Link Configuration for Fidelity-Constrained Entanglement Routing in Quantum Networks

Qiaolun Zhang*, Nicola Di Cicco*, Mërmëdhe Ibrahimi*, Raul C. Almeida Jr.‡, Alberto Gatto*, Raouf Boutaba†, Massimo Tornatore*
* Politecnico di Milano, Italy † University of Waterloo, Canada ‡ Federal University of Pernambuco
Q. Zhang and N. Di Cicco are co-first authors of this paper. Corresponding author: qiaolun.zhang@mail.polimi.it

IEEE INFOCOM 2025 - IEEE Conference on Computer Communications | 979-8-3315-4305-1/25/$31.00 ©2025 IEEE | DOI: 10.1109/INFOCOM55648.2025.11044462

***Abstract*—Entanglement routing (ER) in quantum networks must guarantee entanglement fidelity, a property that is crucial for applications such as quantum key distribution, quantum computation, and quantum sensing. Conventional ER approaches assume that network links can only generate entanglements with a fixed fidelity, and then they rely on purification to improve end-to-end fidelities. However, recent advances in entanglement generation technologies show that quantum links can be configured by choosing among different *fidelity/entanglement-rate combinations (defined in this paper as link configurations)*, hence enabling a more flexible assignment of quantum-network resources for meeting specific application requirements. To exploit this opportunity, we introduce the problem of link configuration for fidelity-constrained routing and purification (LC-FCRP) in Quantum Networks. We first formulate a simplified FCRP version as a Mixed Integer Linear Programming (MILP) model, where the link fidelity can be adjusted within a finite set. Then, to explore the full space of possible link configurations, we propose a link configuration algorithm based on a novel shortest-path-based fidelity determination (SPFD) algorithm w/o Bayesian Optimization, which can be applied on top of any existing ER algorithm. Numerical results demonstrate that link configuration improves the acceptance ratio of existing ER algorithms by 87%.**

***Index Terms*—Quantum Networks, Entanglement Routing, Integer Linear Programming, Optical Networks**

I. Introduction

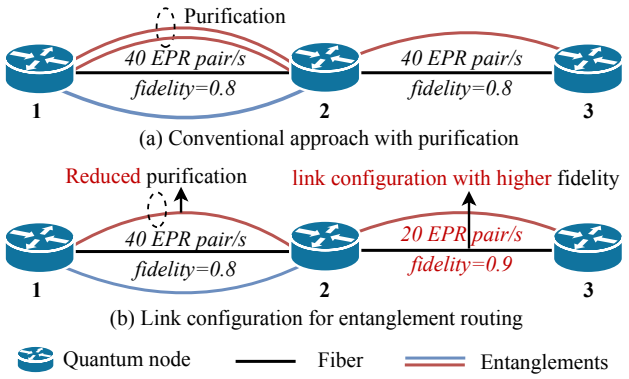
Quantum networking has attracted considerable attention in recent years thanks to its potential for several applications, such as quantum key distribution [1]–[3], quantum computation [4], and quantum sensing [5]. To support these applications, quantum networks need to generate and distribute entangled states across quantum nodes in the network. Specifically, different applications may not only require a specific *entanglement generation rate*, but also a minimum level of *end-to-end fidelity*, i.e., similarity with respect to the intended entangled state [6]. This work focuses on distributing maximally entangled states, specifically Bell pairs [7], between two qubits. These entangled states are also known as Einstein-Podolsky-Rosen pairs, or EPR pairs [8]. To cope with the future demand for quantum entanglements driven by emerging applications, there is a pressing need to investigate efficient strategies for the distribution of quantum entanglement under fidelity constraints [9]–[12].

A quantum network comprises quantum nodes and quantum links. Quantum nodes serve as sources, destinations, or quantum repeaters for entanglement distribution. A typical quantum link is an optical fiber, which is used to generate entanglements between adjacent nodes (defined as link-level entanglements) via heralded entanglement generation techniques [13]. To establish entanglement between non-adjacent nodes, intermediate nodes along the path may perform an operation called *entanglement swapping* on the link-level entanglements [14]. Since the distributed entanglements may be consumed by applications or lost due to decoherence, quantum networks must continuously generate entanglements between adjacent nodes and distribute them across non-adjacent nodes via entanglement swapping. This process is known as the *entanglement routing* (ER) problem.

One of the most fundamental questions in ER is how to distribute entanglements with the guaranteed quality required by quantum applications. The quality of entanglements can be measured with the *fidelity* metric, which is defined as the probability that the set of qubits (two qubits in both end nodes) is in the desired state (entangled state) [15]. Note that fidelity is a real number between 0 and 1, and higher fidelity represents better quality of entanglements. Most entanglement-routing algorithms assume that quantum links can only create entanglements with a *fixed fidelity*, and address the fidelity constraint through *entanglement purification*, which can distill high-fidelity entanglement starting from low-fidelity entanglements [9], [10], [16]. However, recent pioneering research has demonstrated that quantum links can create higher fidelity entanglements by trading off the entanglement generation rate, and vice versa [13], [17]. As a result, quantum links can be configured by selecting different *fidelity/entanglement-rate combinations (defined as link configuration)*. Note that the cost to enable link configuration is only an operational cost (e.g., tuning the microwave pulse applied in quantum nodes), and does not involve any additional hardware [13]. This flexibility allows for more tailored assignment of quantum-network resources to meet specific application requirements.

We illustrate how link configuration can be beneficial for ER via an example in Fig. 1. Assume that we have two requests, one between node pair (1,2) and the other between (1,3). Fig. 1 (a) assumes that each link can only create entanglements with a fixed fidelity of 0.8 and a generation rate of 40 EPR pair/s. The entanglements for requests $(1, 2)$ and $(1, 3)$ are marked with blue lines and red lines, respectively. To guarantee the fidelity constraint of request $(1, 3)$, purification is performed in link $(1, 2)$. In Fig. 1 (b), link $(2, 3)$ is configured to generate entanglements with a fidelity of 0.9 and a generation rate of 20

EPR pairs/s. Since the end-to-end fidelity of entanglements is positively correlated with the fidelity involved in swapping, the fidelity constraint of request $(1, 3)$ might be satisfied without purification in link $(1, 2)$, leading to less resource (i.e. EPR pair rate) consumption in link $(1, 2)$. As a consequence, more entanglements can be used to serve the request $(1, 2)$. Note that, in this example, link $(1, 2)$ is not configured to generate entanglements with higher fidelity, as this would result in a lower generation rate, and, consequently, not enough entanglements would be available for the request between node pair $(1, 2)$. In summary, link configuration can be utilized to improve the performance of entanglement routing, but choosing good link configurations is a non-trivial decision, as it is closely tied to the application needs and the capability of generating entanglements in each link.



(a) Conventional approach with purification

(b) Link configuration for entanglement routing

Fig. 1: Example of link configuration for entanglement routing.

Although link configuration has great potential for improving the fidelity of distributed entanglements, no work has investigated how to set link configurations to support effective ER with purification under end-to-end fidelity constraints. This paper aims to investigate how to utilize link configuration to improve the performance of fidelity-constrained ER in quantum networks.

We summarize our main contributions as follows:

- We propose and investigate, for the first time to the best of our knowledge, the problem of link configuration for fidelity-constrained routing and purification (LC-FCRP) in quantum networks, which aims to guarantee the fidelity of distributed entanglements with jointly tune link configuration and entanglement routing with purification.
- We formulate a simplified version of the FCRP problem as a Mixed Integer Linear Programming (MILP) model, assuming that link fidelities can be configured within a finite set, jointly optimizing the link-link configuration and purification decisions.
- We design a novel link configuration algorithm based on a shortest-path-based fidelity determination (SPFD) algorithm w/o Bayesian Optimization, which can be applied on top of any existing ER algorithm.

II. Related Work

Early works on ER in quantum networks primarily focused on designing optimized resource allocation algorithms to maximize the entanglement generation rate or minimize latency without the fidelity constraint [18]–[20]. More recent works have started to consider the problem of ER with guaranteed fidelity with optimized routing and purification decisions [9], [10], [12], [16], [21]–[25]. Specifically, Ref. [21] proposes to perform purification on each link to achieve a defined fidelity threshold, while Refs. [9], [21] suggest purifying only critical links for resource efficiency. Ref. [23] proposes scheduling the sequence of purification and swapping operations to reduce delay and system overhead. Ref. [24] addresses the probabilistic nature of purification by performing multiple purification attempts until the purification operation succeeds. Unlike these heuristic-based approaches, Ref. [22] and Ref. [12] derive theoretical bounds on achievable fidelity and entanglement generation rates for general-topology quantum networks and quantum repeater chains, respectively. Moreover, Ref. [25] presents optimal and near-optimal algorithms to optimize entanglement throughput with guaranteed fidelity between two distant nodes, and Ref. [16] formulates the ER and purification problem into a convex optimization problem, assuming an initial fidelity extremely close to 0.5, hence necessitating numerous purification operations to achieve high fidelity.

Despite the progress made, most of these works assume quantum links can only generate entanglements with a fixed fidelity [9], [10], [12], [16], [21]–[25]. Instead, recent pioneering experimental and theoretical works [13], [17], [26] have shown that quantum links can be configured to generate entanglements choosing among different fidelity/entanglement-rate combinations. Specifically, Ref. [13] verifies experimentally how to configure the links with different fidelity/entanglement-rate combinations. Ref. [17] and Ref. [26] present theoretical models for link configuration, but do not explore link configuration for entanglement routing with purification. In this work, we build upon these recent technical advances and propose to optimize link configuration for fidelity-constrained ER. By optimizing link configuration, we aim to provide a flexible and efficient link-configuration solution that can be applied on top of different ER algorithms to enhance the performance of quantum networks.

III. System Model and Problem Statement

A. Technical Preliminaries on Quantum Networks

1) Network Elements and Operations of Quantum Networks: Quantum networks can be modeled as a graph $G = (N, E)$, where N is the set of quantum nodes and E is the set of quantum links. Each quantum link can generate entanglements between adjacent node pairs, with each entanglement consuming one qubit of memory from both nodes in the pair. Each quantum node has a limited memory capacity. Current experiments have demonstrated a memory capacity of 225 qubits [27]. This work envisions technological advances and

assumes a larger capacity (e.g., 12,000 qubits) as in Ref. [11]. The main operations in quantum networks are listed below.

Entanglement purification: is an operation serving to distill higher-fidelity entangled pairs from a set of lower-fidelity ones.

Entanglement swapping: is the operation to establish entanglement between non-adjacent nodes. Assume that node 1 and node 2 have one entangled pair, as well as node 2 and node 3. By performing a Bell State Measurement on node 2 and subsequently carrying out some classical communication and operations, an entangled pair can be established between node 1 and node 3. This process is called *entanglement swapping*.

2) Link Configuration: When quantum links are configured to generate entanglements with higher fidelity, one needs to pay a cost in the generation rate of entanglements [17]. This work assumes that the entanglements between adjacent nodes are generated with the state-of-the-art single-photon scheme, as in Ref. [13]. The generated state ρ_w in the considered single-photon scheme has the following form:

$$\rho_w = w|\Psi^+\rangle\langle\Psi^+| + (1-w)\frac{\mathbb{I}_4}{4} \tag{1}$$

Here $|\Psi^+\rangle$ is a Bell state, and $\mathbb{I}_4$ is a completely mixed state representing the noise. w is defined as the Werner parameter, which denotes the extent to which the Werner state retains the entangled state $|\Psi^+\rangle$ versus being mixed state $\mathbb{I}_4$. $(1-w)$ equals the bright state population, which represents the probability of one end node successfully emitting a photon in one attempt to generate entanglements. *The link configuration is performed by configuring the bright state population to the desired value*. The configuration of the bright state population does not necessarily require additional resources. For instance, the experimental work in Ref. [13] demonstrates how to obtain the desired bright state population by changing the microwave pulse applied to the node, which is realized with nitrogen-vacancy (NV) centers.

By tuning the bright state population $(1-w)$, the link can be configured with different fidelity/entanglement-rate configurations as follows. The fidelity of state $|\Psi^+\rangle$ can be derived with Eqn.(2) by measuring the probability of the state ρ_w is in the desired state $|\Psi^+\rangle$ [17]. Moreover, the recent experimental and theoretical works [13], [17] have shown that the generation rate of a link using the single-photon approach can be expressed as $d(1-w)$ using a constant d and Werner parameter w. The constant d is related to various system inefficiencies (e.g., fiber loss) and repetition time (defined as the time between entanglement generation attempts) [17]. In summary, the generation rate, denoted as $d(1-w)$, shows a positive correlation with the bright state population $(1-w)$; instead, the fidelity f of generated entanglements is negatively correlated with the bright state population.

$$f = \langle\Psi^+|\rho_w|\Psi^+\rangle = \frac{3w+1}{4} \tag{2}$$

3) Fidelity Improvement with Entanglement Purification: We adopt nested purification, as shown in Fig. 2. Each iteration of purification is referred to as one round. The entanglements in round 0 are the entanglements generated before any purification process. Then, in each of the following rounds, every two entanglements can be purified to generate one entanglement with higher fidelity with a certain success probability. Assume that the fidelity of entanglements in round z is f_z, the success probability of entanglement purification in round z and the fidelity of successfully purified entanglements are expressed with Eqn. (3) and Eqn. (4), respectively, as in Ref. [28]. The average number of entanglements needed to perform z rounds of success purification steps is denoted with η_z, which can be calculated with Eqn. (5) according to Ref. [29].

$$p_{z+1} = f_z^2 + \frac{2}{3}f_z(1-f_z) + \frac{5}{9}(1-f_z)^2 \tag{3}$$

$$f_{z+1} = \frac{f_z^2 + \frac{1}{9}(1-f_z)^2}{f_z^2 + \frac{2}{3}f_z(1-f_z) + \frac{5}{9}(1-f_z)^2} \tag{4}$$

$$\eta_z = \prod_{i=1}^{z}\frac{2}{p_i} \tag{5}$$

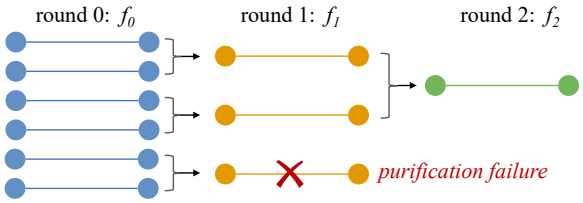
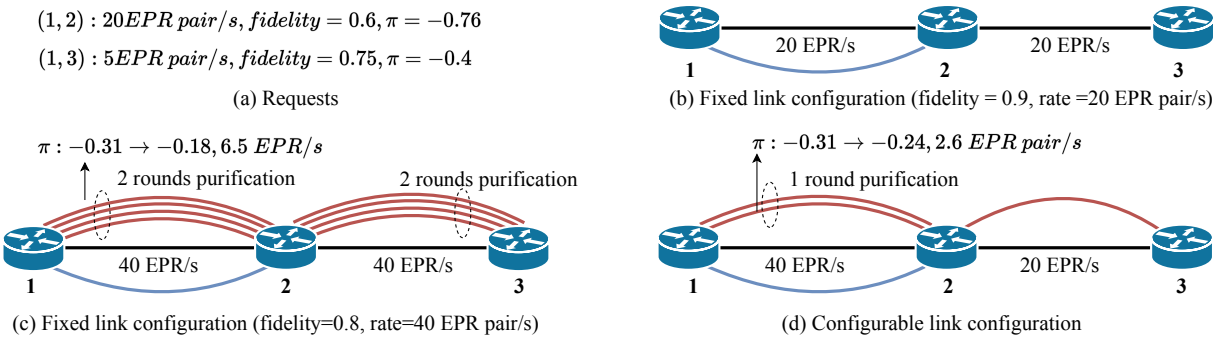


Fig. 2: Examples of two rounds of nested purification.

4) End-to-End Fidelity with Entanglement Swapping: After performing entanglement swapping, the fidelity of a newly-created entanglement between non-adjacent nodes can be calculated through the Werner parameters. Assume that we have a path ϕ and the Werner parameter of the entanglement in link e traversed by path ϕ is w_e. The Werner parameter of the newly created entanglement with swapping in the path is denoted with w_ϕ^{e2e}, which can be calculated with Eqn. (6) as proved in Ref. [30]. Assume that the requested fidelity is f^{req} and the corresponding Werner parameter w^{req} can be obtained from Eq. (2). To ensure that the generated entanglement has a fidelity higher than the requested fidelity, we could compare them using Werner parameters. The comparison of Werner parameters can be performed by checking if Eqn. (7) can be satisfied. If the logarithm of the Werner parameter in Eqn. (7) is pre-calculated, Eqn. (7) becomes an equation comprising only summations, which makes it possible to model the fidelity constraint with a MILP formulation.

$$w_\phi^{e2e} = \prod_{e\in k} w_e \tag{6}$$

$$\log(w^{req}) \leq \sum_{e\in k}\log(w_e) \tag{7}$$



(a) Requests

(b) Fixed link configuration (fidelity = 0.9, rate =20 EPR pair/s)

(c) Fixed link configuration (fidelity=0.8, rate=40 EPR pair/s)

(d) Configurable link configuration

Fig. 3: Illustration of a motivating example.

B. Problem Statement

The problem of link configuration for fidelity-constrained routing and purification in quantum networks can be summarized as: **Given** a quantum network topology and node memory capacity, a set of possible link configurations, a set of time-slots, the relation between fidelity and the number of required entanglements per purification round (in the form of a table), set of requests with the desired generation rate and fidelity, along with swapping success probability, **decide** the link configuration, routing and purification in each edge for each request, **constrained by** the generation rate of each link, node memory capacity constraint, and the fidelity constraint, with the **objective** of maximizing the served generation rate. Note that, since entanglements are susceptible to decoherence over time, a maximum storage time (defined as cutoff) that entanglements can be stored is typically enforced, as in Ref. [11], [30]. This work primarily focuses on selecting optimized link configurations rather than proposing new ER algorithms; therefore, for brevity, we assume the time duration considered is shorter than the cutoff time described in Ref. [11].

C. Motivating Illustrative Example

To emphasize the benefits of joint assignment of link configuration and per-link purification rounds, let us consider the example in Fig. 3 for a simple 3-node quantum network topology. In Fig. 3 (a), we consider two requests between: i) nodes (1, 2) with a generation rate of 20 ebits/s, fidelity equal to 0.6, and Werner parameter (in log scale) equal to -0.76; and ii) nodes (1,3) with 5 ebits/s generation rate, fidelity equal to 0.75 and Werner parameter (in log scale) equal to -0.4. As for possible link configurations, let us assume that links $(1,2)$ and $(2,3)$ can generate entanglements operating with fidelity of either 0.8 or 0.9 (whose Werner parameter is equal to -0.31 and -0.14) with a corresponding generation rate of 40 ebits/s and 20 ebits/s, respectively. We first show the cases with fixed initial fidelity as in Fig. 3 (b) and Fig. 3 (c). Specifically, Fig. 3 (b) shows that, for initial link fidelities fixed to 0.9, only request $(1,2)$ can be served since it already consumes all generation rate between node pair $(1,2)$. Fig. 3 (c) shows a scenario where the initial link fidelities are fixed to 0.8. In this case, 20 ebit/s of generation rate are consumed to serve request $(1,2)$. To serve request $(1,3)$, entanglements between both node pair $(1,2)$ and $(2,3)$ must perform two rounds of purification. Then, the Werner parameter changes from -0.31 to -0.18, requiring, on average, 6.5 ebits/s generation to create 1 ebits/s purified entanglements. In this case, only 3 ebits/s generation rate of request $(1,3)$ are consumed. Finally, Fig. 3 (d) shows the case that link can be configured with different link configurations. Let us assume that the entanglement fidelity between node pair $(1,2)$ and $(2,3)$ are set to 0.8 and 0.9, respectively. Request $(1,2)$ can be served with the entanglements generated between node pair $(1,2)$ since the logarithm of the Werner parameter of the generated entanglements (-0.31) is greater than the logarithm of the Werner parameter of the requested entanglements (-0.76), consuming 20 ebits/s of the generation rate between node pair $(1,2)$. The request between node pair $(1,3)$ cannot be served without purification because the logarithm of the Werner parameter of entanglements created for node pair $(1,3)$ after swapping equals (-0.31) + (-0.14) = -0.44, which is smaller than the Werner parameter of the request (-0.4). After performing 1 round of purification for the entanglements between node pair $(1,2)$, the Werner parameter changes from -0.31 to -0.24, requiring on average 2.6 ebits/s to get 1 ebits/s purified entanglements. Serving the request $(1,3)$ requires having 5 ebits/s generation rate of entanglements after purification, which consumes 13 ebits/s generation rate for node pair $(1,2)$. The total consumed generation rate for node pair $(1,2)$ equals 33 ebits/s. In this case, all requests are served.

In summary, this example illustrates that joint per-link assignment of link fidelity and rounds of purification can, in general, enable higher acceptance rates in fidelity-constrained quantum networks, compared to conventional scenarios where the starting link fidelities before purification are fixed.

IV. Mixed Integer Linear Programming Model for Quantum Entanglement Routing

This Section presents the proposed MILP model for the LC-FCRP problem.

A. Decision Variables and Objective Function

Sets, parameters, and decision variables of the proposed MILP are illustrated in Table I and Table II, respectively. The opposite direction of link e is denoted with $\overline{e}$.

Objective function: maximize the served generation rate:

$$\max \quad \sum_{d\in D}\sum_{k\in\Phi_d}\sum_{t\in T} x_k^{d,t} \tag{8}$$

TABLE I: Sets and Parameters for the MILP Model

Params	Description
N	Set of physical nodes
E	Set of unidirectional physical links
M_e	Set of indexes of possible link configurations for link e
D	Set of node pairs with requests
P	Set of rounds for purification
Φ_d	Set of paths for request $d \in D$
T	Set of time-slots
ζ_i	Memory capacity of node $i \in N$
σ_d^t	Requested entanglement generation rate for request $d \in D$ at time-slot $t \in T$
$\overline{\pi}_d^t$	Werner parameter in logarithmic scale for the minimum required fidelity of request d
$\pi_e^{m,p}$	Werner parameter in logarithmic scale for fidelity of link e if we choose link configuration m and purified p rounds
$\beta_e^{m,p}$	Success probability in logarithmic scale of purification in link $e \in E$ with link configuration m and purification of p rounds
$\delta_{k,e}^d$	Equals to 1 if the link $e \in E$ is in the path $k \in \Phi_d$ of request $d \in D$
$\eta_e^{m,p}$	Average number of entanglements required by link $e \in E$ with link configuration $m \in M_e$ to perform $p \in P$ rounds of purification successfully
$\xi_{e,m}$	Maximum entanglement generation rate of link $e \in E$ if e uses link configuration $m \in M_e$
γ_ϕ	Success probability of swapping in path $\phi \in \bigcup_{d\in D}\Phi_d$
θ	Time duration of one time-slot

TABLE II: Variables for the MILP Model

Variable	Description
$v_{e,m}$	Binary, equals to 1 if link $e \in E$ is configured with fidelity $m \in M_e$
$c_{e,m}^t$	Entanglement generation rate of link $e \in E$ if e uses link configuration $m \in M_e$
$q_{k,e}^{d,t}$	Binary, equals to 1 if request $d \in D$ uses path $k \in \Phi_d$ for swapping at time-slot $t \in T$ and path k traverses link $e \in E$
$r_k^{d,t}$	Binary, equals to 1 if path $k \in \Phi_d$ is chosen for request $d \in D$ at time-slot $t \in T$
$x_k^{d,t}$	Entanglement generation rate for request $d \in D$ in path $k \in \Phi_d$ at time-slot $t \in T$
h_e^t	Number of entanglement stored between node pair $e \in E$ at the beginning of time-slot $t \in T$
$g_{e,m,p}^{d,k,t}$	Binary, equals to 1 if request $d \in D$ is served by path $k \in \Phi_d$ at time-slot $t \in T$, where link $e \in E$ uses fidelity $m \in M_e$ and is purified for p rounds
$y_{e,m,p}^{d,k,t}$	Consumed entanglement generation rate in link e of path k for request d at time-slot $t \in T$, where entanglements in link e is created with fidelity $m \in M_e$ and purified p rounds

B. Constraints

1) Link configuration: Eqn. (9) ensures that each physical link $e \in E$ is configured with exactly one link configuration. Eqn. (10) ensures that the entanglement generation rate with fidelity $m \in M_e$ in edge $e \in E$ at time-slot $t \in T$ should be lower than the maximum generation rate of link e with link configuration m. Eqn. (10) ensures that the generation rate in both directions of a physical link is the same. Eqn. (12) ensures that for each node $i \in N$, the sum of the number of entanglements stored between this node and other nodes at the beginning of the time-slot and the entanglements generated during the time-slot should be lower than the memory capacity at each time-slot $t \in T$.

$$\sum_{m\in M_e} v_{e,m} = 1 \quad \forall e \in E \tag{9}$$

$$c_{e,m}^t \le \xi_{e,m} v_{e,m} \quad \forall e \in E, m \in M_e, t \in T \tag{10}$$

$$c_{e,m}^t = c_{\overline{e},m}^t \quad \forall e \in E, m \in M_e, t \in T \tag{11}$$

$$\sum_{e\in\delta^+(i)} h_e^t + \sum_{e\in\delta^+(i)}\sum_{m\in M_e} c_{e,m}^t * \theta \le \zeta_i \quad \forall i \in N, t \in T \tag{12}$$

2) Route formulation to generate entanglements: Eqn. (13) ensures that $q_{k,e}^{d,t}$ equals to 1 if path $k \in \Phi_d$ is selected for request $d \in D$ at time-slot $t \in T$ and physical link $e \in E$ is in path k.

$$q_{k,e}^{d,t} = r_k^{d,t} * \delta_{k,e}^d \quad \forall d \in D, k \in \Phi_d, e \in E, t \in T \tag{13}$$

3) Fidelity Constraints: Eqn. (14) determines the link configuration m and the rounds of purification for entanglement in link $e \in E$ to serve request $d \in D$ at each time-slot $t \in T$. Eqn. (15) ensures fidelity of link e for request d should be consistent with the link configuration for link e at each time-slot $t \in T$. Eqn. (16) ensures that the fidelity of the entanglement created for request d should be greater or equal to the required fidelity at each time-slot $t \in T$.

$$\sum_{m\in M_e}\sum_{p\in P} g_{e,m,p}^{d,k,t} = q_{k,e}^{d,t} \quad \forall d \in D, k \in \Phi_d, e \in E, t \in T \tag{14}$$

$$g_{e,m,p}^{d,k,t} \le v_{e,m} \quad \forall d \in D, k \in \Phi_d, e \in E, m \in M_e, p \in P, t \in T \tag{15}$$

$$\sum_{e\in E}\sum_{m\in M_e}\sum_{p\in P} \pi_e^{m,p} * g_{e,m,p}^{d,k,t} \ge \overline{\pi}_d * r_k^{d,t} \quad \forall d \in D, k \in \Phi_d, t \in T \tag{16}$$

4) Capacity Constraints: Eqn. (17) ensures that the entanglement generation rate provided in each edge e for path k should be larger or equal to the requested generation rate in the path for entanglement swapping. Eqn. (18) ensures that the served entanglement generation rate for request d should be less than or equal to the requested generation rate σ_d. Eqn. (19) ensures that for each physical link $e \in E$, the sum of the entanglement rate used by all the requests should not exceed the entanglement rate of the link e. In this work, we assume that the Bell Statement Measurement (BSM) is perfect [31],

and hence the entanglement swapping operation does not have failure probability. Note that our MILP formulation is extensible to the case with imperfect BSM. Specifically, $\eta_e^{m,p}$ in Eqn. (19) needs to be divided with the success probability y_ϕ to get the expected number of entanglements required. Since the success probability y_ϕ can be pre-calculated as a constant as in Ref. [19], Eqn. (19) is still a linear constraint in case of imperfect BSM.

$$\sum_{m\in M_e}\sum_{p\in P}\frac{y_{e,m,p}^{d,k,t}}{\eta_e^{m,p}} = x_k^{d,t}\delta_{k,e}^d \quad \forall d\in D, k\in\Phi_d, e\in E, t\in T \tag{17}$$

$$\sum_{k\in\Phi_d} x_k^{d,t} \le \sigma_d^t \quad \forall d\in D, t\in T \tag{18}$$

$$\sum_{d\in D}\sum_{k\in\Phi_d}\sum_{m\in M_e}\sum_{p\in P}(y_{e,m,p}^{d,k,t}+y_{\bar{e},m,p}^{d,k,t}) \le h_e^t + \sum_{m\in M_e} c_{e,m}^t \quad \forall e\in E, t\in T \tag{19}$$

5) Quantum Memory Constraint: Eqn. (20) initializes the entanglements stored in the quantum memory. Eqn. (21) ensures the entanglements stored between one node pair $e \in E$ must be greater or equal to 0. Eqn. (22) ensures that the number of entanglements stored between node pair e is consistent with the number of entanglements stored between the node pair represented with opposite direction, namely, $\bar{e}$. Eqn. (23) ensures that the number of entanglements available at the beginning of the time-slot t equals the number of entanglements available in the previous time-slot minus the entanglements used to use requests.

$$h_e^1 = 0 \quad \forall e\in E \tag{20}$$

$$h_e^t \ge 0 \quad \forall e\in E, t\in T\setminus\{1\} \tag{21}$$

$$h_e^t = h_{\bar{e}}^t \quad e\in E, t\in T \tag{22}$$

$$h_e^t \le h_e^{t-1} - \sum_{d\in D}\sum_{k\in\Phi_d}\sum_{m\in M_e}\sum_{p\in P}\theta*(y_{e,m,p}^{d,k,t-1}+y_{\bar{e},m,p}^{d,k,t-1}) + \sum_{m\in M_e}\theta*c_{e,m}^{t-1} \quad \forall e\in E, t\in T\setminus\{1\} \tag{23}$$

V. Optimized Link Configuration Algorithm for Entanglement Routing

Since the MILP model reported in Section IV applies only to a finite number of link configurations and has well-known scalability limitations, we propose a novel link configuration algorithm to determine the link configuration and then perform entanglement routing with purification.

A. Generalized Procedure of Entanglement Routing

Extending the MILP formulations to the continuous decision space of fidelity in link configuration is challenging due to the following two main reasons. First, the purification process is an iterative process given an initial fidelity before purification, which is infeasible to be formulated into fidelity-constrained ER in closed form since there are an infinite number of possible initial link configurations for links. Second, even with a fixed link configuration and no purification, the LC-FCRP problem is a multi-commodity flow problem, which is NP-hard [32]. Considering the continuous decision space for fidelity, the LC-FCRP problem becomes even more challenging than the already NP-hard multi-commodity flow problem.

To solve the LC-FCRP problem, we first develop a shortest-path-based fidelity determination (SPFD) algorithm to obtain an optimized initial link configuration that improves the performance of ER. Then, the link configuration obtained from the SPFD algorithm is utilized by a subsequent step based on Bayesian Optimization (BO) to improve the link configuration for ER. We apply BO to improve the link configuration since our problem fits into the criteria of BO (no closed-form expression, expensive to evaluate, no available derivatives, etc.) [33], and BO is widely used to tune parameters for various systems. BO iteratively searches for a good link configuration. Specifically, in each iteration, our Bayesian Optimizer suggests a promising link configuration, and then the suggested link configuration is evaluated with any ER algorithm.

B. Link-Configuration Determination with SPFD

The SPFD algorithm is shown in Algorithm 1, which determines link configuration greedily to serve as many requests as possible. The SPFD algorithm contains two steps. The first step estimates the *resource demand* (requested generation rate and the corresponding fidelity) of each edge with a shortest-path-based approach (line 1-9). The second step greedily determines the link configurations for each link based on the resource demand (line 10-25). Specifically, in the first step, the resource demand of one request is assigned to the edges traversed by the shortest path for the request. More specifically, after finding the shortest path and obtaining the number of hops of the path (line 3-4), SPFD associates the average requested generation rate σ_d over all time-slots and the required fidelity to each traversed edge (line 5-8). The fidelity constraint is guaranteed through constraints with the Werner parameter, and the requested minimum value of the Werner parameter in each edge of path k is approximated with $\overline{\pi}_d/h_d$ (line 7). In the second step, the algorithm finds the link fidelity that maximizes the sum of the entanglement generation rates of all requests $\overline{y}_e^{max}$ as follows. For each link e, the SPFD algorithm first initializes $\overline{y}_e^{max}$ to 0, and a set of candidate link configurations M_e, where the fidelity of each configuration has the associate Werner parameter equal to $\overline{\pi}_d/h_d$ (line 11-12). Then, for each link configuration m, SPFD tries to iterate over each request $\overline{d}$ and get the served generation rate $y_{e,m,p}^{d,k}$ for d on shortest path k (line 15-19). If the served generation rate with link configuration m is greater than $\overline{y}_e^{max}$, SPFD sets the link e with link configuration m (line 20-23).

C. Improvement of Link Configuration with BO

The link configuration of the SPFD algorithm is set as a starting point for BO to improve the link configuration further as in Algorithm. 2. In our problem, the Bayesian Optimizer

Algorithm 1: SPFD to determine link configuration

Input: $G = (N, E), D$
Output: Link configurations for all the links
1 Initialize $Dic[e] = []$ for each edge e
2 **for** *each request* $d \in D$ **do**
3 Find the shortest path k for request d
4 Get the number of hops h_d for the shortest path
5 **for** *each edge* e *in the shortest path* k **do**
6 Let $\sigma_d \leftarrow \sum_{t \in T} \sigma_d^t / |T|$
7 Append $(\sigma_d, \overline{\pi}_d / h_d)$ to $Dic[e]$
8 **end**
9 **end**
10 **for** *each edge* e **do**
11 Get D_e, the requests traversing edge e from $Dic[e]$
12 Let $\overline{y}_e^{max} \leftarrow 0$, initialize M_e for each link configuration with logarithm of Werner Parameter equal to $\overline{\pi}_d / h_d$
13 **for** *each* $m \in M_e$ **do**
14 Calculate the generation rate $\xi_{e,m}$ and π_e^m
15 **for** *each request* $d \in D_e$ **do**
16 **if** $\xi_{e,m} = 0$ **then** Break
17 Get minimum p such that $\pi_e^{m,p} \geq \overline{\pi}_d / h_d$
18 Let $y_{e,m,p}^{d,k} \leftarrow min(\sigma_d, \xi_{e,m} / \eta_{e,m,p}^{d,k})$, $\xi_{e,m} \leftarrow \xi_{e,m} - \eta_{e,m,p}^{d,k} * y_{e,m,p}^{d,k}$
19 **end**
20 **if** $\sum_{d \in D_e} \sum_{p \in P} y_{e,m,p}^{d,k} \geq \overline{y}_e^{max}$ **then**
21 Set the fidelity of link e to link configuration m
22 Let $\overline{y}_e^{max} \leftarrow \sum_{d \in D_e} \sum_{p \in P} y_{e,m,p}^{d,k}$
23 **end**
24 **end**
25 **end**

Algorithm 2: BO for improved link configuration.

Input: $G = (N, E), D$
Output: Link configuration for all the edges
1 Initialize Bayesian Optimizer with $|E|$ variables
2 Set the bounds of link fidelity to $(0.5, 1)$
3 Get initial link configuration with SPFD algorithm
4 Obtain generation rate with **ER algorithm**
5 Feed the initial link configuration and generation rate to Bayesian Optimizer
6 Initialize $i \leftarrow 0$
7 **while** $i <$ *maximum number of iterations* **do**
8 Obtain a new link configuration m from Bayesian Optimizer
9 Get the generation rate using m with **ER algorithm**
10 Feed the generation rate to Bayesian Optimizer
11 Perform domain reduction to adapt the bounds of fidelity
12 Let $i \leftarrow i + 1$
13 **end**

considers $|E|$ continuous variables, corresponding to the fidelity values to be set for each edge (line 1-2). Specifically, we set the initial fidelity values to the ones returned by the SPFD algorithm, and we define the bounds of fidelity for link configuration as $(0.5, 1)$, since only fidelity values above 0.5 are useful for practical purposes [7]. Then the link configuration obtained with SPFD is fed to BO as the starting point (line 3-5). Then the BO iteratively searches link configurations (line 6-13). The LC-FCRP problem, being highly non-linear, often leads to stability and convergence issues in naive Bayesian Optimization. To address these challenges, we apply Domain Reduction (line 11) as detailed in [34], which refines the search space into manageable sub-regions.

VI. Illustrative Numerical Results

A. Simulation Setup

The simulations are performed on a server with Intel(R) Xeon(R) CPU E5-2660 v4 CPU (14 Cores @ 2.00GHz) and 125 GB of memory. We implement the MILP formulation with AMPL (A Mathematical Programming Language) [35] and solve it with CPLEX MIP solver 22.1.1. The proposed link configuration algorithm and the ER algorithms are implemented with Python 3.8. In our evaluations, we consider two topologies, a small topology (the German topology in Ref. [36] with 7 nodes and 11 links) and a large topology (the USnet topology in Ref. [37] with 24 nodes and 43 links). We assume each link of both topologies can generate entanglements with a fidelity of 0.8 with a generation rate uniformly distributed in [250, 500] EPR pair/s similar to Ref. [11], and the generation rate and fidelity in different link combinations can be derived according to Sec. III-B. The memory storage is set to 12,000 as in Ref. [11]. The time duration of each time slot is 10 seconds. The number of considered time slots for German and USnet topology is set to 2 and 8, respectively. For German topology, we consider that the entanglement requests are randomly generated between 75% of the node pairs in the network with an average requested fidelity of 0.9. To extend our evaluation, we compare different algorithms in a larger topology where requests are randomly generated among 50% of the node pairs. We first evaluate performance under varying loads, with an average request requested fidelity of 0.9. Then, we evaluated three fidelity scenarios under a requested generation rate of around 700 EPR pair/s: low fidelity scenario, medium fidelity scenario, and high fidelity scenario, with average requested fidelities of 0.80, 0.85, and 0.90, respectively.

We evaluate the improvement brought by link configuration on the performance of three baseline ER algorithms, namely, *Progressive Filling* (PF) algorithm [21], *Low-complexity Routing* (Qleap) algorithm [10], and *Purification-enabled Iterative Routing* (Qpath) [10] algorithm. Specifically, PF performs purification on each link to achieve the same pre-defined fidelity threshold on all the links, while Qleap performs purification on links to achieve a fidelity threshold according to the fidelity required by each request. Moreover, Qpath is an advanced version of Qleap, which performs purification on critical links that have the highest improvement in fidelity with the least entanglements. For these baseline ER algorithms, we assume that all the links generate entanglements with a fixed fidelity sampled from $\mathcal{N}[0.8, 0.1]$ as in Ref. [10]. We evaluate setting link configuration with a combination of the proposed SPFD algorithm with/without BO and one of the existing ER algorithms. The cases where SPFD is applied without BO and together with PF, Qpath and Qleap are named SPFD-PF, SPFD-Qleap, and SPFD-Qpath, respectively. When BO is incorporated to SPFD, their combination with

PF, Qpath, and Qleap are named BO-PF, BO-Qleap, and BO-Qpath, respectively. The simulation results are averaged over 10 instances.

B. Performance Evaluation on a Small Topology

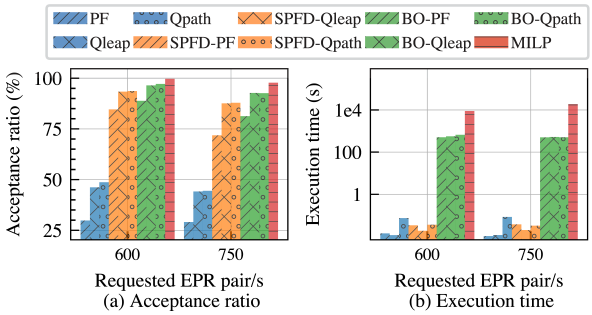


Fig. 4: Performance evaluation in German topology.

Fig. 4 (a) shows the acceptance ratio (defined as the percentage of served generation rate over the total requested generation rate) under two different requested generation rates for the German topology. It is observed that link configuration with SPFD can improve the acceptance ratio achieved with PF, Qleap, and Qpath by up to around 55%, 47%, and 45%, respectively. Moreover, the incorporation of BO to SPFD can further improve the acceptance ratio of SPFD-PF, SPFD-Qleap, and SPFD-Qpath by up to around 9%, 5%, and 5%, respectively. Among all the heuristic algorithms, the highest acceptance ratio is achieved with BO-Qpath, which has an optimality gap of less than 6%.

The low optimality gap of the proposed approach is achieved with a significant execution time reduction compared to MILP, as shown in Fig. 4 (b). For instance, the execution time of ILP is reduced from around 19000 seconds to around 500 seconds under SPFD link configuration with BO, and less than 1 second for link configuration with SPFD alone.

C. Performance Evaluation on a Larger Topology

1) Evaluation with Varying Loads: In Fig. 5 (a), 5 (b) and Fig. 5 (c), we show the impact of the heuristics on, respectively, the network acceptance ratio, connection end-to-end fidelity, and resource utilization under a large range of requested generation rates. As it can be observed in Fig. 5 (a), the employment of link configuration with SFPD can significantly improve the acceptance ratio of the three baseline ER algorithms (PF, Qleap, and Qpath), where values up to around 83%, 81%, and 77%, respectively, were observed. Moreover, link configuration with BO can further improve the acceptance achieved with SPFD-PF, SPFD-Qleap, and SPFD-Qpath by up to around 3.5%, 1.3%, and 1.4%, respectively. When the entanglement routing approaches are compared, the performance of the three baseline ER algorithms keeps the same trend for the cases w/o link configuration for both cases with or without BO. For instance, with link configuration using BO, BO-Qpath achieves an acceptance ratio of around 12% and 5% higher than *BO-PF* and *BO-Qleap*, respectively. This occurs because Qpath has more advanced schemes for purification.

Fig. 5 (b) shows the end-to-end fidelity when varying the requested generation rate. When the requested generation rate is low, the end-to-end fidelity for all cases with link configuration using both SPFD and BO is close to 0.94, significantly higher than the end-to-end fidelity achieved with three baseline ER algorithms. For instance, the end-to-end fidelity achieved with SPFD-Qpath is up to 0.04 higher than that achieved with Qpath. This is because to serve requests between node pairs with a large number of hops in their routes, the link must be properly configured to generate entanglements with high fidelity, as the end-to-end fidelity decays with the increase in the number of entanglement swapping occurrences. When the requested generation rate increases, SPFD w/o BO tends to set the link configuration to generate entanglements with lower fidelity but a higher generation rate to achieve a higher acceptance ratio, as the lower fidelity is sufficient to meet the end-to-end fidelity requirements for requests between node pairs with fewer hops. Note that the end-to-end fidelity achieved by PF does not follow a decreasing trend as in the other cases because PF performs purification on each edge to meet a pre-defined fidelity threshold, rather than adjusting purification based on request variations like the other baseline ER algorithms. Consequently, PF maintains a consistent end-to-end fidelity of around 0.92. This is because the fidelity before swapping also aligns with the set threshold, and PF prioritizes requests with fewer hops, resulting in a relatively stable number of hops for served requests.

The resource utilization under different loads is shown in Fig. 5 (c). The resource utilization increases for all cases as network load rises because resources that cannot serve requests between node pairs with a larger number of hops (where fidelity deterioration prevents meeting the required fidelity constraints) can instead be used to attend an increased number of demands between node pairs with few hops. However, the relative performance of the different ER algorithms using link configuration might be different from that of baseline ER algorithms (i.e., without link configuration), as the resource efficiency of different algorithms might change with the employment of link configuration. For instance, Qleap has a lower resource utilization and a lower acceptance ratio than Qpath, while BO-Qleap has a higher resource utilization than BO-Qpath, although BO-Qleap has a lower acceptance ratio than BO-Qpath. This is because BO-Qleap and BO-Qpath tend to select link configurations that generate entanglement with a higher fidelity. In this case, not all the links need to be purified to guarantee the fidelity constraint, while Qpath still purifies all the links in a path for a request in order to achieve the fidelity threshold, leading to unnecessary purifications. This is also confirmed in Fig. 5 (b), where the end-to-end fidelity achieved by BO-Qleap is higher than that of BO-Qpath.

2) Evaluation with Different Fidelity Scenarios: Fig. 6 shows the performance under different fidelity scenarios. When the requested fidelity increases, the acceptance ratio of all approaches decreases and BO-Qpath always achieves

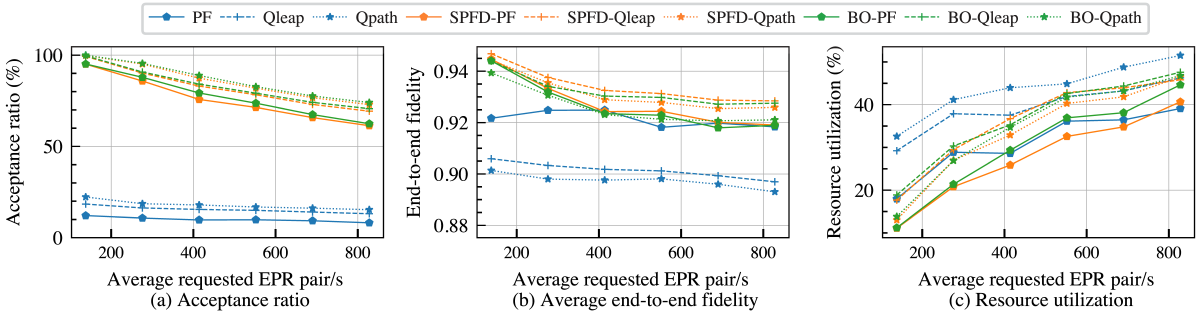


Fig. 5: Performance of link configuration for different load scenarios in the USnet topology.

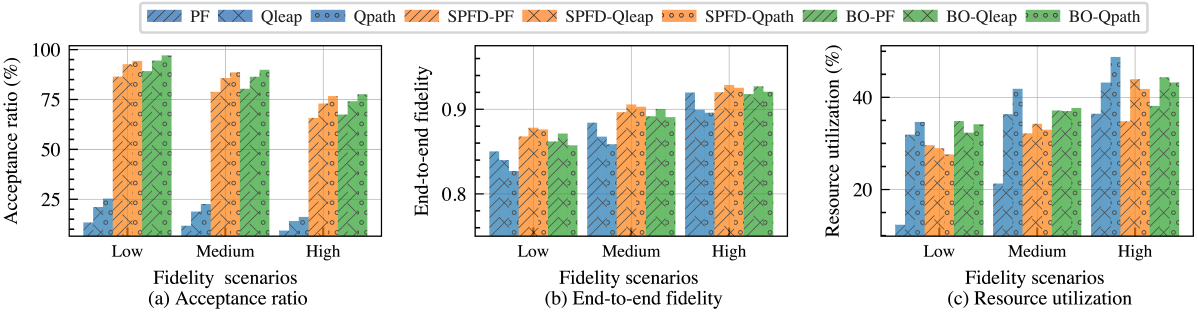


Fig. 6: Performance of link configuration for different fidelity scenarios in the USnet topology.

superior acceptance ratios, as shown in Fig. 6 (a). For instance, under the low-fidelity scenario, the acceptance ratio of BO-Qpath is around 97%, which is around 72% higher than that of Qpath. In the high-fidelity scenario, although the acceptance ratio of BO-Qpath is reduced to around 78%, it is still around 62% higher than that of Qpath. In addition to the improvement in acceptance ratio, the ER algorithms with link configuration also have higher end-to-end fidelity in all fidelity scenarios, as shown in Fig. 6 (b), which highlights the benefits of setting proper link configurations. For instance, in the medium-fidelity scenario, the end-to-end fidelity of BO-Qpath and SPFD-Qpath is approximately 0.03 and 0.04 higher than that of Qpath, respectively. It is worth noticing that the end-to-end fidelity achieved with Qpath is always smaller than that achieved with Qleap regardless of the adoption of link configuration, as Qpath has optimized purification and performs purification only when it is necessary to provide end-to-end fidelity higher than the requested one. In addition, PF achieves a higher end-to-end fidelity than Qleap and Qpath, while PF with link configuration may achieve a lower end-to-end fidelity than Qleap and Qpath with link configuration. For instance, under the medium-fidelity scenario, the end-to-end fidelity achieved by SPFD-PF is around 0.01 lower than those of SPFD-Qleap and SPFD-Qpath. This is because SPFD-Qleap and SPFD-Qpath configure links to generate high-fidelity entanglements for requests with many hops. The fidelity of these entanglements often exceeds the threshold required by SPFD-PF, resulting in less purification by SPFD-PF and thus lower end-to-end fidelity. The resource utilization for different fidelity scenarios is shown in Fig. 6 (c). When the requested fidelity increases, the resource utilization of all ER algorithms with and without link configuration also increases. In addition, similar to the results in Fig. 5 (c), the relative performance of different ER algorithms with or without SPFD or BO may vary, as the resource efficiency of these algorithms is influenced by different initial link configurations.

VII. Conclusion

This paper introduced the Link Configuration for Fidelity-Constrained Routing and Purification (LC-FCRP) problem in quantum networks, addressing the challenge of improving the quality of distributed entanglements. Through link configurations, we greatly improve the performance of the entanglement routing process. We formulated a simplified version of the LC-FCRP problem as a MILP model and proposed a scalable two-phase link configuration algorithm with a low optimality gap. Numerical results indicate that the proposed link configuration algorithm can improve the acceptance ratio achieved by existing entanglement routing algorithms by up to 87%. Future work involves considering real application scenarios and exploring the performance of the proposed approaches under realistic requirements of fidelity.

Acknowledgment

This work was supported in part by funding from the Innovation for Defence Excellence and Security (IDEaS) program from the Canadian Department of National Defence (DND), and in part by the project SERICS (PE00000014) under the MUR National Recovery and Resilience Plan funded by the European Union - NextGenerationEU.

REFERENCES

[1] C. H. Bennett and G. Brassard, "Quantum cryptography: Public key distribution and coin tossing," *Theoretical Computer Science*, vol. 560, pp. 7–11, 2014.

[2] A. K. Ekert, "Quantum cryptography based on bell's theorem," *Physical review letters*, vol. 67, no. 6, p. 661, 1991.

[3] Q. Zhang, O. Ayoub, A. Gatto, J. Wu, F. Musumeci, and M. Tornatore, "Routing, channel, key-rate, and time-slot assignment for qkd in optical networks," *IEEE Transactions on Network and Service Management*, vol. 21, no. 1, pp. 148–160, 2024.

[4] J. I. Cirac, A. Ekert, S. F. Huelga, and C. Macchiavello, "Distributed quantum computation over noisy channels," *Physical Review A*, vol. 59, no. 6, p. 4249, 1999.

[5] V. Giovannetti, S. Lloyd, and L. Maccone, "Quantum-enhanced measurements: beating the standard quantum limit," *Science*, vol. 306, no. 5700, pp. 1330–1336, 2004.

[6] T. Vidick and S. Wehner, *Introduction to quantum cryptography*. Cambridge University Press, 2023.

[7] M. A. Nielsen and I. L. Chuang, *Quantum computation and quantum information*. Cambridge university press Cambridge, 2001, vol. 2.

[8] A. Einstein, B. Podolsky, and N. Rosen, "Can quantum-mechanical description of physical reality be considered complete?" *Phys. Rev.*, vol. 47, pp. 777–780, May 1935.

[9] Y. Zhao, G. Zhao, and C. Qiao, "E2e fidelity aware routing and purification for throughput maximization in quantum networks," in *IEEE Conference on Computer Communications (INFOCOM)*. IEEE, 2022, pp. 480–489.

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

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

[12] Z. Liu, S. Marano, and M. Z. Win, "Establishing high-fidelity entanglement in quantum repeater chains," *IEEE Journal on Selected Areas in Communications*, vol. 42, no. 7, pp. 1763–1778, 2024.

[13] P. C. Humphreys, N. Kalb, J. P. Morits, R. N. Schouten, R. F. Vermeulen, D. J. Twitchen, M. Markham, and R. Hanson, "Deterministic delivery of remote entanglement on a quantum network," *Nature*, vol. 558, no. 7709, pp. 268–273, 2018.

[14] J. Illiano, M. Caleffi, A. Manzalini, and A. S. Cacciapuoti, "Quantum internet protocol stack: A comprehensive survey," *Computer Networks*, p. 109092, 2022.

[15] R. Van Meter, *Quantum networking*. John Wiley & Sons, 2014.

[16] S. Santos, F. A. Monteiro, B. C. Coutinho, and Y. Omar, "Shortest path finding in quantum networks with quasi-linear complexity," *IEEE Access*, 2023.

[17] G. Vardoyan and S. Wehner, "Quantum network utility maximization," in *IEEE International Conference on Quantum Computing and Engineering (QCE)*, vol. 1. IEEE, 2023, pp. 1238–1248.

[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] Y. Zhao and C. Qiao, "Redundant entanglement provisioning and selection for throughput maximization in quantum networks," in *IEEE Conference on Computer Communications (INFOCOM)*. IEEE, 2021, pp. 1–10.

[20] A. Farahbakhsh and C. Feng, "Opportunistic routing in quantum networks," in *IEEE Conference on Computer Communications (INFOCOM)*. IEEE, 2022, pp. 490–499.

[21] C. Li, T. Li, Y.-X. Liu, and P. Cappellaro, "Effective routing design for remote entanglement generation on quantum networks," *npj Quantum Information*, vol. 7, no. 1, p. 10, 2021.

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

[23] Z. Wang, J. Li, K. Xue, D. S. L. Wei, R. Li, N. Yu, Q. Sun, and J. Lu, "An efficient scheduling scheme of swapping and purification operations for end-to-end entanglement distribution in quantum networks," *IEEE Transactions on Network Science and Engineering*, vol. 11, no. 1, pp. 380–391, 2024.

[24] H. Hu, H. Lun, Z. Deng, J. Tang, J. Li, Y. Cao, Y. Wang, Y. Liu, D. Wu, H. Yu, X. Wang, J. Wei, and L. Shi, "High-fidelity entanglement routing in quantum networks," *Results in Physics*, vol. 60, p. 107682, 2024.

[25] L. Chen and Z. Jia, "On optimum entanglement purification scheduling in quantum networks," *IEEE Journal on Selected Areas in Communications*, vol. 42, no. 7, pp. 1779–1792, 2024.

[26] B. C. Coutinho, R. Monteiro, L. Bugalho, and F. A. Monteiro, "Entanglement routing based on fidelity curves for quantum photonics channels," *arXiv preprint arXiv:2303.12864*, 2023.

[27] Y. Pu, N. Jiang, W. Chang, H. Yang, C. Li, and L. Duan, "Experimental realization of a multiplexed quantum memory with 225 individually accessible memory cells," *Nature communications*, vol. 8, no. 1, p. 15359, 2017.

[28] C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. A. Smolin, and W. K. Wootters, "Purification of noisy entanglement and faithful teleportation via noisy channels," *Physical review letters*, vol. 76, no. 5, p. 722, 1996.

[29] W. Dür, H.-J. Briegel, J. I. Cirac, and P. Zoller, "Quantum repeaters based on entanglement purification," *Physical Review A*, vol. 59, no. 1, p. 169, 1999.

[30] Á. G. Iñesta, G. Vardoyan, L. Scavuzzo, and S. Wehner, "Optimal entanglement distribution policies in homogeneous repeater chains with cutoffs," *npj Quantum Information*, vol. 9, no. 1, p. 46, 2023.

[31] M. Pompili, S. L. Hermans, S. Baier, H. K. Beukers, P. C. Humphreys, R. N. Schouten, R. F. Vermeulen, M. J. Tiggelman, L. dos Santos Martins, B. Dirkse *et al.*, "Realization of a multinode quantum network of remote solid-state qubits," *Science*, vol. 372, no. 6539, pp. 259–264, 2021.

[32] 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.

[33] Z. Wang, M. Zoghi, F. Hutter, D. Matheson, N. De Freitas *et al.*, "Bayesian optimization in high dimensions via random embeddings." in *IJCAI*, vol. 13, 2013, pp. 1778–1784.

[34] N. Stander and K. J. Craig, "On the robustness of a simple domain reduction scheme for simulation-based optimization," *Engineering Computations*, vol. 19, no. 4, pp. 431–450, 2002.

[35] R. Fourer, D. M. Gay, and B. W. Kernighan, "Ampl: A mathematical programming language," *Management Science*, vol. 36, no. 5, pp. 519–554, 1990.

[36] Q. Zhang, O. Ayoub, J. Wu, F. Musumeci, G. Li, and M. Tornatore, "Progressive slice recovery with guaranteed slice connectivity after massive failures," *IEEE/ACM Transactions on Networking*, vol. 30, no. 2, pp. 826–839, 2021.

[37] S. Ferdousi, M. Tornatore, F. Dikbiyik, C. U. Martel, S. Xu, Y. Hirota, Y. Awaji, and B. Mukherjee, "Joint progressive network and datacenter recovery after large-scale disasters," *IEEE Transactions on Network and Service Management*, vol. 17, no. 3, pp. 1501–1514, 2020.