

Received 16 November 2023; revised 20 February 2024; accepted 6 April 2024; date of publication 16 April 2024; date of current version 6 May 2024.

Digital Object Identifier 10.1109/TQE.2024.3386753

# Multiobjective Optimization and Network Routing With Near-Term Quantum Computers

SHAO-HEN CHIEW<sup>1</sup> , KILIAN POIRIER<sup>1</sup> , RAJESH MISHRA<sup>1,2</sup> ,  
ULRIKE BORNHEIMER<sup>1</sup> , EWAN MUNRO<sup>1</sup> , SI HAN FOON<sup>3</sup>,  
CHRISTOPHER WANRU CHEN<sup>3</sup>, WEI SHENG LIM<sup>3</sup> , AND CHEE WEI NGA<sup>3</sup>

<sup>1</sup>Entropica Labs, Singapore 068632

<sup>2</sup>Department of Physics, University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA

<sup>3</sup>Defence Science and Technology Agency, Singapore 109679

Corresponding author: Shao-Hen Chiew (e-mail: shaohen@entropicalabs.com).

This work was supported in part by the Ministry of Defense of Singapore and by the Defence Science and Technology Agency.

**ABSTRACT** Multiobjective optimization is a ubiquitous problem that arises naturally in many scientific and industrial areas. Network routing optimization with multiobjective performance demands falls into this problem class, and finding good quality solutions at large scales is generally challenging. In this work, we develop a scheme with which near-term quantum computers can be applied to solve multiobjective combinatorial optimization problems. We study the application of this scheme to the network routing problem in detail, by first mapping it to the multiobjective shortest-path problem. Focusing on an implementation based on the quantum approximate optimization algorithm (QAOA)—the go-to approach for tackling optimization problems on near-term quantum computers—we examine the Pareto plot that results from the scheme and qualitatively analyze its ability to produce Pareto-optimal solutions. We further provide theoretical and numerical scaling analyses of the resource requirements and performance of QAOA and identify key challenges associated with this approach. Finally, through Amazon Braket, we execute small-scale implementations of our scheme on the IonQ Harmony 11-qubit quantum computer.

**INDEX TERMS** Approximation algorithms, hardware, networks, optimization, quantum circuit, quantum computing.

## I. INTRODUCTION

Multiobjective optimization problems (MOOPs) arise naturally in many scientific and industrial areas, where the interplay between multiple conflicting objectives gives rise to a set of optimal solutions, rather than a unique one. They are especially prevalent in engineering contexts, where complex systems involving multiple objectives are encountered [1], [2], [3]. In particular, we focus on the case where the state space and, hence, the set of optimal solutions are discrete, where they are referred to as multiobjective combinatorial optimization problems (MCOPs). An example of such an MCOP is the multiobjective network routing problem, which asks for paths between specified source and destination nodes in a graph that is optimal with respect to multiple objectives. This optimization problem arises in contexts, such as the design of wireless ad-hoc/sensor networks [4], [5], [6], [7] and next-generation communication networks [8], [9],

[10], where large-scale networks with multiple requirements have to be simultaneously satisfied to meet performance demands.

However, the solution to such problems is generally difficult, as finding the global optima—also known as Pareto-optimal solutions—of general MCOPs is NP-hard [11], [12]. Direct methods based on sorting [13], [14], [15] require computational times that scale polynomially with the size of the state space, which becomes intractable for large-scale problems. On the other hand, physics- and biology-inspired metaheuristic classical algorithms have been particularly successful in this area in the past two decades [15], [16], [17] due to their scalable computational costs and are the go-to methods for industrial use-cases. Even so, due to the importance and complexity of MCOPs, there is a need to develop theoretical and algorithmic tools to solve large-scale problems in more memory and time-efficient ways.

Besides developments in classical computers, quantum computers are currently experiencing explosive growth, both in theory and realization. Noisy intermediate-scale quantum (NISQ) computers with qubit counts in the small hundreds are now available, and there is a growing body of work investigating their ability to solve challenging optimization problems in areas ranging from chemistry [18] to finance [19], with the hope that they can outperform existing classical algorithms in the near future. In particular, there has been a strong focus on variational quantum algorithms (VQAs) such as the quantum approximate optimization algorithm (QAOA) [20], which can be used to tackle combinatorial optimization problems with existing quantum computers. A natural question is whether similar methods can be applied to efficiently obtain high-quality solutions to MCOPs such as the network routing problem, to satisfy the performance demands of next-generation wireless networks.

In this context, a significant challenge is the fact that, for a given use-case, investigating the performance of VQAs such as QAOA typically requires an empirical approach. However, given the relatively small sizes of currently available quantum computers and the complexity of simulating quantum computations classically—an affair that also depends on the density and connectivity of the underlying problem [21]—the scope for experimentation is currently rather limited [22], [23], [24]. Nonetheless, in this work, we adopt a pragmatic approach, using the standard form of QAOA [20] as a solution method for MCOPs. This permits us to: 1) use small-scale problem instances to draw insights on how the QAOA solution relates to Pareto-optimal solutions of the MCOP and 2) analyze the scalability of the QAOA approach against known generic limitations of VQAs.

The key results of our work are summarized as follows.

- 1) We develop a general framework with which near-term quantum computers can be used to solve MCOPs, by producing multiple Pareto-optimal solutions in both a priori (with the preferences of the decision maker taken into account prior to optimization) and a posteriori (independent of the decision maker's preferences) manners. This is achieved by casting the multiple objectives and constraints of an MCOP in a quadratic unconstrained binary optimization (QUBO) form and scalarizing it to obtain a cost function, which can be variationally optimized with a VQA. For small problem instances, visualization of the output quantum state on Pareto plots provides insights, which we analyze and explain in a qualitative manner.
- 2) We provide a formulation of the network routing problem that is amenable to implementation with QAOA on near-term quantum computers. Using results from graph theory, we determine its resource requirements and provide a scaling analysis. In particular, we show that this encoding scheme possesses resource requirements that scale mildly with the connectivity of the

underlying graph, which is, in principle, compatible with resources available on current NISQ hardware.

- 3) Numerical simulations of small problem instances using standard QAOA (as described in [20]) show that this framework can produce high-quality solutions efficiently. Concretely, by increasing the circuit depth, we observe a correspondingly proportional increase in the probability of successfully obtaining Pareto-optimal solutions. However, as we explain in Section IV-A, in terms of problem size, the efficacy of the QUBO-based QAOA approach is limited by the large fraction of infeasible solutions in the underlying search space, resulting from the presence of optimization constraints.
- 4) We run a number of demonstrative test cases on the 11-qubit IonQ Harmony quantum computer, accessed through Amazon Braket. The results obtained are in clear agreement with those of numerical simulations.

The rest of the article is organized as follows. We begin by fully describing our approach to obtain optimal solutions of an MCOP and discuss interpretations of resulting Pareto plots in Section II. We then describe the network routing problem, a concrete example of a generally difficult MCOP, which we study throughout the article, along with relevant objectives that arise from reasonable assumptions in Section III. This is followed by theoretical and numerical analyses on the scaling of resources and performance of the procedure in Section IV, and experiments on quantum computers in Section V.

We refer the reader to the appendices for a survey of relevant work in the usage of near-term quantum algorithms in solving MCOPs and the shortest-path problem (see Appendix A), a review of MOOPs (see Appendix B), the mathematical formulation of the multiobjective shortest-path problem (see Appendix C), graph theoretic ideas used in this work (see Appendix D), the formulation of the network routing problem along with complete specifications of physical parameters used to model its objectives (see Appendix E), specifications of problem instances considered in our numerical calculations and experiments (see Appendices F and H), additional numerical results for the chosen parameter initialization scheme (see Appendix G), and additional discussions on the Pareto plot (see Appendix I).

Numerical simulations and the set-up of quantum computations in this work were performed through OpenQAOA [25], an open-source Python package tailored for QAOA and its variants.

## II. SOLVING MCOPS WITH QUANTUM COMPUTERS

In this section, we describe our approach to solve MCOPs using near-term quantum computers with VQAs, such as the QAOA, starting with brief introductions to the QAOA, MCOPs, and combinatorial optimization/QUBO problems. For more in-depth discussions on the previous topics, we refer the reader to the work in [2], [3], [20], and [26], respectively, along with the appendices.

A key component of our procedure is the QAOA. Belonging to the class of VQAs, it was introduced to allow NISQ computers to provide approximate solutions to combinatorial optimization problems such as graph partitioning, coloring, and constraint satisfaction problems [20]. A QUBO problem can be defined by an upper triangular matrix with real values. It is specified within QAOA through a corresponding cost Hamiltonian  $H_c$  expressed as a sum of  $l$  local terms  $H_c = \sum_{i=1}^l h_i$ , which can be obtained by converting the binary variables of the QUBO problem to Ising variables [26]. This allows the quantum state  $|\psi(\vec{\beta}, \vec{\gamma})\rangle$  to be prepared by a  $p$ -layer state-preparation ansatz  $U_{\text{QAOA}}(\vec{\beta}, \vec{\gamma})$

$$\begin{aligned} |\psi(\vec{\beta}, \vec{\gamma})\rangle &= U_{\text{QAOA}}(\vec{\beta}, \vec{\gamma})|+\dots+\rangle \\ &= \prod_{i=1}^p e^{-i\beta_i \sum_{j=1}^n X_j} e^{-i\gamma_i H_c} |+\dots+\rangle \end{aligned} \quad (1)$$

where the real vectors  $\vec{\beta} = (\beta_1, \dots, \beta_p)$  and  $\vec{\gamma} = (\gamma_1, \dots, \gamma_p)$  are variational parameters and  $|+\dots+\rangle$  is the initial quantum state. By sampling from a quantum computer, the expectation values  $\langle \psi(\vec{\beta}, \vec{\gamma}) | H_c | \psi(\vec{\beta}, \vec{\gamma}) \rangle$  can be computed and taken as the cost function for the optimization problem

$$(\vec{\beta}^*, \vec{\gamma}^*) = \underset{\vec{\beta}, \vec{\gamma}}{\operatorname{argmin}} \langle \psi(\vec{\beta}, \vec{\gamma}) | H_c | \psi(\vec{\beta}, \vec{\gamma}) \rangle. \quad (2)$$

In practice, the optimization problem is solved in a variational manner with the usage of a classical optimization algorithm, yielding a quantum state  $|\psi(\vec{\beta}^*, \vec{\gamma}^*)\rangle$  that approximates the ground state of  $H_c$ . This, in turn, allows the solution of the underlying QUBO problem to be extracted. Increasing the number of layers  $p$  leads to an increase in the expressibility of the ansatz and, hence, the potential quality of the output solution, at the expense of longer computation times [27]. We have chosen to investigate the original form of QAOA described in [20] in our work, motivated by its relative simplicity and amenability to the implementation on currently available quantum computers, with a few tens of qubits. In principle, variants of QAOA (such as recursive-QAOA [28], [29], and tailored mixer and initialization choices [30], [31], [32]) are expected to yield better results and can be straightforwardly adapted for our framework to solve MCOPs.

Next, we provide a general description of MCOPs. Given  $L$  objective functions that map states from a state space to real numbers, multiobjective optimization asks for states that are optimal with respect to all of the  $L$  objectives. Since the objectives generally produce competing effects with one another, states that are simultaneously optimal in all objectives generally do not exist. In this case, we ask instead for a set of Pareto-optimal/efficient states, which are optimal in the sense that no other states, which improve on at least one individual objective without deteriorating in others can be found. On top of the objectives, a number of additional constraints may also be present, imposing further complications.

More concretely, we require that the  $L$  individual objectives and  $K$  constraints of the problem can be cast into a QUBO problem. This yields an encoding of the states  $x = (x_1, \dots, x_n)$  as vectors of  $n$  binary decision variables  $x_1, \dots, x_n \in \{0, 1\}$  in a state space  $S = \{0, 1\}^n$ ,  $L$  quadratic objective functions  $C_i : S \rightarrow \mathbb{R}$  of the form

$$C_i(x) = x^\top \mathbf{Q}_i x, \quad i = 1, \dots, L \quad (3)$$

and  $K$  quadratic penalty functions  $P_j : S \rightarrow \mathbb{R}$  that enforce the  $K$  constraints

$$P_j(x) = x^\top \mathbf{P}_j x, \quad j = 1, \dots, K \quad (4)$$

where  $\mathbf{Q}_i$ 's and  $\mathbf{P}_j$ 's are upper triangular matrices with real entries. To compute these functions with quantum computers, we write the objective and penalty functions in terms of Ising variables  $s = (s_1, \dots, s_n) \in S' = \{-1, 1\}^n$ , with  $s_i = 2x_i - 1 \in \{-1, 1\}$ , to eventually convert them into equivalent Hamiltonians and denote the corresponding objective and penalty functions in terms of  $s$  as  $E_i^C(s)$  and  $E_j^P(s)$ , respectively. The MCOP is then

$$\min_{s \in S'} (E_1^C(s), \dots, E_L^C(s)) \quad (5)$$

subject to the minimization of  $E_1^P(s), \dots, E_K^P(s)$ .

Numerous classical methods exist to solve the previous optimization problem. Of particular interest to us is linear scalarization, which seeks for a solution of (5) by first solving a simpler problem obtained by aggregating the objective and penalty functions of the full MCOP in a linear manner

$$\min_{s \in S'} \left( \sum_{i=1}^L w_i E_i^C(s) + w_P \sum_{j=1}^K E_j^P(s) \right) \quad (6)$$

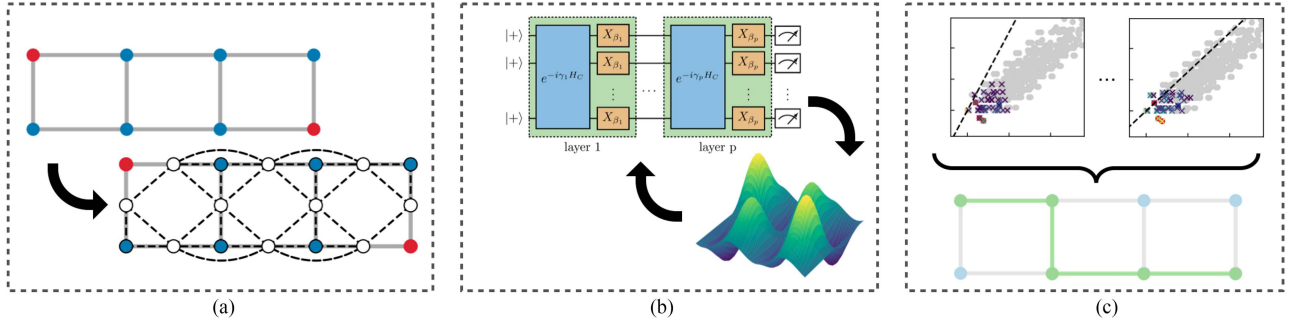
where the scalarization weights  $w_i \geq 0$ , and the penalty weight  $w_P > 0$  controls the degree at which infeasible solutions are penalized. A choice of  $w_P$  that is sufficiently large ensures that the solution of (6) is necessarily a Pareto-optimal solution of (5) that satisfies the  $P$  constraints imposed (i.e., a feasible solution) [3], [33], [34]. In practice, smaller values can be used, and we set  $w_P = 1$  throughout this work in an empirical manner, and set  $w_i \in [0, 1]$  without the loss of generality (since only the relative weight  $w_P/w_i$  matters).

For further details and references on MCOPs and scalarization, we refer the reader to Appendix B. Furthermore, Appendix A presents a survey of relevant work in the usage of near-term quantum algorithms in solving MCOPs and the shortest-path problem.

## A. OBTAINING THE PARETO FRONT WITH THE QAOA

We now describe our procedure for solving MCOPs with near-term quantum computers, which is summarized in Fig. 1.

Following (6), the QUBO objective vector  $\vec{E}^C(s) = (E_1^C(s), \dots, E_L^C(s))$  is first scalarized by aggregating them as a convex sum with weights  $\vec{w} = (w_1, \dots, w_L)$  with  $w_i \in [0, 1]$  and  $\sum_{i=1}^L w_i = 1$ . This is added to the penalty terms



**FIGURE 1.** Summary of workflow to approximate the Pareto front of an MCOP. (a) Visualization of the conversion from the problem's underlying graph to the Hamiltonian corresponding to the QUBO formulation of the MCOP. The graph (with solid lines denoting its edges and circles denoting its nodes) is converted to the Hamiltonian (8) that incorporates the objective and constraints of the problem. The Hamiltonian itself can be represented by a graph,  $G_H$  (bottom right graph, with dotted lines denoting edges and circles denoting qubits). The illustration describes the conversion for the multiobjective routing problem, which is elaborated in Section III. (b) Quantum-classical optimization with QAOA to solve the optimization problem (2) for the scalarized cost Hamiltonian (8). The  $p$ -layer QAOA circuit is run at different angles according to the feedback of a classical optimizer, navigating the cost landscape until it converges to a minima. (c) Aggregation over results of one or more scalarization choices, determination of Pareto front with classical solver, and extraction of final result(s).

$\sum_{i=1}^K E_i^P(s)$  with each penalty function weighted equally by 1, resulting in a scalarized QUBO cost function

$$E_{\text{scalar}}(s) = \vec{w} \cdot \vec{E}^C(s) + \sum_{i=1}^K E_i^P(s). \quad (7)$$

Minimization of this cost function then yields a feasible solution that corresponds to a Pareto-optimal solution  $s^*$  of the MCOP (5). This optimization can be performed on a quantum computer by constructing equivalent objective Hamiltonians  $H_1^C, \dots, H_L^C$  and penalty Hamiltonians  $H_1^P, \dots, H_K^P$  via transformation to Ising variables, and aggregating them to result in an analogous scalarized cost Hamiltonian

$$H_{\text{scalar}} = \vec{w} \cdot \vec{H}^C + \sum_{i=1}^K H_i^P \quad (8)$$

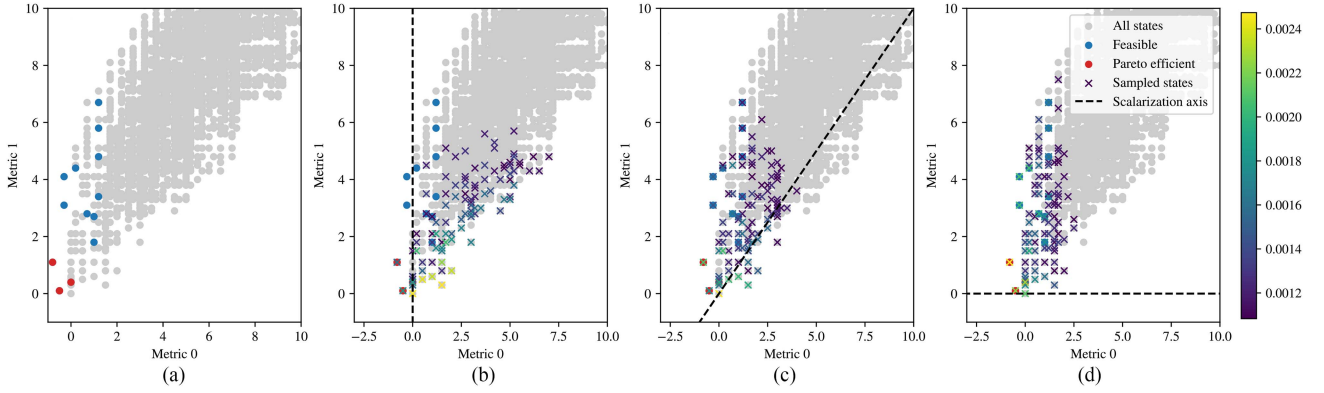
whose ground state encodes the same solution  $s^*$ , where  $\vec{H}^C = (H_1^C, \dots, H_L^C)$ . We perform the ground state search with a VQA, chosen to be QAOA in our work due to its simplicity and amenability by current quantum computers. That is, we variationally optimize the QAOA circuit of (1) to produce an optimized output quantum state  $|\psi(\vec{\beta}^*, \vec{\gamma}^*)\rangle$  that approximates the ground state of  $H_{\text{scalar}}$ , which is then sampled from  $k$  times to yield a set of candidate solutions  $\mathbf{B} = \{b_1, \dots, b_k\}$ , with  $k$  always chosen to be at most polynomial in the number of qubits  $n$ . The set  $\mathbf{B}$  then constitutes a reduced search space within which Pareto-optimality can be efficiently checked with any classical method (such as brute force). The output of this procedure is therefore expected to be a set of feasible solutions that are Pareto-optimal within  $\mathbf{B}$ . As the number of QAOA layers is increased to  $\infty$ , the QAOA ansatz supports an increasingly better approximation of the ground state of (8). Provided that it can be found in the variational procedure, sampling this state would then return solutions that are also Pareto-optimal within  $S$ , with high probability.

The aforementioned scalarization procedure involves the linear aggregation of objectives. The convex weights  $w_i$  can be interpreted as a priori preferences that the decision maker can select, which biases the resulting solution according to  $\vec{w}$ . While other scalarization schemes that involve higher-order terms (such as quadratic scalarization) and inequality constraints (such as Chebyshev scalarization) are available, we focus on the linear scheme, as it preserves the QUBO form of the individual objectives and constraints. The resulting scalarized cost function (7) then contains only linear and quadratic terms. We discuss possible limitations and extensions of this approach in Appendix B.

So far, this procedure belongs to the class of a priori methods, where the decision makers' preferences are specified before the optimization to bias the output of the optimization procedure (cf. Appendix B for details). To recover the entire Pareto front in an unbiased way, this procedure can be converted into a completely a posteriori method by repeating the procedure with numerous choices of  $\vec{w}$  in a problem-agnostic manner, with, e.g., uniform random sampling of the weights or a discretization over all possible weights. The full procedure to recover the Pareto front with QAOA is illustrated in Fig. 1:  $M$  different scalarization weights are chosen corresponding to  $M$  QAOA runs, resulting in  $M$  sets of candidate solutions  $\mathbf{B}^1, \dots, \mathbf{B}^M$ , which are aggregated as  $\bigcup_{i=1}^M \mathbf{B}^i$ . A classical method then checks for Pareto-optimality within this set, yielding a set of feasible solutions that approximates the Pareto front.

## B. VISUALIZATION OF SOLUTIONS IN THE PARETO PLOT

The state space of an  $L$ -objective MCOP can be visualized in a Pareto plot [2], [3], which is an  $L$ -dimensional plot with  $2^{|S|}$  points. Each point corresponds to a state/bitstring  $s$ , with the value of the  $L$  objectives as its coordinates. Pareto-optimal solutions then constitute points located at the boundary of the region populated by the  $2^{|S|}$  points [2].



**FIGURE 2.** Pareto plot of a four-objective problem projected on two of the metrics. Shown in (a), each point in the plot corresponds to a possible state/solution (for a total of  $2^n$  points), with feasible solutions marked as blue and Pareto-optimal solutions marked as red. We also select the top  $k = 100$  most probable states from  $|\psi(\vec{\beta}^*, \vec{\gamma}^*)\rangle$  and color code them according to their probability (colored crosses). We plot the results for three different scalarization choices, with the gradient of the scalarization weight vector visualized (black dotted lines passing through the origin). The scalarization vectors are  $\vec{w} = (0, 1, 0, 0)$  for (b),  $\vec{w} = (1/4, 1/4, 1/4, 1/4)$  for (c), and  $\vec{w} = (1, 0, 0, 0)$  for (d). The underlying network is a triangular lattice graph with 6 nodes and 9 edges, requiring 13 qubits to encode.

In our case, we must also account for the penalty terms  $E_1^P(s), \dots, E_K^P(s)$ , whose contribution should be reflected in the Pareto plot in a consistent way. This can be achieved by defining a point  $\vec{r} = (r_1, \dots, r_L)$  in the  $L$ -dimensional Pareto plot as follows:

$$r_i(s) = \sum_j^K E_j^P(s) + E_i^C(s), \quad i = 1, \dots, L. \quad (9)$$

This definition has the property that the cost function arising from linear scalarization with weights  $\vec{w}$  can be interpreted as the projection of points in the Pareto plot onto the vector  $\vec{w}$ , since

$$\vec{r} \cdot \vec{w} = \sum_j^K E_j^P(s) + \vec{w} \cdot \vec{E}^C(s) \quad (10)$$

which is exactly the scalarized cost function (7) that is minimized in our procedure.

We now give a brief qualitative discussion of some notable features of the Pareto plot obtained from the application of QAOA to an MCOP. We remark that the precise details of a Pareto plot will naturally differ depending on the specific problem instance, but the observations we make in the following were general to all the cases we have considered in this work. We consider a small network routing problem that can be cast into a QUBO problem in the form of (3) and (4) (to be detailed in Section III), defined on 13-qubits for illustration. The underlying network consists of six nodes arranged as a single row of a triangular lattice, with the source and destination nodes located at endpoints of the row, with parameters specified in Table 6.

Fig. 2(a) shows the Pareto plot for our 13-qubit network routing problem, resulting from noiseless classical simulations of the algorithm for different scalarization choices. The problem involves four objective functions, and the figures show a projection onto the plane of two of those objectives.

Each point in the plot corresponds to a possible state (a bitstring), with feasible states marked as blue and Pareto-optimal solutions marked as red. The top- $k$  most probable states from  $|\psi(\vec{\beta}^*, \vec{\gamma}^*)\rangle$  are marked with crosses and color coded according to their probability (colored crosses). In Fig. 2(b)–(d), we illustrate the solution returned by QAOA with different scalarization weights [ $\vec{w} = (0, 1, 0, 0)$  for (a),  $\vec{w} = (1/4, 1/4, 1/4, 1/4)$  for (b), and  $\vec{w} = (1, 0, 0, 0)$  for (c)].

A few important observations from these plots are the following.

- 1) *The states with highest sampled probability in the output state  $|\psi(\vec{\beta}^*, \vec{\gamma}^*)\rangle$  are in the low-energy sector of the Pareto plot:* This is a direct consequence of the quantum-classical optimization loop that results in an output quantum state  $|\psi(\vec{\beta}^*, \vec{\gamma}^*)\rangle$  that has minimal scalarized energy when optimized. Sampling from  $|\psi(\vec{\beta}^*, \vec{\gamma}^*)\rangle$  then leads to a set of candidate states  $B^i$  that are biased to contain high-quality/low-energy states of the scalarized problem, which we can observe from the clustering of the colored crosses in the bottom left region of the Pareto plot.
- 2) *States sampled from  $|\psi(\vec{\beta}^*, \vec{\gamma}^*)\rangle$  are squeezed in the direction of the scalarization line:* Scalarization amounts to projecting points on the Pareto plot to the line passing through the origin with a gradient equal to the scalarization weight vector (black dotted line). Different choices of scalarization, therefore, result in different regions in the Pareto plot that are favored. This feature can be observed clearly through the Pareto plots by the clustering of high probability states (yellow, light-colored crosses) along the direction of the scalarization line. As an example, for the scalarization choice corresponding to weighting Metric 2 with unit

weight (the second of the four plots in Fig. 2), the sampled states can be observed to be squeezed vertically, toward the bottom of the Pareto plot.

- 3) *Points in the Pareto front corresponding to solutions of other scalarization choices can be sampled:* If there are other Pareto-optimal states close to the solution of the scalarized problem, in the sense that they are low-energy eigenstates of the scalarized problem, the set of candidate states  $\mathbf{B}^i$  sampled from  $|\psi(\vec{\beta}^*, \vec{\gamma}^*)\rangle$  may contain these nearby Pareto-optimal solutions with relatively high probability. In principle, this allows multiple Pareto-optimal states to be obtained from a single iteration of the procedure, which is a desirable feature in situations where a thorough coverage of the entire Pareto front is required, instead of only a single solution [3]. For example, in Fig. 2, the rightmost red point (a Pareto-optimal solution) is sampled with relatively high probability in all three scalarization choices; even though it is not a solution of any of these scalarized Hamiltonians, it remains a low-lying excited eigenstate and can, thus, be sampled. We observe that this is often the case in problem instances we consider.
- 4) *Points in the Pareto front located in locally concave regions can be sampled:* Again, because states near the target Pareto optimal point can be sampled with high probability from  $|\psi(\vec{\beta}^*, \vec{\gamma}^*)\rangle$ , states located at locally concave regions of the Pareto front can in principle be sampled as long as they are near the target (in the Pareto/objective space), even with linear scalarization. This allows Pareto-optimal points located at concave regions to be identified, even though they can never correspond to solutions of scalarized cost functions. Using the rightmost red point in Fig. 2 as an example again, we observe that it is sampled with relatively high probability, despite being located at a locally concave region, which cannot be obtained as the solution of any scalarized Hamiltonian.

The latter two observations indicate a degree of robustness to the choice of scalarization and the concavity of the Pareto region, due to QAOA being a sampling-based algorithm, which requires sampling from the optimized quantum state in order to extract information it. This introduces a tradeoff between the ability of QAOA to approximate the ground state (by increasing  $p$ ), and for  $|\psi(\vec{\beta}^*, \vec{\gamma}^*)\rangle$  to remain as a superposition of energy eigenstates so that low-energy states can be sampled (by keeping  $p$  small), potentially imposing a less stringent scaling of  $p$  with problem size.

Finally, we remark that most points in the Pareto plots correspond to infeasible solutions, which violate at least one of the problem constraints encoded in the penalty terms of (8). As we explain in Appendix I, in general, the fraction of feasible to infeasible solutions decreases rapidly with the network size. This fact represents a challenge for QUBO approaches where the encoding of constraints as penalty terms is unable to restrict the search through solutions to remain

solely within the feasible subspace [30], [31], [32]. We discuss this point further in Section IV-A in the specific context of QAOA.

### III. APPLICATION TO THE NETWORK ROUTING PROBLEM

With a general description of our algorithm for MCOPs established, we now consider the concrete example of the multiobjective network routing problem.

We consider a generic multihop wireless network with relay stations distributed throughout a geographical area. This can be modeled as an undirected weighted graph  $G = (V(G), E(G))$  (which we also refer to as the network's graph) with nodes  $V(G)$  that correspond to relay stations and edges  $E(G)$  defined by possible transmission paths between stations (directed graphs can be included with a more general formulation [33]). Two nodes are specified to be the source and destination nodes (denoted  $s$  and  $d$ , respectively). The routing problem then asks for data transmission routes between  $s$  and  $d$  that are Pareto optimal with respect to multiple objectives, which can be written as functions of the node and edge weights of the graph. This is an instance of the multiobjective shortest-path problem, which is a combinatorial optimization problem known to be NP-hard [35]. Its solution is relevant across a wide range of network design and optimization tasks [4], [5], [6], [8], [9], [10], especially beyond current wireless network protocols based on individual objectives such as the optimized link state routing protocol [36] which only considers hop count and is, thus, unable to maximize network resource utilization.

The form of the objectives depends on specific performance demand requirements. We consider the following four objectives in our work.

- 1) *Path loss:* Transmission between two stations incurs an energy cost that is dependent on their physical distance.
- 2) *Node Delay:* Signal processing at each station incurs a time delay, which sums up to an overall delay.
- 3) *Data rate:* The total data output rate of a transmission path is determined by the minimum data rate along its path.
- 4) *Bit error:* Bit errors occur with finite probability during transmission between two stations, which depends on the path loss and the channel's noise profile.

To solve this problem with our approach, a QUBO formulation of the multiobjective shortest-path problem with the previous objectives is needed. This can be achieved by first considering the QUBO formulation of the constraints of the shortest-path problem [in the form of (3)], and formulating the objectives defined earlier as node or edge weights to be scalarized [in the form of (4)]. The QUBO formulation of the shortest-path problem is well-studied, especially in the context of quantum annealing [33], [34], [37]. Following the work in [33], we provide a complete description in terms of binary variables together with the form of the quadratic cost function in Appendix C.

Subsequently, we provide formulations of the above four objectives in Appendix E. Along with reasonable assumptions on network parameters based on software-defined radio use cases [38], the associated objective and penalty Hamiltonians  $H_i^C$  and  $H_i^P$  can then be constructed explicitly, thereby fully specifying the inputs for our procedure in Fig. 1. As we will detail in Section IV, the resource requirements of the resulting quantum circuit is

$$\begin{aligned} n &= |V(G)| + |E(G)| - 2, \\ D_1 &\leq 2\Delta_G, \end{aligned} \quad (11)$$

where  $n$  is the number of qubits required,  $D_1$  the depth of each layer of the QAOA circuit, and  $\Delta_G$  the maximum degree of  $G$ .

Importantly, we observe that for a network with graph  $G$ , the graph of its corresponding scalarized Hamiltonian  $G_H$ —that is, the graph with connectivity defined by the quadratic terms of (8)—can be interpreted as the network graph’s *middle graph* [39]. Denoted as  $M(G)$ , the middle graph of  $G = (V(G), E(G))$  [39] is the graph  $M(G) = (V(G) \cup E(G), E')$ , with vertices  $u, v$  that are adjacent if either:

- a)  $u$  is a vertex in  $G$  and  $v$  is an edge in  $G$  incident to  $u$ , or
- b)  $u$  and  $v$  are edges in  $G$  that are both incident on the same vertex.

Using known properties of  $M(G)$ , we will exploit this correspondence in the next sections to bound the resources needed to execute the algorithm. Hamiltonian’s construction is illustrated in Fig. 1(a). The network’s graph  $G$  (top left graph, with solid lines denoting its connectivity, red circles denoting source and destination nodes, and blue circles denoting remaining nodes) is converted to a Hamiltonian  $H$ , with connectivity defined by the graph  $G_H$ , which takes the form of  $M(G)$  (bottom right graph, with dotted lines denoting the connectivity of  $G_H$  or  $M(G)$ , and circles denoting nodes). We elaborate on points regarding  $M(G)$ , including its definition and properties, in Appendix D.

#### IV. THEORETICAL AND NUMERICAL SCALING ANALYSES

We now turn to analyze in detail the scalability and performance of the QAOA approach to the network routing problem.

In Section IV-A, we begin with a short qualitative discussion of the role of the QAOA mixer and its initial state in determining the efficacy of the algorithm. Subsequently, in Section IV-B, we discuss and provide a pragmatic analysis of the hardware resources required to implement a problem instance of a given size and compare these requirements against known limitations of VQAs. This allows us to distinguish between problem instances that are infeasible for our approach, from those that are potentially amenable. Finally, in Section IV-C, we present the results of numerical simulations that explore the scaling of several success metrics as a function of problem size and the QAOA depth  $p$ . We remark that due to limitations on the system sizes that can be

simulated on a classical computer, strong conclusions on the asymptotic performance of the algorithm cannot be drawn at this stage, and is a subject of ongoing study even for more generic applications of QAOA [22].

All analyses performed here are based on a single scalarization (i.e., we apply the algorithm in an a priori manner by specifying the scalarization weights  $\vec{w}$  beforehand). As described in Section II-A, if the intent is to recover the entire Pareto front instead, the total run time of the procedure depends on the number of scalarizations, which introduces an additional multiplicative overhead in the time required to obtain the Pareto front. As is also the case with classical approaches, this additional complexity is highly problem-dependent.

#### A. QAOA MIXER AND INITIAL STATE

We begin with a qualitative discussion of the role of the QAOA mixer and circuit initial state in determining the efficacy of the approach. In this work, we have used the standard QAOA initial state and mixer pair [20], where the circuit is initially prepared in an equal superposition of all solutions, i.e.,  $|+\rangle^{\otimes n}$ , and the mixer Hamiltonian  $H_m = -\sum_i X_i$  drives bit flips across the register. Constraints are enforced through energy penalties in the cost Hamiltonian (8), with the search ideally converging toward low-energy feasible solutions. However, the solution space is increasingly dominated by infeasible configurations as the network size grows (see Appendix I), motivating the need for alternative ways of searching the solution space [30], [31].

In the context of the network routing problem, we leave the question of designing improved initial state and mixer Hamiltonian pairings for future work (see also our remarks in Section VI), noting that they do not alter our underlying scheme to solve MCOPs. However, we remark that the goal of such strategies is to reduce the size of the solution space to be searched, either by entirely avoiding infeasible solutions or by avoiding some subset of them. In the former case, where the search takes place through feasible solutions only, the penalty terms in (8) can be eliminated. In the latter case (which may arise if a strategy to search only feasible solutions cannot be found, or carries impractical resource requirements), the penalty terms would still be necessary to enforce constraints indirectly (i.e., through the objective function).

#### B. COST HAMILTONIAN: THEORETICAL RESOURCE ESTIMATIONS

To determine the resource requirements of our QAOA approach, we examine the runtime and number of qubits necessary for its implementation on quantum hardware. This is determined by considering the resources needed to implement each layer of the QAOA circuit, and the number of repeated executions required to estimate the cost function up to an error  $\epsilon$ .

We remark that for a given error tolerance  $\epsilon$ , a complete performance analysis would involve exposing the dependence of the number of required QAOA layers  $p$  on problem size, which is challenging from both an analytic and numerical perspective. This is further related to the issue of trainability and existence of barren plateaus of VQAs [40], [41], [42], which depend strongly on problem class, presence of noise, and ansatz choice, and is a subject of ongoing study. Specifically, if the physical circuit depth for an application scales super-linearly in  $n$ , the optimization procedure will suffer from a *noise-induced barren plateau* [40], implying expensive gradient computations that scale exponentially in  $n$ , a phenomenon that is conceptually similar to the issue of vanishing gradients that previously plagued the training of neural networks in classical machine learning [43].

We begin by outlining a few assumptions and simplifications that we will make. First, the quantum circuit of a  $p$ -layer QAOA consists of  $p$  alternating mixer and cost unitary operations (1). Since elementary gates in the cost unitary  $e^{-i\gamma H_c}$  fully commute with one another, the 1-qubit RZ gates can be scheduled to be executed first, in parallel. Together with the fact that the mixer unitary consists only of 1-qubit RX gates, which can also be executed simultaneously, and that 1-qubit gates execute significantly faster than 2-qubit gates, we only need to consider contributions from 2-qubit gates. Second, we neglect compilation overheads arising from qubit routing, i.e., the need to include additional SWAP gates to carry out 2-qubit gates that are not natively executable on a quantum processor. While this is a valid assumption for quantum processors with a fully-connected topology (such as currently available small-scaled ion trap devices, which we utilize in Section V), the mismatch between the quantum processor and the cost Hamiltonian's topologies will generally require routing (or ion shuttling, in the case of larger ion trap devices [44]). The routed output depends on the degree of mismatch and the routing strategy employed, among many other factors that are beyond the scope of our discussion [45]. Nonetheless, we remark that the inclusion of  $n - 1$  additional layers of SWAP gates is a naive upper bound [46], [47], incurring at most  $O(n)$  additional layers of quantum gates. Third, to streamline our argument we ignore quadratic terms arising from the source and destinations constraints, which only incur a negligible, constant number of terms (detailed in Appendix C).

Finally, we remark that in our formulation of the problem, increasing the number of objectives solely results in additional linear terms in the cost Hamiltonian (8), since only the penalties contain quadratic terms. An increase in the number of objectives, therefore, does not directly incur additional time and qubit costs.

## 1) NUMBER OF QUBITS

Since the graph of the cost Hamiltonian  $G_H$  can be mapped to the middle graph  $M(G)$ , the number of qubits involved is equal to the number of nodes of  $M(G)$ , resulting in  $n =$

$|V(G)| + |E(G)| = O(|V(G)|\Delta_G)$ , where  $\Delta_G$  is the maximum degree of the network's graph  $G$  [see Appendix D, (25)]. This implies that problem instances with a maximum degree independent of problem size are expected to be more efficient, with  $n = O(|V(G)|)$ .

## 2) COST COMPUTATION TIME

The computation time of the  $p$ -layer QAOA cost function is determined by two factors: 1) The number of repetitions  $n_{\text{rep}}$  required to estimate the cost function expectation value, given some specified error tolerance and 2) the depth of each executed circuit, which we denote  $D_p$  for  $p$  layers. These two factors combine to give an execution time of  $O(D_p n_{\text{rep}})$ , or  $O(pD_1 n_{\text{rep}})$  in terms of the depth of a single QAOA layer  $D_1$ .

$D_1$  depends on the number and connectivity of the 2-qubit terms in the cost Hamiltonian. As they are Ising gates that are diagonal in the computational basis, they fully commute with one another and can be scheduled to maximize parallelization. The determination of such a schedule amounts to the solution of an edge-coloring problem on  $G_H$ , where edges (i.e., gates) of the same color are scheduled to be executed simultaneously. The minimal number of colors required is commonly termed the edge chromatic number  $\chi'(G_H)$ , so we have  $D_1 = \chi'(G_H)$  layers in total. To estimate  $\chi'(G_H)$ , note that the structure of  $G_H$  is completely determined by (18)–(20). Intuitively, we expect  $G_H$  to inherit the local structure of  $G$ , since the squares in (18)–(20) only lead to interactions between nodes (of  $G_H$ ) representing nodes and edges incident to them (i.e., the  $x_i x_{ij}$  terms), and between edges that are incident on the same node (i.e., the  $x_{ij} x_{ik}$  terms). It turns out that this is precisely the relationship between a graph  $G$  and its middle graph  $M(G)$ —in other words,  $G_H = M(G)$ , and hence,  $\chi'(G_H) = \chi'(M(G))$ . This realization allows us to leverage on known properties of  $M(G)$  to show that

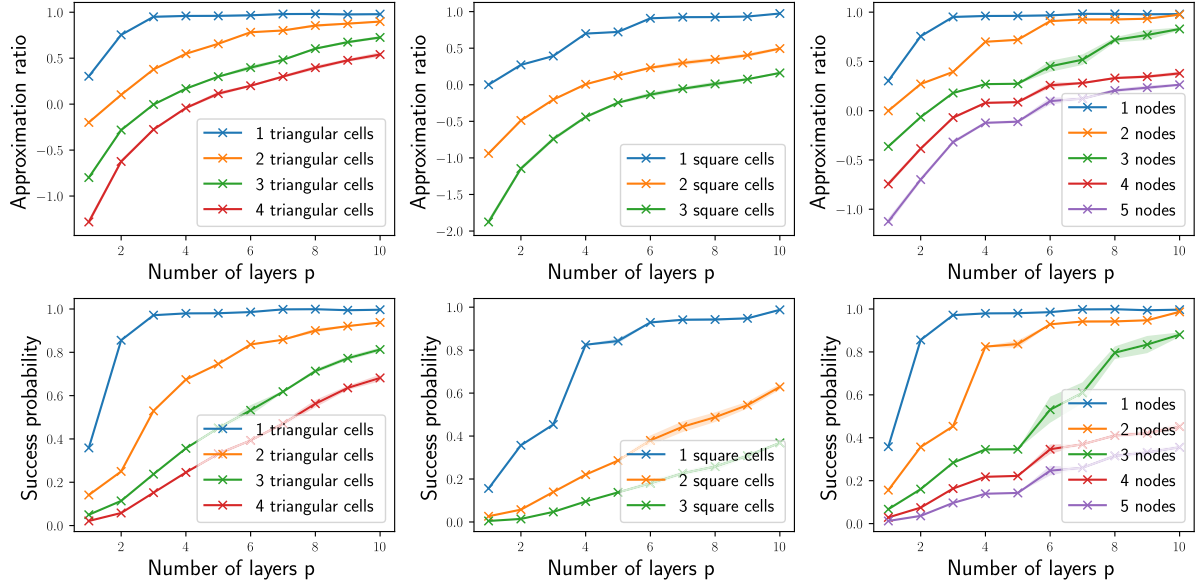
$$\chi'(M(G)) \leq 2\Delta_G \quad (12)$$

where we delegate the proof and a review of relevant ideas to Appendix D. We conclude that the depth  $D_1$  of each QAOA layer scales linearly with the maximum degree of the network's graph  $G$ , i.e.,  $D_1 = O(\Delta_G)$ .

On the other hand, it is known that  $n_{\text{rep}} = O(L/\epsilon^2)$  measurements are needed to compute expectation values up to an error tolerance of  $\epsilon$  by operator averaging [48], where  $L$  is equal to the number of Pauli terms in the operator (which is the cost Hamiltonian in our case). Of the  $L$  terms,  $|V(G)| + |E(G)| \leq \frac{1}{2}|V(G)|(2 + \Delta_G)$  terms are linear (since the number of node and edge weights is equal to the total number of nodes and edges of  $G$ ) while  $|E^{M(G)}| \leq \frac{1}{2}|V(G)|\Delta_G^2$  terms are quadratic (26), because  $G_H = M(G)$ , for a total of  $L \leq \frac{1}{2}|V(G)|(2 + \Delta_G + \Delta_G^2)$  terms. Thus,  $n_{\text{rep}}$  scales asymptotically as  $O(|V(G)|\Delta_G^2/\epsilon^2)$ .

Summarizing, the computation time of a QAOA cost function up to an error  $\epsilon$  scales as  $O(p|V(G)|\Delta_G^3/\epsilon^2)$ , which can be parallelized up to a factor of  $n_{\text{rep}}$ .





**FIGURE 3.** Scaling of the approximation ratio and success probability with the number of layers for triangular lattices, square lattices, and cycle graphs (left to right columns). Different colors indicate different problem sizes, with standard deviations displayed in shaded regions. For cycle graphs (rightmost column), the legend indicates the number of nodes besides the source and destination nodes. The data were produced after optimization with linear ramp initialization, two metrics (node based and edge based, respectively), with node/edge weights uniformly randomized between  $-1$  and  $1$ , averaged over 50 instances. Scalarization weights are chosen to be equal for both objectives, i.e.,  $\vec{w} = (0.5, 0.5)$ .

We remark that this implies efficiency for problem classes with a connectivity or maximum degree that is independent of the problem size (such as  $k$ -regular/lattice-type graphs), with only a linear dependence of the computation time in problem size. This is likely the scenario for large-scale applications, since the cost of supporting a dense network (such as fully-connected mesh networks) over large geographical regions will otherwise be high, rendering it practically infeasible. In that case, the computation time (and the number of qubits needed) becomes at most quadratic in problem size.

### C. NUMERICAL SCALING ANALYSES

#### 1) PERFORMANCE METRICS

We now introduce several measures that quantify the degree to which our scheme has successfully solved the task at hand, which is to obtain Pareto-optimal points of a multiobjective problem.

A standard measure of the quality of solution output by QAOA is the *approximation ratio*, defined as the ratio between the energy of the output state and the ground state

$$r_{\text{approx}} := \frac{\langle \psi(\vec{\beta}^*, \vec{\gamma}^*) | H | \psi(\vec{\beta}^*, \vec{\gamma}^*) \rangle}{E_{\min}}. \quad (13)$$

An approximation ratio of 1, thus, implies that QAOA has found the exact ground state of  $H$ .

In our multiobjective context, however, we ask for the set of Pareto-optimal solutions, which generally cannot be encoded as ground states of a single Hamiltonian. As a more pragmatic measure in a multiobjective context, we supplement analyses of  $r_{\text{approx}}$  by defining the *success probability*, which is the probability that a Pareto-optimal state is sampled

from  $|\psi(\vec{\beta}^*, \vec{\gamma}^*)\rangle$

$$p_{\text{success}} := \sum_{x_i \in \text{Pareto}^* \text{ front}} |\langle \psi(\vec{\beta}^*, \vec{\gamma}^*) | x_i \rangle|^2. \quad (14)$$

A success probability of 1 implies that sampling from  $|\psi(\vec{\beta}^*, \vec{\gamma}^*)\rangle$  will always yield a Pareto-optimal state. Note that it does not contain information about whether it contains points from the entire Pareto front.

#### 2) SCALING WITH SYSTEM SIZE AND NUMBER OF LAYERS

In this section, we present results from classical simulations of QAOA to study the scalability of the approach with respect to different problem sizes, graph geometries, and the number of QAOA layers  $p$ . These simulations assume an absence of both sampling and hardware noise. Taking into account limitations on the classical simulation and optimization of general QAOA circuits for large  $n$  and  $p$ , and the need to average over multiple problem instances, we consider problems of sizes up to 16 qubits and  $p = 10$  layers.

Fig. 3 displays the scaling of the approximation ratio and success probability as a function of  $p$ , for different problem sizes chosen from different graph geometries, namely triangular lattices, square lattices, and cycle graphs. For all examples, we consider two linear objectives, one involving only edge variables [i.e., of the type (17)] and another only involving node variables [i.e., of the type (16)], with node/edge weights uniformly randomized between  $-1$  and  $1$ , averaged over 50 instances. The scalarization weights are also chosen to be equal for both objectives, i.e.,  $\vec{w} = (0.5, 0.5)$ . Finally, initial parameters for QAOA are chosen according to a linear ramp initialization scheme, which is a heuristic choice that

linearly ramps up  $\gamma$ 's and ramps down  $\beta$ 's, based on the analogy between QAOA and quantum annealing [27]. Additional numerical results for the performance of this initialization scheme can be found in Appendix G.

As first observations, we see that both metrics increase monotonically as a function of  $p$ . This is in accordance with the fact that as  $p \rightarrow \infty$  with QAOA, we recover the limit of infinitesimally slow adiabatic quantum annealing and recover the ground state/optimal solution of scalarized cost Hamiltonians with unit probability. For any fixed value of  $p$ , both metrics also achieve higher values for smaller problems, again mirroring the increase in annealing time needed for larger problems. Finally, we observe that problem classes with higher average connectivity tend to reach saturation at a slower rate (number of layers to reach saturation, in descending order: cycle graphs, triangular lattice, and square lattice). These conclusions hold for generic problem instances of the same class, due to the average over numerous instances.

Focusing on the success probability (lower row of Fig. 3), we find that it increases monotonically with  $p$ , with indications of saturation at the maximum value of 1 for the limited problem sizes we have considered. These results suggest that, at least for the problem instances considered here, the procedure indeed outputs Pareto-optimal solutions with increasing frequency as we increase  $p$ . This justifies the pragmatic approach of increasing  $p$  to improve the quality of the procedure.

Nonetheless, we note that the aforementioned difficulty in classically simulating and optimizing large QAOA circuits prevents further scaling conclusions from being drawn. We leave a more complete investigation of these issues, involving the large-scale benchmarking of our scheme on actual quantum computers, and more accurate classical simulations under noisy conditions, as further work. As a final remark, a positive indication is provided in the closely related context of quantum annealing, by the polynomially vanishing energy gap between the ground and the first excited state of the Hamiltonian (8) [34], which is strongly correlated to a better performance of QAOA [27].

## VI. COMPUTATIONS ON QUANTUM COMPUTERS

In this section, we describe an implementation of our scheme for a small-scale network routing problem on the 11-qubit IonQ Harmony quantum computer, accessed through the Amazon Braket cloud quantum computing service with the OpenQAOA Python SDK [25]. This device is composed of an 11-qubit chain of  $^{171}\text{Yb}^+$  ions that can be entangled in an all-to-all manner via XX gates [49] and can, therefore, implement QAOA circuits without the need for SWAP gates.

These experiments serve to empirically verify that the scheme can yield Pareto-optimal solutions to the network routing problem, in a manner consistent with our description in Section II. We consider a 4-node fully-connected network and a 6-node square-lattice network, requiring 8 and 11 qubits, respectively, to encode in QUBO form. For both problem instances, we consider four objectives relevant to the

network routing problem: 1) data rate; 2) path loss; 3) node delay; and 4) bit-error-rate. Further details on the problem instances can be found in Appendix F, including parameters used to generate the problem instances in Tables 2–5.

We run the algorithm for the scalarization choice  $\vec{w} = (1/4, 1/4, 1/4, 1/4)$ , with  $p = 1$  and 2000 shots per circuit execution, and the standard QAOA circuit involving standard single-qubit  $RX$  mixers and uniform computational basis state initialization. Starting at initially suboptimal parameters, COBYLA [50] is chosen as the classical optimization algorithm.

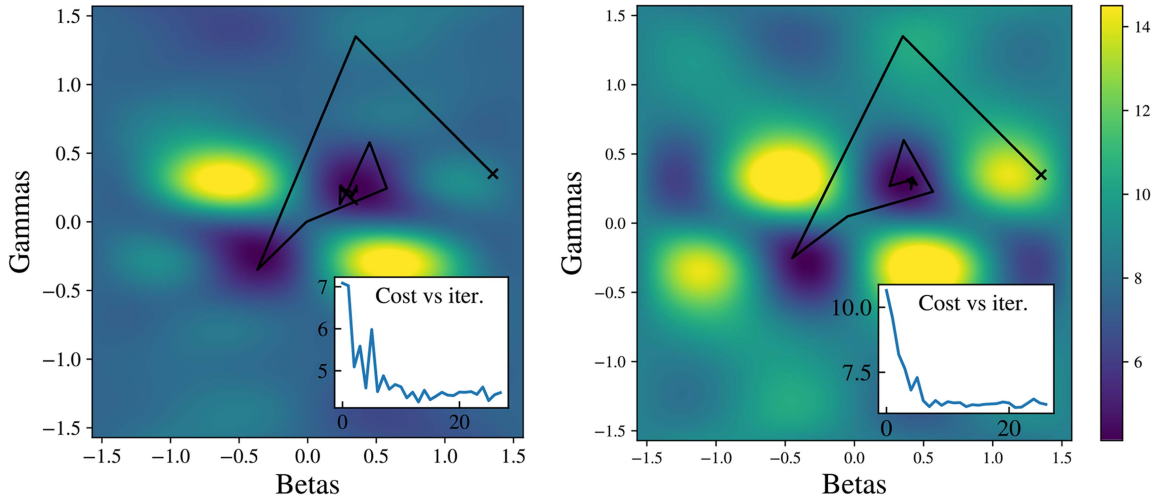
The results of the optimization are shown in Fig. 4, which displays the trajectory of the parameters  $\beta$  and  $\gamma$  (black line), plotted on top of the cost function landscape obtained by classical simulation. The inset at each plot shows the evolution of the cost function value during optimization. Starting from suboptimal initial parameters, we observe convergence to local minima for both problems.

To obtain Pareto-optimal solutions to the network routing problem, we sample from the optimized quantum states. Fig. 5 displays the Pareto plot for two pairs of objectives obtained from the IonQ device, with the top  $k = 125$  most probable states marked as colored crosses (we limit  $k$  for visual clarity). Consistent with our discussions in Section II, we observe the clustering of high-probability states close to the Pareto front in these plots. The observed success probabilities are also computed to be  $\approx 0.008$  for the 11-qubit problem and  $\approx 0.04$  for the 8-qubit problem, which are higher compared to random sampling ( $3/2^{11} \approx 0.001$  for the 11-qubit problem and  $4/2^8 \approx 0.016$  for the 8-qubit problem).

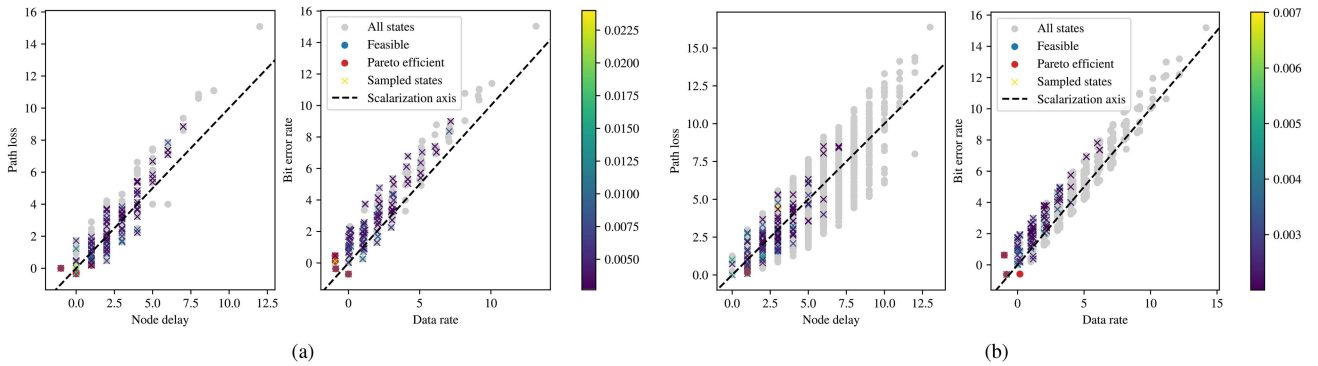
## VI. CONCLUSION AND OUTLOOK

This work proposes and studies a procedure to obtain Pareto-optimal solutions of MCOPs in a manner amenable to near-term quantum computers via VQAs. By scalarizing the multiobjective cost function in QUBO form and variationally optimizing it with a quantum computer, the procedure allows the Pareto front of MCOPs to be recovered and visualized in an intuitive manner. Focusing on the practically relevant network routing problem, which is an instance of the multiobjective shortest-path problem, efficient QUBO formulations for practically relevant objectives were detailed. Analytical scaling analyses on the resources required for this procedure point toward implementation feasibility for real-world problem instances, for networks with bounded maximum degrees. This is supplemented by numerical simulations at small scales, which show that high-quality results can be obtained by systematically increasing the depth of the QAOA ansatz. Finally, we tested our approach on an ion trap quantum computer, obtaining consistent results that verify its applicability, constituting one of the first studies of multiobjective optimization on actual digital quantum computers.

Our suggested framework can be readily generalized to incorporate other MCOPs and VQAs, as long as QUBO



**FIGURE 4.** Trajectory of the QAOA parameters during optimization (black line) for both the 8-qubit fully-connected network (left plot) and 11-qubit square lattice networks (right plot), overlaid on the full cost function landscape obtained by classical simulation of the QAOA circuits. Initial parameters are displayed as black crosses. The inset shows cost versus iteration plots of the optimization.



**FIGURE 5.** Pareto plots obtained from sampling the optimized quantum state of QAOA with the IonQ device for both the 8-qubit fully-connected network (left two plots) and 11-qubit square lattice networks (right two plots). Each point in the plot corresponds to a possible state, with feasible solutions marked as blue and Pareto-optimal solutions marked as red. The top  $k = 125$  most probable states sampled are plotted as crosses and color coded according to its probability. The gradient of the scalarization weight vector is also visualized (black dotted lines). (a) Fully-connected graph, 8 qubits. (b) Square lattice, 11 qubits.

formulations of the individual objectives are available. Additionally, to further tailor the scheme we have developed for MCOPs, more general scalarization choices beyond the simple linear scalarization approach could be explored. Examples are the Chebyshev and epsilon-constraint scalarization schemes, which are more powerful and versatile but incur higher computational costs in the form of additional inequality constraints [1], [3].

Another improvement is in the direction of tailored state initializations and mixer designs [30], [31]. Due to the large fraction of infeasible solutions in the state space, the default initial state of the QAOA procedure—an equal superposition of all computational basis states—is highly sub-optimal. Ideally one would like to be able to: (a) prepare an initial state that is an equal superposition over all feasible solutions and (b) implement a mixer Hamiltonian that does not induce transitions to infeasible solutions, so that the search is fully restricted to the feasible subspace. In the context of the network routing problem we leave

this as a question for future work, and note that Zhang et al. [32] describe QAOA mixer Hamiltonians for network flow optimization. We also remark that the two aforementioned requirements (a) and (b) may carry a significant resource footprint, depending on the complexity of preparing the initial state and the polynomial degree of the mixer Hamiltonian.

We can also ask whether it is possible to more directly leverage the fact that QAOA works with quantum superpositions of candidate solutions to the optimization problem. Rather than optimizing based on the energy directly, it may be fruitful to consider modifying the cost function to one that favors states, which contain Pareto-optimal configurations with high probability, which shares conceptual similarities with recently developed methods such as the conditional value-at-risk technique [51]. Such a modification can be performed efficiently on the sample since the Hamiltonian corresponding to QUBO problems only contains terms that fully commute with one another.

Even on classical computers, MCOPs are generally NP-hard optimization problems that are solved with metaheuristic optimization algorithms, such as particle swarm and genetic algorithms (cf. Appendix B). Can our scheme, or VQAs in general, outperform them in some respect at large scales that are challenging for classical approaches. Due to the difficulty to theoretically analyze the performance of QAOA at large scales, and the intractability of classical simulations of quantum computations, general claims about performance likely must be made empirically, on real quantum computers. While we were able to implement and obtain solutions for the procedure for small-scale problems, applications in practical settings will require larger and more reliable quantum computers than those available today. Our work serves as a first step toward answering this question, and alongside existing studies [22], [23], [24], motivate the need for large-scale experiments on quantum computers, and benchmarks against existing classical methods.

## APPENDIX A SURVEY OF RELEVANT WORK

This section briefly discusses relevant work in the usage of near-term quantum algorithms in solving MCOPs and the shortest-path problem.

Applications of quantum annealing to the shortest-path problem were first studied in [33] and [37], which explored different QUBO formulations of the shortest-path problem. Krauss and McCollum [33] further provided numerical and empirical evidence on its performance on a quantum annealer. With the same formalism, Hauke et al. [34] studied a related problem in the context of chemistry, by again solving a relatively large problem on a quantum annealer, providing numerical evidence on its efficiency with increasing system size through a polynomially (rather than exponentially) vanishing ground state energy gap. In the context of VQAs, Fan et al. [52] studied the application of QAOA to the shortest-path problem.

On the other hand, the application of near-term quantum algorithms to MCOPs is a relatively unexplored area. Barán and Villagra [53] provide theoretical arguments verifying that adiabatic quantum algorithms can be applied to find Pareto-optimal points corresponding to the solution of scalarized objective functions. Most recently, Urgelles et al. [54] applied QAOA to the multiobjective routing problem in the context of 6G communication networks by sequentially solving single-objective problems with a lexicographic ordering method. After the announcement of our work, Ekstrom et al. [55] further studied a variational approach that can simultaneously obtain multiple points on the Pareto front.

In relation to the previous references, our work explores the solution of the multiobjective routing problem with QAOA through scalarization and provides an in-depth analysis on the approach's scalability and performance. The approach can be generalized to VQAs beyond QAOA, MCOPs beyond the routing problem, and techniques beyond linear scalarization. To the best of our knowledge, experiments

on actual quantum hardware to solve MOOPs (described in Section V) are new.

## APPENDIX B MULTIOBJECTIVE OPTIMIZATION

This section provides a brief review of relevant notions in MOOPs and their solutions.

Given  $m$  objective functions  $f_i : S \rightarrow \mathbb{R}, i = 1, \dots, m$  that map objects from a state space  $S$  to real numbers, multi-objective optimization asks for states that are optimal with respect to all of the objectives  $\{f_i\}$ . Depending on their forms, the objectives generally produce competing effects with one another, so states that are simultaneously optimal in all objectives generally do not exist. In this case, we ask instead for a set of *Pareto-optimal/efficient* states, which are optimal in the sense that no other states that improve on at least one individual objective without deteriorating in others can be found. The set of all Pareto-optimal states is also conventionally referred to as the *Pareto front*, and they are always located at the boundary of the region occupied by the set of states [2]. Depending on the form of the objectives, the resulting Pareto front of the problem will either be convex or concave in the objectives [2], [3], with the latter posing a higher difficulty to solve. When the state space  $S$  is finite—as is the case for problems we consider—this is called an MCOP. Depending on the form of the objectives, MCOPs are generally difficult, belonging to the class of NP-hard problems [11], [12].

Classically, the most common approach to solve MOOPs is with scalarization techniques. Instead of optimizing the set of objectives  $\{f_i\}$  directly, scalarization techniques aggregate them into a single-objective function, thereby converting a multiobjective problem into a single-objective one. This amounts to projecting an  $m$ -dimensional vector onto a line, and finding the optimal state within it. Depending on the forms of  $\{f_i\}$  and the resources required to compute the objectives, different scalarization techniques are employed—this is a well-studied subject in the literature [1], [3]. For our purposes, we consider linear scalarization, where the final aggregated objective is a convex sum of the  $m$  objectives

$$f_{\text{linear}}(x) = \sum_{i=1}^m w_i f_i(x) \quad x \in S \quad (15)$$

where the weights  $w_i$  are real numbers. This corresponds to the projection of the  $m$ -dimensional vector onto a straight line with gradient  $\vec{w}$ . A property of this procedure is that a solution of the linearized problem is also a point on the Pareto front, guaranteeing the Pareto-optimality of the linearized problem's solution. Therefore, by optimizing over different choices of the weights, we can, in principle, build up a set of different points in the Pareto front.

While this is one of the simplest scalarization techniques, it preserves the quadratic form of the objectives and constraints in our MCOP and also allows the scalarization weights  $w_i$  to be interpreted as a priori preferences that the decision maker can select. In the case where the objectives

result in a convex Pareto front, it can be easily shown that every state in the Pareto front corresponds to a solution of a linearized problem. However, this simple linear scalarization may not be sufficient to capture the entire Pareto front if the Pareto front is concave, where some points in the Pareto front may never be captured as solutions to any linearized problem. This can be overcome with more advanced scalarization techniques such as Chebyshev scalarization [1], [3], at the expense of introducing inequality constraints. In the context of our work, if a nonlinear scalarization scheme is used, the resulting cost Hamiltonian may contain higher-order terms, resulting in cost unitary circuit with greater depth due to the need to implement higher-order gates. Alternatively, one can consider a more general variational ansatz, and optimize over a nonlinear cost Hamiltonian.

Finally, to optimize the scalarized objective classically, suitable heuristic optimization algorithms are usually employed. Algorithms that work very well in practice are population (e.g., ant colony methods) and evolutionary or genetic-based algorithms, such as NSGA-II [1], [3].

More broadly, approaches to solve MOOPs can be broadly categorized into the following three classes.

- 1) *A priori methods* take the preferences (e.g., the weights/relative importance of each objective) of the engineer/decision maker into account prior to the optimization and adapts the optimization process based on this preference. Scalarization falls into this class of methods, where the weights of the objectives are specified in an a priori manner.
- 2) *A posteriori methods* aim to solve for either a representative subset or all possible Pareto-optimal solutions, only taking preferences of the decision maker into account after the optimization process.
- 3) *Interactive methods* are adaptive and iterative methods that require the continuous interaction of the decision maker at each step of the optimization process.

As discussed in the main text, our scheme belongs to the class of a priori or a posteriori methods, depending on whether the choice of scalarization weights is made explicitly.

### APPENDIX C QUBO FORMULATION OF THE MULTIOBJECTIVE SHORTEST-PATH PROBLEM

Following the problem statement in Section III, this appendix details the QUBO formulation of the multiobjective shortest-path problem. We begin by providing an encoding of the state space of the problem in terms of binary variables, before detailing the form of the quadratic cost function to be optimized. More detailed analyses in the context of annealing, including an extension to directed graphs, can be found in [33].

An encoding of the problem can be achieved with  $|V| + |E| - 2$  binary variables representing each state  $x$  in the state space [33]. (We will use  $x_i$ 's and  $x_{ij}$ 's interchangeably with

$x$  to denote states when the context is clear from now on.) The first  $|V| - 2$  variables  $x_1, x_2, \dots, x_{|V|-2}$  correspond to the nodes in the graph excluding the source and destination nodes while the remaining  $|E|$  variables  $x_{ij}$ , where  $(i, j) \in E$ , correspond to the edges. With this encoding, we consider objectives that take the following forms.

- 1) *Node cost*: Associates each node in the graph with a cost, depending on the node's weight  $V_i$

$$E_{\text{node}}(x) = \sum_{i \in V} V_i x_i. \quad (16)$$

- 2) *Edge cost*: Associates each edge in the graph with a cost, depending on the edge's weight  $E_{ij}$

$$E_{\text{edge}}(x) = \sum_{(i,j) \in E} E_{ij} x_{ij}. \quad (17)$$

With only one objective (for instance, the minimization of the distance between two points of a graph, which is an edge cost), the problem reduces to a single-objective shortest-path problem, which can be solved efficiently in polynomial time by Dijkstra's algorithm [56], which takes  $O(|E| + |V| \log |V|)$  steps. The presence of multiple objectives constitutes an MCOP.

Equations (16) and (17) allow us to compute the cost vector associated with a state  $x$ . However, of the  $2^{|V|+|E|-2}$  possible states, not all of them represent actual paths. A valid path is one that: 1) starts from a specified source node  $s$ , 2) ends at a specified destination node  $d$ , and 3) has no broken links or branches along the path from  $s$  to  $d$ . States that satisfy these criteria are called *feasible*, and *infeasible* otherwise. These three constraints can be enforced as quadratic penalty terms added to our cost function, so that infeasible solutions have higher total energies than feasible ones.

- 1) *Source constraint*: Penalizes paths that do not have exactly one edge connected to the source node

$$E_s(x) = -x_s^2 + \left( x_s - \sum_j x_{sj} \right)^2 \quad (18)$$

where  $s$  is the index of the source node, and the sum is over all edges that are connected to node  $s$ . This term has a minimum value of  $-1$ , which occurs for states where the source node is used ( $x_s = 1$ ), and there is only one way of leaving the source node (the bracketed term is equal to zero).

- 2) *Destination constraint*: Penalizes paths that do not have exactly one edge connected to the destination node

$$E_d(x) = -x_d^2 + \left( x_d - \sum_j x_{dj} \right)^2 \quad (19)$$

where  $d$  is the index of the destination node, and the sum is over all edges that are connected to node  $d$ . This

term has a minimum value of  $-1$ , similar to the source constraint.

- 3) *Path constraint*: Penalizes paths that do not have exactly two edges connected to intermediate nodes:  $E_{\text{path}} = \sum_{i \in V} E_i$ , such that for each intermediate node  $i$

$$E_i(x) = \left( 2x_i - \sum_j x_{ij} \right)^2 \quad (20)$$

where the sum is over all edges that are connected to node  $i$ . This has a minimum value of 0, which occurs for states where for all intermediate nodes used (for which  $x_i = 1$ ), the node degree is equal to 2.

A solution of the multiobjective shortest-path problem is then a path  $x$  that achieves the minimum of the penalties (18), (19), and (20) and is Pareto-optimal with respect to the node and edge costs.

As mentioned in the main text, the connectivity of the graph associated with the sum of the path constraints is precisely that of the problem's middle graph, which we review and exploit in Appendices D-B and D-C, respectively. Furthermore, the constraints (18) and (19) only incur a small number of linear and quadratic terms (dependent on the local connectivity of the source and destination nodes), which we safely ignore for conciseness.

#### APPENDIX D DETAILS ON GRAPH THEORY, THE MIDDLE GRAPH, AND THE PROOF OF (12)

Here, we list down notions in graph theory used in the resource estimation part of the main text and provide a proof for (12).

As mentioned in Section III, the correspondence  $G_H = M(G)$  implies that the solution of a shortest-path problem defined on a network  $G$  can be mapped to the ground state of a Hamiltonian/Ising model  $H$  with graph  $G_H = M(G)$ . Ignoring the quadratic terms resulting from the source and destination nodes, which only contribute a negligible constant overhead, (12) implies that the implementation of the cost unitary is efficient, depending only linearly on the maximum degree of the network  $G$ .

##### A. NOTIONS IN GRAPH THEORY

For an undirected graph  $G = (V(G), E(G))$ , the degree of a vertex  $u \in V(G)$ , denoted  $\deg(u)$ , is the number of edges that are incident to the vertex. The maximum degree, denoted  $\Delta_G$ , is the maximum of its vertices' degrees  $\Delta_G = \max_{u \in V(G)} \deg(u)$ . The degree sum formula relates them to the number of edges of  $G$

$$|E(G)| = \frac{1}{2} \sum_{u \in V(G)} \deg(u) \leq \frac{1}{2} |V(G)| \Delta_G. \quad (21)$$

The line graph of  $G = (V(G), E(G))$  is the graph  $L(G) = (E(G), E')$  with the edges of  $G$  as vertices, such that they are

adjacent if and only if their corresponding edges are incident on the same vertex. The maximum degrees of  $G$  and  $L(G)$  are related by

$$\Delta_{L(G)} \leq 2\Delta_G - 2. \quad (22)$$

The middle graph of  $G = (V(G), E(G))$  [39] is the graph  $M(G) = (V(G) \cup E(G), E')$ , with vertices  $u, v$  that are adjacent if either of the following statements applies:

- $u$  is a vertex in  $G$  and  $v$  is an edge in  $G$  incident to  $u$ ;
- $u$  and  $v$  are edges in  $G$  that are both incident on the same vertex.

The endline graph of  $G$ , denoted  $G^+$ , is defined as the graph obtained from  $G$  by adding to each of its nodes an end-vertex (i.e., an edge that is connected to a single node). The attachment of an end-vertex to each node immediately implies that

$$\Delta_{G^+} = \Delta_G + 1. \quad (23)$$

An edge coloring of a graph  $G$  is an assignment of colors to its edges, such that no two edges sharing a vertex can have the same color. The edge chromatic number/chromatic index of  $G$ , denoted  $\chi'(G)$ , is the minimum number of such colors needed. Vizing's theorem relates the edge chromatic number of  $G$  with the maximum degree for any graph

$$\Delta_G \leq \chi'(G) \leq \Delta_G + 1. \quad (24)$$

##### B. PROPERTIES OF THE MIDDLE GRAPH

We introduce the following properties on  $M(G)$  that will be used.

- The number of vertices and edges of  $M(G)$  are, respectively,

$$|V(M(G))| = |V(G)| + |E(G)| \leq \frac{1}{2} |V(G)| (2 + \Delta_G) \quad (25)$$

$$|E(M(G))| = \frac{1}{2} \sum_{u \in V(G)} \deg(u)^2 \leq \frac{1}{2} |V(G)| \Delta_G^2. \quad (26)$$

- The middle graph and the endline graph of the line graph of  $G$  are isomorphic [39]

$$M(G) \cong L(G)^+. \quad (27)$$

This immediately implies that their maximum degree is also equal

$$\Delta_{M(G)} = \Delta_{L(G)^+}. \quad (28)$$

##### C. PROOF OF $\chi'(M(G)) \leq 2\Delta_G$

The following chain of (in)equalities is true:

$$\chi'(M(G)) \leq \Delta_{M(G)} + 1 \quad (29)$$

$$= \Delta_{L(G)^+} + 1 \quad (30)$$

$$= \Delta_{L(G)} + 2 \quad (31)$$

**TABLE 1. Network parameters**

Parameters	Values
Network coverage area (Fully-connected)	1000m × 1000m
Network coverage area (Square lattice)	1000m × 2000m
Node delay, $\Delta$	1 ms
Path loss exponent, $\alpha$	2.7
Transmitting power, $P_T$	50 W
Carrier wavelength, $\lambda_c$	1.2 m
Carrier frequency, $f_c$	250 MHz
Modulation	QPSK
BER Gaussian noise mean, $\mu$	-90 dBm
BER Gaussian noise std. dev., $\sigma$	10 dBm

$$\leq 2\Delta_G \quad (32)$$

where applying Vizing's theorem yields the first inequality,  $M(G) \cong L(G)^+$  yields the second equality, (23) yields the third inequality, and (22) yields the final equality.

## APPENDIX E QUBO FORMULATION OF THE NETWORK ROUTING PROBLEM

Following Section III and Appendix C, this section details how the objectives of the multiobjective network routing problem (path loss, node delay, data rate, and bit-error) are formulated in our work, in particular how each objective can be modeled as either a node, (16), or edge, (17), cost to be minimized. Network parameters used to generate problem instances considered in numerical simulations and experiments on quantum computers are summarized in Table 1.

### A. NODE DELAY

The node delay quantifies the total time delay incurred by intermediate processing steps at relay stations along a signal's trajectory from sender to receiver and can be written as a linear objective function

$$E_{\Delta}(x) = \sum_{i \in V} \Delta_i x_i \quad (33)$$

where  $\Delta_i$  is the time delay incurred by node  $i$ . In our numerical simulations, we further assume that every node in the network incurs the same node delay of  $\Delta_i = \Delta = 1$  ms, so that the node delay is proportional to the number of nodes traversed by the signal.

### B. PATH LOSS

The path loss metric accounts for the loss in the amplitude of the electromagnetic signal during transmission between stations, which increases with distance traveled [57], [58]. Between nodes  $i$  and  $j$ , it is given by

$$L_{ij} = \frac{P_T}{P_R} = \left( \frac{4\pi d_{ij}}{\lambda_c} \right)^\alpha \quad (34)$$

where  $P_T = 50$  W is the power at the transmitting station,  $P_R$  the power at the receiving station,  $\alpha = 2.7$  the path loss exponent (typically between 2 in free space and 4 in a lossy environment),  $d_{ij}$  the distance between nodes  $i$  and  $j$ , and  $\lambda_c = 1.2$  m the carrier wavelength (resulting in the carrier

frequency of  $f_c = 250$  MHz, in the VHF regime). The resulting total path loss can be written as the sum of linear terms of the form

$$E_L(x) = \sum_{(i,j) \in E} L_{ij} x_{ij}. \quad (35)$$

### C. SYMMETRIZED BIT ERROR RATE

The bit error rate accounts for the probability of introducing an error into the message during the signal processing phase at a node. This error takes the shape of bit flip in the message and depends on the power of the signal and the noise at the receiving node

$$p_{ij} = \frac{1}{2} \left( 1 - \sqrt{\frac{R_{ij}}{R_{ij} + 1}} \right) \quad (36)$$

where  $R_{ij} = \frac{P_{R,ij}}{\log_2(M)P_{N,ij}}$ ,  $M = 4$  for QPSK modulation, and  $P_{R,ij} = P_T \left( \frac{\lambda_c}{4\pi d_{ij}} \right)^\alpha$ . We model the noise at each node  $P_{N,i}$  to follow a normal distribution with mean  $\mu = -90$  dBm and standard deviation  $\sigma = 10$  dBm, and additionally average (and thus symmetrize) the noise at neighboring nodes, i.e.,  $P_{N,ij} = \frac{1}{2}(P_{N,i} + P_{N,j})$ . This yields a linear cost function associated with the connecting edge of the form

$$E_{\text{BER}} = \sum_{(i,j) \in E} p_{ij} x_{ij} \quad (37)$$

so that the total error probability is the sum of the individual error probabilities for each selected edge. It is a first-order approximation of the exact total error (equivalent to assuming errors do not occur twice on the same bit).

### D. DATA RATE

The data rate metric quantifies the maximum data transfer rate that can be achieved across a transmission path. The Nyquist data rate for noiseless signals can be described by  $\Gamma_{ij} = 2B_{ij} \log_2(M)$ , with  $M = 4$  the modulation level for QPSK and  $B_{ij}$  the bandwidth of the signal transmitting between nodes  $i$  and  $j$ . We model the total available data rate between two nodes to be the base rate of 5 Mb/s minus the utilized data rate, such that it results in total available data rates that are uniformly distributed between 0 and 5 Mb/s, in steps of 50 kb/s.

The link with the minimum data rate determines the overall data rate along a transmission path, forming a bottleneck. In other words, we wish to penalize transmission paths with low minimum data rates. To express this metric in QUBO form as a linear sum of edge costs, we use a negative exponential function to weight the data rates  $\Gamma_{ij}$

$$E_{\Gamma} = \sum_{(i,j) \in E} x_{ij} \cdot e^{-\beta \Gamma_{ij}} \quad (38)$$

so that suboptimal links with low data rates are penalized heavily. Here,  $\beta$  is a graph-dependent coefficient that should be chosen such that for any choice of transmission path  $x$ , the dominant contribution to  $E_{\Gamma}$  corresponds

**TABLE 2. Node Properties and Parameters of the Fully Connected Network**

node $i$	coordinates [m]	noise level $P_{N,i}$ [dBm]	node delay $\Delta$ [ms]
0	(-181, 88)	-93.665	1.0
1	(-82, 674)	-94.609	1.0
2	(589, 84)	-81.811	1.0
3	(645, 664)	-85.622	1.0

**TABLE 3. Edge Properties and Parameters of the Fully Connected Network**

edge $(i, j)$	distance $d_{ij}$ [m]	data rate $\Gamma_{ij}$ [kb/s]	path loss $L_{ij}$ [dB]	$R_{ij}$ [dB]	bit error $p_{ij}$
(0, 1)	594.757	4250	102.448	35.643	6.82e-05
(0, 2)	770.565	2350	105.485	23.042	1.24e-03
(0, 3)	1007.517	1801	108.629	23.351	1.15e-03
(1, 2)	893.738	3850	107.223	21.355	1.82e-03
(1, 3)	727.298	1350	104.807	27.289	4.66e-04
(2, 3)	582.504	1900	102.204	25.087	7.73e-04

**TABLE 4. Node Properties and Parameters of the Square Lattice Network**

node $i$	coordinates [m]	noise level $P_{N,i}$ [dBm]	node delay $\Delta$ [ms]
0	(157, -93)	-85.058	1.0
1	(73, 316)	-92.388	1.0
2	(425, -102)	-91.186	1.0
3	(408, 469)	-93.737	1.0
4	(980, -57)	-84.012	1.0
5	(999, 635)	-91.733	1.0

**TABLE 5. Edge Properties and Parameters of the Square Lattice Network**

edge $(i, j)$	distance $d_{ij}$ [m]	data rate $\Gamma_{ij}$ [kb/s]	path loss $L_{ij}$ [dB]	$R_{ij}$ $p_{ij}$ [dB]	bit error
(0, 1)	417.987	300	98.312	32.999	1.25e-04
(0, 2)	268.148	3950	93.107	37.993	3.97e-05
(1, 3)	368.637	3600	96.839	40.151	2.41e-05
(2, 3)	571.589	2200	101.982	34.274	9.34e-05
(2, 4)	556.17	3250	101.661	28.579	3.46e-04
(3, 5)	614.043	650	102.822	33.777	1.05e-04
(4, 5)	693.11	3000	104.242	26.081	6.15e-04

to the minimum data rate along the path. It can be estimated by either considering the link in  $G$  with the smallest possible data rate or by determining the value of  $c(\beta) = \sum_{x \in \text{Paths}} (e^{-\beta R_{\min}^x} - \sum_{j \neq \min} e^{-\beta R_j^x})$  by sampling over different paths.  $R_{\min}^x$  is the minimum data rate of path  $x$ , and  $R_j^x$  are the data rates on all other edges in  $x$ .

## APPENDIX F SPECIFICATIONS OF PROBLEM INSTANCES IN QPU COMPUTATIONS

This section specifies the parameters used to generate problem instances solved in Section V with the IonQ quantum computer, with the four objectives defined in Appendix E. Nodes of the network are distributed within a (1000 m  $\times$  1000 m) area for the fully-connected networks and (1000 m  $\times$  2000 m) for the square lattice networks.

### A. FULLY CONNECTED NETWORK

The fully-connected network is a 4-node graph with five feasible solutions, of which 4 are Pareto-optimal. In total, 8 qubits are required to encode this problem.

### B. SQUARE LATTICE NETWORK

The square lattice network is a 6-node graph with 4 feasible solutions, of which 3 are Pareto-optimal. 11 qubits are required to encode this problem.

## APPENDIX G ADDITIONAL NUMERICAL RESULTS ON PARAMETER INITIALIZATION

In this section, we briefly verify the validity of the linear ramp parameter initialization scheme used in our numerical simulations, which is a heuristic choice that linearly ramps up  $\gamma$ 's and ramps down  $\beta$ 's, based on the analogy between QAOA and quantum annealing [27]. Fig. 6 displays numerical scaling results comparing this initialization (labeled as “linear ramp”) with random initializations for different values of  $p$ . The number of random initializations is scaled as  $100p$ , which in our numerical examples is sufficient to capture the global minima in most cases. The best result from the  $100p$  random initializations (labeled as “random max” to refer to the maximum approximation ratio achieved) can, therefore, be taken as the global optimum of the problem. For illustration, the problem instance is chosen to be a 5-node triangular lattice with 1 column and 3 rows, an 8-qubit problem.

We observe that the linear ramp initialization, indeed, greatly outperforms random sampling on average—both the approximation ratio and success probabilities approach the global optima in a monotonic fashion as  $p$  increases, compared to an average random initialization run, which saturates and even deteriorates for large  $p$ . We will, therefore, proceed to conduct all experiments presented in this work with the linear ramp initialization, unless stated otherwise.

## APPENDIX H SPECIFICATION OF PROBLEM INSTANCE OF FIG. 2

Table 6 provides the complete specification of the triangular lattice graph and the multiobjective cost function for the problem instance used in Fig. 2. That is, each of the four objective functions takes the form

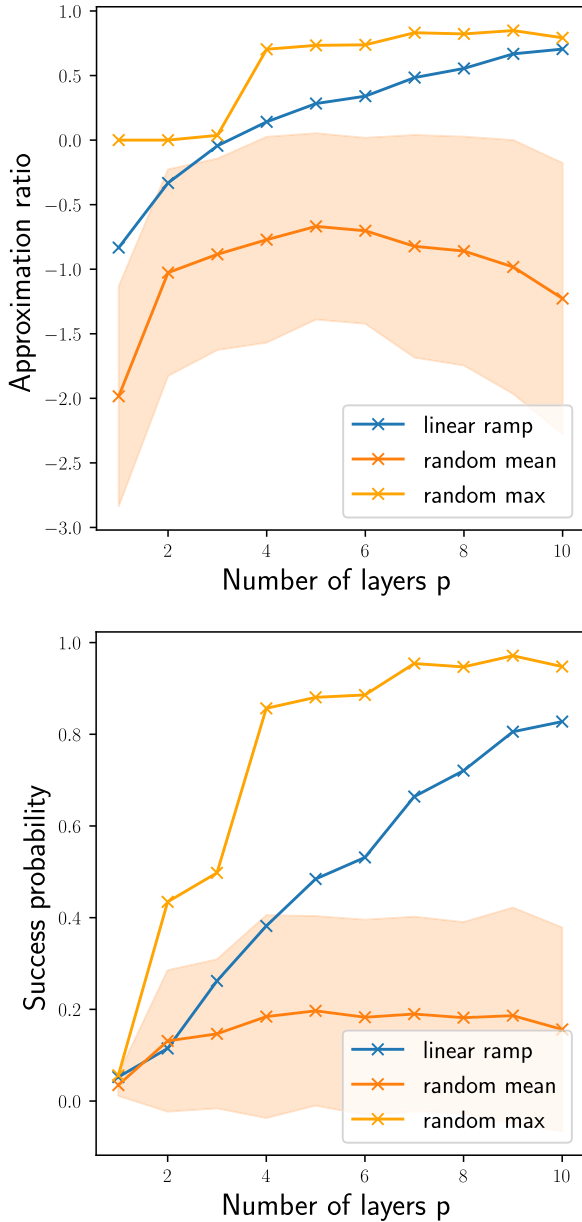
$$E_i^C(x) = \sum_i h_i x_i + \sum_{(i,j)} J_{ij} x_i x_j \quad (39)$$

where  $h_i$  corresponds to node weights and  $J_{ij}$  corresponds to edge weights with values as assigned in the table.

## APPENDIX I FRACTION OF FEASIBLE VERSUS INFEASIBLE CONFIGURATIONS

Here, we provide an intuitive explanation for the observation made in Section II-B that the ratio of the number of feasible solutions to the number of infeasible solutions is a decreasing function of the network size. We will illustrate this from the standpoint of the constraint described by (20), which enforces that any intermediate vertex used in a path through the network must have degree 2. All feasible solutions must, as a necessary condition, satisfy this particular constraint.





**FIGURE 6.** Scaling of the approximation ratio and success probability with the number of layers for different initializations, for one specific problem. To produce the random initializations curves, we perform  $100p$  random initializations and optimize with QAOA. The lighter region displays a standard deviation around the mean. The average result from  $100p$  random initializations is labeled as “random mean,” the best result from the  $100p$  random initializations is labeled as “random max” (to refer to the maximum approximation ratio achieved), and the (one) result from the linear ramp initialization is labeled as “linear ramp.”

Our argument will be based on how we can construct infeasible configurations that violate this constraint, starting from feasible ones.

Suppose we have a feasible configuration that uses intermediate vertex  $j$  (implying  $x_j = 1$ ), with the path entering vertex  $j$  through the edge  $(i, j)$ , and leaving through the edge  $(j, k)$  (implying that  $x_{ij} = x_{jk} = 1$ ). There are multiple infeasible configurations in the state space that can be obtained from this feasible configuration by simply flipping the

**TABLE 6.** Weights of Cost Terms of the Triangular Lattice Problem Instance of Fig. 2 for All Four Objectives, Labeled According to Their Corresponding Edges or Nodes

Node (i)	Obj. 1	Obj. 2	Obj. 3	Obj. 4
(0)	1.5	0	0.5	0
(1)	0.5	0	1	0
(2)	1.5	0	1.5	0
(3)	0.2	0	2	0
(4)	1	0	1	0
(5)	0.5	0	1.5	0
Edge (i,j)	Obj. 1	Obj. 2	Obj. 3	Obj. 4
(0,1)	0	0.5	0	2.5
(0,3)	0	1.5	0	0.1
(1,2)	0	1.6	0	0.5
(1,4)	0	1	0	1.5
(1,3)	0	3	0	0.5
(2,5)	0	1.6	0	0.6
(2,4)	0	2	0	0.5
(3,4)	0	1	0	0.5
(4,5)	0	0.6	0	1.6

Six node weights and nine edge weights are present. The problem can then be specified via (39).

value of one variable. For instance, if we set  $x_j = 0$  in the example before, and leave all other variables unchanged, we have an infeasible configuration that describes a path using the edges  $(i, j)$  and  $(j, k)$ , but which never enters the vertex  $j$ . Similarly, if we instead set  $x_{ij} = 0$ , and leave everything else unchanged, we would have a path that leaves vertex  $j$ , but never enters it. If there are more than just two possible paths in or out of vertex  $j$ , we could also have solutions where the degree of the vertex is larger than two, which would also represent an infeasible configuration.

The examples just discussed would violate the constraint of (20), and hence, they represent infeasible configurations. As we see, for every feasible configuration, we can minimally “perturb” it by flipping the value of individual variables to obtain infeasible configurations. However, we could build “even worse” infeasible configurations by introducing such perturbations in multiple locations simultaneously, not only around a single vertex or edge. The number of combinations of infeasible configurations we can generate in this way grows exponentially in the number of variables (i.e., in the number of vertices and edges). We, therefore, conclude that the fraction of feasible to infeasible configurations shrinks exponentially in the problem size.

## ACKNOWLEDGMENT

The authors acknowledge the support of Amazon Web Services for the computations on quantum computers executed using Amazon Braket. The authors also would like to thank Adrien Vandembroucq, Tommaso Demarie, Too Huseh Tien, and Leow Kian Siang for valuable discussions and feedback on this article, Chui Chee Cheon and Soh Wei Siang Jonathan for their inputs on wireless technology, and Leonardo Disilvestro for his help in the experiments on quantum computers.

## REFERENCES

- [1] Y. Collette and P. Siarry, *Multiojective Optimization: Principles and Case Studies*. Berlin, Germany: Springer, 2004, doi: [10.1007/978-3-662-08883-8](https://doi.org/10.1007/978-3-662-08883-8).
- [2] S. P. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004, doi: [10.1017/CBO9780511804441](https://doi.org/10.1017/CBO9780511804441).
- [3] M. Emmerich and A. H. Deutz, "A tutorial on multiobjective optimization: Fundamentals and evolutionary methods," *Natural Comput.*, vol. 17, no. 3, pp. 585–609, 2018, doi: [10.1007/s11047-018-9685-y](https://doi.org/10.1007/s11047-018-9685-y).
- [4] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," *IEEE Commun. Mag.*, vol. 40, no. 8, pp. 102–114, Aug. 2002, doi: [10.1109/MCOM.2002.1024422](https://doi.org/10.1109/MCOM.2002.1024422).
- [5] N. A. Pantazis and D. D. Vergados, "A survey on power control issues in wireless sensor networks," *IEEE Commun. Surveys Tut.*, vol. 9, no. 4, pp. 86–107, Oct.–Dec. 2007, doi: [10.1109/COMST.2007.4444752](https://doi.org/10.1109/COMST.2007.4444752).
- [6] J. Sarangapani, *Wireless Ad Hoc and Sensor Networks: Protocols, Performance, and Control*. Boca Raton, FL, USA: CRC, 2017, doi: [10.1201/9781420015317](https://doi.org/10.1201/9781420015317).
- [7] H. Karl and A. Willig, *Protocols and Architectures for Wireless Sensor Networks*. Hoboken, NJ, USA: Wiley, 2007, doi: [10.1002/0470095121](https://doi.org/10.1002/0470095121).
- [8] R. V. Kulkarni, A. Förster, and G. K. Venayagamoorthy, "Computational intelligence in wireless sensor networks: A survey," *IEEE Commun. Surveys Tut.*, vol. 13, no. 1, pp. 68–96, Jan.–Mar. 2011, doi: [10.1109/SURV.2011.040310.00002](https://doi.org/10.1109/SURV.2011.040310.00002).
- [9] C. Bockelmann et al., "Towards massive connectivity support for scalable MMTC communications in 5G networks," *IEEE Access*, vol. 6, pp. 28969–28992, 2018, doi: [10.1109/ACCESS.2018.2837382](https://doi.org/10.1109/ACCESS.2018.2837382).
- [10] S. K. Sharma and X. Wang, "Toward massive machine type communications in ultra-dense cellular IoT networks: Current issues and machine learning-assisted solutions," *IEEE Commun. Surveys Tut.*, vol. 22, no. 1, pp. 426–471, Jan.–Mar. 2020, doi: [10.1109/COMST.2019.2916177](https://doi.org/10.1109/COMST.2019.2916177).
- [11] T. Back, *Evolutionary Algorithms in Theory and Practice: Evolution Strategies, Evolutionary Programming, Genetic Algorithms*. London, U.K.: Oxford Univ. Press, 1996, doi: [10.1093/oso/9780195099713.001.0001](https://doi.org/10.1093/oso/9780195099713.001.0001).
- [12] C. A. C. Coello et al. *Evolutionary Algorithms for Solving Multi-Objective Problems*, vol. 5. Berlin, Germany: Springer, 2007, doi: [10.1007/978-0-387-36797-2](https://doi.org/10.1007/978-0-387-36797-2).
- [13] J. Horn, N. Nafpliotis, and D. E. Goldberg, "A niched Pareto genetic algorithm for multiobjective optimization," in *Proc. 1st IEEE Conf. Evol. Comput., IEEE World Congr. Comput. Intell.*, 1994, pp. 82–87, doi: [10.1109/ICEC.1994.350037](https://doi.org/10.1109/ICEC.1994.350037).
- [14] J. D. Knowles and D. W. Corne, "Approximating the nondominated front using the Pareto archived evolution strategy," *Evol. Comput.*, vol. 8, no. 2, pp. 149–172, Jun. 2000, doi: [10.1162/106365600568167](https://doi.org/10.1162/106365600568167).
- [15] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Trans. Evol. Comput.*, vol. 6, no. 2, pp. 182–197, Apr. 2002, doi: [10.1109/4235.996017](https://doi.org/10.1109/4235.996017).
- [16] M. Dorigo, M. Birattari, and T. Stutzle, "Ant colony optimization," *IEEE Comput. Intell. Mag.*, vol. 1, no. 4, pp. 28–39, Nov. 2006, doi: [10.1109/MCI.2006.329691](https://doi.org/10.1109/MCI.2006.329691).
- [17] J. Kennedy, "Swarm intelligence," in *Handbook of Nature-Inspired and Innovative Computing*. Berlin, Germany: Springer, 2006, pp. 187–219, doi: [10.1007/0-387-27705-6\\_6](https://doi.org/10.1007/0-387-27705-6_6).
- [18] J. Tilly et al., "The variational quantum eigensolver: A review of methods and best practices," *Phys. Rep.*, vol. 986, pp. 1–128, 2022, doi: [10.1016/j.physrep.2022.08.003](https://doi.org/10.1016/j.physrep.2022.08.003).
- [19] D. J. Egger et al., "Quantum computing for finance: State-of-the-art and future prospects," *IEEE Trans. Quantum Eng.*, vol. 1, 2020, Art. no. 3101724, doi: [10.1109/TQE.2020.3030314](https://doi.org/10.1109/TQE.2020.3030314).
- [20] E. Farhi, J. Goldstone, and S. Gutmann, "A quantum approximate optimization algorithm," 2014, *arXiv:1411.4028*, doi: [10.48550/arXiv.1411.4028](https://doi.org/10.48550/arXiv.1411.4028).
- [21] E. Pelofske, A. Bärttschi, L. Cincio, J. Golden, and S. Eidenbenz, "Scaling whole-chip QAOA for higher-order Ising spin glass models on heavy-hex graphs," 2023, *arXiv:2312.00997*, doi: [10.48550/arXiv.2312.00997](https://doi.org/10.48550/arXiv.2312.00997).
- [22] J. Weidenfeller et al., "Scaling of the quantum approximate optimization algorithm on superconducting qubit based hardware," *Quantum*, vol. 6, 2022, Art. no. 870, doi: [10.22331/q-2022-12-07-870](https://doi.org/10.22331/q-2022-12-07-870).
- [23] S. V. Barron et al., "Provable bounds for noise-free expectation values computed from noisy samples," 2023, *arXiv:2312.00733*, doi: [10.48550/arXiv.2312.00733](https://doi.org/10.48550/arXiv.2312.00733).
- [24] E. Pelofske, A. Bärttschi, and S. Eidenbenz, "Quantum annealing vs. QAOA: 127 qubit higher-order Ising problems on NISQ computers," in *Proc. Int. Conf. High Perform. Comput.*, 2023, pp. 240–258, doi: [10.1007/978-3-031-32041-5\\_13](https://doi.org/10.1007/978-3-031-32041-5_13).
- [25] V. Sharma et al., "OpenQAOA—An SDK for QAOA," 2022, *arXiv:2210.08695*, doi: [10.48550/arXiv.2210.08695](https://doi.org/10.48550/arXiv.2210.08695).
- [26] A. Lucas, "Ising formulations of many np problems," *Front. Phys.*, vol. 2, 2014, Art. no. 5, doi: [10.3389/fphys.2014.00005](https://doi.org/10.3389/fphys.2014.00005).
- [27] L. Zhou, S.-T. Wang, S. Choi, H. Pichler, and M. D. Lukin, "Quantum approximate optimization algorithm: Performance, mechanism, and implementation on near-term devices," *Phys. Rev. X*, vol. 10, no. 2, 2020, Art. no. 021067, doi: [10.1103/PhysRevX.10.021067](https://doi.org/10.1103/PhysRevX.10.021067).
- [28] S. Bravyi, A. Kliesch, R. Koenig, and E. Tang, "Obstacles to variational quantum optimization from symmetry protection," *Phys. Rev. Lett.*, vol. 125, 2020, Art. no. 260505, doi: [10.1103/PhysRevLett.125.260505](https://doi.org/10.1103/PhysRevLett.125.260505).
- [29] S. Bravyi, A. Kliesch, R. Koenig, and E. Tang, "Hybrid quantum-classical algorithms for approximate graph coloring," *Quantum*, vol. 6, 2022, Art. no. 678, doi: [10.22331/q-2022-03-30-678](https://doi.org/10.22331/q-2022-03-30-678).
- [30] Z. Wang, N. C. Rubin, J. M. Dominy, and E. G. Rieffel, "XY mixers: Analytical and numerical results for the quantum alternating operator ansatz," *Phys. Rev. A*, vol. 101, Jan. 2020, Art. no. 012320, doi: [10.1103/PhysRevA.101.012320](https://doi.org/10.1103/PhysRevA.101.012320).
- [31] F. G. Fuchs, K. O. Lye, H. M. Nilsen, A. J. Stasik, and G. Sartor, "Constraint preserving mixers for the quantum approximate optimization algorithm," *Algorithms*, vol. 15, Jun. 2022, Art. no. 202, doi: [10.3390/a15060202](https://doi.org/10.3390/a15060202).
- [32] Y. Zhang, R. Zhang, and A. C. Potter, "QED driven QAOA for network-flow optimization," *Quantum*, vol. 5, 2021, Art. no. 510, doi: [10.22331/q-2021-07-27-510](https://doi.org/10.22331/q-2021-07-27-510).
- [33] T. Krauss and J. McCollum, "Solving the network shortest path problem on a quantum annealer," *IEEE Trans. Quantum Eng.*, vol. 1, 2020, Art. no. 3101512, doi: [10.1109/TQE.2020.3021921](https://doi.org/10.1109/TQE.2020.3021921).
- [34] P. Hauke, G. Mantiotti, and P. Facchioli, "Dominant reaction pathways by quantum computing," *Phys. Rev. Lett.*, vol. 126, no. 2, 2021, Art. no. 028104, doi: [10.1103/PhysRevLett.126.028104](https://doi.org/10.1103/PhysRevLett.126.028104).
- [35] P. Serafini, "Some considerations about computational complexity for multi objective combinatorial problems," in *Recent Advances and Historical Development of Vector Optimization*. Berlin, Germany: Springer, 1987, pp. 222–232, doi: [10.1007/978-3-642-46618-2\\_15](https://doi.org/10.1007/978-3-642-46618-2_15).
- [36] P. Jacquet, P. Muhlethaler, T. Clausen, A. Laouiti, A. Qayyum, and L. Viennot, "Optimized link state routing protocol for ad hoc networks," in *Proc. IEEE Int. Multi Topic Conf.*, 2001, pp. 62–68, doi: [10.1109/INMIC.2001.995315](https://doi.org/10.1109/INMIC.2001.995315).
- [37] C. Bauckhage, E. Brito, K. Cvejoksi, C. Ojeda, J. Schucker, and R. Sifa, "Towards shortest paths via adiabatic quantum computing," in *Proc. 14th Int. Workshop Mining Learn. Graphs*, 2018. [Online]. Available: [https://www.mlgworkshop.org/2018/papers/MLG2018\\_paper\\_7.pdf](https://www.mlgworkshop.org/2018/papers/MLG2018_paper_7.pdf)
- [38] D. Alanis, P. Botsinis, S. X. Ng, and L. Hanzo, "Quantum-assisted routing optimization for self-organizing networks," *IEEE Access*, vol. 2, pp. 614–632, 2014, doi: [10.1109/ACCESS.2014.2327596](https://doi.org/10.1109/ACCESS.2014.2327596).
- [39] T. Hamada and I. Yoshimura, "Traversability and connectivity of the middle graph of a graph," *Discrete Math.*, vol. 14, no. 3, pp. 247–255, 1976, doi: [10.1016/0012-365X\(76\)90037-6](https://doi.org/10.1016/0012-365X(76)90037-6).
- [40] S. Wang et al., "Noise-induced barren plateaus in variational quantum algorithms," *Nature Commun.*, vol. 12, no. 1, 2021, Art. no. 6961, doi: [10.1038/s41467-021-27045-6](https://doi.org/10.1038/s41467-021-27045-6).
- [41] J. R. McClean, S. Boixo, V. N. Smelyanskiy, R. Babbush, and H. Neven, "Barren plateaus in quantum neural network training landscapes," *Nature Commun.*, vol. 9, no. 1, 2018, Art. no. 4812, doi: [10.1038/s41467-018-07090-4](https://doi.org/10.1038/s41467-018-07090-4).
- [42] M. Cerezo, A. Sone, T. Volkoff, L. Cincio, and P. J. Coles, "Cost function dependent barren plateaus in shallow parametrized quantum circuits," *Nature Commun.*, vol. 12, no. 1, 2021, Art. no. 1791, doi: [10.1038/s41467-021-21728-w](https://doi.org/10.1038/s41467-021-21728-w).
- [43] S. Hochreiter, "The vanishing gradient problem during learning recurrent neural nets and problem solutions," *Int. J. Uncertainty, Fuzziness, Knowl.-Based Syst.*, vol. 6, no. 2, pp. 107–116, 1998, doi: [10.1142/S0218488598000094](https://doi.org/10.1142/S0218488598000094).

- [44] D. Kielpinski, C. Monroe, and D. J. Wineland, "Architecture for a large-scale ion-trap quantum computer," *Nature*, vol. 417, no. 6890, pp. 709–711, 2002, doi: [10.1038/nature00784](https://doi.org/10.1038/nature00784).
- [45] J. Weidenfeller et al., "Scaling of the quantum approximate optimization algorithm on superconducting qubit based hardware," *Quantum*, vol. 6, Dec. 2022, Art. no. 870, doi: [10.22331/q-2022-12-07-870](https://doi.org/10.22331/q-2022-12-07-870).
- [46] I. D. Kivlichan et al., "Quantum simulation of electronic structure with linear depth and connectivity," *Phys. Rev. Lett.*, vol. 120, no. 11, 2018, Art. no. 110501, doi: [10.1103/PhysRevLett.120.110501](https://doi.org/10.1103/PhysRevLett.120.110501).
- [47] B. O’Gorman, W. J. Huggins, E. G. Rieffel, and K. B. Whaley, "Generalized swap networks for near-term quantum computing," 2019, *arXiv:1905.05118*, doi: [10.48550/arXiv.1905.05118](https://doi.org/10.48550/arXiv.1905.05118).
- [48] A. Peruzzo et al., "A variational eigenvalue solver on a photonic quantum processor," *Nature Commun.*, vol. 5, no. 1, 2014, Art. no. 4213, doi: [10.1038/ncomms5213](https://doi.org/10.1038/ncomms5213).
- [49] K. Wright et al., "Benchmarking an 11-qubit quantum computer," *Nature Commun.*, vol. 10, no. 1, 2019, Art. no. 5464, doi: [10.1038/s41467-019-13534-2](https://doi.org/10.1038/s41467-019-13534-2).
- [50] M. J. D. Powell, *A Direct Search Optimization Method That Models the Objective and Constraint Functions by Linear Interpolation*. Dordrecht, The Netherlands: Springer, 1994, pp. 51–67, doi: [10.1007/978-94-015-8330-5\\_4](https://doi.org/10.1007/978-94-015-8330-5_4).
- [51] P. K. Barkoutsos, G. Nannicini, A. Robert, I. Tavernelli, and S. Woerner, "Improving variational quantum optimization using CVAR," *Quantum*, vol. 4, 2020, Art. no. 256, doi: [10.22331/q-2020-04-20-256](https://doi.org/10.22331/q-2020-04-20-256).
- [52] Z. Fan, J. Xu, G. Shu, X. Ding, H. Lian, and Z. Shan, "Solving the shortest path problem with QAOA," in *Spin*, vol. 13. Singapore: World Scientific, 2023, Art. no. 2350002, doi: [10.1142/S2010324723500029](https://doi.org/10.1142/S2010324723500029).
- [53] B. Barán and M. Villagra, "Multiobjective optimization in a quantum adiabatic computer," *Electron. Notes Theor. Comput. Sci.*, vol. 329, pp. 27–38, 2016, doi: [10.1016/j.entcs.2016.12.003](https://doi.org/10.1016/j.entcs.2016.12.003).
- [54] H. Urgelles, P. Picazo-Martinez, D. Garcia-Roger, and J. F. Monserrat, "Multi-objective routing optimization for 6G communication networks using a quantum approximate optimization algorithm," *Sensors*, vol. 22, no. 19, 2022, Art. no. 7570, doi: [10.3390/s22197570](https://doi.org/10.3390/s22197570).
- [55] L. Ekstrom, H. Wang, and S. Schmitt, "Variational quantum multi-objective optimization," 2023, *arXiv:2312.14151*, doi: [10.48550/arXiv.2312.14151](https://doi.org/10.48550/arXiv.2312.14151).
- [56] E. Dijkstra, "A note on two problems in connexion with graphs," *Numerische Mathematik*, vol. 1, pp. 269–271, 1959, doi: [10.1007/BF01386390](https://doi.org/10.1007/BF01386390).
- [57] R. Steele and L. Hanzo, *Mobile Radio Communications: Second and Third Generation Cellular and WATM Systems*, 2nd ed. Hoboken, NJ, USA: Wiley, 1999. [Online]. Available: <https://www.wiley.com/en-us/Mobile+Radio+Communications,+2nd+Edition-p-9780471978060>
- [58] H. Fernández, L. Rubio, V. M. Rodrigo-Peñarrocha, and J. Reig, "Path loss characterization for vehicular communications at 700 MHz and 5.9 GHz under LOS and NLOS conditions," *IEEE Antennas Wireless Propag. Lett.*, vol. 13, pp. 931–934, 2014, doi: [10.1109/LAWP.2014.2322261](https://doi.org/10.1109/LAWP.2014.2322261).