IEEE Transactions on Quantum Engineering

Received 10 July 2023; revised 16 January 2024; accepted 8 February 2024; date of publication 16 February 2024; date of current version 29 March 2024.

Digital Object Identifier 10.1109/TQE.2024.3366696

On the Bipartite Entanglement Capacity of Quantum Networks

GAYANE VARDOYAN[1,3], EMILY VAN MILLIGEN[2], SAIKAT GUHA[2], STEPHANIE WEHNER[1], AND DON TOWSLEY[3] (Life Fellow, IEEE)

[1]Delft University of Technology, 2628 CD Delft, The Netherlands
[2]The University of Arizona, Tucson, AZ 85721 USA
[3]University of Massachusetts, Amherst, MA 01003 USA

Corresponding author: Gayane Vardoyan (e-mail: gvardoyan@cs.umass.edu).

The work of S. Guha and D. Towsley was supported in part by the National Science Foundation (NSF) under Grant CNS-1955744, in part by the NSF Engineering Research Center for Quantum Networks under Grant EEC-1941583, and in part by Multidisciplinary University Initiative Army Research Office under Grant W911NF2110325. This work was supported in part by the QuTech Dutch Research Council (NWO) funding 2020–2024—Part I "Fundamental Research," Project Number 601.QT.001-1, financed by the Dutch Research Council. The work of G. Vardoyan was supported by the NWO ZK Quantum Software Consortium (QSC) Ada Lovelace Fellowship and NWO QSC under Grant BGR2 17.269.

ABSTRACT We consider the problem of multipath entanglement distribution to a pair of nodes in a quantum network consisting of devices with nondeterministic entanglement swapping capabilities. Multipath entanglement distribution enables a network to establish end-to-end entangled links across any number of available paths with preestablished link-level entanglement. Probabilistic entanglement swapping, on the other hand, limits the amount of entanglement that is shared between the nodes; this is especially the case when, due to practical constraints, swaps must be performed in temporal proximity to each other. Limiting our focus to the case where only bipartite entanglement is generated across the network, we cast the problem as an instance of generalized flow maximization between two quantum end nodes wishing to communicate. We propose a mixed-integer quadratically constrained program (MIQCP) to solve this flow problem for networks with arbitrary topology. We then compute the overall network capacity, defined as the maximum number of Einstein–Podolsky–Rosen (EPR) states distributed to users per time unit, by solving the flow problem for all possible network states generated by probabilistic entangled link presence and absence, and subsequently by averaging over all network state capacities. The MIQCP can also be applied to networks with multiplexed links. While our approach for computing the overall network capacity has the undesirable property that the total number of states grows exponentially with link multiplexing capability, it nevertheless yields an exact solution that serves as an upper bound comparison basis for the throughput performance of more easily implementable yet nonoptimal entanglement routing algorithms.

INDEX TERMS Entanglement distribution and routing, mixed-integer quadratically constrained program (MIQCP), quantum network.

I. INTRODUCTION

Quantum networks consisting of first- and second-generation quantum repeaters [1], [2] rely on entanglement distribution to enable distributed quantum applications between distant nodes. In the most basic scenario, bipartite entangled states are first established at the physical (elementary) link level between neighboring nodes, and later "connected" via entangling measurements. The latter, a process also known as entanglement swapping, serves to extend entanglement across a longer distance, but for certain quantum architectures succeeds only probabilistically. Fig. 1 illustrates a successful execution of this process using the state $|\Psi^+\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$ as an example. In large, complex quantum networks with arbitrary topology, entanglement routing algorithms are necessary to determine the precise subset of all available swapping operations that must be performed in order to link two nodes that wish to communicate. Efficient use of entanglement—a resource that is created probabilistically and whose usefulness carries an expiration date—is thus an advantageous property for such algorithms. The number of choices that confront an entanglement routing algorithm further increases when there is an option of utilizing multiple paths to distribute as many bipartite entangled states to a pair of users as possible, and when entanglement

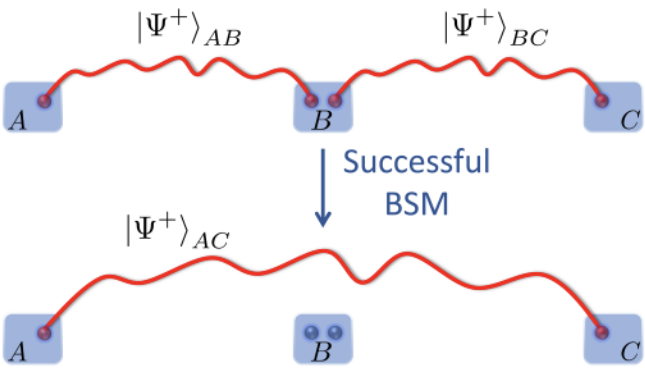

FIGURE 1. Entanglement swapping. Quantum network nodes *A*, *B*, and *C* begin by generating elementary link-level entangled states $|\Psi^+\rangle_{AB}$ and $|\Psi^+\rangle_{BC}$. A successful swap/BSM (Bell state measurement) at node *B* results in an entangled state $|\Psi^+\rangle_{AC}$ shared between *A* and *C*.

swapping success probability varies from node to node.

A metric of interest for the multipath entanglement distribution problem is the overall *bipartite entanglement capacity* (BEC) of a given network, defined here as the average number of bipartite entangled quantum states that the network can provide to a specific pair of users or nodes, s and t (a mathematical definition is provided later in this section). This quantity is especially relevant to applications which have either stringent rate requirements or a necessity to generate multiple contemporaneous Einstein–Podolsky–Rosen (EPR) pairs. An example of the former is a scenario wherein a pair of users carrying out a quantum key distribution (QKD) [3], [4], [5] protocol wish to generate a secret key by a given deadline. An example of the latter is a blind quantum computation (BQC) protocol [6], [7] wherein typically, a user and server must establish several EPR pairs close together in time, in order to be able to carry out each round of the protocol at all. The ability to utilize multiple paths along which to generate entanglement may thus not only boost an application's performance but also make it feasible for implementation on a network. Computation of the BEC provides insight into a network's ability to support such quantum applications. We note that in this work, the BEC is always defined with respect to a node pair (s, t); and further, that "capacity" here incorporates only the rate at which end-to-end entanglement is generated (in contrast to other definitions of capacity commonly found in quantum literature).

Our goal in this article is to compute the BEC for an arbitrary network serving a single user pair. We approach this problem by computing the capacities of all possible network states that result from probabilistic elementary link[1] entanglement generation, and subsequently take the expectation of these values. In this article, "network state" is defined in terms of the presence and location of entangled links; this is not to be confused with *quantum* states. We formulate the (sub)problem of computing the capacity of each network state as an *s-t* flow optimization problem, where the flow through the network may not split or merge, and the network nodes "leak" a certain amount of flow corresponding to their entanglement swapping failure probabilities. We cast this as a mixed-integer quadratically constrained program (MIQCP), which can be directly applied to non-multiplexed networks, i.e., those in which each link is able to generate at most one entangled state per allotted time interval. With a straightforward modification to the graph corresponding to a given network state, the MIQCP can also be applied to a multiplexed network (one in which each elementary link may generate multiple entangled states per time interval). A favorable consequence of the MIQCP approach is that its solution for a given network state yields not only the capacity of the state but also a way to obtain it: each variable's value in the solution specifies the entanglement routing procedure, i.e., the optimal set of swaps to be performed.

A quantum network may be represented as an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with $\mathcal{V}$ comprising a set of nodes (e.g., quantum-equipped users, quantum repeaters, switches, etc.) interconnected via physical links that make up the edge set $\mathcal{E}$. For a given point in time, the *network state* is also represented by a graph, $\mathcal{G}' = (\mathcal{V}, \mathcal{E}')$, where $\mathcal{E}' \subseteq \mathcal{E}$ indicates successfully generated link-level entanglement. A model commonly used in literature is a network that operates in a synchronized, time-slotted manner, i.e., time is discretized into equal-length slots wherein network state may change according to new events or actions. For the type of network we wish to analyze, we are motivated to adopt such a model as well: specifically, we assume that at the beginning of each time slot, all links attempt entanglement generation, and that network state is reset to the empty state (one in which no entangled links are present) at the end of each time slot. Time slot length is assumed to be long enough for all links to be able to perform the necessary number of attempts. For instance, in the absence of multiplexing or with, e.g., spatial multiplexing, time slot length need only be as long as the propagation delay of the longest link in the network $\tau_{\max}$. Certain temporal multiplexing proposals, on the other hand, might require a time slot length that is a multiple of $\tau_{\max}$. Entanglement swapping, namely Bell state measurements (BSMs), is performed at the end of each time slot, before elementary link-level entanglement expires. This model is descriptive of a quantum network with highly limited quantum memory coherence times—for an example, networks with atomic ensemble memories [8] necessitate an almost immediate use of entanglement once it has been generated [9], [10]. Finally, we assume in our model that each BSM succeeds with an *entanglement swapping success probability* $q \in (0, 1]$.

Fig. 2 illustrates an example quantum network with two nodes, s and t, wishing to share as many entangled states as possible. In a heterogeneous network, such as this one, each link (i, j) has its own entanglement generation suc-

[1] An elementary link is one that exists between two neighboring network nodes, i.e., nodes which have a direct physical connection (e.g., optical fiber).

IEEE Transactions on Quantum Engineering

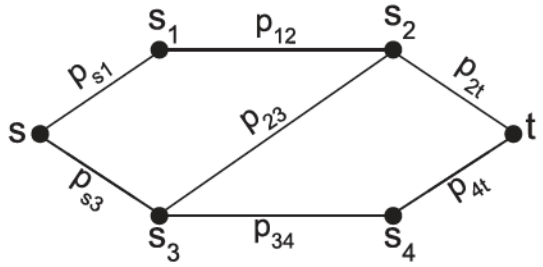

FIGURE 2. Example of a quantum network where two nodes, s and t, wish to share entanglement. Link (i, j) generates entanglement with probability p_{ij}, and node n_i performs entanglement swapping with success probability q_i.

cess probability p_{ij}, and each network node n_i has its own entanglement swapping success probability q_i. Even if the network in Fig. 2 were nonmultiplexed, so that each pair of neighboring nodes shares at most one bipartite entangled state at any time, we may nevertheless already encounter a nontrivial decision-making scenario. Consider, for example, a time instance when all entangled links are present: the network then has the option of connecting s and t via (a) two paths, $s - s_1 - s_2 - t$ and $s - s_3 - s_4 - t$, or via (b) a single path, $s - s_3 - s_2 - t$. Even though option (a) may result in two entangled states while option (b) produces at most one, the optimal choice depends on the swapping success probabilities of the nodes: if, for instance, q_1 and q_4 are extremely low compared to q_2 and q_3, then one may opt for (b) to avoid actions that result, with high probability, in no s-t entanglement at all.

More concretely, the BEC of the aforementioned network state is given by

$$\max\{q_1q_2 + q_3q_4, q_2q_3\}$$

where the capacity of each path is defined as the product of entanglement swapping probabilities of the nodes on that path. The capacity of a network state S may then be defined as

$$C_S := \max_{R\in\mathcal{D}} \sum_{r\in R} f(r) \prod_{\substack{i\in r,\\ i\neq s,t}} q_i \tag{1}$$

where $\mathcal{D}$ is the set of all possible combinations of link-disjoint s-t path sets in the network state, $r \in R$ is an s-t path that is link-disjoint with all other paths in R, $i \in r$ is a node that lies on path r; and $f(r)$ is a function defined on path r. $f(r)$ may, in principle, incorporate an entanglement measure (e.g., secret key rate, entanglement entropy, distillable entanglement, etc.) that is most relevant to the application that is being executed by nodes s and t. As such entanglement measures typically depend on the specific hardware of the network nodes, link lengths, noise processes, entanglement generation protocols, among other factors (which collectively determine the type of bipartite entanglement that results on each link, as well as how each state evolves with time), we take here $f(r) = 1$, for all s-t paths $r \in R$ to keep the analysis both simple and general. In other words, we compute only the expected *number* of entangled pairs shared by s and t at the end of swapping. We remark that in the context of entanglement routing, a more general definition of network state capacity C_S may exist: (1) need not be additive, for instance.

Given our definition of network state capacity above, the overall BEC of the network (with reference to nodes s, t) is given by

$$C = \sum_{S\in\mathcal{S}} P_S C_S \tag{2}$$

where $\mathcal{S}$ is the set of all possible network states, and P_S is the probability of observing the network in state S (see more on this in Section III). While in this work we compute C exactly for a number of networks, it is in principle possible to approximate it by computing the expectation over a subset of $\mathcal{S}$ corresponding to the most likely network states (with highest values of P_S). Arbitrarily tight upper and lower bounds on C can be obtained by increasing the size of this subset, and keeping in mind that the minimum and maximum C_S are given by 0 and $C_{S_{\text{full}}}$, respectively, where S_{full} is the network state with all elementary link-level entanglement successfully established.

We emphasize that the definition of network capacity provided in (2) is given for the specific setup described earlier in the section, namely, for the case in which quantum nodes are only capable of performing BSMs. This formula is a special instance of certain information-theoretic definitions found in literature that consider arbitrary and adaptive local operations and classical communication (LOCC); see Section II for further discussion.

The rest of this article is organized as follows. In Section II, we discuss relevant background and related work. In Section III, we state our assumptions, describe our model of a quantum network, and formulate the problem in terms of s-t flow in a directed graph. In Section IV, we introduce the MIQCP for nonmultiplexed networks and prove its validity, while in Section V we extend the framework to multiplexed networks. We present numerical examples in Section VI. Finally, Section VII concludes this article.

II. RELATED WORK

Request scheduling, path selection, and entanglement routing—problems which are closely interrelated in quantum networking—have all previously received attention in literature. In [11], Caleffi studied optimal entanglement routing, where the goal was to determine the best (single) path between two nodes in a quantum network. Van Meter et al. [12] adapted Dijkstra's algorithm to quantum networks and demonstrated its usefulness for path selection. Gyongyosi et al. [13] proposed a decentralized routing scheme for finding the shortest path in a quantum network. A number of virtual graph construction and routing techniques were also proposed, e.g., [14], [15], and [16]. Such approaches are especially useful for reducing quantum memory requirements

on quantum nodes, as well as lowering latency during the routing process.

As noted in the previous section, a host of studies take an information-theoretic approach in their analyses of quantum network capacity and development of entanglement distribution protocols. Pirandola [17] and Azuma et al. [18] established upper bounds on the maximum rate at which entanglement can be distributed to two nodes in a quantum network. The authors in [17] and [19] considered both single- and multipath routing scenarios, with the latter focusing on realistic networks, e.g., those with imperfect repeaters. In [20], the authors study entanglement distribution protocols (again from an information-theoretic perspective), for a family of quantum networks wherein independent repeater schemes run on each link. At the link level, details describing specific processes that generate entanglement, including, e.g., purification, are abstracted away. The authors use a graph-theoretic result known as Menger's Theorem, which quantifies the number of edge-disjoint s-t paths in a network, to derive protocols that can achieve information-theoretic limits. While the work mainly focuses on deterministic entanglement swapping, the authors provide a brief treatment of the probabilistic-swap case by proposing that swaps can be performed along each s-t path in a "knockout tournament" manner. In the context of our problem, neither this technique, nor Menger's Theorem, can be applied directly: the swapping probabilities play a critical role in determining the optimal set of edge-disjoint paths to be utilized by s and t.

In [21], the authors presented linear programs (LPs) based on the max-flow min-cut theorem, with the goal of efficiently computing bounds on entanglement distribution rates. We remark that the max-flow min-cut theorem is also inapplicable to our problem, as it does not offer assurances that flow will not split or merge. A number of studies obtained exact results only in idealized situations, or specific network topologies: for instance, the authors in [17] analyzed routing with ideal repeater nodes (namely, with perfect BSMs) and reduced the problem to finding the maximum flow through a quantum network. Dai et al. [22] proposed an LP for optimal remote entanglement distribution, in a scenario where quantum memories have the capability to store qubits for an infinite amount of time. In [23], Cicconetti et al. optimized request scheduling and path selection in quantum networks with multiple user requests over a finite time horizon, but did not consider entanglement swap failures. As mentioned in the previous section, we consider networks which reset the state at the end of each time slot; this means that there is no need to consider a time horizon beyond a single time slot.

Yet another collection of studies opted for heuristic strategies. Le et al. [24] applied a deep neural network to address the problem of servicing multiple end-to-end entanglement requests in a quantum network, where entanglement swapping always succeeds. In [25], Victora et al. optimized distillable entanglement in quantum networks; here, multiple paths were considered, although entanglement swapping was assumed to succeed with unit probability. Several other quality-of-service (QoS)-conscious entanglement routing algorithms have been proposed [26], [27], [28], [29], [30], [31], [32], [33], [34], [35], with examples of QoS being bandwidth or throughput, latency, efficient resource use, fidelity, secret key rate; and number of and fairness to concurrent flows. Of these studies, perhaps the most similar *formulation* to that of our work is found in [32], although the problem is posed with an emphasis on resource allocation: in particular, redundant resource provisioning is proposed before any entanglement generation attempts take place. This work improved that of Shi et al. [36] which introduced another (in general nonoptimal) algorithm for concurrent entanglement routing.

Another study with a similar formulation to that of ours (albeit, with the broader aim of distributing entanglement to multiple simultaneously communicating source-destination pairs in a network), is that of [37], wherein the authors constructed an LP to solve the entanglement distribution problem using a multicommodity flow approach. It was assumed that entanglement swapping succeeds probabilistically although all repeaters were assumed to have equal BSM success probabilities. Users were assumed to have minimum end-to-end fidelity requirements which translated into a maximum permitted number of swapping operations. The rate maximization problem was solved using an edge-based formulation which involved the construction of an extended graph, resulting in an LP that scales efficiently with the number of variables within the optimization problem. The objective function of the optimization framework then summed over the flow out of all source nodes in the extended graph, but the assumption of equal-probability entanglement swaps allowed the authors to treat same-length paths in this graph (i.e., paths with the same number of hops) equally. Specifically, any path of length j has a gain factor of q^{j-1}, where q is the BSM success probability. The reason why we cannot (at least, directly or through simple/obvious modifications) adopt our problem as a special case of this multicommodity flow approach, is that in our formulation, node heterogeneity (arising from potentially different swapping success probabilities) does not allow us to decouple gain factors q_i from the flow variables—two equal-length paths could have substantially different flow capacities, making the q_i's a more fundamental aspect of the problem.

Pant et al. [38] investigated multipath entanglement routing in a setting where BSMs succeed probabilistically. The authors proposed routing protocols based on global or local knowledge of the network topology and link state. For a grid network topology, the authors introduced a greedy algorithm, which while not optimal in general, yields a rate that is within a constant factor of the optimum. As the setup and assumptions in this work are most similar to those in our article, we use a similar algorithm as a comparison basis of our results. In [39], Patil et al. improved on the results of [38], by

allowing n-GHZ measurements at network nodes, for $n \geq 3$. Leone et al. [40] also experimented with multipath routing algorithms but without optimality guarantees. On a final note, entanglement routing has also been studied within the context of entanglement percolation, as in [41]; this, however, lies outside the scope of the current study.

III. PROBLEM FORMULATION

In this work, we consider a quantum network that operates under the following assumptions.

1) The network is comprised of quantum repeater nodes interconnected by physical links (e.g., optical fiber) in an arbitrary topology.
2) During execution, time is discretized into slots. A time slot is divided into two stages: I, wherein all links attempt entanglement generation, and II, wherein all nodes attempt entanglement swapping. We will sometimes refer to the network state immediately after stage I as a *snapshot*. In stage II, only BSMs may occur—nodes are incapable of other types of entangling measurements.
3) A network node has a number of links, each of which is allocated quantum storage wherein entangled states may be stored. If the link implements multiplexing (e.g., temporal, spatial, or frequency), so that it may generate up to $k > 1$ entangled states per time slot, then the link has k dedicated storage qubits; otherwise, it has a single storage qubit.
4) At the end of a time slot (after entanglement swapping has taken place), any unconsumed entanglement—entangled states that were not involved in swapping operations—is discarded.
5) End-to-end entanglement, when generated successfully across any s-t path in the network using the processes described above, is assumed to be of sufficiently high fidelity for the application(s) at nodes s and t desiring the Bell pair.

The first part of assumption 2) implies that the network is operating in a coordinated manner: first, entanglement is attempted, and subsequently entanglement routing decisions are executed using global knowledge of the network via entanglement swapping. Such schemes, which may be accomplished with the help of a centralized controller, have been proposed as a means of scaling up quantum networks and ensuring QoS to applications, e.g., delivery of high-fidelity end-to-end entanglement. See, e.g., [42], which proposes such an architecture as a way of circumventing some of the limitations of near-term multiuser quantum networks, [43] which adopts the aforementioned, and [44] where the authors suggest a software-defined networking (SDN) approach toward such ends.

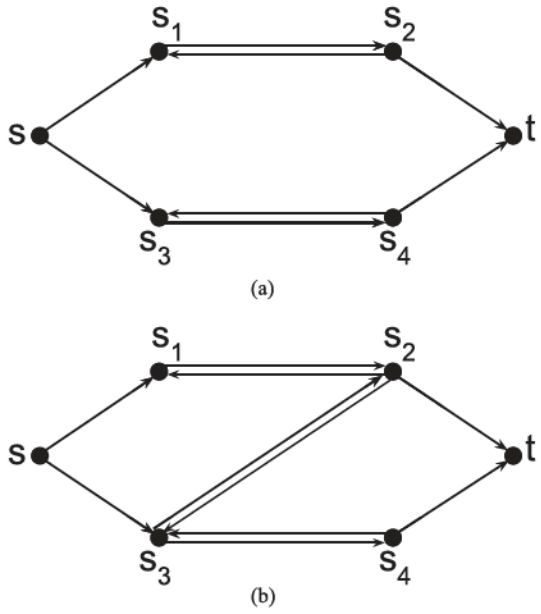

FIGURE 3. Two possible network snapshots after link-level entanglement generation attempts. (a) Snapshot 1. (b) Snapshot 2.

IV. "LEAKY" DISJOINT s-t FLOW IN GRAPHS WITH UNIT EDGE CAPACITY

We focus now on solving the flow problem for a nonmultiplexed network snapshot. While entanglement distribution is inherently symmetric (with no difference in treatment of the end nodes), when cast as a flow problem necessity arises for one node to be designated the source (s) and the other the terminal (t): flow *has* directionality. Either node may assume either role—see discussion at the end of this section. To proceed with this formulation, however, we must devise a directed representation of the undirected graph $\mathcal{G}$ representing a network state (snapshot) immediately after stage I (link-level entanglement generation).

From now on, we may assume that the snapshot $\mathcal{G}$ contains at least one path between s and t (since, otherwise, the capacity of the snapshot is trivially zero). Let $G = (V, E)$ represent the corresponding directed graph for the snapshot. We construct this graph by converting each undirected edge (u, v) of $\mathcal{G}$ into a set of directed edges (u, v) and (v, u), for any $u, v \notin \{s, t\}$. For undirected edges (s, u), it suffices to include only directed edges (s, u) within E, and for any undirected edges (t, u), it suffices to include only directed edges (u, t) within E. Fig. 3 depicts two possible directed snapshots resulting from entanglement generation attempt stages of the network shown in Fig. 2.

In addition to the basic flow conservation constraint frequently encountered in flow problems, our formulation must also account for the following two phenomena: 1) entanglement swapping is subject to failures and 2) each swap operation consumes two entangled links. Incorporating 1) can be accomplished simply by modeling each node in the

snapshot as lossy, with a gain factor equal to the entanglement swapping success probability. Incorporating 2) on the other hand is less straightforward, as it involves designing constraints that ensure the following:

1) two opposing flows cannot coexist, i.e., edges (i, j) and (j, i) cannot simultaneously carry flow—this corresponds to the assumption that any two neighboring nodes have at most one entangled link between them;
2) flows do not split into two or more portions as they propagate through the network;
3) each flow remains disjoint from other flows (flows do not merge); in the special case of a network with unit edge capacities, this amounts to ensuring that each edge is used by at most one flow.

We call the second constraint *flow intactness* and the third *flow disjointness*. As we shall see, these requirements may be satisfied with the help of indicator variables that help direct and restrict the flow as necessary. We begin by introducing an MIQCP that solves this constrained flow problem in Section IV-A, where we also intuitively explain the purpose of each constraint. In Section IV-B, we prove more rigorously the validity of the MIQCP.

A. MIQCP FORMULATION

Let $F_{ij}, i \neq t, j \neq s$, represent flow from node i to j. Consider the following objective function and set of constraints:

$$\max \sum_{(j,t)\in E} F_{jt} \tag{3}$$

$$0 \leq F_{ij} \leq 1 \quad \forall (i, j) \in E \tag{4}$$

$$x_{ijk} \in \{0, 1\} \quad \forall (i, j), (j, k) \in E / \, i \neq k, \, j \neq s, t \tag{5}$$

$$\sum_k (x_{ijk} + x_{kji}) \leq 1 \quad \forall (i, j) \in E \tag{6}$$

$$x_{ijk}(F_{ij}q_j - F_{jk}) = 0 \quad \forall (i, j), (j, k) \in E \tag{7}$$

$$\sum_k x_{ijk} \geq F_{ij} \quad \forall (i, j) \in E / j \neq t \tag{8}$$

$$\sum_{j\in V} (F_{ij} - F_{ji}q_i) = 0 \quad \forall i \in V, i \neq s, t. \tag{9}$$

An intuitive explanation of each line is as follows: first, the objective function maximizes the flow to node t, and since flow can only originate at node s, the solution maximizes the overall s-t flow. In Section IV-B, we explain why this is equivalent to obtaining the capacity of our snapshot graph. Next, we have the constraint on the flow variables F_{ij}: the upper bound of one is due to the fact that we have one quantum memory per physical link interface, and $F_{ij} \in [0, 1]$ is due to the fact that our "internal" network nodes are lossy (internal nodes are all nodes in V except s and t). In the next line, we introduce binary variables x_{ijk} defined for an internal node j and two of its physical neighbors i and k. These variables help us perform a constrained form of bipartite matching between the incoming and the outgoing edges to/from node j.

Then, constraint (6) enacts the flow disjointness and intactness constraints, which consist of ensuring that an edge admits flow from at most one other node, and that an edge releases flow to at most one other node. Further, constraint (6) ensures that two neighboring nodes i and j are only allowed to share flow in one direction. Next, constraint (7) ensures that whenever flow travels from i to j and into k, it must be conserved. Note that this is a more refined flow constraint than that of (9), which conserves flow at each node but does not necessarily ensure the proper amount of flow on each edge, as dictated by the intactness and disjointness constraints. Finally, constraint (8) ensures that flow cannot exist on an edge unless a bipartite matching constraint allows it, i.e., for a flow to be nonzero, there must be a variable x_{ijk} indicating that the edge is actually being used to carry flow.

B. MIQCP VALIDITY

We next study the validity of the MIQCP given by (3)–(9). This involves proving the following:

1) that neighboring nodes share flow in one direction at a time

$$F_{ij} > 0 \Rightarrow F_{ji} = 0, \ \forall (i, j) \in E \tag{10}$$

2) that flow intactness and disjointness hold. Formally, this means the following: given a node $j \neq t$, first construct the sets

$$\mathcal{I}_j = \{F_{ij} : (i, j) \in E, F_{ij} > 0\} \tag{11}$$

$$\mathcal{O}_j = \{F_{jk} : (j, k) \in E, F_{jk} > 0\}. \tag{12}$$

In other words, $\mathcal{I}_j$ is the set of nonzero incoming flows and $\mathcal{O}_j$ is the set of outgoing flows with respect to j. We require the existence of a one-to-one correspondence between these sets, which satisfies the condition that for a flow $F_{ij} \in \mathcal{I}_j$, $F_{ij}q_j = F_{jk}$ for a flow $F_{jk} \in \mathcal{O}_j$;
3) that flows in the network follow valid s-t paths. A valid, l-hop path is defined by a sequence of positive flows $F_{i_0,i_1}, F_{i_1,i_2}, \ldots, F_{i_l,i_{l+1}}$ satisfying $F_{i_k,i_{k+1}}q_{i_{k+1}} = F_{i_{k+1}}$, $k = 1, 2, \ldots, l$, where the set of nodes $\{i_1, i_2, \ldots, i_l\} \subseteq V$. Note that nodes in the path may in principle repeat, although this would not yield optimal results with subunit gain factors. A valid s-t path has the additional property that $i_0 = s$ and $i_{l+1} = t$;
4) that each constraint is necessary.

Having defined path validity, as well as flow intactness and disjointness, we are now able to establish a direct connection between the snapshot capacity (1) with $f(r) = 1 \quad \forall r \in R$, and the MIQCP objective (3). Namely, let us assume that all three conditions hold for any flow in the graph. Then,

IEEE Transactions on Quantum Engineering

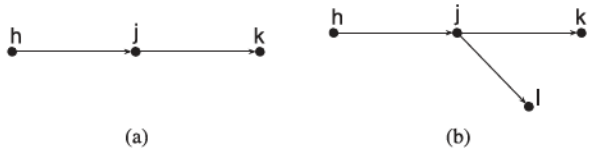

FIGURE 4. Different scenarios relevant to Lemma 1 proof. Directed arrows indicate presence of flow in that direction.

for a given $F_{jt} > 0$, the flow must have originated at node s and followed a valid path $s \to n_1 \to n_2 \to \cdots \to n_k \to t$ through the network, all the while staying intact and disjoint. Since the objective is to maximize flow—a value that lies in [0,1], it follows that $F_{s,n_1} = 1$. Combining this with the aforementioned flow properties allows us to deduce that $F_{jt} = q_1, q_2 \ldots, q_k$, where q_i, $i \in \{1, \ldots, k\}$, is the gain factor associated with node n_i. This expression corresponds to a single term in the sum of (1), and the flow maximization objective then results in finding the optimal set of link-disjoint paths in the graph. We thus conclude that the value of the MIQCP objective function for a given snapshot is equivalent to the snapshot's capacity.

1) NEIGHBORING NODES SHARE FLOW IN AT MOST ONE DIRECTION

We begin with some useful lemmas.

Lemma 1: For any edge (j, k) s.t. $j \neq s$

$$\sum_i x_{ijk} = 0 \Rightarrow F_{jk} = 0.$$

Intuitively, this result states that if no edge (i, j) carries flow into (j, k), then the flow value on (j, k) is zero.

Proof: Our proof is by contradiction. Assume that $F_{jk} > 0$, for a node $j \neq s$, while $\sum_i x_{ijk} = 0$. Recall that there are no outgoing edges at node t by our definition of G, so we know that j is an internal node. Then, by condition (9) on node j, we see that

$$F_{jk} + \sum_{i \neq k} F_{ji} - q_j \sum_i F_{ij} = 0 \tag{13}$$

which implies the existence of some node h such that $F_{hj} > 0$ (since if no such node exists, (9) is violated due to the left-hand side being strictly positive). Thus, we may rewrite (13) as

$$F_{jk} - q_j F_{hj} + \sum_{i \neq k} F_{ji} - q_j \sum_{i \neq h} F_{ij} = 0. \tag{14}$$

We do not require that $h \neq k$ for our proof. Fig. 4(a) depicts the scenario, where $h \neq k$.

Since $F_{hj} > 0$, by condition (8) there exists a node l such that $x_{\text{hjl}} = 1$, and by condition (6) applied to edge (h, j), this l is unique. Further, given the way that the x_{ijk} variables are defined, we can be sure that $l \neq h$. Thus, if $h = k$, i.e., $x_{\text{kjl}} = 1$, then we know by the definition of the variables x_{ijk} that $l \neq k$. On the other hand, if $h \neq k$, then by our assumption that $\sum_i x_{\text{ijk}} = 0$, it follows that $x_{\text{hjk}} = 0$, so that $l \neq k$, i.e., in either case, $l \neq k$—an important observation that, as we shall see shortly, means that flow F_{jk} will remain in the equation.

By condition (7) applied to edges (h, j) and (j, l), we must have

$$F_{hj} q_j = F_{jl} \tag{15}$$

which implies that F_{jl} is strictly positive, so that we have flow from j to l, as depicted in Fig. 4(b), for the case where $h \neq k$. Thus, we may write (14) as

$$F_{jk} - q_j F_{hj} + F_{jl} + \sum_{i \neq k,l} F_{ji} - q_j \sum_{i \neq h} F_{ij} = 0 \tag{16}$$

and by (15)

$$F_{jk} + \sum_{i \neq k,l} F_{ji} - q_j \sum_{i \neq h} F_{ij} = 0. \tag{17}$$

Comparing (13) and (17), we see that the only difference is that in the latter, two components on the left-hand side have cancelled each other out: one from each of the sums, while the F_{jk} component has not changed at all. If we now attempt to balance out F_{jk} using another (negative) component from the second sum above, say F_{gj} $(g \neq h)$, then conditions (8) and (6) ensure the existence of a unique $m \neq g$ for which $x_{gjm} = 1$. Note that $m \neq k$ since either (i) $g \neq k$, so that if $m = k$ we would have $x_{gjk} = 1$, which contradicts our assumption that $x_{ijk} = 0\ \forall i$, or (ii) $g = k$, so that x_{kjm} does not allow $m = k$ by definition. Further, $m \neq l$, since we already have $x_{hjl} = 1$, so we cannot also have $x_{gjl} = 1$ (this is prohibited by condition (6) applied to edge (j, l)). By condition (7) we thus have positive flow F_{jm}, for $m \neq k, l$, which must come from the first sum in (17).

It is evident that any time we attempt to balance out flow F_{jk} with a negative component from (13), we force another (strictly positive) flow into node k, and this positive flow cancels the negative component, leaving F_{jk} unaffected. We can continue this cancellation process until there are no longer any negative components left in the flow balance equation, which would leave us with a strictly positive value on the left-hand side that is $\geq F_{jk}$. From this, we conclude that $\sum_i x_{ijk} = 0 \Rightarrow F_{jk} = 0$. □

Lemma 2: For any edge $(i, j) \in E$, $F_{ij} > 0 \Rightarrow F_{ji} = 0$.

Proof: Note that by G's construction, edges (j, s) and (t, j) do not exist, so we are only concerned with edges (i, j) for which $i, j \notin \{s, t\}$. If $F_{ij} > 0$, then by condition (8), there exists a node k' for which $x_{ijk'} = 1$. Then, by condition (6), $\sum_k x_{kji} = 0$. By Lemma 1, $F_{ji} = 0$. □

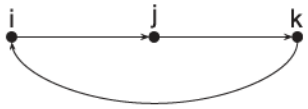

FIGURE 5. Valid flow scenario described in Section IV-B3.

2) FLOW INTACTNESS AND DISJOINTNESS

Lemma 3: Flow cannot split or merge. In other words, there is a one-to-one correspondence between incoming and outgoing flow for any node's flow balance equation. A consequence of this is that there is an even number of components for any flow conservation equation (9).

Proof: For any node $i \neq s, t$, the flow conservation constraint equation (or flow balance equation) (9) can be rewritten as

$$\sum_{j' \in V / F_{ij'} > 0} F_{ij'} = q_i \sum_{j \in V / F_{ji} > 0} F_{ji} \tag{18}$$

where, by Lemma 2, no j' on the left-hand side may be equal to a j on the right-hand side. Consider any F_{ji}. Since it is positive (and since $i \neq t$), we must have, by constraint (8), some node k for which $x_{\text{jik}} = 1$, and by constraint (6) this k must be unique. Then, by constraint (7), we must have $F_{ji} q_i = F_{ik}$, implying that $F_{ik} > 0$. Then, Lemma 1 implies that there exists some node l for which $x_{\text{lik}} = 1$, and constraint (6) ensures that this l is unique. But since we already have $x_{\text{jik}} = 1$, it must be that $l = j$. Thus, for every flow F_{ji} on the right-hand side of (18) there is a unique corresponding flow F_{ik} on the left-hand side of (18), proving the lemma. □

Lemma 3 implies flow intactness and disjointness.

3) PATH VALIDITY

First, we show that flows in the network originate at node s and terminate at node t. To show this, we consider a flow $F_{ij} > 0$, with $i, j \neq s, t$, respectively. Let us first see where the flow ends: since $j \neq t$, there exists a node $k \neq i$ for which $x_{ijk} = 1$ [condition (8)], which then implies a positive F_{jk} [condition (7)]. If $k = t$, we are done, as condition (8) no longer applies, and there are no outgoing edges from t. Otherwise, condition (8) is applicable: $\exists l \neq j$ s.t. $x_{jkl} = 1$, implying a positive flow F_{kl}. We note here that in principle, l could be the same node as i, as shown in Fig. 5, albeit likely a suboptimal option, especially if i is a lossy node (i.e., where $q_i < 1$). Continued applications of conditions (8) and (7), subject to constraint (6), continue until the sole stopping condition—positive flow that ends at node t.

We now focus on the origin of flow F_{ij}, by backtracking it from node i with repeated applications of Lemma 1. First, note that by the contrapositive of the lemma's hypothesis, $F_{ij} > 0 \Rightarrow \sum_l x_{\text{lij}} \neq 0$, implying that $x_{\text{hij}} = 1$ for some $h \neq j$. By condition (7), $F_{hi} > 0$, and if $h = s$, we are done, since Lemma 1 no longer applies, and there are no edges that go into s. Otherwise, the lemma does apply, and we continue its application along with condition (7) until the only stopping condition is reached, i.e., until we find a flow originating at node s. Finally, recall that, by Lemma 3, no flow in the network may split or merge—together with the discussion above, it is evident that the paths are valid, i.e., they are disjoint, intact paths that originate at s and end at t.

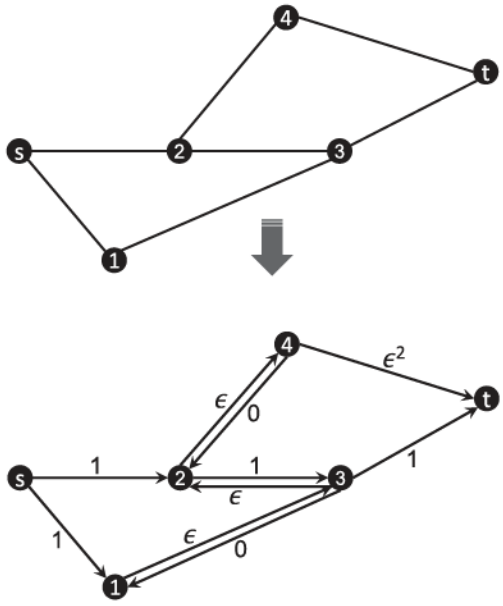

FIGURE 6. Example illustrating the necessity of constraint (6). Nodes 1 and 4 have gain factors ϵ, while nodes 2 and 3 have unit gain factors.

4) NECESSITY OF EACH CONSTRAINT

We consider here the nontrivial constraints (6)–(9). While they were all previously helpful in arguing certain flow properties, we have yet to show that each of these constraints is *necessary*. To do so, we consider the MIQCP of Section IV-A without each constraint individually, and provide counterexamples that illustrate undesirable flow behavior.First, suppose that (6) is removed from the MIQCP. Fig. 6 presents an example where the absence of this constraint allows two opposing flows to coexist—edge (2,3) carries unit flow, while simultaneously edge (3,2) carries an ϵ amount of flow; we explain how this happened shortly. In this network, nodes 2 and 3 have unit gain factors while nodes 1 and 4 have gain factors of ϵ. From the undirected snapshot (top), it is evident that depending on the value of ϵ, it may be optimal to either use the two completely disjoint paths $s-1-3-t$ and $s-2-4-t$, or simply the path $s-2-3-t$. The former yields a snapshot capacity of 2ϵ, while the latter yields 1. Without constraint (6), however, it is also possible to obtain a snapshot capacity of $1+\epsilon^2$, if one uses paths $s-2-3-t$ and $s-1-3-2-4-t$. Table 1 presents all nonzero variables and their assigned values for this network. Note that constraints (7)–(9) are satisfied, but without (6) path $s-1-3-2-4-t$, which is not disjoint with path $s-2-3-t$, is active. Constraint (6) would have ensured

IEEE Transactions on Quantum Engineering

TABLE 1 Variables and Their Assigned Values for the Network in Fig. 6 (Variables Assigned 0 Not Shown)

Flow on (i,j)	Corresp. x_{ijk}
$F_{s1}=1$	$x_{s13}=1$
$F_{13}=\epsilon$	$x_{132}=1$
$F_{32}=\epsilon$	$x_{324}=1$
$F_{24}=\epsilon$	$x_{24t}=1$
$F_{4t}=\epsilon^2$	N/A
$F_{s2}=1$	$x_{s23}=1$
$F_{23}=1$	$x_{23t}=1$
$F_{3t}=1$	N/A

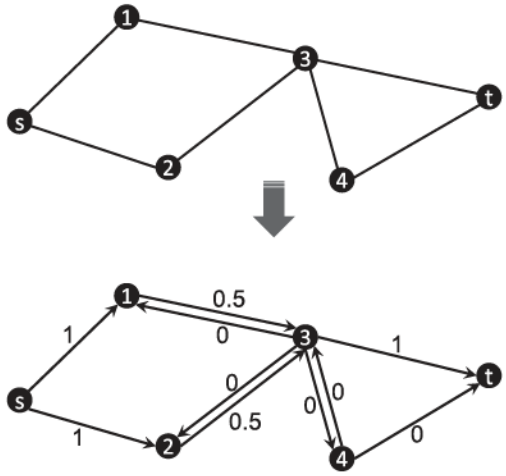

FIGURE 7. Example illustrating the necessity of constraint (7). Nodes 1 and 2 have gain factors of 1/2, node 3 a unit gain factor, and node 4 a gain factor of $\epsilon \ll 1$.

TABLE 2 Variables and Their Assigned Values for the Network in Fig. 7 (Variables Assigned 0 Not Shown)

Flow on (i,j)	Corresp. x_{ijk}
$F_{s1}=1$	$x_{s13}=1$
$F_{13}=0.5$	$x_{13t}=1$
$F_{3t}=1$	N/A
$F_{s2}=1$	$x_{s23}=1$
$F_{23}=0.5$	$x_{234}=1$

that x_{s23} and x_{324} are not both set to one (similarly for x_{23t} and x_{132}).

We next move on to constraint (7): to see why it is necessary, consider the network in Fig. 7. Here, nodes 1 and 2 have gain factors of 1/2, node 3 a unit gain factor, and node 4 a gain factor of $\epsilon \ll 1$. The true optimal value of the objective is $0.5+0.5\epsilon<1$, using for example the two edge-disjoint paths $s-1-3-t$ and $s-2-3-4-t$. Without constraint (7); however, it is possible to route flow so that the optimal value is 1, as shown in Fig. 7, bottom. First, note that with these flow assignments, flow at each internal node is balanced [constraint (9)]. Table 2 makes it easy to see that (8) is also satisfied; the second column of the table can also be used to verify that conditions (6) are met. If (7) were included, it would have ensured that flows do not merge at node 3.

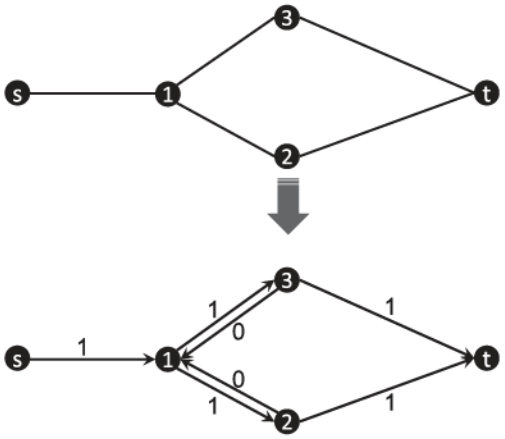

FIGURE 8. Example illustrating the necessity of constraint (9). All internal nodes have unit gain factors.

TABLE 3 Variables and Their Assigned Values for the Network in Fig. 8 (Variables Assigned 0 Not Shown)

Flow on (i,j)	Corresp. x_{ijk}
$F_{s1}=1$	$x_{s13}=1$
$F_{13}=1$	$x_{13t}=1$
$F_{3t}=1$	N/A
$F_{12}=1$	$x_{12t}=1$
$F_{2t}=1$	N/A

Finally, we address constraint (9). Fig. 8 presents a network snapshot (top) and corresponding flow formulation (bottom) where the sink node t receives more flow than is released from the source s. In this network, all internal nodes have unit gain factors. Table 3 presents the nonzero flows and their corresponding x_{ijk}'s [each set to one, to ensure that constraints (8) are satisfied]. It can be checked from the right column of the table that constraints (6) are also satisfied. Since x_{s13}, x_{13t}, and x_{12t} are set to one, we have $F_{s1}=F_{13}$, $F_{13}=F_{3t}$, and $F_{12}=F_{2t}$, respectively, to ensure that (7) are satisfied. Without constraint (9); however, nothing enforces overall flow conservation at node 1.

We conclude this section with a final remark: that switching the roles of s and t so that s is the sink and t the source, i.e., reversing the flow within the network, yields the same solution to the snapshot flow problem. To see this, recall that flows follow edge-disjoint s-t paths, with the additional requirement that if positive flow exists on edge (i,j), then edge (j,i) carries zero flow. This means that to obtain the same solution on a reversed network, one simply reverses each of the s-t paths obtained on the original network. This reversal preserves all necessary requirements on the flow, and the amount of flow that reaches node s in the modified network is the same as the amount that would have reached t in the original network (the multiplication order of q_i's is inconsequential). Obtaining a better solution on the reversed network is impossible, as a consequence of the same argument as above—one could obtain the same solution on the original network using path reversal.

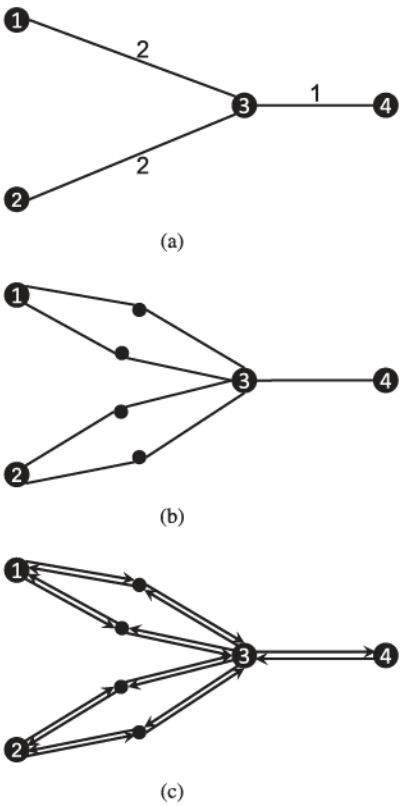

FIGURE 9. Example multiplexed network snapshot snippet, (a), is transformed by first placing a node of unit gain factor in the middle of each entangled link as shown in (b), and then by converting to a directed graph, (c). (a) Initial snapshot. (b) Entangled link transformation. (c) Conversion to a directed graph.

V. EXTENSION TO MULTIPLEXED LINKS

In this section, we extend our approach to compute the capacity of a multiplexed quantum network, i.e., one in which neighboring nodes may establish multiple entangled links in a single time slot. There are two approaches of extending the problem formulation to include multiplexing at the link level. The first is to modify the MIQCP from the previous section to account for the possible presence of more than one Bell pair on each link. Another approach is to instead transform the graph in a way that would conform with the MIQCP in Section IV. While the first method may be more efficient in yielding a solution (it uses fewer decision variables), in this work, we adapt the latter approach as it has the advantage of not requiring any changes to the MIQCP, thus making it straightforward to check its correctness.

We propose to carry out such a transformation as follows: consider a link that has capacity $c = 2, 3, \ldots$ We first replace the link with c undirected links, each with capacity one. We then transform each such entangled link (i, j) into two edges and a new node n with unit gain factor, so that the result is a snapshot with edges (i, n) and (n, j), and no direct edge between i and j. An example of this step is shown in Fig. 9(b), for the multiplexed network snippet in 9(a), where all nodes are assumed to be internal ones. This step is only necessary for snapshot edges with capacities of at least two—notice that the edge (3,4) in the example is unchanged. Note further that this transformation does not change the solution of the problem, in particular because each new node has a unit gain factor and is only connected to two nodes which no longer have a direct connection. Hence, the resulting snapshot is equivalent to the original one within the context of our flow problem, but has the advantageous property that each edge is of unit capacity. The next step is thus to convert the transformed snapshot into a directed graph, as shown in Fig. 9(c) (for a snippet of a snapshot), enabling us to apply the MIQCP to obtain a solution.

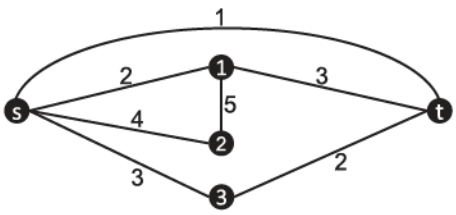

FIGURE 10. Five-node network snapshot with multiplexing. Each edge has a capacity, or the number of available Bell pairs that were successfully generated for this snapshot. The loss factors for the internal nodes 1, 2, and 3 are 0.5, 0.27, and 0.64, respectively.

Fig. 10 shows an example of a simple network snapshot with multiplexing. The solution for this example is easy to obtain by hand: first, note that there is a direct edge between s and t, of capacity 1. Then, there are two paths $s-3-t$, contributing $0.64 \times 2 = 1.28$ to the total capacity. The next obvious paths are the two $s-1-t$ paths (each of which necessarily yields a higher capacity than the longer path $s-2-1-t$), with an overall capacity of $0.5 \times 2 = 1$. Finally, we may use one of the $s-2-1-t$ paths, with a capacity of $0.27 \times 0.5 = 0.135$. This yields a snapshot capacity of $1 + 1.28 + 1 + 0.135 = 3.415$, which is also the output of the MIQCP.

VI. NUMERICAL EXAMPLES

We evaluate our capacity computation approach on several unit edge capacity and multiplexed quantum network scenarios. For nonmultiplexed networks, we validate the results of the MIQCP using a brute-force capacity computation method applied to every possible network snapshot. For multiplexed networks, whose brute-force validation is infeasible due to both the total number of snapshots as well as their complexities (i.e., a large number of disjoint path set combinations to be evaluated), we apply the brute-force method to a small subset of the snapshots. These results are further validated by the capacity predictions of the local-knowledge-based algorithm introduced by Pant et al. [38]. In this work, the authors study multipath entanglement routing in homogeneous networks, wherein all links and nodes have identical entanglement generation and swapping success probabilities, respectively. The proposed routing algorithm uses a Euclidean distance-based metric, as well as local link state knowledge

IEEE Transactions on Quantum Engineering

to inform each individual repeater what entanglement swaps to perform. To compare against our formulation, the distance metric was updated to use hop distances which was seen to perform better on our arbitrary-topology, nonhomogeneous networks. Additional modifications were made to account for multiplexed edges, before Monte Carlo simulations were performed to find the average capacity; a full description of the algorithm is given in [45]. As this algorithm provides no guarantee of yielding an optimal solution on a network snapshot, we expect a worse performance than that achieved with the MIQCP.

For most of our examples, entanglement generation success probabilities are computed using the formula $p = c\eta$, where η is the transmissivity of optical fiber and c is a factor that accounts for various losses other than the transmission loss in fiber. For a specific link l, the former is given by

$$\eta_l = 10^{-0.1\beta L_l} \tag{19}$$

where β is the fiber attenuation coefficient and L_l is the length of the link. We take $c = 0.9$ and $\beta = 0.2$ dB/km. Note that this value for c is optimistic for near-term quantum networks, but using it does allow us to explore more interesting regimes (as well as to avoid numerical instability issues).

Two of the networks which we study in this section, namely, the Abilene and the NSFNet, which formerly stretched across the contiguous U.S., are scaled down when we evaluate our capacity computation method on them: the NSFNet to a metropolitan area, and the Abilene to a campus-sized area. The reason for this is threefold as follows.

1) A first- or second-generation quantum network with this topology stretched across the U.S. would require additional repeater chains to implement long connections, as direct connections over these distances would result in prohibitively high transmission losses.
2) Ignoring the practical aspect of the item above, high transmission losses translate to very small entanglement generation success probabilities, potentially resulting in numerical instability during capacity computation.
3) Studying metropolitan and campus-area-sized networks is more relevant to the near term.

We solve each snapshot's corresponding MIQCP using the Gurobi Optimizer (version 9.1.2, with default solver options) [46]. Our code is available at [47]. Throughout this section, network capacities are in ebits per second, unless otherwise noted.

A. SIMPLE FIVE-NODE NETWORK

The first scenario we consider is the simple five-node network in Fig. 10. The link-level entanglement generation success probabilities for this network have been randomly chosen, and are 0.5679, 0.5179, 0.4723, 0.2479, 0.7839, 0.5423, and 0.4308 for edges (0,1), (0,2), (0,3), (0,4), (1,2), (1,4), and (3,4), respectively. The entanglement swapping success probabilities, also randomly chosen, are 0.5, 0.27, and 0.64 for nodes 1, 2, and 3, respectively. The network capacity of this network is 1.2121 ebits per time unit, which can be compared to the maximum possible 3.415 when all links are present. The network capacity value obtained via MIQCP matches that of the brute-force algorithm. The local knowledge algorithm produces an average capacity of 1.18019, similar to our calculated results.

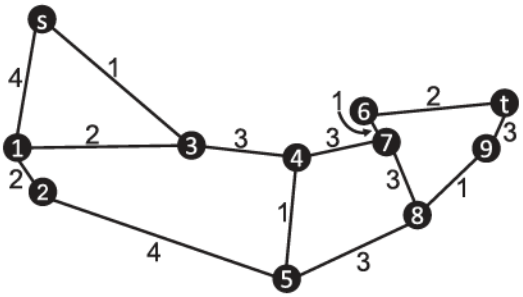

FIGURE 11. Abilene network topology, scaled down to a metropolitan area (see Table 4 for link lengths and entanglement generation success probabilities). Edges are labeled with link capacities for the multiplexed version of the network.

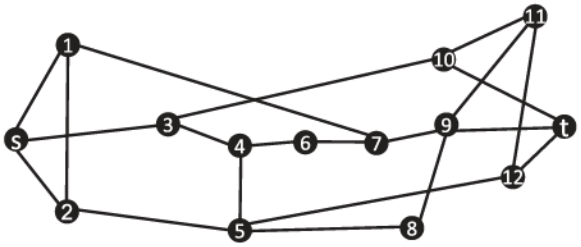

FIGURE 12. NSFNet topology.

TABLE 4 Description of the Abilene Network, With Success Probabilities Corresponding to the Scaled-Down Distances

Edge	Original dist. (km)	Scaled dist. (km)	Success prob.
$(s,1)$	1138	1.138	0.8540
$(s,3)$	1641	1.641	0.8345
$(1,2)$	503	0.503	0.8794
$(1,3)$	1504	1.504	0.8398
$(2,5)$	2206	2.206	0.8131
$(3,4)$	896	0.896	0.8636
$(4,5)$	1041	1.041	0.8579
$(4,7)$	727	0.727	0.8704
$(5,8)$	1128	1.128	0.8544
$(6,7)$	265	0.265	0.8891
$(6,t)$	1144	1.144	0.8538
$(7,8)$	688	0.688	0.8719
$(8,9)$	872	0.872	0.8646
$(9,t)$	328	0.328	0.8865

B. NETWORKS BASED ON THE ABILENE NETWORK

Launched in the late 90s, the Abilene Network connected several institutions across the United States and was used extensively for research and development until its retirement in 2007. The topology of the network is shown in Fig. 11; Table 4 provides a more detailed description of the network. Each node's BSM success probability was sampled uniformly at random from (0.9,1), resulting in $\{q_1, \ldots, q_9\} = \{0.99, 0.96, 0.92, 0.93, 0.91, 0.99, 0.97, 0.92, 0.98\}$.

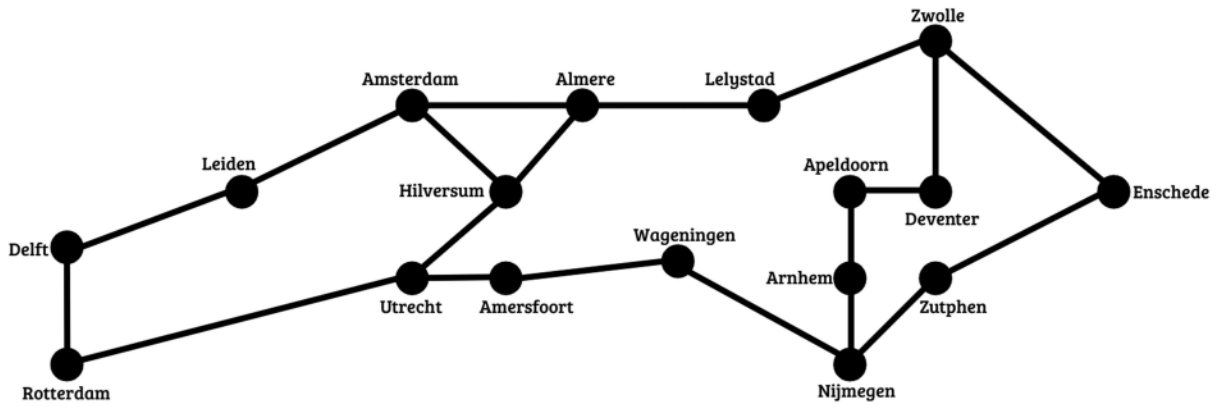

FIGURE 13. Pruned version of the SURFnet topology.

The capacity of the nonmultiplexed scaled-down Abilene Network, as defined by (2), is 0.8301 (this number matches that of the brute-force search algorithm output). This value can be contrasted with the capacity of the snapshot with all entangled links present: 1.607. The local knowledge algorithm gives an average capacity of 0.85271. For the multiplexed Abilene Network, we leave the link generation probabilities unchanged, i.e., in each time slot, there may be up to c_l EPR pairs on link l, each generated with probability p_l. The swapping success probabilities are also left unchanged. The capacity of the multiplexed Abilene Network with edge capacities c_l, as shown in Fig. 11 is 1.387. This can be compared to a capacity of 0.95198 provided by the local-knowledge algorithm. For a multiplexed Abilene Network with $c_l = 2\ \forall l$, capacity increases to 1.983. This increase in capacity compared to that of the network in Fig. 11 is due to the possibility of generating extra links on bottleneck edges (6,7) and (8,9). This can be compared to a capacity of 1.62778 given by the local-knowledge algorithm.

C. NETWORK BASED ON THE NSFNET

NSFNET was a backbone network that linked several U.S. national supercomputing centers and later played a part in developing a portion of the Internet backbone. A version of the NSFNET topology is presented in Fig. 12. Table 5 describes the original approximate link distances, adopted from [48], as well as scaled distances. Success probabilities for link-level entanglement generation correspond to the scaled distances. The BSM success probabilities, sampled uniformly at random from (0.5,1) and rounded down, are set to $\{q_1, \ldots, q_{12}\} = \{0.9, 0.9, 0.8, 0.5, 0.5, 0.5, 0.5, 0.5, 0.7, 0.5, 0.7, 0.5\}$. The network capacity is 0.1013397 (brute-force algorithm validates this). The local knowledge algorithm gives an average capacity of 0.08136.

D. NETWORK BASED ON SURFNET

SURFnet is a backbone network used for research and education purposes in the Netherlands. Fig. 13 depicts a pruned version of the SURFnet topology. The full topology is provided in [49] and a detailed dataset is located at [50]. Note that, contrary to what is shown in Fig. 2 of [49], according to SURFnet topological data there is no link between nodes *Delft 1* and *R'dam 1:* to make the (*Delft*, *Rotterdam*) link, we assume there is no *Delft 2* node, and sum the fiber lengths between *Delft 1*, *Delft 2*, and *R'dam 1*. In a similar manner, we also eliminate nodes *Utrecht 2*, *Wageningen 1*, *Nijmegen 2*, *Zwolle 2*, and *Enschede 1*. The pruned topology is described in more detail in Table 6. Swapping success probabilities are set to $\{q_1, \ldots, q_{15}\} = \{0.87, 0.74, 0.79, 0.62, 0.73, 0.98, 0.77, 0.76, 0.62, 0.74, 0.81, 0.84, 0.7, 0.68, 0.99\}$. Our goal is to determine the *s-t* flow between Delft and Enschede. We determine this value to be 1.0762×10^{-7} (validated by the brute-force algorithm). The local knowledge based algorithm gave a network capacity of 0. This could be in part because it was calculated over a finite number of samples, and is believed to agree with our results. We remark that the reason for the

TABLE 5 Description of the NSFNet Network, With Success Probabilities Corresponding to Rescaled Metropolitan-Area Distances

Edge	Original dist. (km)	Scaled dist. (km)	Success prob.
$(s, 1)$	1100	11	0.5423
$(s, 2)$	600	6	0.6827
$(s, 3)$	1000	10	0.5679
$(1, 2)$	1600	16	0.4308
$(1, 7)$	2800	28	0.2479
$(2, 5)$	2000	20	0.3583
$(3, 4)$	600	6	0.6827
$(3, 10)$	2400	24	0.2980
$(4, 5)$	1100	11	0.5423
$(4, 6)$	800	8	0.6226
$(5, 8)$	1200	12	0.5179
$(5, 12)$	2000	20	0.3583
$(6, 7)$	700	7	0.6520
$(7, 9)$	700	7	0.6520
$(8, 9)$	900	9	0.5946
$(9, 11)$	500	5	0.7149
$(9, t)$	500	5	0.7149
$(10, 11)$	800	8	0.6226
$(10, t)$	1000	10	0.5679
$(11, 12)$	500	5	0.7149
$(12, t)$	300	3	0.7839

IEEE Transactions on Quantum Engineering

TABLE 6 Link Lengths and Corresponding Link-Level Entanglement Generation Success Probabilities for the SURF Network

Edge	Distance (km)	Success prob.
(Delft, Rotterdam) $\equiv (s, 1)$	16.8	0.4152
(Delft, Leiden) $\equiv (s, 2)$	30.6	0.2199
(Leiden, Amsterdam) $\equiv (2, 3)$	60.4	0.0557
(Rotterdam, Utrecht) $\equiv (1, 5)$	70	0.0358
(Amsterdam, Hilversum) $\equiv (3, 4)$	30.2	0.2240
(Amsterdam, Almere) $\equiv (3, 7)$	38.9	0.1501
(Utrecht, Hilversum) $\equiv (4, 5)$	36.7	0.1661
(Hilversum, Almere) $\equiv (4, 7)$	35.4	0.1763
(Utrecht, Amersfoort) $\equiv (5, 6)$	33.8	0.1898
(Almere, Lelystad) $\equiv (7, 9)$	44.2	0.1176
(Amersfoort, Wageningen) $\equiv (6, 8)$	62.5	0.0506
(Wageningen, Nijmegen) $\equiv (8, 12)$	66.3	0.0425
(Nijmegen, Arnhem) $\equiv (11, 12)$	25.7	0.2756
(Nijmegen, Zutphen) $\equiv (12, 15)$	58.1	0.0620
(Arnhem, Apeldoorn) $\equiv (10, 11)$	45.3	0.1117
(Apeldoorn, Deventer) $\equiv (10, 14)$	24.4	0.2926
(Lelystad, Zwolle) $\equiv (9, 13)$	47.7	0.1001
(Deventer, Zwolle) $\equiv (13, 14)$	44.7	0.1149
(Zwolle, Enschede) $\equiv (13, t)$	78.7	0.0240
(Zutphen, Enschede) $\equiv (15, t)$	60	0.0568

low capacity compared to the other examples is due to both the comparably large number of hops between s and t, as well as the (in general) much lower link-level entanglement generation rates. Based on these results, we conclude that this network would benefit from multiplexing, or if unavailable, an alternative method of boosting link-level rates, e.g., via a single-photon entanglement generation scheme [51], [52].

VII. CONCLUSION

In this work, we studied the problem of multipath entanglement routing between two nodes in a quantum network. The links in the network may have multiplexing capabilities, but repeater nodes may not be able to perform deterministic entanglement swapping. To solve the problem, we proposed an MIQCP, the solution of which yields the optimal throughput for a given time interval. By averaging over all possible network states, we were able to obtain the average capacity of various networks included in our case studies.

A possible future direction of our work is to explore approximation methods to the capacity computation. A simple way to achieve this would be to compute the capacities of only the most likely network snapshots (those occurring with relatively high probabilities). Other approximations, possibly with much lower time complexities, are also of interest. An extension of our work would be to consider a time horizon beyond one time slot, and to allow the repeater nodes to perform n-GHZ measurements, for $n > 2$ instead of, or in addition to BSMs.

ACKNOWLEDGMENT

G. Vardoyan would like to thank Subhransu Maji, Guus Avis, and Ashlesha Patil for useful discussions. This work was performed in part using high-performance computing equipment obtained under a grant from the Collaborative R&D Fund managed by the Massachusetts Technology Collaborative.

REFERENCES

[1] W. J. Munro, K. Azuma, K. Tamaki, and K. Nemoto, "Inside quantum repeaters," *IEEE J. Sel. Topics Quantum Electron.*, vol. 21, no. 3, pp. 78–90, May/Jun. 2015, doi: 10.1109/JSTQE.2015.2392076.

[2] S. Muralidharan, L. Li, J. Kim, N. Lütkenhaus, M. D. Lukin, and L. Jiang, "Optimal architectures for long distance quantum communication," *Sci. Rep.*, vol. 6, no. 1, 2016, Art. no. 20463, doi: 10.1038/srep20463.

[3] C. H. Bennett and G. Brassard, "Quantum cryptography: Public key distribution and coin tossing," in *Proc. IEEE Int. Conf. Comput., Syst., Signal Process.*, 1984, pp. 175–179, doi: 10.1016/j.tcs.2014.05.025.

[4] A. K. Ekert, "Quantum cryptography based on Bell's theorem," *Phys. Rev. Lett.*, vol. 67, no. 6, 1991, Art. no. 661, doi: 10.1103/PhysRevLett.67.661.

[5] C. H. Bennett, G. Brassard, and N. D. Mermin, "Quantum cryptography without Bell's theorem," *Phys. Rev. Lett.*, vol. 68, no. 5, 1992, Art. no. 557, doi: 10.1103/PhysRevLett.68.557.

[6] A. Broadbent, J. Fitzsimons, and E. Kashefi, "Universal blind quantum computation," in *Proc. 50th Annu. IEEE Symp. Foundations Comput. Sci.*, 2009, pp. 517–526, doi: 10.1109/FOCS.2009.36.

[7] D. Leichtle, L. Music, E. Kashefi, and H. Ollivier, "Verifying BQP computations on noisy devices with minimal overhead," *PRX Quantum*, vol. 2, no. 4, 2021, Art. no. 040302, doi: 10.1103/PRXQuantum.2.040302.

[8] L.-M. Duan, M. D. Lukin, J. I. Cirac, and P. Zoller, "Long-distance quantum communication with atomic ensembles and linear optics," *Nature*, vol. 414, no. 6862, pp. 413–418, 2001, doi: 10.1038/35106500.

[9] D. Maier, "Investigating the scalability of quantum repeater protocols based on atomic ensembles," Master's Thesis, Faculty Physics, Ludwig Maximilian Univ. of Munich, Munich, Germany, and QuTech, Delft TU, Delft, The Netherlands, 2020. [Online]. Available: https://repository.tudelft.nl/islandora/object/uuid:04b9f054-2139-4b30-ba67-4b7b4752ce86

[10] J. Rabbie, "Simulation model for atomic ensemble based quantum repeaters and the optimization of their positioning," Master's Thesis, Delft Tech. Univ., Delft, The Netherlands, 2020. [Online]. Available: https://repository.tudelft.nl/islandora/object/uuid%3Ad687b10f-3948-4185-9418-53044e15c0ab

[11] M. Caleffi, "Optimal routing for quantum networks," *IEEE Access*, vol. 5, pp. 22299–22312, 2017, doi: 10.1109/ACCESS.2016.2641474.

[12] R. V. Meter, T. Satoh, T. D. Ladd, W. J. Munro, and K. Nemoto, "Path selection for quantum repeater networks," *Netw. Sci.*, vol. 3, pp. 82–95, 2013, doi: 10.1007/s13119-013-0026-2.

[13] L. Gyongyosi and S. Imre, "Entanglement-gradient routing for quantum networks," *Sci. Rep.*, vol. 7, no. 1, 2017, Art. no. 14255, doi: 10.1038/s41598-017-14394-w.

[14] E. Schoute, L. Mancinska, T. Islam, I. Kerenidis, and S. Wehner, "Shortcuts to quantum network routing," 2016, *arXiv:1610.05238*, doi: 10.48550/arXiv.1610.05238.

[15] L. Gyongyosi and S. Imre, "Decentralized base-graph routing for the quantum Internet," *Phys. Rev. A*, vol. 98, no. 2, 2018, Art. no. 022310, doi: 10.1103/PhysRevA.98.022310.

[16] K. Chakraborty, F. Rozpedek, A. Dahlberg, and S. Wehner, "Distributed routing in a quantum Internet," 2019, *arXiv:1907.11630*, doi: 10.48550/arXiv.1907.11630.

[17] S. Pirandola, "End-to-end capacities of a quantum communication network," *Commun. Phys.*, vol. 2, no. 1, 2019, Art. no. 51, doi: 10.1038/s42005-019-0147-3.

[18] K. Azuma, A. Mizutani, and H.-K. Lo, "Fundamental rate-loss tradeoff for the quantum Internet," *Nature Commun.*, vol. 7, no. 1, 2016, Art. no. 13523, doi: 10.1038/ncomms13523.

[19] C. Harney and S. Pirandola, "End-to-end capacities of imperfect-repeater quantum networks," *Quantum Sci. Technol.*, vol. 7, no. 4, 2022, Art. no. 045009, doi: 10.1088/2058-9565/ac7ba0.

[20] K. Azuma and G. Kato, "Aggregating quantum repeaters for the quantum Internet," *Phys. Rev. A*, vol. 96, no. 3, 2017, Art. no. 032332, doi: 10.1103/PhysRevA.96.032332.

[21] S. Bäuml, K. Azuma, G. Kato, and D. Elkouss, "Linear programs for entanglement and key distribution in the quantum Internet," *Commun. Phys.*, vol. 3, no. 1, 2020, Art. no. 55, doi: 10.1038/s42005-020-0318-2.

[22] W. Dai, T. Peng, and M. Z. Win, "Optimal remote entanglement distribution," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 3, pp. 540–556, Mar. 2020, doi: 10.1109/JSAC.2020.2969005.

[23] C. Cicconetti, M. Conti, and A. Passarella, "Request scheduling in quantum networks," *IEEE Trans. Quantum Eng.*, vol. 2, pp. 2–17, 2021, doi: 10.1109/TQE.2021.3090532.

[24] L. Le and T. N. Nguyen, "DQRA: Deep quantum routing agent for entanglement routing in quantum networks," *IEEE Trans. Quantum Eng.*, vol. 3, pp. 1–12, 2022, doi: 10.1109/TQE.2022.3148667.

[25] M. Victora, S. Tserkis, S. Krastanov, A. S. d. l. Cerda, S. Willis, and P. Narang, "Entanglement purification on quantum networks," *Phys. Rev. Res.*, vol. 5, no. 3, Sep. 2023, Art. no. 033171, doi: 10.1103/PhysRevResearch.5.033171.

[26] L. Yang, Y. Zhao, H. Xu, and C. Qiao, "Online entanglement routing in quantum networks," in *Proc. IEEE/ACM 30th Int. Symp. Qual. Service*, 2022, pp. 1–10, doi: 10.1109/IWQoS54832.2022.9812920.

[27] T. N. Nguyen, K. J. Ambarani, L. Le, I. Djordjevic, and Z.-L. Zhang, "A multiple-entanglement routing framework for quantum networks," 2022, *arXiv:2207.11817*, doi: 10.48550/arXiv.2207.11817.

[28] Y. Zeng, J. Zhang, J. Liu, Z. Liu, and Y. Yang, "Multi-entanglement routing design over quantum networks," in *Proc. INFOCOM-IEEE Conf. Comput. Commun.*, 2022, pp. 510–519, doi: 10.1109/INFOCOM48880.2022.9796810.

[29] L. Zhang, S.-X. Ye, Q. Liu, and H. Chen, "Multipath concurrent entanglement routing in quantum networks based on virtual circuit," in *Proc. 4th Int. Conf. Adv. Comput. Technol., Inf. Sci. Commun.*, 2022, pp. 1–5, doi: 10.1109/CTISC54888.2022.9849807.

[30] T. N. Nguyen, D. H. Nguyen, D. H. Pham, B.-H. Liu, and H. N. Nguyen, "LP relaxation-based approximation algorithms for maximizing entangled quantum routing rate," in *Proc. IEEE Int. Conf. Commun.*, 2022, pp. 3269–3274, doi: 10.1109/ICC45855.2022.9838825.

[31] O. Amer, W. O. Krawec, and B. Wang, "Efficient routing for quantum key distribution networks," in *Proc. IEEE Int. Conf. Quantum Comput. Eng.*, 2020, pp. 137–147, doi: 10.1109/QCE49297.2020.00027.

[32] Y. Zhao and C. Qiao, "Redundant entanglement provisioning and selection for throughput maximization in quantum networks," in *Proc. INFOCOM IEEE Conf. Comput. Commun.*, 2021, pp. 1–10, doi: 10.1109/INFOCOM42981.2021.9488850.

[33] M. Ghaderibaneh, H. Gupta, C. Ramakrishnan, and E. Luo, "Pre-distribution of entanglements in quantum networks," in *Proc. IEEE Int. Conf. Quantum Comput. Eng.*, 2022, pp. 426–436, doi: 10.1109/QCE53715.2022.00064.

[34] J. Li et al., "Fidelity-guaranteed entanglement routing in quantum networks," *IEEE Trans. Commun.*, vol. 70, no. 10, pp. 6748–6763, Oct. 2022, doi: 10.1109/TCOMM.2022.3200115.

[35] C. Li, T. Li, Y.-X. Liu, and P. Cappellaro, "Effective routing design for remote entanglement generation on quantum networks," *npj Quantum Inf.*, vol. 7, no. 1, pp. 1–12, 2021, doi: 10.1038/s41534-020-00344-4.

[36] S. Shi and C. Qian, "Concurrent entanglement routing for quantum networks: Model and designs," in *Proc. Annu. Conf. ACM Special Int. Group Data Commun. Appl., Technol., Architectures, Protoc. Comput. Commun.*, 2020, pp. 62–75, doi: 10.1145/3387514.3405853.

[37] K. Chakraborty, D. Elkouss, B. Rijsman, and S. Wehner, "Entanglement distribution in a quantum network: A multicommodity flow-based approach," *IEEE Trans. Quantum Eng.*, vol. 1, 2020, Art. no. 4101321, doi: 10.1109/TQE.2020.3028172.

[38] M. Pant et al., "Routing entanglement in the quantum Internet," *npj Quantum Inf.*, vol. 5, no. 1, 2019, Art. no. 25, doi: 10.1038/s41534-019-0139-x.

[39] A. Patil, M. Pant, D. Englund, D. Towsley, and S. Guha, "Entanglement generation in a quantum network at distance-independent rate," *npj Quantum Inf.*, vol. 8, no. 1, 2022, Art. no. 51, doi: 10.1038/s41534-022-00536-0.

[40] H. Leone, N. R. Miller, D. Singh, N. K. Langford, and P. P. Rohde, "QuNet: Cost vector analysis & multi-path entanglement routing in quantum networks," 2021, doi: 10.48550/arXiv.2105.00418.

[41] A. Acín, J. I. Cirac, and M. Lewenstein, "Entanglement percolation in quantum networks," *Nature Phys.*, vol. 3, no. 4, pp. 256–259, 2007, doi: 10.1038/nphys549.

[42] M. Skrzypczyk and S. Wehner, "An architecture for meeting quality-of-service requirements in multi-user quantum networks," 2021, doi: 10.48550/arXiv.2111.13124.

[43] M. Pompili et al., "Experimental demonstration of entanglement delivery using a quantum network stack," *npj Quantum Inf.*, vol. 8, no. 1, 2022, Art. no. 121, doi: 10.1038/s41534-022-00631-2.

[44] W. Kozlowski and S. Wehner, "Towards large-scale quantum networks," in *Proc. 6th Annu. ACM Int. Conf. Nanoscale Comput. Commun.*, 2019, pp. 1–7, doi: 10.1145/3345312.3345497.

[45] E. A. V. Milligen, E. Jacobson, A. Patil, G. Vardoyan, D. Towsley, and S. Guha, "Entanglement routing over networks with time multiplexed repeaters," 2023, doi: 10.48550/arXiv.2308.15028.

[46] Gurobi Optimization, LLC, "Gurobi optimizer reference manual," 2023. [Online]. Available: https://www.gurobi.com

[47] G. Vardoyan, "Quantum network capacity," 2023. [Online]. Available: https://github.com/gvardoyan/QuantumNetworkCapacity

[48] S. Ghose, R. Kumar, N. Banerjee, and R. Datta, "Multihop virtual topology design in WDM optical networks for self-similar traffic," *Photonic Netw. Commun.*, vol. 10, no. 2, pp. 199–214, 2005, doi: 10.1007/s11107-005-2484-2.

[49] J. Rabbie, K. Chakraborty, G. Avis, and S. Wehner, "Designing quantum networks using preexisting infrastructure," *npj Quantum Inf.*, vol. 8, no. 1, pp. 1–12, 2022, doi: 10.1038/s41534-021-00501-3.

[50] J. Rabbie and G. Avis, "RepAlloc," 2020. [Online]. Available: https://github.com/jtrabbie/RepAlloc

[51] C. Cabrillo, J. I. Cirac, P. Garcia-Fernandez, and P. Zoller, "Creation of entangled states of distant atoms by interference," *Phys. Rev. A*, vol. 59, no. 2, 1999, Art. no. 1025, doi: 10.1103/PhysRevA.59.1025.

[52] P. C. Humphreys et al., "Deterministic delivery of remote entanglement on a quantum network," *Nature*, vol. 558, no. 7709, pp. 268–273, 2018, doi: 10.1038/s41586-018-0200-5.