

Optimizing Satellite-based Entanglement Distribution in Quantum Networks via Quantum-Assisted Approaches

Xinliang Wei*, Lei Fan^{†‡}, Yuanxiong Guo[§], Zhu Han[‡], Yu Wang*

*Department of Computer and Information Sciences, Temple University, Philadelphia, PA, USA

[†]Department of Engineering Technology, University of Houston, Houston, TX, USA

[‡]Department of Electrical and Computer Engineering, University of Houston, Houston, TX, USA

[§]Department of Information Systems and Cyber Security, University of Texas at San Antonio, San Antonio, TX, USA
{xinliang.wei, wangyu}@temple.edu, {lfan8, zhan2}@uh.edu, yuanxiong.guo@utsa.edu.

Abstract—Satellite-based quantum networks leverage the advantageous properties of optical signals from satellites to ground stations, to enable the distribution of high-fidelity quantum entanglements over vast distances, thus circumventing the limitations of traditional terrestrial-based systems. However, the satellite-based entanglement distribution coupled with terrestrial quantum swapping in the space-terrestrial integrated network becomes very complex when considering the joint optimization with satellite assignment, resource allocation, and path selections. To tackle this issue, we introduce a hybrid quantum-classical Dantzig-Wolfe decomposition technique by leveraging the strengths of both quantum and classical computing, to solve the joint optimization problem. Through a series of experiments, the paper demonstrates the efficiency and robustness of the proposed methods in addressing large-scale network optimization and balancing qubit usage. The insights generated by this work offer valuable guidance for the design and implementation of satellite-based entanglement distribution for quantum networks, thereby laying the groundwork for the realization of a secure quantum communication infrastructure on a global scale.

Index Terms—Entanglement distribution, quantum swapping, quantum networks, hybrid quantum-classical optimization, space-terrestrial integrated network

I. INTRODUCTION

Quantum entanglement distribution has been one of the foundational components for all quantum communication protocols and systems. However, ground-based distribution methods have faced limitations and drawbacks that hinder their effectiveness. These limitations include susceptibility to environmental disturbances such as atmospheric interference and fiber attenuation, which can degrade the fidelity of entangled states and limit the achievable distances for entanglement distribution. Moreover, ground-based systems are also vulnerable to eavesdropping. An alternative approach

The work is partially supported by the US National Science Foundation (under Grant No. CNS-2006604, CNS-2106761, CNS-2107216, CNS-2128368, CNS-2128378, CMMI-2222670, CMMI-2222810, and ECCS-2302469), the US Department of Transportation, Toyota, Amazon, and Japan Science and Technology Agency (JST) Adopting Sustainable Partnerships for Innovative Research Ecosystem (ASPIRE) under JPMJAP2326.

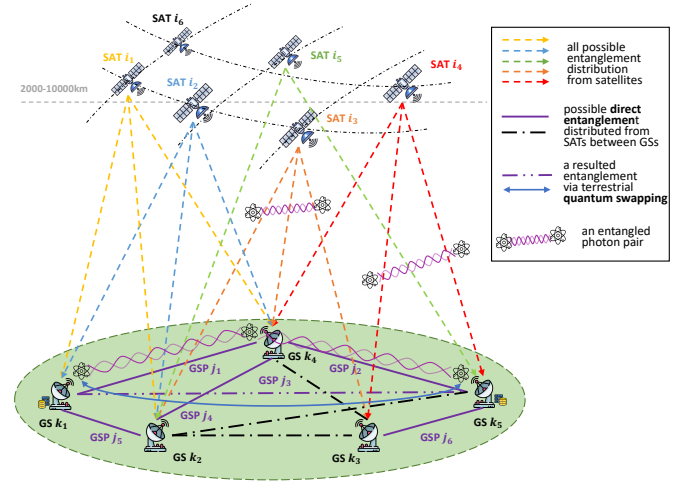


Fig. 1. The overall satellite-based entanglement distribution and terrestrial quantum swapping architecture.

that has gained attention in recent years is the use of satellite-based quantum entanglement distribution, as shown in Fig. 1. By leveraging space-based platforms, this approach offers several advantages over ground-based methods. First, satellite-based distribution can bypass many environmental factors that degrade entanglement fidelity on the ground, enabling the creation of long-distance entangled links with higher fidelity and greater security. Second, the capability to distribute entangled states from space allows for a global reach, overcoming the geographical constraints of ground-based systems. Generally, the architecture of a satellite-based quantum network involves the deployment of quantum-enabled satellites equipped with entangled photon sources and communication modules (e.g., photo transmitters). These satellites serve as relay nodes for the distribution of entangled states between ground stations equipped with photon receivers, enabling the creation of a global-scale quantum network. The satellite-based architecture enables the establishment of quantum links between geograph-

ically distant locations, facilitating secure communication and quantum key distribution on a global scale [1]–[3].

While the potential of satellite-based quantum entanglement distribution is promising, existing works [4]–[6] have faced limitations and challenges. For example, Khatri *et al.* [4] thoroughly examined the double down-link architecture for a satellite constellation in polar orbits, analyzing various satellite configurations to reduce the number of deployed satellites and increase entanglement distribution rates between ground stations. However, their proposed method is a heuristic greedy algorithm and ignores some resource constraints (e.g., number of transmitters/receivers) in their problem formulation. Following that, Panigrahy *et al.* [5] further considered various resource constraints at the satellites and ground stations. Still, they only solved them efficiently for some specific settings by converting the problem into either a maximum weight independent set problem or a maximum weight bipartite matching problem. Therefore, a more universal and flexible scenario needs to be considered, such as leveraging terrestrial quantum swapping to support the satellite-based quantum entanglement distribution.

Given the limited resources in satellite-based quantum networks, including the number of transmitters per satellite, photon sources, and receivers at ground stations, it is crucial to efficiently allocate network resources and schedule transmissions. In this paper, we investigate a joint satellite assignment, resource allocation, and path selection scenario to build quantum entanglements for ground stations by leveraging satellite-based entanglement distribution and terrestrial quantum swapping within the space-terrestrial integrated network (STIN). By integrating terrestrial quantum swapping into satellite-based entanglement distribution, it can serve more traffic ground station pairs and generate better quantum entanglement for further quantum applications. We then propose a new joint optimization model and adopt a novel hybrid quantum-classical Dantzig-Wolfe (HQCDW) decomposition technique [7] by leveraging the strengths of both quantum and classical computing, to balance the network resources and handle a large-scale network setting. Our proposed approaches aim to overcome the limitations of existing works and address the challenges associated with satellite-based entanglement distribution and quantum communication.

In summary, the highlights of this paper are as follows.

- We consider a satellite-based entanglement distribution with terrestrial quantum swapping to cooperatively generate entanglement for traffic ground station pairs and formulate a joint satellite assignment, resource allocation, and path selection problem modeled as integer non-linear programming (INLP), within STIN to maximize the total utility of all traffic ground station pairs. (Section III)
- We adopt a novel hybrid quantum-classical Dantzig-Wolfe decomposition algorithm (HQCDW) by leveraging the advantage of both quantum and classical computing to solve the complex INLP problem. The original optimization problem is decomposed into a master problem that is solved in a classical computer and several subproblems that are processed in a quantum annealer. (Section IV)

- We conduct extensive simulations using the commercial quantum annealer to evaluate our proposed algorithms. Numerous experiments have demonstrated that our proposed HQCDW can handle more complex network settings compared to the non-decomposed manner, and achieve the same result as the classical schemes but with shorter solver accessing time, which demonstrates the quantum advantage. (Section V)

The remainder of this paper is organized as follows. Section II introduces our system model and the joint optimization problem is formulated in Section III. Section IV presents our proposed HQCDW. Evaluations of the proposed method are provided in Section V and Section VI concludes the paper.

II. SYSTEM MODEL

Network Model. We consider a space-terrestrial integrated network (STIN) that consists of $|M|$ satellites (SAT), $|N|$ ground stations (GS), and total $|O|$ ground station pairs (GSP) as shown in Fig. 1, where $M = \{1, \dots, |M|\}$ indexed by i , $N = \{1, \dots, |N|\}$ indexed by k , $O = \{1, \dots, |O|\}$ indexed by j . We also define $O' \subseteq O$ as the traffic GSP set that needs to generate the entanglement link. In addition, let $O_k \subset O$ be the GSP set of GS k . We consider the polar satellite constellation and assume that each satellite has a quantum memory that can store at most sm_i entangled photon pairs. Each satellite is also equipped with tr_i transmitters that can transmit entangled photon pairs to multiple GSPs. Let $sm = \{sm_1, sm_2, \dots, sm_{|M|}\}$ be the quantum memory set of all satellites and we can find the maximal one $sm_{max} = \max\{sm_i, \forall i \in M\}$. Each GS also has a quantum memory gm_k for any quantum application such as quantum key distribution (QKD). Additionally, each GS owns rr_k receivers to receive the photon from the satellite or generate the entanglement link from other GSs by using quantum swapping. We also define an elevation angle threshold θ_e for any satellite and GSP. The entangled photon can be successfully received at the GS as long as the elevation angle between the satellite and the horizon at the GS exceeds this threshold.

Channel Model. In a STIN, there are typically multiple quantum communication paths as illustrated in [8], such as downlink and uplink channels between a ground station and a satellite and lateral paths between two ground stations or two satellites. This work mainly focuses on the dual down-link architecture for photonic entanglement distribution with spontaneous parametric down-conversion (SPDC) [9]. Free-space optical transmission is a crucial aspect to consider in analyzing such links. Therefore, we take into account the optical channel's characteristics, including transmission loss and noise. The transmission of photons from satellites to ground stations can be effectively modeled using a bosonic pure-loss channel with transmittance. Based on previous works [4], [5], the space-to-ground transmittance $\eta_{i,k}^{sg}$ between SAT i and GS k consists of two parts: the free-space transmittance $\eta_{i,k}^{fs}$ and the atmospheric transmittance $\eta_{i,k}^{atm}$, and can be defined as

$$\eta_{i,k}^{sg} = \eta_{i,k}^{fs} \cdot \eta_{i,k}^{atm}. \quad (1)$$

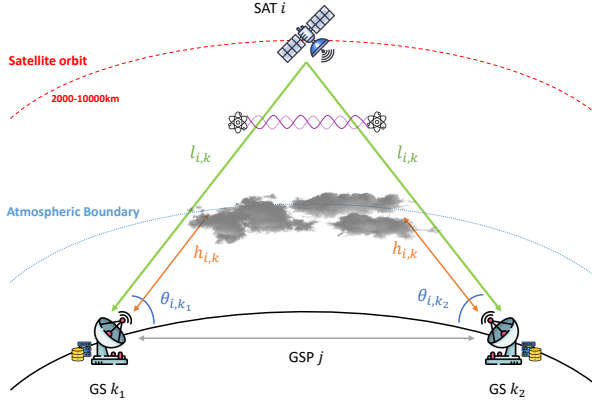


Fig. 2. An example of the space-to-ground transmittance between one SAT and two GSs for the entanglement distribution to this GSP.

Here, $\eta_{i,k}^{fs}$ depends on the orbital parameters, such as altitude and zenith angle, while $\eta_{i,k}^{atm}$ depends on the atmospheric conditions, e.g., turbulence and weather conditions. As shown in Fig. 2, let $l_{i,k}$ be the distance length between SAT i and GS k , and $h_{i,k}$ be the distance height between GS k and atmospheric boundary when connected to SAT i . In our analysis, we take into account the circular apertures of diameter d_i^T and d_k^R for the transmitter and receiver telescopes at SAT i and GS k , respectively. These telescopes operate at a specific wavelength λ . Hence, the free-space transmittance and the atmospheric transmittance can be calculated by

$$\eta_{i,k}^{fs} = \frac{(\pi(d_i^T/2)^2)(\pi(d_k^R/2)^2)}{(\lambda l_{i,k})^2}, \quad \eta_{i,k}^{atm} = e^{-\hat{\alpha} h_{i,k}}, \quad (2)$$

where $\hat{\alpha}$ is the atmospheric extinction coefficient.

Next, we examine the transmission of photons in the presence of background photons. Let $\bar{n}_k$ be the number of background photons received by GS k and F_0 be the initial fidelity of the entangled photon pair. We consider an initially imperfect Bell state with the assumption that $\eta_{i,k_1}^{sg} = \eta_{i,k_2}^{sg} = \eta_{i,j}^{sg}$ and $\bar{n}_{k_1} = \bar{n}_{k_2} = \bar{n}_j$. Therefore, the fidelity of entangled photon pair after transmission can be approximated by

$$F_{i,j} = \begin{cases} \frac{1}{4} \left(1 + \frac{4F_0-1}{(1+\frac{\bar{n}_j}{\eta_{i,j}^{sg}})^2} \right), & \text{if } \theta_{i,k_1}, \theta_{i,k_2} \geq \theta_e, (k_1, k_2) \in j, \\ 0, & \text{otherwise.} \end{cases} \quad (3)$$

Terrestrial Quantum Swapping. An intermediate GS, equipped with quantum memories entangled with memories on two remote GSs, is used to perform entanglement swapping. The successful execution of this process creates an extended entanglement “link” between the remote GSs. The probability of a successful entanglement swapping is represented by ps , while the probability of failure is denoted by $1 - ps$. If the process fails, the two Bell pairs involved are discarded. We consider a perfect quantum swapping and two depolarized Bell states (Werner states) will be swapped to a new Werner state.

Then the fidelities of the new states after swapping are defined below [10]–[12]

$$F_{sw} = F_1 F_2 + \frac{(1 - F_1)(1 - F_2)}{3}, \quad (4)$$

where F_1 and F_2 are the fidelities of two entanglement links, respectively. Note that by considering terrestrial quantum swapping, our entanglement distribution can serve more GSPs (even serving some GSPs which cannot be served by any SAT directly due to either visibility or fidelity issues).

III. PROBLEM FORMULATION

The goal of this paper is to solve a joint satellite assignment, resource allocation, and path selection problem for entanglement distribution in STIN. In this problem, each satellite acts as an entangled photon source (EPS) and distributes entangled photon pairs to the GSP. The objective is to maximize the overall utility of all traffic GSPs by optimally assigning the satellites and allocating photon resources to the GSP when a satellite is visible to both GSs within the same GSP.

In cases where no satellite is visible to the GSP, we explore ground-based quantum swapping using the entanglement links generated by satellite-based entanglement distribution. Moreover, we need to determine the optimal path selection for the remaining GSPs. Specifically, we are interested in two-hop quantum swapping and denote P as all two-hop paths from the graph indexed by p . Let $R_{p,j}^{gsp} = \{j_1, j_2\}$ be the GSP set in path p for GSP j , and $R_{p,j}^{sat} = \{i_1, i_2\}$ be the related satellite set in path p for GSP j . Moreover, let $P_k \subset P$ be the path set of GS k , $P_j \subset P$ be the path set of GSP j , and $P_i \subset P$ be the path set of SAT i . Denote $w_{i,j}$ as the weight or utility associated with the assignment of SAT i to GSP j and $w_{p,j}$ as the utility associated with the selection of path p for GSP j . We also assume that each satellite can generate at most sm_i entangled photon pairs.

Decision Variables. Let $x_{i,j}$ be a binary variable to indicate whether satellite i ($i \in M$) is assigned to GSP j ($j \in O$) and $y_{i,j}$ be an integer variable range from 0 to sm_{max} to indicate the entangled photon pair allocation between satellite i and GSP j . Denote $z_{p,j}$ by a binary variable to indicate whether path p ($p \in P$) is selected for GSP j for quantum swapping. We also define $\hat{y}_{p,j}$ as an integer variable ranging from 0 to sm_{max} to indicate the entangled photon pair allocation for GSP j using the path p . For any GSP $j \in O'$, if it can be served by the satellite, then we need to determine which satellite is optimal and allocate appropriate resources to this GSP j . The utility for this case can be defined as $U_1 = \sum_{j \in O'} \sum_{i \in M} w_{i,j} x_{i,j} y_{i,j}$. If no satellite can serve the GSP j , then we consider the two-hop quantum swapping by leveraging other available GSPs. Therefore, we need to determine which path is optimal and also allocate appropriate resources to the two GSPs in this path. Then, the utility for this case can be defined as $U_2 = \sum_{j \in O'} \sum_{p \in P_j} w_{p,j} z_{p,j} \hat{y}_{p,j}$.

Thus, the joint satellite assignment, resource allocation, and path selection problem is formulated as follows

$$\max_{x,y,z,\hat{y}} \sum_{j \in O'} \left(\sum_{i \in M} w_{i,j} x_{i,j} y_{i,j} + \sum_{p \in P_j} w_{p,j} z_{p,j} \hat{y}_{p,j} \right) \quad (5)$$

$$\text{s.t.} \sum_{i \in M} x_{i,j} + \sum_{p \in P_j} z_{p,j} \leq 1, \quad \forall j \in O', \quad (5a)$$

$$\sum_{i \in M} \sum_{j \in O_k} x_{i,j} \leq rr_k, \quad \forall k \in N, \quad (5b)$$

$$\sum_{j \in O} x_{i,j} \leq tr_i, \quad \forall i \in M, \quad (5c)$$

$$\sum_{i \in M} \sum_{j \in O_k} x_{i,j} y_{i,j} + \sum_{j \in O'} \sum_{\hat{j} \in O_k} \sum_{p \in P_j} z_{p,j} \hat{y}_{p,j} \leq gm_k, \quad \forall k, \quad (5d)$$

$$\sum_{j \in O'} (x_{i,j} y_{i,j} + \sum_{p \in P_i} z_{p,j} \hat{y}_{p,j}) \leq sm_i, \quad \forall i \in M, \quad (5e)$$

$$\sum_{p \in P_j} (z_{p,j} \sum_{j' \in R_{p,j}^{gsp}} \sum_{i' \in R_{p,j}^{sat}} x_{i',j'}) = (1 - \sum_{i \in M} x_{i,j}) \cdot 2, \quad \forall j, \quad (5f)$$

$$\sum_{i \in M} x_{i,j} (F_{i,j} - F^{th}) + \sum_{p \in P_j} z_{p,j} (F_{p,j} - F^{th}) \geq 0, \quad \forall j, \quad (5g)$$

$$x_{i,j}, z_{p,j} \in \{0, 1\}, y_{i,j}, \hat{y}_{p,j} \in \{0, \dots, sm_{max}\}. \quad (5h)$$

Here, constraint (5a) ensures that each GSP $j \in O'$ only connects to one satellite or only selects one swapping path if the satellite assignment is not available. Constraint (5b) means that a GS k can be part of multiple GSPs and thus is not allowed to be allocated to more than rr_k satellites. Constraint (5c) ensures that SAT i does not get allocated to more than tr_i GSPs. Constraint (5d) makes sure that the total entangled photon received from different satellites cannot exceed the maximal quantum memory of GS k . Constraint (5e) guarantees that the total entangled photon pair allocation fraction of SAT i cannot exceed sm_i . Constraint (5f) emphasizes that if a swapping path is selected, then two entanglement links along this path must be established based on the satellite assignment. Constraint (5g) confirms that the fidelity of entanglement links after transmitting or swapping is larger than the fidelity threshold. It is difficult to obtain the optimal solution to the optimization problem since this is a quadratic constrained quadratic discrete optimization problem, which is NP-hard and challenging to solve with classical computing when the problem scale increases. Note that our problem is much harder and more general than the one in [5].

IV. QUANTUM-ASSISTED ALGORITHMS

To tackle the complex optimization in our formulated entanglement distribution problem, we now leverage recent advanced QC [7], [13]–[19] to design two quantum-assisted algorithms. We first show how to convert the problem into a

quadratic unconstrained binary optimization (QUBO) format so that it can be solved by quantum annealer directly. Then, we introduce the Dantzig-Wolfe decomposition technique to break the original problem into smaller-scale problems, so that a hybrid quantum-classical solution (HQCDW) can be used to solve the optimization more efficiently.

A. Quadratic Unconstrained Binary Formulation

Inspired by the superiority of quantum annealing (QA) in solving large-scale complex optimization problems, we leverage QA to obtain the solution for our formulated joint optimization problem. To effectively solve optimization problems using the quantum annealer, these problems need to be formulated as an Ising model or a QUBO model. In a QUBO problem, there is typically a set of binary variables represented by vector $\mathbf{x}$ and an upper-diagonal matrix denoted as $\mathbf{Q}$, which is an $N' \times N'$ matrix with upper-triangular properties. The objective of QUBO is to minimize the following function:

$$\min_{\mathbf{x} \in \{0,1\}^{N'}} \mathbf{x}^\top \mathbf{Q} \mathbf{x}. \quad (6)$$

Note that problem (5) is a quadratic problem and all integer variables $y, \hat{y}$ can be expressed as a vector of binary variables. Hence, problem (5) can further represent the quadratic binary optimization problem. Next, we need to deal with all constraints to convert the problem into the pure QUBO form. To do so, we introduce a penalty for each constraint. The idea behind this is to find an optimal penalty coefficient of the constraints. Here, we leverage the binary search method to iteratively determine the optimal penalty for each constraint. For example, constraint (5a) is converted as follows

$$(5a) \Rightarrow \vartheta_1 : P^1 \left(\sum_{i \in M} x_{i,j} + \sum_{p \in P_j} z_{p,j} - 1 + \sum_{l=0}^{\bar{l}^1} 2^l s_l^1 \right)^2, \\ \text{where } \bar{l}^1 = \lceil \log_2 [1 - \min_{x,z} \left(\sum_{i \in M} x_{i,j} + \sum_{p \in P_j} z_{p,j} \right)] \rceil.$$

Here, P^* is the predefined penalty coefficient when the corresponding constraint is violated. s_l^* is a binary slack variable and $\bar{l}^*$ is the upper bound of number of slack variables. Similarly, we add a penalty ϑ_k for each of the seven constraints in problem (5).

Then, the original problem in (5) can be rewritten in the QUBO form as follows

$$\max_{x,y,z,\hat{y}} \sum_{j \in O'} \left(\sum_{i \in M} w_{i,j} x_{i,j} y_{i,j} + \sum_{p \in P_j} w_{p,j} z_{p,j} \hat{y}_{p,j} \right) + \vartheta_1 + \vartheta_2 + \vartheta_3 + \vartheta_4 + \vartheta_5 + \vartheta_6 + \vartheta_7. \quad (7)$$

Now, we can send the reformulated problem in (7) to the quantum annealer to calculate the optimal solution. It is worth noting that the quantum annealer currently has restricted qubits and can not perform the execution when the model scale further increases. Therefore, we apply Dantzig-Wolfe decomposition to decompose the original problem into the master problem and several subproblems to reduce the model

scale. In this case, the quantum annealer can solve the smaller-scale subproblems parallelly and separately.

B. Dantzig-Wolfe Decomposition

The DW decomposition algorithm is a method used to solve linear programming problems that have a specific structure (i.e., a block-angular or block-diagonal arrangement in the constraint matrix). It employs a delayed column generation (CG) technique to make solving large-scale linear programs more manageable. When applied to integer linear programming (ILP) problems, the DW algorithm often has most columns (variables) outside of the basis. The basis refers to a set of linearly independent columns from the constraint matrix that makes up the current active solution. In this approach, a master problem includes the currently active columns (basis), and a subproblem or subproblems are utilized to generate additional columns to enter the basis, thereby improving the objective function. The difference between the CG technique and DW decomposition is that in the CG process, a column is added to the master problem while an extreme point or extreme ray is added to the master problem in the DW process.

Recall that the problem in (5) is non-linear, but the DW decomposition is available to solve ILP problems. Therefore, we need to linearize the original problem first. Due to the existence of non-linear terms $x_{i,j}y_{i,j}$ and $z_{p,j}\hat{y}_{p,j}$ in the objective function and constraints, we introduce additional variables to linearize them. After full linearization, the original problem in (5) is defined as follows

$$\max_{x,y,z,\hat{y},\phi,\varphi,\psi} \sum_{j \in O'} \left(\sum_{i \in M} w_{i,j} \phi_{i,j} + \sum_{p \in P_j} w_{p,j} \varphi_{p,j} \right) \quad (8)$$

s.t. (5a) – (5h),

$$\phi_{i,j} \leq y_{i,j}, \quad \forall i \in M, j \in O', \quad (8a)$$

$$\phi_{i,j} \leq x_{i,j} sm_{max}, \quad \forall i \in M, j \in O', \quad (8b)$$

$$x_{i,j} + y_{i,j} - sm_{max} \leq \phi_{i,j}, \quad \forall i \in M, j \in O', \quad (8c)$$

$$\varphi_{p,j} \leq \hat{y}_{p,j}, \quad \forall j \in O', p \in P_j, \quad (8d)$$

$$\varphi_{p,j} \leq z_{p,j} sm_{max}, \quad \forall j \in O', p \in P_j, \quad (8e)$$

$$z_{p,j} + \hat{y}_{p,j} - sm_{max} \leq \varphi_{p,j}, \quad \forall j \in O', p \in P_j, \quad (8f)$$

$$\psi_{p,j,i',j'} \leq z_{p,j}, \quad \forall j \in O', p \in P_j, i' \in R_{p,j}^{sat}, j' \in R_{p,j}^{gsp}, \quad (8g)$$

$$\psi_{p,j,i',j'} \leq x_{i',j'}, \quad \forall j, p, i', j', \quad (8h)$$

$$z_{p,j} + x_{i',j'} - 1 \leq \psi_{p,j,i',j'}, \quad \forall j, p, i', j', \quad (8i)$$

$$x_{i,j}, z_{p,j}, \psi_{p,j,i',j'} \in \{0, 1\}, \quad (8j)$$

$$y_{i,j}, \hat{y}_{p,j}, \phi_{i,j}, \varphi_{p,j} \in \{0, \dots, sm_{max}\}. \quad (8k)$$

Here, $\phi_{i,j}$ and $\varphi_{p,j}$ are auxiliary integer variables, $\psi_{p,j,i',j'}$ is the auxiliary binary variable. Constraints (8a)-(8c) are the linearization of non-linear term $x_{i,j}y_{i,j}$. Constraints (8d)-(8f) are the linearization of non-linear term $z_{p,j}\hat{y}_{p,j}$, while Constraints (8g)-(8i) are the linearization of non-linear term $z_{p,j}x_{i',j'}$. Since problem (8) and constraints are all linear, the problem can be further expressed as a general form below

$$\max_{\mathbf{X}, \mathbf{Y}} \mathbf{h}^\top \mathbf{Y} \quad (9)$$

$$\text{s.t. } \mathbf{B}_1 \mathbf{X} \leq \mathbf{b}_1, \quad (9a)$$

$$\mathbf{B}_2 \mathbf{X} = \mathbf{b}_2, \quad (9b)$$

$$\mathbf{C} \mathbf{Y} \leq \mathbf{b}_3, \quad (9c)$$

$$\mathbf{A} \mathbf{X} + \mathbf{G} \mathbf{Y} \leq \mathbf{b}_4, \quad (9d)$$

$$\mathbf{X} = [x, z, \psi]^\top, \quad \mathbf{X} \in \mathcal{X}, \quad (9e)$$

$$\mathbf{Y} = [y, \hat{y}, \phi, \varphi]^\top, \quad \mathbf{Y} \in \mathcal{Y}, \quad (9f)$$

where $\mathbf{h}$ is the coefficient for integer variables in the objective function. $\mathbf{A}, \mathbf{B}_1, \mathbf{B}_2, \mathbf{C}$, and $\mathbf{G}$ are coefficients in the constraints while $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3$ and $\mathbf{b}_4$ are constant vectors. Let $\dim_{\mathbf{X}} = |M| \times |O| + |P| \times |O| + |P| \times |O| \times |M| \times |O|$ and $\dim_{\mathbf{Y}} = (|M| \times |O| + |P| \times |O|) \times 2$.

Next, we reformulate the original problem by increasing variable constraints to reduce the number of inequality or equality constraints. Define $\mathcal{X}$ and $\mathcal{Y}$ as

$$\mathcal{X} = \{\mathbf{X} \in \{0, 1\}^{\dim_{\mathbf{X}}} : \mathbf{B}_1 \mathbf{X} \leq \mathbf{b}_1 \text{ and } \mathbf{B}_2 \mathbf{X} = \mathbf{b}_2\}, \quad (10)$$

$$\mathcal{Y} = \{\mathbf{Y} \in \{0, \dots, sm_{max}\}^{\dim_{\mathbf{Y}}} : \mathbf{C} \mathbf{Y} \leq \mathbf{b}_3\}. \quad (11)$$

Then optimization problem (9) is reformulated as

$$\max_{\mathbf{X}, \mathbf{Y}} \mathbf{h}^\top \mathbf{Y} \quad (12)$$

$$\text{s.t. } \mathbf{A} \mathbf{X} + \mathbf{G} \mathbf{Y} \leq \mathbf{b}_4, \quad (12a)$$

$$\mathbf{X} \in \mathcal{X}, \quad (12b)$$

$$\mathbf{Y} \in \mathcal{Y}. \quad (12c)$$

Let $\mathcal{U}$ be the feasible region of (12). A feasible region $\mathcal{U}$ is the set of all possible points of (12) that satisfy the problem's constraints:

$$\mathcal{U} = \{\mathbf{X} \in \mathcal{X}, \mathbf{Y} \in \mathcal{Y} : \mathbf{A} \mathbf{X} + \mathbf{G} \mathbf{Y} \leq \mathbf{b}_4\}. \quad (13)$$

Note that every polyhedron $\mathcal{U}$ can be written as the sum of finitely many extreme points and extreme rays. Thus, we denote its subsets of extreme points with $\mathcal{P}_{\mathcal{X}} = \{\mathbf{X}^{(i)}, \forall i \in \mathcal{I}'\}$ and $\mathcal{P}_{\mathcal{Y}} = \{\mathbf{Y}^{(j)}, \forall j \in \mathcal{J}'\}$. Then we can express the problem (12) as the linear combination of its extreme points:

$$\max_{\substack{\mu_i, \forall i \in \mathcal{I}' \\ \nu_j, \forall j \in \mathcal{J}'}} \sum_{j \in \mathcal{J}'} (\mathbf{h}^\top \mathbf{Y}^{(j)}) \nu_j \quad (14)$$

$$\text{s.t. } \sum_{i \in \mathcal{I}'} (\mathbf{A} \mathbf{X}^{(i)}) \mu_i + \sum_{j \in \mathcal{J}'} (\mathbf{G} \mathbf{Y}^{(j)}) \nu_j \leq \mathbf{b}_4, \quad (14a)$$

$$\sum_{i \in \mathcal{I}'} \mu_i = 1, \quad (14b)$$

$$\sum_{j \in \mathcal{J}'} \nu_j = 1, \quad (14c)$$

$$\mu_i \in [0, 1], \quad \forall i \in \mathcal{I}', \quad (14d)$$

$$\nu_j \in [0, 1], \quad \forall j \in \mathcal{J}'. \quad (14e)$$

where $\mu_i \in \mathbb{R}$ and $\nu_j \in \mathbb{R}$ represent the weights of each extreme point for binary and integer variables.

Furthermore, we introduce the Lagrangian relaxation, and then the Lagrangian dual problem can be expressed as

$$\min_{\alpha, \xi, \zeta} \quad \alpha \mathbf{b}_4 + \xi + \zeta \quad (15)$$

$$\text{s.t.} \quad -\alpha \mathbf{A} \mathbf{X}^{(i)} - \xi \geq 0, \quad \forall i \in \mathcal{I}', \quad (15a)$$

$$(\mathbf{h}^\top - \alpha \mathbf{G}) \mathbf{Y}^{(j)} - \zeta \geq 0, \quad \forall j \in \mathcal{J}', \quad (15b)$$

where $\alpha \in \mathbb{R}_+$ is the row vector (dual variable) of constraint (14a) and $\xi, \zeta \in \mathbb{R}$ are Lagrangian multipliers for constraints (14b) and (14c), respectively. The above optimization problem is called *dual restricted master problem*. At every step t , we generate extreme points $\mathbf{X}^{(i)}$ and $\mathbf{Y}^{(j)}$. These extreme points are incorporated into the master problem, necessitating the addition of new μ_i and ν_j columns. Constraints (15a) and (15b) are called *reduced cost*. Then the two *pricing problems* (subproblems) are given as

$$\max_{\mathbf{X} \in \mathcal{X}} \quad -\alpha^{(t)} \mathbf{A} \mathbf{X} \quad (16)$$

$$\max_{\mathbf{Y} \in \mathcal{Y}} \quad (\mathbf{h}^\top - \alpha^{(t)} \mathbf{G}) \mathbf{Y}, \quad (17)$$

where $\alpha^{(t)}$ is the dual variables of constraint (14a). If the solution of (16) is larger than $\xi^{(t)}$, then we set $\mathcal{I}' \leftarrow \mathcal{X}^{(t)}$. Similarly, if the solution of (17) is larger than $\zeta^{(t)}$, then we set $\mathcal{J}' \leftarrow \mathcal{Y}^{(t)}$.

C. Hybrid Quantum-Classical Solution: HQCDW

We now introduce the hybrid quantum-classical algorithm for solving our original problem (5) using DW decomposition. Recall that the DW process involves solving a master problem and several subproblems iteratively. As discussed in Section IV-B, the restricted master problem is defined as

$$\begin{aligned} \text{Master:} \quad & \max_{\substack{\mu_i, \forall i \in \mathcal{I}' \\ \nu_j, \forall j \in \mathcal{J}'}} \sum_{j \in \mathcal{J}'} (\mathbf{h}^\top \mathbf{Y}^{(j)}) \nu_j \\ \text{s.t.} \quad & (14a) - (14c). \end{aligned} \quad (18)$$

This restricted master problem is easier to solve and can provide initial solutions for the original master problem. Hence, we can solve it by using a classical solver (e.g., Gurobi, Scipy) running in the classical CPU computer. Subsequently, two subproblems (pricing problems) are given by

$$\text{Subproblem 1:} \quad \max_{\mathbf{X}} \quad -\alpha^{(t)} \mathbf{A} \mathbf{X} \quad (19)$$

$$\text{s.t.} \quad \mathbf{B}_1 \mathbf{X} \leq \mathbf{b}_1, \quad (19a)$$

$$\mathbf{B}_2 \mathbf{X} = \mathbf{b}_2, \quad (19b)$$

$$\mathbf{X} \in \{0, 1\}^{dim_{\mathbf{X}}}. \quad (19c)$$

$$\text{Subproblem 2:} \quad \max_{\mathbf{Y}} \quad (\mathbf{h}^\top - \alpha^{(t)} \mathbf{G}) \mathbf{Y} \quad (20)$$

$$\text{s.t.} \quad \mathbf{C} \mathbf{Y} \leq \mathbf{b}_3, \quad (20a)$$

$$\mathbf{Y} \in \{0, \dots, sm_{max}\}^{dim_{\mathbf{Y}}}. \quad (20b)$$

Here, Subproblems 1 and 2 are pure linear binary or integer problems, which can be conveniently mapped into the QUBO form as discussed in Section IV-A, and we can solve them by

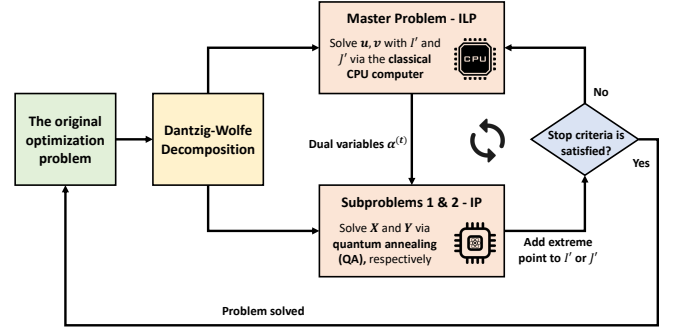


Fig. 3. Proposed HQCDW framework for joint satellite-based quantum entanglement distribution and terrestrial quantum swapping.

Algorithm 1 Hybrid Quantum-Classical Dantzig-Wolfe Decomposition (HQCDW) Algorithm

Input: Related information $M, N, O, sm_i, tr_i, sm_{max}, gm_k, rr_k, \theta_e, w_{i,j}, w_{p,j}$, loss and noise parameters.

Output: $x_{i,j}, z_{p,j}, y_{i,j}, \hat{y}_{p,j}$

- 1: Get $\mathbf{A}, \mathbf{B}_1, \mathbf{B}_2, \mathbf{C}, \mathbf{G}, \mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3$ and $\mathbf{b}_4$
- 2: $\mathbf{X}^{(0)}, \mathbf{Y}^{(0)} \leftarrow$ Initialize the extreme points
- 3: $\mathcal{I}', \mathcal{J}' \leftarrow \mathbf{X}^{(0)}, \mathbf{Y}^{(0)}$
- 4: $\alpha^{(0)}, \xi^{(0)}, \zeta^{(0)} \leftarrow$ Initialize the dual values
- 5: $max_itr \leftarrow 100, t \leftarrow 0, s1, s2 \leftarrow 0, 0$
- 6: **while** stop_criteria_is_not_satisfied() **do**
- 7: $s1, \mathbf{X}^{(t)} \leftarrow$ solve (19) with $\alpha^{(t)}$ ▷ quantum
- 8: **if** $s1 > \xi^{(t)}$ **then**
- 9: $\mathcal{I}' \leftarrow \mathbf{X}^{(t)}$ ▷ add the extreme point to set $\mathcal{I}'$
- 10: **end if**
- 11: $s2, \mathbf{Y}^{(t)} \leftarrow$ solve (20) with $\alpha^{(t)}$ ▷ quantum
- 12: **if** $s2 > \zeta^{(t)}$ **then**
- 13: $\mathcal{J}' \leftarrow \mathbf{Y}^{(t)}$ ▷ add the extreme point to set $\mathcal{J}'$
- 14: **end if**
- 15: $\mu, \nu \leftarrow$ solve (18) with $\mathbf{X}^{(t)}, \mathbf{Y}^{(t)}$ ▷ classical
- 16: $t \leftarrow t + 1$
- 17: $\alpha^{(t)}, \xi^{(t)}, \zeta^{(t)} \leftarrow$ get dual solution from (18)
- 18: **end while**
- 19: $x_{i,j}, z_{p,j}, y_{i,j}, \hat{y}_{p,j} \leftarrow$ extract from $\mathcal{I}', \mathcal{J}'$
- 20: **return** $x_{i,j}, z_{p,j}, y_{i,j}, \hat{y}_{p,j}$

using QA parallelly. Since the model scale of problems (19) and (20) are smaller than the original problem in (5), and they are all linear functions compared with the quadratic term in the original problem in (5), we can solve a larger-scale model compared with using pure QUBO solution. Consequently, the proposed HQCDW algorithm iteratively solves the restricted master problem and the pricing problems until convergence is reached. The detailed procedure and algorithm are shown in Fig. 3 and Algorithm 1, respectively.

V. PERFORMANCE EVALUATION

In this section, we evaluate the proposed algorithms using extensive simulations on a hybrid D-Wave quantum processing unit (QPU) with around 5,000 qubits accessed through the

Leap quantum cloud service [20]. For classical computations, we used an LP solver on a CPU with an Intel(R) Core(TM) i7-6700HQ running at 2.60 GHz and 16GB RAM.

A. Experiment Setup

Network architecture. We consider a polar satellite constellation as discussed in [4], [5] where there are 10 spaced rings of satellites in polar orbits and 10 satellites are deployed in each ring within an altitude of 2,000 km to 10,000 km. A static scenario is considered where only a few satellites are visible to a specific ground station at any given time. This allows for the application of alternative satellite constellations. Each satellite is equipped with a quantum memory and a random number of transmitters between 6 and 10. Several major cities are designated as GSs, with their locations based on real GPS coordinates. The total number of GSs is capped at 36, with a subset chosen as traffic GSs. For each GS, the quantum memory and the number of receivers are randomly selected from ranges of 10 to 20 and 2 to 6, respectively. Experiment settings are denoted by the number of satellites and GSs used; for instance, 14-5 represents a scenario with 14 satellites and 5 ground stations. For the loss and noise parameters in the transmission of photons from satellites to GS, we set them according to previous works [4], [5].

Comparison methods. To verify the effectiveness and advantage of our algorithms, we compare our methods (QUBO and HQCDW) with the following schemes:

- *Steepest Descent (SD)* [20]: SD is a solver for binary quadratic models by D-Wave Systems with the best move determined through a local minimization.
- *Random Steepest Descent (RASD)* [20]: RASD combines random sampling and steepest descent, with the algorithm using random sampling to generate initial states for SD.
- *Classical Optimizer (COPT)*: COPT solves the original problem (5) by using a classical optimizer (e.g., Gurobi, Scipy) in a classical CPU computer.
- *Classical DW Decomposition (CDWD)*: CDWD decouples the original problem using DW decomposition and solves all problems with a classical optimizer.

B. Experiment Results and Analysis

1) *Performance of Number of Samples*: In this study, we investigate the impact of sample size on a quantum annealer's performance in solving a problem. The number of samples corresponds to the output solutions generated by the quantum annealer. Results from Fig. 4(a) show that the COPT algorithm consistently outperforms various QUBO versions with sample sizes of 1, 20, 50, 100, and 200 across different network settings. As network complexity grows, QUBO with higher sample sizes exhibits better QUBO objective values, with QUBO-200 achieving the highest values, followed by QUBO-100. QUBO-1 consistently yields the lowest values, indicating that a single sample is insufficient for effective optimization. This pattern holds across all network settings, with QUBO approaching COPT's performance with sufficient samples, typically 200 or more.

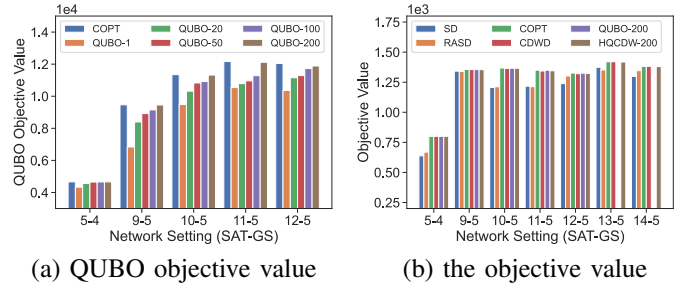


Fig. 4. (a) The QUBO objective value (i.e., problem (7)) under different network settings for COPT and different numbers of QUBO sampling. (b) The objective value (i.e., problem (5)) under different network settings with two DW subproblems (linear).

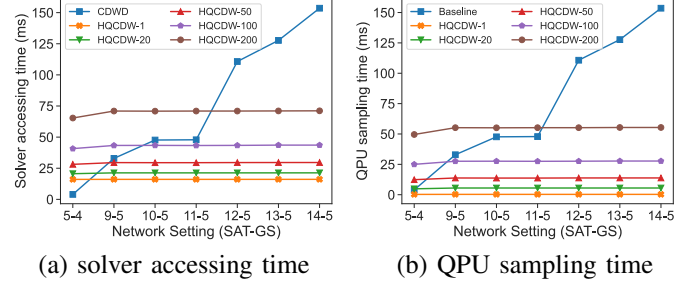


Fig. 5. (a) The solver accessing time under various network settings for CDWD and different numbers of HQCDW sampling using partial linearization. (b) The QPU sampling time under various network settings for different numbers of HQCDW sampling using partial linearization.

2) *Performance in Objective Value*: We now compare the methods based on the objective value. The objective value is initially presented under various network settings using full linearization. To ensure a fair comparison of quantum computation, we set the number of samples to 200 for QUBO and HQCDW, as shown in Fig. 4(b). SD and RASD perform poorly across all network settings compared to the other methods, possibly due to SD getting stuck in local optima. COPT, CDWD, QUBO-200, and HQCDW-200 generally achieve similar results across most network settings, except for settings 13-5 and 14-5. However, as the problem size increases beyond setting 12-5, QUBO-200 struggles due to the limited qubits in QA, while HQCDW-200 can handle problems up to setting 14-5. This highlights the effectiveness and robustness of using DW decomposition to tackle larger network scales.

3) *Performance in Solver Accessing Time*: We study the performance of our proposed methods regarding solver accessing time, which specifically refers to the time taken by the QPU solver and local solver without considering other overheads like variable setting and parameter transmission time. We examine the solver accessing time across different network settings for CDWD and various numbers of HQCDW sampling, as shown in Fig. 5(a). CDWD experiences a significant increase in solver accessing time as the network settings progress, surpassing HQCDW-200 beyond a certain point. Furthermore, utilizing the DW decomposition enables problem-solving up to 14-5 compared to 12-5 of QUBO without it. A similar pattern is observed for QPU sampling time, as illustrated in Fig. 5(b). This analysis highlights

TABLE I

TOTAL QUBITS USAGE OF DIFFERENT QUANTUM-BASED APPROACHES UNDER VARIOUS NETWORK SETTINGS.

Setting	Non-DW				DW			
	Quadratic		Linear		Linear			
	LQ	RQ	LQ	RQ	Sub X		Sub Y	
5-4	336	399	670	670	LQ	RQ	LQ	RQ
9-5	1750	2083	2856	2856	324	324	2532	2532
10-5	1960	2278	3149	3149	359	359	2790	2790
11-5	2170	2490	3484	3484	394	394	3090	3090
12-5	3360	4549	5176	5176	568	568	4608	4608
13-5	-	-	-	-	638	638	5058	5058
14-5	-	-	-	-	708	708	5574	5574

LQ: Logic Qubit, RQ: Real Qubit, "-": Not available.

the advantage of using the DW decomposition to overcome constraints imposed by the maximum number of qubits in the quantum machine.

4) *Qubit Usage Comparison*: We compare the total qubit usage for two DW subproblems across different network settings in Table I. The table presents a detailed comparison of qubit usage for settings such as 5-4, 9-5, and 10-5, dividing columns into Non-DW and DW categories. Non-DW refers to the original problem without decomposition, while DW signifies decomposition into two subproblems. Subcategories include Quadratic and Linear for Non-DW, and Linear for DW, delineating Logic Qubits (LQ) and Real Qubits (RQ) used. Linearization impacts qubit usage, with larger network settings requiring more qubits for representation. Linear uses the most qubits while Quadratic uses the fewest due to handling quadratic terms effectively. Linear has equal LQ and RQ counts, unlike Quadratic. Non-DW is limited to 12-5 due to qubit constraints. In DW subproblems, Sub Y's qubit usage is crucial, acting as a bottleneck. DW decomposition allows solving up to scale 14-5 but faces challenges beyond due to the significant number of variables in the master problem. We leave further study methods to improve DW based method by addressing this bottleneck at the master problem. In summary, this result highlights the benefits of DW decomposition for larger-scale problems.

VI. CONCLUSION

In this work, we delved into the intricacies of joint satellite assignment, resource allocation, and path selection for entanglement distribution within a space-terrestrial integrated network. By harnessing the capabilities of satellite-based entanglement distribution and terrestrial quantum swapping, we formulated and addressed a complex joint optimization problem. The novel hybrid quantum-classical Dantzig-Wolfe decomposition technique showcased the potential of leveraging both quantum and classical computing to effectively solve large-scale network optimization challenges and balance qubit usage. The series of experiments provided compelling evidence of the efficiency and robustness of the proposed methods, with a stable solver accessing time and consistent results compared to classical optimizers, thereby laying a strong foundation for future advancements in this field.

REFERENCES

- [1] J. Yin, Y. Cao, Y.-H. Li, S.-K. Liao, L. Zhang, J.-G. Ren, W.-Q. Cai, W.-Y. Liu, B. Li, *et al.*, "Satellite-based entanglement distribution over 1200 kilometers," *Science*, vol. 356, no. 6343, pp. 1140-1144, Jun. 2017.
- [2] Z. Wang, R. Malaney, and J. Green, "Satellite-based entanglement distribution using orbital angular momentum of light," in *IEEE International Conference on Communications Workshops*, Virtual, Jun. 2020.
- [3] J. Yin, Y.-H. Li, S.-K. Liao, M. Yang, Y. Cao, L. Zhang, J.-G. Ren, W.-Q. Cai, W.-Y. Liu, S.-L. Li, *et al.*, "Entanglement-based secure quantum cryptography over 1,120 kilometres," *Nature*, vol. 582, no. 7813, pp. 501-505, Jun. 2020.
- [4] S. Khatri, A. J. Brady, R. A. Desporte, M. P. Bart, and J. P. Dowling, "Spooky action at a global distance: analysis of space-based entanglement distribution for the quantum Internet," *npj Quantum Information*, vol. 7, no. 4, pp. 1-15, Jan. 2021.
- [5] N. K. Panigrahy, P. Dhara, D. Towsley, S. Guha, and L. Tassiulas, "Optimal Entanglement Distribution using Satellite Based Quantum Networks," in *IEEE Conference on Computer Communications (INFOCOM) Workshops*, Virtual, May 2022.
- [6] X. Wei, J. Liu, L. Fan, Y. Guo, Z. Han, and Y. Wang, "Optimal entanglement distribution problem in satellite-based quantum networks," *IEEE Network*, 2024, to appear.
- [7] X. Wei, J. Liu, L. Fan, Y. Guo, Z. Han, and Y. Wang, "Hybrid quantum-classical computing via Dantzig-Wolfe decomposition for integer linear programming," in *33rd IEEE International Conference on Computer Communications and Networks*, Big Island, HI, Jul. 2024.
- [8] T. Gonzalez-Raya, S. Pirandola, and M. Sanz, "Satellite-based entanglement distribution and quantum teleportation with continuous variables," *arXiv preprint arXiv:2303.17224*, Mar. 2023.
- [9] P. Dhara, S. J. Johnson, C. N. Gagatsos, P. G. Kwiat, and S. Guha, "Heralded Multiplexed High-Efficiency Cascaded Source of Dual-Rail Entangled Photon Pairs Using Spontaneous Parametric Down-Conversion," *Physical Review Applied*, vol. 17, no. 034071, pp. 1-24, Mar. 2022.
- [10] H. Gu, Z. Li, R. Yu, X. Wang, F. Zhou, and J. Liu, "FENDI: High-Fidelity Entanglement Distribution in the Quantum Internet," *arXiv preprint arXiv:2301.08269*, Jan. 2023.
- [11] N. K. Panigrahy, T. Vasantham, D. Towsley, and L. Tassiulas, "On the Capacity Region of a Quantum Switch with Entanglement Purification," *arXiv preprint arXiv:2212.01463*, Dec. 2022.
- [12] A. Zang, X. Chen, A. Kolar, and J. Chung, "Entanglement Distribution in Quantum Repeater with Purification and Optimized Buffer Time," in *IEEE Conference on Computer Communications Workshops*, New York, NY, May 2023.
- [13] T. Tran, M. Do, E. Rieffel, J. Frank, Z. Wang, B. O'Gorman, D. Venturelli, and J. Beck, "A hybrid quantum-classical approach to solving scheduling problems," in *Proceedings of the International Symposium on Combinatorial Search*, New York, NY, Jul. 2016.
- [14] Z. Zhao, L. Fan, and Z. Han, "Hybrid quantum benders' decomposition for mixed-integer linear programming," in *IEEE Wireless Communications and Networking Conference*, Austin, TX, Apr. 2022.
- [15] A. Ajagekar, K. Al Hamoud, and F. You, "Hybrid classical-quantum optimization techniques for solving mixed-integer programming problems in production scheduling," *IEEE Transactions on Quantum Engineering*, vol. 3, no. 3102216, pp. 1-16, Jun. 2022.
- [16] L. Fan and Z. Han, "Hybrid quantum-classical computing for future network optimization," *IEEE Network*, vol. 36, no. 5, pp. 72-76, Nov. 2022.
- [17] Z. Zhao, L. Fan, Y. Guo, Y. Wang, Z. Han, and L. Hanzo, "QAOA-assisted benders' decomposition for mixed-integer linear programming," in *Proceedings of IEEE International Conference on Communications*, Denver, CO, Jun. 2024.
- [18] X. Wei, L. Fan, Y. Guo, Y. Gong, Z. Han, and Y. Wang, "Quantum assisted scheduling algorithm for federated learning in distributed networks," in *32nd IEEE International Conference on Computer Communications and Networks*, Honolulu, HI, Jul. 2023.
- [19] X. Wei, L. Fan, Y. Guo, Y. Gong, Z. Han, and Y. Wang, "Hybrid quantum-classical Benders' decomposition for federated learning scheduling in distributed networks," *IEEE Transactions on Network Science and Engineering (TNSE)*, 2024, to appear.
- [20] D-Wave Systems Inc., "Ocean software documentation: dwave-samplers, steepestdescentsolver, randomnessampler," Oct. 2023. [Online]. Available: https://docs.ocean.dwavesys.com/en/stable/docs_samplers/reference.html