if** $Dist_{u_k} = \infty$ **then**
8: Break
9: **end if**
10: $Visit_{u_k} \leftarrow 1$
11: **for all** $e_{u_k,u_h} \in \mathcal{E}, Q_{u_h} \geq 2$ **do**
12: **if** $Dist_{u_h} > Dist_{u_k} + \alpha L_{u_k,u_h} - \ln q$ **then**
13: $Dist_{u_h} \leftarrow Dist_{u_k} + \alpha L_{u_k,u_h} - \ln q$
14: $Prev_{u_h} \leftarrow u_k$
15: **end if**
16: **end for**
17: **end while**
18: **if** $Dist_{u_j} = \infty$ **then**
19: No valid channel from u_i to u_j
20: **else**
21: $Pos \leftarrow u_j$
22: $A_{u_i,u_j} \leftarrow \{u_j\}$
23: **while** $Pos \neq u_i$ **do**
24: $Pos \leftarrow Prev_{Pos}$
25: $A_{u_i,u_j} \leftarrow A_{u_i,u_j} \cup Pos$
26: **end while**
27: $RATE_{u_i,u_j} \leftarrow \exp(-\ln q - Dist_{u_j})$
28: **end if**

After selecting the channels, their product yields the entanglement rate of MUERP. By extracting the $-\ln q$ item from each channel, we can assign the weight of each edge to $\alpha L_{u_k,u_h} - \ln q$ to constitute the sum of all edge weights of a channel as the channel length. When obtaining the final result, the weight of all selected x channels is set to $Total$, and the entanglement rate of the corresponding solution is $\exp(-x \ln q - Total)$. In the following algorithms and discussions, when we refer to "the weight of MUERP", the weight of the solution should be transformed through the formula above into the corresponding formal solution that maximizes the entanglement rate.

B. A Special Case Study: A Sufficient Condition with the Optimal Solution

As the *MUERP* is NP-Hard and determining its feasible solution existence problem is NP-Complete, we first consider a special case assuming $Q_i \geq 2|\mathcal{U}|, \forall i \in \mathcal{R}$. This condition can ensure that any switch has sufficient capacity to serve the entanglement among quantum users when the channels of all pairs of quantum users pass through this switch.

We propose a two-step approach in Algorithm 2 to address this problem. First, the algorithm finds all potential quantum channels with maximum entanglement rates between all pairs of quantum users. Second, the algorithm selects channels to span quantum users as an entanglement tree, while maximizing the entanglement rate.

Algorithm 2 Optimal Algorithm Design under a Special Case

Input: $G = (\mathcal{V} = (\mathcal{U} \cup \mathcal{R}), \mathcal{E})$
Output: $\mathcal{A}, MaxRate$
1: $\mathcal{A} = \emptyset, MaxRate \leftarrow 1$
2: All quantum users are not in the same union
3: **for all** $< u_i, u_j >, u_i, u_j \in \mathcal{U}, i \neq j$ **do**
4: Find the channel A_{u_i,u_j} with the maximum entanglement rate $RATE_{u_i,u_j}$
5: **end for**
6: Sort all A_{u_i,u_j} in descending order of $RATE_{u_i,u_j}$
7: **for all** A_{u_i,u_j} in descending order of $RATE_{u_i,u_j}$ **do**
8: **if** u_i and u_j are not in the same union **then**
9: Merge u_i and u_j into the same union
10: $MaxRate \leftarrow MaxRate \times RATE_{u_i,u_j}$
11: $\mathcal{A} \leftarrow \mathcal{A} + A_{u_i,u_j}$
12: **end if**
13: **end for**

The first step, outlined in Lines 1 to 5 in Algorithm 2, involves finding all quantum channels between each pair of quantum users with the maximum entanglement rate based on Algorithm 1. The resultant channels form a set containing all quantum channels with the maximum entanglement rate for all potential user pairs $u_i, u_j \in \mathcal{U}$.

In the second step, Algorithm 2 selects channels from $\mathcal{A}$ to form a tree spanning users in $\mathcal{U}$ while maximizing the entanglement rate. Line 6 sorts all channels in descending order of the entanglement rate. The algorithm proceeds by adding channels, beginning with the one with the maximum entanglement rate. More specifically, if a newly added channel connects quantum users that are not connected in the current channel set, the algorithm incorporates it into the set and updates the connectivity relationships among quantum users. We utilize a union-find data structure [46] to maintain the connectivity of different quantum users. This data structure aids the algorithm in determining if two quantum users are connected within the same union.

Theorem 3. *When $Q_{v_i} \geq 2|\mathcal{U}|, \forall v_i \in \mathcal{R}$, Algorithm 2 outputs the optimal solution of the MUERP.*

Proof. Due to the space limitation, a brief proof is provided. Given that each switch has sufficient capacity for the entanglement of all quantum users and users are originally connected through optical fibers, a feasible solution must exist where all quantum users can be spanned by channels forming a tree.

We use a contradiction method to prove that the output of Algorithm 2 is the optimal solution. Suppose there exists a solution with a larger entanglement rate. Given all quantum users are connected in a tree structure by quantum channels, removing any quantum link from a channel would disrupt the connectivity between a pair of quantum users. Therefore, there must be a quantum channel with a higher entanglement rate that was not included in the set and connects a pair of users. However, in Algorithm 2, the channels are selected in descending order of the entanglement rate, which contradicts the assumption that a superior solution exists. □

Theorem 3 presents a loose bound of the switch capacity to ensure a feasible solution exists for the *MUERP*. This is an interesting direction to find a necessary and sufficient condition that tightens this boundary for solution existence. Based on the condition, Theorem 3 proves that Algorithm 2 can produce the optimal solution of the *MUERP* while all nodes have enough qubits for all quantum pairs.

Time Complexity. The time complexity of the first step is $O(|\mathcal{U}|^2(|\mathcal{E}| + |\mathcal{V}| \log |\mathcal{V}|))$ initially. To optimize this, we can run Algorithm 1 once for each source node u_i rather than for all pairs $< u_i, u_j >$. After running for a fixed u_i, we can recover the channels A_{u_i,u_j} for all destinations u_j through the $Prev$ array. This reduces the time complexity of the first step to $O(|\mathcal{U}|(|\mathcal{E}| + |\mathcal{V}| \log |\mathcal{V}|))$. The second step has a time complexity of $O(|\mathcal{U}|^2 \log |\mathcal{U}|^2)$ due to the sort operation on all A_{u_i,u_j} and the union-find data structure operations [46]. Since $|\mathcal{U}| \log |\mathcal{U}| \leq |\mathcal{V}| \log |\mathcal{V}|$, the overall time complexity of Algorithm 1 simplifies to $O(|\mathcal{U}|(|\mathcal{E}| + |\mathcal{V}| \log |\mathcal{V}|))$.

C. Heuristic Approach I: Dealing Capacity Conflicts

In this subsection, we provide a heuristic algorithm of the *MUERP* based on the solution of Algorithm 2.

The key distinction and challenge here arise from the capacity limitation of switches, implying that a switch might lack sufficient capacity to facilitate entanglement among quantum users. Building upon the entanglement routing design from Algorithm 2, we develop Algorithm 3 to address the conflicts arising from switches exceeding their capacity.

In Algorithm 3, the procedure begins by examining each switch to determine if the number of quantum channels passing through it exceeds the switch's capacity. If that's the case, some of the passing quantum channels need to be removed, and the corresponding users will have to use alternative channels for entanglement. Similarly to Algorithm 2, we utilize a union-find data structure [46] to maintain the connectivity of different quantum users. After the removal of certain channels, the quantum users may be split into several unions, necessitating additional channels to reconnect the divided unions.

To find these channels, the algorithm calculates the possible channels that can connect two distinct unions, and adds the most suitable channel, until all quantum users are part of the same union. Moreover, if the algorithm is unable to find feasible channels to connect unions with users in separate unions, it will consequently terminate.

There are two key decision-making processes. The first is to determine which channel should be removed when a switch exceeds its capacity. The second is to decide which channel should be employed when connecting two separate unions. Since one of the primary objectives of our algorithm is to maximize the entanglement rate, we adopt a greedy strategy that always opts to retain the channel with the maximum entanglement rate. This implies that we keep the channels with the maximum entanglement rate, and we utilize the channels

Algorithm 3 Conflict-free Algorithm

Input: $G = (\mathcal{V} = (\mathcal{U} \cup \mathcal{R}), \mathcal{E}), \mathcal{Q}, \mathcal{A}$
Output: $\mathcal{A}', MaxRate$

$\mathcal{A}' \leftarrow \emptyset, MaxRate \leftarrow 1$
All quantum users are not in the same union
Sort $A_{u_i,u_j} \in \mathcal{A}$ in descending order of entanglement rate $RATE_{u_i,u_j}$
for all $A_{u_i,u_j} \in \mathcal{A}$ **do**
  if $Q_{r_k} \geq 2, \forall r_k \in A_{u_i,u_j}\ \&\ r_k \in \mathcal{R}$ **then**
    for all $r_k \in A_{u_i,u_j}\ \&\ r_k \in \mathcal{R}$ **do**
      $Q_{r_k} \leftarrow Q_{r_k} - 2$
    end for
    Merge u_i and u_j into the same union
    $MaxRate \leftarrow MaxRate \times RATE_{u_i,u_j}$
    $\mathcal{A}' \leftarrow \mathcal{A}' \cup A_{u_i,u_j}$
  else
    u_i and u_j are not in the same union
  end if
end for
while $\exists u_i, u_j \in \mathcal{U}$ are not in the same union **do**
  $CurrentRate \leftarrow 0$
  for all $u_i, u_j \in \mathcal{U}$ are not in the same union **do**
    Find the channel A_{u_i,u_j} with maximum entanglement rate between u_i, u_j
    if $RATE_{u_i,u_j} > CurrentRate$ **then**
      $CurrentRate \leftarrow RATE_{u_i,u_j}, u_1 \leftarrow u_i, u_2 \leftarrow u_j$
    end if
  end for
  if $CurrentRate = 0$ **then**
    Cannot find a feasible entanglement tree, terminate the algorithm
  end if
  for all $r_k \in A_{u_1,u_2}\ \&\ r_k \in \mathcal{R}$ **do**
    $Q_{r_k} \leftarrow Q_{r_k} - 2$
  end for
  Merge u_i and u_j into the same union
  $MaxRate \leftarrow MaxRate \times CurrentRate$
  $\mathcal{A}' \leftarrow \mathcal{A}' \cup A_{u_1,u_2}$
end while

with the maximum entanglement rate to connect different connected unions.

Time Complexity. The time complexity of the sorting process is $O(|\mathcal{U}|^2 \log |\mathcal{U}|^2)$. The subsequent algorithm identifies $|\mathcal{U}| - 1$ channels. Before finding each channel, the algorithm recalculates all potential channels between quantum users who are not in the same union. The time complexity for recalculating the maximum entanglement rate channels is $O(|\mathcal{U}|(|\mathcal{E}| + |\mathcal{V}| \log |\mathcal{V}|))$. Thus, the overall time complexity is $O(|\mathcal{U}|^2(|\mathcal{E}| + |\mathcal{V}| \log |\mathcal{V}|))$.

D. Heuristic Approach II: Prim-based

Algorithm 3 needs the output of Algorithm 2 as the input to find a solution. In this subsection, we propose Algorithm 4 based on the principle of Prim Algorithm [42] to find a solution directly. The Prim Algorithm is typically proposed to find solutions for combinatorial problems which are NP-Hard or NP-Complete [42].

Initially, we have two user sets, $\mathcal{U}_1$ and $\mathcal{U}_2$. We randomly select a quantum user u_i to be included in $\mathcal{U}_1$, and define $\mathcal{U}_2$ as the remaining quantum users, i.e., $\mathcal{U}_2 = \{\mathcal{U} \backslash u_i\}$. The algorithm then repeats the following process for $|\mathcal{U}| - 1$ rounds:

Algorithm 4 Prim-based Algorithm

Input: $G = (\mathcal{V} = (\mathcal{U} \cup \mathcal{R}), \mathcal{E}), \mathcal{Q}$
Output: $\mathcal{A}'', MaxRate$

$\mathcal{A}'' \leftarrow \emptyset, MaxRate \leftarrow 1$
Randomly pick $u_0 \in \mathcal{U}$
$\mathcal{U}_1 \leftarrow \{u_0\}, \mathcal{U}_2 \leftarrow \mathcal{U} \backslash u_0$
while $\mathcal{U}_2 \neq \emptyset$ **do**
  $CurrentRate \leftarrow 0$
  for all $< u_i, u_j >, u_i \in \mathcal{U}_1, u_j \in \mathcal{U}_2$ **do**
    Find the channel A_{u_i,u_j} with maximum entanglement rate between u_i, u_j
    if $RATE_{u_i,u_j} > CurrentRate$ **then**
      $CurrentRate \leftarrow RATE_{u_i,u_j}, u_1 \leftarrow u_i, u_2 \leftarrow u_j$
    end if
  end for
  if $CurrentRate = 0$ **then**
    Cannot find a feasible entanglement tree, terminate the algorithm
  end if
  for all $r_k \in A_{u_1,u_2}\ \&\ r_k \in \mathcal{R}$ **do**
    $Q_{r_k} \leftarrow Q_{r_k} - 2$
  end for
  $MaxRate \leftarrow MaxRate \times CurrentRate$
  $\mathcal{U}_1 \leftarrow \mathcal{U}_1 + u_2, \mathcal{U}_2 \leftarrow \mathcal{U}_2 - u_2$
  $\mathcal{A}'' \leftarrow \mathcal{A}'' \cup A_{u_1,u_2}$
end while

During each round, the algorithm calculates the entanglement channel with the maximum entanglement rate from each user in $\mathcal{U}_1$ to each user in $\mathcal{U}_2$. When finding these channels, any channel involving switches without enough qubits is excluded. Assuming the quantum channel with the maximum entanglement rate connects $u_i \in \mathcal{U}_1$ and $u_j \in \mathcal{U}_2$, the algorithm transfers u_j from $\mathcal{U}_2$ to $\mathcal{U}_1$, and subtracts the corresponding qubits from the switches in the channel.

After $|\mathcal{U}| - 1$ iterations, $\mathcal{U}_1$ contains all quantum users, i.e., $\mathcal{U}_1 = \mathcal{U}$, and $\mathcal{U}_2$ is empty. This means that all quantum users are connected through quantum channels. If the algorithm cannot find a feasible solution, it will terminate.

Time Complexity. The algorithm identifies $|\mathcal{U}| - 1$ channels across $|\mathcal{U}| - 1$ rounds. In each round, the algorithm recalculates the channels with the maximum entanglement rate. Thus, similar to Algorithm 3, the total time complexity is $O(|\mathcal{U}|^2(|\mathcal{E}| + |\mathcal{V}| \log |\mathcal{V}|))$.

V. SIMULATION RESULTS

In this section, we conduct a comprehensive set of simulations to evaluate the performance of our proposed algorithms. We manipulate multiple parameters within these simulations to enhance their reliability.

A. Simulation Setup

Network Topology: To evaluate the efficiency of our proposed algorithms, we employ three different network generation methods: Waxman method [47], Watts-Strogatz method [48], and Volchenkov method [49]. These methods are widely utilized to create random networks mirroring the complexity and topology of real-world networks, such as the Internet [12]. Our quantum network covers an area of $10k \times 10k$ square units, with each unit approximating 1 kilometer [12]. Switches and quantum users are placed randomly within this area. By default, we select the Waxman method [47] for network generation and configure the network with 50 switches and 10 quantum users. We determine the total number of edges based on an average degree D of nodes, set to 6. Each switch has 4 qubits, with a successful swapping rate of 0.9. The simulation's metric is the entanglement rate of users, as defined in Eq. (2). If a channel in the entanglement tree cannot be established due to network constraints, the entanglement rate becomes zero. To reduce the impact of network topology randomness, we generate 20 random networks and compute the average of the observed results. The constant dependent on the physical material, α, is set to 10^{-4} [12].

Comparative Benchmarks:

- Extended Q-CAST (E-Q-CAST): In [12], the algorithm only considers pairs of users, making it unfeasible for multi-user cases. Therefore, We extended the algorithm in [12] incorporating additional pairs to ensure connectivity. For example, we establish entanglement channels $< u_1, u_2 >$, $< u_2, u_3 >$, $< u_3, u_4 >$ to entangle $\{u_1, u_2, u_3, u_4\}$.
- N-FUSION: We consider the MP-P algorithm in Ref. [32] that covers the cases in Ref. [31], [33]. The main difference between N-FUSION and MP-P is the switch capacity. Switches in N-FUSION have limited capacity, whereas those in MP-P possess infinite capacity. Therefore, N-FUSION considers a central user connecting all users (like Tree B in Figure 3 of Ref. [32]).

B. Results

Based on our simulation results, our proposed algorithms excel across different topologies and surpass current methods. In particular, Algorithms 2, 3, and 4 can boost the entanglement rate by up to 5347%, 3180%, and 3155% respectively when compared to N-FUSION, and by 5068%, 3014%, and 2990% respectively when compared to E-Q-CAST. The detailed discussion is as follows.

Impact of network topology: Fig. 5 presents the results under different topologies. We observe variations in algorithm performance across different network topologies. For instance, N-FUSION fails to entangle users in the graphs generated by the Watts-Strogatz method. This implies that the network topology has a significant impact on the entanglement, presenting an intriguing area that needs further in-depth exploration. Meanwhile, the proposed algorithms surpass the performance of existing baselines. We can therefore conclude that our proposed algorithm adapts well to a variety of networks, enhancing entanglement efficiency.

Impact of the network scale: Fig. 6 and Fig. 7 display the effects of network scale. In Fig. 6a, the entanglement rate decreases as the number of users to be entangled increases. This can be attributed to the need for more quantum channels to entangle additional users, thus lowering the product

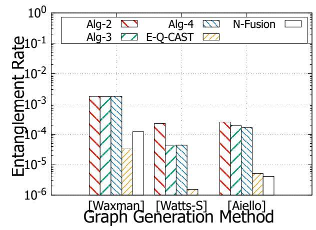
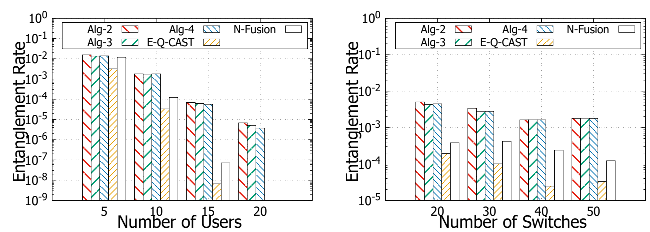


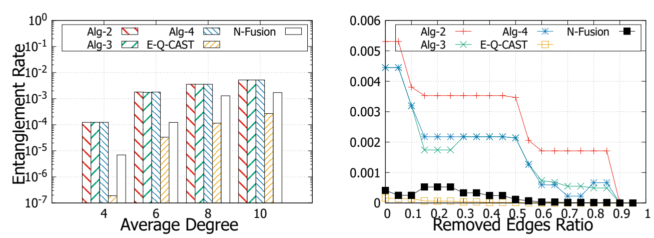
Fig. 5. Entanglement rate v.s. Network topology.

of entanglement rates of channels. Fig. 7a shows that the entanglement rate increases when the average degree of a switch increases, as the network provides a wider selection for channel assignment. Contrarily, in Fig. 6b, the entanglement rate declines as the number of switches increases. This is due to quantum channels having to pass through more switches, consequently reducing the entanglement rate. When the number of switches increases from 40 to 50, the entanglement rate for part of the algorithms increases. This is because the length of the channel does not increase significantly with switches increasing from 40 to 50. Instead, it provides more potential quantum channels, which in turn boosts the entanglement rate.



(a) (b)

Fig. 6. (a) Entanglement rate v.s. The number of switches. (b) Entanglement rate v.s. The number of users.



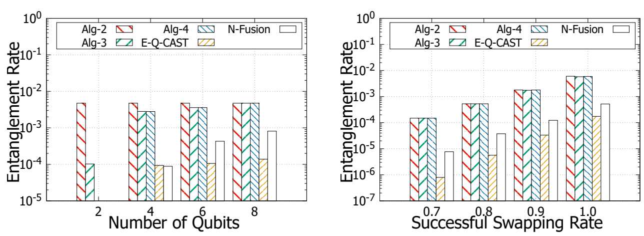
(a) (b)

Fig. 7. (a) Entanglement rate v.s. The average degree of a switch. (b) Entanglement rate v.s. Removed edges ratio.

In Fig. 7b, we examine the edge (optical fiber) boundary condition necessary to support entanglement. We construct a graph with 10 quantum users, 50 switches, and 600 optical fibers, with each quantum switch holding 4 qubits. We proceed by randomly removing 30 optical fibers from the graph and repeating this process until no feasible quantum channels remain that can entangle all quantum users.

We observe that: (1) In most cases, the entanglement rate reduces as more optical fibers are removed. (2) Each algorithm's results can stay constant when some optical fibers are removed, as they depend on a few 'critical' edges. Even with a 5% reduction, the outcome won't alter if these crucial ones are preserved. (3) Algorithm performance might improve after removing certain optical fibers. This can occur when these optical fibers lead the algorithm to less efficient quantum channels. Removing such ones could enable the use of more optimized quantum channels, enhancing overall results.

Impact of the switch: Figure 8 illustrates the impact of the quantum switch. In Fig. 8a, We vary the number of qubits in the switch Q_i from 2 to 8. Algorithm 2 is not constrained by this. The switches in Algorithm 2 has $2|\mathcal{U}| = 20$ qubits. It's observed that when $Q_i = 2$, Algorithm 3 is the only one capable of supporting entanglement. As Q_i increases to 6, the network has sufficient qubits for Algorithm 3 and Algorithm 4 to serve entanglement. The entanglement rate of the two baselines continues to rise when $Q_i = 8$. This suggests that our proposed algorithms can utilize the qubits in switches more effectively. Fig. 8b tests the impact of the successful swapping rate of the switch q. As q increases, so does the entanglement rate, indicating that a more reliable switch with a higher swapping rate can enhance the reliability of user entanglement.



(a) (b)

Fig. 8. (a) Entanglement rate v.s. The number of qubits in a switch. (b) Entanglement rate v.s. The successful entanglement-swapping rate of a switch.

VI. Related Work

The entanglement routing problem has been drawing great attention in recent years. The majority of research focuses on two-user entanglement routing. There are two primary types of these studies.

1. Studies that assume perfect, noise-free switches. Pant *et al.* [11] first proposed the entanglement routing problem, primarily focusing on routing protocols for one quantum-user pair. Shi and Qian [12] developed heuristic algorithms for multiple quantum user pairs by sequentially selecting the user pair with the highest throughput. Zhao *et al.* [18] designed two routing protocols to distribute qubits in quantum data networks. Farahbakhsh and Feng [14] proposed an asynchronous entanglement routing method that does not require simultaneous entanglement generation along the path.

2. Studies that consider imperfect switches with noise. These research efforts consider fidelity as the constraint limiting the length of quantum channels, with different optimization objectives and specific network topologies. For example, Li *et al.* [15] considered a lattice network to optimize various routing metrics such as fairness and the entanglement rate. Liu *et al.* [16] focused on the design of entanglement protocols for communication. Ghaderibaneh *et al.* [17] examined a tree structure to determine the swapping policy. Ref. [18] and Ref. [19] aimed to maximize the throughput of multiple quantum user pairs. The two-user entanglement paradigm prevalent in these studies is primarily limited to quantum communication. As such, it falls short of addressing the numerous quantum applications that require multi-user entanglement.

Some papers consider multi-user entanglement routing to form GHZ states among users. Bugalho *et al.* [31] proposed a protocol for distributing a 3 qubit Greenberger–Horne–Zeilinger (GHZ) state, assuming that a switch possesses precisely 3 qubits. In [32], Sutcliffe and Beghelli presented protocols to increase the successful entanglement rate of multi-user applications by leveraging multi-path routing. Avis *el al.* [33] analyzed the performance of a single switch that connects multiple users to form a GHZ state. However, these studies tend to oversimplify the network model under consideration and rely on an unreliable entanglement-swapping method, $n-$fusion.

VII. Conclusion

In this paper, we consider the multi-user entanglement routing problem in the quantum Internet, aiming to maximize the entanglement rate. This issue is modeled as a novel graph routing problem to capture the fundamental essence of the quantum Internet. As determining the existence of a feasible solution to the problem is NP-Complete, we first propose a sufficient condition guaranteeing the existence of a feasible solution, followed by an algorithm offering an optimal solution. Given the NP-Hardness of the problem, we propose two heuristic algorithms for effective solution derivation. Our simulations demonstrate the superior performance of these algorithms over existing ones. This work provides fundamental insights for designing efficient multi-user entanglement routing in the quantum Internet under complex scenarios, such as accounting for fidelity, or simultaneous routing of multiple independent entanglement groups. The developed algorithms can serve as a foundation for other related quantum networking research, including quantum mapping and the architectural design of the quantum Internet. This will facilitate the practical implementation of large-scale, multi-user quantum applications in the future.

Acknowledgment

This work is supported in part by the US National Science Foundation under grant numbers 1717731, 1730291, 2231040, 2230620, 2214980, 2046444, 2106027, and 2146909.

References

[1] P. W. Shor, "Algorithms for quantum computation: discrete logarithms and factoring," in *Proceedings 35th annual symposium on foundations of computer science*. IEEE, 1994, pp. 124–134.

[2] A. W. Harrow, A. Hassidim, and S. Lloyd, "Quantum algorithm for linear systems of equations," *Physical review letters*, vol. 103, no. 15, p. 150502, 2009.

[3] N. Gisin and R. Thew, "Quantum communication," *Nature photonics*, vol. 1, no. 3, pp. 165–171, 2007.
[4] Y.-A. Chen, Q. Zhang, T.-Y. Chen, W.-Q. Cai, S.-K. Liao, J. Zhang, K. Chen, J. Yin, J.-G. Ren, Z. Chen *et al.*, "An integrated space-to-ground quantum communication network over 4,600 kilometres," *Nature*, vol. 589, no. 7841, pp. 214–219, 2021.
[5] R. Valivarthi, S. I. Davis, C. Peña, S. Xie, N. Lauk, L. Narváez, J. P. Allmaras, A. D. Beyer, Y. Gim, M. Hussein *et al.*, "Teleportation systems toward a quantum internet," *PRX Quantum*, vol. 1, no. 2, p. 020317, 2020.
[6] M. A. Nielsen and I. L. Chuang, "Quantum computation and quantum information," *Phys. Today*, vol. 54, no. 2, p. 60, 2001.
[7] A. Dahlberg, M. Skrzypczyk, T. Coopmans, L. Wubben, F. Rozpędek, M. Pompili, A. Stolk, P. Pawełczak, R. Knegjens, J. de Oliveira Filho *et al.*, "A link layer protocol for quantum networks," in *Proceedings of the ACM special interest group on data communication*, 2019, pp. 159–173.
[8] S. Wehner, D. Elkouss, and R. Hanson, "Quantum internet: A vision for the road ahead," *Science*, vol. 362, no. 6412, p. eaam9288, 2018.
[9] J.-W. Pan, D. Bouwmeester, H. Weinfurter, and A. Zeilinger, "Experimental entanglement swapping: entangling photons that never interacted," *Physical review letters*, vol. 80, no. 18, p. 3891, 1998.
[10] M. Pompili, S. L. Hermans, S. Baier, H. K. Beukers, P. C. Humphreys, R. N. Schouten, R. F. Vermeulen, M. J. Tiggelman, L. dos Santos Martins, B. Dirkse *et al.*, "Realization of a multinode quantum network of remote solid-state qubits," *Science*, vol. 372, no. 6539, pp. 259–264, 2021.
[11] M. Pant, H. Krovi, D. Towsley, L. Tassiulas, L. Jiang, P. Basu, D. Englund, and S. Guha, "Routing entanglement in the quantum internet," *npj Quantum Information*, vol. 5, no. 1, pp. 1–9, 2019.
[12] S. Shi and C. Qian, "Concurrent entanglement routing for quantum networks: Model and designs," in *Proceedings of the Annual conference of the ACM Special Interest Group on Data Communication on the applications, technologies, architectures, and protocols for computer communication*, 2020, pp. 62–75.
[13] Y. Zhao and C. Qiao, "Distributed transport protocols for quantum data networks," *IEEE/ACM Transactions on Networking*, 2023.
[14] A. Farahbakhsh and C. Feng, "Opportunistic routing in quantum networks," in *IEEE INFOCOM 2022-IEEE Conference on Computer Communications*. IEEE, 2022, pp. 490–499.
[15] 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, p. 10, 2021.
[16] M. Liu, J. Allcock, K. Cai, S. Zhang, and J. C. Lui, "Quantum networks with multiple service providers: Transport layer protocols and research opportunities," *IEEE Network*, vol. 36, no. 5, pp. 56–62, 2022.
[17] M. Ghaderibaneh, C. Zhan, H. Gupta, and C. Ramakrishnan, "Efficient quantum network communication using optimized entanglement swapping trees," *IEEE Transactions on Quantum Engineering*, vol. 3, pp. 1–20, 2022.
[18] Y. Zhao, G. Zhao, and C. Qiao, "E2e fidelity aware routing and purification for throughput maximization in quantum networks," in *IEEE INFOCOM 2022-IEEE Conference on Computer Communications*. IEEE, 2022, pp. 480–489.
[19] 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.
[20] T. Aoki, G. Takahashi, T. Kajiya, J.-i. Yoshikawa, S. L. Braunstein, P. Van Loock, and A. Furusawa, "Quantum error correction beyond qubits," *Nature Physics*, vol. 5, no. 8, pp. 541–546, 2009.
[21] L. Xiao, G. L. Long, F.-G. Deng, and J.-W. Pan, "Efficient multiparty quantum-secret-sharing schemes," *Physical Review A*, vol. 69, no. 5, p. 052307, 2004.
[22] J. Kempe, "Multiparticle entanglement and its applications to cryptography," *Physical Review A*, vol. 60, no. 2, p. 910, 1999.
[23] H. Zhang, A. Wu, Y. Wang, G. Li, H. Shapourian, A. Shabani, and Y. Ding, "Oneq: A compilation framework for photonic one-way quantum computation," in *Proceedings of the 50th Annual International Symposium on Computer Architecture*, 2023, pp. 1–14.
[24] M. Caleffi, M. Amoretti, D. Ferrari, D. Cuomo, J. Illiano, A. Manzalini, and A. S. Cacciapuoti, "Distributed quantum computing: a survey," *arXiv preprint arXiv:2212.10609*, 2022.
[25] X.-Y. Chang, D.-L. Deng, X.-X. Yuan, P.-Y. Hou, Y.-Y. Huang, and L.-M. Duan, "Experimental realization of an entanglement access network and secure multi-party computation," *Scientific reports*, vol. 6, no. 1, p. 29453, 2016.
[26] N. Friis, O. Marty, C. Maier, C. Hempel, M. Holzäpfel, P. Jurcevic, M. B. Plenio, M. Huber, C. Roos, R. Blatt *et al.*, "Observation of entangled states of a fully controlled 20-qubit system," *Physical Review X*, vol. 8, no. 2, p. 021012, 2018.
[27] J. Chow, O. Dial, and J. Gambetta, "Ibm quantum breaks the 100-qubit processor barrier," *IBM Research Blog*, vol. 2, 2021.
[28] R. Shaydulin, H. Ushijima-Mwesigwa, C. F. Negre, I. Safro, S. M. Mniszewski, and Y. Alexeev, "A hybrid approach for solving optimization problems on small quantum computers," *Computer*, vol. 52, no. 6, pp. 18–26, 2019.
[29] N. H. Nguyen, E. C. Behrman, and J. E. Steck, "Quantum learning with noise and decoherence: a robust quantum neural network," *Quantum Machine Intelligence*, vol. 2, no. 1, pp. 1–15, 2020.
[30] B. Bauer, S. Bravyi, M. Motta, and G. K.-L. Chan, "Quantum algorithms for quantum chemistry and quantum materials science," *Chemical Reviews*, vol. 120, no. 22, pp. 12 685–12 717, 2020.
[31] L. Bugalho, B. C. Coutinho, F. A. Monteiro, and Y. Omar, "Distributing multipartite entanglement over noisy quantum networks," *quantum*, vol. 7, p. 920, 2023.
[32] E. Sutcliffe and A. Beghelli, "Multi-user entanglement distribution in quantum networks using multipath routing," *arXiv preprint arXiv:2303.03334*, 2023.
[33] G. Avis, F. Rozpędek, and S. Wehner, "Analysis of multipartite entanglement distribution using a central quantum-network node," *Physical Review A*, vol. 107, no. 1, p. 012609, 2023.
[34] M. Ghaderibaneh, H. Gupta, and C. Ramakrishnan, "Generation and distribution of ghz states in quantum networks," in *2023 IEEE International Conference on Quantum Computing and Engineering (QCE)*, vol. 1. IEEE, 2023, pp. 1120–1131.
[35] V. Mannalath and A. Pathak, "Multiparty entanglement routing in quantum networks," *Physical Review A*, vol. 108, no. 6, p. 062614, 2023.
[36] A. Patil, M. Pant, D. Englund, D. Towsley, and S. Guha, "Entanglement generation in a quantum network at distance-independent rate," *npj Quantum Information*, vol. 8, no. 1, p. 51, 2022.
[37] C.-Y. Lu, T. Yang, and J.-W. Pan, "Experimental multiparticle entanglement swapping for quantum networking," *Physical review letters*, vol. 103, no. 2, p. 020501, 2009.
[38] A. Borras, A. Plastino, J. Batle, C. Zander, M. Casas, and A. Plastino, "Multiqubit systems: highly entangled states and entanglement distribution," *Journal of Physics A: Mathematical and Theoretical*, vol. 40, no. 44, p. 13407, 2007.
[39] J.-W. Pan, Z.-B. Chen, C.-Y. Lu, H. Weinfurter, A. Zeilinger, and M. Żukowski, "Multiphoton entanglement and interferometry," *Reviews of Modern Physics*, vol. 84, no. 2, p. 777, 2012.
[40] W. Dür and H.-J. Briegel, "Stability of macroscopic entanglement under decoherence," *Physical review letters*, vol. 92, no. 18, p. 180403, 2004.
[41] G. B. Xavier and G. Lima, "Quantum information processing with space-division multiplexing optical fibres," *Communications Physics*, vol. 3, no. 1, p. 9, 2020.
[42] D. B. West *et al.*, *Introduction to graph theory*. Prentice hall Upper Saddle River, 2001, vol. 2.
[43] F. K. Hwang and D. S. Richards, "Steiner tree problems," *Networks*, vol. 22, no. 1, pp. 55–89, 1992.
[44] J. Hartmanis, "Computers and intractability: A guide to the theory of np-completeness (michael r. garey and david s. johnson)," *SIAM Review*, vol. 24, no. 1, pp. 90–91, 1982.
[45] T. N. Bui and C. M. Zrncic, "An ant-based algorithm for finding degree-constrained minimum spanning tree," in *Proceedings of the 8th annual conference on Genetic and evolutionary computation*, 2006, pp. 11–18.
[46] S. Conchon and J.-C. Filliâtre, "A persistent union-find data structure," in *Proceedings of the 2007 workshop on Workshop on ML*, 2007, pp. 37–46.
[47] B. M. Waxman, "Routing of multipoint connections," *IEEE journal on selected areas in communications*, vol. 6, no. 9, pp. 1617–1622, 1988.
[48] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world'networks," *nature*, vol. 393, no. 6684, pp. 440–442, 1998.
[49] D. Volchenkov and P. Blanchard, "An algorithm generating random graphs with power law degree distributions," *Physica A: Statistical Mechanics and its Applications*, vol. 315, no. 3-4, pp. 677–690, 2002.