Solving optimization problems with local light-shift encoding on Rydberg quantum annealers

Kapil Goswami [1,*] Rick Mukherjee,[1,†] Herwig Ott,[2] and Peter Schmelcher[1,3]

[1]*Zentrum für Optische Quantentechnologien, Universität Hamburg, Luruper Chaussee 149, 22761 Hamburg, Germany*
[2]*Department of Physics and Research Center OPTIMAS, Rheinland-Pfälzische Technische Universität Kaiserslautern-Landau, 67663 Kaiserslautern, Germany*
[3]*The Hamburg Centre for Ultrafast Imaging, Universität Hamburg, Luruper Chaussee 149, 22761 Hamburg, Germany*

We provide a non-unit-disk framework to solve combinatorial optimization problems such as maximum cut and maximum independent set on a Rydberg quantum annealer. Our setup consists of a many-body interacting Rydberg system where locally controllable light shifts are applied to individual qubits in order to map the graph problem onto the Ising spin model. Exploiting the flexibility that optical tweezers offer in terms of spatial arrangement, our numerical simulations implement the local-detuning protocol while globally driving the Rydberg annealer to the desired many-body ground state, which is also the solution to the optimization problem. Using optimal control methods, these solutions are obtained for prototype graphs with varying sizes at timescales well within the system lifetime and with approximation ratios close to one. The nonblockade approach facilitates the encoding of graph problems with specific topologies that can be realized in two-dimensional Rydberg configurations and is applicable to both unweighted as well as weighted graphs. A comparative analysis with fast simulated annealing is provided which highlights the advantages of our scheme in terms of system size, hardness of the graph, and the number of iterations required to converge to the solution.

I. INTRODUCTION

Using quantum computing for applications requires a fault-tolerant scalable architecture [1–5]. The current so-called noisy intermediate-scale quantum (NISQ) era is dictated by noisy qubits and gates with comparatively low fidelities [6,7]. Despite these limitations, there is ongoing research in trying to find suitable problems for which the quantum approach is superior [7–10]. In particular, there is a class of problems whose exact solutions are hard to obtain in polynomial time, and at best, in certain cases, there exist only approximate solutions to them. They fall under the category of NP-hard problems [11,12] and are regularly investigated beyond the realm of computational complexity theory in order to get better insight into the performance of quantum algorithms [13–15]. In this work, we consider two such NP-hard problems, namely, maximum cut (Max-Cut) and maximum independent set (MIS), which are combinatorial optimization problems that have many practical applications [16–20].

When solving a combinatorial optimization problem with quantum systems, there are two pertinent challenges that arise: (i) the choice of quantum hardware and (ii) the choice of

---
*kgoswami@physnet.uni-hamburg.de
†rick.mukherjee@physnet.uni-hamburg.de

quantum algorithm to be implemented. An ideal quantum hardware should allow scalability with the number of qubits. Qubits of superconducting devices still suffer from noise and rely on a fixed architecture with localized connectivity [21], while trapped ions are still too sensitive to external fields to allow scalability for more than 50 qubits [22,23]. Motivated by recent experiments where hundreds of neutral Rydberg atoms can be trapped and whose interactions can be manipulated [24–26], there is a tendency and perspective to use Rydberg platforms to solve optimization problems, which is also the focus of this work.

Very recent developments in solving optimization problems on Rydberg platforms include solving the weighted MIS graph problem on a Rydberg simulator [27] and the Max-Cut problem using Rydberg gates [15] whose approximation ratio suffers due to noisy gates. In both works, the quantum approximate optimization algorithm (QAOA) approach is used to approximate the optimal solution [28,29]. In general, variational methods require finding optimal values for a large number of variational parameters (e.g., initial state ansatz, choice of the mixer, and number of layers), which by itself is found to be an NP-hard problem [30,31]. This can make the implementation of variational algorithms to solve a particular optimization problem fairly cumbersome. For this purpose, we propose an optimized quantum annealing protocol that is implemented on the Rydberg platform.

Another issue is the choice of encoding which is a scheme by which a real-world optimization problem is mapped onto a system of interacting qubits. The encoding scheme depends on the choice of both hardware and algorithm. Max-Cut and MIS are optimization problems that can be represented in a

Quadratic Unconstrained Binary Optimization (QUBO) form. In this representation, the objective function along with its constraints has a quadratic form with variables that take binary values [32,33]. The advantage of formulating the problem in QUBO form is that it can naturally map to interacting spin models [34]. However, it should be mentioned that depending on the type of optimization problem, this mapping is neither unique nor trivial to be physically realized [27,35–38]. For example, in Ref. [27], the MIS problem was encoded on an interacting spin system using the unit-disk encoding [39] that relies on the perfect implementation of the Rydberg blockade effect [40] between the atoms. Based on the unit-disk encoding, a problem graph with N vertices will require $\sim 4N^2$ physical spins which can be a substantial overhead of resources [41].

In this work, by taking advantage of localized light shifts on individual atoms, we encode and solve the Max-Cut and MIS problems using a non-unit-disk mapping to spin models. The proposed scheme has the benefit of tackling weighted as well as unweighted graphs within the same framework and can in principle be generalized to other QUBO problems. By combining gradient and nongradient optimal control methods, we implement the temporal evolution of the laser parameters that make the annealing process on the Rydberg platform efficient to solve optimization problems. Thus we obtain optimal solutions for different graphs (with sizes 5–15, up to degree 5) with varying *hardness* within 3–5 μs with an approximation ratio close to one. The hardness parameter measures the complexity of the problem graph and serves as a convenient metric for benchmarking the performance of our optimal quantum annealing method. For graphs with similar size and complexity, we find that our protocol outperforms the simulated annealing algorithm [42] in terms of accuracy. Our encoding scheme should be applicable on any quantum platform that allows local qubit control along with global driving of the many-body system to its ground state.

We proceed as follows. In Sec. II we outline the theoretical description for the Rydberg simulator and introduce the Max-Cut and MIS problems. We propose a three-step scheme for non-unit-disk encoding and solving graph problems using the Rydberg annealer architecture in Sec. III. In Sec. IV we discuss the tools to characterize the complexity of the problems and the quality of the solutions. Finally, in Sec. V we analyze our results, and detail the physical implementation of our scheme in Sec. VI. Section VII contains our conclusions and outlook.

II. THEORY

In Sec. II A, we provide a brief introduction to the many-body Rydberg Hamiltonian which provides the setup for quantum annealing while Sec. II B defines the Max-Cut and MIS problems.

A. Hamiltonian of Rydberg simulator

We consider ultracold atoms trapped in optical tweezers [43,44]. They are described as two-level systems consisting of a ground state ($|g\rangle$) and a Rydberg excited state ($|e\rangle$) that are optically coupled using a laser [45,46]. The Hamiltonian

expressed in the atomic basis reads as follows:

$$
\hat{H}_{\mathrm{Ryd}} = \frac{\Omega}{2} \sum_j |e\rangle_j \langle g| + |g\rangle_j \langle e| - \sum_j \Delta_j |e\rangle_j \langle e|
$$
$$
+ \sum_{k<j} V_{kj}(|e\rangle_k \langle e| \otimes |e\rangle_j \langle e|), \tag{1}
$$

where $V_{kj} = C_6/|\mathbf{r}_j - \mathbf{r}_k|^6$ is the van der Waals ($C_6 > 0$) interaction between the pair of Rydberg atoms $|e\rangle_k$ and $|e\rangle_j$ where C_6 is the associated dispersion coefficient. Here $\mathbf{r}_j$ and $\mathbf{r}_k$ are the positions of the two atoms labeled as j and k. The detuning Δ_j is the site-dependent laser parameter describing the difference between the frequency of the applied field and the natural frequency associated with the atomic transition. The Rabi frequency Ω is the global laser parameter that couples the two states and is proportional to the intensity of the driving field and the dipole moment associated with the atomic transition. Using the Pauli spin operator representation and setting $\Omega = 0$, the above equation reduces to the Ising Hamiltonian with a longitudinal field (up to a constant $C = \frac{1}{4} \sum_{j=1}^{N-1} \sum_{k=j+1}^{N} V_{jk}$), given as

$$
\hat{H}_{\mathrm{Ising}} = \sum_{j=1}^{N} \left(\frac{\Delta_j}{2} + \frac{1}{4} \sum_{k=1, j\neq k}^{N} V_{jk} \right) \hat{\sigma}_j^z + \frac{1}{4} \sum_{j=1}^{N-1} \sum_{k=j+1}^{N} V_{jk} \hat{\sigma}_j^z \hat{\sigma}_k^z, \tag{2}
$$

where the negative sign is absorbed in Δ_j. The first term is the local longitudinal field which depends on the detuning as well as $\sum_{k=1, j\neq k}^{N} V_{jk}$, while the second term corresponds to the interaction between spins. The above Hamiltonian is useful for the connection to optimization problems that can be formulated in the QUBO framework.

B. Max-cut and MIS problem definitions

QUBO problems are often represented by a graph $G(V, E)$ with vertices V and edges E which take different representations depending on the problem, as shown in Fig. 1. Below we define the classical cost functions for the Max-Cut and MIS problems, respectively.

Max-Cut. Given a graph G with weights w_{jk} associated with each edge (j, k), a subset of vertices from the graph is referred to as a *cut*. This *cut* becomes a Max-Cut when the vertices are split into two sets in a way that maximizes the total weight (Max-Cut value) of edges between them. The problem graph is illustrated in Figs. 1(a) and 1(b). Figure 1(a) displays a weighted graph of size 4 and Fig. 1(b) shows a dashed curve representing the cut that divides the graph into two sets, blue and red. Both sets are interchangeable as the solution is degenerate. The Max-Cut cost function, which is to be maximized, is expressed as

$$
C_{\mathrm{Max\text{-}Cut}} = \sum_{(j,k)\in E} w_{j,k}(X_j(1 - X_k) + X_k(1 - X_j)), \tag{3}
$$

where the sum is over all the edges in the graph G. The variable $X_j \in \{0, 1\}$ represents the two sets in a cut.

MIS. Given a graph G with weights w_j associated with each vertex j, a subset of mutually nonadjacent vertices is said to be a *maximum independent set* if it has the largest possible sum of weights w_j over all vertices in the subset. Figures 1(d)
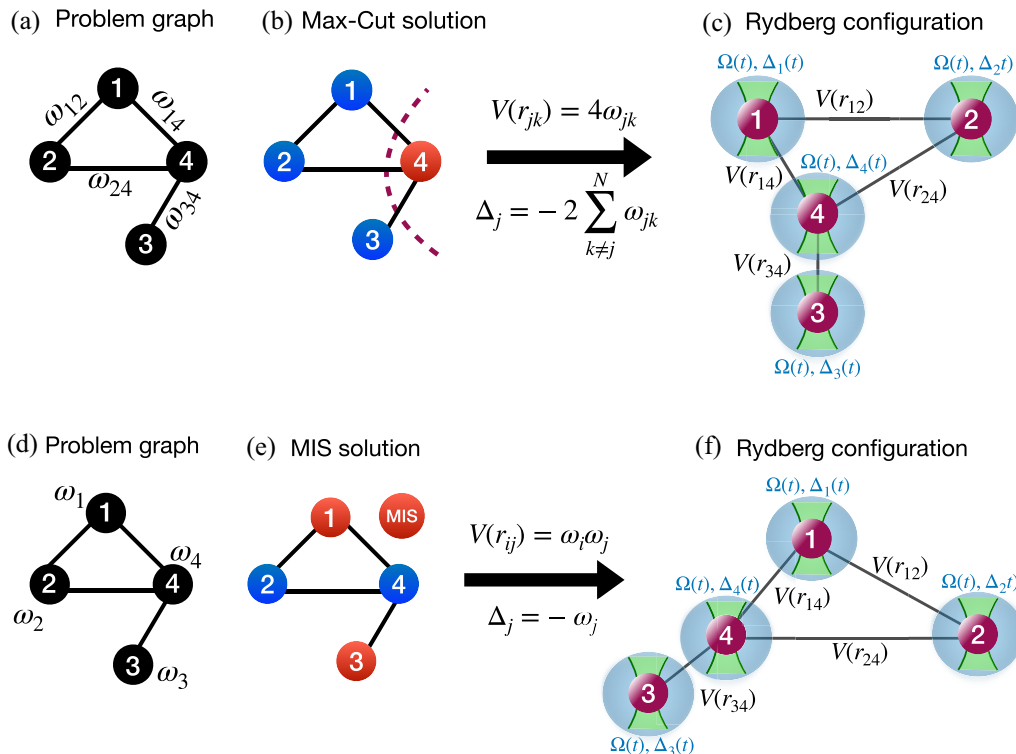
FIG. 1. Setup of weighted Max-Cut and MIS problems. (a) The problem graph for Max-Cut with weights w_{ij} on the edges (such that $\omega_{34} > \omega_{14} > \omega_{24} > \omega_{12}$) and (d) the problem graph for MIS with weights w_i on the vertices (such that $\omega_3 > \omega_1 > \omega_4 > \omega_2$). Panels (b) and (e) show the solution to the corresponding problems. The dashed curve in (b) indicates the cut, dividing the graph into two sets, red and blue vertices. In (e), red vertices constitute the MIS. Panels (c) and (f) correspond to an atomic configuration where the shaded blue region around each atom indicates the local detuning and the global Rabi frequency used in the setup. Each atom is subjected to specific values of detuning depending on the weights which are encoded in the distance-dependent interactions between the atoms as indicated with the solid black arrow. Distance between atoms in the Rydberg configuration in the case of the Max-Cut problem follows $r_{34} < r_{14} < r_{24} < r_{12}$ and in the case of the MIS problem follows $r_{34} < r_{14} < r_{12} < r_{24}$.

and 1(e) depict the MIS problem. The MIS can be found by maximizing the following cost function:

$$C_{\text{MIS}} = \sum_{j \in V} w_j X_j - \sum_{(j,k) \in E} w_j w_k X_j X_k, \qquad (4)$$

where the sum is over all the vertices V in the first term, and in the second term the sum is over all the edges E. Here, $X_j \in \{0, 1\}$ where $X_j = 1$ indicates that a vertex j is in the independent set. The first term in the cost function contributes only when $X_j = 1$, and the second term ensures that no two vertices with an edge between them belong to the independent set. It is done by suppressing the occurrence of simultaneous $X_j = 1$ and $X_k = 1$ along an edge.

III. NON-BLOCKADE-BASED PROTOCOL FOR SOLVING MAX-CUT AND MIS ON RYDBERG ANNEALER

This section discusses the protocol adopted in this work to solve the QUBO problems. The first step in solving the Max-Cut and MIS problems using Rydberg annealers is to encode the problem graph onto the physical spins (qubits) of the Rydberg setup. In principle, this encoding scheme is not unique and the most widely used approach, in the context of Rydberg simulators, is the unit disk (UD) encoding [39,47]. Although it is the natural choice of encoding owing to the

fact that the Rydberg blockade effect successfully implements unit disks in a straightforward manner [27,35,41], it has its drawbacks. In order to solve a particular QUBO problem (in this case Max-Cut or MIS) graph with N vertices, it requires solving a graph with $\sim 4N^2$ unit disks. This significant overhead along with the issue of unwanted interactions calls for alternative encoding schemes.

One of the highlights of this work is to explore an alternative, non-blockade-based encoding scheme that allows us to map QUBO to Rydberg annealers both for weighted and unweighted graphs using a single framework and it possesses a linear scaling with respect to N. In steps 1 and 2, we outline our scheme of encoding the Max-Cut and MIS problem graphs onto the Rydberg spin model, while in step 3, we discuss the implementation of optimal quantum annealing using Rydberg atoms.

Step 1: Mapping of cost functions using local detuning. The cost functions $C_{\text{Max-Cut}}$ and C_{MIS} defined in the earlier section can be directly mapped to the spin Hamiltonians using the standard definitions of the Pauli operators. The details are provided in Appendix A and the corresponding spin Hamiltonians with N spins are given in the following, where

$$\hat{H}_{\text{Max-Cut}} = \sum_{j=1}^{N-1} \sum_{k=j+1}^{N} w_{jk} \hat{\sigma}_j^z \hat{\sigma}_k^z \qquad (5)$$

is the spin Hamiltonian for the Max-Cut cost function [Eq. (3)]. Comparing Eqs. (2) and (5), we define the detunings Δ_j in a manner such that they cancel the effective longitudinal field in $\hat{H}_{\text{Ising}}$. Thus, for an atom at the jth position, the local light shift is chosen to be $\Delta_j = -1/2 \sum_{k \neq j} V_{jk}$. The edge weights w_{jk} from the graph are identified as the interaction strength V_{jk} between the atoms labeled as j and k as $V_{jk} = 4w_{jk}$. This implies that the value of the detunings is related to the weights according to

$$\Delta_j = -2 \sum_{k \neq j}^{N} w_{jk}. \quad (6)$$

Similarly for the MIS problem, we get the Hamiltonian for interacting spins to be

$$\hat{H}_{\text{MIS}} = \sum_{j=1}^{N} \left(\frac{-w_j}{2} + \frac{1}{4} \sum_{k=1, k \neq j}^{N} e_{jk} w_j w_k \right) \hat{\sigma}_j^z$$
$$+ \frac{1}{4} \sum_{j=1}^{N-1} \sum_{k=j+1}^{N} e_{jk} w_j w_k \hat{\sigma}_j^z \hat{\sigma}_k^z. \quad (7)$$

In the case of the MIS problem, $e_{jk} = 1$ if and only if there is an edge between vertices j and k, otherwise $e_{jk} = 0$. On comparing Eq. (2) with Eq. (7), the vertex weights w_j from the graph are related to the interaction strength V_{jk} between the atoms labeled as j and k as $V_{jk} = e_{j,k} w_j w_k$. The detunings Δ_j get chosen to represent the effective longitudinal field for an atom at the jth position. This gives the following relationship:

$$\Delta_j = -w_j. \quad (8)$$

Thus, the mathematical problem of maximizing Eqs. (3) and (4) is reduced to a physical problem of finding the many-body ground state (MBGS) of the spin Hamiltonians given by Eqs. (5) and (7), respectively. The scheme involves temporally varying the laser parameters till the specific choice of detuning values are obtained at the end of the protocol, all the while minimizing the energy of the full system. This will be elaborated in step 3.

The mapping of classical cost function to the Rydberg Hamiltonian described here is a *direct encoding* type which results in the linear scaling of the number of atoms needed to represent the vertices of the graph. Consider a problem graph $G(\mathbf{V}, \mathbf{E}; \mathbf{W})$ where $n(\mathbf{V})$ and $n(\mathbf{E})$ are the number of vertices and edges, respectively. Thus, for an all-to-all connected graph, $n(\mathbf{E})$ takes the maximum value, which is

$$n(\mathbf{E}_{\text{All}}) = \frac{n(\mathbf{V})(n(\mathbf{V}) - 1)}{2}.$$

However, the number of edges for a general graph can vary from 1 to $n(\mathbf{E}_{\text{All}})$ depending on the problem that needs to be solved. This implies that the dependence of $n(\mathbf{E})$ on $n(\mathbf{V})$ can be expressed as

$$n(\mathbf{E}) \propto n(\mathbf{V})^{\alpha},$$

where $\alpha \in [0, 2]$. In some sense, the degrees of freedom for a graph can be understood in terms of $n(\mathbf{E})$. Hence, the number of independent degrees of freedom in w_{jk} for generic graphs does not always scale quadratically with $n(\mathbf{V})$ but can have

$\alpha < 2$. For a direct encoding, the number of atoms required is the same as the number of vertices in the problem graph. If the direct encoding of the graph with N vertices were to be implemented using only interactions between atoms, then unwanted interactions cannot be ignored. This leads to unavoidably creating unwanted edges that did not exist in the original problem graph.

Step 2: Spatial arrangement of atoms. As seen in the previous step, the distance-dependent interactions encode the information about the weights of the graph. Therefore, before any evolution of the many-body Hamiltonian, the initial arrangement of the atoms is crucial to the encoding scheme as is represented schematically in Fig. 1. For example, the case of a weighted Max-Cut graph with weights $\omega_{34} > \omega_{14} > \omega_{24} > \omega_{12}$ will correspond to atoms (purple balls) arranged in a configuration such that $r_{34} < r_{14} < r_{24} < r_{12}$ which are trapped in tweezers (shown in green). Similarly, the weighted MIS problem with weight relationship given by $\omega_3 > \omega_1 > \omega_4 > \omega_2$ is associated with a Rydberg configuration where $r_{34} < r_{14} < r_{12} < r_{24}$. Since the weights ω_{ij} are related to the interactions $V(r_{ij})$, which in turn is reflected in the choice of the r_{ij}, the atoms are arranged in a manner that provides a true representation of the graph. However, it is possible to have interactions between atoms that do not share an edge in the original graph problem. Such *unwanted interactions* become a serious issue for graphs with higher degrees. For this purpose, in two-dimensional (2D) geometry, the maximum degree of a single node in the graph is limited to five, for which we numerically checked that the unwanted interactions do not play a significant role. Apart from the unwanted interactions, the arrangement of the atoms which do have an edge in the problem graph is also limited. This case appears when the weights in the graph are such that the resulting relative distances between the atoms cannot be realized on a 2D plane. Although the graphs that can be represented by the geometrical arrangement of atoms are limited, we emphasize that our scheme is applicable to a large class of graphs out of which only a few are demonstrated in this work. One of the ways to generate the family of accessible graphs is by considering geometrically implementable graphs as the basis and connecting them through an additional node between them. However, adding a third dimension can also increase the value of the allowed degree and the flexibility of the atom arrangement.

Step 3: Optimal quantum annealing with Rydberg atoms. The previous two steps outlined the encoding of the graph problem to the Rydberg spin Hamiltonian. The goal is to find accurate solutions to a problem graph as efficiently as possible with respect to the number of iterations and run time. This is achieved by numerically solving the spin dynamics using the following Rydberg Hamiltonian for a specific arrangement of atoms:

$$\hat{H}_{\text{Ryd}} = \frac{\Omega(t)}{2} \sum_j \hat{\sigma}_j^x - \sum_j \Delta_j(t) \hat{n}_j^e + \sum_{k,j} V_{kj} \hat{n}_k^e \hat{n}_j^e, \quad (9)$$

where $\hat{\sigma}_j^x = |e\rangle_j \langle g| + |g\rangle_j \langle e|$ and excitation number operator $\hat{n}_j^e = \frac{1}{2}(\hat{\sigma}_j^z + \mathbb{1})$ as $|e\rangle_j \langle e|$. The objective is to reach the target Hamiltonian $\hat{H}_{\text{Max-Cut,MIS}}$ [Eqs. (5) and (7)] while minimizing the energy of the system to obtain the instantaneous

ground state whose configuration provides the solution to the optimization problem. Initially, all atoms are in the ground state $|gg \cdots g\rangle$ which corresponds to a large nonzero detuning value $\Delta_j(t=0) \neq 0$. During the protocol, the detuning value on each atom, $\Delta_j(t)$, varies but attains a specific value at the end of the protocol as defined by Eqs. (6) and (8). The initial and final values of the Rabi frequency are set to zero. At intermediate times, nonzero values of the Rabi frequency $\Omega(t)$ provide a transverse field for the above Ising Hamiltonian. This generates quantum fluctuations and causes different many-body states to couple with each other, thereby accessing a larger part of the Hilbert space. This is key to the quantum annealing process as it explores different configurations through the application of the nonzero transverse field while traversing the energy landscape to get to the desired many-body ground-state configuration (i.e., the solution to the problem graph). The idea is to use well-established techniques of optimal control theory applied to quantum systems [48–54] in order to reach this state in an efficient manner well within the system lifetime. In particular, a combination of gradient- and non-gradient-based methods (see Appendix C for details) is used to shape the pulses in time which allows us to optimally steer the many-body state towards the true solution. The initial guesses for detuning and Rabi frequency are inspired by the adiabatic evolution and are then optimized in time. The objective function that needs to be minimized during the optimal control is the expectation value of the target Hamiltonian $\hat{H}_{\text{Max-Cut,MIS}}$ with respect to the instantaneous many-body state $|\psi_{\text{inst}}(t)\rangle$ which is given as

$$E = \langle \psi_{\text{inst}}(t)| \hat{H}_{\text{Max-Cut,MIS}} |\psi_{\text{inst}}(t)\rangle. \tag{10}$$

E is used as the cost function for the optimization but is not the typical energy of the instantaneous many-body state. Apart from E, we also define the overlap between the instantaneous many-body state with the many-body ground state $|\psi_g\rangle$ of the problem Hamiltonian $\hat{H}_{\text{Max-Cut,MIS}}$ during the protocol given as

$$F(t) = \sum_g |\langle \psi_{\text{inst}}(t)|\psi_g\rangle|^2. \tag{11}$$

The fidelity F is calculated over all the degenerate ground states for an *a posteriori* analysis of the quality of the solution obtained. More details about the control methods are provided in Appendix C.

IV. CHARACTERIZATION OF MAX-CUT AND MIS PROBLEMS

In order to benchmark our protocol (algorithm), we need to ascertain how close is the solution provided by our algorithm to the true solution for problem graphs and compare it with other methods. This is encapsulated in a quantity referred to as an *approximation ratio* which is discussed in this section. With regards to the efficiency of obtaining the solution for a given algorithm, it is possible to have a scenario where finding a solution for a particular graph is faster in run time than compared to another graph of similar size. This indicates that the latter graph is more *complex* or *hard* to solve which is an inherent property of the problem graph. However, it is nontrivial to characterize the complexity of arbitrary graphs.

TABLE I. The table shows a one-to-one correspondence between the degenerate solutions (DSs) of the original Max-Cut and MIS problems to the MBGS of the corresponding Rydberg Hamiltonians given by Eq. (5) and (7), respectively.

Max-Cut[a]		MIS[b]	
DS	MBGS	DS	MBGS
10100	$|egegg\rangle$	10001	$|eggge\rangle$
01011	$|gegee\rangle$	10100	$|egegg\rangle$
01010	$|gegeg\rangle$	01010	$|gegeg\rangle$
10101	$|egege\rangle$	01001	$|gegge\rangle$

[a]For the unweighted case of the graph in Fig. 2(a).
[a]For the unweighted case of the graph in Fig. 2(f).

Motivated by Refs. [27,41], we generalize their hardness parameter to include a broader class of problems using the notion of degenerate subspaces.

A. Approximation ratio

In general, the approximation ratio quantifies the worst-case performance of an algorithm for solving a particular problem. In certain cases, this ratio can be evaluated analytically. This is the case for the Goemans-Williamson algorithm [55] solving the Max-Cut problem which is considered to be the best classical approximate algorithm with an approximation ratio of 0.878. Any quantum algorithm outperforming this ratio is suggestive of having an advantage over the classical case. In order to benchmark quantum algorithms, it is more convenient to evaluate this ratio numerically as defined in Ref. [56] and given as

$$R = \frac{C_{\text{obt}}}{C_{\text{opt}}}, \tag{12}$$

where C_{opt} is the optimal value of the cost function and C_{obt} is the obtained value of the cost function evaluated through our method. In order to evaluate C_{opt}, the exact solution to the given problem should be available *a priori*. In the Results section, we evaluate R for different graphs.

B. Hardness parameter

For Max-Cut and MIS problems, there are certain graphs such as weighted ones that are generally more challenging to solve computationally [57]. One way to identify the complexity in a graph is by looking at the symmetries in the adjacency matrix [58]. However, a physically more intuitive way is to study the degeneracy in the many-body ground state of the interacting spin systems. As a result of mapping the cost functions using local detuning described earlier, there is a one-to-one correspondence between the degenerate solutions of the original Max-Cut and MIS problem to the MBGS of the corresponding Rydberg Hamiltonians as shown in Table I. The larger the degeneracy for the many-body ground state (solution space), the higher the probability to get to the optimal solution during the dynamics. But if the orthogonal subspaces (not belonging to the solution space) also have large degeneracy, then there is a possibility for the solution to get stuck in one of these unwanted subspaces.

Using this notion of comparing degeneracy between relevant subspaces, a hardness parameter was defined only in the context of the MIS problem [27], $HP = D_{|MIS|-1}/(|MIS|D_{|MIS|})$. Here $|MIS|$ and $D_{|MIS|}$ correspond to the size of the MIS and the degeneracy of the solution space. $D_{|MIS|-1}$ defines the degeneracy of the subspace of suboptimal independent sets which has one less arbitrary element (vertex) from the MIS solution set. This hardness parameter is limited as it does not take into account all the other subspaces (with sizes $< |MIS| - 1$) with relatively higher degeneracy. Furthermore, this hardness parameter is not well defined for the solution to the Max-Cut problem.

Thus, in this work, a more general hardness parameter is introduced for both Max-Cut and MIS. This parameter includes information about the degenerate subspaces and is defined as

$$HP = \frac{\sum_{D > D_{\text{cutoff}}} D}{C_{\text{opt}} D_{\text{opt}}}, \qquad (13)$$

where D_{opt} represents the degeneracy of the solution space, C_{opt} is the optimal value of the cost function, and D represents the degeneracy of a subspace. The sum is taken over all orthogonal subspaces whose degeneracy is greater than a specific cutoff value. The cutoff value is arbitrarily chosen until the value of the hardness parameter HP shows convergence. It should be noted that if $D_{\text{opt}} > D_{\text{cutoff}}$ then this degenerate configuration is excluded from the numerator sum. The hardness defined for a problem instance is an inherent property of the problem, independent of the algorithm. This is because the hardness here characterizes if one instance of the problem has a more nonconvex optimization landscape as compared to the other instance. If a problem has more degenerate subspaces, the optimization landscape will be more nonconvex and hence harder to navigate for any classical algorithm. We numerically find that our hardness parameter faithfully captures the complexity of the problem graphs for various cases as shown in the Results section.

V. RESULTS

As mentioned in step 2 of Sec. III, although we consider prototypical graphs with a degree not more than five and with restricted connectivity that is realizable in a two-dimensional array of atoms in tweezers, we have solved for a large class of graphs out of which only a few are shown here. These other graphs can, e.g., be generated by combining smaller graphs that are known to be solvable through an additional node between them. Refer to Appendix B for numerical simulation details.

To understand the working principle of our optimal quantum annealing protocol, we consider a relatively simple case of the weighted graph of size 5 and degree 3 for which the Max-Cut and MIS solutions are shown in Figs. 2(a)–2(e) and Figs. 2(f)–2(j), respectively. The optimized shapes of the laser parameters (dotted black symbols for detunings, solid dark red for Rabi frequency) are shown in Figs. 2(b) and 2(g), plots of fidelity (dashed green) and energy (solid blue) in Figs. 2(c) and 2(h), plots of ordered energies $E_i(t)$ along with basis state contribution in the instantaneous eigenstates (indicated by green color bar) in Figs. 2(d) and 2(i), and the population of basis states at three different time steps in

Figs. 2(e) and 2(j). The solutions for both problem graphs are obtained efficiently ($\sim$3.5 μs) and with high fidelity ($\sim$0.99). By turning on the Rabi pulse, one obtains multiple avoided crossings in the many-body energy spectrum. The rate at which the individual detunings are swept controls the Landau-Zener transitions across multi-level crossings. This results in a nontrivial path of dynamics for the many-body system to its target state. Let $C = [\Delta_1(T), \Delta_2(T), \ldots, \Delta_N(T)]$ be the set of final local detunings for N atoms determined *a priori* by the weights on the graph given by Eqs. (6) and (8). The ratio between the set of detunings $\Delta_1(T) : \Delta_2(T) : \cdots : \Delta_N(T)$ is kept fixed during the entire protocol. Thus, a single time-dependent parameter $\Delta_G(t)$ is defined such that the temporal variation of each individual detuning at each atom j is given as $\Delta_j(t) = \Delta_G(t)\Delta_j(T)$ as shown in Figs. 2(b) and 2(g). This factor $\Delta_G(t)$ is initialized such that it starts at a negative value at $t = 0$ and increases to 1 at $t = T$, thereby providing the desired set of detuning values at the end of the protocol. The physical implementation of the local-detuning protocol is elaborated in Sec. VI. The quantity E [defined by Eq. (10)] is minimized and this is reflected in F [defined by Eq. (11)], shown in Figs. 2(c) and 2(h). The approximation ratios R [defined by Eq. (12)] of the final state are also indicated. For a general Max-Cut problem graph, every cut has at least twofold degeneracy because of the symmetry between up and down spins. This is also the case for the chosen graph and thus the Max-Cut problem has more degenerate states compared to the MIS problem. This key feature is reflected in the initial part of the dynamics. The change in E during the initial part of the dynamics is synonymous with the change in the instantaneous eigenstates [Figs. 2(h) and 2(i)] for the MIS problem. However, this is not the case for the Max-Cut problem. For Max-Cut protocols, we find that E stays constant for a more extended initial period, signifying that the energy of the instantaneous state is close to the high energy manifold of the $\hat{H}_{\text{Max-Cut}}$ and remains there longer as a result of the degeneracy. Both for Max-Cut and MIS, the initial state $|ggggg\rangle$ is populated as shown in the leftmost panels of Figs. 2(e) and 2(j). At $t = T$ for Max-Cut [rightmost panel of Fig. 2(e)], the degenerate ground states ($|egegg\rangle$ and $|gegee\rangle$) are populated while for the MIS case, a single state $|eggge\rangle$ is populated [rightmost panel of Fig. 2(j)]. At intermediate times [middle panels of Figs. 2(e) and 2(j)] due to the nonzero values of the transverse field ($\Omega \neq 0$), multiple states get populated. Specifically, 12 out of 32 in Fig. 2(e) and 9 out of 32 in Fig. 2(j), highlighting the fact that the optimal protocol dynamics across the energy landscape is nonintuitive. This aspect is reinforced in Fig. 3 where we have 15 spins.

Figure 3 demonstrates the flexibility and scaling of our encoding for a more complex case involving a weighted graph of size 15 and degree 5. Figures 3(a) and 3(e) show the graph topology and the exact solution for Max-Cut and MIS. For this graph, there will be in general 15 final detuning values $\Delta_j(T)$ at time T. The set of local detunings will all vary with the same fixed ratios between them. For illustration purposes, we show the variation of the factor $\Delta_G(t)$ in Figs. 3(b) and 3(f). Since the complexity of the problem is significantly increased, we have a different choice for the initial parameters. In particular, a linear guess is used for $\Delta_G(t)$ in Fig. 3(f) while a linear ramp with a flat top is used in Fig. 3(b). Figures 3(c) and 3(g)
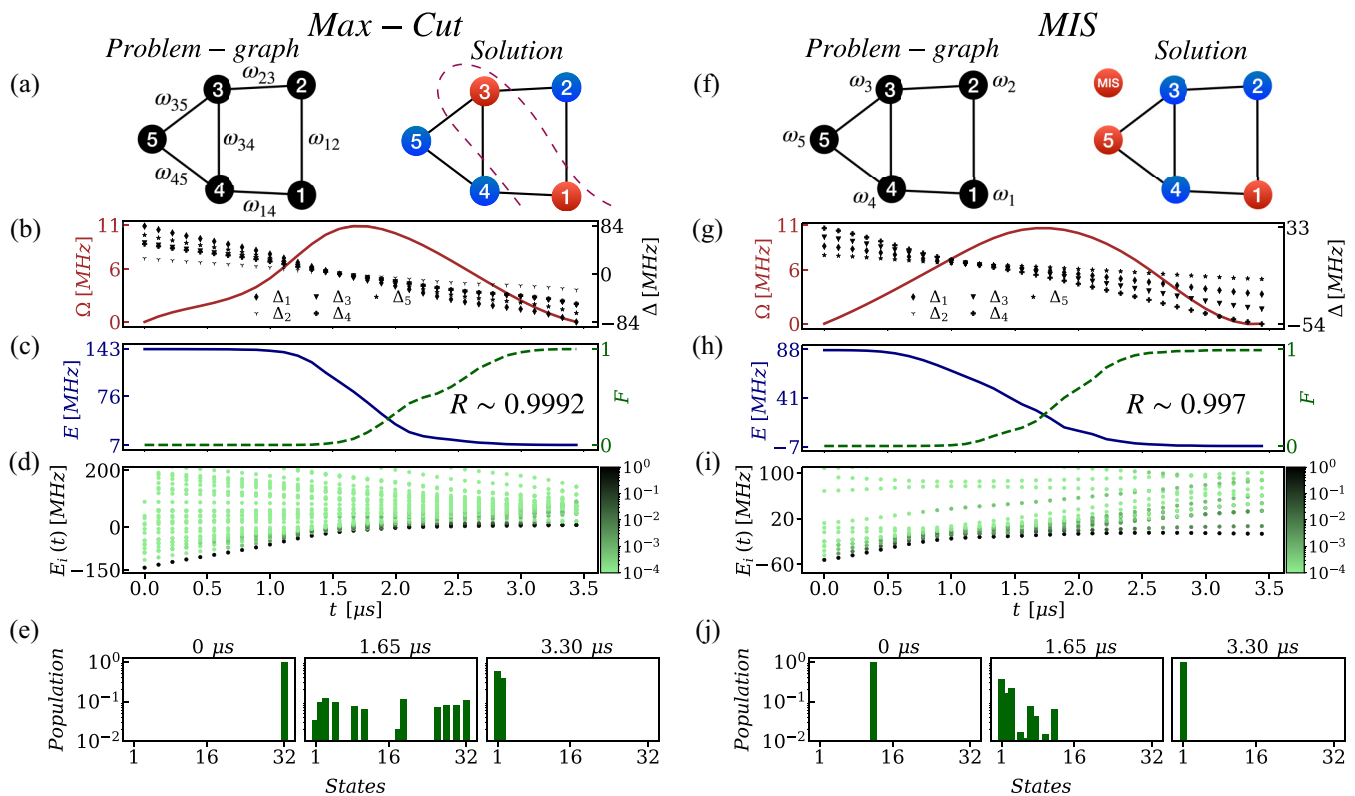
FIG. 2. [(a)–(e)] The Max-Cut problem and [(f)–(j)] the MIS problem. [(a) and (f)] Weighted prototype graphs of size 5 with corresponding solution graphs. [(b) and (g)] Optimal protocols for the Rabi frequency (solid dark red) and the local detunings (dotted black symbols) with time. The local detuning of each atom is controlled by varying a single time-dependent parameter $\Delta_G(t)$ which is explained in the main text. Maximum and minimum speeds in detuning change for the Max-Cut problem are 47.4 and 15.4 MHz/µs, respectively, and for the MIS problem, they are 24.5 and 4.7 MHz/µs, respectively. The expectation value E of the problem Hamiltonian with respect to the instantaneous state (solid blue) and the fidelity F (dashed green) of the instantaneous state with respect to the ground state are shown in (c) and (h), where R indicates the approximation ratio. [(d) and (i)] The ordered energies of the instantaneous eigenstates, with a color bar indicating the population of the basis states during the protocol. The population of the basis states of the Hamiltonian at final time $t = T$ is shown at three different times during the protocol in (e) and (j). As shown in the middle panels of (e) and (j), 12 states and 9 states out of 32 are populated in the middle of the protocol. The output at the end of the protocol captures all the degenerate states corresponding to all the degenerate Max-Cut and MIS solutions.

show the corresponding variation of the expectation value E and fidelity F during the protocol. Another signature of the complexity of this problem graph can be seen in Figs. 3(d) and 3(h), where we find the significant population of a large number of basis states out of the 2^{15} states, especially at intermediate times. In Fig. 3(d), there are 1380 populated states at $t = 3.50$ µs, and 310 populated states at $t = 7.10$ µs. Similarly, in Fig. 3(h), there are 678 populated states at $t = 3.50$ µs, and 1029 populated states at $t = 7.10$ µs. It is remarkable that, at the final time, we end up with a small set of degenerate MBGSs. This is a direct consequence of the optimizer minimizing E which forces the system to transfer the population to the lower energy levels of the target Hamiltonian. Despite the population of so many basis states, the optimal protocol ensures efficient transfer to the optimal solution ($R \sim 0.99$), making the relevance of energy gaps redundant in the context of multilevel dynamics. One may speculate that this intuition will hold for larger systems.

Figure 4 compares the performance of our protocol (LO-QAL) with fast simulated annealing (SA) for different graphs. More details about SA are provided in Appendix D. In

particular, we illustrate the variation of the approximation ratio error $1 - R$ [R is defined by Eq. (12)] with varying system size $N = 2$–15 and hardness parameter HP [defined by Eq. (13)]. In the case of Max-Cut, the approximation ratio error ($1 - R$) shows similar trends with respect to the system size N and the hardness parameter HP as shown in Figs. 4(a) and 4(b). As system size increase, so does the number of degenerate orthogonal subspaces (which affects the numerator of HP). But since the degeneracy of the solution space for the Max-Cut is always twofold (this affects the denominator of HP), the overall hardness parameter increases with system size. For the MIS problem, the SA gives a general upward trend with oscillations in the approximation ratio error with increasing system size N as seen in Fig. 4(c). These oscillations are attributed to the degeneracy of the solution space being linked to the system size which by construction is twofold for even N and onefold for odd N for the graphs considered in this study. The larger the degeneracy of the solution space, the easier it is for the system to reach one of the solutions. This oscillatory behavior is not reflected in the hardness parameter because our definition of HP takes into account the
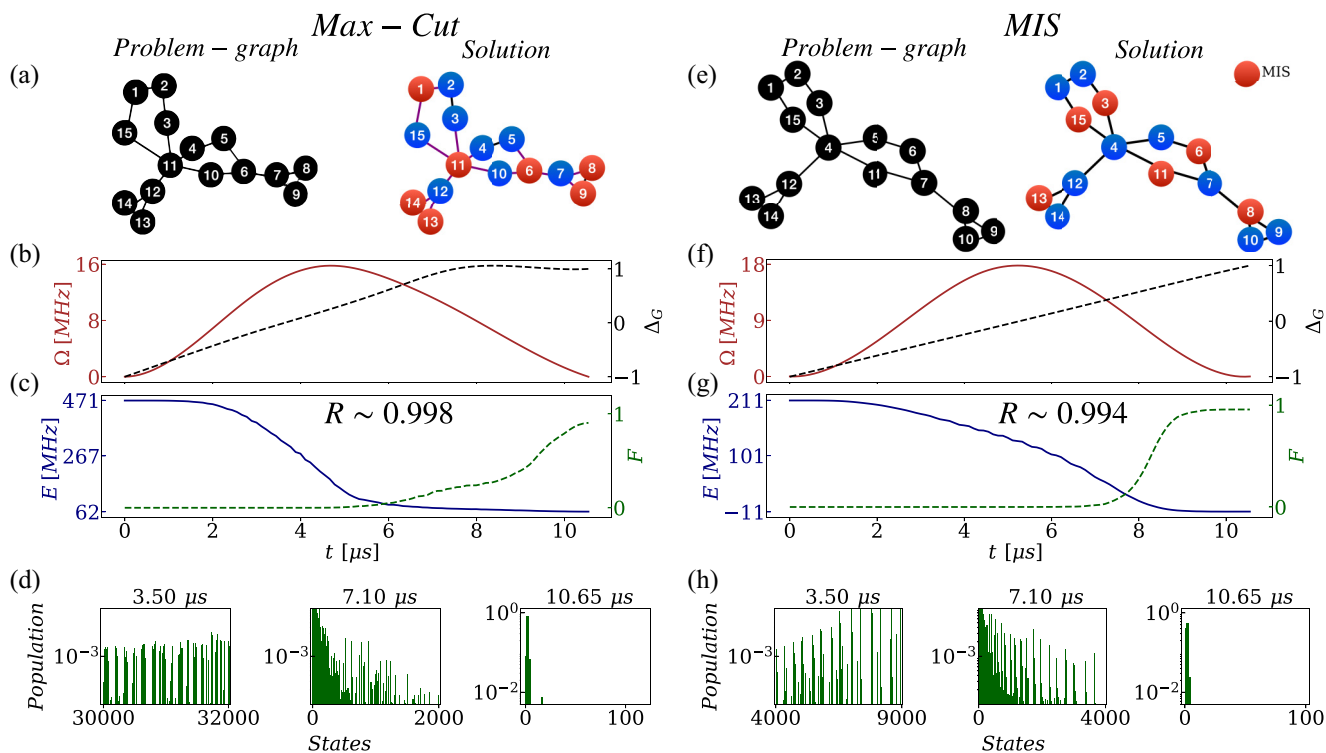
FIG. 3. Solutions to the [(a)–(d)] Max-Cut and [(e)–(h)] MIS problems for a graph of size 15 and degree 5 using the optimal quantum annealing. Weighted prototype problem graphs with solution graphs for the Max-Cut problem are shown in (a) and are shown for the MIS problem in (e). (a) Vertex 11 has a degree of 5; (e) vertex 4 has a degree of 5. [(b) and (f)] The optimal protocol for the Rabi frequency depicted by the solid dark red curve and $\Delta_G(t)$ (defined similar to Fig. 2) depicted by the dashed black curve, with time. The maximum and minimum speeds for detuning change, for Max-Cut protocols, are 28.8 and 5.1 MHz/μs, respectively, and, for MIS protocols are 1.8 and 0.8 MHz/μs, respectively. The expectation value E (solid blue) of the problem Hamiltonian with respect to the instantaneous state and fidelity F (dashed green) of the instantaneous state with respect to the ground state are shown in (c) and (g), where R indicates the approximation ratio. The population of states at three different times ($t = 3.50$ μs, $t = 7.10$ μs, and $t = 10.65$ μs) are shown in (d) and (h).
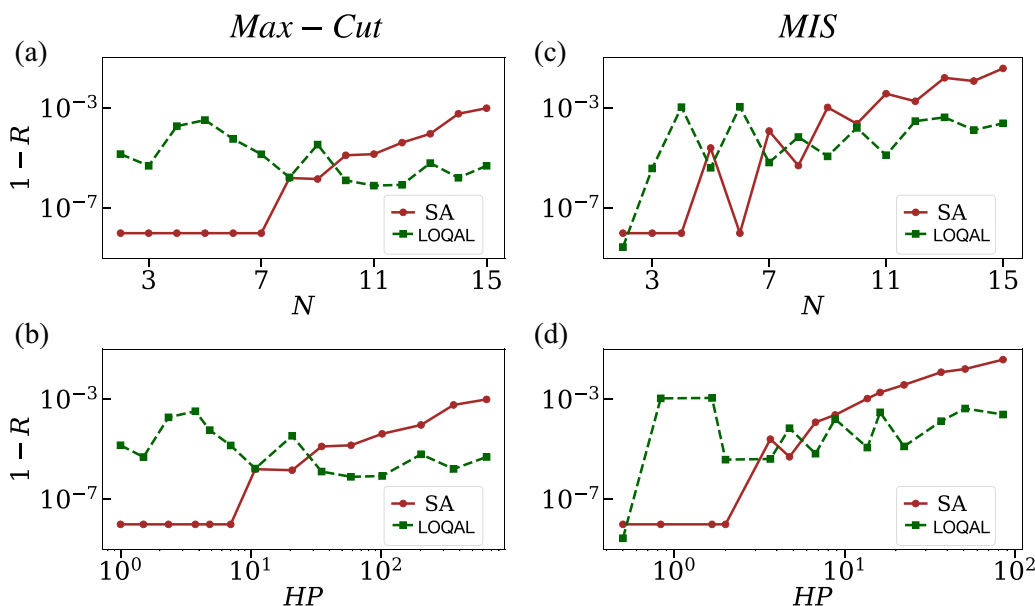

FIG. 4. Comparison of optimized simulated annealing (SA) and localized optimal control for quantum annealing in a loop (LOQAL) for [(a) and (b)] Max-Cut and [(c) and (d)] MIS problems. Approximation ratio error $1 - R$ with respect to system size N is shown in (a) and (c), and for the hardness parameter HP in (b) and (d).
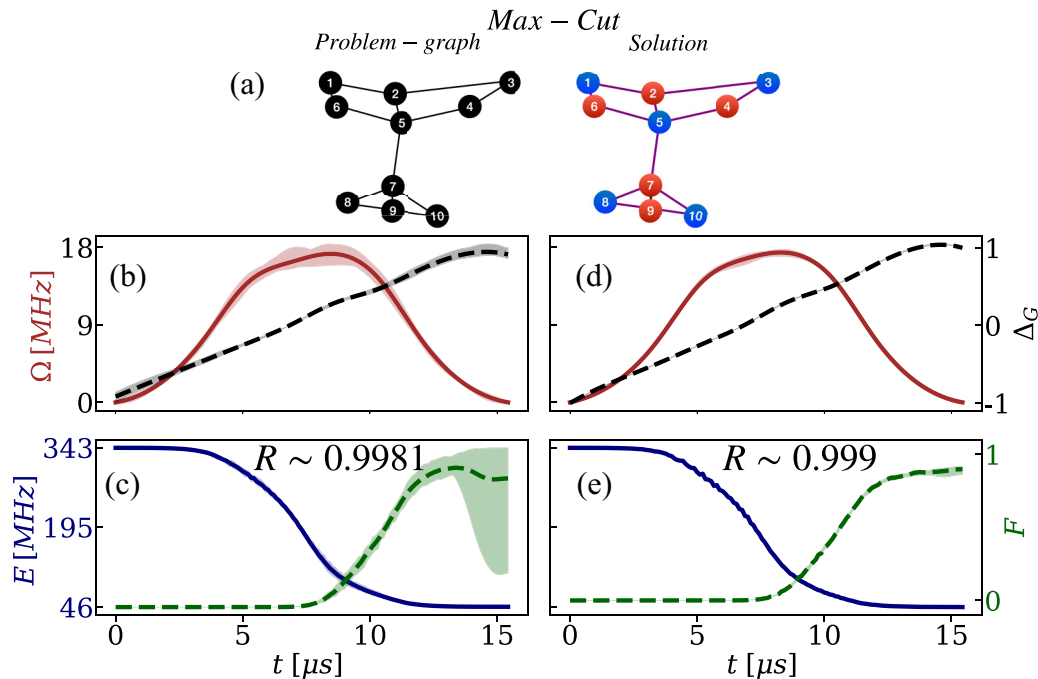
FIG. 5. Results using the optimal quantum annealing with noisy protocols to solve the Max-Cut problem for a graph of size 10. The weighted prototype graph for the Max-Cut problem is shown in (a) along with the solution graph. The optimal protocol for the Rabi frequency (solid dark red) and the factor controlling the detunings Δ_G (dashed black) are shown in (b), where noise is added to the laser parameters of the preoptimized protocol and in (d) where noise is added during the optimization of the parameters. The shaded regions represent the fluctuations in the laser parameters for each run that were chosen from a random distribution. [(c) and (e)] Corresponding expectation value (solid blue) of the problem Hamiltonian with respect to the instantaneous state (E) and the fidelity F (dashed green) of the instantaneous state with respect to the ground state for both cases, where R indicates the approximation ratio.

degeneracy of all the relevant subspaces and thereby successfully characterizes the problem graphs. For the graphs considered in Fig. 4, it clearly shows the quantum algorithm is more robust and performs better than SA.

All results so far did not include any noise in the dynamics. Using Max-Cut as an example, we apply the optimal quantum annealing protocol to a noisy system as is expected in real experiments. Figure 5 displays the Max-Cut problem of size 10 with noise up to 8% in the laser parameters chosen randomly which is shown as shaded regions across the bold lines (mean value) in Figs. 5(b) and 5(d). Two different approaches are adopted to simulate the noisy model. In Fig. 5(b) the noise is added after the optimization of the pulses while in Fig. 5(d) it is added during the optimization procedure and the optimizer adapts to the noisy protocol. The shaded region of the fidelity F is broader than that of E (which is barely visible) and is shown in Fig. 5(c), indicating that slight variations in the parameters bring a small change in E, but drastic changes in F. This is due to the fact that the gap in the lower energy levels of the target Hamiltonian is vanishing. That results in the system occupying low-energy excited states close to the MBGS, at the same time dropping the fidelity as it is a quantity highly sensitive to the occupied states. In order to better simulate real experiments, a random error of up to 8% is added to the laser parameters during each run which is shown in Fig. 5(d). This causes the optimization landscape to adapt to the random variations in the parameters. Thus, in Fig. 5(d), the shaded region of the pulses is not as broad as in Fg. 5(b). In

Fig. 5(e) both the expectation value E and the fidelity F show no significant effect of the noise, thereby being more resilient to the laser noise.

VI. PHYSICAL IMPLEMENTATION OF THE PROTOCOL

For the numerical simulations shown in this work, we considered Rydberg states $60S_{1/2}$ of the Cs atoms which have a van der Waals coefficient $C_6 \sim 139$ GHz μm^6 and a radiative lifetime of ~ 234 μs [25,59,60]. An essential aspect of the protocol is to realize an array of trapped atoms with adjustable interatomic distances ranging from 1 to 7 μm, which should be achievable with the current state-of-the-art optical tweezer technology [43]. An ingredient of our protocol is the optimal control of the laser parameters. Apart from one global near-resonant laser that couples the ground-state atom to its Rydberg state with Rabi frequency Ω, we have another laser whose intensity will be distributed over the atoms selectively using a spatial light modulator [61]. This will provide individual local light shifts for the ground-Rydberg transition, thereby implementing specific local detunings. At the end of the experimental sequence, measurements of the distribution of the Rydberg excitations can be performed by fluorescence imaging. The whole process will be repeated multiple times to efficiently calculate and classically minimize the expectation value of the problem Hamiltonian [Eq. (10)].

We considered the case of noise in laser parameters in our simulations but the experiments can suffer from a variety of

errors, introducing different types of noise to the system which includes motional dynamics, interaction induced dephasing [62], blackbody radiation [63], and state preparation and measurement (SPAM) errors [64], all of which could lead to a significant drop in readout fidelity. There are ways to address and control errors on the Rydberg platform, including converting them into erasures [63] and transforming complicated error models into Pauli-Z errors through the use of ancillary atoms [65]. Future works will involve optimizing for noisy models using our protocol.

VII. CONCLUSIONS AND OUTLOOK

Rydberg atoms can achieve the required scalability with arbitrary connectivity between qubits, making them highly desirable platforms to solve optimization problems. The protocol presented in this work is fairly universal in the sense that it can handle both weighted and unweighted graphs within the same framework, which is not always obvious for other schemes and possibly could be generalized to other QUBO problems beyond Max-Cut and MIS. Although the focus of this work is on a Rydberg annealer, any quantum device that permits local qubit control along with the global driving of the many-body system can implement our protocol. A promising aspect of the optimized dynamics is that it goes beyond the energy gap dependence that would usually limit the adiabatic protocols for large system sizes, which need to be verified for larger systems.

The motivation for presenting an alternative encoding scheme that is implementable on an optimized Rydberg annealer is to have a more favorable scaling of the required number of qubits for problem graphs with N vertices. For the sizes and complexity of graphs considered here, we do find accurate solutions for both the Max-Cut and MIS problem graphs with $N = 5$ and 15 vertices that do have an $O(N)$ scaling. However, the topology of the graphs is limited due to the issue of unwanted interactions. Such effects can be mitigated using a three-dimensional arrangement of atoms but further investigations are needed to generalize our encoding scheme for graphs with arbitrary connectivity and degree. Despite having shown the advantage of our protocol with respect to fast simulated annealing, benchmarking our method with other classical algorithms can shed more insights into the performance of our method. A more thorough analysis of optimal dynamics that can adapt to different types of noise and in particular where Bayesian methods [53,54,66] can prove to be useful will be the scope of future work.

ACKNOWLEDGMENTS

APPENDIX A: MAX-CUT AND MAXIMUM INDEPENDENT SET ENCODING

The mathematical mappings between the classical cost functions [67] of Max-Cut and maximum independent set to the target Hamiltonian as given in the main article are provided in this section.

1. Max-Cut

The classical Max-Cut cost function which is to be maximized is given by

$$C_{\text{Max-Cut}} = \sum_{(j,k) \in E} w_{jk}(X_j(1 - X_k) + X_k(1 - X_j)), \quad \text{(A1)}$$

where $X_j \in \{0, 1\}$ and the sum is over all the edges E with weights W in the graph G. By replacing X_j to $Z_j = 2X_j - 1$ such that $Z_j \in \{-1, 1\}$, the Max-Cut cost function $C_{\text{Max-Cut}}$ becomes

$$C_{\text{Max-Cut}} = \sum_{(j,k) \in E} w_{jk}\left(\frac{1 + Z_j}{2}\left(1 - \frac{1 + Z_k}{2}\right)\right.$$
$$\left. + \frac{1 + Z_k}{2}\left(1 - \frac{1 + Z_j}{2}\right)\right) \quad \text{(A2)}$$

$$= \frac{1}{4} \sum_{(j,k) \in E} w_{jk}((1 + Z_j)(1 - Z_k)$$
$$+ (1 + Z_k)(1 - Z_j)) \quad \text{(A3)}$$

$$= \frac{1}{2} \sum_{(j,k) \in E} w_{jk}(1 - Z_j Z_k). \quad \text{(A4)}$$

The $1/2$ in w_{jk} is absorbed and the sum $\sum_{(j,k) \in E}$ over edges is changed to a double sum $\sum_{j=1}^{N-1} \sum_{k=j+1}^{N}$, where N is the total number of vertices. This gives

$$C_{\text{Max-Cut}} = \sum_{j=1}^{N-1} \sum_{k=j+1}^{N} w_{jk}(1 - Z_j Z_k), \quad \text{(A5)}$$

where $w_{jk} = 0$, if there is no edge between the vertices j and k. The above cost function can be represented as an Ising-type Hamiltonian with the help of Pauli matrices,

$$\hat{H}_I = \sum_{j=1}^{N-1} \sum_{k=j+1}^{N} w_{jk}\left(1 - \hat{\sigma}_j^z \hat{\sigma}_k^z\right), \quad \text{(A6)}$$

such that the expectation value of $\hat{H}_I$ is the same as $C_{\text{Max-Cut}}$. The sum over the first term in $\hat{H}_I$ [Eq. (A6)] is just a constant, so finding the maximum $\langle \hat{H}_I \rangle$ can be formulated in terms of the Hamiltonian,

$$\hat{H}_{\text{Max-Cut}} = \sum_{j=1}^{N-1} \sum_{k=j+1}^{N} w_{jk} \hat{\sigma}_j^z \hat{\sigma}_k^z, \quad \text{(A7)}$$

where minimization of $\langle \hat{H}_{\text{Max-Cut}} \rangle$ leads to the maximization of $\langle \hat{H}_I \rangle$. In this way, the mathematical problem of maximizing $C_{\text{Max-Cut}}$ in Eq. (A1) is reduced to a physical problem of finding the many-body ground state of the Hamiltonian given by Eq. (A7).

2. Maximum independent set

The MIS can be found by maximizing the following cost function:

$$C_{\mathrm{MIS}} = \sum_{j \in V} w_j X_j - \sum_{(j,k) \in E} w_j w_k X_j X_k, \qquad (A8)$$

where $X_j \in \{0, 1\}$. The sum is over all the vertices V with weights W in the first term, and in the second term the sum is over all the edges E. Similar to the case of Max-Cut, X_j is replaced by Z_j and the cost function C_{MIS} becomes

$$C_{\mathrm{MIS}} = \sum_{j \in V} w_j \frac{(1 + Z_j)}{2} - \sum_{(j,k) \in E} w_j w_k \frac{(1 + Z_j)}{2} \frac{(1 + Z_k)}{2} \qquad (A9)$$

$$= \sum_{j \in V} \frac{w_j}{2} + \sum_{j \in V} \frac{w_j Z_j}{2} - \frac{1}{4} \sum_{(j,k) \in E} w_j w_k - \frac{1}{4} \sum_{(j,k) \in E} w_j w_k (Z_j + Z_k) - \frac{1}{4} \sum_{(j,k) \in E} w_j w_k (Z_j Z_k) \qquad (A10)$$

$$= \sum_{j \in V} \frac{w_j}{2} - \frac{1}{4} \sum_{(j,k) \in E} w_j w_k + \sum_{j \in V} \frac{w_j Z_j}{2} - \frac{1}{4} \sum_{(j,k) \in E} w_j w_k (Z_j + Z_k) - \frac{1}{4} \sum_{(j,k) \in E} w_j w_k (Z_j Z_k). \qquad (A11)$$

Now $\sum_{(j,k) \in E} w_j w_k (Z_j + Z_k)$ contains the sum over all the edges connected to the jth vertex. So for each vertex, the contribution is coming from the neighbors alone,

$$\sum_{(j,k) \in E} w_j w_k (Z_j + Z_k) = \sum_{j \in V} w_j Z_j \left(\sum_{k \in S_j} w_k \right), \qquad (A12)$$

where S_j is the set consisting of the neighbors of the jth vertex. The cost function then becomes

$$C_{\mathrm{MIS}} = \sum_{j \in V} \frac{w_j}{2} - \frac{1}{4} \sum_{(j,k) \in E} w_j w_k + \sum_{j \in V} \frac{w_j Z_j}{2} - \frac{1}{4} \sum_{j \in V} w_j Z_j \left(\sum_{k \in S_j} w_k \right) - \frac{1}{4} \sum_{(j,k) \in E} w_j w_k (Z_j Z_k) \qquad (A13)$$

$$= \sum_{j \in V} \frac{w_j}{2} - \frac{1}{4} \sum_{(j,k) \in E} w_j w_k + \sum_{j \in V} \left(\frac{1}{2} - \frac{\sum_{j \in S_j} w_k}{4} \right) w_j Z_j - \frac{1}{4} \sum_{(j,k) \in E} w_j w_k (Z_j Z_k). \qquad (A14)$$

The first two terms are just constants. If Z_i is replaced by $\hat{\sigma}_i^z$, the problem of finding the maximum of C (or minimizing $-C$) is equivalent to finding the many-body ground state of $\hat{H}_{\mathrm{WMIS}}$,

$$\hat{H}_{\mathrm{MIS}} = \sum_{j \in V} \left(\frac{\sum_{j \in S_j} w_k}{4} - \frac{1}{2} \right) w_j \hat{\sigma}_j^z + \frac{1}{4} \sum_{(j,k) \in E} w_j w_k \hat{\sigma}_j^z \hat{\sigma}_k^z. \qquad (A15)$$

The sum $\sum_{(j,k) \in E}$ in the second term in Eq. (A15) can be converted to a double sum as follows:

$$\hat{H}_{\mathrm{MIS}} = \sum_{j=1}^{N} \left(\frac{-w_j}{2} + \frac{1}{4} \sum_{k=1, k \neq j}^{N} e_{jk} w_j w_k \right) \hat{\sigma}_j^z + \frac{1}{4} \sum_{j=1}^{N-1} \sum_{k=j+1}^{N} e_{jk} w_j w_k \hat{\sigma}_j^z \hat{\sigma}_k^z, \qquad (A16)$$

where $e_{ij} = 1$, if and only if there is an edge between vertices i and j; otherwise, $e_{ij} = 0$. The parameter e_{ij} captures the information about the neighbors of the vertex i. The sum $\sum_{j \in S_i}$ is over all the neighbors of spin i and is replaced by $\sum_{j=1, j \neq i}^{N} e_{ij}$.

APPENDIX B: NUMERICAL DETAILS OF RYDBERG ANNEALER

The time-dependent Rydberg Hamiltonian provided by Eq. (9) is implemented using the *sesolve* function from the QuTip library [68]. The objective is to find the many-body ground state of the target Hamiltonians given by Eq. (5) for Max-Cut and Eq. (7) for MIS (see main text). The expectation value E of the target Hamiltonian with respect to the

instantaneous many-body state is minimized during the optimal quantum annealing protocol. The profile of a global Rabi frequency $\Omega(t)$ is optimized in time while the initial and final values remain zero. An initial guess for $\Omega(t)$ is provided such that $\Omega(t) = A(1 - cos(\pi t/T))^2$, where T is the total time of the protocol and A is of the order of a few megahertz. A localized detuning $\Delta_j(t)$ on each atom is also optimized in time where each $\Delta_j(t)$ at time T is given by Eqs. (6) and (8). For a particular graph, the ratios between individual $\Delta_j(T)$ are kept fixed during the entire protocol; hence, optimizing a single prefactor $\Delta_G(t)$ will result in the optimization of individual detunings given by $\Delta_j(t) = \Delta_G(t)\Delta_j(T)$. Δ_G takes a negative value at $t = 0$ such that the many-body ground state of the Hamiltonian at $t = 0$ is $|gg \cdots g\rangle$, which means all atoms are in the ground state. Δ_G is then varied in time

and reaches one at $t = T$; the many-body ground state of the Hamiltonian at $t = T$ corresponds to the solution of Max-Cut and MIS problems. The total time T is divided into subparts by selecting eight points; $\Omega(t)$ and Δ_G at these eight time points are optimized during the optimal quantum annealing to reduce E and are connected by b splines. After the system reaches its final state, fidelity F is also calculated to measure the accuracy of the obtained many-body state. F represents the probability of finding the system in one of the ground states, which differs from the approximation ratio. The approximation ratio R is also calculated at the end of the protocol to measure the quality of the solution.

APPENDIX C: OPTIMAL CONTROL THEORY METHODS

Optimal control theory is a mathematical framework that helps determine the best way to control a dynamical system by finding parameters that extremize a specific objective function [69]. It involves solving an optimization problem by adjusting control inputs over time while considering system dynamics and constraints. In physics, it has been used for shaping laser pulses, gate operations, and controlling chemical reactions [70–72]. For such optimization problems, gradient [73,74] and nongradient [75,76] methods can be used. Gradient-based methods rely on calculating the gradient of the objective function with respect to the parameters and on updating them iteratively in the direction of the negative gradient to reach an optimal solution. In contrast, nongradient methods are typically heuristic or evolutionary algorithms that iteratively explore the search space to reach an optimal solution. In this work, we use a combination of Broyden-Fletcher-Goldfarb-Shanno (BFGS; gradient-based) [77–80] and Nelder-Mead (non-gradient-based) methods [81]. A description of both methods is given below.

Broyden-Fletcher-Goldfarb-Shanno (BFGS). The BFGS method [77–80] is one such gradient-based method that approximates the inverse of the Hessian matrix of the objective function using information from the gradients of the function. At each iteration, the BFGS method calculates the change in the gradient and uses it to update the current estimate of the inverse Hessian. The new estimate of the inverse Hessian is then used to determine the search direction for the next iteration. The update formula is given by

$$H_{k+1} = \left(I - \rho_k s_k y_k^T\right) H_k \left(I - \rho_k y_k s_k^T\right) + \rho_k s_k s_k^T, \quad \text{(C1)}$$

where H_k is the inverse Hessian approximation at iteration k, $s_k = x_{k+1} - x_k$ is the difference between the current and previous estimates of the parameters, $y_k = \nabla f(x_{k+1}) - \nabla f(x_k)$ is the difference between the current and previous estimates of the gradient, $\rho_k = 1/(y_k^T s_k)$, and I is the identity matrix. The BFGS method typically starts with an initial inverse Hessian approximation, H_0, and iteratively updates the approximation until convergence is reached. The search direction at each iteration is given by $d_k = -H_k \nabla f(x_k)$, and the step size is determined using a line search method.

Nelder-Mead (NM). As for non-gradient-based methods, NM [81] is one such method: it iteratively searches for the minimum of an objective function. The NM method relies on exploring the simplex, a geometric figure with $n + 1$ vertices

in n dimensions. At each iteration, the algorithm evaluates the objective function at the vertices of the simplex and then performs a set of operations to update the simplex, such as reflection, expansion, contraction, or shrinkage, based on the values of the function. The method continues until a stopping criterion is met, such as reaching a maximum number of iterations or when the function value change is small. Initialization is a step in which a simplex is defined based on an initial set of $n + 1$ points in n-dimensional space. The function values at the vertices of the simplex are evaluated. The worst vertex of the simplex (i.e., the vertex with the highest function value) is then reflected through the centroid of the remaining vertices. If the reflected vertex has a lower function value than the second-worst vertex, the simplex is expanded in that direction. Otherwise, the reflected vertex is retained. If the reflected vertex has a higher function value than the worst vertex, the simplex is contracted towards the best vertex. If the contracted vertex has a lower function value than the worst vertex, it is retained. Otherwise, the simplex is shrunk towards the best vertex. If none of the previous steps improves the function value, the simplex is reduced toward the best vertex. Finally, the algorithm terminates when a stopping criterion is met, such as a maximum number of iterations or a slight change in the function value.

BFGS when compared to NM, can converge faster for smooth, convex functions. But at the same time, BFGS may get stuck in local minima or saddle points if the function has many local optima. Consequently, a combination of both methods is used in this work. The methods were applied sequentially in the order of BFGS-NM-BFGS (B-Nel-B), balancing exploration and exploitation. Starting with BFGS, the algorithm can quickly converge to a local minimum, and then NM is used to explore other regions of the objective function and get closer to the global optimal value. Finally, using BFGS can refine the solution and potentially converge to a better optimal value.

Parameters for NM and BFGS. Implementation in the code is done by the *optimize.minimize* function from the PYTHON library SciPy [82]. For both methods a convergence factor of 10^{-4} was set. The first layer of BFGS has a maximum iteration variable which was set to 3–6 depending on the size of the problem and was set to 2 for the last layer. In NM, both maximum iterations and maximum number of objective function evolution need to be fixed and were given a value of 300 to get convergence. These values were set based on the competition between the time it takes to evaluate one layer and the quality of the solution. In the case when the system was noisy, three times the number of runs were required as compared to the noiseless case.

APPENDIX D: SIMULATED ANNEALING

In this section, details of simulated annealing (SA) are discussed, which is used for benchmarking the quantum control method used in this work for the specific graphs. All the simulations for this method are performed by using an optimized SA from the SciPy library [82].

Method. Simulated annealing [83] is an algorithm used to solve optimization problems through a global search approach. It works by simulating the process of heating and
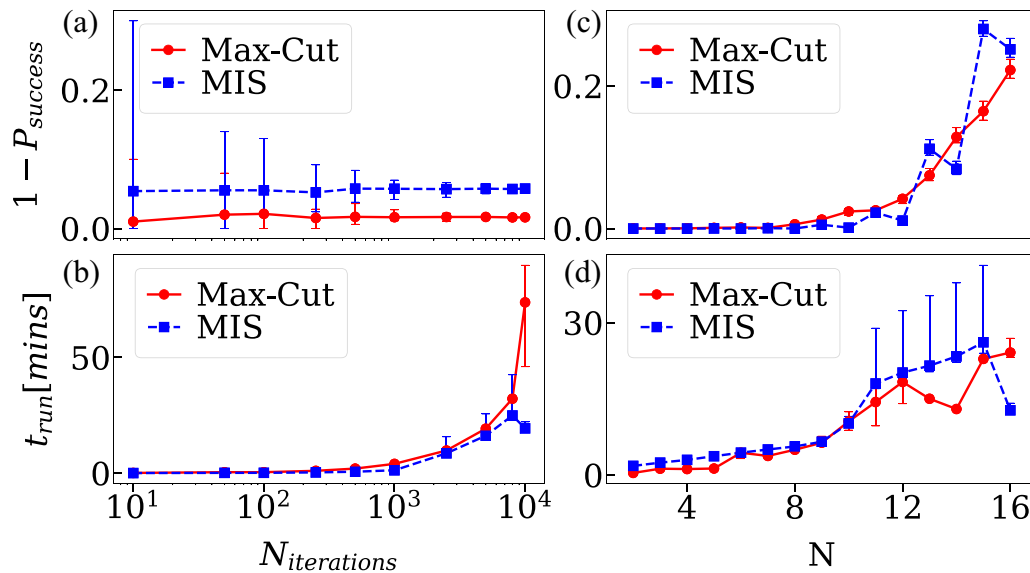
FIG. 6. Optimized simulated annealing for Max-Cut and MIS problems. (a) The plot of the probability of failure ($1 - P_{\text{success}}$) and (b) the run time (t_{run}) in minutes as a function of the number of iterations, $N_{\text{iterations}}$ (varying from 10 to 10 000). In (a) and (b), a weighted problem graph with 10 vertices and 13 edges is chosen. [(c) and (d)] The plot of the probability of failure $1 - P_{\text{success}}$ and run time t_{run} in minutes as a function of the number of vertices, N (varying from 2 to 16), for the unweighted problem graphs.

cooling a material to reduce defects and minimize energy. With each function call, the algorithm searches for a new solution point in the search space. If the new point has lower energy than the previous point, it is accepted as the new optimal value with a probability of 1. If not, the probability of acceptance depends on the temperature. As the temperature decreases, the algorithm becomes more selective and only accepts better solutions. The algorithm consists of multiple cycles, each defined by a one-time cooling process to lower the temperature from an initial to a final value. While the algorithm is stochastic, it can be optimized through a local search approach to reduce the search space and find an optimal solution. Time for the procedure can further be optimized by including the annealing schedule from fast simulated annealing (FSA) [42], which consists of semilocal searches with occasional long jumps.

Setting up of the numerics for SA. For both MIS and Max-Cut problems, the cost function was defined to provide the optimum solution at the function's minima. In this work for SA, the temperature is lowered from 0.4 to 0.01 under a distorted Cauchy-Lorentz distribution schedule (FSA) [42,84]. During the simulations, a limit of 2000 function calls per iteration and 50 cycles was fixed. However, the maximum

number of function calls was never reached and all cycles were completed. Each full run consists of multiple iterations, an optimum is reached at the end of each full run, and 50 such full runs were performed for statistics. The probability of success, P_{success}, and run time t_{run} to reach the optimum are measured as a function of the number of iterations for a weighted problem graph to find an optimum number of iterations for a single run as shown in Figs. 6(a) and 6(b). A total of 5000 iterations was chosen for subsequent simulations, as it is enough to decrease the error in the statistics of P_{success} and have a *reasonable* t_{run}. After fixing the number of iterations, 15 unweighted graphs were chosen of a varying number of vertices from 2 to 16, to study the effect of system size in finding the optimum and scaling of run time. As shown in Figs. 6(c) and 6(d), a general trend of increase in run time and a decrease in the probability of success is observed for both problems. An oscillatory behavior in the probability of success is also observed with varying system sizes for even and odd numbers of vertices in the graph. This behavior is attributed to the degeneracy of the solution space. If the optimum is highly degenerate, there is an increase in the probability of reaching the optimum stochastically as the solution space is larger.

[1] P. W. Shor, *Fault-Tolerant Quantum Computation* (IEEE, New York, 1996), pp. 56–65.

[2] D. Aharonov and M. Ben-Or, Fault-tolerant quantum computation with constant error, in *Proceedings of the 29th Annual ACM Symposium on Theory of Computing* (Association for Computing Machinery, New York, 1997), pp. 176–188.

[3] J. Preskill, Fault-tolerant quantum computation, in *Introduction to Quantum Computation and Information* (World Scientific, Singapore, 1998), pp. 213–269.

[4] D. Gottesman, Theory of fault-tolerant quantum computation, Phys. Rev. A **57**, 127 (1998).

[5] F. Bova, A. Goldfarb, and R. G. Melko, Commercial applications of quantum computing, EPJ Quantum Technol. **8**, 2 (2021).

[6] J. Preskill, Quantum computing in the NISQ era and beyond, Quantum **2**, 79 (2018).

[7] K. Bharti, A. Cervera-Lierta, T. H. Kyaw, T. Haug, S. Alperin-Lea, A. Anand, M. Degroote, H. Heimonen, J. S. Kottmann,

T. Menke *et al.*, Noisy intermediate-scale quantum algorithms, Rev. Mod. Phys. **94**, 015004 (2022).

[8] S. Aaronson, BQP and the polynomial hierarchy, in *Proceedings of the 42nd ACM Symposium on Theory of Computing* (Association for Computing Machinery, New York, 2010), pp. 141–150.

[9] A. J. Daley, I. Bloch, C. Kokail, S. Flannigan, N. Pearson, M. Troyer, and P. Zoller, Practical quantum advantage in quantum simulation, Nature (London) **607**, 667 (2022).

[10] S. Boixo, S. V. Isakov, V. N. Smelyanskiy, R. Babbush, N. Ding, Z. Jiang, M. J. Bremner, J. M. Martinis, and H. Neven, Characterizing quantum supremacy in near-term devices, Nat. Phys. **14**, 595 (2018).

[11] R. M. Karp, *On the Computational Complexity of Combinatorial Problems* (Wiley, New York, 1975), pp. 45–68.

[12] C. Papadimitriou and M. Yannakakis, Optimization, approximation, and complexity classes, in *Proceedings of the 20th Annual ACM Symposium on Theory of Computing* (Association for Computing Machinery, New York, 1988), pp. 229–234.

[13] M. P. Harrigan, K. J. Sung, M. Neeley, K. J. Satzinger, F. Arute, K. Arya, J. Atalaya, J. C. Bardin, R. Barends, S. Boixo *et al.*, Quantum approximate optimization of non-planar graph problems on a planar superconducting processor, Nat. Phys. **17**, 332 (2021).

[14] G. Pagano, A. Bapat, P. Becker, K. S. Collins, A. De, P. W. Hess, H. B. Kaplan, A. Kyprianidis, W. L. Tan, C. Baldwin *et al.*, Quantum approximate optimization of the long-range Ising model with a trapped-ion quantum simulator, Proc. Natl. Acad. Sci. USA **117**, 25396 (2020).

[15] T. Graham, Y. Song, J. Scott, C. Poole, L. Phuttitarn, K. Jooya, P. Eichler, X. Jiang, A. Marra, B. Grinkemeyer *et al.*, Multiqubit entanglement and algorithms on a neutral-atom quantum computer, Nature (London) **604**, 457 (2022).

[16] J.-D. Cho, S. Raje, and M. Sarrafzadeh, Fast approximation algorithms on maxcut, k-coloring, and k-color ordering for VLSI applications, IEEE Trans. Comput. **47**, 1253 (1998).

[17] D. S. Hochbaum and A. Pathria, Analysis of the greedy approach in problems of maximum k-coverage, Naval Res. Logistics **45**, 615 (1998).

[18] M. Perelshtein, A. Sagingalieva, K. Pinto, V. Shete, A. Pakhomchik, A. Melnikov, F. Neukart, G. Gesek, A. Melnikov, and V. Vinokur, Practical application-specific advantage through hybrid quantum computing, arXiv:2205.04858.

[19] S. Butenko, Maximum independent set and related problems, with applications, Ph.D. thesis, University of Florida, 2003.

[20] C. Scheideler, A. Richa, and P. Santi, An $O(\log n)$ dominating set protocol for wireless ad-hoc networks under the physical interference model, in *Proceedings of the Ninth ACM International Symposium on Mobile Ad Hoc Networking and Computing* (Association for Computing Machinery, New York, 2008), pp. 91–100.

[21] M. H. Devoret, A. Wallraff, and J. M. Martinis, Superconducting qubits: A short review, arXiv:cond-mat/0411174.

[22] A. Härter, A. Krükow, A. Brunner, and J. Hecker Denschlag, Long-term drifts of stray electric fields in a Paul trap, Appl. Phys. B **114**, 275 (2014).

[23] C. D. Bruzewicz, J. Chiaverini, R. McConnell, and J. M. Sage, Trapped-ion quantum computing: Progress and challenges, Appl. Phys. Rev. **6**, 021314 (2019).

[24] A. Browaeys and T. Lahaye, Many-body physics with individually controlled Rydberg atoms, Nat. Phys. **16**, 132 (2020).

[25] T. F. Gallagher, Rydberg atoms, Rep. Prog. Phys. **51**, 143 (1988).

[26] C. Gross and I. Bloch, Quantum simulations with ultracold atoms in optical lattices, Science **357**, 995 (2017).

[27] S. Ebadi, A. Keesling, M. Cain, T. T. Wang, H. Levine, D. Bluvstein, G. Semeghini, A. Omran, J.-G. Liu, R. Samajdar *et al.*, Quantum optimization of maximum independent set using Rydberg atom arrays, Science **376**, 1209 (2022).

[28] E. Farhi, J. Goldstone, and S. Gutmann, A quantum approximate optimization algorithm, arXiv:1411.4028.

[29] S. Hadfield, Z. Wang, B. O'gorman, E. G. Rieffel, D. Venturelli, and R. Biswas, From the quantum approximate optimization algorithm to a quantum alternating operator ansatz, Algorithms **12**, 34 (2019).

[30] L. Bittel and M. Kliesch, Training variational quantum algorithms is NP-hard, Phys. Rev. Lett. **127**, 120502 (2021).

[31] L. Bittel, S. Gharibian, and M. Kliesch, The optimal depth of variational quantum algorithms is QCMA-hard to approximate, in *Proceedings of the 38th Computational Complexity Conference (CCC 2023)*, Leibniz International Proceedings in Informatics (LIPIcs) Vol. 264 (Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2023), pp. 34:1–34:24.

[32] F. Barahona, M. Grötschel, M. Jünger, and G. Reinelt, An application of combinatorial optimization to statistical physics and circuit layout design, Oper. Res. **36**, 493 (1988).

[33] F. Glover, G. Kochenberger, and Y. Du, A tutorial on formulating and using QUBO models, arXiv:1811.11538.

[34] A. Lucas, Ising formulations of many np problems, Front. Phys. **2**, 5 (2014).

[35] M. Kim, K. Kim, J. Hwang, E.-G. Moon, and J. Ahn, Rydberg quantum wires for maximum independent set problems, Nat. Phys. **18**, 755 (2022).

[36] M. Lanthaler, C. Dlaska, K. Ender, and W. Lechner, Rydberg-blockade-based parity quantum optimization, Phys. Rev. Lett. **130**, 220601 (2023).

[37] X. Qiu, P. Zoller, and X. Li, Programmable quantum annealing architectures with Ising quantum wires, PRX Quantum **1**, 020311 (2020).

[38] W. Lechner, P. Hauke, and P. Zoller, A quantum annealing architecture with all-to-all connectivity from local interactions, Sci. Adv. **1**, e1500838 (2015).

[39] B. N. Clark, C. J. Colbourn, and D. S. Johnson, Unit disk graphs, Discr. Math. **86**, 165 (1990).

[40] E. Urban, T. A. Johnson, T. Henage, L. Isenhower, D. Yavuz, T. Walker, and M. Saffman, Observation of Rydberg blockade between two atoms, Nat. Phys. **5**, 110 (2009).

[41] M.-T. Nguyen, J.-G. Liu, J. Wurtz, M. D. Lukin, S.-T. Wang, and H. Pichler, Quantum optimization with arbitrary connectivity using Rydberg atom arrays, PRX Quantum **4**, 010316 (2023).

[42] H. Szu and R. Hartley, Fast simulated annealing, Phys. Lett. A **122**, 157 (1987).

[43] A. M. Kaufman and K.-K. Ni, Quantum science with optical tweezer arrays of ultracold atoms and molecules, Nat. Phys. **17**, 1324 (2021).

[44] J. Zeiher, R. Van Bijnen, P. Schauß, S. Hild, J.-y. Choi, T. Pohl, I. Bloch, and C. Gross, Many-body interferometry of a Rydberg-dressed spin lattice, Nat. Phys. **12**, 1095 (2016).

[45] A. Gaëtan, Y. Miroshnychenko, T. Wilk, A. Chotia, M. Viteau, D. Comparat, P. Pillet, A. Browaeys, and P. Grangier, Observation of collective excitation of two individual atoms in the Rydberg blockade regime, Nat. Phys. **5**, 115 (2009).

[46] R. Löw, H. Weimer, J. Nipper, J. B. Balewski, B. Butscher, H. P. Büchler, and T. Pfau, An experimental and theoretical guide to strongly interacting Rydberg gases, J. Phys. B **45**, 113001 (2012).

[47] H. Pichler, S.-T. Wang, L. Zhou, S. Choi, and M. D. Lukin, Computational complexity of the Rydberg blockade in two dimensions, arXiv:1809.04954.

[48] H. Rabitz, R. de Vivie-Riedle, M. Motzkus, and K. Kompa, Whither the future of controlling quantum phenomena?, Science **288**, 824 (2000).

[49] J. Kelly, R. Barends, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, A. G. Fowler, I.-C. Hoi, E. Jeffrey, A. Megrant, J. Mutus, C. Neill, P. J. J. O'Malley, C. Quintana, P. Roushan, D. Sank, A. Vainsencher, J. Wenner, T. C. White *et al.*, Optimal quantum control using randomized benchmarking, Phys. Rev. Lett. **112**, 240504 (2014).

[50] J. Li, X. Yang, X. Peng, and C.-P. Sun, Hybrid quantum-classical approach to quantum optimal control, Phys. Rev. Lett. **118**, 150503 (2017).

[51] T. Caneva, T. Calarco, and S. Montangero, Chopped random-basis quantum optimization, Phys. Rev. A **84**, 022326 (2011).

[52] C. Ferrie and O. Moussa, Robust and efficient in situ quantum control, Phys. Rev. A **91**, 052306 (2015).

[53] R. Mukherjee, F. Sauvage, H. Xie, R. Löw, and F. Mintert, Preparation of ordered states in ultra-cold gases using Bayesian optimization, New J. Phys. **22**, 075001 (2020).

[54] R. Mukherjee, H. Xie, and F. Mintert, Bayesian optimal control of Greenberger-Horne-Zeilinger states in Rydberg lattices, Phys. Rev. Lett. **125**, 203603 (2020).

[55] M. X. Goemans and D. P. Williamson, .879-approximation algorithms for MAXCUT and MAX2SAT, in *Proceedings of the 26th Annual ACM Symposium on Theory of Computing* (Association of Computing Machinery, New York, 1994), pp. 422–431.

[56] V. V. Vazirani, *Approximation Algorithms* (Springer, Berlin, 2001), pp. 2–5.

[57] B. Mohar and S. Poljak, Eigenvalues and the max-cut problem, Czech. Math. J. **40**, 343 (1990).

[58] S. Ben-David, A. M. Childs, A. Gilyén, W. Kretschmer, S. Podder, and D. Wang, Symmetries, graph properties, and quantum speedups, in *Proceedings of the 2020 IEEE 61st Annual Symposium on Foundations of Computer Science (FOCS)* (IEEE, New York, 2020), pp. 649–660.

[59] Z.-G. Feng, L.-J. Zhang, J.-M. Zhao, C.-Y. Li, and S.-T. Jia, Lifetime measurement of ultracold caesium Rydberg states, J. Phys. B **42**, 145303 (2009).

[60] N. Šibalić, J. D. Pritchard, C. S. Adams, and K. J. Weatherill, ARC: An open-source library for calculating properties of alkali Rydberg atoms, Comput. Phys. Commun. **220**, 319 (2017).

[61] C. Maurer, A. Jesacher, S. Bernet, and M. Ritsch-Marte, What spatial light modulators can do for optical microscopy, Laser Photonics Rev. **5**, 81 (2011).

[62] M. Saffman, T. G. Walker, and K. Mølmer, Quantum information with Rydberg atoms, Rev. Mod. Phys. **82**, 2313 (2010).

[63] Y. Wu, S. Kolkowitz, S. Puri, and J. D. Thompson, Erasure conversion for fault-tolerant quantum computing in alkaline earth Rydberg atom arrays, Nat. Commun. **13**, 4657 (2022).

[64] T. M. Graham, M. Kwon, B. Grinkemeyer, Z. Marra, X. Jiang, M. T. Lichtman, Y. Sun, M. Ebert, and M. Saffman, Rydberg-mediated entanglement in a two-dimensional neutral atom qubit array, Phys. Rev. Lett. **123**, 230501 (2019).

[65] I. Cong, H. Levine, A. Keesling, D. Bluvstein, S.-T. Wang, and M. D. Lukin, Hardware-efficient, fault-tolerant quantum computation with Rydberg atoms, Phys. Rev. X **12**, 021049 (2022).

[66] F. Sauvage and F. Mintert, Optimal quantum control with poor statistics, PRX Quantum **1**, 020322 (2020).

[67] E. Boros and P. L. Hammer, Pseudo-Boolean optimization, Disc. Appl. Math. **123**, 155 (2002).

[68] J. R. Johansson, P. D. Nation, and F. Nori, Qutip: An open-source PYTHON framework for the dynamics of open quantum systems, Comput. Phys. Commun. **183**, 1760 (2012).

[69] F. L. Lewis, D. Vrabie, and V. L. Syrmos, *Optimal Control* (John Wiley & Sons, New York, 2012).

[70] J. Werschnik and E. Gross, Quantum optimal control theory, J. Phys. B **40**, R175 (2007).

[71] F. Dolde, V. Bergholm, Y. Wang, I. Jakobi, B. Naydenov, S. Pezzagna, J. Meijer, F. Jelezko, P. Neumann, T. Schulte-Herbrüggen *et al.*, High-fidelity spin entanglement using optimal control, Nat. Commun. **5**, 3371 (2014).

[72] C. Winterfeldt, C. Spielmann, and G. Gerber, Colloquium: Optimal control of high-harmonic generation, Rev. Mod. Phys. **80**, 117 (2008).

[73] E. Polak, *Optimization: Algorithms and Consistent Approximations* (Springer Science & Business Media, 2012), Vol. 124.

[74] L. Hasdorff, *Gradient Optimization and Nonlinear Control* (Wiley, New York, 1976).

[75] A. R. Conn, K. Scheinberg, and L. N. Vicente, *Introduction to Derivative-Free Optimization* (SIAM, Philadelphia, 2009).

[76] W. Hare, J. Nutini, and S. Tesfamariam, A survey of non-gradient optimization methods in structural engineering, Adv. Eng. Software **59**, 19 (2013).

[77] C. G. Broyden, The convergence of a class of double-rank minimization algorithms 1. General considerations, IMA J. Appl. Math. **6**, 76 (1970).

[78] R. Fletcher, A new approach to variable metric algorithms, Comput. J. **13**, 317 (1970).

[79] D. Goldfarb, A family of variable metric updates derived by variational means, Math. Comput. **24**, 23 (1970).

[80] D. F. Shanno, Conditioning of quasi-Newton methods for function minimization, Math. Comput. **24**, 647 (1970).

[81] J. A. Nelder and R. Mead, A simplex method for function minimization, Comput. J. **7**, 308 (1965).

[82] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright *et al.*, SciPy 1.0: Fundamental algorithms for scientific computing in PYTHON, Nat. Methods **17**, 261 (2020).

[83] D. Bertsimas and J. Tsitsiklis, Simulated annealing, Stat. Sci. **8**, 10 (1993).

[84] Y. Xiang, D. Sun, W. Fan, and X. Gong, Generalized simulated annealing algorithm and its application to the Thomson model, Phys. Lett. A **233**, 216 (1997).