


Hybrid Quantum-Classical Heuristic to Solve Large-Scale Integer Linear Programs

Marika Svensson^{1,2,*}, Martin Andersson,¹ Mattias Grönkvist,¹ Pontus Vikstål,³ Devdatt Dubhashi,² Giulia Ferrini,³ and Göran Johansson³

¹Jeppesen, Gothenburg 411 03, Sweden

²Department of Computer Science and Engineering, Chalmers University of Technology, Gothenburg 412 96, Sweden

³Wallenberg Centre for Quantum Technology, Department of Microtechnology and Nanoscience, Chalmers University of Technology, Gothenburg 412 96, Sweden

 (Received 13 July 2021; revised 23 September 2022; accepted 11 July 2023; published 26 September 2023)

We present a method that integrates any quantum algorithm capable of finding solutions to integer linear programs into the branch-and-price algorithm, which is regularly used to solve large-scale integer linear programs with a specific structure. The role of the quantum algorithm is to find integer solutions to subproblems appearing in branch-and-price. Obtaining optimal or near-optimal integer solutions to these subproblems can increase the quality of solutions and reduce the depth and branching factor of the branch-and-price algorithm and hence reduce the overall running time. We investigate the viability of the approach by considering the tail assignment problem and the quantum approximate optimization algorithm (QAOA). Here, the master problem is the optimization problem set partitioning or its decision version exact cover and can be expressed as finding the ground state of an Ising spin glass Hamiltonian. For exact cover, our numerical results indicate that the required algorithm depth decreases with the number of feasible solutions for a given success probability of finding a feasible solution. For set partitioning, on the other hand, we find that for a given success probability of finding the optimal solution, the required algorithm depth can increase with the number of feasible solutions if the Hamiltonian is balanced poorly, which in the worst case is exponential in the problem size. We therefore address the significance of properly balancing the objective and constraint parts of the Hamiltonian. We empirically find that the approach is viable with QAOA if polynomial algorithm depth can be realized on quantum devices.

DOI: [10.1103/PhysRevApplied.20.034062](https://doi.org/10.1103/PhysRevApplied.20.034062)

I. INTRODUCTION

Large-scale integer linear programs (ILPs) appear in the real world frequently as they model problems, such as planning, scheduling, and resource allocation. These problems are characterized by their large size, a linear cost function, affine inequality and/or equality constraints, as well as variables required to be integers.

Airline planning problems such as crew rostering, crew pairing [1,2], and tail assignment [3,4] fall into this category. These problems are made more difficult by very complex rules and regulations imposed by aviation authorities, airlines, and unions [1]. These rules can even be hard to express in mathematical optimization models and the

models can furthermore have objective functions that are nonlinear in some optimization formulations [5,6]. One way to address these difficulties is to formulate the optimization problem with a very large number of variables and to separate the problem into a generation problem and a selection problem. With this formulation, standard approaches such as branch-and-bound or branch-and-cut [7] cannot be used directly to solve these problems due to their large size, where even enumerating the legal decision variables can require exponential time and space, see Sec. A 2 for a more detailed explanation. Instead, by starting with an empty set of variables, the generation problem is responsible for generating alternative variables (aircraft routes in the tail assignment problem) to the selection problem (an ILP for the tail assignment problem). The task of the selection problem is to find the subset of the generated variables that in the most cost-effective way satisfy all the constraints in the ILP (in the most basic tail assignment problem, this corresponds to having each flight in the schedule covered by exactly one aircraft). This process is generalized in the branch-and-price algorithm [6], which

*marika.svensson@jeppesen.com

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Funded by [Bibsam](https://www.bibsam.com/).

combines branch-and-bound [8] and column generation [9,10] and has generally been successful for large-scale ILPs with this type of structure. The benefit of separating the problem is that the complex rules affect only the generation problem, whereas the selection problem is often a pure set cover or set partitioning problem.

In the column generation algorithm, the generation and selection problems are solved iteratively until optimal conditions hold. In this context, the selection problem is called the restricted master problem (RMP) and the generation problem is called the pricing problem (PP). The RMP, which contains only a subset of the decision variables of the original problem, is solved as a linear program (LP). Column generation is generally insufficient to solve the original ILP since the solution is most likely fractional. To remedy this, column generation is combined with branch-and-bound for finding the integer solution. For readers unfamiliar with branch-and-price, details are given in Appendix A.

With the results for factoring with Shor’s algorithm [11] and unstructured database search with Grover’s algorithm [12], providing subexponential and quadratic speedup, respectively, it is natural to ask if quantum algorithms also can provide speedup for ILPs even though superpolynomial speedup for these problems is not expected. The adiabatic quantum algorithm [13] and quantum annealing [14] have subsequently been proposed. Other quantum algorithms for combinatorial optimization problems [15,16] such as Grover’s adaptive search algorithm [17] have also been proposed. In recent years, much interest has been given to the quantum approximate optimization algorithm (QAOA) [18] for solving combinatorial optimization problems, as it may be a suitable algorithm to run on near-term gate-based quantum computers and to demonstrate quantum advantage or quantum supremacy [19].

Experiments performed in Ref. [20] have reported to demonstrate quantum supremacy for a problem that is not related to optimization. Such devices can be classified as noisy intermediate-scale quantum (NISQ) computers, where qubits are controlled imperfectly and quantum error correction is generally not considered [21]. Moreover, QAOA was demonstrated in Ref. [22] for the Sherrington-Kirkpatrick model and maxcut, where experiments agree well with simulations. Such results further motivate investigating QAOA for ILPs and distinctly large-scale ILPs.

Here, we address the open question of whether quantum algorithms can provide any advantage for large-scale ILPs, where we stress that these problems can require exponential time and space even to *generate* the full ILP or the continuous relaxation counterpart. The large number of decision variables therefore in practice rules out a direct application of any quantum algorithm capable of solving an ILP, as well as standard classical algorithms for ILPs and the continuous relaxation. Building instead on the above mentioned branch-and-price algorithm, we propose

to augment it by using a quantum co-processor to find optimal or near-optimal solutions to RMP instances. For real-world large-scale ILPs, optimality is often intractable and finding good bounds can be hard. The exact definition of a near-optimal or “good-enough” solution will thus be a balance between running time and the quality of the solution. We can then ensure that the number of decision variables does not exceed the capabilities of the quantum processor when we generate the RMP. Here, we also would like to note that the proposed hybrid algorithm is a heuristic and the performance for real-world ILP problems cannot be properly evaluated until the hardware reaches between 10^3 to 10^4 qubits, corresponding to the same amount of decision variables in the RMP problem. At that point, we argue that this method has the potential to reduce the time to solution and improve solution quality. The detailed description of where a NISQ co-processor could first be used to test for possible speedup in finding near-optimal solutions to ILPs, is one of the main results of this paper. Furthermore, we investigate the method numerically by considering QAOA and the real-world problem tail assignment that generalizes set partitioning and its decision version exact cover, which are NP-hard and NP-complete problems [23]. Here, we are naturally limited to instances with up to 20 decision variables, but we explore these instances to learn how to balance the cost and constraint parts of the Hamiltonian to achieve a large probability of finding the optimal solution, when there are multiple feasible solutions.

The results have been obtained by simulating ideal QAOA circuits applied to instances with one or more feasible solutions, extracted from a heuristic branch-and-price algorithm [3]. The numerical results expand on Ref. [24], where QAOA was applied to instances with a single feasible solution and mapped as an exact cover problem (the decision version of the optimization problem set partitioning), also extracted from tail assignment. The exact cover version of this problem has also been investigated experimentally on a quantum processor [25] and on a quantum annealer [26]. The current paper thus gives a broader ILP context and background for these works as well as a first extension to cases with multiple feasible solutions. The paper is organized as follows. In Sec. II we introduce the tail assignment problem. We present the method for integrating a quantum algorithm with branch-and-price in Sec. III. In Sec. IV we review QAOA and the chosen mapping of exact cover and set partitioning to an Ising spin glass Hamiltonian. In Sec. V the extracted RMP instances are presented. We present and motivate the chosen optimization strategy for studying larger algorithm depths in Sec. VI. Results are given in Sec. VII first for exact cover and second for set partitioning. Last, in Sec. VIII we summarize the findings and discuss interesting open questions that are beyond the scope of this work.

II. TAIL ASSIGNMENT: AN EXAMPLE OF A REAL-WORLD LARGE-SCALE INTEGER LINEAR PROGRAM

Airlines regularly face several large NP-hard planning problems such as fleet assignment, crew pairing, crew rostering, and tail assignment in the planning process [3,27]. For tail assignment, the task is to determine, given a set of flights and a set of aircraft, what flights are operated by which individual aircraft and what order under the constraint that each flight is flown exactly once such that some objective is optimized. Operational constraints, such as minimum connection times, airport curfews, maintenance, and preassigned activities must also be respected, and can be considered part of the input to tail assignment. A set of flights operated by an aircraft is referred to as a route, where the operational constraints distinguish legal routes from illegal routes. This means that a solution consists of a set of legal routes that cover all flights exactly once in the most cost-effective way. As an example, an airline can encounter problems with one thousand flights per day with hundreds of aircraft, where the aircraft are of ten different types [3]. In the worst case, this means that the number of possible routes to determine if they are legal or illegal would be $2^{|F|}$, where F is the set of flights. By considering restrictions such as the arrival time must be less than the departure time of two flights following each other in a route the combinatorial explosion can be decreased. However, typically the number of legal routes will be very large and too large to solve without separating the problem into a selection problem and a generation problem.

Tail assignment can thus be classified as a large-scale ILP, where we refer the readers to Refs. [7,28] for a comprehensive view of established algorithms for solving ILPs and to Refs. [6,10,29,30] for large-scale ILPs. The classical algorithm we consider here used to find optimal or near-optimal solutions to tail assignment in Ref. [3] is a heuristic branch-and-price. The heuristic branch-and-price can be understood as the branch-and-price algorithm where the branching step is replaced with a fixing step that is better suited for tail assignment by diving into a branch of the full search tree.

For consistency, we give the details of the algorithms branch-and-bound, column generation, branch-and-price, and the heuristic branch-and-price in Appendix A.

A. The set partitioning problem and the exact cover problem

We define a simple path-based model of tail assignment as a set partitioning problem

$$\text{minimize } \sum_{r \in R} c_r x_r, \quad (1)$$

$$\text{subject to } \sum_{r \in R} a_{fr} x_r = 1 \quad \forall f \in F, \quad (2)$$

$$x_r \in \{0, 1\} \quad \forall r \in R, \quad (3)$$

where F is the set of flights and R is the set of legal aircraft routes. In the linear objective function, Eq. (1), $c_r \in \mathbb{Z}$ corresponds to the cost of using route r . The entries $a_{fr} \in \{0, 1\}$ are elements of a constraint matrix A indicating if flight f is part of route r . A column in the constraint matrix is therefore a route. Furthermore, Eq. (2) enforces the requirement that the set of routes in a solution should contain flight f exactly once. Finally, the decision variables $x_r \quad \forall r \in R$ indicate which routes are used. The tail assignment problem can, in practice, also be described by the decision problem exact cover, for cases where the objective is to find any feasible solution and not the optimal solution necessarily. The exact cover problem can be modeled as an ILP where the objective function in Eq. (1) is ignored and set to 0 for any assignment of the decision variables.

We now define the set S_{feasible} to be the set of feasible solutions to the set partitioning problem and the exact cover problem as

$$S_{\text{feasible}} = \left\{ \vec{x} \in \{0, 1\}^{|R|} : \sum_{r \in R} a_{fr} x_r = 1 \quad \forall f \in F \right\}. \quad (4)$$

If we consider a linear system of equations modulus 2

$$A\vec{x} = \vec{b} \pmod{2}, \quad (5)$$

where the matrix A is of dimension $|F| \times |R|$, \vec{x} is a column vector with $|R|$ unknown variables and \vec{b} is a column vector with $|F|$ entries. The elements of A , \vec{b} , and \vec{x} are either 0 or 1, respectively. The system of equations has

$$2^{|R| - \text{rank}(A)} \quad (6)$$

number of solutions as long as the linear system of equations in Eq. (5) has at least one solution [31]. For set partitioning $2^{|R| - \text{rank}(A)}$ constitutes an upper bound on the number of feasible solutions $|S_{\text{feasible}}|$ [32], since any feasible solution to set partitioning is also a solution modulus 2 to the system of equations in Eq. (5) where all entries in \vec{b} is set to one. It is therefore possible that the number of feasible solutions is significantly smaller than the upper bound. Furthermore, as the counting version of exact cover and set partitioning is #P-complete [33], obtaining the actual number of feasible solutions for typical instances for tail assignment becomes intractable.

We have investigated the number of feasible solutions for generated RMP instances of tail assignment with CPLEX [34]. We find that the number of feasible solutions for two sets of generated instances can be

larger than 5×10^6 for problems with 700–800 decision variables. We can therefore not rule out that the number of feasible solutions can be very large in practice, and the consequence is to investigate if a large feasible set is a limiting factor in the performance for QAOA.

III. INTEGRATING A QUANTUM ALGORITHM WITH BRANCH-AND-PRICE

In this section we present the method where the branch-and-price algorithm is augmented by integrating any quantum algorithm capable of finding optimal or near-optimal integer solutions to RMP instances. The integrated branch-and-price algorithm is depicted in Fig. 1 where branch-and-price is distinguished with green and blue colored boxes, and dotted and dashed borders. The green boxes with dotted borders highlight the column generation algorithm, and the blue boxes with dashed borders are distinctive for the branch-and-bound algorithm. The red boxes with solid borders give the integration of a quantum algorithm. This hybrid algorithm is one of the main results of this paper.

The integrated method utilizes a quantum algorithm for each column generation iteration if the RMP is deemed promising. We remind the reader that since routes are generated dynamically by the column generation algorithm each iteration corresponds to another ILP instance, which means that each iteration provides a possibility to find an alternative integer solution to the problem via a quantum (or classical) algorithm. For example, we might want to avoid using a quantum algorithm in the beginning of the column generation process as it will, in general, be more likely to find good integer solutions in later iterations. However, determining how often to use a quantum algorithm will be a trade-off that depends on if the RMP instance is expected to contain integer solutions with reasonable quality, the runtime of the algorithm for practical instances, the quality of solutions the quantum algorithm can find, and its potential to be used in parallel with the branch-and-price algorithm. Additionally, prior to utilizing a quantum algorithm, classical preprocessing techniques are applied to the RMP instance and the output of the quantum algorithm is used as input to classical postprocessing techniques. We note that the method is similar to those explored in Ref. [35] and shares similarities to the use of a quantum device for scheduling problems in Ref. [36]. However, our proposed method considers the hybrid classical and quantum approach for large-scale ILPs and is inspired by the integration of classical IP solvers for 0-1 integer programs into a generation and selection approach for large-scale ILPs in Ref. [5].

The addition of a quantum algorithm can improve the classical algorithm in several ways. Firstly, the quantum algorithm can provide a set of optimal or near-optimal

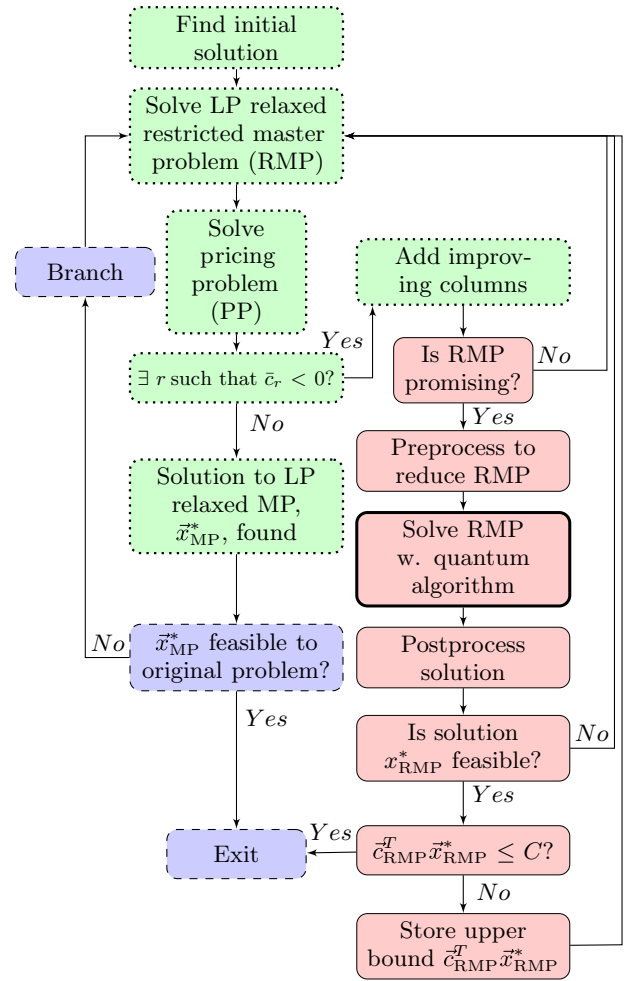


FIG. 1. High-level depiction of the branch-and-price algorithm integrated with a quantum algorithm capable of finding solutions to ILPs. The variable \bar{c}_r is the reduced cost of route r , \bar{c}_{RMP} is the cost vector of an RMP instance, and its entries correspond to for tail assignment to the entries c_r for $r = 1, \dots, |R|$ in Eqs. (1)–(3). The solution provided by a quantum algorithm with postprocessing is the vector \bar{x}_{RMP}^* , and the constant C is a threshold for the accepted quality of a solution. This hybrid algorithm is one of the main results of this paper.

integer solutions to RMP instances, which means that the quantum algorithm can be used as a primal heuristic in the column generation algorithm. This technique is sometimes referred to as the restricted master heuristic [37]. In the restricted master heuristic, a subset, which is a fixed number of columns and variables, is chosen from the RMP and the resulting problem is solved as a static integer program (IP). However, we do not wish to restrict the number of variables and columns to solve as a static IP. Instead, we propose to use the whole RMP instance unless we are required to leave out variables due to limitations in the size of a quantum device. Such heuristics can improve the solution quality as observed in Ref. [38] by simply obtaining

optimal or near-optimal solutions to RMP instances. Furthermore, as primal heuristics have been shown to be key for solving mixed integer programs, heuristics that leverage a quantum algorithm seems to be a natural step for branch-and-price.

Moreover, by finding a set of integer solutions, some flexibility is introduced as it is possible to compare the quality of several solutions with respect to more parameters than each solution's cost. This is mainly an advantage for a real-world problem, where buffers occurring in solutions can improve the sensitivity to disruptions.

Secondly, the quantum algorithm can provide tighter upper bounds in the branching step, which can be utilized in pruning decisions directly without sacrificing optimality. When we have access to tighter upper bounds, these bounds are compared to the lower bounds found in the column generation algorithm. If the lower bound is greater or equal to the upper bound, we can discard the subproblem as we can prune by bound. If we do not have access to these tighter upper bounds, more subproblems are created and explored. This means that the upper bounds can reduce the search tree's size, which leads to a reduced running time of the algorithm. The upper bounds can also reduce the number of iterations required in the column generation algorithm as noted in Ref. [35] by computing the Lagrangian lower bound, where the stopping criteria is given when the Lagrangian lower bound is greater than the best known upper bound. We can also consider introducing heuristic pruning rules that can reduce the running time of branch-and-price. We remark that finding a good solution fast can be preferable to finding the optimal solution for real-world problems. Heuristic pruning rules guided by optimal or near-optimal solutions to RMP instances can therefore be beneficial. However, as even optimal integer solutions to RMP instances do not guarantee an optimal solution to the subproblem in branch-and-price, the pruning decisions will be heuristic and do not guarantee an optimal solution.

By introducing heuristic pruning rules, the goal is thus to obtain high-quality solutions faster. The pruning decision can be determined by comparing the solution quality for different RMP instances by monitoring the iterative change in the objective and the LP lower bound gap. If the branch-and-price is based on variable fixing decisions, the solutions from a quantum algorithm can indicate if certain variables can be chosen to be fixed. The procedure of fixing a variable is such that if a variable x_i is set to 1 for a majority of the obtained solutions, the variable can be fixed to 1 and the branch-and-bound algorithm dives into this particular branch of the search tree. Further techniques as in RQAOA in Ref. [39] can also be utilized where it is possible to find relations between two decision variables $z_i = \sigma_{ij} z_j$ where $\sigma_{ij} = \text{sign}(\langle \vec{\gamma}^*, \vec{\beta}^* | \hat{\sigma}_i^z \hat{\sigma}_j^z | \vec{\gamma}^*, \vec{\beta}^* \rangle)$ and (i, j) is an edge in the graph $G = (V, E)$ such that

$(i, j) = \text{argmax}_{(i', j') \in E} \{ |\langle \vec{\gamma}^*, \vec{\beta}^* | \hat{\sigma}_{i'}^z \hat{\sigma}_{j'}^z | \vec{\gamma}^*, \vec{\beta}^* \rangle| \}$ of an Ising model that encodes an ILP. Such heuristic pruning rules would be similar to the ones of diving heuristics (which can be greedy, random, or based on rounding strategies) or local branching heuristics [40].

Finally, the quantum algorithm can reduce the running time if it finds some integer solution below a given threshold or sufficiently close to the lower bound of the original problem as the algorithm, in that case, stops even though the search tree of branch-and-price is not explored fully.

Since the method is heuristic, the running time can best be evaluated by executing it on real problems and quantum devices, which is currently intractable due to the current size of quantum computers. We expect that as quantum hardware matures, such experiments will be of interest.

We can, on the other hand, note that the general branch-and-bound algorithm has worst-case running time $\mathcal{O}(Mb^d)$, where b is branching factor, d is the search depth, and M is the upper bound on the running time to explore a subproblem fully. If we can obtain optimal or near-optimal integer solutions to subproblems, the number of nodes we can prune is larger and thus reduces the algorithm's running time.

Furthermore, as the augmented algorithm is valid for any branch-and-price algorithm applied to problems with master problems (MPs) possible to solve by some quantum algorithm, the method can be suitable for a large class of ILPs. In particular, this framework can be employed for airline planning problems such as tail assignment, crew pairing, and crew rostering but also other large-scale ILPs, such as vehicle routing problems [41]. Whilst this approach prohibits applying a quantum algorithm to the tail assignment problem and other large-scale ILPs directly, it reduces the number of required decision variables and qubits. In particular, the MP for tail assignment, which is a set partitioning problem, is mapped such that the number of decision variables corresponds exactly to the required number of qubits (this is true also for other MPs that are 0-1 variable LPs with equality constraints). We could map the tail assignment problem directly to an Ising model using an arc-based formulation [see in Ref. [3] Eqs. (4.1)–(4.7)], but this would require 10^7 qubits for a problem with 10^3 flights and ten aircraft prior preprocessing. For typical RMP instances, we instead expect to require around $10^3 - 10^4$ decision variables for the path-based formulation in Eqs. (1)–(3). The proposed method is thus much more suitable for NISQ computers. The arc-based formulation has an additional disadvantage beyond the resource requirement of qubits for problems as tail assignment, which are the recursive maintenance requirements. These are nontrivial to map to an Ising model, and removing the constraints would likely result in infeasible solutions.

Moreover, we expect that if RMP instances can be solved approximately with sufficiently shallow circuit depth, the circuits can be realized on NISQ [42] computers.

As mentioned earlier in this section, we propose a preprocessing step using classical integer programming techniques [43,44] in order to reduce the number of variables and constraints of the problem prior to utilizing a quantum computer. Reducing the number of variables (required qubits) and constraints (problem graph connectivity) is useful for the limited NISQ computers to be able to address real-world problems. The level of sophistication can range from very basic to very advanced techniques and the level of sophistication used will be a trade-off between the computational time of the preprocessor and the size and performance of the quantum computer. We also consider classical postprocessing of the output from a quantum algorithm, where additional local searches can be done and we can combine good RMP solutions to obtain improved solutions with standard or specialized classical solvers. Infeasible solutions can additionally be attempted to be corrected to feasible solutions by heuristic classical algorithms.

We stress that the benefit in separating the original problem with the branch-and-price algorithm is that the master problem often is a pure set partitioning or set cover problem without any additional side constraints. The PP, on the other hand, is often a resource-constrained shortest-path problem that considers the complex rules. Thus, the method is not based on being more suitable for NISQ devices but is based on known successful methods for solving complex large-scale ILPs. Furthermore, by simplifying a real-world problem to a pure set cover or set partitioning problem we also avoid tackling an ILP with potentially many complicated side constraints with quantum algorithms. This also means that the intricate task of balancing multiple constraint penalties is simplified.

If the method is favorable for large-scale ILPs depends on how complicated the constraints are and the resource requirements of various formulations. The method proposed here can be expected to provide constant speedup and improve the quality of the solutions. However, it is unclear if the method can provide polynomial speedup as the addition of a quantum algorithm provides no guarantee for a speedup and is tied to the column generation algorithm, which limits the possible speedup we can expect. If it is possible to use an alternate formulation that is not required to be separated into a generation problem and a selection problem, it might be beneficial to map the problem directly to an Ising spin glass Hamiltonian. However, as we have pointed out, this often requires significantly more decision variables and qubits to be applicable to real problems and will be more challenging for NISQ devices.

IV. THE QUANTUM APPROXIMATE OPTIMIZATION ALGORITHM

Farhi, Goldstone, and Gutmann presented in Ref. [18] QAOA, which is a hybrid classical and quantum variational algorithm capable of finding approximate solutions to combinatorial optimization problems. The algorithm is inspired by the adiabatic quantum algorithm but is designed for gate-based quantum computers. Furthermore, evidence that a classical computer cannot simulate a QAOA circuit without exponential overhead was presented later in Ref. [19]. The algorithm consists of alternating the operators $e^{-i\gamma_k \hat{H}_f}$ and $e^{-i\beta_k \hat{H}_M}$ for $k = 1, 2, \dots, p$, where p is the depth of the algorithm. An ideal QAOA circuit applied to the initial state $|+\rangle = 1/\sqrt{2^n} \sum_{i=0}^{2^n-1} |i\rangle$ gives the QAOA state for depth p

$$|\vec{\gamma}, \vec{\beta}\rangle = e^{-i\beta_p \hat{H}_M} e^{-i\gamma_p \hat{H}_f} \dots e^{-i\beta_1 \hat{H}_M} e^{-i\gamma_1 \hat{H}_f} |+\rangle,$$

where $\hat{H}_M = \sum_{i=1}^n \hat{\sigma}_i^x$ is the mixing Hamiltonian and $\hat{H}_f = \sum_{\vec{x} \in \{0,1\}^n} f(\vec{x}) |\vec{x}\rangle \langle \vec{x}|$ is a diagonal cost Hamiltonian with respect to the computational basis. The cost Hamiltonian encodes an objective function $f(\vec{x})$, which represents a combinatorial optimization problem. With optimal angles $\vec{\gamma}^*$ and $\vec{\beta}^*$ and sufficiently large algorithm depth, the QAOA state should have a large proportion in states that are close to the ground state and equal to the ground state. By repeating the process of constructing the state and performing measurements in the computational basis, a solution that is equal or close to the ground state of the cost Hamiltonian can be found.

The total running time to execute the quantum circuit, as well as the implementation of the gates associated to the cost Hamiltonian is both graph and hardware architecture dependent. If we let the algorithm depth go to infinity and restrict the angles to be small, the algorithm becomes exact [18].

For an ILP problem \hat{H}_f will consist of one partial Hamiltonian that corresponds to the objective function and another that corresponds to constraints, not unlike common penalty methods [45]. If $f(\vec{x})$ represents a minimization problem the optimal angles $\vec{\gamma}^*$ and $\vec{\beta}^*$ can be found by solving the classical optimization problem

$$\operatorname{argmin} \langle \vec{\gamma}, \vec{\beta} | \hat{H}_f | \vec{\gamma}, \vec{\beta} \rangle, \quad (7)$$

$$\text{subject to } \gamma_i \in [0, 2\pi] \forall i = 1, \dots, p, \quad (8)$$

$$\beta_i \in [0, \pi] \forall i = 1, \dots, p \quad (9)$$

as $\langle \vec{\gamma}, \vec{\beta} | \hat{H}_f | \vec{\gamma}, \vec{\beta} \rangle = f(\vec{x}^*)$ if $|\vec{\gamma}, \vec{\beta}\rangle = |\vec{x}^*\rangle$ where \vec{x}^* is the optimal solution to the problem $f(\vec{x})$ represents. The function in Eq. (7) is the expectation value function and can be referred to as the energy landscape. The domain in Eqs. (8) and (9) holds for Hamiltonian \hat{H}_f with integer eigenvalues [24].

As far as we know, instances extracted from the real-world problem tail assignment has previously only been studied for QAOA in the context of exact cover in Ref. [24] where success probabilities close to unity for instances up to 25 qubits with one feasible solution could be obtained for $p \leq 20$ for ideal QAOA circuits. Recently, the vehicle routing problem was also studied up to 20 qubits [46] where a clear dependency was established between the problem size and the performance of QAOA. On the other hand, real-world problems have been studied for quantum annealing, such as for flight gate assignment in Ref. [47], where the authors address the issue of bin packing the cost vector of the objective function. However, the complication of degenerate problem instances have not been discussed to a large extent in the context of QAOA, nor has much focus been given to how suitable weights are found to balance the constraints and the objective part of the Hamiltonian \hat{H}_f . In Sec. VII, we focus on the effect of choosing suitable weights on the required algorithm depth given a success probability and if having a large feasible space is a limiting factor for the performance.

A. Mapping set partitioning and exact cover

It is possible to map the set partitioning and exact cover problem to the Ising spin glass Hamiltonian with an underlying graph $G = (V, E)$ with nodes given by the set V and the edges given by the set E , where the Hamiltonian is $\hat{H} = \sum_{i=1}^{|V|} h_i \hat{\sigma}_i^z + \sum_{(i,j) \in E} J_{ij} \hat{\sigma}_i^z \hat{\sigma}_j^z$ as presented in Ref. [48]. In this case, the Hamiltonian has at most two spin interaction terms $\hat{\sigma}_i^z \otimes \hat{\sigma}_j^z$ and is 2-local [49], albeit this does not correspond to a geometric locality with respect to hardware architecture.

By introducing a quadratic penalty on the constraints in Eq. (2) a nonlinear integer optimization problem is obtained. The quadratic penalty results in a Hamiltonian which has two parts (when ignoring a constant energy shift), a Hamiltonian, which is related to the objective function and a Hamiltonian related to the constraints. These parts are weighted with constants μ_1 and μ_2 accordingly

$$\begin{aligned} \hat{H}^{\text{set partitioning}} = & \sum_{r \in R} [\mu_1 \times h_r^{\text{objective}} \\ & + \mu_2 \times h_r^{\text{exact cover}}] \hat{\sigma}_r^z \\ & + \mu_2 \times \sum_{r' > r} J_{rr'}^{\text{exact cover}} \hat{\sigma}_r^z \hat{\sigma}_{r'}^z, \end{aligned}$$

where

$$\begin{aligned} h_r^{\text{objective}} &= \frac{c_r}{2}, \\ h_r^{\text{exact cover}} &= \sum_{f \in F} a_{fr} \left(\sum_{r' \in R} \frac{a_{fr'}}{2} - 1 \right) \end{aligned}$$

and

$$J_{rr'}^{\text{exact cover}} = \sum_{f \in F} \frac{a_{fr} a_{fr'}}{2}.$$

We observe that the terms $h_r^{\text{objective}}$ are given by the objective function in Eq. (1) and therefore indicate the cost of an assignment of the decision variables $\vec{x} \in \{0, 1\}^{|R|}$. The terms $h_r^{\text{exact cover}}$ and $J_{rr'}^{\text{exact cover}}$ are due to the constraints in Eq. (2), where $J_{rr'}^{\text{exact cover}}$ gives a penalty for each overlapping flight in route r and r' and with the terms $h_r^{\text{exact cover}}$ gives a penalty if the combination of routes in an assignment does not cover all flights.

The problem graph $G = (V, E)$ is given by the coefficients $h_r^{\text{exact cover}}$, $h_r^{\text{objective}}$ and $J_{rr'}^{\text{exact cover}}$ in the Hamiltonian where the graph itself can be thought also as a conflict graph of the variables. Finally, the detailed mapping of exact cover to an Ising spin glass model was presented in Ref. [24] and further expanded for the mapping of the set partitioning problem in Appendix B. Mappings for other minimization problems common for large-scale ILPs such as set cover can also be found in Ref. [48].

V. PROBLEM INSTANCES

The instances [50] have been extracted from the real-world problem tail assignment by finding a set of different integer solutions when executing the heuristic branch-and-price algorithm. The different solutions are found by permuting the cost of routes randomly during the execution of the algorithm. From this set, 35 instances have been constructed with a varying number of routes and number of feasible solutions by combining complete and partial solutions. Typically, the instances have very large costs and can be as large as 10^6 , making the energy landscape numerically hard to search. The objective function has therefore been further simplified to study qualitative differences in the performance of RMP instances for QAOA. The costs have been simplified such that the smallest cost c_r^{min} is set to 1, larger costs have been modified such that each cost c_r has a unique value and that the optimal solution is unique. For real instances this is not a proposed methodology, as it can disturb the order of the solutions with respect to quality significantly. An option for real instances is to either increase the weight for the penalty of the constraints, which results in a numerically challenging energy landscape to optimize or we can disturb the costs such that they are easier to handle but preserves the objective function with some accuracy.

We can modify the costs by subtracting all costs with a constant and dividing all costs with another constant, finally the costs are rounded to integers. There is a limit to how much we can disturb the costs such that the order of solutions with respect to cost is not changed significantly. One should choose to divide by a constant that separates the costs c_r by at least a constant integer, which results in

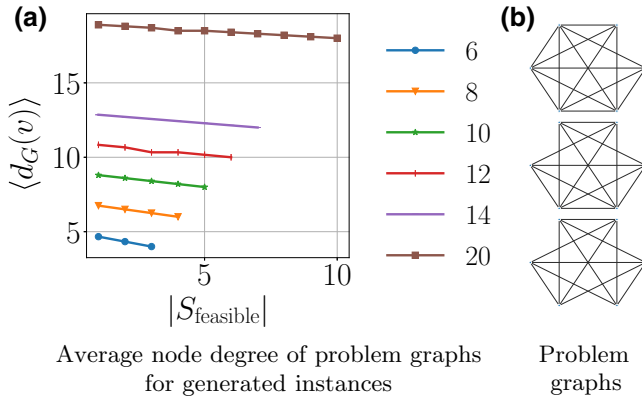


FIG. 2. Average node degree of the problem graphs are shown in Fig. 2(a). Problem graphs $\{G_r^s\}_{r=6}^{s=1,2,3}$ are depicted in Fig. 2(b).

a better preservation of the objective function compared to choosing a larger constant to divide the costs by. Here we have assumed a simple objective function to study the performance of QAOA.

The instances are identified by the number of decision variables $|R|$ and the number of feasible solutions $|\mathcal{S}_{\text{feasible}}|$. The number of decision variables are 6, 8, 10, 12, 14, and 20. The number of feasible solutions vary from 1 to $|R|/2$. We denote a problem graph associated to an instance $G_{r=|R|}^{s=|\mathcal{S}_{\text{feasible}}|}$, which gives the set of graphs as

$$\{G_r^s\}_{r=6}^{s=\lfloor r/2 \rfloor}.$$

Additionally, in Ref. [24] it was observed that the average node degree of the problem graphs affects the performance of QAOA, in that obtaining near unity success probability requires greater algorithm depth as the average node degree, $\langle d_G(v) \rangle$, of the problem graphs increases. The effect of the average node degree was found to dominate over the problem size such that for a given success probability, the required algorithm depth was greater for instances with 15 qubits compared to instances with 25

qubits. We have extracted the average node degree of each problem graph, depicted in Fig. 2. It can be noted that the average node degree increases with the problem size and decreases as the number of feasible solutions increases. We further noted that the problem graphs are close to being complete graphs, i.e., each node's degree is $|R| - 1$ or $|R| - 2$. It is thus expected that such instances are hard for QAOA to solve with respect to problem size.

VI. OPTIMIZATION STRATEGY

Finding the solution to the optimization problem in Eqs. (7)–(9) is NP-hard [51,52] in itself. Furthermore, each query of the function in Eq. (7) requires either executing the QAOA circuit on a quantum device or a simulation on a classical computer. As we are currently prohibited from executing QAOA for the problem instances with sufficient algorithm depths on a quantum device the remaining option is to simulate the algorithm with a classical computer. Moreover, since simulating the quantum circuits is exponential in the number of qubits, the consequence is that a function evaluation is computationally expensive. Furthermore, in order to study the performance of QAOA more accurately, we wish to study intermediate to large algorithm depths, which makes the simulations even more expensive as the dimension of the expectation value function in Eq. (7) is 2 times the algorithm depth. Compared to problems as maxcut with uniform weights set to 1 or versions thereof [18,53,54] the set partitioning problem and exact cover problem have coefficients in the Hamiltonian h_i and J_{ij} that are governed by the constraint matrix and objective function that grow with the chosen weights. These coefficients are thus not constrained to 0,1 or -1 and can be large. The difference in coefficients results in complicated energy landscapes, that oscillate rapidly, to optimize with multiple local minima. Moreover, we can see this from the closed form expression of the energy landscape for $p = 1$ for an Ising spin glass Hamiltonian, associated to a graph $G = (V, E)$ with edge weights J_{ij} and node weights h_i , which is given by

$$\begin{aligned} \langle \gamma\beta | \hat{H} | \gamma\beta \rangle &= \sum_{i=1}^n h_i \sin(2\beta) \sin(2\gamma h_i) \prod_{j:(i,j) \in E} \cos(2\gamma J_{ij}) + \sum_{(i,j) \in E} \frac{J_{ij}}{2} \left(\sin^2(2\beta) \prod_{\substack{(i,k) \in E \\ (j,k) \notin E}} \cos(2\gamma J_{ik}) \prod_{\substack{(j,k) \in E \\ (i,k) \notin E}} \cos(2\gamma J_{jk}) \right. \\ &\times \left[\cos(2\gamma(h_i - h_j)) \prod_{\substack{(i,k) \in E \\ (j,k) \in E}} \cos(2\gamma(J_{ik} - J_{jk})) - \cos(2\gamma(h_i + h_j)) \prod_{\substack{(j,k) \in E \\ (i,k) \in E}} \cos(2\gamma(J_{ik} + J_{jk})) \right] \\ &\left. + \sin(4\beta) \sin(2\gamma J_{ij}) \left[\cos(2\gamma h_i) \prod_{k \neq j:(i,k) \in E} \cos(2\gamma J_{ik}) + \cos(2\gamma h_j) \prod_{l \neq i:(j,l) \in E} \cos(2\gamma J_{jl}) \right] \right), \quad (10) \end{aligned}$$

as presented in Ref. [55]. We derive the expression for consistency in Appendix C.

The complicated energy landscape underlies our motivation to focus on obtaining good locally optimal angles via the interpolation strategy presented by Zhou, in Ref. [56], in order to study the success probability for QAOA with intermediate to large algorithm depth p . The first step in the interpolation algorithm is to perform global optimization for algorithm depth $k = 1$ and for algorithm depth $k > 1$ locally optimal angles $(\vec{\gamma}^{L^*}, \vec{\beta}^{L^*})$ angles are found by providing a good starting point $(\vec{\gamma}^L, \vec{\beta}^L)$ to a local search algorithm. The starting point for local search is determined by interpolating previously found locally optimal angles. The algorithm iterates for $k = 2, \dots, p$. The following definition gives the interpolation in each step:

$$\eta_{k+1,i}^L = \begin{cases} \eta_{k,1}^{L^*} & \text{if } i = 1 \\ \frac{i-1}{k} \eta_{k,i-1}^{L^*} + \frac{k-i+1}{k} \eta_{k,i}^{L^*} & \text{if } i = 2, \dots, k, \\ \eta_{k,k}^{L^*} & \text{if } i = k + 1 \end{cases}$$

where η is γ or β . The index i denotes the i :th element of locally optimal angles found for algorithm depth k and index k denotes the best found angles of algorithm depth k . The distinction between L and L^* is the separation of the starting point and angles found after a local search. In our case, the global optimization was performed with PYTHON's differential evolution routine. The local optimization was performed with L-BFGS-B, which is also a standard solver in PYTHON.

VII. NUMERICAL RESULTS FOR RESTRICTED MASTER PROBLEM INSTANCES

We present the numerical results obtained for ideal QAOA circuits where the variational parameters have been obtained via the interpolation strategy first for exact cover in Sec. VII A and second for set partitioning in Sec. VII B.

A. Solving the exact cover problem

For exact cover, we require only to obtain a feasible solution $\vec{x}_i \in \mathcal{S}_{\text{feasible}}$. For such a purpose, the most natural choice of mapping is by ignoring the objective part of the Hamiltonian, i.e., the cost Hamiltonian is expressed as

$$\hat{H}_f = \hat{H}^{\text{exact cover}}.$$

Furthermore, it is straightforward to define the success probability as the probability of obtaining any of the feasible solutions

$$P_{\text{success}}^{\text{exact cover}} = \sum_{\vec{x}_i \in \mathcal{S}_{\text{feasible}}} |(\vec{x}_i | \vec{\gamma}^{L^*}, \vec{\beta}^{L^*})|^2.$$

The success probabilities for QAOA applied to the exact cover instances are plotted in Fig. 3. We remark that the

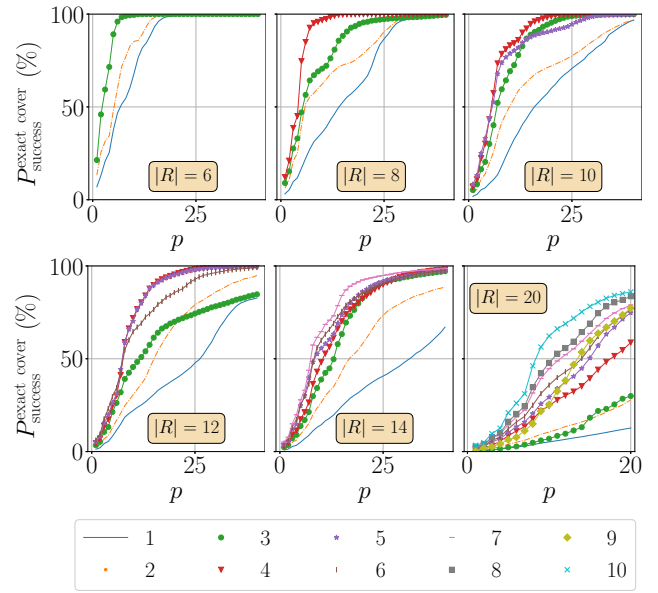


FIG. 3. Success probabilities for exact cover. The colors and markers indicate $|\mathcal{S}_{\text{feasible}}|$.

required algorithm depth decreases for a given success probability as the number of feasible solutions increases in general, albeit increases when the problem size increases. These results correspond to results found in Ref. [24], wherein Fig. 2 we presented the average node degree of the instances, that decreases with the number of feasible solutions whilst increasing more significantly as the problem size increases.

As the most challenging cases are those where the number of feasible solutions is small, we observe that obtaining a success probability above 50% can require an algorithm depth that is more than $|R|$ by at least a constant, where $|R|$ is the number of decision variables and qubits of the instance. Therefore, it is unknown how well QAOA can perform for instances with $10^3 - 10^4$ decision variables when executed on a NISQ device as decoherence is a limiting factor currently.

B. Solving the set partitioning problem

When we consider applying QAOA to the set partitioning problem, two additional aspects are of interest. The first aspect is how one should choose good weights that balance the objective part of the Hamiltonian and the exact cover (constraints given by a quadratic penalty) part of the Hamiltonian. The second aspect is a consequence of the first, namely how the chosen weights affect the required algorithm depth for a given success probability. The total cost Hamiltonian is a combination of the two partial Hamiltonians accordingly

$$\hat{H}_f = \hat{H}^{\text{set partitioning}} = \mu_1 \hat{H}^{\text{objective}} + \mu_2 \hat{H}^{\text{exact cover}}.$$

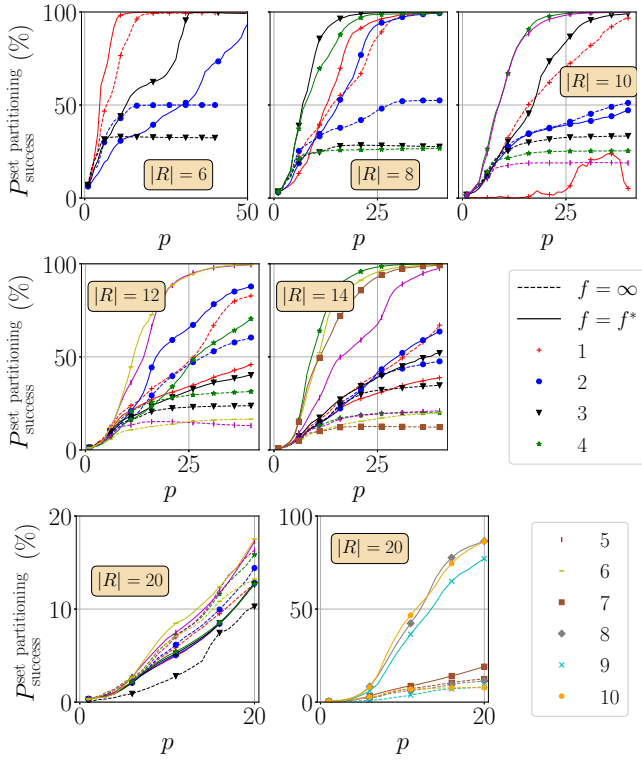


FIG. 4. Success probability for solving the set partitioning problem depending on the choice of weights μ_1 and μ_2 , the dashed lines correspond to $f = \infty$ and solid lines correspond to the best found factors $f = f^*$. The colors and markers of the lines indicate the value of $|S_{\text{feasible}}|$.

We have chosen the weight $\mu_1 \in \{\mathbb{Z}^+ \cup \{0\}\}$ depending on a factor f

$$\mu_1 = \begin{cases} 0 & \text{if } f = \infty \\ 1 & \text{otherwise} \end{cases},$$

and $\mu_2 \in \mathbb{Z}^+$ depending on the largest eigenvalues of the partial objective and exact cover Hamiltonians, and factor f

$$\mu_2 = \begin{cases} 1 & \text{if } f = \infty \\ \left\lceil f \times \frac{\lambda_{\text{objective}}^{\max}}{\lambda_{\text{exact cover}}^{\max}} \right\rceil & \text{otherwise} \end{cases}.$$

By choosing the weights to be integers, the domain is preserved in the optimization problem defined in Eqs. (7)–(9). Thus, $f = \infty$ corresponds to the mapping where $\hat{H}^{\text{set partitioning}} = \hat{H}^{\text{exact cover}}$. We then define the success probability as the probability of finding the optimal solution

$$P_{\text{success}}^{\text{set partitioning}} = |\langle \vec{x}^* | \vec{\gamma}^{L^*}, \vec{\beta}^{L^*} \rangle|^2,$$

where \vec{x}^* is the solution to the set partitioning problem, i.e., the binary vector that corresponds the minimal value of Eq. (1) such that $\vec{x}^* \in S_{\text{feasible}}$. The success probabilities of set partitioning are plotted in Fig. 4 for ideal

TABLE I. Success probabilities for QAOA applied to set partitioning for problem sizes 6–20, given algorithm depth p for multiple choices of factor f .

$ R $	p	$ S_{\text{feasible}} $	$P^{f=\infty}$	$P^{f=100}$	$P^{f=10}$	$P^{f=1}$
6	40	1	99.55	99.52	99.96	99.98
		2	50.	51.36	59.93	72.67
		3	32.39	36.	67.43	99.71
8	40	1	99.68	99.54	99.83	99.83
		2	52.47	57.29	95.46	99.33
		3	27.62	38.48	99.22	99.99
		4	26.69	96.68	99.75	99.4
10	40	1	96.75	5.19		
		2	51.15	0.28		47.18
		3	33.37	38.33	99.03	
		4	25.38	39.52	99.95	
		5	18.89	89.3	99.91	
12	40	1	82.72	45.85		
		2	60.41	64.52	87.83	
		3	23.78	27.19		40.21
		4	31.35	37.62		
		5	13.06	20.15	99.5	
		6	16.54	24.31	99.53	
14	40	1	67.07	38.76		
		2	47.59	63.63	21.06	
		3	34.73	43.78	52.09	
		4	19.94	42.91	99.92	
		5	20.99	31.98	97.46	
		6	20.06	36.09	99.63	
		7	12.25	25.68	99.22	
20	20	1	12.71	12.79		
		2	14.42	12.8		
		3	10.27	12.64		
		4	15.84	12.68		
		5	16.34	17.26		
		6	13.23	17.53		
		7	12.44	19.09		
		8	11.32	86.52		
		9	8.19	77.11		
		10	7.75	86.39		

QAOA circuits. Dashed lines distinguish the lines for factor $f = \infty$ and the best found factors f^* are distinguished by the solid lines. Furthermore, success probabilities are tabulated for additional factors for a given algorithm depth in Appendix D, where the factors have been chosen to construct cost Hamiltonians with the constraint that the ground state corresponds to the optimal solution \vec{x}^* .

It is clear from the results of the numerical simulations in Fig. 4 and Table I that the success probability of solving set partitioning can be increased (and thus reducing the required algorithm depth) with a suitable choice of weights μ_1 and μ_2 for 22 instances of the 29 instances with more than one feasible solution. We also observe that a good choice of weights for instances with a single feasible solution corresponds to $f = \infty$ for all problem sizes. We observe that the success probability can decrease with the number of feasible solutions to $P_{\text{success}}^{\text{set partitioning}} \approx 1/|S_{\text{feasible}}|$ if the weights are chosen poorly, which in the worst case is exponential in the problem size. To avoid requiring a considerable algorithm depth, finding good weights is thus required to solve the optimization problem with NISQ devices.

Moreover, the regret (the difference between the minimum expectation value found during the optimization procedure and the optimal solution) of the expectation value function is depicted in Fig. 5 for instances with six routes with varying weights.

We observe for $f = \infty$ that the regret is reduced to near zero, whilst failing to increase the success probability significantly above $1/|S_{\text{feasible}}|$. For factors 10 and 100, the regret is greater compared to the best found factor for a given algorithm depth. The difference in regret corresponds to decreased required algorithm depth for the best found factor compared to factors 10 and 100 to achieve near-unity success probability for set partitioning, see Table I.

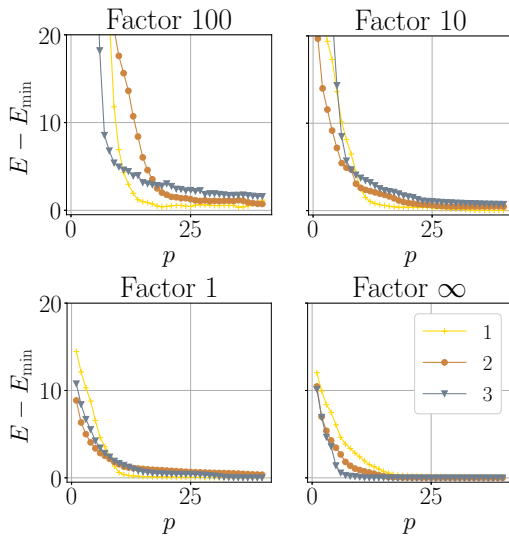


FIG. 5. Regret of the expectation value function for instances corresponding to problem graphs $G_6^{1,2,3}$ for factors 1, 10, 100, and ∞ . The regret is defined as the difference between $E = \langle \vec{\gamma}^{L*}, \vec{\beta}^{L*} | \hat{H}_f | \vec{\gamma}^{L*}, \vec{\beta}^{L*} \rangle$ where $(\vec{\gamma}^{L*}, \vec{\beta}^{L*})$ are the locally optimal angles found by the interpolation strategy and $E_{\min} = \langle \vec{x}^* | \hat{H}_f | \vec{x}^* \rangle$.

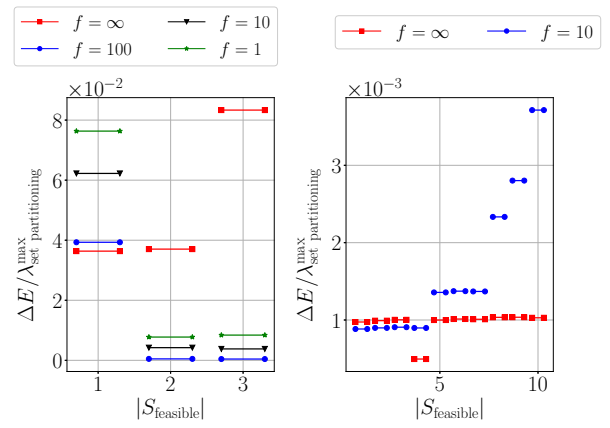


FIG. 6. Minimum nonzero energy gap ΔE for cost Hamiltonian $\hat{H}^{\text{set partitioning}}$ depending on the factor f , as a ratio of the maximum eigenvalue.

Since we observed that choosing a factor other than ∞ fails to increase the success probability for seven instances with more than one feasible solution, we have extracted the smallest nonzero energy gap ratio with respect to the maximum eigenvalue. Figure 6 shows the ratio for instances with 6 and 20 routes. The graphs show for instances with six routes that the ratio can be increased for G_6^1 but not for $G_6^{2,3}$ by choosing a factor that considers the cost function. The lack of increased ratio corresponds to the increased required algorithm depth to obtain near-unity success probability for $G_6^{2,3}$ compared to G_6^1 . Furthermore, the choice $f = 10$ compared to $f = \infty$ results for instances G_{20}^{1-3} in decreased ratios. Whereas the ratio is increased for G_{20}^{4-7} and more distinctly for $G_{20}^{8,9,10}$. We note that as the ratio increases for the choice of factor f , the required algorithm depth is decreased for a given success probability here as well. We conclude from these results that a suitable choice for weights is such that the nonzero energy gap is as large as possible as a ratio of the maximum eigenvalue of the cost Hamiltonian. Moreover, when we attempt to balance the objective and constraint parts of the cost Hamiltonian the smallest eigenvalues are not guaranteed to correspond to feasible solutions. This means that in the pursuit of finding the optimal solution, we can decrease the probability of finding a feasible solution. However, if we find weights such that the smallest eigenvalues correspond to feasible solutions, we do not sacrifice the probability of finding good feasible solutions for finding the optimal solution.

These results indicate that if NISQ devices are limited in algorithm depth, finding suitable weights will be crucial, requiring more computational effort. The task of finding suitable weights for set partitioning via the quadratic penalty method typically requires that several subproblems are solved, where each subproblem corresponds to a choice of weights. Typically, with the quadratic penalty method,

the weight for the objective part is set to 1 and the weight for the quadratic penalty is set to be small initially. The weight of the quadratic penalty is then increased for a number of iterations or until convergence is reached. The quadratic penalty method could be executed with QAOA. We could also consider solving the problem with a classical computer, where the integer requirement could be relaxed to provide a good guess for the weights. An initial starting point for the weights can also be chosen as $\mu_1 = 1$ and $\mu_2 > \max_{(i,j) \in E} \{|c_i - c_j|\}$ if we assume that the smallest penalty is 1 for exchanging variables x_i and x_j or $\mu_2 > \sum_{r=1}^{|R|} |c_r|$ [57] where μ_2 is bounded from above. An alternative method to obtain suitable weights can be to initially attempt to solve exact cover where the weight is zero for the objective part and one for the constraint part. For the second iteration, QAOA with equal penalties set to one for the objective and constraint part of the Hamiltonian is then executed. If the solution degrades to be infeasible compared to the first solution we can assume that the objective part of the Hamiltonian dominates the constraint part. In that case, we need to increase the penalty for the constraint part for a number of iterations or until we reach a convergence. If we, on the other hand, find that we obtain a solution of similar cost as when we attempted to solve exact cover, we can increase the penalty for the objective part of the Hamiltonian for a number of iterations until we observe convergence for the solutions or until the solution degrades again such that it is infeasible. Since each choice of weights corresponds to a subproblem to be solved with QAOA it implies a computational overhead. However, if QAOA itself is executed in polynomial time the overhead should not change the overall complexity of the algorithm.

Finally, we conclude that the required algorithm depth of QAOA can be expected to grow with the problem size and increase as the number of feasible solutions decreases (assuming that we have identified suitable weights). Figure 4 shows that we can expect to require at least $|R|$ in algorithm depth to achieve success probability above 50%.

VIII. CONCLUSIONS

We have proposed a method that can leverage quantum algorithms for large-scale ILPs and investigated the method by considering the quantum algorithm QAOA and the problem tail assignment. The method is useful for problems that are typically solved via column generation techniques, where a direct application to the problem (typically in a path-based formulation) requires in the worst case exponentially many qubits.

The method can also be useful for NISQ devices as our method require less quantum resources compared to the arc-based formulations for problems as vehicle routing and tail assignment [defined in Ref. [3] as model TAS in Eqs. (4.1)–(4.7)]. For crew pairing and crew rostering, in particular, some constraints are not suited to be expressed

in mathematical terms as noted in Ref. [5], utilizing a quantum algorithm in the branch-and-price framework for solving RMP instances can thus be the only viable option. Furthermore, for tail assignment, some constraints are recursive and nontrivial to express as an Ising model, limiting the potential to apply a quantum algorithm to the arc-based formulation directly.

The numerical results expand on the results in Ref. [24] by considering more diverse and realistic, albeit small instances. The results indicate that the required algorithm depth decreases for a given success probability as the number of feasible solutions increases for exact cover, where we find the opposite results for set partitioning if the cost Hamiltonian is weighted poorly. Moreover, the reduction in success probability for set partitioning can be significant as the number of feasible solutions can be very large. However, we also found that it is possible for most instances to find a suitable choice of weights such that the algorithm depth is significantly reduced to obtain a success probability above 50%, in particular, for instances where the number of solutions is larger. Even with suitable weights, we expect that instances can require an algorithm depth that grows with the problem size and node degree, where harder instances are those with few feasible solutions for QAOA with respect to both set partitioning and exact cover. Especially hard set partitioning instances for QAOA are expected to be those where the minimum nonzero energy gap is small with respect to the largest eigenvalue for any weights we choose and where the minimum eigenvalues no longer correspond to the feasible solutions (whilst the ground state is still the optimal solution). These instances are more difficult because the probability of finding a feasible solution degrades in these cases whilst favoring the optimal solution.

Moreover, we have chosen to follow the mapping for both problems as presented in Ref. [48]. Since there exists no evidence that suggests that this particular mapping, although obvious, is optimal there can exist some other more suitable mapping. Since it was observed that the node degree of the graphs affects the required algorithm depth, there might exist some more suitable mapping to be explored where the average node degree of the problem graphs can be reduced. However, exploring alternative mappings for exact cover and set partitioning has been omitted in this work and left as a potential future challenge to consider.

It can further be observed that common sizes of RMP instances of tail assignment require approximately 10^3 – 10^4 for practical problems. As NISQ computers were suggested to typically have 50–100 qubits initially, we would like to address this discrepancy. We remark that the quantum hardware is improving and promises of NISQ devices with 1000 qubits by companies, such as IBM in 2023 [58], implies that the method will become applicable on NISQ devices in the near future. For future work, it would

therefore be interesting to run QAOA on such devices for larger instances. Instances of interest to consider are generated RMP instances in branch-and-price frameworks for real-world problems and other hard ILP instances publicly available in operational research and mathematical optimization libraries. A remaining challenge for NISQ devices will be to realize QAOA circuits with the desired number of qubits for polynomial algorithm depths.

For future work, it could also be interesting to study if it is possible to reduce the RMP instance to be better suited for NISQ devices. For example, one could attempt to choose a subset of decision variables in RMP instances to construct smaller RMPs. However, such a reduction corresponds to options with a combinatorial behavior. Reducing the size of RMP instances can therefore require more advanced preprocessing techniques. Further techniques as those explored in Ref. [59] can also be valuable to consider.

We note that whilst our method provides a possibility to leverage quantum algorithms to an advantage for large-scale ILPs, any quantum algorithm under consideration must be capable of either providing significant speedup in finding solutions of similar quality as the best classical solvers or capable of finding solutions of improved quality compared to classical solvers during the same execution time.

The numerical experiments we have considered in this paper for QAOA cannot answer these open questions fully. However, it should be observed that as the average node degree of the generated instances are large, we can therefore consider that the results in Secs. VII A and VII B to correspond to hard instances for QAOA with respect to problem size. Larger instances that are sparse can therefore have a reduced requirement on the algorithm depth, which further motivates studying instances with lower node degrees by both numerical simulations and executions on quantum devices.

Finally, we conclude that it is possible to integrate QAOA with a branch-and-price algorithm, where we achieve reasonably high success probabilities for RMP instances with a polynomial algorithm depth. In obtaining high-quality integer solutions to RMP instances, the runtime of the general and heuristic branch-and-price algorithms can therefore be reduced and improve solution quality.

ACKNOWLEDGMENTS

This work was supported from the Knut and Alice Wallenberg Foundation through the Wallenberg Center for Quantum Technology (WACQT). G.F. acknowledges financial support from the Swedish Research Council through the VR project QUACVA.

APPENDIX A: THE HEURISTIC BRANCH-AND-PRICE ALGORITHM FOR SOLVING TAIL ASSIGNMENT

The branch-and-price algorithm is designed to solve large-scale ILPs and combines the algorithms column generation and branch-and-bound. In this section, we review first branch-and-bound and second the column generation algorithm. Last, we review the branch-and-price algorithm and the fixing heuristic presented in Ref. [3] subject to being integrated with a quantum algorithm.

1. Branch-and-bound

The branch-and-bound algorithm, given in Ref. [8] and surveyed in Ref. [60] more recently, provides a framework for finding the optimal solution to ILPs. As the feasible region is restricted to integer points and not convex, algorithms applicable for LPs can not solve ILPs generally. The distinction here is that LPs can be solved efficiently, whereas ILPs are NP-hard problems.

The algorithm, given in pseudocode in Algorithm 1, decomposes the original ILP into subproblems recursively that can be visualized with a tree structure. Exhaustive search is avoided by pruning nodes of the tree giving more

```

1:  $\bar{x}^* \leftarrow \emptyset$ 
2:  $z^* \leftarrow \infty$ 
3:  $z_U^* \leftarrow \infty$ 
4:  $\mathcal{L} \leftarrow \{S\}$ 
5: while  $|\mathcal{L}| > 0$  do
6:    $S_i \leftarrow \text{chooseSubProblem}(\mathcal{L})$ 
7:    $\mathcal{L} \leftarrow \mathcal{L} \setminus \{S_i\}$ 
8:   if  $S_i$  has feasible solution to LP relaxation then
9:      $(z_L, \bar{x}_L) \leftarrow \text{solveLPrelaxation}(S_i)$ 
10:    if  $z_L < z_U^*$  then
11:      if  $\bar{x}_L$  feasible to ILP then
12:         $\triangleright$  Prune by integrality
13:      if  $z_L < z^*$  then
14:         $\bar{x}^* \leftarrow \bar{x}_L$ 
15:         $z^* \leftarrow z_L$ 
16:         $z_U^* \leftarrow z^*$ 
17:      end if
18:    else
19:       $(z_U, \bar{x}_U) \leftarrow \text{getFeasibleSolution}(S_i)$ 
20:       $z_U^* \leftarrow \min(z_U^*, z_U)$ 
21:       $\{S_{i_1}, \dots, S_{i_k}\} \leftarrow \text{partition}(S_i)$ 
22:       $\mathcal{L} \leftarrow \mathcal{L} \cup \{S_{i_1}, \dots, S_{i_k}\}$ 
23:    end if
24:  else
25:     $\triangleright$  Prune by bound
26:  end if
27: else
28:   $\triangleright$  Prune by infeasibility
29: end if
30: end while
31: return  $(z^*, \bar{x}^*)$ 

```

Algorithm 1. Branch-and-bound (S).

acceptable running times in practice. Each node in the tree represents a subproblem, which is the original ILP with a reduced feasible space. Each subproblem can be relaxed, i.e., the decision variables are not discrete but continuous, yielding either a lower bound (if a minimization problem), an integer solution, or that the subproblem is infeasible.

Consider here that we are applying branch-and-bound to an integer linear program

$$I = \min \left\{ \sum_{i=1}^n c_i x_i : \vec{x} \in S \right\},$$

where $S = \{\vec{x} \in \mathbb{Z}_+^n : \sum_{i=1}^n a_{ji} x_i \geq b_j \ \forall j = 1, \dots, m\}$. The LP relaxation of the ILP is

$$L = \min \left\{ \sum_{i=1}^n c_i x_i : \vec{x} \in P \right\},$$

where $P = \{\vec{x} \in \mathbb{R}_+^n : \sum_{i=1}^n a_{ji} x_i \geq b_j \ \forall j = 1, \dots, m\}$. We know from linear programming theory that the LP relaxation of an ILP gives the relation $L < I$. A partition of the ILPs feasible space S yields two subproblems

$$I_1 = \min \left\{ \sum_{i=1}^n c_i x_i : \vec{x} \in S_1 \right\},$$

$$I_2 = \min \left\{ \sum_{i=1}^n c_i x_i : \vec{x} \in S_2 \right\},$$

where S_1 and S_2 are disjoint sets that partition S by a constraint on variable x_j such that $S_1 = \{\vec{x} \in S : x_j \leq \lfloor x_j^0 \rfloor\}$ and $S_2 = \{\vec{x} \in S : x_j \geq \lceil x_j^0 \rceil\}$. The variable $x_j^0 \in \vec{x}^0$ has some fractional value and \vec{x}^0 is an optimal solution to LP. We further know from linear programming theory that either ILP_1 or ILP_2 has the optimal solution to ILP. Similarly, the two subproblems can be related to alternative problems that correspond to the LP relaxation of ILP_1 and ILP_2 , which provides lower bounds, can show that there exists no feasible integer point or can find an optimal integer solution. The three problems, ILP, ILP_1 , and ILP_2 , can be visualized as a tree with a parent node and two child nodes, see Fig. 7. Clearly, ILP_1 and ILP_2 can be partitioned further into subproblems giving the tree structure rooted in a node representing the original ILP. If an LP relaxed subproblem is found to be infeasible, the node is pruned, i.e., the branch is not explored further and we say that the node is *pruned by infeasibility*.

When the algorithm recursively explores subproblems, an incumbent, z^* , is maintained which is the current best feasible solution found to the ILP. Whenever a subproblem yields a solution greater or equal to the incumbent, this region cannot contain any integer solutions that would

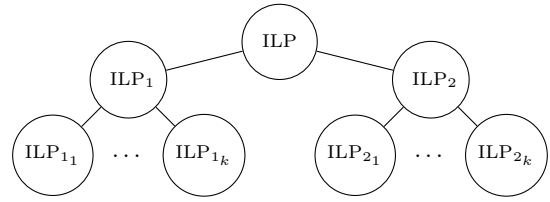


FIG. 7. Conceptual search tree of branch-and-bound.

improve upon the one we already have and this particular node is pruned. We say that the node is *pruned by bound*.

If we find that a solution to a subproblem is integral, we also prune this node as we have found an optimal partial solution or candidate incumbent z_i for this specific region. We say that the node is *pruned by integrality*. If $z_i < z^*$ the incumbent is updated.

Finally, if a subproblem cannot be pruned by infeasibility, bound or integrality the subproblem is partitioned into $k \geq 2$ nodes representing k subproblems, which are children to the current subproblem we are exploring in the tree. The k subproblems are then added to a list of unexplored subproblems and another subproblem is chosen to be explored. When there are no unexplored subproblems left the algorithm terminates and returns the incumbent solution and the corresponding assignment.

2. Column generation

In the previous section we mentioned that the LP relaxation of an ILP could be efficiently solved. However, consider the case where the number of variables is exponentially large so that even generating the LP would take exponential time and space. This is exactly the case for large-scale ILPs as the tail assignment formulation in Ref. [3], which has an exponential number of possible routes in the worst case.

The column generation algorithm [10], depicted with green colored boxes with dotted borders in Fig. 1 and presented in pseudocode in Algorithm 2, is based on well-known duality concepts from linear programming theory. It has been proved successful for both linear programs and ILPs, particularly when the number of decision variables is very large. Instead of attempting to construct and solve the complete problem it is decomposed into a master problem (MP)

$$z_{MP}^* = \text{minimize } \sum_{j \in J} c_j x_j,$$

$$\text{subject to } \sum_{j \in J} a_{ij} x_j \geq b_i \ \forall i \in I$$

$$x_j \geq 0 \ \forall j \in J$$

and a pricing problem (PP)

```

1:  $RMP \leftarrow FindInitialSolution(F, T)$ 
2:  $\bar{\pi}, \bar{x} \leftarrow SolveRestrictedMasterProblem(RMP)$ 
3:  $negativeReducedCosts \leftarrow solvePricingProblem(\bar{\pi})$ 
4:  $exitingVariables \leftarrow findExitingVariables(RMP)$ 
5: while  $negativeReducedCosts \neq \emptyset$  do
6:    $RMP \leftarrow RMP \cup \{negativeReducedCosts\}$ 
7:    $RMP \leftarrow RMP \setminus \{exitingVariables\}$ 
8:    $\bar{\pi}, \bar{x} \leftarrow SolveRestrictedMasterProblem(RMP)$ 
9:    $negativeReducedCosts \leftarrow solvePricingProblem(\bar{\pi})$ 
10:   $exitingVariables \leftarrow findExitingVariables(RMP)$ 
11: end while
12: return  $z = \bar{c}^T \bar{x}, \bar{x}$ 

```

Algorithm 2. Column generation (F, T).

$$\operatorname{argmin} \left\{ \bar{c}_j = c_j - \sum_{i \in I} a_{ij} \pi_i : j \in J \right\},$$

where π_i are the dual variables that correspond to the primal variables, x_j , found by solving the MP. The PP often encapsulates most of the problem-specific details and difficult constraints and generates alternative columns, also referred to as entering variables.

Since the number of decision variables is very large, the MP is further reduced to a restricted version, denoted the restricted master problem (RMP), meaning that the number of decision variables is smaller, often much smaller, than the original problem. The reduced size of the RMP is tractable to solve with some LP solver such as the dual simplex [61] or primal simplex [62] algorithm, compared to the MP.

The decomposition results in an iterative algorithm where the RMP and the PP are solved for a number of iterations or until optimal conditions hold. For each iteration, we attempt to find entering and exiting variables where the exiting variables are removed from the RMP and the entering variables are added to the RMP, resulting in other RMP and PP instances.

The PP is thus some problem that when solved can generate improving columns and decision variables to the RMP, based on given input of the dual variables from the RMP, such that the cost of the solution, which at this point is not guaranteed to be integral, is improved. Improving columns are identified by having a negative reduced cost \bar{c}_j and optimal conditions hold when no variables with negative reduced cost can be found, which is the same condition as in the simplex algorithm.

If the original problem is an ILP, the MP is the LP relaxation of the ILP. In the case of tail assignment the RMP corresponds to a restricted and LP relaxed set partitioning or exact cover problem, see Sec. II A, where the decision variables are continuous real variables. The PP can thus be

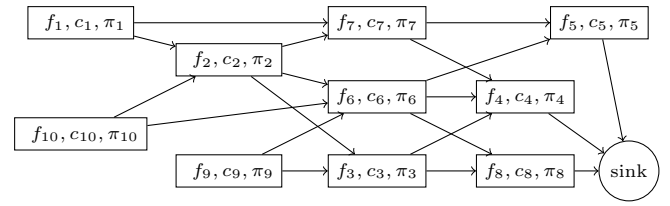


FIG. 8. Pricing problem.

defined as

$$\operatorname{argmin} \left\{ \bar{c}_r = c_r - \sum_{f \in F} a_{fr} \pi_f : r \in R \right\} \quad (\text{A1})$$

for tail assignment, where π_f is the dual variable of flight f obtained when solving the RMP.

To be noted, the first step of column generation is to construct an initial RMP, which for tail assignment can be $A = \mathbb{1}_{|F| \times |F|}$ where the costs c_r are set to some large number and thus unlikely to be part of a solution. Variables can be chosen as exiting variables when the value of the reduced cost is above a given threshold, however, removing variables from the RMP does not necessarily improve convergence as removing variables also removes dual information. Further investigations in deleting columns can be found in Ref. [3], in Sec. 6.4.

Furthermore, solving the PP at first glance appears intractable as the number of reduced costs can be exponentially large. By formulating the problem as a resource-constrained shortest-path problem (RCSP) we avoid explicitly constructing all routes. The RCSP is described by a connection network, depicted in Fig. 8 with a unique sink vertex and other vertices representing flights with edges that represent legal connections where the nodes are associated with a flight cost c_f and a dual variable π_f found by solving the RMP. The problem depicted in Fig. 8 is a shortest-path problem, where the objective is given by Eq. (A1) and additional costs for each edge, i.e., flight connection. The problem becomes an RCSP problem when we introduce cumulative constraints as resources, where a resource is a value accumulated throughout the route and is required to not go above some limit, hence giving a resource constraint. Resources and subsequently resource constraints are introduced for each maintenance type where a requirement can be given by the maximum flying hours that are allowed prior to a check or the maximum number of landings an aircraft is allowed to make prior to a check.

The task is then to find the shortest path with respect to the reduced costs in the network and the resource constraints. The PP is NP-hard where, for example, a label-setting algorithm [63] can be applied to solve instances.

3. Branch-and-price and fixing heuristics

Since only integral solutions are accepted for ILPs (and the original formulation of tail assignment), the column generation algorithm is typically augmented to branch-and-price [6], by combining column generation and branch-and-bound. In branch-and-price, we add an additional branching step, which occurs when no columns with a negative reduced cost can be found via solving the PP and the optimal solution is not integral. The fractional solution from the column generation provides a lower bound, if we are considering a minimization problem, as the algorithm solves the LP relaxed subproblem in branch-and-bound. In the branching step, the search space is partitioned, where the column generation algorithm is executed for each subproblem created. Therefore, we point out that branch-and-price can be thought of as branch-and-bound where column generation is utilized as a subroutine to compute bounds, show infeasibility or find an integer solution.

Moreover, Grönkvist [3] noticed that branch-and-price might be unnecessarily slow when applied to tail assignment and introduced a fixing heuristic where the branching step is replaced. The fixing heuristic finds the variable x_i closest to 1 and fixes it to 1, which forces the corresponding route to be part of the solution. It can be noted that the difference between the fixing heuristic and the typical branching is that the search space is restricted and not partitioned, meaning that the fixing heuristic is a dive into a specific branch of the search tree. Additional backtracking methods are utilized but are beyond this section's scope where such further information can be found in Ref. [3]. We denote the modified branch-and-price algorithm as the heuristic branch-and-price and depict the algorithm with the blue and green colored boxes with dotted and dashed borders in Fig. 1 subject to be integrated with a quantum algorithm in Sec. III.

APPENDIX B: MAPPING PROBLEMS TO THE ISING SPIN GLASS MODEL

If we consider the set partitioning problem in Eqs. (1)–(3) and apply a quadratic penalty on the constraints we obtain a nonlinear integer optimization problem. If we further assume constants $\mu_1 \in \{\mathbb{Z}^+ \cup \{0\}\}$, $\mu_2 \in \mathbb{Z}^+$ that balance the objective function and the constraints we obtain an alternative optimization problem

$$\min. \mu_1 \sum_{r \in R} c_r x_r + \mu_2 \sum_{f \in F} \left(\left[\sum_{r \in R} a_{fr} x_r \right] - 1 \right)^2, \quad (\text{B1})$$

$$\text{such that} \quad x_r \in \{0, 1\} \quad \forall r \in R. \quad (\text{B2})$$

The alternative optimization problem in Eqs. (B1) and (B2) can subsequently be modified to have variables $s_r \in$

$\{-1, 1\}$ by replacing the variables $x_r = 1 + s_r/2$, as presented by Lucas for several combinatorial optimization problems [48]. The variable change results in the following classical Hamiltonian:

$$\begin{aligned} H(s_1, \dots, s_{|R|}) &= \mu_1 \times \sum_{r \in R} c_r \frac{1 + s_r}{2} \\ &+ \mu_2 \times \sum_{f \in F} \left(\left[\sum_{r \in R} a_{fr} \frac{1 + s_r}{2} \right] - 1 \right)^2 \\ &= \mu_1 H^{\text{objective}}(s_1, \dots, s_{|R|}) \\ &+ \mu_2 H^{\text{exact cover}}(s_1, \dots, s_{|R|}), \end{aligned}$$

which we expand separately for the objective Hamiltonian and the exact cover Hamiltonian, where the exact cover Hamiltonian can be referred to as the constraint Hamiltonian. For the objective part we obtain

$$\begin{aligned} H^{\text{objective}}(s_1, \dots, s_{|R|}) &= \sum_{r \in R} h_r^{\text{objective}} s_r + \sum_{r' > r} J_{rr'}^{\text{objective}} s_r s_{r'} \\ &= \sum_{r \in R} \frac{c_r}{2} s_r + \sum_{r \in R} \frac{c_r}{2} = \sum_{r \in R} \frac{c_r}{2} s_r, \end{aligned}$$

by ignoring the constant energy shift. Thus

$$\begin{aligned} h_r^{\text{objective}} &= \frac{c_r}{2}, \\ J_{rr'}^{\text{objective}} &= 0. \end{aligned}$$

For the constraints, i.e., the exact cover Hamiltonian, it was showed in Ref. [24] that the classical Hamiltonian takes the form

$$\begin{aligned} H^{\text{exact cover}}(s_1, \dots, s_{|R|}) &= \sum_{r \in R} h_r^{\text{exact cover}} s_r \\ &+ \sum_{r' > r} J_{rr'}^{\text{exact cover}} s_r s_{r'}, \end{aligned}$$

where

$$\begin{aligned} h_r^{\text{exact cover}} &= \sum_{f \in F} a_{fr} \left(\sum_{r' \in R} \frac{a_{fr'}}{2} - 1 \right), \\ J_{rr'}^{\text{exact cover}} &= \sum_{f \in F} \frac{a_{fr} a_{fr'}}{2}. \end{aligned}$$

For the set partitioning problem we then obtain the following Hamiltonian:

$$\begin{aligned} H^{\text{set partitioning}}(s_1, \dots, s_{|R|}) \\ = \sum_{r \in R} [\mu_1 \times h_r^{\text{objective}} + \mu_2 \times h_r^{\text{exact cover}}] s_r \\ + \mu_2 \times \sum_{r' > r} J_{rr'}^{\text{exact cover}} s_r s_{r'}. \end{aligned}$$

Finally, the quantum Hamiltonian is obtained by promoting s_r to $\hat{\sigma}_r^z$

$$\begin{aligned} \hat{H}^{\text{set partitioning}}(\hat{\sigma}_1^z, \dots, \hat{\sigma}_{|R|}^z) \\ = \sum_{r \in R} [\mu_1 \times h_r^{\text{objective}} + \mu_2 \times h_r^{\text{exact cover}}] \hat{\sigma}_r^z \\ + \mu_2 \times \sum_{r' > r} J_{rr'}^{\text{exact cover}} \hat{\sigma}_r^z \hat{\sigma}_{r'}^z. \end{aligned}$$

It can be noted that the mapping holds for any ILP of the form

$$\begin{aligned} \text{minimize } & \sum_{r \in R} c_r x_r, \\ \text{subject to } & \sum_{r \in R} a_{fr} x_r = b_f \quad \forall f \in F, \\ & x_r \in \{0, 1\} \quad \forall r \in R, \end{aligned}$$

if $h_r^{\text{exact cover}}$ is modified to

$$h_r^{\text{exact cover}} = \sum_{f \in F} a_{fr} \left(\sum_{r' \in R} \frac{a_{fr'}}{2} - b_f \right).$$

APPENDIX C: EXPECTATION VALUE FOR ALGORITHM DEPTH ONE

In this section we derive the expression of the expectation value in Eq. (10) for algorithm depth $p = 1$ of QAOA. The expectation value of a general Ising spin glass Hamiltonian $\hat{H} = \sum_{i=1}^n h_i \hat{\sigma}_i^z + \sum_{(i,j) \in E} J_{ij} \hat{\sigma}_i^z \hat{\sigma}_j^z$ associated to an undirected graph $G = (V, E)$ with $n = |V|$ nodes and $|E|$ edges can be computed accordingly

$$\langle E \rangle = \text{Tr}[\rho \hat{H}] = \sum_{i=1}^n h_i \text{Tr}[\rho \hat{\sigma}_i^z] + \sum_{(i,j) \in E} J_{ij} \text{Tr}[\rho \hat{\sigma}_i^z \hat{\sigma}_j^z].$$

The undirected graph G has no self loops, which means that no edge (i, i) is present in the graph. We furthermore consider the edge (i, j) as identical to edge (j, i) and the sum over edges thus include the edge between node i and j exactly once. In other words, the edges are unordered pairs that connect the two nodes without a particular direction,

hence in graph G that we consider (j, i) is simply another way of referring to edge (i, j) , which means that $J_{ij} = J_{ji}$. The density matrix in the expression for the expectation value is $\rho = U_M(\beta) U_c(\gamma) |+\rangle \langle +| U_c^\dagger(\gamma) U_M^\dagger(\beta)$, where QAOA operators are defined as

$$U_M(\beta) = \prod_{i=1}^n e^{-i\beta \hat{\sigma}_i^x},$$

$$U_c(\gamma) = \prod_{i=1}^n e^{-i\gamma h_i \hat{\sigma}_i^z} \prod_{(i,j) \in E} e^{-i\gamma J_{ij} \hat{\sigma}_i^z \hat{\sigma}_j^z} = U_c^1(\gamma) U_c^2(\gamma).$$

We can rewrite the expectation value as

$$\begin{aligned} \langle E \rangle &= \sum_{i=1}^n h_i \text{Tr} \left[|+\rangle \langle +| U_c^\dagger(\gamma) U_M^\dagger(\beta) \hat{\sigma}_i^z U_M(\beta) U_c(\gamma) \right] \\ &+ \sum_{(i,j) \in E} J_{ij} \text{Tr} \left[|+\rangle \langle +| U_c^\dagger(\gamma) U_M^\dagger(\beta) \right. \\ &\quad \left. \times \hat{\sigma}_i^z \hat{\sigma}_j^z U_M(\beta) U_c(\gamma) \right] \\ &= \sum_{i=1}^n h_i \langle E_i \rangle + \sum_{(i,j) \in E} J_{ij} \langle E_{ij} \rangle \end{aligned}$$

by the cyclic property of the trace. We remark that partial terms $\text{Tr} [|+\rangle \langle +| \hat{a}]$ of the expectation value contribute if \hat{a} is a combination of $\hat{\sigma}^x$ and/or $\mathbb{1}$. The resulting value for terms $\langle E_i \rangle$ and $\langle E_{ij} \rangle$ have been derived for triangle free graphs in Ref. [53]. However, the resulting value for a graph with triangles was shown via MATHEMATICA in Ref. [55]. In this section we show the same general form of $\langle E_{ij} \rangle$ by analytical means.

We begin by considering some edge (i, j) , clearly all terms in $U_M(\beta)$ commute with $\hat{\sigma}_i^z \hat{\sigma}_j^z$ except for $e^{-i\beta \hat{\sigma}_i^x}$ and $e^{-i\beta \hat{\sigma}_j^x}$. We use the following relation:

$$\begin{aligned} F(\hat{a}, \eta \hat{b}) &= e^{\eta \hat{b}} \hat{a} e^{-\eta \hat{b}} \\ &= c_\eta^2 \hat{a} + s_\eta^2 \hat{b} \hat{a} \hat{b} + i \frac{s_{2\eta}}{2} [\hat{b}, \hat{a}], \end{aligned} \quad (\text{C1})$$

where $c_x^y = \cos^y(x)$ and $s_x^y = \sin^y(x)$ for convenience. The terms resulting from the mixing operator $U_M(\gamma)$ are therefore

$$\begin{aligned} U_M^\dagger(\beta) \hat{\sigma}_i^z \hat{\sigma}_j^z U_M(\beta) &= F(\hat{\sigma}_i^z, \beta \hat{\sigma}_i^x) F(\hat{\sigma}_j^z, \beta \hat{\sigma}_j^x) \\ &= c_{2\beta}^2 \hat{\sigma}_i^z \hat{\sigma}_j^z \\ &\quad + c_{2\beta} s_{2\beta} [\hat{\sigma}_i^z \hat{\sigma}_j^y + \hat{\sigma}_i^y \hat{\sigma}_j^z] \\ &\quad + s_{2\beta}^2 \hat{\sigma}_i^y \hat{\sigma}_j^y \end{aligned}$$

by evaluating Eq. (C1). For the ease of future derivations, we separate these parts as $\langle E_{ij} \rangle = \langle E_{ij}^{zz} \rangle + \langle E_{ij}^{zy} \rangle +$

$\langle E_{ij}^{yz} \rangle + \langle E_{ij}^{yy} \rangle$ where the sinus and cosinus terms are temporarily ignored. To clarify, here we defined $\langle E_{ij}^{ab} \rangle = \text{Tr} \left[|+\rangle \langle +| U_c^\dagger(\gamma) \hat{\sigma}_i^a \hat{\sigma}_j^b U_c(\gamma) \right]$.

We note that $[U_c(\gamma), \hat{\sigma}_i^z \hat{\sigma}_j^z] = 0$ and hence $\langle E_{ij}^{zz} \rangle$ does not contribute to the overall expectation value $\langle E \rangle$. However, for $\hat{\sigma}_i^y \hat{\sigma}_j^z$ all terms in $U_c^2(\gamma)$ with operators corresponding to edges with node i contribute, i.e., $[e^{-i\gamma J_{kp} \hat{\sigma}_k^z \hat{\sigma}_p^z}, \hat{\sigma}_i^y \hat{\sigma}_j^z] \neq 0 \forall p : (k=i, p) \in E$ and only $[e^{-i\gamma h_i \hat{\sigma}_i^z}, \hat{\sigma}_i^y \hat{\sigma}_j^z] \neq 0$ of all operator terms corresponding to nodes in $U_c^1(\gamma)$. Similarly for $\hat{\sigma}_i^z \hat{\sigma}_j^y$, all operator terms in $U_c^2(\gamma)$ corresponding to edges that include node j and the term $e^{-i\gamma h_j \hat{\sigma}_j^z}$ in $U_c^1(\gamma)$ contribute. For $\hat{\sigma}_i^y \hat{\sigma}_j^y$ we note that all operator terms in $U_c^2(\gamma)$ for edges that include node i or j contributes as $[e^{-i\gamma J_{kp} \hat{\sigma}_k^z \hat{\sigma}_p^z}, \hat{\sigma}_i^y \hat{\sigma}_j^y] \neq 0 \forall p \neq i : (k=i, p) \in E$ and $\forall p \neq j : (k=i, p) \in E$. Furthermore, both terms in $U_c^1(\gamma)$ that correspond to node i and j contribute to the expectation value as well.

We now wish to evaluate the terms $\langle E_{ij}^{yy} \rangle$, $\langle E_{ij}^{zy} \rangle$, and $\langle E_{ij}^{yz} \rangle$. We begin with the most complex case, $\langle E_{ij}^{yz} \rangle$, which is the only term that changes if triangles are present in the graph compared to the expression given in Ref. [53]. Since the only terms in $U_c^1(\gamma)$ that do not commute with $\hat{\sigma}_i^y \hat{\sigma}_j^y$ are $e^{-i\gamma h_i \hat{\sigma}_i^z}$ and $e^{-i\gamma h_j \hat{\sigma}_j^z}$, the terms that we obtain from $U_c^1(\gamma)$ are thus

$$\begin{aligned} U_c^{\dagger 1}(\gamma) \hat{\sigma}_i^y \hat{\sigma}_j^y U_c^1(\gamma) &= F(\hat{\sigma}_i^y, \gamma h_i \hat{\sigma}_i^z) F(\hat{\sigma}_j^y, \gamma h_j \hat{\sigma}_j^z) \\ &= c_{2h_i\gamma} c_{2h_j\gamma} \hat{\sigma}_i^y \hat{\sigma}_j^y + s_{2h_i\gamma} c_{2h_j\gamma} \hat{\sigma}_i^x \hat{\sigma}_j^y \\ &\quad + c_{2h_i\gamma} s_{2h_j\gamma} \hat{\sigma}_i^y \hat{\sigma}_j^x + s_{2h_i\gamma} s_{2h_j\gamma} \hat{\sigma}_i^x \hat{\sigma}_j^x, \end{aligned}$$

which gives us four terms to consider. We will now use the following relation:

$$\begin{aligned} G(\hat{c}, \eta_a \hat{a}, \eta_b \hat{b}) &= e^{i\eta_a \hat{a}} e^{i\eta_b \hat{b}} \hat{c} e^{-i\eta_b \hat{b}} e^{-i\eta_a \hat{a}} \\ &= e^{i\eta_a \hat{a}} (c_{\eta_b}^2 \hat{c} + s_{\eta_b}^2 \hat{b} \hat{c} \hat{b} + i \frac{S_{2\eta_b}}{2} [\hat{b}, \hat{c}]) e^{-i\eta_a \hat{a}} \\ &= c_{\eta_b}^2 [c_{\eta_a}^2 \hat{c} + s_{\eta_a}^2 \hat{a} \hat{c} \hat{a} + i \frac{S_{2\eta_a}}{2} [\hat{a}, \hat{c}]] \\ &\quad + s_{\eta_b}^2 [c_{\eta_a}^2 \hat{b} \hat{c} \hat{b} + s_{\eta_a}^2 \hat{a} (\hat{b} \hat{c} \hat{b}) \hat{a} + i \frac{S_{2\eta_a}}{2} [\hat{a}, \hat{b} \hat{c} \hat{b}]] \\ &\quad + i \frac{S_{2\eta_b}}{2} [c_{\eta_a}^2 [\hat{b}, \hat{c}] + s_{\eta_a}^2 \hat{a} [\hat{b}, \hat{c}] \hat{a} + i \frac{S_{2\eta_a}}{2} [\hat{a}, [\hat{b}, \hat{c}]]] \end{aligned} \quad (\text{C2})$$

to evaluate the contributing terms when $U_c^2(\gamma)$ is applied. It then becomes clear that the operators $\hat{\sigma}_i^x \hat{\sigma}_j^y$ and $\hat{\sigma}_i^y \hat{\sigma}_j^x$ do not contribute to $\langle E_{ij}^{yz} \rangle$ as there are no terms that can result in a pure $\hat{\sigma}^x$ and/or $\mathbb{1}$ combination when evaluating Eq. (C2). We explicitly give the expressions for all terms

below for the case of $\hat{\sigma}_i^y \hat{\sigma}_j^y$ (which by symmetry also allow us to throw away $\hat{\sigma}_i^y \hat{\sigma}_j^x$):

$$\begin{aligned} \hat{a} &= \hat{\sigma}_i^z \hat{\sigma}_k^z, & \hat{b} &= \hat{\sigma}_j^z \hat{\sigma}_p^z, & \hat{c} &= \hat{\sigma}_i^x \hat{\sigma}_j^y, \\ \hat{a} \hat{c} \hat{a} &\propto \hat{\sigma}_i^x \hat{\sigma}_j^y, & [\hat{a}, \hat{c}] &\propto \hat{\sigma}_i^y \hat{\sigma}_j^y \hat{\sigma}_k^z, & \hat{b} \hat{c} \hat{b} &\propto \hat{\sigma}_i^x \hat{\sigma}_j^y, \\ \hat{a} (\hat{b} \hat{c} \hat{b}) \hat{a} &\propto \hat{\sigma}_i^x \hat{\sigma}_j^y, & [\hat{a}, \hat{b} \hat{c} \hat{b}] &\propto \hat{\sigma}_i^y \hat{\sigma}_j^y \hat{\sigma}_k^z, & [\hat{b}, \hat{c}] &\propto \hat{\sigma}_i^x \hat{\sigma}_j^y \hat{\sigma}_p^z, \\ \hat{a} [\hat{b}, \hat{c}] \hat{a} &\propto \begin{cases} \hat{\sigma}_i^x \hat{\sigma}_j^x \hat{\sigma}_k^z & \text{if } k=p \\ \hat{\sigma}_i^x \hat{\sigma}_j^x \hat{\sigma}_p^z & \text{else} \end{cases}, \\ [\hat{a}, [\hat{b}, \hat{c}]] &\propto \begin{cases} \hat{\sigma}_i^y \hat{\sigma}_j^x & \text{if } k=p \\ \hat{\sigma}_i^y \hat{\sigma}_j^x \hat{\sigma}_k^z \hat{\sigma}_p^z & \text{else} \end{cases}. \end{aligned}$$

This means that we need only to consider $\hat{\sigma}_i^y \hat{\sigma}_j^y$ and $\hat{\sigma}_i^x \hat{\sigma}_j^x$. We can rewrite the contributing terms as

$$\frac{1}{2} \left(c_{2(h_i-h_j)\gamma} [\hat{\sigma}_i^y \hat{\sigma}_j^y + \hat{\sigma}_i^x \hat{\sigma}_j^x] - c_{2(h_i+h_j)\gamma} [\hat{\sigma}_i^x \hat{\sigma}_j^x - \hat{\sigma}_i^y \hat{\sigma}_j^y] \right)$$

by using the relations

$$s_{2h_i\gamma} s_{2h_j\gamma} = \frac{1}{2} [c_{2(h_i-h_j)\gamma} - c_{2(h_i+h_j)\gamma}], \quad (\text{C3})$$

$$c_{2h_i\gamma} c_{2h_j\gamma} = \frac{1}{2} [c_{2(h_i-h_j)\gamma} + c_{2(h_i+h_j)\gamma}]. \quad (\text{C4})$$

For the contributing terms we now consider how terms in $U_c^2(\gamma)$ corresponding to a triangle (i, j, p) act on $\hat{\sigma}_i^y \hat{\sigma}_j^y$ and $\hat{\sigma}_i^x \hat{\sigma}_j^x$. By evaluating the function in Eq. (C2) for $\hat{\sigma}_i^x \hat{\sigma}_j^x$ we get the following contributing terms:

$$\begin{aligned} G(\hat{\sigma}_i^x \hat{\sigma}_j^x, \gamma J_{ip} \hat{\sigma}_i^z \hat{\sigma}_p^z, \gamma J_{jp} \hat{\sigma}_j^z \hat{\sigma}_p^z) &= c_{2\gamma J_{jp}} c_{2\gamma J_{ip}} \hat{\sigma}_i^x \hat{\sigma}_j^x \\ &\quad + s_{2\gamma J_{jp}} s_{2\gamma J_{ip}} \hat{\sigma}_i^y \hat{\sigma}_j^y \end{aligned}$$

since

$$\begin{aligned} \hat{a} &= \hat{\sigma}_i^z \hat{\sigma}_p^z, & \hat{b} &= \hat{\sigma}_j^z \hat{\sigma}_p^z, & \hat{c} &= \hat{\sigma}_i^x \hat{\sigma}_j^x, \\ \hat{a} \hat{c} \hat{a} &= -\hat{\sigma}_i^x \hat{\sigma}_j^x, & [\hat{a}, \hat{c}] &= 2i \hat{\sigma}_i^y \hat{\sigma}_j^y \hat{\sigma}_p^z, & \hat{b} \hat{c} \hat{b} &= -\hat{\sigma}_i^x \hat{\sigma}_j^x, \\ \hat{a} (\hat{b} \hat{c} \hat{b}) \hat{a} &= \hat{\sigma}_i^x \hat{\sigma}_j^x, & [\hat{a}, \hat{b} \hat{c} \hat{b}] &= -2i \hat{\sigma}_i^y \hat{\sigma}_j^y \hat{\sigma}_p^z, \\ [\hat{b}, \hat{c}] &= 2i \hat{\sigma}_i^x \hat{\sigma}_j^y \hat{\sigma}_p^z, \end{aligned}$$

$$\hat{a} [\hat{b}, \hat{c}] \hat{a} = -2i \hat{\sigma}_i^x \hat{\sigma}_j^y \hat{\sigma}_p^z, \quad [\hat{a}, [\hat{b}, \hat{c}]] = -4 \hat{\sigma}_i^y \hat{\sigma}_j^y.$$

The final expression for $\hat{\sigma}_i^x \hat{\sigma}_j^x$ is found by the relation $c_{2a} = (c_a^2 - s_a^2)$. For $\hat{\sigma}_i^y \hat{\sigma}_j^y$ we get the following contributing

terms:

$$G(\hat{\sigma}_i^y \hat{\sigma}_j^y, \gamma J_{ip} \hat{\sigma}_i^z \hat{\sigma}_p^z, \gamma J_{jp} \hat{\sigma}_j^z \hat{\sigma}_p^z) = c_{2\gamma J_{ip}} c_{2\gamma J_{jp}} \hat{\sigma}_i^y \hat{\sigma}_j^y + s_{2\gamma J_{ip}} s_{2\gamma J_{jp}} \hat{\sigma}_i^x \hat{\sigma}_j^x.$$

by inspecting the partial terms we obtain by considering Eq. (C2):

$$\begin{aligned} \hat{a} &= \hat{\sigma}_i^z \hat{\sigma}_p^z, & \hat{b} &= \hat{\sigma}_j^z \hat{\sigma}_p^z, & \hat{c} &= \hat{\sigma}_i^y \hat{\sigma}_j^y, \\ \hat{a}\hat{c}\hat{a} &= -\hat{\sigma}_i^y \hat{\sigma}_j^y, & [\hat{a}, \hat{c}] &= -2i\hat{\sigma}_i^x \hat{\sigma}_j^y \hat{\sigma}_p^z, & \hat{b}\hat{c}\hat{b} &= -\hat{\sigma}_i^y \hat{\sigma}_j^y, \\ \hat{a}(\hat{b}\hat{c}\hat{b})\hat{a} &= \hat{\sigma}_i^y \hat{\sigma}_j^y, & [\hat{a}, \hat{b}\hat{c}\hat{b}] &= 2i\hat{\sigma}_i^x \hat{\sigma}_j^y \hat{\sigma}_p^z, & [\hat{b}, \hat{c}] &= \\ &= -2i\hat{\sigma}_i^y \hat{\sigma}_j^x \hat{\sigma}_p^z, \\ \hat{a}[\hat{b}, \hat{c}]\hat{a} &= 2i\hat{\sigma}_i^y \hat{\sigma}_j^x \hat{\sigma}_p^z, & [\hat{a}, [\hat{b}, \hat{c}]] &= -4\hat{\sigma}_i^x \hat{\sigma}_j^x. \end{aligned}$$

By using trigonometric relations

$$c_{a+b} = c_a c_b - s_a s_b, \quad (C5)$$

$$c_{a-b} = c_a c_b + s_a s_b, \quad (C6)$$

it is clear that the operators corresponding to a triangle (i, j, p) act on $\hat{\sigma}_i^y \hat{\sigma}_j^y + \hat{\sigma}_i^x \hat{\sigma}_j^x$ and $\hat{\sigma}_i^x \hat{\sigma}_j^x - \hat{\sigma}_i^y \hat{\sigma}_j^y$ as

$$\begin{aligned} G(\hat{\sigma}_i^y \hat{\sigma}_j^y + \hat{\sigma}_i^x \hat{\sigma}_j^x, \gamma J_{ip} \hat{\sigma}_i^z \hat{\sigma}_p^z, \gamma J_{jp} \hat{\sigma}_j^z \hat{\sigma}_p^z) \\ &= c_{2\gamma(J_{ip}-J_{jp})} [\hat{\sigma}_i^y \hat{\sigma}_j^y + \hat{\sigma}_i^x \hat{\sigma}_j^x], \\ G(\hat{\sigma}_i^x \hat{\sigma}_j^x - \hat{\sigma}_i^y \hat{\sigma}_j^y, \gamma J_{ip} \hat{\sigma}_i^z \hat{\sigma}_p^z, \gamma J_{jp} \hat{\sigma}_j^z \hat{\sigma}_p^z) \\ &= c_{2\gamma(J_{ip}+J_{jp})} [\hat{\sigma}_i^x \hat{\sigma}_j^x - \hat{\sigma}_i^y \hat{\sigma}_j^y]. \end{aligned}$$

Furthermore, other terms in $U_c^2(\gamma)$ corresponding to edges (i, k) and (j, l) that are not part of a triangle will give rise to the following expressions for $\hat{\sigma}_i^x \hat{\sigma}_j^x$:

$$\begin{aligned} F(\hat{\sigma}_i^x \hat{\sigma}_j^x, \gamma J_{ik} \hat{\sigma}_i^z \hat{\sigma}_k^z) &= c_{2\gamma J_{ik}} \hat{\sigma}_i^x \hat{\sigma}_j^x - s_{2\gamma J_{ik}} \hat{\sigma}_i^y \hat{\sigma}_j^y \hat{\sigma}_k^z, \\ F(\hat{\sigma}_i^x \hat{\sigma}_j^x, \gamma J_{jl} \hat{\sigma}_j^z \hat{\sigma}_l^z) &= c_{2\gamma J_{jl}} \hat{\sigma}_i^x \hat{\sigma}_j^x - s_{2\gamma J_{jl}} \hat{\sigma}_i^y \hat{\sigma}_j^y \hat{\sigma}_l^z \end{aligned}$$

and for $\hat{\sigma}_i^y \hat{\sigma}_j^y$ we get

$$\begin{aligned} F(\hat{\sigma}_i^y \hat{\sigma}_j^y, \gamma J_{ik} \hat{\sigma}_i^z \hat{\sigma}_k^z) &= c_{2\gamma J_{ik}} \hat{\sigma}_i^y \hat{\sigma}_j^y + s_{2\gamma J_{ik}} \hat{\sigma}_i^x \hat{\sigma}_j^x \hat{\sigma}_k^z, \\ F(\hat{\sigma}_i^y \hat{\sigma}_j^y, \gamma J_{jl} \hat{\sigma}_j^z \hat{\sigma}_l^z) &= c_{2\gamma J_{jl}} \hat{\sigma}_i^y \hat{\sigma}_j^y + s_{2\gamma J_{jl}} \hat{\sigma}_i^x \hat{\sigma}_j^x \hat{\sigma}_l^z. \end{aligned}$$

Thus, for $\langle E_{ij}^{yy} \rangle$ we get the following contributing parts to the expectation value:

$$\begin{aligned} \langle E_{ij}^{yy} \rangle &= \frac{1}{2} s_{2\beta}^2 \prod_{\substack{(i,k) \in E \\ (j,k) \notin E}} c_{2\gamma J_{ik}} \prod_{\substack{(j,l) \in E \\ (i,l) \notin E}} c_{2\gamma J_{jl}} \\ &\times \left[c_{2(h_i-h_j)\gamma} \prod_{\substack{(i,p) \in E \\ (j,p) \in E}} c_{2(J_{ip}-J_{jp})\gamma} \right. \\ &\quad \left. - c_{2(h_i+h_j)\gamma} \prod_{\substack{(j,p) \in E \\ (i,p) \in E}} c_{2(J_{ip}+J_{jp})\gamma} \right] \quad (C7) \end{aligned}$$

since only the $\hat{\sigma}_i^x \hat{\sigma}_j^x$ terms are nonzero when the trace is taken.

For $\langle E_{ij}^{yz} \rangle$, on the other hand, the only noncommuting term of $U_c^1(\gamma)$ with $\hat{\sigma}_i^y \hat{\sigma}_j^z$ is $e^{-i\gamma h_i \hat{\sigma}_i^z}$, which results in

$$F(\hat{\sigma}_i^y \hat{\sigma}_j^z, \gamma h_i \hat{\sigma}_i^z) = c_{2\gamma h_i} \hat{\sigma}_i^y \hat{\sigma}_j^z + s_{2\gamma h_i} \hat{\sigma}_i^x \hat{\sigma}_j^z.$$

For the operator, $e^{-iJ_{ij}\gamma\hat{\sigma}_i^z\hat{\sigma}_j^z}$, corresponding to edge (i, j) , we have that

$$\begin{aligned} F(\hat{\sigma}_i^y \hat{\sigma}_j^z, \gamma J_{ij} \hat{\sigma}_i^z \hat{\sigma}_j^z) &= c_{2\gamma J_{ij}} \hat{\sigma}_i^y \hat{\sigma}_j^z + s_{2\gamma J_{ij}} \hat{\sigma}_i^x, \\ F(\hat{\sigma}_i^x \hat{\sigma}_j^z, \gamma J_{ij} \hat{\sigma}_i^z \hat{\sigma}_j^z) &= c_{2\gamma J_{ij}} \hat{\sigma}_i^x \hat{\sigma}_j^z - s_{2\gamma J_{ij}} \hat{\sigma}_i^y. \end{aligned}$$

Other operators of $U_c^2(\gamma)$ corresponding to edges that include node i further gives the expressions

$$\begin{aligned} F(\hat{\sigma}_i^y \hat{\sigma}_j^z, \gamma J_{ip} \hat{\sigma}_i^z \hat{\sigma}_p^z) &= c_{2\gamma J_{ip}} \hat{\sigma}_i^y \hat{\sigma}_j^z + s_{2\gamma J_{ip}} \hat{\sigma}_i^x \hat{\sigma}_j^z \hat{\sigma}_p^z, \\ F(\hat{\sigma}_i^x, \gamma J_{ip} \hat{\sigma}_i^z \hat{\sigma}_p^z) &= c_{2\gamma J_{ip}} \hat{\sigma}_i^x - s_{2\gamma J_{ip}} \hat{\sigma}_i^y \hat{\sigma}_p^z, \\ F(\hat{\sigma}_i^x \hat{\sigma}_j^z, \gamma J_{ip} \hat{\sigma}_i^z \hat{\sigma}_p^z) &= c_{2\gamma J_{ip}} \hat{\sigma}_i^x \hat{\sigma}_j^z - s_{2\gamma J_{ip}} \hat{\sigma}_i^y \hat{\sigma}_j^z \hat{\sigma}_p^z, \\ F(\hat{\sigma}_i^y, \gamma J_{ip} \hat{\sigma}_i^z \hat{\sigma}_p^z) &= c_{2\gamma J_{ip}} \hat{\sigma}_i^y + s_{2\gamma J_{ip}} \hat{\sigma}_i^x \hat{\sigma}_p^z. \end{aligned}$$

We can again conclude that since only the $\hat{\sigma}_i^x$ term contributes here, the final contribution to the expectation value is

$$\langle E_{ij}^{yz} \rangle = \frac{s_{4\beta}}{2} c_{2\gamma h_i} s_{2\gamma J_{ij}} \prod_{p \neq i: (i,p) \in E} c_{2\gamma J_{ip}} \quad (C8)$$

and

$$\langle E_{ij}^{zy} \rangle = \frac{s_{4\beta}}{2} c_{2\gamma h_j} s_{2\gamma J_{ij}} \prod_{p \neq i: (j,p) \in E} c_{2\gamma J_{jp}} \quad (C9)$$

by exchanging index i and j .

Finally, for $\langle E_i \rangle$ the only noncommuting term of $U_M(\beta)$ is $e^{-i\beta\hat{\sigma}_i^x}$, which results in

$$F(\hat{\sigma}_i^z, e^{-i\beta\hat{\sigma}_i^x}) = c_{2\beta}\hat{\sigma}_i^z + s_{2\beta}\hat{\sigma}_i^y.$$

We can exclude $\hat{\sigma}_i^z$ as it commutes with $U_c(\gamma)$. On the other hand, $\hat{\sigma}_i^y$ does not commute with $e^{-i\gamma h_i \hat{\sigma}_i^z}$. The operator $U_c^1(\gamma)$ therefore gives rise to the following expression:

$$F(\hat{\sigma}_i^y, \gamma h_i \hat{\sigma}_i^z) = c_{2\gamma h_i} \hat{\sigma}_i^y + s_{2\gamma h_i} \hat{\sigma}_i^x.$$

When we act with $e^{-i\gamma J_{ip} \hat{\sigma}_i^z \hat{\sigma}_p^z}$ for an edge $p : (i, p) \in E$ we get

$$\begin{aligned} F(\hat{\sigma}_i^y, \gamma J_{ip} \hat{\sigma}_i^z \hat{\sigma}_p^z) &= c_{2\gamma J_{ip}} \hat{\sigma}_i^y + s_{2\gamma J_{ip}} \hat{\sigma}_i^x \hat{\sigma}_p^z, \\ F(\hat{\sigma}_i^x, \gamma J_{ip} \hat{\sigma}_i^z \hat{\sigma}_p^z) &= c_{2\gamma J_{ip}} \hat{\sigma}_i^x - s_{2\gamma J_{ip}} \hat{\sigma}_i^y \hat{\sigma}_p^z, \end{aligned}$$

which contributes only with $c_{2\gamma J_{ip}} \hat{\sigma}_i^x$. The resulting contribution of $\langle E_i \rangle$ to the overall expectation value is therefore

$$\langle E_i \rangle = s_{2\beta} s_{2\gamma h_i} \prod_{p:(i,p) \in E} c_{2\gamma J_{ip}}. \quad (\text{C10})$$

We conclude by noting that if we add all the terms in Eqs. (C7)–(C10), with their coefficients h_i and J_{ij} we get the expression in Eq. (10).

APPENDIX D: SUCCESS PROBABILITIES OF SET PARTITIONING

In this section, a summary is given of the results of ideal simulations of QAOA circuits for all instances applied to the set partitioning problem. Table I shows success probabilities for Hamiltonians constructed for factors $f = \infty$, f^* and intermediate choices. As a shorthand, $P_f^{\text{set partitioning}}$ denotes given a set partitioning Hamiltonian with weights μ_1 and μ_2 for a factor f .

[1] F. Quesnel, G. Desaulniers, and F. Soumis, A branch-and-price heuristic for the crew pairing problem with language constraints, *Eur. J. Oper. Res.* **283**, 1040 (2020).
 [2] M. Deveci and N. Çetin Demirel, A survey of the literature on airline crew scheduling, *Eng. Appl. Artif. Intell.* **74**, 54 (2018).
 [3] M. Grönkvist and J. Kjerrström, *Tail Assignment in Practice* (Springer Berlin Heidelberg, Berlin, Heidelberg, 2005).
 [4] M. Grönkvist and J. Kjerrström, *Tail Assignment in Practice* (Springer Berlin Heidelberg, 2005).
 [5] D. Wedelin, An algorithm for large scale 0-1 integer programming with application to airline crew scheduling, *Ann. Oper. Res.* **57**, 283 (1995).
 [6] C. Barnhart, E. Johnson, G. Nemhauser, M. Savelsbergh, and P. Vance, Branch-and-price: Column generation for solving huge integer programs, *Oper. Res.* **46**, 316 (1970).

[7] M. Conforti, G. Cornuéjols, and G. Zambelli, *Integer Programming, Graduate Texts in Mathematics* (Springer International Publishing, New York, 2014).
 [8] A. H. Land and A. G. Doig, An automatic method for solving discrete programming problems, *Econometrica* **28**, 497 (1960).
 [9] G. B. Dantzig and P. Wolfe, Decomposition principle for linear programs, *Oper. Res.* **8**, 101 (1960).
 [10] M. E. Lübbecke and J. Desrosiers, Selected topics in column generation, *Oper. Res.* **53**, 1007 (2005).
 [11] P. W. Shor, Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer, *SIAM J. Comput.* **26**, 1484 (1997).
 [12] L. K. Grover, in *Proceedings of the Twenty-Eighth Annual ACM Symposium on Theory of Computing*, STOC '96 (Association for Computing Machinery, New York, NY, USA, 1996), p. 212.
 [13] E. Farhi, J. Goldstone, S. Gutmann, J. Lapan, A. Lundgren, and D. Preda, A quantum adiabatic evolution algorithm applied to random instances of an NP-complete problem, *Science* **292**, 472 (2001).
 [14] T. Kadowaki and H. Nishimori, Quantum annealing in the transverse Ising model, *Phys. Rev. E* **58**, 5355 (1998).
 [15] E. Zahedinejad and A. Zaribafiyani, Combinatorial optimization on gate model quantum computers: A survey (2017), [ArXiv:1708.05294](https://arxiv.org/abs/1708.05294).
 [16] A. Montanaro, Quantum speedup of branch-and-bound algorithms, *Phys. Rev. Res.* **2**, 013056 (2020).
 [17] A. Gilliam, S. Woerner, and C. Gonciulea, Grover adaptive search for constrained polynomial binary optimization, *Quantum* **5**, 428 (2021).
 [18] E. Farhi, J. Goldstone, and S. Gutmann, A quantum approximate optimization algorithm (2014), [ArXiv:1411.4028](https://arxiv.org/abs/1411.4028).
 [19] E. Farhi and A. W. Harrow, Quantum supremacy through the quantum approximate optimization algorithm (2019), [ArXiv:1602.07674](https://arxiv.org/abs/1602.07674).
 [20] F. Arute, K. Arya, R. Babbush, D. Bacon, J. Bardin, R. Barends, R. Biswas, S. Boixo, F. Brandao, D. Buell, B. Burkett, Y. Chen, Z. Chen, B. Chiaro, R. Collins, W. Courtney, A. Dunsworth, E. Farhi, B. Foxen, and J. Martinis, Quantum supremacy using a programmable superconducting processor, *Nature* **574**, 505 (2019).
 [21] J. Preskill, Quantum computing in the NISQ era and beyond, *Quantum* **2**, 79 (2018).
 [22] M. P. Harrigan, K. J. Sung, M. Neeley, K. J. Satzinger, F. Arute, K. Arya, J. Atalaya, J. C. Bardin, R. Barends, and S. Boixo, *et al.*, Quantum approximate optimization of non-planar graph problems on a planar superconducting processor, *Nat. Phys.* **17**, 332 (2021).
 [23] M. R. Garey and D. S. Johnson, *Computers and Intractability; A Guide to the Theory of NP-Completeness* (W. H. Freeman & Co., USA, 1990).
 [24] P. Vikstål, M. Grönkvist, M. Svensson, M. Andersson, G. Johansson, and G. Ferrini, Applying the Quantum Approximate Optimization Algorithm to the Tail-Assignment Problem, *Phys. Rev. Appl.* **14**, 034009 (2020).
 [25] A. Bengtsson, P. Vikstål, C. Warren, M. Svensson, X. Gu, A. Frisk Kockum, P. Krantz, C. Križan, D. Shiri, I.-M. Svensson, G. Tancredi, G. Johansson, P. Delsing, G. Ferrini, and J. Bylander, Improved Success Probability with

- Greater Circuit Depth for the Quantum Approximate Optimization Algorithm, *Phys. Rev. Appl.* **14**, 034010 (2020).
- [26] D. Willsch, M. Willsch, C. D. Gonzalez Calaza, F. Jin, H. De Raedt, M. Svensson, and K. Michielsen, Benchmarking advantage and D-wave 2000Q quantum annealers with exact cover problems, *Quantum Inf. Process.* **21**, 141 (2022).
- [27] P. Belobaba, A. Odoni, and C. Barnhart, *The Global Airline Industry*, Aerospace Series (Wiley, 2009), Chap. 6-7, p. 153.
- [28] G. L. Nemhauser, L. A. Wolsey, eds., *Integer and Combinatorial Optimization* (John Wiley & Sons, Inc., 1988).
- [29] G. Desaulniers, J. Desrosiers, and M. M. Solomon, *Column Generation*. [Electronic Resource] (Springer US, New York, 2005).
- [30] L. S. Lasdon, *Optimization Theory for Large Systems* (Dover Publications, Mineola (N.Y), 2002).
- [31] M. Mezard and A. Montanari, *Information, physics, and computation* (Oxford University Press, Inc., USA, 2009), Chap. 18, p. 403.
- [32] A. V. Seliverstov, *Binary Solutions to Some Systems of Linear Equations* (Springer, Cham, 2018).
- [33] N. Livne, A note on #P-completeness of NP-witnessing relations, *Inf. Process. Lett.* **109**, 259 (2009).
- [34] Cplex, IBM ILOG: V12. 1: User's manual for cplex, International Business Machines Corporation **46**, 157 (2009).
- [35] E. Danna and C. Le Pape, in *Column Generation*, edited by G. Desaulniers, J. Desrosiers, and M. M. Solomon (Springer US, Boston, MA, 2005), p. 99.
- [36] T. Tran, M. Do, E. Rieffel, J. Frank, Z. Wang, B. O'Gorman, D. Venturelli, and J. Beck, in *SOCS* (2016).
- [37] C. Joncour, S. Michel, R. Sadykov, D. Sverdlov, and F. Vanderbeck, Column generation based primal heuristics, *Electron. Notes Discrete Math.* **36**, 695 (2010).
- [38] J. Núñez Ares, H. de Vries, and D. Huisman, A column generation approach for locating roadside clinics in Africa based on effectiveness and equity, *Eur. J. Oper. Res.* **254**, 1002 (2016).
- [39] S. Bravyi, A. Kliesch, R. Koenig, and E. Tang, Obstacles to Variational Quantum Optimization from Symmetry Protection, *Phys. Rev. Lett.* **125**, 260505 (2020).
- [40] M. Fischetti and A. Lodi, Local branching, *Math. Program.* **98**, 23 (2003).
- [41] D. Feillet, A tutorial on column generation and branch-and-price for vehicle routing problems, *4OR* **8**, 407 (2010).
- [42] F. Leymann and J. Barzen, The bitter truth about gate-based quantum algorithms in the NISQ era, *Quantum Sci. Technol.* **5**, 044007 (2020).
- [43] M.-T. Kong and N. Shah, Preprocessing rules for integer programming solutions to the generalised assignment problem, *J. Oper. Res. Soc.* **52**, 567 (2001).
- [44] T. Achterberg, R. Bixby, Z. Gu, E. Rothberg, and D. Weninger, Presolve reductions in mixed integer programming, *INFORMS J. Comput.* **32**, 473 (2019).
- [45] J. Nocedal and S. Wright, *Numerical Optimization*. [electronic resource]., Springer Series in Operations Research and Financial Engineering (Springer New York, 2006), Chap. 17, p. 497.
- [46] U. Azad, B. Behera, E. Ahmed, P. Panigrahi, and A. Farouk, Solving vehicle routing problem using quantum approximate optimization algorithm, *IEEE Trans. Intell. Transp. Syst.* **PP**, 1 (2022).
- [47] T. Stollenwerk, E. Lobe, and M. Jung, in *Quantum Technology and Optimization Problems*, edited by S. Feld and C. Linnhoff-Popien (Springer International Publishing, Cham, 2019), p. 99.
- [48] A. Lucas, Ising formulations of many NP problems, *Front. Phys.* **2**, 5 (2014).
- [49] S. Gharibian, Y. Huang, Z. Landau, and S. W. Shin, Quantum hamiltonian complexity, *Found. Trends® Theor. Comput. Sci.* **10**, 159 (2015).
- [50] M. Svensson, Extracted-dainstances-for-tailassignment, <https://github.com/marikasvenssonjeppesen/Extracted-dainstances-for-Tailassignment> (2021).
- [51] E. G. Rieffel, S. Hadfield, T. Hogg, S. Mandrà, J. Marshall, G. Mossi, B. O'Gorman, E. Plamadeala, N. M. Tubman, D. Venturelli, W. Vinci, Z. Wang, M. Wilson, F. Wudarski, and R. Biswas, From ansätze to z-gates: A NASA view of quantum computing, *Adv. Parallel Comput.* **34**, 133 (2019).
- [52] L. Bittel and M. Kliesch, Training Variational Quantum Algorithms is NP-Hard, *Phys. Rev. Lett.* **127**, 120502 (2021).
- [53] M. Willsch, D. Willsch, F. Jin, H. De Raedt, and K. Michielsen, Benchmarking the quantum approximate optimization algorithm, *Quantum Inf. Process.* **19**, 197 (2020).
- [54] Z. Wang, S. Hadfield, Z. Jiang, and E. G. Rieffel, Quantum approximate optimization algorithm for maxcut: A fermionic view, *Phys. Rev. A* **97**, 022304 (2018).
- [55] A. Ozaeta, W. van Dam, and P. L. McMahon, Expectation values from the single-layer quantum approximate optimization algorithm on Ising problems, *Quantum Sci. Technol.* **7**, 045036 (2022).
- [56] 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* **10**, 021067 (2020).
- [57] S. Harwood, C. Gambella, D. Tenev, A. Simonetto, D. Neira, and D. Greenberg, Formulating and solving routing problems on quantum computers, *IEEE Trans. Quantum Eng.* **PP**, 1 (2021).
- [58] A. Cho, IBM promises 1000-qubit quantum computer—a milestone—by 2023, *Science* (2020).
- [59] A. W. Harrow, Small quantum computers and large classical data sets (2020), [ArXiv:2004.00026](https://arxiv.org/abs/2004.00026).
- [60] D. R. Morrison, S. H. Jacobson, J. J. Sauppe, and E. C. Sewell, Branch-and-bound algorithms: A survey of recent advances in searching, branching, and pruning, *Discrete Optim.* **19**, 79 (2016).
- [61] C. E. Lemke, The dual method of solving the linear programming problem, *Naval Res. Logist. Q.* **1**, 36 (1954).
- [62] G. B. Dantzig, The general simplex method for minimizing a linear form under inequality constraints, *Pac. J. Math.* **5**, 183 (1955).
- [63] M. Desrochers and F. Soumis, A generalized permanent labeling algorithm for the shortest path problem with time windows, *Inf. Syst. Res. - ISR* **26**, 191 (1988).