Multi-Entanglement Routing Design over Quantum Networks

Yiming Zeng[‡], Jiarui Zhang[‡], Ji Liu, Zhenhua Liu, Yuanyuan Yang

Stony Brook University, Stony Brook, NY 11794, USA

{yiming.zeng, jiarui.zhang.2, zhenhua.liu, yuanyuan.yang}@stonybook.edu

Abstract—**Quantum networks are considered as a promising future platform for quantum information exchange and quantum applications, which have capabilities far beyond the traditional communication networks. Remote quantum entanglement is an essential component of a quantum network. How to efficiently design a multi-routing entanglement protocol is a fundamental yet challenging problem. In this paper, we study a quantum entanglement routing problem to simultaneously maximize the number of quantum-user pairs and their expected throughput. Our approach is to formulate the problem as two sequential integer programming steps. We propose efficient entanglement routing algorithms for the two integer programming steps and analyze their time complexity and performance bounds. Results of evaluation highlight that our approach outperforms existing solutions in both served quantum-user pairs numbers and the network expected throughput.**

Index Terms—**Quantum Networks; Entanglement Routing; Integer Programming**

I. Introduction

Quantum networks enable to generate, transmit and compute quantum information (qubits) in addition to classical data between quantum (ebits) processors [1]. It supports massive quantum applications in both quantum computing and quantum communication systems, such as distributed quantum computing [2], [3], quantum communication [4], quantum machine learning [5] and quantum key distribution [6]. Several quantum systems have been constructed, such as long-distance link (40 kilometers) teleportation over the fiber link [7], the mobile quantum network [8], the integrated entanglement system through satellites which can support the entanglement over 4600 kilometers [9].

Entanglement is an essential component of most quantum applications mentioned above. For example, the quantum key distribution system has provable security for the distributed information [1] by taking advantage of the entanglement and no-cloning theorem [10]. Supporting the long-distance entanglement is fundamental for quantum networks. However, the entanglement process is probabilistic and not stable. Different from binary ebits in traditional communication, qubits created by photons are extremely fragile. The successful entanglement rate among qubits decreases exponentially with the transmission length. Hence, to enable long-distance entanglement of quantum users in the quantum network, quantum switches are placed in the network as relays to supply end-to-end entanglements for multiple quantum users that demand

them [11], [12]. Quantum switches are equipped with quantum memories (qubits) and have the ability to perform multi-qubits measurement (swapping) [12].

The *entanglement routing* problem about *how to build long-distance entanglement through quantum switches* is crucial in the quantum network. Thoughtful design for the entanglement routing in the quantum network can boost the network performance by efficiently utilizing resources, e.g., switch memories.

While large-scale quantum networks have not been implemented out of the lab due to physical and experimental challenges, it is still valuable to investigate the entanglement routing problem from the network layer for the future. The entanglement routing problem has been drawing great attention in previous studies. [12]–[17] study the entanglement routing problem or theoretical entanglement performance on the special network topologies such as a single switch, single entanglement path, rings, grids, or spheres. [18], [19] consider a general quantum network for multiple quantum users pairs entanglement. However, their strategy is a greedy algorithm to maximize the throughput of the quantum user pair one by one which might assign too many resources to limited quantum users, and other quantum users are neglected. The proposed algorithm incurs high time complexity and lacks the performance guarantee.

Moreover, most existing works treat the transmission link capacity as the main bottleneck of the network. However, the switch resource (the number of qubits) is the limitation of the quantum network in reality instead of the transmission link capacity. A most recent quantum processor can only have up to 8 qubits [20]. An optical fiber cable can contain up to 25 cores, each core can be used as an independent link for the entanglement. Multiple optical fiber cables can be placed between quantum switches. Hence, the transmission link has enough capacity to serve the entanglement demands for the quantum users in current quantum networks.

In this paper, we consider a general quantum network structure and present a comprehensive entanglement process for multiple pairs of quantum users. Our goal is to *maximize the number of quantum-user pairs and the expected network throughput at the same time*. Our contributions are as follows:

1) We describe the detailed multi-entanglement routing process for multiple quantum-user pairs as the offline and the online stage.
2) We formulate the problem as two integer linear programming problems that both are NP-Complete.

[‡] Both authors contributed equally to this research.

3) We design the routing protocol by proposing efficient algorithms with lower time complexity and performance guarantees.
4) Results of evaluation highlight that our approach can improve the number of served quantum-user pairs 85% and the expected throughput 27% in average compared with existing works.

To the best of our knowledge, this is the first paper to maximize the network served quantum-user pairs number and expected throughput simultaneously.

The organization of the paper is as follows: we first introduce the background of the quantum network and the multi-entanglement routing process in Section II. Then, we present the quantum network model and formulate the routing entanglement process as two integer linear programming problems in Section III based on the routing process introduced in Section II. The entanglement routing algorithms are proposed in Section IV and Section V for two integer linear programming problems, respectively. We conduct extensive simulations to discuss and analyze the performance of our proposed algorithms and compare them with previous work in Section VI, followed by related work in Section VII the conclusion in Section VIII.

II. QUANTUM NETWORK BACKGROUND

In this section, we introduce some basic quantum network backgrounds, including quantum network components and multi-routing entanglement processes.

A. Quantum Communication

1) Qubit: In the quantum network or quantum computing, a qubit is the basic unit to represent quantum information. A qubit can be an electron or a photon or a nucleus from an atom. A qubit is described by its state [1]. Different from an ebit in the classical Internet representing 0 or 1, a qubit can present a coherent superposition of both.

2) Entanglement: Entanglement is a phenomenon that a group of qubits expresses a high correlation state which can not be explained by individual qubits states. In this paper, we consider the simplest case of two qubits entanglement which is bipartite entangled states. In the quantum physic, a simple way to entangle two independent qubits is by using CNOT gate [21]. When the entanglement qubits number is two, Bell-state measures (BSMs) can be applied to measure the entanglement.

3) Teleportation: If a pair of entanglement quibts are shared by two nodes, the secret information can be transmitted from one node to another one with the help of quantum measurement. This process is called teleportation. An example is illustrated in Figure 1.

Fig. 1. A Teleportation Example. The source node teleport a qubit by a pair of entanglement qubits.

4) Entanglement Swapping: Figure 2 presents an example of swapping. If Alice shares an entangled qubit pair (Bell pair) with the middle node Carol, and Carol shares another entangled qubit pair with Bob, Carol can teleport its qubit entangled with Alice to Bob, then Alice and Bob are entangled directly [22].

Fig. 2. A Swapping Example.

B. Quantum Network Components

With these basic concepts, we first introduce several important components in the quantum network.

1) Quantum Users: A quantum user has demands to entangle with another user in the quantum network for quantum communication. We assume that there is a direction between a pair of quantum users intending to be entangled. The user who intends to entangle with another user is called a source node. Another user who tried to be entangled is called a destination node.

2) Quantum Switches: The quantum switch is a node with quantum memories to work as relays for the entanglement process in the quantum network [12], [23]. They can either transmit qubits or establish the entanglement at distant nodes without physically sending an entangled qubit by swapping.

3) Quantum Links: Quantum links are the links used for connecting quantum switches and quantum users. In this paper, we assume that the quantum network is connected by optical fiber cables among quantum switches and quantum users. The successful entanglement generation probability is related to the material and the length of the quantum link, i.e., $p = e^{-\alpha L}$, where α is a positive constant related to the material of the quantum link and L is the length of the quantum link.

4) The Traditional Internet (The Cloud): The quantum network co-works with the traditional Internet together for quantum users' entanglement routing. The Internet is responsible for exchanging information among the networks. For each node including quantum users and quantum switches, they are equipped with traditional computing devices (e.g., computers) and can communicate through the traditional Internet.

We list several the most important roles of the traditional Internet (the cloud) in the quantum network, but not including all.

- The cloud is the center of the network that knows detailed information of quantum network including quantum-user pairing information, the quantum network topology, the quantum switch capacity, and so on.
- The cloud computes the offline routing paths of quantum-user pairs with network information available.
- The cloud shares network information through the Internet such as quantum-user pairing information, routing paths to quantum switches.

- During the entanglement process, adjacent switches (e.g., the graph distance between switches is small) communicate through the Internet to inform each other about link and switch states.

C. Entanglement Process

[18] presents the detailed quantum network entanglement process for one quantum-user pair. Here, we summarize the routing entanglement process for multiple quantum-user pairs as a two-stage process including an *offline stage* and an *online stage*.

1) Offline Stage: In *offline stage*, the main tasks of the quantum network are offline entanglement routing design for quantum-user pairs and transmitting the routing paths to switches for the entanglement in *online stage*.

The offline routing protocol design is conducted by the cloud. We assume that the following offline information of the network is known by the cloud: the quantum-user pairing information; the network topology (switches placement and connection); switches information (the number of qubits in each switch). With all information available, the cloud computes the routing paths for quantum-user pairs with the limitation of switches capacity, and the detailed computing process is discussed in Section IV and Section V. After that, the routing paths computed by the cloud are transmitted through the Internet to switches for the entanglement.

2) Online Stage: In *online stage*, the switches try to generate entanglement among links with the routing paths sent from the cloud, and then swap in the interiors.

The entanglement and swapping process is probabilistic, e.g., the successful entanglement rate over an optical fiber is typically 0.01% [20]. The duration of the entanglement over a link is short, e.g., 1.46s [20]. The entanglement generation time of one attempt is usually 165 μs [20]. All the entanglement and swapping processes over a path should be processed in the duration of the entanglement T. The short duration of T requires the entanglement and swapping process to be carefully considered.

The detailed entanglement process is as follows.

- First, all the switches are time-synchronized through the Internet [13] which can ensure the whole quantum network start entanglement at the same time.
- Second, all the switches try to process entanglement over links and swap in the interiors given the routing paths of all quantum-user pairs. Each switch can try multiple times until the entanglement is generated or the time out (greater than T).
- Third, some switches may fail to generate entanglement over part of links to build a path for quantum-user pairs. Then, the switches will try to build a recovery path for quantum-user pairs locally. Link states (entanglement or not) and swapping states cannot be efficiently sent to the cloud for rescheduling in T due to the Internet delay. The switch can access link states near it through communication with nearby switches with the Internet. The transmission delay from a switch to other switches

TABLE I
TABLE OF NOTATIONS USED IN THE NETWORK MODEL

Notation	Definition
$\mathcal{M}$	The set of $\langle S, D \rangle$ pairs
$\mathcal{A}$	The path set of $\langle S, D \rangle$ pairs
$\mathcal{A}'$	The path set of $\langle S, D \rangle$ pairs with M^2 shortest distance paths
$\hat{\mathcal{M}}$	The set of $\langle S, D \rangle$ pairs selected in STEP I
$\mathcal{S}$	The set of source $\{s_1, s_2, \ldots, s_M\}$
$\mathcal{D}$	The set of destinations $\{d_1, d_2, \ldots, d_M\}$
$\mathcal{V}$	The set of switch nodes $\{v_1, v_2, \ldots, v_N\}$
$\mathcal{E}$	The set of connection links between switches $\{e_{ij}, v_i, v_j \in V\}$
L_{ij}	Length of link e_{ij}
Q_i	The number of qubits contained by switch v_i
$\tilde{Q}_i$	The number of available qubits of switch v_i after STEP I
p_{ij}	The successful entanglement rate of edge eij
α	Link transmission efficiency
$\mathcal{A}_m$	The set of all paths for $\langle s_m, d_m \rangle$
A_m	A path belongs to $\mathcal{A}_m$
$\mathcal{A}'_m$	The set of M shortest distance paths for $\langle s_m, d_m \rangle$
A'_m	A path belongs to $\mathcal{A}'_m$
$\mathcal{P}_{A_m}$	The expected output qubits of path A_m
Q^{A_m}	The number of qubits assigned to path A_m
$x_{A_m} \in \{0, 1\}$	Binary variable indicates whether path A_m is selected in the network

in a few hops is acceptable compared with T. The exact number of hops depends on the Internet latency condition. A typical communication time between two switches within one hop is around 1 ms [20]. With the link states and swapping states available, the switches decide the recovery paths for quantum-user pairs locally.

In this paper, we focus on the entanglement routing design in *offline stage*. The recovery path design in *online stage* has been addressed in [18], [19] efficiently in an online manner which can be applied to our model directly.

III. QUANTUM NETWORK MODEL

In this section, we first describe the quantum network model, and then formulate the routing entanglement problem with the goal to maximize the number of quantum-user pairs that can be served by the network and their expected throughput. The network model described here follows real quantum network entanglement experiments [24]–[26] and previous studies about quantum entanglement routing [11]–[13], [18]. Figure 3 shows an example of the proposed quantum network. The key notations are summarized in Table I.
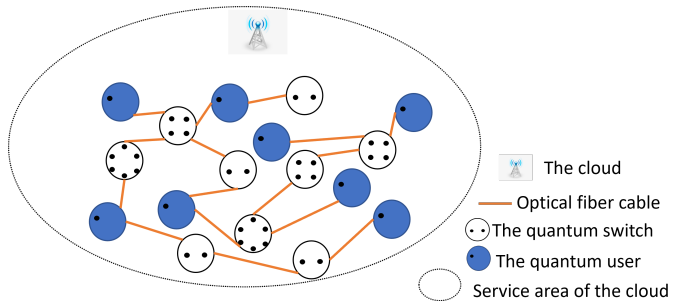

Fig. 3. An Example of Network.

A. Network Model

Quantum Users: The Quantum user set $\mathcal{M}$ consists of M quantum-user pairs $\langle s_1, d_1 \rangle, \langle s_2, d_2 \rangle, \cdots, \langle s_M, d_M \rangle$. $\mathcal{S} = \{s_1, s_2, \cdots, s_m\}$ denotes the set of sources, and $\mathcal{D} = \{d_1, d_2, \cdots, d_m\}$ denotes the set of destinations. They are connected with the quantum network and request for the entanglement. In this paper, we assume that quantum users do not involve other quantum-user pairs entanglement process as switches, and all switches are honest and controlled by the cloud to serve the network.

Network Graph: The transmission graph consists of quantum switches and quantum links. The network is abstracted as an undirected graph which is denoted as $G = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \{v_i\}_{i=1}^N$ denotes the set of quantum switches, and $\mathcal{E} = \{e_{ij}\} \subset \{(v_i, v_j) : v_i, v_j \in \mathcal{V}\}$ denotes the set of the quantum links.

Quantum Switch: Each quantum switch $v_i \in \mathcal{V}$ has Q_i qubits that can be assigned for the entanglement. We focus on 2-qubit entanglement, and the switch uses Bell-state measurements (BSMs). Since exact 2 qubits will be involved in the swapping process, we assume that Q_i is a positive even number. This assumption also fits the real switch design [20]. The successful swapping rate in each switch for any pair of qubits is uniform and denoted as $q \in [0, 1]$.

Quantum Link: e_{ij} is an edge which is an optical fiber cable connecting v_i and v_j for transmitting qubits. In each cable, there are several cores. Each core can be used as a quantum link for the entanglement of a pair of qubits. Therefore, multiple qubits can be assigned at one edge for the entanglement at the same time. We assume that the optical fiber cable contains enough cores for the entanglement between switches. The length of e_{ij} is denoted as L_{ij}. The success rate of each attempt to generate an entanglement over e_{ij} is $p_{ij} = e^{-\alpha L_{ij}}$, where α is a positive constant depending on the physical material. Since p_{ij} only depends on the link length and link material, successful entanglement rates for different pairs of qubits over different cores at the same edge are the same. If a pair of qubits from a quantum-user pair successfully generate the entanglement, there will be a quantum channel between qubits. Each channel can transmit an ebit at each time.

B. Routing Matrices

We use the expected throughput of a path as routing matrix to evaluate the performance of the quantum network.

For a quantum-user pair $\langle s, d \rangle$, let $\mathcal{A}$ denote the set of all paths between s and d. Fix a path $A \in \mathcal{A}$, where $A = \{v_0, v_1, v_2, \cdots, v_{l-1}, v_l\}$, where $v_0 = s$, $v_l = d$, and l denotes the distance of A, i.e., the number of its edges. The nodes in A are listed as the order in path from the source s to the destination d, and the adjacent nodes are connected by one quantum link. Every switch in path A assigns Q^A qubits for the entanglement, which implies the number of parallel quantum channels in path A can be up to $\frac{Q^A}{2}$.

From Section II-C, to build a quantum channel successful for a quantum-user pair along a path requires all links to gener-

ate entanglement and switches to swap successfully during the fixed time period. The probability of one attempt to generate the entanglement successfully of all links in a quantum channel at the same time is the product of the successful entanglement rate of every single link in the channel, i.e., $\Pi_{i=0}^l p_{i(i+1)}$. The probability of one attempt to swap successfully in all switches of a channel at the same time is the product of every switch's successful swapping rate in the channel, i.e., q^{l-1}. Then, the successful probability to build a quantum channel for the entanglement is $\Pi_{i=0}^l p_{i(i+1)} q^{l-1}$. Formally, the routing matrices are defined as the expected throughput of path A with $\frac{Q^A}{2}$ quantum channels for the quantum-user pair $\langle s, d \rangle$:

$$P = \frac{Q^A}{2} \cdot \Pi_{i=0}^l p_{i(i+1)} \cdot q^{l-1} = \frac{Q^A}{2} e^{-\alpha \sum_{i=0}^l L_{i,i+1}} q^{l-1}, \tag{1}$$

which indicates the expected number of ebits can be transmitted from the source to the destination in the fixed time period. In the current setting, the routing matrices also correspond to the expected number of entangled pairs of qubits along the path that are successfully established during the fixed time period.

C. Problem Formulation

We divide our objectives into two steps, named as STEP I and STEP II. In STEP I, our goal is to maximize the number quantum-user pairs that can be served by the network, and a main routing path is selected for every chosen quantum-user pair. In STEP II, we aim to maximize the expected throughput of all selected quantum-user pairs from STEP I. The main paths selected from STEP I is kept to ensure there is at least one routing path for each quantum-user pair selected in STEP I.

STEP I: We first formulate the problem of STEP I. To maximize the number of quantum-user pairs, we assume that $Q^A = 1$ for any path $A \in \mathcal{A}$, and at most one path can be selected for a quantum-user pair. Let the binary variable $x_{A_m} \in \{0, 1\}$ denote whether the path A_m of $\langle s_m, d_m \rangle$ is chosen to be entangled in the network or not. The formulation to maximize the number of quantum-user pairs is as follows:

$$\textbf{Problem } \mathbf{S}_1 : \max \sum_{m \in \mathcal{M}} x_{A_m}, \tag{2}$$

subject to :

$$x_{A_m} \in \{0, 1\}, \ \forall m \in \mathcal{M} \tag{3}$$

$$\sum_{A_m \in \mathcal{A}_m} x_{A_m} \leq 1, \ \forall m \in \mathcal{M} \tag{4}$$

$$\sum_{m \in \mathcal{M}, x_{A_m}=1} |v_i \cap (A_m x_{A_m})| \leq \frac{Q_i}{2}, \ \forall v_i \in \mathcal{V}, \tag{5}$$

$$A_m \in \mathcal{A}_m, \forall m \in \mathcal{M}, \tag{6}$$

where $\mathcal{A}_m$ denotes the set of all paths between $\langle s_m, d_m \rangle$, and A_m is a path of $\mathcal{A}_m$. Constraint (4) denotes that at least one path can be selected for each quantum-user pair which can ensure the network serve quantum-user pairs as many as

possible. Constraint (5) indicates that for any switch $v_i \in \mathcal{V}$, the total number of qubits assigned for all path through v_i cannot over its capacity Q_i, $|\cdot|$ in (5) denotes the number of elements in the set.

STEP II: Next, we formulate the problem in STEP II to maximize the expected throughput of selected quantum-user pairs from STEP I by determining the qubits assigned to possible paths from the path set. We first reserve the qubits in the network assigned for the main paths selected in STEP II, and maximize the expected throughput for quantum-user pairs from STEP I in the residual graph. Let $\hat{\mathcal{M}}$ denote the set of quantum-user pairs selected from STEP I, and $\hat{M}$ denote the number of pairs in $\hat{\mathcal{M}}$. $\hat{Q}_i$ denotes the available qubits of switch v_i after STEP I. The formulation is as follows:

$$\textbf{Problem } \mathbf{S}_2 : \max_{Q^{A_{\hat{m}}}} \sum_{\hat{m}=1}^{\hat{M}} \sum_{A_{\hat{m}} \in \mathcal{A}_{\hat{m}}} P_{A_{\hat{m}}}, \tag{7}$$

subject to

$$A_{\hat{m}} \in \mathcal{A}_{\hat{m}}, \forall \hat{m} \in \hat{\mathcal{M}}, \tag{8}$$

$$Q^{A_{\hat{m}}} \in \mathbb{N}, \ \forall A_{\hat{m}} \in \mathcal{A}_{\hat{m}}, \ \hat{m} \in \hat{\mathcal{M}}, \tag{9}$$

$$0 \le Q^{A_{\hat{m}}} \le \frac{\hat{Q}_i}{2}, \ \ \forall A_{\hat{m}} \in \mathcal{A}_m, \ \hat{m} \in \hat{\mathcal{M}}, \ \forall v_i \in \mathcal{V}, \tag{10}$$

$$\sum_{m \in \mathcal{M}} Q^{A_{\hat{m}}} |v_i \cap A_{\hat{m}}| \le \frac{\hat{Q}_i}{2}, \ \ \forall v_i \in \mathcal{V}, \tag{11}$$

where $P_{A_{\hat{m}}}$ are the expected throughput of path $A_{\hat{m}}$ defined in (1) and $\mathbb{N}$ denotes the set of non-negative integers. $Q^{A_{\hat{m}}}$ is the qubits assigned for path $A_{\hat{m}}$. (10) means that switch v_i cannot assign the qubits to the path over its capacity. (11) indicates that the number of qubits in the switch is the main limitation for paths selection of quantum-user pairs.

IV. Entanglement Routing Algorithm of STEP I

We first propose algorithms to solve Problem $\mathbf{S}_1$ in STEP I, and analyze their performance and time complexity. There are two parts to solve Problem $\mathbf{S}_1$. First, we relax the binary variable x_{A_m} from $\{0, 1\}$ to $[0, 1]$. Let Problem $\hat{\mathbf{S}}_1$ denote the relaxed problem which is a standard linear programming. However, the time complexity to solve Problem $\hat{\mathbf{S}}_1$ is extremely high because of the huge size of the path set (the detailed analyses are presented in Section IV-A). Hence, we construct a smaller path set that contains sufficient paths to reduce the complexity to solve Problem $\hat{\mathbf{S}}_1$. Second, we derive the feasible integer solution from the solution of Problem $\hat{\mathbf{S}}_1$.

A. Complexity

The Problem $\mathbf{S}_1$ in STEP I is an binary multi-commodity flow problem. It has been proved that the problem is NP-Complete [27]. When relaxing the binary variable to be continuous, the fractional solution can be solved by the standard Linear-Programming techniques such as simplex [28].

However, the overhead for computing the paths set is not considered in the previous papers [19], [29]. An inevitable prerequisite to solve Problem $\mathbf{S}_1$ is that the routing paths set $\mathcal{A}_m, \forall m \in \mathcal{M}$ should be calculated. This will add extra

extremely huge computing complexity to solve Problem $\mathbf{S}_1$. More specifically, there could be up to $|\mathcal{E}|!$ paths between one quantum-user pair in a complete graph (the switches can be selected multiple times), where $|\mathcal{E}|$ is the number of edges in $\mathcal{G}$.

The huge paths sets will cause great computational overhead to solve relaxed Problem $\hat{\mathbf{S}}_1$ by using standard linear programming techniques not to mention the integer solution. The computing complexity will be unacceptable to solve Problem $\mathbf{S}_1$ directly.

B. Problem $\hat{\mathbf{S}}_1$ Solution

As we have discussed above, using the standard linear programming technique to solve Problem $\hat{\mathbf{S}}_1$ with huge paths sets will bring unacceptable complexity. To address this challenge, we select the shortest distance (i.e., the number of its edges) paths of quantum-user pairs as the path set instead of all possible paths. Choosing shortest distance paths can consume fewer resources (e.g., the qubits in switches) to satisfy more commodities. This is because the network sources are limited, it is preferred to choose a shorter distance path that consumes fewer resources from the feasible path set to maximize the number of quantum-user pairs in the network. More accurate proofs are shown in [30], [31].

Algorithm 1 Selective Paths Algorithm

Input: $\mathcal{G} = (\mathcal{V}, \mathcal{E}), \mathcal{S}, \mathcal{D}, \mathcal{M}$
Output: $\mathcal{A}'$
1: $\mathcal{A}' = \emptyset$
2: **for all** $m \in \mathcal{M}$ **do**
3:     Obtain M^2 shortest distance paths of the pair $\langle s_m, d_m \rangle$ by Yen's algorithm, $\mathcal{A}'_m = \{A'^k_m\}, k \in [1, M^2]$
4:     $\mathcal{A}' = \mathcal{A}' \cup \mathcal{A}'_m$
5: **end for**
6: Sort paths in $\mathcal{A}'$ by ascending order of length
7: Remove the path with largest length in $\mathcal{A}'$ until $|\mathcal{A}'| = M^2$
8: **for all** $m \in \mathcal{M}$ **do**
9:     Remove the path with largest length in $\mathcal{A}'_m$ until $|\mathcal{A}'_m| = M$
10:     $\mathcal{A}' = \mathcal{A}' \cup \mathcal{A}'_m$
11: **end for**

The detailed path set selection algorithm is summarized in Algorithm 1. We explain how it runs as follows. The goal is to construct a new smaller feasible set $\mathcal{A}'$ with total M^2 paths for Problem $\hat{\mathbf{S}}_1$, and the path set for each quantum-user pair has M paths. We first compute M^2 shortest distance paths by Yen's algorithm [32] for each quantum-user pair. The reasons are as follows. First, M^2 is a large enough number to ensure a source-destination pair has a set with sufficient paths. Meanwhile, M^2 does not bring a huge impact on the time complexity to solve the problem that will be discussed in detail later. Second, if we find shortest distance paths of all quantum-user pairs directly instead of for the individual pair one by one, some pairs with a small number of shortest distance paths will be less likely considered. From the fairness aspect, we choose the path set of each quantum-user pair one by one.

Then, we sort those M^3 paths by ascending order of their distance and add M^2 paths with the shortest distance to $\mathcal{A}'$. We check whether each quantum-user pair has M paths. If not, we continue to add paths of that quantum-user pair with the shortest distance to $\mathcal{A}'$ until $\mathcal{A}'$ includes M paths of each quantum-user pair.

With a smaller path set $\mathcal{A}'$, Problem $\hat{\mathbf{S}}_1$ can be solved by the standard linear programming techniques [28], [33] with the acceptable time complexity. Let $\tilde{x} = \{\tilde{x}_{A'_m} \in [0,1] | \forall A'_m \in \mathcal{A}'_m, m \in \mathcal{M}\}$ denote the set of Problem $\hat{\mathbf{S}}_1$ solution.

C. Integer Solution Recovery

The solution of Problem $\hat{\mathbf{S}}_1$ may be fractional which is not feasible to Problem $\mathbf{S}_1$. Hence, we recover the feasible integer solution of Problem $\mathbf{S}_1$ from $\tilde{x}$ and select one main routing path for each selected quantum-user pair. Let $x^{\dagger}_{A'_m}$ denote the recovered integer solution from Algorithm 2.

The detailed algorithm is described in Algorithm 2. We first add the paths for $\tilde{x}_{A'_m} = 1$. Then, we implement the branch and price strategy [31] to derive feasible integer solution. Let $x^{\dagger}_{A'_m}$ denote the temporal integer solution iterated in Algorithm 2 and Algorithm 3.

Algorithm 2 STEP I Integer Solution Recovery Algorithm

Input: LP solution to STEP I, $\tilde{x}_{A'_m} \forall A'_m \in \mathcal{A}'_m, m \in \mathcal{M}$
Output: Integer solution to STEP I, $x^{\dagger}_{A'_m}, \forall A'_m \in \mathcal{A}'_m, m \in \mathcal{M}$
1: $x^{\dagger}_{A'_m} = 0, \bar{x}_{A'_m} = 0, \forall A'_m \in \mathcal{A}'_m, m \in \mathcal{M}$
2: Sort $\tilde{x}_{A'_m}$ in descending order
3: **for all** $A'_m, m \in \mathcal{M}$ **do**
4:     **if** $\tilde{x}_{A'_m} = 1$ **then**
5:         $\bar{x}_{A'_m} = 1$, mark the pair $\langle s_m, d_m \rangle$
6:     **end if**
7: **end for**
8: Find the maximum $\tilde{x}_{A'_m} < 1, \forall A'_m \in \mathcal{A}'_m, m \in \mathcal{M}$ that satisfies the corresponding $\langle s_m, d_m \rangle$ is not entangled
9: Branch-and-price ($\emptyset$, m, $\tilde{x}_{A'_m}, \bar{x}_{A'_m}$)

Branch and Price Algorithm: The basic idea of the branch and price strategy is to compare the results from different search branches. To accelerate the search process, we optimize the search order and cut some poorly performing branches. We only choose two branches that maximize the number of entanglement pairs as the integer solution instead of all branches. The detailed process is summarized in Algorithm 3.

To start the search process, we first sort the $\tilde{x}_{A'_m}$ in the descending order and select the quantum-user pair $\langle s_m, d_m \rangle$ with the highest $\tilde{x}_{A'_m}$ as the initial pair in the algorithm.

Then, we search the feasible integer path for this pair $\langle s_m, d_m \rangle$. Start from s_m, we search along the path until we find a branch node v_{b_m}, e.g. $A^1 = \{s_m, v_1, v_2\}$, $A^2 = \{s_m, v_1, v_3\}$, $A^3 = \{s_m, v_1, v_4\}$, the branch node is v_1. With multiple paths to select, the preference is to search for the paths with the larger x value to reduce branches. For example, the current path to the branch node v_{b_m} is denoted as $A'_m = \{s_m, v_1, v_2, \cdots, v_{b_m}\}$. For every possible next node v_i which could be added in the path, we append v_i to A'_m,

---

Algorithm 3 Branch and Price Algorithm

Input: Current path $A'_m = \{s_m, v_1, v_2, ..., v_l\}$, m, $\tilde{x}_{A'_m}, \bar{x}_{A'_m}$
Output: $x^{\dagger}_{A'_m}$
1: $\mathcal{A}'' = \emptyset$
2: **for all** $A'_m \in \mathcal{A}'_m$ **do**
3:     **if** $A_m \cap A'_m = A'_m$ and A'_m is feasible **then**
4:         $\mathcal{A}'' = \mathcal{A}'' \cup A'_m$
5:     **end if**
6: **end for**
7: **if** $|\mathcal{A}''| \leq 1$ **then**
8:     Mark the pair $\langle s_m, d_m \rangle$
9:     **if** $|\mathcal{A}''| = 1$ **then**
10:         $\bar{x}_{A''_m} = 1, A''_m \in \mathcal{A}''$
11:     **end if**
12:     Compare $\bar{x}_{A'_m}$ and $x^{\dagger}_{A'_m}$, update $x^{\dagger}_{A'_m}$ if necessary
13:     Find the maximum $\tilde{x}_{A'_{m'}} < 1, \forall A'_{m'} \in \mathcal{A}'_{m'}, m' \in \mathcal{M}$ such that $A'_{m'}$ is feasible and $\langle s_{m'}, d_{m'} \rangle$ is not marked
14:     Branch-and-price ($\emptyset$, m', $\tilde{x}_{A'_m}, \bar{x}_{A'_m}$)
15:     Unmark the pair $\langle s_m, d_m \rangle$, $\bar{x}_{A''_m} = 0$
16: **else**
17:     Find the minimum i s.t. exist two paths $A'_m(v_i)(j), A'_m(v'_i)(j) \in \mathcal{A}'$, satisfies $v_i \neq v'_i$, $v_i \in A'_m(v_i)(j)$, $v'_i \in A'_m(v'_i)(j)$
18:     Choose any $A'_m \in \mathcal{A}'_m$, append $v_{l+1}, ..., v_{i-1}$ to A'_m
19:     **for all** $A'_m(v_i)(j) \in \mathcal{A}''$ **do**
20:         $c_{v_i} = c_{v_i} + \tilde{x}'_{A'_m(v_i)(j)}$
21:     **end for**
22:     Find two maximum $c_{v_i}, c_{v'_i}$
23:     Branch-and-price($A'_m \cup v_i, m, \tilde{x}_{A'_m}, \bar{x}_{A'_m}$)
24:     Branch-and-price($A'_m \cup v'_i, m, \tilde{x}_{A'_m}, \bar{x}_{A'_m}$)
25:     Choose the better solution
26: **end if**

---

and count the total $x_{A'_m}$ value that satisfies A'_m is feasible. In Algorithm 3, among all possible nodes could be added to A'_m, we select two nodes denoted as v_i and v'_i with the top two x values. Next, we continue to build paths in the following two branches denoted as $A'_m(v_i)$ and $A'_m(v'_i)$, respectively.

We repeat this process to construct the path from s_m to d_m until there is only one feasible path or no feasible paths exist. If there is only one feasible path A'_m, then $\bar{x}_{A'_m} = 1$, otherwise $\bar{x}_{A'_m} = 0$. Then, We mark the current entanglement pair $\langle s_m, d_m \rangle$ as searched.

After traversing branches for $\langle s_m, d_m \rangle$, we choose the next quantum-user pair with the largest x value which has not been searched. The search process will end if there are no quantum-user pairs to be searched.

D. Performance Analyses and Discussion

1) Time Complexity to solve Problem $\hat{\mathbf{S}}_1$ (Algorithm 1): The size of the new selected path set $\mathcal{A}'$ is $O(M^2)$, thus there are M^2 variables in the linear programming. The total time complexity of our Algorithm 1 and solving the corresponding linear programming is $O(M^3 N(N^2 + N \log N)) + O((M + M^2)^{2.373}) = O(M^3 N^3 + M^{4.746})$ (when Problem $\hat{\mathbf{S}}_1$ is solved by [33], the linear programming solution with lowest time complexity is $O((M + M^2)^{2.373}$ as far as we know), where N is the number of switches.

2) Algorithm 2 and Algorithm 3: Algorithm 3 is a sub-function in Algorithm 2, so we analyze them together.

Algorithm 3 is a modified branch and price algorithm with the recursion, the complexity is almost impossible to track [31]. Hence, we only analyze the performance guarantee here. Let $f^*(x^*_{A'_m}) = \sum_{m \in \mathcal{M}} x^*_{A'_m}$ and $x^*_{A'_m}$ denote the optimal result and the optimal solution of Problem $\hat{\mathbf{S}}_1$ with path set $\mathcal{A}'$, respectively. Let $f^\dagger(x^\dagger_{A'_m}) = \sum_{m \in \mathcal{M}} x^\dagger_{A'_m}$ and $x^\dagger_{A'_m}$ denote the integer result and the integer solution from the Algorithm 2, respectively. The relationship between $f^*(x^*_{A'_m})$ and $f^\dagger(x^\dagger_{A'_m})$ is stated in the following theorem.

Theorem 1. *Algorithm 2 is an approximation algorithm to Problem $\mathbf{S}_1$ with path set $\mathcal{A}'$, and it achieves an approximation ratio of 2, i.e., $f^*(x^*_{A'_m}) \leq 2f^\dagger(x^\dagger_{A'_m})$.*

The detailed proof will be provided in the extended version of this paper due to the space limitation.

V. EXPECTED THROUGHPUT MAXIMIZATION OF STEP II

In STEP I, we have determined the maximum quantum-user pairs number that can be served by the network and selected one major path for each of them. We then reserve the qubits in the network assigned for the main paths. In STEP II, we aim to maximize the expected throughput of selected quantum-user pairs in STEP I in the residual graph by optimizing the qubits assigned to each path in Problem $\mathbf{S}_2$. The updated formulation of Problem $\mathbf{S}_2$ is,

$$\textbf{Problem } \mathbf{S}_3 : \max_{Q^{A'_{\hat{m}}}} \sum_{\hat{m}=1}^{\hat{M}} \sum_{A'_{\hat{m}} \in \mathcal{A}'_{\hat{m}}} P_{A'_{\hat{m}}},$$

subject to

$$A'_{\hat{m}} \in \mathcal{A}'_{\hat{m}}, \forall \hat{m} \in \hat{\mathcal{M}},$$

$$Q^{A'_{\hat{m}}} \in \mathbb{N}, \ \forall A'_{\hat{m}} \in \mathcal{A}'_{\hat{m}}, \ \hat{m} \in \hat{\mathcal{M}},$$

$$0 \leq Q^{A'_{\hat{m}}} \leq \frac{\hat{Q}^i}{2}, \quad \forall A'_{\hat{m}} \in \mathcal{A}'_m, \ \hat{m} \in \hat{\mathcal{M}}, \ \forall v_i \in \mathcal{V},$$

$$\sum_{\hat{m} \in \hat{\mathcal{M}}} Q^{A'_{\hat{m}}} |v_i \cap A'_{\hat{m}}| \leq \frac{\hat{Q}^i}{2}, \quad \forall v_i \in \mathcal{V}.$$

Problem $\mathbf{S}_3$ is an integer multi-commodity flow problem which is NP-Complete [27]. The formulation of Problem $\mathbf{S}_3$ is similar with Problem $\mathbf{S}_1$ except constraints (3) and (4). Hence, we modify algorithms in STEP I to address the problem in STEP II. The path set of Problem $\mathbf{S}_3$ is constrained by the newly constructed path set $\mathcal{A}'$.

First, we relax $Q^{A'_{\hat{m}}} \in \mathbb{N}$ to be a continuous non-negative real number. The relaxed problem of Problem $\mathbf{S}_3$ is denoted as Problem $\hat{\mathbf{S}}_3$. Problem $\hat{\mathbf{S}}_3$ is a continuous linear programming which can be solved by the standard linear programming method [28], [33]. Let $\tilde{Q}^{A'_{\hat{m}}}$ denote the solution solved from Problem $\hat{\mathbf{S}}_3$. Because $\tilde{Q}^{A'_{\hat{m}}}$ could be fractional which is not feasible, we design an integer recovery algorithm to derive the feasible integer solution. Let $Q^{A'_{\hat{m}}\dagger}$ denote the recovered integer solution, and $\bar{Q}^{A'_{\hat{m}}}$ denote the temporary integer solution iterated in Algorithm 4 and Algorithm 5.

Algorithm 4 STEP II Integer Solution Algorithm

Input: LP solution of STEP II, $\tilde{Q}^{A'_{\hat{m}}}, \forall A'_{\hat{m}} \in \mathcal{A}'_{\hat{m}}, \hat{m} \in \hat{\mathcal{M}}$
Output: Integer solution to step II, $Q^\dagger_{A'_{\hat{m}}}, \forall A'_{\hat{m}} \in \mathcal{A}'_{\hat{m}}, \hat{m} \in \hat{\mathcal{M}}$

1: $Q^\dagger_{A'_{\hat{m}}} = 0, \forall A'_{\hat{m}} \in \mathcal{A}^I_m, \hat{m} \in \hat{\mathcal{M}}$
2: Sort $\tilde{Q}^{A'_{\hat{m}}}$ in descending order
3: **for all** $A'_{\hat{m}} \in \mathcal{M}$ **do**
4:      **while** $\tilde{Q}^{A'_{\hat{m}}} > 1$ **do**
5:          $\bar{Q}^{A'_{\hat{m}}} = \bar{Q}^{A'_{\hat{m}}} + 1, \tilde{Q}^{A'_{\hat{m}}} = \tilde{Q}^{A'_{\hat{m}}} - 1$
6:          Remove the corresponding qubits
7:      **end while**
8: **end for**
9: Find the maximum $\tilde{Q}^{A'_{\hat{m}}} < 1, \forall A_{\hat{m}} \in \mathcal{A}_{\hat{m}}, \hat{m} \in \hat{\mathcal{M}}$ that satisfies $A'_{\hat{m}}$ is feasible
10: Branch-and-price2 $(\tilde{Q}^{A'_{\hat{m}}}, \bar{Q}^{A'_{\hat{m}}})$

Algorithm 5 Branch-and-price2 Algorithm

Input: $\tilde{Q}^{A'_{\hat{m}}}, \bar{Q}^{A'_{\hat{m}}}$
Output: $Q^\dagger_{A'_{\hat{m}}}$

1: Find the maximum $\tilde{Q}^{A'_{\hat{m}}} > 0$, where $A'_{\hat{m}}$ is a feasible and unmarked path
2: **if** Found such $\tilde{Q}^{A'_{\hat{m}}}$ **then**
3:      Mark the path $A'_{\hat{m}}$
4:      $\bar{Q}^{A'_{\hat{m}}} = \bar{Q}^{A'_{\hat{m}}} + 1$, remove the corresponding qubits
5:      Branch-and-price2 $(\tilde{Q}^{A'_{\hat{m}}}, \bar{Q}^{A'_{\hat{m}}})$
6:      $\bar{Q}^{A'_{\hat{m}}} = \bar{Q}^{A'_{\hat{m}}} - 1$, add the corresponding qubits
7:      Branch-and-price2 $(\tilde{Q}^{A'_{\hat{m}}}, \bar{Q}^{A'_{\hat{m}}})$
8:      Unmark the path $A'_{\hat{m}}$
9: **else**
10:      Compare $\bar{Q}^{A'_{\hat{m}}}$ and $Q^\dagger_{A'_{\hat{m}}}$, update $Q^\dagger_{A'_{\hat{m}}}$ if necessary
11: **end if**

Algorithm 4 first determines $\tilde{Q}^{A'_{\hat{m}}}$ that is equal or greater than 1. $\bar{Q}^{A'_{\hat{m}}}$ equals to the integer part of $\tilde{Q}^{A'_{\hat{m}}}$ solved from Problem $\hat{\mathbf{S}}_3$. The main difference needed to be dealt in Algorithm 4 compared with Algorithm 2 is that the range of $\tilde{Q}^{A'_{\hat{m}}}$ is $[0, \frac{Q_i}{2}]$ instead of $[0, 1]$. This indicates that $\bar{Q}^{A'_{\hat{m}}}$ can be added greater than 1 in process (5th row in Algorithm 4).

Then, we use Branch and price algorithm to deal with remaining fractional part in which $\tilde{Q}^{A'_{\hat{m}}} < 1$, and try to recover the feasible integer solution from this part. The Branch and price algorithm is presented in Algorithm 5.

Algorithm 5 is similar but simpler compared with Algorithm 3 because Problem $\mathbf{S}_3$ does not have the constraint to limit the maximum path number of one quantum-user pair. For each iteration, the algorithm finds a feasible path for iteration and decides whether to occupy the current path. The algorithm marks the path to avoid repeatedly accessing the same path. When there are no paths to continue the iteration, Algorithm 5 updates the optimal solution. Hence, Algorithm 4 can derive the optimal solution.

Theorem 2. *The output of Algorithm 4 $\{Q^{A'_{\hat{m}}\dagger}, \forall A_{\hat{m}} \in \mathcal{A}_{\hat{m}}, \hat{m} \in \hat{\mathcal{M}}\}$ is the optimal solution of Problem $\mathbf{S}_3$.*

The proof will be provided in the extended version due to the space limitation. The optimality of Algorithm 4 is guaranteed by the optimality of Algorithm 5 which can be

proved the contradiction. The performance difference between Algorithm 5 and Algorithm 3 is that Algorithm 3 only searches two branches and Algorithm 5 searches every possible branch. This explains why the output of Algorithm 3 is not optimal.

Algorithm 4 can be implemented as an independent algorithm to maximize the network expected throughput without considering selected quantum-user pairs in STEP I. We conduction simulations about Algorithm 4 to maximize expected throughput directly in Section VI and the results reveal that Algorithm 4 outperforms existing works.

VI. SIMULATION RESULTS

In this section, we implement the proposed algorithms on a randomly generated network topology and compare the performance from the number of quantum-user pairs served by the network and the expected throughput with existing works.

A. Network Topology

We generate random networks without the fixed topology. Considering the randomness of the network topology, we generate 5 random networks and take the average value of the measured value, i.e., the expected throughput and the network served quantum-user pairs. The area of the quantum network is set as $10k \times 10k$ unit square, each unit could be 1 kilometer. The number of switches N is set as 50 and the number of quantum-user pairs M is set as 20 by default. These nodes are randomly placed in the area. Figure 4 and Figure 6 present results with varying M and N, respectively. The edge generation follows the work [34]. The number of edges is determined by the average degree of nodes which is set as 10 by default. The length of each edge is at least $\leq \frac{50}{\sqrt{N}}$. The edge capacity does not have the limitation according to our assumption in the model. Quantum-user nodes do not connect with other quantum-user nodes directly, and they are connected with switches directly. We vary the network average degree D in Figure 7b. The number of qubits in each switch Q_i is set as 2, and we assume each quantum user has enough qubits for the entanglement. Figure 5 tests different Q_i. The fiber link material parameter α is deduced by setting the successful entanglement rate of a single link as 0.01%. The successful swapping rate $q = 0.9$, and we vary q in Figure 7a.

B. Algorithm Benchmarks

Our proposed routing design consisted of STEP I and STEP II is denoted as MULTI-R. We compare MULTI-R with the following algorithms and routing matrices:

- ALGORITHM 4: We skip STEP I and implement Algorithm 4 directly over the path set $\mathcal{A}'$ to maximize the network expected throughput.
- FER [19]: First sort quantum-user pairs as the descending order of the expected throughput, then select the pair with the largest expected throughput until no feasible paths exist.
- Q-PASS [18]: Q-PASS is a similar greedy algorithm with FER that uses $\sum \frac{1}{p_{i(i+1)}}$ as the routing matrix, where $p_{i(i+1)}$ is the successful entanglement rate of edge $e_{i(i+1)}$.

It indicates the summation of each link creation rate in a path.
- BASELINE-1(B1): we use the number of hops of a path (i.e., l in (1)) as the evaluation matrices, and run the greedy selection similar with Q-PASS.

C. Performance Evaluation

Number of served quantum-user pairs. Figure 4a, 5a, 6a, 7a present the number of served quantum-user pairs with different numbers of quantum users, numbers of switches, numbers of qubits in a switch, respectively. MULTI-R can improve the network served quantum-user pairs number up to 85%, 329%, 356%, 25% more than FER, Q-PASS, B1 and ALGORITHM 4, respectively. It reveals that our routing design to maximize the number of served quantum-user pairs can effectively let the network serve more quantum users compared with existing works.

Expected Throughput. The unit of expected throughput is ebits per time slot. Figure 4b, 5b, 6b, 7b compare the expected throughput with varying numbers of quantum users, numbers of switches, numbers of qubits in a switch. MULTI-R has similar expected throughput compared with FER. This is because we reserve main paths and assign one qubit for each main path in STEP I, some of them have small expected throughput and will not be selected in STEP II. However, they still reserve a lot resources, which impacts the expected throughput of the network. Compared with Q-PASS and $B1$, MULTI-R can improve the network expected throughput up to $16\times$ times. ALGORITHM 4 has the largest expected throughput (improve up to 27% compared with FER) among them. This benefits from the fact that ALGORITHM 4 is based on optimization instead of the greedy design. Optimizing the qubits assigned to every possible path can effectively utilize the resources of the network.

Impact of the network parameters. Figure 4 and Figure 7b vary numbers of switches N and numbers of average degree D of the network. We can observe that adding more switches and edges can improve the expected throughput and enable the network to serve more quantum-user pairs. When the qubits numbers of a switch is limited by the physical challenges, increasing the density of switches and edges (optical fiber cables) can improve the network performance and serving capability. Figure 5 shows the change of expected throughput and the number of served quantum-user pairs with different numbers of qubits in a switch. Even though the number of qubits in a switch is relatively small, improving the capacity of a switch can enable the network to serve more quantum-user pairs and improve their throughput. Figure 7a tests the expected throughput with different successful swapping rate q. It indicates that improving the q can increase the network expected throughput.

MULTI-R and ALGORITHM 4 perform better than existing baselines in the network served quantum-user pairs number and the expected throughput, respectively, when the capacity of the network (e.g., the number of qubits of a switch, the number of switches, the number of edges) increases.
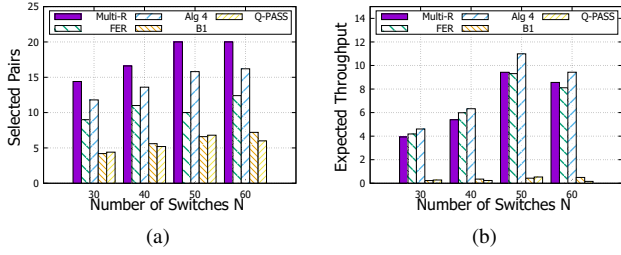
Fig. 4. (a) The selected quantum-user pairs with different switches numbers in network. (b) The expected throughput with different switches numbers in network.
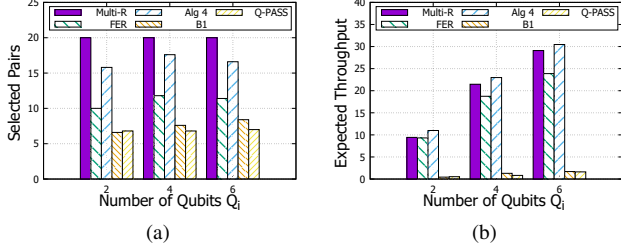

Fig. 5. (a) The selected quantum-user pairs with varying qubits number in each switch. (b) The expected throughput with varying qubits number in each switch.

Impact of the quantum-user pairs number. We vary the number of quantum-user pairs M in Figure 6. The quantum-user pairs number denotes the network demand for the entanglement. When M increases, the demands of the network is larger which leads to the increase of network expected throughput (shown in Figure 6b). MULTI-R can let the network serve more quantum-user pairs compared with FER, Q-PASS, B1 and ALGORITHM 4, respectively. ALGORITHM 4 has obvious improvement in expected throughput compared with existing works.
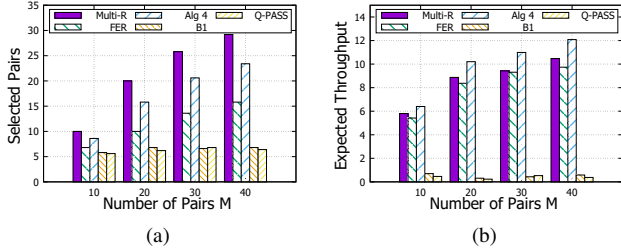

Fig. 6. (a) The selected quantum-user pairs with different quantum-user pairs number M. (b) The expected throughput with different quantum-user pairs number M.
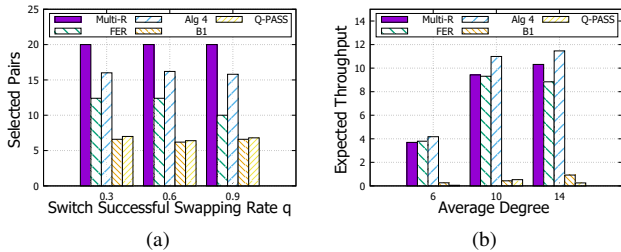

Fig. 7. (a) The selected quantum-user pairs with varying successful swapping rates. (b) The expected throughput with varying network average degrees.

VII. RELATED WORK

Quantum networks and their applications have drawn great attention. Several trials for constructing real quantum net-

works have been conducted, such as DARPA Quantum Network [35], SECOQC Vienna QKD network [36], Tokyo QKD network [37], the mobile quantum network [8], the integrated satellites [9]. These trial networks aim to distribute quantum keys or transmit real qubits for communication. However, it is still not close to be widely applied in the large-scale quantum network in reality because of the physical and hardware limitation.

A few studies have been conducted on the theoretical network layer for the future large-scale quantum network. Numerical evaluations or simulations on the virtual simulator are the main methods to justify the efficiency. Vardoyan *et al.* [12] studied theoretical performance about the switch capacity, the memory occupancy distribution for a single switch with multiple quantum users. Shchhukin *et al.* [16] analyzed the average waiting time for a single entanglement path based on Markov chain theory. Pant *et al.* [13] proposed a local routing policy for independent switch both in single flow and multi-flow. Das *et al.* [17] presented a routing protocol for two groups of quantum users in a Bravais lattice topology. Li *et al.* [14] studied the flow based network performance in a lattice network. Chakraborty *et al.* [15] proposed a greedy routing design in ring and grid networks. These papers considered the routing design in quantum networks with special topologies. These topologies may bring the advantage for the efficient design of routing protocol but they can not fit arbitrary graphs that are more common in reality. Shi *et al.* [18] proposed the routing protocol in a random graph. Their protocol was to add the path one by one with the largest expected throughput. [19] enhanced the performance by using the remaining qubits in the network. However, their protocol assigned too many resources for limited quantum-user pairs that may waste the network resources and limits the number of quantum-user pairs that can be served. Their algorithms were greedy-based without considering the time complexity of choosing paths set and lacked performance guarantee. [29] considered the fidelity as the main limitation for the entanglement which had high-level requirements for the capacity of the network.

VIII. CONCLUSION

In this paper, we have proposed an effective routing protocol for multi-entanglement routing in quantum networks to maximize the number of quantum-user pairs and their throughput at the same time. We have formulated our goal as two sequential integer programming steps and proposed efficient algorithms with low computational complexity and performance guarantees. We have conducted simulations to show that our proposed algorithms have better performance compared with existing algorithms.

ACKNOWLEDGMENT

REFERENCES

[1] R. Van Meter, *Quantum networking*. John Wiley & Sons, 2014.
[2] A. S. Cacciapuoti, M. Caleffi, F. Tafuri, F. S. Cataliotti, S. Gherardini, and G. Bianchi, "Quantum internet: networking challenges in distributed quantum computing," *IEEE Network*, vol. 34, no. 1, pp. 137–143, 2019.
[3] M. Caleffi, A. S. Cacciapuoti, and G. Bianchi, "Quantum internet: from communication to distributed computing!" in *Proceedings of the 5th ACM International Conference on Nanoscale Computing and Communication*, 2018, pp. 1–4.
[4] N. Gisin and R. Thew, "Quantum communication," *Nature photonics*, vol. 1, no. 3, pp. 165–171, 2007.
[5] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd, "Quantum machine learning," *Nature*, vol. 549, no. 7671, pp. 195–202, 2017.
[6] V. Scarani, H. Bechmann-Pasquinucci, N. J. Cerf, M. Dušek, N. Lütkenhaus, and M. Peev, "The security of practical quantum key distribution," *Reviews of modern physics*, vol. 81, no. 3, p. 1301, 2009.
[7] 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.
[8] H.-Y. Liu, X.-H. Tian, C. Gu, P. Fan, X. Ni, R. Yang, J.-N. Zhang, M. Hu, J. Guo, X. Cao *et al.*, "Optical-relayed entanglement distribution using drones as mobile nodes," *Physical Review Letters*, vol. 126, no. 2, p. 020503, 2021.
[9] 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.
[10] W. K. Wootters and W. H. Zurek, "The no-cloning theorem," *Physics Today*, vol. 62, no. 2, pp. 76–77, 2009.
[11] R. Van Meter and J. Touch, "Designing quantum repeater networks," *IEEE Communications Magazine*, vol. 51, no. 8, pp. 64–71, 2013.
[12] G. Vardoyan, S. Guha, P. Nain, and D. Towsley, "On the stochastic analysis of a quantum entanglement switch," *ACM SIGMETRICS Performance Evaluation Review*, vol. 47, no. 2, pp. 27–29, 2019.
[13] 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.
[14] 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.
[15] K. Chakraborty, F. Rozpedek, A. Dahlberg, and S. Wehner, "Distributed routing in a quantum internet," *arXiv preprint arXiv:1907.11630*, 2019.
[16] E. Shchukin, F. Schmidt, and P. van Loock, "Waiting time in quantum repeaters with probabilistic entanglement swapping," *Physical Review A*, vol. 100, no. 3, p. 032322, 2019.
[17] S. Das, S. Khatri, and J. P. Dowling, "Robust quantum network architectures and topologies for entanglement distribution," *Physical Review A*, vol. 97, no. 1, p. 012335, 2018.
[18] 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.
[19] S. Zhang, S. Shi, C. Qian, and K. L. Yeung, "Fragmentation-aware entanglement routing for quantum networks," *Journal of Lightwave Technology*, 2021.
[20] A. Dahlberg, M. Skrzypczyk, T. Coopmans, L. Wubben, F. Rozpedek, 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.
[21] M. A. Nielsen and I. Chuang, "Quantum computation and quantum information," 2002.
[22] B. Coecke, "The logic of entanglement," in *Horizons of the Mind. A Tribute to Prakash Panangaden*. Springer, 2014, pp. 250–267.
[23] H.-J. Briegel, W. Dür, J. I. Cirac, and P. Zoller, "Quantum repeaters: the role of imperfect local operations in quantum communication," *Physical Review Letters*, vol. 81, no. 26, p. 5932, 1998.
[24] 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.

[25] W.-H. Zhang, G. Chen, X.-X. Peng, X.-J. Ye, P. Yin, X.-Y. Xu, J.-S. Xu, C.-F. Li, and G.-C. Guo, "Experimental realization of robust self-testing of bell state measurements," *Physical review letters*, vol. 122, no. 9, p. 090402, 2019.
[26] D. Bouwmeester, J.-W. Pan, K. Mattle, M. Eibl, H. Weinfurter, and A. Zeilinger, "Experimental quantum teleportation," *Nature*, vol. 390, no. 6660, pp. 575–579, 1997.
[27] S. Even, A. Itai, and A. Shamir, "On the complexity of time table and multi-commodity flow problems," in *16th Annual Symposium on Foundations of Computer Science (sfcs 1975)*. IEEE, 1975, pp. 184–193.
[28] V. Chvatal, V. Chvatal *et al.*, *Linear programming*. Macmillan, 1983.
[29] K. Chakraborty, D. Elkouss, B. Rijsman, and S. Wehner, "Entanglement distribution in a quantum network: A multicommodity flow-based approach," *IEEE Transactions on Quantum Engineering*, vol. 1, pp. 1–21, 2020.
[30] L. R. Ford Jr and D. R. Fulkerson, "A suggested computation for maximal multi-commodity network flows," *Management Science*, vol. 5, no. 1, pp. 97–101, 1958.
[31] 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.
[32] J. Y. Yen, "Finding the k shortest loopless paths in a network," *management Science*, vol. 17, no. 11, pp. 712–716, 1971.
[33] M. B. Cohen, Y. T. Lee, and Z. Song, "Solving linear programs in the current matrix multiplication time," *Journal of the ACM (JACM)*, vol. 68, no. 1, pp. 1–39, 2021.
[34] B. M. Waxman, "Routing of multipoint connections," *IEEE journal on selected areas in communications*, vol. 6, no. 9, pp. 1617–1622, 1988.
[35] C. Elliott, A. Colvin, D. Pearson, O. Pikalo, J. Schlafer, and H. Yeh, "Current status of the darpa quantum network," in *Quantum Information and computation III*, vol. 5815. International Society for Optics and Photonics, 2005, pp. 138–149.
[36] M. Peev, C. Pacher, R. Alléaume, C. Barreiro, J. Bouda, W. Boxleitner, T. Debuisschert, E. Diamanti, M. Dianati, J. Dynes *et al.*, "The secoqc quantum key distribution network in vienna," *New Journal of Physics*, vol. 11, no. 7, p. 075001, 2009.
[37] M. Sasaki, M. Fujiwara, H. Ishizuka, W. Klaus, K. Wakui, M. Takeoka, S. Miki, T. Yamashita, Z. Wang, A. Tanaka *et al.*, "Field test of quantum key distribution in the tokyo qkd network," *Optics express*, vol. 19, no. 11, pp. 10 387–10 409, 2011.