

SISSA

**Scuola
Internazionale
Superiore di
Studi Avanzati**

Physics Area - PhD course in
Theory and Numerical Simulation of Condensed Matter

Quantum Optimal Control and Variational Algorithms for Optimization

Thesis submitted for the degree of
Doctor Philosophiae

Candidate:
Ruiyi Wang

Advisor:
Giuseppe E. Santoro

Academic Year 2025-2026



Abstract

Recent advances in quantum technologies have intensified the search for algorithms that can use quantum resources effectively. Quantum annealing and hybrid quantum–classical variational algorithms are two prominent approaches in this setting. Conventional quantum annealing follows a prescribed Hamiltonian interpolation and may therefore be limited by small spectral gaps along that path. This Thesis investigates how algorithmic performance changes when three elements are varied: the annealing schedule (Chapter 2), the independent parameters of a digitized evolution (Chapter 3), and the encoding of the problem (Chapter 4).

For weighted MaxCut instances with anomalously small annealing gaps, optimized smooth schedules outperform linear digitized annealing at equal circuit depth. Rather than restoring adiabaticity or enlarging the gap, the dynamics deliberately leaves the instantaneous ground state and later returns the population to the target state. Smoothness allows these digital solutions to retain their quality as continuous-time controls and when transferred between related instances.

The schedule restriction is then removed for a frustrated Ising ring whose annealing gap closes exponentially with system size. Optimizing all circuit angles independently yields numerically exact ground-state preparation at the critical depth $P_1^{\text{cr}} = (N^2 - 1)/4$. This quadratic threshold is explained by the dimension and symmetries of the reachable free-fermion manifold, demonstrating that controllability, rather than the gap of a particular interpolation, determines the required number of digital controls.

Finally, the same ring family is reformulated through the parity mapping. For the original frustrated ring, Parity-QAOA reaches the exact ground state using only single-qubit rotations, without applying the constraint unitary. The number of layers equals the number of distinct coupling values and does not grow with system size. To clarify the role of constraints, an equal-magnitude signed ring is also studied, for which the constraint unitary is essential for preparing the target ground state.

Together, these results identify the shape, controllability, and encoding of a quantum protocol as resources for overcoming limitations imposed by more restricted routes to ground-state preparation.

Acknowledgements

I am deeply grateful to my supervisor, Giuseppe E. Santoro, for his scientific guidance, careful feedback, and continued support throughout my doctoral studies. My thanks also go to the faculty and my fellow students in the SISSA Condensed Matter group, who patiently helped me build much of my background in physics from the foundations. I would like to thank Giovanni Pecci for guiding me through my first research project, and Pietro Torta, Vincenzo Roberto Arezzo, and Kiran Thengil for their collaboration and discussions. I am grateful to Wolfgang Lechner for hosting my research visit to his group in Innsbruck, and to Glen Bigan Mbeng and Anita Weidinger for their collaboration and discussions during and after it.

I would like to thank my parents, Xiangbin Wang and Jing Chen, for being my constant source of support and enabling me to focus fully on the path I have chosen. I also wish to acknowledge my therapist, Minqian Zhang, for her practical strategies and consistent availability during my doctoral studies. I would also like to express my gratitude to my master's supervisor, Achilleas Lazarides, whose mentorship and encouragement have continued well beyond my master's studies.

This thesis is also dedicated to the memory of my grandfathers and my great-grandfather, whose values and wisdom I will carry forward.

Publications

This thesis reports results from the following publications:

- 1) G. Pecci, R. Wang, P. Torta, G. B. Mbeng, and G. E. Santoro, *Beyond quantum annealing: optimal control solutions to MaxCut problems*, Quantum Sci. Technol. **9**, 045013 (2024).
- 2) R. Wang, V. R. Arezzo, K. Thengil, G. Pecci, and G. E. Santoro, *From exponential to quadratic: optimal control for a frustrated Ising ring model*, Quantum Sci. Technol. **10**, 035052 (2025).
- 3) R. Wang, A. Weidinger, G. B. Mbeng, W. Lechner, and G. E. Santoro, *Parity Mapping for Quantum Optimization on Frustrated Ising Rings*, arXiv:2608.12213 (2026).

During the course of the PhD, I have further co-authored the following publication:

- 1) V. R. Arezzo, R. Wang, K. Thengil, G. Pecci, and G. E. Santoro, *Digital controllability of transverse field Ising chains*, Phys. Rev. A **113**, 012610 (2026).

Contents

1. Introduction	3
1.1. Quantum computing and quantum optimization	3
1.1.1. Example problem: MaxCut	6
1.1.2. Computational complexity	7
1.2. From continuous-time to digitized protocols	8
1.2.1. Quantum annealing and the adiabatic theorem	8
1.2.2. Grover search: the schedule as a resource	13
1.2.3. Shortcuts to adiabaticity	16
1.2.4. Variational quantum algorithms, dQA, and QAOA	17
1.2.5. The uniform Ising chain: continuous annealing, dQA, and QAOA	20
1.3. Quantum optimal control	22
1.3.1. Theory of quantum optimal control	23
1.3.2. Numerical methods: GRAPE and CRAB	26
1.4. Towards more hardware-aware quantum algorithms	28
1.5. Outline of the Thesis	29
2. Optimal Control Solutions to MaxCut Problems	31
2.1. Introduction	32
2.2. From QA to variational quantum algorithms	34
2.2.1. QA and optimal control	34
2.2.2. QAOA	38
2.3. The MaxCut problem	41
2.4. Results	43
2.4.1. Comparison of different methods	43
2.4.2. Eigenstate population and shortcut to adiabaticity	46
2.4.3. Transferability among hard instances	49
2.4.4. Smoothness of the solutions and continuous-time annealing	50
2.5. Conclusions	53
3. Optimal Control of a Frustrated Ising Ring Model	57
3.1. Introduction	57
3.2. The frustrated-ring model	59
3.3. Methodology	61
3.3.1. Digitized QA dynamics with CRAB	63

3.3.2. QAOA	65
3.4. Results	66
3.4.1. Digital controllability through QAOA-RAND	66
3.4.2. dQA-CRAB smooth optimal schedules	68
3.4.3. Populations of the instantaneous eigenstates in dQA-CRAB	69
3.4.4. Annealing time comparison	73
3.5. Conclusions and perspectives	73
4. Parity Mapping for Quantum Optimization	77
4.1. Introduction	77
4.2. The Parity Transformation	78
4.3. Parity-QAOA	81
4.4. Frustrated Ising ring models	82
4.5. Results	83
4.5.1. Minimal Energy gap in QA	84
4.5.2. Parity-QAOA	85
4.6. Conclusions	89
5. Conclusions	93
A. Technical details for Chapter 2	97
A.1. Gradient-based optimization	97
A.2. The FOURIER approach of Zhou <i>et al.</i>	98
A.3. Local neighborhood of different optimal solutions	100
A.4. Performance over different instances	101
B. Technical details for Chapter 3	103
B.1. Low-energy spectrum in the Nambu formalism	103
B.2. Digitized dynamics in the Nambu formalism	104
B.3. Residual energy	108
B.4. Gradients of the residual energy	109
C. Technical details for Chapter 4	111
C.1. Simulation details	111
C.2. Ground state probability for Parity-QAOA at arbitrary size	111
C.3. Local controllability	112
C.4. Optimal parameters for standard QAOA	115

List of Acronyms

Acronym	Full Form
NISQ	Noisy Intermediate-Scale Quantum
NP	Nondeterministic Polynomial-time
QA	Quantum Annealing
AQC	Adiabatic Quantum Computation
dQA	digitized Quantum Annealing
QAOA	Quantum Approximate Optimization Algorithm
VQA	Variational Quantum Algorithm
PQC	Parameterized Quantum Circuit
CRAB	Chopped RAndom Basis
dCRAB	dressed Chopped RAndom Basis
STA	Shortcut To Adiabaticity
CD	Counter-Diabatic
QSL	Quantum Speed Limit
LZ	Landau–Zener
QUBO	Quadratic Unconstrained Binary Optimization
MaxCut	Maximum Cut
LHZ	Lechner–Hauke–Zoller (parity architecture)
TFIM	Transverse-Field Ising Model
BdG	Bogoliubov–de Gennes
BCS	Bardeen–Cooper–Schrieffer

1. Introduction

1.1. Quantum computing and quantum optimization

The modern formulation of quantum mechanics dates back a century, to Heisenberg’s matrix mechanics in 1925 and Schrödinger’s wave equation in 1926 [1, 2]. The state of a system can be expanded in a basis of classical configurations. For N spins,

$$|\Psi\rangle = \sum_{\sigma_1, \dots, \sigma_N = \uparrow, \downarrow} c_{\sigma_1 \dots \sigma_N} |\sigma_1 \dots \sigma_N\rangle, \quad (1.1)$$

where each configuration is assigned a complex amplitude $c_{\sigma_1 \dots \sigma_N}$, and a measurement returns the configuration $(\sigma_1, \dots, \sigma_N)$ with probability $P_{\sigma_1 \dots \sigma_N} = |c_{\sigma_1 \dots \sigma_N}|^2$ [3, 4]. Explicitly storing a generic pure state on a classical computer therefore requires 2^N complex numbers. This exponential growth makes brute-force state-vector simulation impractical beyond a few tens of spins, although special states with sufficient structure can admit compact classical representations.

In the 1980s, Feynman proposed to invert this difficulty: since quantum systems are hard to simulate classically, a controllable quantum system could itself serve as the simulator [5]. Deutsch subsequently showed that such a machine can be made universal, defining a full model of computation rather than a dedicated simulator [6]. A decisive step came in 1994, when Shor found a quantum algorithm that factors integers in polynomial time [7]. The fact that factoring is not a quantum-mechanical problem, yet admits a superpolynomial quantum speedup over the best currently known classical factoring algorithms, demonstrated that quantum computers could address problems beyond quantum simulation and helped establish quantum computing as an independent research field.

Realizing such a machine experimentally is a formidable task. A qubit must be isolated well enough to remain coherent, yet coupled well enough to be controlled and measured, and these two requirements are in mutual tension. Four representative leading platforms illustrate the different trade-offs currently faced by quantum hardware.

Trapped ions. Ions confined in electromagnetic traps encode a qubit in two long-lived internal states. Single-qubit rotations are driven by laser pulses, while two ions are entangled through the vibrational modes of the chain they share, following the scheme of Cirac and Zoller [8]. Trapped-ion systems combine long coherence times

with high-fidelity gates [9]; in 2026, a 98-qubit processor with all-to-all connectivity reported a two-qubit gate fidelity of 99.92% [10]. Gate times are comparatively long, however, and scaling requires architectures that shuttle ions between interaction zones or connect separate traps through photonic links.

Superconducting circuits. In this platform the qubit is not an atom but an electrical circuit, cooled to millikelvin temperatures, whose two lowest energy levels serve as the logical states; the transmon is the standard design [11]. Gates are driven by microwave pulses on nanosecond time scales, and the circuits are fabricated with standard lithographic techniques [12]. Superconducting processors have produced both the sampling experiment of 2019 [13] and recent demonstrations of error correction below threshold [14, 15]. Coherence times remain short compared to atomic qubits, and each qubit requires dedicated control wiring, so that the classical electronics surrounding the chip poses a significant engineering challenge.

Photons. A photon can carry a qubit in its polarization or in its optical path. Since photons interact only weakly with their environment, they can propagate over long distances and are natural carriers for quantum communication. Optical processing can take place at room temperature, although high-performance sources and detectors may require cooling. The same weak interaction is also a limitation: in standard linear-optical schemes, entangling gates are induced through measurement and succeed probabilistically [16]. Photonic experiments have nevertheless reached the regime of quantum computational advantage in boson sampling [17].

Neutral atoms. Single atoms held in arrays of optical tweezers can be arranged in nearly arbitrary geometries. In 2025, a tweezer array trapped more than 6,100 atomic qubits while demonstrating long coherence and global single-qubit control [18]. Entangling gates are realized by exciting atoms to Rydberg states, which switches on a strong interaction between neighbors. The platform has progressed rapidly from many-body experiments [19, 20] to algorithms executed on logical, error-corrected qubits [21, 22]. Atom loss and two-qubit gate fidelity remain the main limitations.

None of these devices is yet a general-purpose fault-tolerant computer. Preskill coined the term noisy intermediate-scale quantum (NISQ) for processors that operate without full error correction and are limited to relatively shallow circuits because every gate adds noise [23]. Their physical-qubit counts now vary widely across platforms, so scale alone does not characterize their computational capability. Algorithms suited to this regime are typically hybrid quantum-classical schemes: a quantum circuit with adjustable parameters prepares a trial state, a classical optimizer updates the parameters, and the loop is iterated until convergence. The variational quantum eigensolver (VQE) [24] and the Quantum Approximate Optimization Algorithm (QAOA) [25] are the best-known examples, and the latter plays a central role in this Thesis.

The route toward large-scale fault-tolerant quantum computing relies on quantum error correction, in which logical qubits are encoded across multiple physical qubits to protect quantum information from errors [26, 27]. Early developments, from Shor’s code using nine qubits through CSS codes and the stabilizer formalism, established the modern theory of quantum error correction [28–31]. Leading approaches include topological surface and color codes, subsystem codes based on gauge degrees of freedom, qLDPC codes with finite rate that can reduce asymptotic qubit overhead, and bosonic codes such as GKP, which encode information in oscillator degrees of freedom [32–36]. Once the physical error rate falls below the error-correction threshold, increasing the code distance can further suppress the logical error rate [37, 38]. Recent experiments have crossed this threshold in superconducting processors [14, 15] and have executed encoded algorithms in neutral-atom arrays [21, 22]. The transition from NISQ to fault-tolerant computing is nevertheless better viewed as a crossover rather than a sharp transition, and the efficiency of algorithms remains essential in either regime.

These considerations motivate the joint development of hardware and algorithms. Advances in hardware expand the range of executable circuits, while advances in algorithms reduce the resources required for a given task through, for instance, improved schedules, shallower circuits, or more efficient encodings. Applications span quantum simulation, quantum chemistry, cryptography, quantum machine learning, and classical optimization [5, 7, 39–41]. In this Thesis, we take quantum optimization to be a prominent testbed for addressing the challenge of using quantum resources effectively.

We start with a classical combinatorial optimization problem: among a finite set of candidate solutions, the task is to find one that minimizes or maximizes a cost function [42, 43]. Encoding each candidate as a bit string makes the cost function a rule that assigns a real value to every string. To formulate this problem quantum mechanically, the bits are rewritten as Ising variables $z_j = \pm 1$ and promoted to $\hat{\sigma}^z$ operators. The resulting target Hamiltonian is diagonal in the computational basis, and its eigenvalues reproduce the cost-function values up to discarded constant terms. Many such problems can be reduced, sometimes using penalty terms, quadratization, or auxiliary variables, to the Ising form [44, 45]

$$\hat{H}_{\text{targ}} = \hat{H}_z = \sum_{j < j'} J_{jj'} \hat{\sigma}_j^z \hat{\sigma}_{j'}^z + \sum_j h_j \hat{\sigma}_j^z, \quad (1.2)$$

whose ground states encode the optimal strings. The substitution $x_j = (1 - z_j)/2$ converts the same quadratic cost function into a quadratic unconstrained binary optimization (QUBO) model, and the inverse substitution recovers the Ising model, up to a constant energy shift [44, 45]. Ising and QUBO are therefore equivalent representations rather than distinct application problems. The classical cost function fixes the target spectrum; the quantum task is to prepare a state with low energy in that spectrum.

1.1.1. Example problem: MaxCut

As a concrete example of the Ising encoding in Eq. (1.2), we consider MaxCut. It places couplings on the edges of a graph. The weighted three-regular instances analyzed in Chapter 2 are specified in Section 2.3 [46].

MaxCut. MaxCut is defined on a graph $\mathcal{G}$ with N vertices and edge set $\mathcal{E}$ [48]. A non-negative weight $J_{jj'}$ is assigned to every edge $(jj') \in \mathcal{E}$. A cut divides the vertices into two sets, and its value is the total weight of the edges whose endpoints belong to different sets. A maximum cut is therefore a cut with the largest possible value for the given graph. Figure 1.1 shows an unweighted graph, with $J_{jj'} = 1$ for every edge $(jj') \in \mathcal{E}$. It presents a maximum cut in which five of the six edges are cut. We assign an Ising variable z_j to each vertex j : $z_j = +1$ if vertex j belongs to one set and $z_j = -1$ if it belongs to the other. This gives the objective function

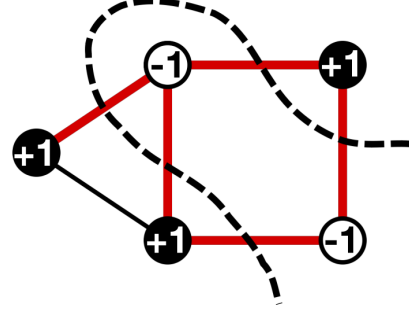


Figure 1.1.: A maximum cut of an unweighted graph. The dashed curve separates the two vertex sets; five of the six edges are cut. Adapted from Fig. 1.2(b) of Ref. [47].

$$C_{\text{cut}}(\mathbf{z}) = \frac{1}{2} \sum_{(jj') \in \mathcal{E}} J_{jj'} (1 - z_j z_{j'}) . \quad (1.3)$$

Each edge contributes its weight precisely when its endpoint labels differ. The objective is to maximize $C_{\text{cut}}(\mathbf{z})$ over all assignments $\mathbf{z} \in \{-1, +1\}^N$, or equivalently over all ways of dividing the vertices into two sets. To write this objective as a quantum problem, let $|\mathbf{z}\rangle$ be the computational-basis state labelled by the Ising variables $z_j = \pm 1$, such that $\hat{\sigma}_j^z |\mathbf{z}\rangle = z_j |\mathbf{z}\rangle$. Replacing each z_j by $\hat{\sigma}_j^z$ gives the cost operator

$$\hat{C}_{\text{cut}} = \frac{1}{2} \sum_{(jj') \in \mathcal{E}} J_{jj'} (1 - \hat{\sigma}_j^z \hat{\sigma}_{j'}^z) . \quad (1.4)$$

For every computational-basis state, $\hat{C}_{\text{cut}} |\mathbf{z}\rangle = C_{\text{cut}}(\mathbf{z}) |\mathbf{z}\rangle$. MaxCut searches for the largest cut value, whereas a ground state has the lowest energy. We therefore define the Hamiltonian as the negative of the cost operator,

$$\hat{H}_{\text{targ}} = \hat{H}_z = -\hat{C}_{\text{cut}} = \frac{1}{2} \sum_{(jj') \in \mathcal{E}} J_{jj'} (\hat{\sigma}_j^z \hat{\sigma}_{j'}^z - 1) . \quad (1.5)$$

It follows that $\hat{H}_z |\mathbf{z}\rangle = -C_{\text{cut}}(\mathbf{z}) |\mathbf{z}\rangle$, so the cuts with the largest classical value are the ground states of the Hamiltonian [44, 49].

1.1.2. Computational complexity

Computational complexity quantifies how the resources required by an algorithm scale with the encoded input length n for a family of instances [50]. We use worst-case running time and call an algorithm efficient when this time is bounded by a polynomial $\mathcal{O}(n^k)$ for some constant k . Complexity classes are normally formulated for decision problems, whose answer is yes or no. Figure 1.2 summarizes the classical and quantum complexity notions used below.

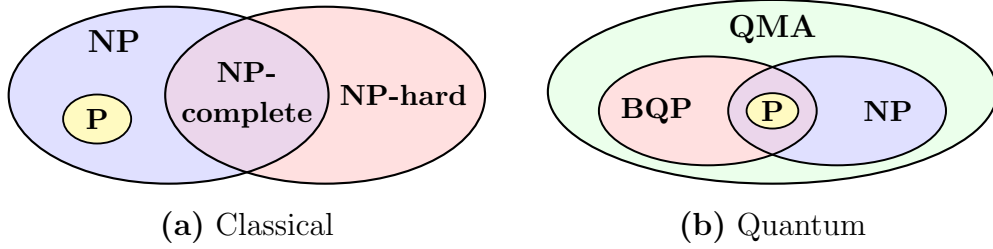


Figure 1.2.: Schematic relations among the complexity notions used in this Thesis. (a) $\mathbf{P}$ is contained in $\mathbf{NP}$; $\mathbf{NP}$ -complete problems are both in $\mathbf{NP}$ and $\mathbf{NP}$ -hard, while $\mathbf{NP}$ -hard optimization problems need not belong to $\mathbf{NP}$ [51]. (b) Both $\mathbf{BQP}$ and $\mathbf{NP}$ are contained in $\mathbf{QMA}$, and $\mathbf{P}$ is contained in the intersection of $\mathbf{BQP}$ and $\mathbf{NP}$ [52]. The shapes indicate only these known containments: they are not to scale and do not assert that any containment is strict or that $\mathbf{BQP} \cap \mathbf{NP} = \mathbf{P}$.

The class $\mathbf{P}$ contains decision problems solvable by a deterministic classical algorithm in polynomial time.

The class $\mathbf{NP}$ contains decision problems for which, given a proposed solution of polynomial size, one can verify in polynomial time whether it proves that the correct answer is yes. For example, the decision version of MaxCut asks whether a given graph has a cut whose total weight is at least a specified target value. A proposed cut divides the n vertices into two groups and requires $\mathcal{O}(n)$ bits. Even for a fully connected graph, its weight can be evaluated in at most $\mathcal{O}(n^2)$ time by summing the weights of the crossing edges and comparing the result with the target. Thus, checking a proposed cut is efficient, even though finding one may be difficult. Since every problem in $\mathbf{P}$ can be solved, and hence verified, in polynomial time, $\mathbf{P} \subseteq \mathbf{NP}$.

A problem is **NP-hard** if every problem in $\mathbf{NP}$ can be reduced to it in polynomial time. Here, reduced means reformulating any instance of an $\mathbf{NP}$ problem as an instance of this problem using only polynomial time, so that solving the reformulated instance also solves the original one. Consequently, an efficient method for solving an $\mathbf{NP}$ -hard problem could be used to solve every problem in $\mathbf{NP}$ efficiently.

A decision problem is **NP-complete** if it is both in $\mathbf{NP}$ and $\mathbf{NP}$ -hard. The decision version of MaxCut described above is $\mathbf{NP}$ -complete, whereas finding the maximum cut is an $\mathbf{NP}$ -hard optimization problem [53, 54]. MaxCut remains $\mathbf{NP}$ -hard for

cubic graphs [55], so the three-regular instances studied in Chapter 2 retain the worst-case hardness of the general problem.

Whether $P = NP$ remains open [53, 56].

BQP (Bounded Error Quantum Polynomial Time) is the class of decision problems that can be solved by a quantum computer in polynomial time with a bounded probability of error, conventionally with a success probability of at least $2/3$ for every input [57, 58]. BQP is regarded as the quantum analogue of P because both contain decision problems that can be solved in polynomial time, using deterministic classical computation in P and quantum computation with bounded error in BQP.

QMA (Quantum Merlin Arthur) contains decision problems for which, given a quantum state on a polynomial number of qubits as a proposed solution, a quantum computer can verify in polynomial time and with bounded error whether it proves that the correct answer is yes [59, 60]. QMA is regarded as the quantum analogue of NP because it replaces the classical proposed solution and classical verification used in NP with a quantum state and quantum verification.

Integer factorization illustrates these complexity classes. Given $N \in \mathbb{N}$, the task is to find primes $p_1, \dots, p_m$ satisfying $N = p_1 p_2 \cdots p_m$. The corresponding decision problem, denoted $\text{FACTOR}(N, k)$, asks whether N has a prime factor at most k . Shor's algorithm places this problem in BQP [7]. Moreover, $\text{FACTOR} \in NP \cap \text{coNP}$: when such a factor exists, a divisor allows this to be checked efficiently; when none exists, a complete prime factorization allows this to be checked, with primality verifiable in P [61].

These complexity classes describe the classical target problem. The spectrum of $\hat{H}_z$ is simply the set of values taken by its classical cost function. Quantum annealing and QAOA instead concern the dynamics used to prepare a low-energy state of the target Hamiltonian. Classical hardness motivates this search but does not by itself guarantee a quantum speedup. The remainder of this chapter therefore develops the dynamical and control tools used in the Thesis, beginning in Section 1.2.1 with quantum annealing and the adiabatic theorem.

1.2. From continuous-time to digitized protocols

1.2.1. Quantum annealing and the adiabatic theorem

Quantum annealing (QA) is a continuous-time approach to quantum optimization [62–65]. The solution of an optimization problem is encoded in the ground state of a target Hamiltonian $\hat{H}_{\text{targ}}$. The system is initialized in the ground state $|\Psi_{\text{mix}}\rangle$ of a mixing, or driver, Hamiltonian $\hat{H}_{\text{mix}}$, which is chosen so that this state

is easy to prepare. It is normally chosen not to commute with $\hat{H}_{\text{targ}}$, so that the evolution can move between candidate configurations. The system is then evolved under a time-dependent Hamiltonian that interpolates between the mixer and the target. A standard QA Hamiltonian is

$$\hat{H}(t) = s(t) \hat{H}_{\text{targ}} + (1 - s(t)) \hat{H}_{\text{mix}}, \quad s(0) = 0, \quad s(\tau) = 1, \quad (1.6)$$

where $s(t)$ is the annealing schedule and τ is the total annealing time. These boundary conditions give $\hat{H}(0) = \hat{H}_{\text{mix}}$ and $\hat{H}(\tau) = \hat{H}_{\text{targ}}$. The conventional choice is the linear schedule $s(t) = t/\tau$, although other schedules can traverse the same Hamiltonian path at a nonuniform rate.

The state at the end of the annealing evolution is

$$|\Psi(\tau)\rangle = \mathcal{T} \exp \left[-\frac{i}{\hbar} \int_0^\tau \hat{H}(t) dt \right] |\Psi_{\text{mix}}\rangle \equiv U(\tau, 0) |\Psi_{\text{mix}}\rangle, \quad (1.7)$$

where $\mathcal{T}$ denotes time ordering. The aim is to obtain a final state with a large overlap with the ground state of $\hat{H}_{\text{targ}}$, so that a measurement returns a solution of the optimization problem with high probability.

Let $E_0(t)$ and $E_1(t)$ denote the instantaneous ground-state and first-excited-state energies. Their separation and its minimum value along the annealing path are

$$\Delta(t) = E_1(t) - E_0(t), \quad \Delta_{\min} = \min_{t \in [0, \tau]} \Delta(t). \quad (1.8)$$

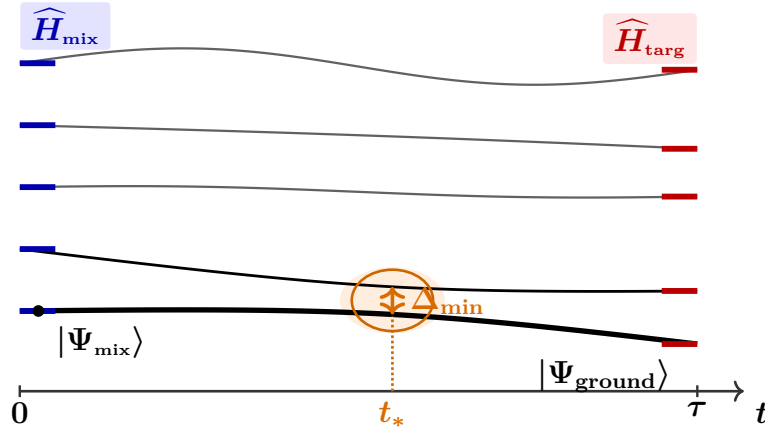


Figure 1.3.: Schematic instantaneous spectrum along a quantum annealing path. Blue and red bars mark the endpoint spectra of $\hat{H}_{\text{mix}}$ and $\hat{H}_{\text{targ}}$. Ideally, the state follows the thick ground branch between their ground states. The minimum gap is marked at an avoided crossing; the higher levels are illustrative.

Figure 1.3 illustrates this spectral picture. The marked avoided crossing is a common source of a small gap, although the detailed spectrum is problem dependent and exact crossings may occur between different symmetry sectors. If the two lowest

levels dominate the dynamics, the vicinity of $\Delta_{\min}$ is where nonadiabatic excitation is most likely [65].

The backbone of conventional QA is the adiabatic theorem. In its simplest form, it states that a system initially prepared in a nondegenerate eigenstate of a slowly varying Hamiltonian remains close to the corresponding instantaneous eigenstate throughout the dynamics. Following the pedagogical development in Ref. [66], the calculation below has two purposes: to derive the additional Berry's phase and to motivate why transitions between different instantaneous eigenstates can be neglected when the Hamiltonian changes sufficiently slowly. The same ingredients are standard textbook material [67].

At every fixed time t , the Hamiltonian can be diagonalized. Denote an orthonormal basis of its instantaneous eigenstates by $\{|n(t)\rangle\}$, so that

$$\hat{H}(t)|n(t)\rangle = E_n(t)|n(t)\rangle, \quad \langle n(t)|m(t)\rangle = \delta_{nm}. \quad (1.9)$$

As $\hat{H}(t)$ changes, the eigenstates themselves form a moving basis. The instantaneous basis then permits the exact state to be written as

$$|\Psi(t)\rangle = \sum_n a_n(t) e^{-i\theta_n(t)} |n(t)\rangle, \quad \theta_n(t) = \frac{1}{\hbar} \int_0^t E_n(t') dt'. \quad (1.10)$$

The explicit exponential factors out the familiar dynamical phase from each basis component. The coefficients $a_n(t)$ retain everything else, including any redistribution of amplitude caused by the motion of the basis. Insert Eq. (1.10) into the time-dependent Schrödinger equation. Applying the product rule to $a_n(t)$, the phase, and $|n(t)\rangle$ gives

$$\begin{aligned} i\hbar|\dot{\Psi}(t)\rangle &= i\hbar \sum_n \left[\dot{a}_n|n\rangle + a_n|\dot{n}\rangle - i\dot{\theta}_n a_n|n\rangle \right] e^{-i\theta_n}, \\ \hat{H}(t)|\Psi(t)\rangle &= \sum_n a_n E_n |n\rangle e^{-i\theta_n}, \quad \dot{\theta}_n = \frac{E_n}{\hbar}. \end{aligned} \quad (1.11)$$

The three terms in the first line arise respectively from the derivative of $a_n(t)$, the moving eigenstate, and the exponential. For the last term, $i\hbar(-i\dot{\theta}_n) = \hbar\dot{\theta}_n = E_n$; it therefore reproduces, component by component, the action of $\hat{H}(t)$ shown in the second line. The energy terms cancel exactly, leaving

$$i\hbar \sum_n (\dot{a}_n|n\rangle + a_n|\dot{n}\rangle) e^{-i\theta_n} = 0. \quad (1.12)$$

Project Eq. (1.12) onto an arbitrary instantaneous eigenstate $\langle m(t)|$ and using orthonormality, we obtain

$$\sum_n [\dot{a}_n \delta_{mn} + a_n \langle m|\dot{n}\rangle] e^{-i\theta_n} = 0. \quad (1.13)$$

Separating the $n = m$ term from the remaining terms then gives

$$\begin{aligned}\dot{a}_m(t) &= - \sum_n a_n(t) \langle m(t) | \dot{n}(t) \rangle e^{i[\theta_m(t) - \theta_n(t)]} \\ &= -a_m(t) \langle m(t) | \dot{m}(t) \rangle - \sum_{n \neq m} a_n(t) \langle m(t) | \dot{n}(t) \rangle e^{i[\theta_m(t) - \theta_n(t)]}.\end{aligned}\quad (1.14)$$

Equation (1.14) remains exact. Its first term on the RHS involves the same level m while each term in the sum connects level m to a different instantaneous level n and is therefore a possible source of transitions.

We now take the time derivative of the instantaneous eigenvalue equation in Eq. (1.9). Projection of the result onto $\langle m(t) |$ yields

$$\begin{aligned}\dot{\hat{H}}|n\rangle + \hat{H}|\dot{n}\rangle &= \dot{E}_n|n\rangle + E_n|\dot{n}\rangle, \\ \langle m|\dot{\hat{H}}|n\rangle + E_m\langle m|\dot{n}\rangle &= \dot{E}_n\delta_{mn} + E_n\langle m|\dot{n}\rangle.\end{aligned}\quad (1.15)$$

For $n \neq m$, the Kronecker delta vanishes. Collecting the two terms proportional to $\langle m|\dot{n}\rangle$ gives $(E_n - E_m)\langle m|\dot{n}\rangle = \langle m|\dot{\hat{H}}|n\rangle$, and hence

$$\langle m(t) | \dot{n}(t) \rangle = \frac{\langle m(t) | \dot{\hat{H}}(t) | n(t) \rangle}{E_n(t) - E_m(t)}, \quad n \neq m. \quad (1.16)$$

This is the key exact equation. It makes the physics structure visible: rapid changes of the Hamiltonian increase coupling, while a large energy separation suppresses it. Substituting it into the second line of Eq. (1.14) gives the exact coefficient equation

$$\dot{a}_m(t) = -a_m(t) \langle m(t) | \dot{m}(t) \rangle - \sum_{n \neq m} a_n(t) \frac{\langle m(t) | \dot{\hat{H}}(t) | n(t) \rangle}{E_n(t) - E_m(t)} e^{i[\theta_m(t) - \theta_n(t)]}. \quad (1.17)$$

All terms up to Eq. (1.17) have followed from the Schrödinger equation without approximation. The first term evolves the coefficient already associated with level m ; the sum can feed it from other levels. The adiabatic approximation is made precisely by neglecting this off-diagonal sum. The claim is that the sum becomes small relative to the diagonal evolution when the relevant matrix elements of $\dot{\hat{H}}$ are small. That is, when the process is slow on the scale set by the energy separations.

Once the off-diagonal sum is omitted, the approximation and corresponding solution read:

$$\begin{aligned}\dot{a}_m(t) &\simeq -a_m(t) \langle m(t) | \dot{m}(t) \rangle, \\ \frac{\dot{a}_m(t)}{a_m(t)} &\simeq -\langle m(t) | \dot{m}(t) \rangle, \\ a_m(t) &= a_m(0) e^{i\gamma_m(t)}, \quad \gamma_m(t) = i \int_0^t \langle m(t') | \dot{m}(t') \rangle dt'.\end{aligned}\quad (1.18)$$

The phase $\gamma_m(t)$ is the geometric, or Berry's phase. The state then becomes

$$|\Psi(t)\rangle \simeq \sum_m a_m(0) e^{-i\theta_m(t)} e^{i\gamma_m(t)} |m(t)\rangle. \quad (1.19)$$

Within this approximation, each occupied component therefore remains on its instantaneous eigenstate branch and accumulates two distinct phases: the dynamical phase $\theta_m(t)$ and Berry's phase $\gamma_m(t)$ arising from the motion of the eigenbasis.

The coefficient equation identifies the transition terms, but it does not yet justify discarding them. A local perturbative argument in Ref. [66] addresses this question. Choose an arbitrary time t_0 and freeze the instantaneous basis at that time, using $\{|r(t_0)\rangle\}$ as a fixed basis throughout the next interval of length ε . Assume that the system has followed level n up to t_0 , so that its fixed-basis coefficients satisfy $a_r(t_0) = \delta_{rn}$. The question we want to ask is: what is $a_m(t_0 + \varepsilon)$ for $m \neq n$?

Over this short interval, the change of the Hamiltonian relative to $\hat{H}(t_0)$ is treated by first-order time-dependent perturbation theory [67]:

$$a_m^{(1)}(t_0 + \varepsilon) = -\frac{i}{\hbar} \int_{t_0}^{t_0 + \varepsilon} H_{mn}(t) e^{i\omega_{mn}(t-t_0)} dt. \quad (1.20)$$

Here $H_{mn}(t) = \langle m(t_0) | \hat{H}(t) | n(t_0) \rangle$ and $\hbar\omega_{mn} = E_m(t_0) - E_n(t_0)$.

We shift the time origin to $t_0 = 0$ and integrate by parts:

$$\begin{aligned} a_m^{(1)}(\varepsilon) = & -\frac{1}{\hbar\omega_{mn}} [H_{mn}(t) e^{i\omega_{mn}t}]_0^\varepsilon \\ & + \frac{1}{\hbar\omega_{mn}} \int_0^\varepsilon \dot{H}_{mn}(t) e^{i\omega_{mn}t} dt. \end{aligned} \quad (1.21)$$

Smoothness of the Hamiltonian over the short interval is assumed, which permits the following expansions:

$$H_{mn}(t) = H_{mn}(0) + \dot{H}_{mn}(0)t + O(t^2), \quad e^{i\omega_{mn}t} = 1 + i\omega_{mn}t + O(t^2). \quad (1.22)$$

Substituting these expansions and the fact that $H_{mn}(0) = 0$ into Eq. (1.21) and retaining terms through first order in ε gives

$$a_m^{(1)}(\varepsilon) = -\frac{\dot{H}_{mn}(0)}{\hbar\omega_{mn}} \varepsilon + \frac{\dot{H}_{mn}(0)}{\hbar\omega_{mn}} \varepsilon + O(\varepsilon^2) = O(\varepsilon^2). \quad (1.23)$$

The boundary contribution and the integral contribution cancel at order ε . Thus, within this local first-order calculation, the transition amplitude has no term linear in the interval length and begins at $O(\varepsilon^2)$. This provides the heuristic motivation for adiabatic following: over a sufficiently short segment of a slowly changing path, the unwanted transition is of higher order than the ordinary change being followed. This local cancellation is an indication of the mechanism, not a complete proof of the adiabatic theorem.

For QA, the dependence of the required annealing time on the minimum gap is often summarized schematically as [68]

$$\tau = O\left(\frac{1}{\Delta_{\min}^2}\right). \quad (1.24)$$

Here $\hbar$ and problem-dependent energy scales are understood to be absorbed into the prefactor hidden by the $O(\cdot)$ notation. This inverse-gap-squared relation conveys the practical consequence of a small minimum gap without adding another derivation to the partial argument above.

The same ground-state construction becomes a model of computation when the Hamiltonian path is designed so that the final ground state encodes the output of a computational problem. More precisely, *adiabatic quantum computation* (AQC) specifies an initial Hamiltonian with an easily prepared ground state, a final Hamiltonian whose ground-state subspace encodes the output, and an interpolation slow enough for the state to follow the corresponding instantaneous ground-state subspace; the output is then obtained by measurement [65, 68–70]. In the adiabatic setting used here, QA is the optimization-oriented specialization of this construction: the target Hamiltonian represents a classical cost function, while a noncommuting driver induces quantum fluctuations and is gradually switched off during the evolution [63–65]. Programmable superconducting QA has progressed from an eight-flux-qubit array to measurements of quantum-critical spin-glass dynamics on a 5,000-qubit processor, while digitized AQC has been demonstrated with up to nine superconducting qubits [71–73]. At the level of computational models, AQC with sufficiently general local Hamiltonians is polynomially equivalent to the circuit model of quantum computation; this equivalence does not imply that every chosen Hamiltonian path can be traversed in polynomial time [65, 74].

Continuous adiabatic evolution can also be approximated by a gate sequence. Section 1.2.4 develops this digitization and its relation to QAOA. Before turning to that general construction, the following Grover example shows concretely that the cost of an adiabatic computation depends not only on the Hamiltonian path and its gap, but also on the schedule used to traverse it.

1.2.2. Grover search: the schedule as a resource

Unstructured search provides the clearest example of schedule shape as an algorithmic resource [47]. Let $\mathcal{X} = \{0, 1, \dots, N_s - 1\}$ be a search space of $N_s = 2^N$ labels, and suppose that exactly one label $m \in \mathcal{X}$ is marked. An *oracle* is a black-box operation that answers a prescribed question about a queried input but reveals no other information about the solution. Here the oracle is the Boolean function $f : \mathcal{X} \rightarrow \{0, 1\}$ satisfying $f(m) = 1$ and $f(x) = 0$ for every $x \neq m$: it reports whether the queried label is the marked one without directly revealing m . The task is to

determine m by querying f , without any structure that indicates where the marked label lies. If m is uniformly distributed and labels are queried without replacement, a classical search uses $(N_s + 1)/2$ queries on average and N_s queries in the worst case; its query complexity is therefore $\Theta(N_s)$.

In Grover's quantum algorithm, each label x is represented by a computational-basis state $|x\rangle$, and the initial state is the uniform superposition. The phase oracle O_f and diffusion operator D are defined by [75]

$$\begin{aligned} |\psi_0\rangle &= \frac{1}{\sqrt{N_s}} \sum_{x=0}^{N_s-1} |x\rangle, & O_f|x\rangle &= (-1)^{f(x)}|x\rangle, \\ D &= 2|\psi_0\rangle\langle\psi_0| - \mathbb{1}, & G &= DO_f. \end{aligned}$$

Thus O_f reverses only the phase of $|m\rangle$, while D reflects all amplitudes about their average. Successive applications of the Grover iteration G rotate the state towards $|m\rangle$, amplifying its amplitude at the expense of the unmarked states. After approximately $k \simeq (\pi/4)\sqrt{N_s}$ iterations, measurement returns m with probability close to one. The resulting $\Theta(\sqrt{N_s})$ query complexity is optimal for unstructured quantum search [76].

The adiabatic Hamiltonians can be derived directly from the eigenspaces of the two reflections used by the circuit algorithm. The phase oracle has eigenvalue -1 on the marked state and $+1$ on every unmarked state. Consequently, $(\mathbb{1} + O_f)/2 = \mathbb{1} - |m\rangle\langle m|$ is the projector onto the unmarked subspace. Assigning energy J to that subspace and zero energy to $|m\rangle$ gives the target Hamiltonian [77]

$$\hat{H}_{\text{targ}} = \frac{J}{2}(\mathbb{1} + O_f) = J(\mathbb{1} - |m\rangle\langle m|). \quad (1.25)$$

The state $|m\rangle$ has energy zero, while every other computational-basis state has energy J . Because all dependence on the unknown label m is contained in this operator, $\hat{H}_{\text{targ}}$ is also called the oracle Hamiltonian.

Similarly, the diffusion operator has eigenvalue $+1$ on $|\psi_0\rangle$ and -1 on its orthogonal complement. Therefore, $(\mathbb{1} - D)/2 = \mathbb{1} - |\psi_0\rangle\langle\psi_0|$ projects onto that complement. Assigning it energy J gives the driver Hamiltonian

$$\hat{H}_{\text{mix}} = \frac{J}{2}(\mathbb{1} - D) = J(\mathbb{1} - |\psi_0\rangle\langle\psi_0|). \quad (1.26)$$

Its ground state $|\psi_0\rangle$ is the uniform superposition and is easy to prepare. Evolution for a time $\pi\hbar/J$ under $\hat{H}_{\text{targ}}$ implements $-O_f$, whereas evolution for the same time under $\hat{H}_{\text{mix}}$ implements D ; the minus sign is only a global phase. The circuit algorithm alternates these two reflections. The adiabatic algorithm instead continuously interpolates between their corresponding Hamiltonians by substituting Eqs. (1.25) and (1.26) into Eq. (1.6).

Although its Hilbert space contains N_s states, the evolution is restricted to the two-dimensional subspace spanned by $|m\rangle$ and the uniform superposition of the unmarked states,

$$|m_\perp\rangle = \frac{1}{\sqrt{N_s-1}} \sum_{x \neq m} |x\rangle. \quad (1.27)$$

The two lowest energies can therefore be obtained exactly:

$$E_{0,1}(s) = \frac{J}{2} \left[1 \mp \sqrt{1 - 4s(1-s) \left(1 - \frac{1}{N_s}\right)} \right]. \quad (1.28)$$

Their gap is [47, 77]

$$\begin{aligned} \Delta(s) &= J \sqrt{1 - 4s(1-s) \left(1 - \frac{1}{N_s}\right)}, \\ \Delta^2(s) &= \Delta_{\min}^2 \left[1 + 4(N_s - 1) \left(s - \frac{1}{2}\right)^2 \right], \quad \Delta_{\min} = \frac{J}{\sqrt{N_s}}. \end{aligned} \quad (1.29)$$

The minimum occurs at $s = 1/2$, where the ground and first excited states form a narrow avoided crossing, as shown in Fig. 1.4(a).

The standard local-adiabatic prescription introduces a small parameter ε and chooses the schedule according to [65, 77]

$$\frac{dt}{ds} \geq \frac{\hbar \|\partial_s \hat{H}(s)\|}{\varepsilon \Delta^2(s)}. \quad (1.30)$$

For a linear schedule, $dt/ds = \tau$ is constant and must be large enough at the minimum gap. Since $\|\partial_s \hat{H}\| = J\sqrt{1 - N_s^{-1}}$, this gives $\tau = O(N_s)$: linear-schedule QA has the same asymptotic scaling as classical search.

Roland and Cerf instead chose the largest rate specified by Eq. (1.30) at every point of the evolution [77]. Replacing the inequality by an equality gives

$$\frac{ds}{dt} = \frac{\varepsilon \Delta^2(s)}{\hbar \|\partial_s \hat{H}(s)\|} = \frac{\varepsilon J}{\hbar \sqrt{1 - N_s^{-1}}} \left[1 - 4 \frac{N_s - 1}{N_s} s(1-s) \right]. \quad (1.31)$$

With $s(0) = 0$ and $s(\tau) = 1$, its solution is

$$s(t) = \frac{1}{2} + \frac{\tan[C(2t/\tau - 1)]}{2 \tan C}, \quad C = \arctan \sqrt{N_s - 1}, \quad (1.32)$$

and the total time is

$$\tau = \frac{\hbar}{J\varepsilon} \sqrt{N_s} \arctan \sqrt{N_s - 1} \xrightarrow{N_s \gg 1} \frac{\pi \hbar}{2J\varepsilon} \sqrt{N_s}. \quad (1.33)$$

The resulting schedule, plotted in Fig. 1.4(b), moves quickly where the gap is large and slows down around $s = 1/2$. It recovers the $O(\sqrt{N_s})$ scaling of Grover search.

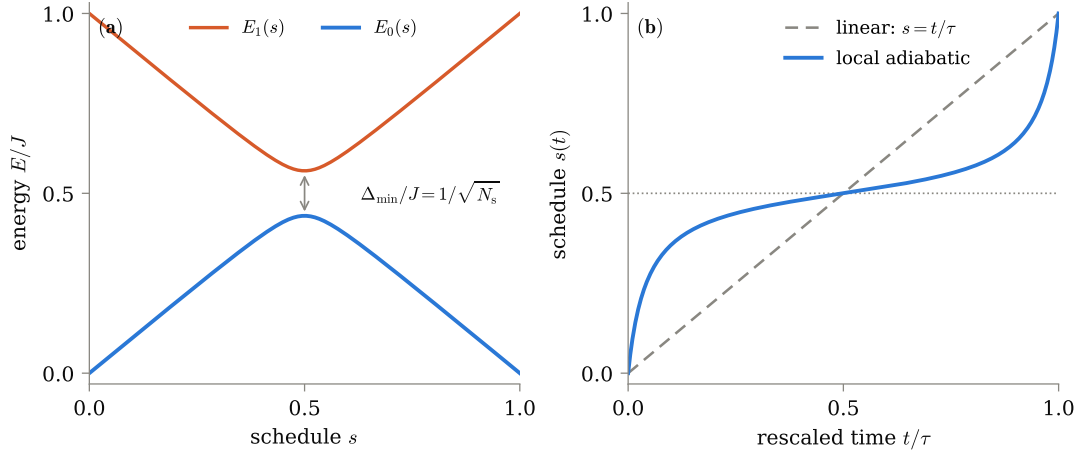


Figure 1.4.: Adiabatic Grover search for $N_s = 64$ states. **(a)** The two lowest energies in units of J . Their avoided crossing at $s = 1/2$ has minimum gap $\Delta_{\min}/J = 1/\sqrt{N_s}$. **(b)** The Roland–Cerf schedule obtained from the local adiabatic condition (solid blue), compared with the linear schedule (dashed gray). The local schedule moves rapidly where the gap is large and slows around the minimum gap. Replotted from the analytic solution of Ref. [77], following the presentation of Ref. [47].

The Hamiltonian path and its endpoints are identical for the linear and Roland–Cerf protocols; only the rate at which the path is traversed is changed. The schedule is therefore an algorithmic resource. This example also shows the main difficulty: constructing the local schedule requires the gap profile and the location of its minimum. The same slowing near a small gap appears as a plateau in optimized smooth protocols, making the Grover schedule a useful reference for the numerical control strategies discussed later in this Thesis [46, 69].

The uniform transverse-field Ising chain supplies a complementary many-body example. Section 1.2.5 uses its standard exact solution to compare continuous annealing, dQA, and QAOA; details of the Jordan–Wigner and Bogoliubov treatment can be found in Ref. [78].

1.2.3. Shortcuts to adiabaticity

Quantum annealing seeks the ground state of $\hat{H}_{\text{targ}}$ by following the instantaneous ground state of an interpolating Hamiltonian $\hat{H}(s)$. The adiabatic route becomes costly when the spectrum contains a small minimum gap: suppressing transitions then requires a long annealing time, which can grow exponentially when the gap closes exponentially with system size [65, 79, 80]. Tailoring the rate $\dot{s}$ can reduce this cost, as the Grover example above illustrates, but doing so generally requires prior knowledge of the gap profile. Shortcuts to adiabaticity (STA) address the same bottleneck from a different direction: they are finite-time protocols designed to reproduce the final outcome of a reference adiabatic evolution without requiring

the actual dynamics to remain adiabatic at intermediate times [81]. Here the relevant outcome is ground-state preparation.

A paradigmatic STA construction is *counterdiabatic*, or transitionless, driving [82, 83]. Let $\hat{H}(s)|n(s)\rangle = E_n(s)|n(s)\rangle$. One supplements this reference Hamiltonian with an auxiliary term,

$$\hat{H}_{\text{tot}}(t) = \hat{H}(s(t)) + \hat{H}_{\text{CD}}(t), \quad \hat{H}_{\text{CD}}(t) = \dot{s} \mathcal{A}_s, \quad (1.34)$$

where $\mathcal{A}_s$ is the *adiabatic gauge potential* (AGP). In a gauge with vanishing diagonal matrix elements, its exact spectral form is

$$\mathcal{A}_s = i\hbar \sum_{m \neq n} \frac{\langle m(s) | \partial_s \hat{H}(s) | n(s) \rangle}{E_n(s) - E_m(s)} |m(s)\rangle \langle n(s)|. \quad (1.35)$$

Equations (1.34) and (1.35) are the transitionless-driving construction of Ref. [83], written for the parameter $s(t)$. The added term cancels transitions between the instantaneous eigenstates, so an initial eigenstate follows the corresponding eigenstate of $\hat{H}(s)$ exactly, up to phase factors. Formally this works for an arbitrarily fast reference schedule, but faster driving demands stronger controls. In an interacting many-body system the exact AGP also requires detailed spectral information and generally contains nonlocal operators [84]. Practical counterdiabatic schemes therefore approximate it within a tractable, often local, variational ansatz [85].

In this Thesis, STA provides an interpretive framework; it is not used to construct the protocols. Neither the MaxCut nor the frustrated-ring protocol adds $\hat{H}_{\text{CD}}$ or reconstructs $\mathcal{A}_s$; the available generators remain $\hat{H}_{\text{targ}}$ and $\hat{H}_{\text{mix}}$. Nevertheless, their population-resolved dynamics display a common pattern: optimization drives population out of the instantaneous ground state before the spectral bottleneck and later returns it coherently to the final ground state [46, 86]. For the digitized MaxCut protocols, Baker–Campbell–Hausdorff commutators provide a qualitative connection to operator directions appearing in approximate AGPs [87], but do not establish an implemented counterdiabatic Hamiltonian. The resulting dynamics are therefore best described as *STA-like*: they realize a finite-time nonadiabatic shortcut to the target state without transitionless tracking of the reference instantaneous ground state.

1.2.4. Variational quantum algorithms, dQA, and QAOA

Quantum annealing describes a continuous evolution controlled by a schedule $s(t)$, whereas gate-based algorithms are expressed as discrete sequences of unitaries. Digitization connects these descriptions. It converts an annealing schedule into an alternating circuit and, conversely, identifies the subset of variational angles that retain a schedule interpretation. This connection underlies the protocol families studied throughout the Thesis.

Digitized quantum annealing

The passage from continuous QA to a gate sequence occurs in two stages: first discretizing the time-dependent evolution, and then applying a Trotter splitting to each short-time propagator.

Digitization. The first step replaces the continuous schedule by a stepwise one [47, 49, 73]. Divide $[0, \tau]$ into P intervals of width $\Delta t = \tau/P$ and sample the schedule at their midpoints, $t_p = (p - \frac{1}{2})\Delta t$ and $s_p = s(t_p)$. Approximating the Hamiltonian as constant within each interval gives, setting $\hbar = 1$,

$$U(\tau, 0) \simeq U_{\text{step}} = e^{-i\Delta t \hat{H}(s_P)} \dots e^{-i\Delta t \hat{H}(s_1)}. \quad (1.36)$$

Equation (1.36) is a time discretization only: each factor still contains the full interpolating Hamiltonian $\hat{H}(s_p) = (1 - s_p)\hat{H}_{\text{mix}} + s_p\hat{H}_{\text{targ}}$. The stepwise propagator converges to the time-ordered continuous evolution as the largest time step tends to zero.

Trotterization. Because $\hat{H}_{\text{mix}}$ and $\hat{H}_{\text{targ}}$ generally do not commute, the exponential of their sum is not the product of their exponentials. For each digitized interval, the first-order Lie–Trotter formula [88, 89] gives

$$e^{-i\Delta t \hat{H}(s_p)} \simeq e^{-i\theta_p^x \hat{H}_{\text{mix}}} e^{-i\theta_p^z \hat{H}_{\text{targ}}} \equiv U_p(\theta_p^x, \theta_p^z), \quad (1.37)$$

with

$$\theta_p^z = \Delta t s_p, \quad \theta_p^x = \Delta t (1 - s_p), \quad p = 1, \dots, P. \quad (1.38)$$

The local error of this nonsymmetric splitting is $\mathcal{O}(\Delta t^2)$ and is controlled by commutators of the two Hamiltonians. Symmetric or higher-order product formulas can reduce this error, but the first-order form already produces the alternating circuit used throughout this Thesis.

The resulting *digitized quantum annealing* (dQA) protocol [49, 90] is

$$U_{\text{dQA}} = U_P \dots U_1, \quad |\Psi_{\text{P}}^{\text{dQA}}\rangle = U_{\text{dQA}} |\Psi_{\text{mix}}\rangle, \quad (1.39)$$

where $|\Psi_{\text{mix}}\rangle$ is the ground state of the mixing Hamiltonian. It contains P alternating pairs, hence $2P$ unitaries. For the uniform digitization above, $\theta_p^x + \theta_p^z = \Delta t$ in every layer and $\sum_p (\theta_p^x + \theta_p^z) = \tau$. Each step of the sampled interpolation is thus replaced by consecutive evolution under $\hat{H}_{\text{targ}}$ and $\hat{H}_{\text{mix}}$, yielding a piecewise-constant, bang–bang realization of the annealing path.

When both angles in each layer are nonnegative, this correspondence can be inverted:

$$s_p = \frac{\theta_p^z}{\theta_p^x + \theta_p^z}, \quad \Delta t_p = \theta_p^x + \theta_p^z, \quad \tau = \sum_{p=1}^P (\theta_p^x + \theta_p^z), \quad (1.40)$$

provided $\theta_p^x + \theta_p^z > 0$. Thus, a nonnegative angle sequence defines both an effective schedule value s_p and a layer duration Δt_p . The durations need not be uniform, and the reconstructed schedule need not be monotonic. By contrast, generic variational angles may be negative or exploit periodicities of the unitaries, in which case they do not admit this direct annealing interpretation. The distinction is important in Chapters 2 and 3, where regular digital solutions are reconstructed as continuous, analog-compatible protocols.

QAOA and the variational viewpoint

Farhi, Goldstone, and Gutmann originally defined QAOA as an alternating sequence of cost and mixing unitaries with independently chosen angles [25]. This ansatz has the same alternating circuit structure as dQA, but its angles are independent variational parameters rather than being fixed by a sampled annealing schedule. At depth P , the prepared state is

$$|\Psi_P(\boldsymbol{\theta})\rangle = e^{-i\theta_P^z \hat{H}_{\text{mix}}} e^{-i\theta_P^x \hat{H}_{\text{targ}}} \dots e^{-i\theta_1^x \hat{H}_{\text{mix}}} e^{-i\theta_1^z \hat{H}_{\text{targ}}} |\Psi_{\text{mix}}\rangle, \quad (1.41)$$

The quantum processor prepares this state and estimates

$$E_P(\boldsymbol{\theta}) = \langle \Psi_P(\boldsymbol{\theta}) | \hat{H}_{\text{targ}} | \Psi_P(\boldsymbol{\theta}) \rangle, \quad (1.42)$$

while a classical optimizer updates the $2P$ components of $\boldsymbol{\theta}$. This quantum–classical loop is the defining structure of a variational quantum algorithm (VQA) [91].

For $P = 1$, the original analysis guarantees an approximation ratio of at least 0.6924 for MaxCut on 3-regular graphs [25].

This variational freedom allows QAOA to explore dynamics that need not admit a direct interpretation as an annealing schedule. For suitable choices of the two generators, alternating-Hamiltonian circuits can even be computationally universal [92].

The target Hamiltonian $\hat{H}_{\text{targ}}$ is diagonal in the computational basis and encodes the classical objective. In the present minimization convention, the standard unconstrained mixer is $\hat{H}_{\text{mix}} = -\sum_j \hat{\sigma}_j^x$, whose ground state is $|\Psi_{\text{mix}}\rangle = |+\rangle^{\otimes N}$ [25]. The mixer is not unique: an XY exchange mixer built from $\hat{\sigma}_j^x \hat{\sigma}_k^x + \hat{\sigma}_j^y \hat{\sigma}_k^y$ preserves Hamming weight, while taking $\hat{H}_{\text{mix}}$ as the adjacency matrix of a connected graph of feasible configurations restricts the induced quantum walk to the feasible subspace [93]. When an equal superposition $|\Psi_F\rangle$ of feasible configurations can be prepared efficiently, the Grover-projector choice $\hat{H}_{\text{mix}}^G = -|\Psi_F\rangle\langle\Psi_F|$ provides another constraint-preserving alternative [94].

Locating good parameters in the nonconvex, $2P$ -dimensional landscape is a central practical difficulty. For MaxCut, Zhou *et al.* observed regular patterns in high-quality angles and proposed two structured strategies: INTERP, which seeds depth

$P + 1$ from an optimized depth- P solution, and FOURIER, which expands the angle sequences in a truncated Fourier basis [95]. In their benchmarks these strategies found quasi-optimal parameters with polynomial effort in P , whereas comparable random-start searches required exponentially many runs; this empirical scaling is not a general guarantee.

1.2.5. The uniform Ising chain: continuous annealing, dQA, and QAOA

The uniform transverse-field Ising chain is an exactly solvable many-body benchmark for comparing continuous and digital approaches to state preparation. We consider an even number N of spins with periodic boundary conditions, $\hat{\sigma}_{N+1}^z = \hat{\sigma}_1^z$, and the target and transverse-field Hamiltonians [49]

$$\hat{H}_z = \sum_{j=1}^N (\hat{\sigma}_j^z \hat{\sigma}_{j+1}^z - 1), \quad \hat{H}_x = - \sum_{j=1}^N \hat{\sigma}_j^x. \quad (1.43)$$

The ground state of $\hat{H}_x$ is the uniform product state $|+\rangle^{\otimes N}$. This simple model provides a common setting in which continuous QA, dQA, and QAOA can be compared analytically.

Continuous-time annealing. For the uniform chain, the interpolation in Eq. (1.6) becomes $\hat{H}(s) = s\hat{H}_z + (1-s)\hat{H}_x$. A schedule satisfying $s(0) = 0$ and $s(\tau) = 1$ therefore changes the Hamiltonian continuously from the transverse field to the Ising interaction. The path crosses its quantum critical point at $s = 1/2$, where the instantaneous gap is smallest and the evolution is most susceptible to excitations.

For a linear schedule, the critical point is crossed at a constant rate and the residual energy follows the Kibble–Zurek scaling $\epsilon^{\text{res}}(\tau) \sim \tau^{-1/2}$. Mbeng *et al.* also compared this baseline with schedules designed to slow the evolution near the known critical point. Over the time window studied, a Roland–Cerf-inspired schedule and an optimized power-law schedule produced approximate decays $\tau^{-0.75}$ and $\tau^{-0.8}$, respectively. These fitted finite-time scalings improve on the linear protocol, but their construction uses prior knowledge of the critical region. This comparison is presented in Sec. IV.C, “Comparison with other QA strategies,” of Ref. [49].

Digitized quantum annealing. As reviewed in Section 1.2.4, digitization proceeds in two stages. First, the time axis is divided into P intervals, which may have unequal durations Δt_m , and the schedule is held at a constant value s_m on the m th interval. The resulting step-QA evolution still uses the full Hamiltonian $\hat{H}(s_m) = s_m\hat{H}_z + (1-s_m)\hat{H}_x$ during each interval. This construction is introduced in Sec. II.B, “Continuous and digitized quantum annealing,” of Ref. [49].

Second, a first-order Trotter splitting replaces this evolution by successive ap-

plications of its two Hamiltonian components. According to Eq. (1.38), the system evolves under $\hat{H}_z$ for a duration $\theta_m^z = s_m \Delta t_m$ and under $\hat{H}_x$ for a duration $\theta_m^x = (1 - s_m) \Delta t_m$. Repeating these two operations for all P intervals gives the dQA protocol. In the fixed-step comparison of Sec. IV.C of Ref. [49], linear dQA retains the $\tau^{-1/2}$ Kibble–Zurek exponent of continuous QA; digitization produces a constant offset in the residual energy without changing that exponent.

The construction of Fig. 1 in Ref. [49] is explained in its Sec. IV.A, “Optimal schedules for $2P < N$ and digitized-QA.” Accordingly, Fig. 1.5 is constructed in the reverse direction, starting from eight pairs of nonnegative durations (θ_m^z, θ_m^x) . Each pair appears in the black curve as evolution at $s = 1$ for a duration θ_m^z , followed by evolution at $s = 0$ for a duration θ_m^x . Thus, the black curve encodes successive applications of $\hat{H}_z$ and $\hat{H}_x$ rather than a pointwise approximation to a continuous schedule. Applying the inverse dictionary of Eq. (1.40) gives the interval duration $\Delta t_m = \theta_m^z + \theta_m^x$ and the corresponding schedule value $s_m = \theta_m^z / (\theta_m^z + \theta_m^x)$. The cumulative intervals are nonuniform, sum to $\tau = 9.76$, and produce the blue step-QA curve. A smooth interpolation associated with these step values then gives the red continuous-time schedule.

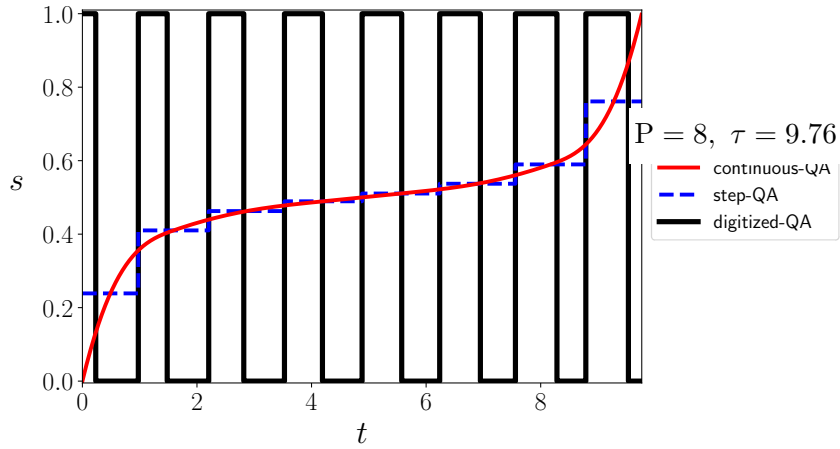


Figure 1.5.: Construction of the continuous representation from a digital protocol at $P = 8$. The target and transverse-field durations define the black digitized-QA protocol. Their sums and ratios determine the nonuniform intervals and values of the blue step-QA schedule, with total duration $\tau = 9.76$; the red line is the associated smooth interpolation. Adapted from Fig. 1 of Ref. [49]; the construction is described in Sec. IV.A of that reference.

QAOA. QAOA keeps the alternating circuit of dQA but releases the schedule constraint, treating its $2P$ angles as independent variational parameters, as reviewed in Section 1.2.4. For an energy E_P , define the spectral-width-normalized residual energy $\epsilon_P^{\text{res}} = (E_P - E_{\min}) / (E_{\max} - E_{\min})$. The uniform even ring obeys the tight

depth-dependent bound [49]

$$\epsilon_P^{\text{res}} \geq \begin{cases} \frac{1}{2P+2}, & 2P < N, \\ 0, & 2P \geq N. \end{cases} \quad (1.44)$$

When $2P < N$, locality restricts the circuit around any measured bond to a reduced chain of $2P + 2$ spins. An antiperiodic boundary condition on this reduced chain supplies the nonzero variational bound, while the Jordan–Wigner solution shows that it is saturated. Consequently, for $2P < N$ the optimum is $(2P + 2)^{-1}$ and is attained by a finite set of degenerate minima. Exact preparation first becomes possible at $2P = N$; for $2P > N$, the exact solutions form a continuous manifold [49, 90].

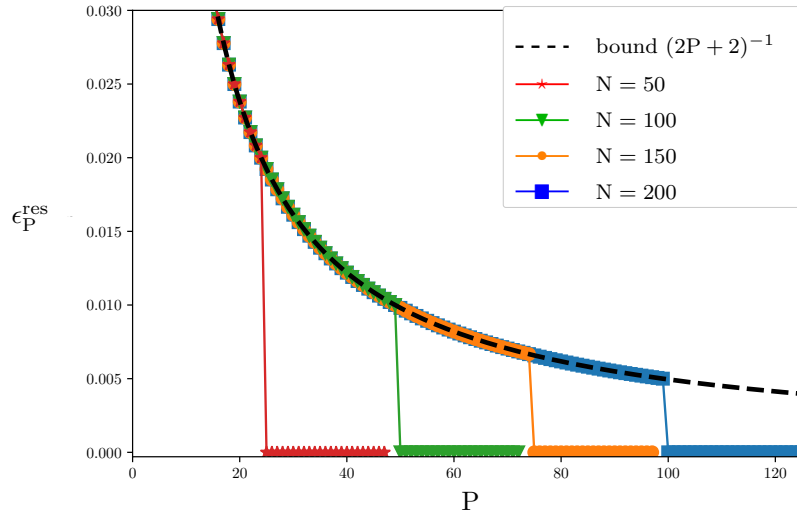


Figure 1.6.: Optimal QAOA residual energy for uniform even Ising rings as a function of circuit depth P . For $2P < N$, the numerical optima for different sizes follow the theoretical value $(2P + 2)^{-1}$ (black dashed line). At $2P \geq N$, the residual energy reaches zero and exact preparation becomes possible. Adapted from Fig. 4 of Ref. [49].

This solvable ring is the starting point for the developments in the following Chapters. Chapter 2 extends smooth schedule optimization to hard weighted MaxCut instances; Chapter 3 asks how controllability and structured digital protocols change when the ring is frustrated; and Chapter 4 studies how a parity encoding changes the spectrum and the digital resource threshold.

1.3. Quantum optimal control

Quantum optimal control (QOC) grew from applying classical optimal-control ideas to the increasingly precise coherent manipulation of quantum systems. Its

modern formulation took shape in molecular physics during the 1980s, when Tannor and Rice, and later Peirce, Dahleh, and Rabitz, cast the design of time-dependent fields that steer quantum dynamics toward a prescribed target as an optimization problem [96, 97]. During the 1990s, monotonically convergent iterative schemes, including Krotov-based and Zhu–Rabitz methods, were developed for numerical pulse design [98, 99]. Closed-loop learning, which connects pulse shaping directly to measured outcomes, was proposed [100] and subsequently demonstrated experimentally [101]. Applications to quantum information included direct optimization of unitary transformations [102], while GRAPE enabled efficient gradient-based optimization of NMR control pulses [103]. From the late 2000s onward, QOC was extended to quantum-speed-limit studies and interacting many-body dynamics, while CRAB provided a reduced-parameter search adapted to costly simulations [104–106]. Contemporary QOC is increasingly hardware-aware, incorporating amplitude and bandwidth bounds, leakage suppression, robustness requirements, and open-system effects [107–109].

1.3.1. Theory of quantum optimal control

Control problem. Consider a finite-dimensional closed system and set $\hbar = 1$ throughout this Section. Its state obeys the time-dependent Schrödinger equation, with the Hamiltonian written as a fixed *drift* plus externally driven *control* terms [109, 110],

$$i \frac{\partial}{\partial t} |\psi(t)\rangle = \hat{H}(t) |\psi(t)\rangle, \quad \hat{H}(t) = \hat{H}_0 + \sum_{k=1}^K u_k(t) \hat{H}_k. \quad (1.45)$$

The drift $\hat{H}_0$ generates the internal dynamics that remain present throughout the protocol. The pair $u_k(t)\hat{H}_k$ defines the k th *control channel*: $u_k(t)$ is an independently adjustable scalar field, and $\hat{H}_k$ is the operator through which it drives the system. The total number of available control channels is denoted by K . A control field may represent a shaped microwave or laser amplitude, a phase, or a tunable coupling. Quantum control asks how these fields can be shaped to produce a desired dynamical outcome. Once the operators, admissible field amplitudes and bandwidths, and total duration τ are specified, reachability and optimality become distinct questions: can the target be reached, and, if so, which admissible control reaches it most effectively?

Common control tasks include the following. In *state-to-state transfer*, one maximizes the final fidelity $F = |\langle \psi_{\text{targ}} | \psi(\tau) \rangle|^2$ [97], whereas *gate synthesis* compares the complete propagator $\hat{U}(\tau)$ with a target unitary $\hat{U}_{\text{targ}}$ [102]. In *observable optimization*, the objective is an expectation value such as $\langle \psi(\tau) | O | \psi(\tau) \rangle$ [111]. The problems studied in this Thesis combine the first and third viewpoints: the operational objective is to minimize the final expectation value of $\hat{H}_{\text{targ}}$, and attaining its ground energy certifies preparation in its ground-state manifold.

Controllability. Having specified the drift and control Hamiltonians, a fundamental question is whether they can generate an arbitrary target unitary U . This property, known as operator controllability, is characterized by the Lie algebra rank condition (LARC) [110, 112]. To state this condition, define the dynamical Lie algebra

$$\mathfrak{g} = \text{Lie}_{\mathbb{R}}\{iH_0, iH_1, \dots, iH_K\}, \quad (1.46)$$

The symbol $\text{Lie}_{\mathbb{R}}$ denotes the smallest real vector space containing the generators and closed under repeated commutators,

$$[A, B] = AB - BA. \quad (1.47)$$

Here a *Lie group* is a continuous group of transformations, and its *Lie algebra* is the vector space of its infinitesimal generators. $U(N)$ is the group of all $N \times N$ unitary matrices, while $SU(N) \subset U(N)$ is the subgroup of unitary matrices with determinant one. Their Lie algebras are $\mathfrak{u}(N)$, the skew-Hermitian matrices $X^\dagger = -X$, and $\mathfrak{su}(N)$, the traceless skew-Hermitian matrices.

If $\mathfrak{g} = \mathfrak{su}(N)$, the system can generate every unitary transformation in $SU(N)$, and is therefore fully operator controllable up to a global phase. If the global phase is also included, full controllability on $U(N)$ requires $\mathfrak{g} = \mathfrak{u}(N)$.

Variational formulation. Quantum optimal control theory (QOCT) turns the quantum control problem into a rigorous variational optimization problem. For the energy-minimization problems considered in this Thesis, a suitable cost functional combines a terminal cost with a quadratic penalty on the control fields [113, 114],

$$J[u] = \langle \psi(\tau) | O | \psi(\tau) \rangle + \lambda \int_0^\tau \sum_k |u_k(t)|^2 dt, \quad J[u] \longrightarrow \min, \quad (1.48)$$

Here $J[u]$ is the cost regarded as a functional of the controls, $|\psi(\tau)\rangle$ is the final state, O is the observable defining the terminal cost, and τ is the protocol duration. The function $u_k(t)$ denotes the k th control field, while the positive coefficient λ sets the weight of the integrated quadratic fluence penalty. Thus, both a large terminal cost and unnecessarily large control fluence increase $J[u]$.

Pontryagin's Maximum Principle in QOC. Within QOC, Pontryagin's Maximum Principle (PMP) provides the principal framework for determining optimal open-loop control laws [115]. Here, *open-loop* means that the time-dependent control fields $u_k(t)$ are determined before the quantum evolution begins and are not adjusted in response to measurements during that evolution.

The goal is to choose the control fields $u_k(t)$ that minimize the cost $J[u]$ in Eq. (1.48), subject to the requirement that the state obeys the Schrödinger equation. Because $|\psi(t)\rangle$ is fixed by the controls and the initial state, it cannot be varied independently. The constraint is handled by introducing a costate $|\chi(t)\rangle$ as a time-dependent Lagrange multiplier and adding the Schrödinger-equation residual

to $J[u]$ to form an augmented functional $\mathcal{L}$. $\mathcal{L}$ is not a new objective; it is used only to derive the state, costate, and gradient equations. In the notation used here, Eqs. (9)–(10) of Ref. [113] become

$$\mathcal{L}[\psi, \chi, u] = J[u] - 2 \operatorname{Im} \int_0^\tau \left\langle \chi(t) \left| \left(i \frac{d}{dt} - \hat{H}(t) \right) \right| \psi(t) \right\rangle dt. \quad (1.49)$$

The added integrand contains the residual of the Schrödinger equation, which vanishes for every admissible trajectory; hence $\mathcal{L}[\psi, \chi, u] = J[u]$ on such trajectories. Variations of $\mathcal{L}$ with respect to χ , ψ , and u_k yield, respectively, the forward Schrödinger equation (1.50), the costate equation (1.52) with terminal condition (1.51), and the cost gradient (1.53). The same construction is given in general state-space notation in Eq. (23) of Ref. [115].

The resulting adjoint-state, or weak-PMP, equations and their boundary conditions are presented below.

Let $u_k^{(n)}(t)$ denote the trial controls at iteration n , and let $\hat{H}^{(n)}(t)$ denote the Hamiltonian $\hat{H}(t)$ of Eq. (1.45) evaluated at these controls. They determine a forward trajectory through

$$i \frac{d}{dt} |\psi^{(n)}(t)\rangle = \hat{H}^{(n)}(t) |\psi^{(n)}(t)\rangle, \quad |\psi^{(n)}(0)\rangle = |\psi_0\rangle, \quad (1.50)$$

which is integrated from $t = 0$ to $t = \tau$. The state reached at the final time fixes the terminal value of the costate. For the terminal expectation value in Eq. (1.48), a convenient normalization is

$$|\chi^{(n)}(\tau)\rangle = O |\psi^{(n)}(\tau)\rangle. \quad (1.51)$$

In this Thesis, $O = \hat{H}_{\text{targ}}$, so the backward propagation starts from $\hat{H}_{\text{targ}} |\psi^{(n)}(\tau)\rangle$. For the frequently used state-transfer infidelity $1 - |\langle \psi_{\text{targ}} | \psi(\tau) \rangle|^2$, the terminal costate is instead proportional to $|\psi_{\text{targ}}\rangle$, with the proportionality factor given by the final overlap $\langle \psi_{\text{targ}} | \psi(\tau) \rangle$ [97, 113, 115]. Some projection-operator algorithms absorb this overlap factor into the field update and therefore propagate $|\psi_{\text{targ}}\rangle$ itself backward [113]; this useful shorthand is not the general terminal condition.

With Eq. (1.51) as its boundary condition, the costate is obtained from

$$i \frac{d}{dt} |\chi^{(n)}(t)\rangle = \hat{H}^{(n)}(t) |\chi^{(n)}(t)\rangle, \quad (1.52)$$

integrated from $t = \tau$ to $t = 0$. Thus, for bilinear Schrödinger dynamics, the state and costate obey the same differential equation, but the state is specified initially and propagated forward, whereas the costate is specified finally and propagated backward.

Once both trajectories are known, the first variation of the cost can be evaluated at every time. With the Hamiltonian convention of Eq. (1.45) and the costate

normalization in Eq. (1.51), it is

$$\frac{\delta J}{\delta u_k(t)} = 2 \operatorname{Im} \langle \chi^{(n)}(t) | \hat{H}_k | \psi^{(n)}(t) \rangle + 2\lambda u_k^{(n)}(t). \quad (1.53)$$

A gradient-descent step then gives

$$u_k^{(n+1)}(t) = u_k^{(n)}(t) - \eta \frac{\delta J}{\delta u_k(t)}, \quad \eta > 0, \quad (1.54)$$

Together, these equations define the iterative forward-state/backward-costate procedure summarized in Algorithm 1.

Algorithm 1 Forward-state/backward-costate optimization cycle

Require: Initial state $|\psi_0\rangle$, trial controls $u_k^{(0)}(t)$, and step size $\eta > 0$

- 1: Set the iteration index $n \leftarrow 0$
- 2: **repeat**
- 3: Propagate $|\psi^{(n)}(t)\rangle$ forward from 0 to τ using Eq. (1.50)
- 4: Set the terminal costate $|\chi^{(n)}(\tau)\rangle$ using Eq. (1.51)
- 5: Propagate $|\chi^{(n)}(t)\rangle$ backward from τ to 0 using Eq. (1.52)
- 6: Evaluate $\delta J / \delta u_k(t)$ from Eq. (1.53)
- 7: Update every control using Eq. (1.54)
- 8: $n \leftarrow n + 1$
- 9: **until** $J[u^{(n)}]$ has converged

Ensure: Stationary open-loop control fields $u_k(t)$

After each update, the forward and backward equations are solved again. Repeating this cycle until $J[u]$ converges yields a stationary control [103, 113, 116]. Alternative sign or normalization conventions for the costate modify the displayed gradient formula correspondingly but leave the control update invariant.

1.3.2. Numerical methods: GRAPE and CRAB

Analytical optimal controls are known only in special cases [117], and many-body problems generally require numerical optimization. Two representative methods are *Gradient Ascent Pulse Engineering* (GRAPE) [103, 116] and the *Chopped Random-Basis* (CRAB) method [105, 106].

The forward-state/backward-costate structure of GRAPE can be obtained from the weak PMP described above [115]. In this sense, GRAPE numerically realizes the same adjoint-gradient conditions, rather than the full pointwise maximization condition of the PMP. Historically, however, GRAPE was introduced through direct differentiation of time-sliced propagators rather than explicitly as a consequence of PMP [103].

GRAPE is a numerical method for continuous-time quantum optimal control. In its standard implementation, the continuous control fields are parameterized as piecewise constant over a finite time grid [103, 116]. Specifically, standard GRAPE divides $[0, \tau]$ into P equal intervals and writes

$$\Delta t = \frac{\tau}{P}, \quad t_p = p\Delta t, \quad u_k(t) = u_{k,p}, \quad t \in [t_{p-1}, t_p], \quad (1.55)$$

for $p = 1, \dots, P$. The dynamics remains continuous within each interval, with

$$\hat{H}_p = \hat{H}_0 + \sum_{k=1}^K u_{k,p} \hat{H}_k, \quad U_p = e^{-i\Delta t \hat{H}_p}. \quad (1.56)$$

The continuous PMP equations for the state, costate, gradient, and control update then take the compact GRAPE form

$$\begin{aligned} |\psi_p\rangle &= U_p |\psi_{p-1}\rangle, & |\psi_0\rangle &= |\psi(0)\rangle, \\ |\chi_P\rangle &= O |\psi_P\rangle, & |\chi_{p-1}\rangle &= U_p^\dagger |\chi_p\rangle, \\ g_{k,p} &\equiv \frac{\partial J}{\partial u_{k,p}} = 2 \operatorname{Re} \left\langle \chi_p \left| \frac{\partial U_p}{\partial u_{k,p}} \right| \psi_{p-1} \right\rangle + 2\lambda \Delta t u_{k,p}, & u_{k,p} &\leftarrow u_{k,p} - \eta g_{k,p}. \end{aligned} \quad (1.57)$$

All quantities on the right-hand side are evaluated for the controls at the current iteration. For the terminal-energy objective used in this Thesis, $O = \hat{H}_{\text{targ}}$ and $\lambda = 0$. For the K control channels defined in Eq. (1.45), the P intervals therefore give KP independently optimized amplitudes. CRAB instead restricts the search to a low-dimensional family of controls. Starting from an initial control $u_k^{(0)}(t)$, a canonical Fourier-CRAB parameterization, introduced in Ref. [105] and detailed in Ref. [106], is

$$\begin{aligned} u_k^{\text{CRAB}}(t) &= u_k^{(0)}(t) \left[1 + \frac{1}{\Lambda_k(t)} \sum_{n=1}^{N_c} [A_{kn} \sin(\omega_n t) + B_{kn} \cos(\omega_n t)] \right], \\ \omega_n &= \frac{2\pi n}{\tau} (1 + r_n), \quad r_n \in [-1/2, 1/2]. \end{aligned} \quad (1.58)$$

The boundary function satisfies $\Lambda_k(t) \rightarrow \infty$ as $t \rightarrow 0$ and $t \rightarrow \tau$, so the multiplicative correction approaches one and preserves the endpoint values of the initial control. Random offsets remove the rigidity of a pure harmonic truncation, while the finite basis limits the number of optimization variables. The coefficients $\{A_{kn}, B_{kn}\}$ can then be optimized by direct search, such as Nelder–Mead, using only repeated evaluations of the objective. This makes CRAB useful when gradients are unavailable or each dynamical simulation is expensive, although the parameterization itself is also compatible with gradient-based optimization.

Therefore, GRAPE and CRAB differ primarily in their control parameterization: GRAPE normally assigns an independently optimized control value $u_{k,p}$ to each channel k and time slice p , whereas CRAB searches over a smaller number of coefficients in a structured functional basis. In *dressed CRAB* (dCRAB), new basis

functions are introduced around the best current solution when the search stalls, opening directions that were absent from the initial truncation [108].

We study CRAB-based optimal-control methods in Chapters 2 and 3. Section 2.2.1 specializes Eq. (1.58) to sine modes with $\omega_n = (\pi n/\tau)(1 + r_n)$ and also considers a randomized Chebyshev basis. Section 3.3.1 further refines the frequency choice for the frustrated ring: the dimensionless frequencies $\omega_n\tau/\pi$ are sampled from Gamma distributions (Fig. 3.3), with different scales in the two dCRAB superiterations. Instead of optimizing every digital angle independently, these methods select structured families of smooth schedules. As the results chapters show, the smooth schedules we obtain can be transferred between digital and continuous implementations.

1.4. Towards more hardware-aware quantum algorithms

A major challenge in implementing QA and QAOA efficiently on current noisy intermediate-scale quantum (NISQ) devices [23] is the limited connectivity of the physical hardware. Target Hamiltonians may involve all-to-all or long-range interactions, whereas most hardware platforms typically support only quasi-local couplings. The Lechner–Hauke–Zoller (LHZ), or Parity, architecture addresses this challenge by mapping interacting pairs of logical spins to parity qubits that encode their relative alignment [118, 119]. After this Parity mapping, logical interactions are encoded in fully programmable local fields on parity qubits, at the cost of introducing parity constraints. These constraints are then enforced by penalty terms in continuous-time algorithms and by additional constraint unitaries in digital algorithms. Implementations of QA and QAOA within the parity architecture, referred to as Parity-QA [118] and Parity-QAOA [120], have demonstrated advantages across various models [121–124]. The Parity mapping restructures the accessible Hilbert space and introduces global constraints that must be enforced either through non-local penalty terms (in QA) or additional constraint unitaries (in QAOA). While enabling more efficient and less error-prone simulations on current devices, this approach typically involves trade-offs such as increased qubit overhead and a greater number of QAOA unitaries on the tested cases.

In Chapter 4, we apply the Parity architecture to the frustrated Ising ring and compare standard and parity formulations of QA and QAOA in terms of spectral gaps, circuit depth, and constraint-related implementation costs.

1.5. Outline of the Thesis

Having introduced the theoretical framework and the control tools used throughout this Thesis, we now summarize the contents and principal results of the following Chapters.

Chapter 2 investigates whether optimizing the annealing schedule can overcome the poor performance of linear digitized annealing on weighted MaxCut instances with anomalously small gaps. At equal circuit depth, we compare smooth CRAB parameterizations with iterative QAOA strategies. The optimized protocols use coherent nonadiabatic population transfer, and the smooth solutions retain their performance more reliably when reconstructed in continuous time or transferred between instances. The results were published in Ref. [46].

Chapter 3 removes the schedule constraint and treats the individual angles of the alternating circuit as independent controls for a frustrated Ising ring. Exact ground-state preparation is first reached at the critical depth $P_1^{\text{cr}} = (N^2 - 1)/4$, despite the exponentially closing gap of the reference annealing path. The result is explained by controllability and parameter counting on the accessible free-fermion manifold, while smooth CRAB schedules remain advantageous below the critical depth. This Chapter is based on Ref. [86] and includes the companion controllability analysis of Ref. [125].

Chapter 4 reformulates the same ring family in the parity architecture introduced in Section 1.4. The new encoding changes the annealing spectrum, the available digital generators, and the resource accounting. For the tested coupling families with distinct field magnitudes, Parity-QAOA reaches the exact target through factorized single-qubit dynamics at a layer count independent of system size. When the constraint unitary is required, four layers achieve $P_{\text{GS}} > 0.99$ for all tested sizes up to $N = 21$, with a total CNOT count linear in N . These results are reported in Ref. [126].

Finally, Chapter 5 draws together the lessons of the three studies, states their limitations, and discusses extensions towards interacting models, realistic hardware, and automated protocol design. The Appendices collect technical derivations and supplementary results for Chapters 2–4.

2. Optimal Control Solutions to MaxCut Problems

As discussed in Chapter 1, standard quantum annealing with a linear schedule can perform poorly on hard optimization instances with small spectral gaps, motivating the search for improved protocols. In this work, we present different techniques for improving on linear-schedule QA along two directions, conceptually distinct but leading to similar outcomes:

- (1) The first approach consists of constructing a Trotter-digitized QA (dQA), introduced in Section 1.2.4, with schedules parameterized in terms of Fourier modes or Chebyshev polynomials, inspired by the Chopped Random Basis algorithm (CRAB), reviewed in Section 1.3.2, for optimal control in continuous time. This method is explained in detail in Section 2.2.1.
- (2) The second approach is technically a Quantum Approximate Optimization Algorithm (QAOA), introduced in Section 1.2.4, whose solutions are found iteratively using linear interpolation or expansion in Fourier modes. This method is explained in detail in Section 2.2.2.

Both approaches emphasize finding smooth optimal schedule parameters, ultimately leading to hybrid quantum-classical variational algorithms of the alternating Hamiltonian Ansatz type.

We apply these techniques to MaxCut problems. A tutorial introduction to the basic MaxCut problem was already given in Section 1.1.1. The specific weighted 3-regular MaxCut problem studied in this Chapter will be introduced in Section 2.3. We focus on graphs with $N = 14$ vertices and on hard instances that exhibit a small spectral gap, for which a standard linear-schedule QA performs poorly.

We characterize the physics behind the optimal protocols for both the dQA and QAOA approaches, discovering *shortcuts to adiabaticity*-like dynamics, as discussed in Section 1.2.3. Furthermore, we study the transferability of such smooth solutions among hard instances of MaxCut at different circuit depths. Finally, we show that the smoothness pattern of these protocols obtained in a digital setting enables us to adapt them to continuous-time evolution, contrarily to generic non-smooth solutions. This procedure results in an optimized quantum annealing schedule that is implementable on analog devices.

2.1. Introduction

We start with quantum annealing (QA) and its standard linear schedule, introduced in Section 1.2.1. As shown through the Grover-search example in Section 1.2.2, linear-schedule QA requires an annealing time $\tau = O(N_s)$, whereas an appropriately tailored schedule recovers the quadratic speedup, $\tau = O(\sqrt{N_s})$. This demonstrates that improving on the linear schedule can be essential to achieve a quantum speedup whenever possible, while determining the appropriate schedule remains a complicated task that we partly address in this Chapter.

On a broader perspective, there is a large body of literature showing that QA faces strong limitations in the presence of exponentially small energy gaps, which can result in exponentially increasing annealing times with system size [79, 80, 127, 128]. Various techniques have been proposed to overcome this limitation within the paradigm of continuous-time Schrödinger evolution, for instance employing the toolbox of Quantum Control (QC) [110].

A general strategy would be to perform a state-dependent Shortcut to Adiabaticity (STA) [129], literally a “fast route” to the ground state of the final target Hamiltonian in a shorter time τ , as compared to a slow QA (adiabatic) evolution. A specific STA technique, known as transitionless or counter-diabatic driving [82, 83], involves adding selected terms to the time-evolving Hamiltonian, which are designed to suppress transitions out of the instantaneous ground state. However, exact implementations of this scheme are difficult in many-body systems, as they require hard-to-devise non-local terms. Variational approaches have been introduced [84, 121, 130–135] to overcome these difficulties.

Other techniques have been specifically introduced within the field of Adiabatic Quantum Computation [65], including the addition of waiting times (pauses) in the schedule [136], reverse quantum annealing [137], or the introduction of Hamiltonian terms that transform the nature of the critical points encountered during the evolution, from 1st to 2nd order [138].

Various methods of Quantum Control appear particularly well-suited for achieving a ground-state STA, notably the Chopped Random Basis (CRAB) algorithm [105, 106] and follow-up developments [108, 109]. In CRAB, the various control functions are expanded in terms of a finite basis set of functions, usually a Fourier basis — but polynomials have also been used [139] —, hence transforming the functional minimization problem into a finite-dimensional optimization. Alternatively, a linear piece-wise decomposition of the schedule between a series of randomly chosen points has been explored [140] to optimize a frustrated Ising ring model [80], which shows an exponentially small gap in the excitation spectrum, thereby posing severe difficulties for QA.

Parallel developments in the field of Variational Quantum Algorithms (VQA) have

approached the problem from the alternative perspective of a digitized quantum evolution [73], whereby a series of different unitary operators are applied to the initial simple-to-construct state $|\psi_0\rangle$. These approaches include the Quantum Approximate Optimization Algorithm (QAOA) [25], originally introduced as an alternative to QA for classical optimization problems, and the Variational Quantum Eigensolver (VQE) [141], specifically aiming at solving quantum chemistry problems and, more generally, quantum ground state preparation. These digitized techniques offer the practical advantage of a native digital setting that is well adapted to universal gate-based [4] quantum computing hardware, and encompass a large class of hybrid quantum-classical algorithms [91, 92, 142].

Yet, the trade-off between the expressivity and the trainability of parameterized quantum circuits remains a substantial challenge for the practical implementation of this framework on real quantum devices [143, 144]. On top of the presence of statistical and physical noise, the optimization landscape of non-trivial (or classically-simulable) parameterized quantum circuits is affected by the presence of low-quality local minima [144] and vanishing gradients (barren plateaus) [145–147]. Several approaches to limit or avoid barren plateaus have been proposed and applied to specific tasks [148–154]. At the same time, developing smart initializations [155] or iterative optimization schemes [95] may prove essential to successfully train a VQA.

Quantum optimal control techniques may help, in this digital framework, to single out optimal or nearly-optimal variational parameters that are smooth functions of the layer index, which have been observed in several works [49, 87, 95, 156–160]. Such smooth optimal solutions exhibit notable properties and have proved effective in avoiding local minima [95] and circumventing barren plateaus [153]. Remarkably, it has been empirically demonstrated [153, 159] that this class of solutions is transferable across different problem sizes, from small to large, for the quantum many-body ground state preparation at any point of the phase diagram, highlighting their versatility and practical applicability in hybrid quantum-classical computing.

In this Chapter, we utilize established quantum optimal control methods to systematically generate smooth optimal solutions for combinatorial optimization problems. For definiteness, our focus is on MaxCut instances of weighted 3-regular graphs, i.e. graphs where each vertex has exactly three neighbors. However, our techniques can be generalized to more complicated graph structures or even to other combinatorial optimization problems. In particular, we select hard MaxCut instances characterized by a small energy gap, a regime in which standard *linear-schedule* QA is expected to yield poor performance. The rationale behind this choice is that a closing gap is expected to occur in most real-world use cases of NP-hard classical optimization problems tackled with QA, especially for growing system size [79]. As our results do not depend explicitly on the regularity properties of the graphs, we expect similar outcomes for regular graphs of different degree.

We characterize the dynamics implemented by such smooth and optimal parameters, unveiling the non-adiabatic mechanism that enables these techniques to outperform adiabatic protocols. Additionally, we observe promising transferability results across different *hard* 3-regular instances with the same number of vertices.

Although these techniques are applied in a discrete-time setting, making them suitable for digital devices, we demonstrate that smooth solutions can be effectively transferred to analog devices with comparable performance. This highlights the adaptability of smooth solutions for continuous-time quantum annealing, yielding an optimized problem-dependent schedule $s(t)$, thereby broadening their range of applicability across different quantum computing platforms.

The Chapter is structured as follows. In Section 2.2 we present the different methods we implemented to generate smooth solutions. In Section 2.3 we describe the MaxCut problem and the procedure utilized to single out hard instances. In Section 2.4 we present our results, and we draw a comparison among the different techniques and their performance. We characterize the non-adiabatic mechanism associated with smooth solutions and prove transferability among different hard instances. Finally, we implement these smooth solutions in a continuous-time evolution framework, showing their suitability for analog quantum devices. In Sec. 2.5 we summarize the significance of our results and discuss possible future work.

2.2. From QA to variational quantum algorithms

2.2.1. QA and optimal control

Building on the reviews of QA, digitization, and quantum optimal control in Sections 1.2.1, 1.2.4, and 1.3.1, we now specialize that framework to the problem studied in this Chapter. Standard QA, including analog implementations [71], interpolates between an easily prepared transverse-field driver and a target Hamiltonian and ideally follows its instantaneous ground state adiabatically [65]; near a critical point or first-order phase transition, however, the required annealing time can become very large [128], while slowing the schedule near an a priori unknown bottleneck is generally infeasible [161]. An optimal-control problem can be similarly formulated for a classical optimization task encoded in a diagonal target Hamiltonian $\hat{H}_{\text{targ}} = \hat{H}_z$, allowing in principle for non-local models [162]. We restrict the present study to MaxCut. We simulate the resulting digitized dynamics using efficient matrix-vector multiplication in the computational basis [163]. The corresponding state obeys

$$i\hbar \frac{d}{dt} |\Psi(t)\rangle = \hat{H}(t) |\Psi(t)\rangle, \quad (2.1)$$

with a time-dependent Hamiltonian of the interpolating form

$$\hat{H}(t) = A^z(t) \hat{H}_z + A^x(t) \hat{H}_x, \quad (2.2)$$

where we impose the boundary conditions $A^z(0) = 0$, $A^x(\tau) = 0$, and $A^z(\tau) > 0$. The goal is to determine the optimal form of the two driving fields $A^z(t)$ and $A^x(t)$ such that the variational energy

$$E^{\text{fin}}(\tau) = \langle \Psi(\tau) | \hat{H}_z | \Psi(\tau) \rangle \quad (2.3)$$

is *minimal*. This optimal control problem can be formulated either for a pre-determined total time τ , or by including the optimization of τ itself as part of the task.

Such a functional minimization is generally intractable. To simplify it, we transform it into a *finite-dimensional* minimization by using the so-called Chopped Random Basis (CRAB) algorithm [105, 106, 108, 164], i.e., we expand the fields in terms of a selected discrete set of N_c basis functions $\{f_n(t)\}$:

$$\begin{cases} A^x(t) = \left(1 - \frac{t}{\tau}\right) \left(1 + \sum_{n=1}^{N_c} C_n^x f_n(t)\right) \\ A^z(t) = \frac{t}{\tau} \left(1 + \sum_{n=1}^{N_c} C_n^z f_n(t)\right) \end{cases} . \quad (2.4)$$

Notice that the fields automatically satisfy the two constraints $A^z(0) = 0$ and $A^x(\tau) = 0$, and reduce to the linear schedule QA form, i.e. $A^z(t) = s(t) = t/\tau$ and $A^x(t) = 1 - s(t)$, if all the expansion coefficients $C_n^{x/z}$ are set to zero. Concerning the basis functions, a common choice is given by Fourier modes. Here we take:

$$f_n(t) = \sin(\omega_n t) \quad \omega_n = \frac{\pi n}{\tau} (1 + r_n) , \quad (2.5)$$

where $r_n \in [-\frac{1}{2}, \frac{1}{2}]$ is a random number, introducing a multiplicative noise in the frequencies, which is useful to improve the expressive power of this finite-dimensional *Ansatz* [106]. We dub this choice F-CRAB, where F stands for Fourier. We have verified empirically that this works better than an additive noise $\omega_n = (\pi/\tau)(n + r_n)$.

An alternative option consists of using

$$f_n(t) = \cos \left(n(1 + r_n) \arccos(t/\tau) \right) , \quad (2.6)$$

where $r_n \in [-\frac{1}{2}, \frac{1}{2}]$ is again a random number. For $r_n = 0$, these functions coincide with the Chebyshev polynomials of the first kind $T_n(t/\tau)$, while for $r_n \neq 0$ they are natural extensions of polynomials with non-integer powers [165–167], often known as signomials. We refer to this choice as C-CRAB.

In both variants of the CRAB algorithm, the variational energy in Eq. (2.3) is optimized with respect to the set of $2N_c$ expansion coefficients $C_n^{x/z}$. Instead of repeatedly performing multiple local optimizations on the same energy landscape from various starting points, the strategy involves generating a new landscape each time. This is achieved by selecting a different set of random frequencies [106].

To classically simulate the quantum dynamics in Eq. 2.1, a time-discretization is necessary: the continuous-time dynamics is discretized by introducing a finite — in principle, sufficiently small — time-step, denoted as $\Delta_t = \tau/P$. We define an associated time grid $t_m = (m - \frac{1}{2}) \Delta_t$, with $m = 1 \cdots P$, at which we evaluate the fields, setting $A_m^{x/z} = A^{x/z}(t_m)$. We then approximate the fields as constant in each interval $[t_m - \Delta_t/2, t_m + \Delta_t/2]$. The ensuing evolution operator over a time-step Δ_t is further approximated with a first-order Trotter splitting:

$$e^{-i(\theta_m^x \hat{H}_x + \theta_m^z \hat{H}_z)} \approx e^{-i\theta_m^x \hat{H}_x} e^{-i\theta_m^z \hat{H}_z} \stackrel{\text{def}}{=} \hat{U}(\theta_m^x, \theta_m^z), \quad (2.7)$$

where we defined (by setting $\hbar = 1$)

$$\begin{cases} \theta_m^x = \Delta_t A_m^x = \frac{P-(m-\frac{1}{2})}{P} \Delta_t \left(1 + \sum_{n=1}^{N_c} C_n^x f_n(t_m) \right) \\ \theta_m^z = \Delta_t A_m^z = \frac{(m-\frac{1}{2})}{P} \Delta_t \left(1 + \sum_{n=1}^{N_c} C_n^z f_n(t_m) \right) \end{cases}. \quad (2.8)$$

The time-discretized and Trotter-split Schrödinger evolution results in a *digitized* dynamics leading to the final state

$$|\Psi_P^{\text{dQA}}(\boldsymbol{\theta})\rangle = \hat{U}(\theta_P^x, \theta_P^z) \cdots \hat{U}(\theta_1^x, \theta_1^z) |\Psi_0\rangle, \quad (2.9)$$

where $\boldsymbol{\theta} = (\boldsymbol{\theta}^x, \boldsymbol{\theta}^z)$. Note that $\boldsymbol{\theta}^x = (\theta_1^x \cdots \theta_P^x)$ and $\boldsymbol{\theta}^z = (\theta_1^z \cdots \theta_P^z)$ are both P -dimensional vectors, each depending on the N_c expansion coefficients $\mathbf{C}^x = (C_1^x \cdots C_{N_c}^x)$ and $\mathbf{C}^z = (C_1^z \cdots C_{N_c}^z)$, respectively. Albeit we denote this state with the apex dQA, an acronym for digitized Quantum Annealing [73, 168], this is actually a digitized version of a generic optimal control dynamics, reducing to the linear-schedule digitized-QA of Ref. [168] only when all the expansion coefficients $C_n^{x/z}$ vanish.

Previous evidence on linear-schedule dQA [168] showed that even a first-order Trotter splitting is adequate up to values of Δ_t of order 1. In principle, the order of the Trotter approximation could be increased by utilizing, e.g., a second or fourth-order expansion. However, the focus here relies on the optimization of the variational parameters in the digitized state rather than on approximating a prescribed continuous-time evolution. Therefore, we keep the form of the state in Eqs. (2.7)-(2.9), since improving the Trotter order would lead to the same alternating *Ansatz* structure [93], only at a higher depth.

Finally, to enlarge the variational manifold, it would be beneficial to include the time-step Δ_t as an additional variational parameter. To keep a *linear* dependence of the angles in Eq. 2.8 on the set of variational parameters, we rather fix $\Delta_t = 1$,

and introduce a new free parameter C_0 as follows:

$$\begin{cases} \theta_m^x = \frac{P - (m - \frac{1}{2})}{P} \left(C_0 + \sum_{n=1}^{N_c} C_n^x f_n(t_m) \right) \\ \theta_m^z = \frac{(m - \frac{1}{2})}{P} \left(C_0 + \sum_{n=1}^{N_c} C_n^z f_n(t_m) \right) \end{cases} . \quad (2.10)$$

This approach leads to a digitized version of quantum optimal control without a constraint on the time-step Δt , while preserving the linear dependence of the angles on the variational parameters. Indeed, an optimal protocol characterized by the $2N_c + 1$ variational parameters $\mathbf{C} = (C_0, \mathbf{C}^x, \mathbf{C}^z)$ is equivalent to the digitized state with angles in Eq. 2.8, by identifying C_0 with Δt and rescaling the other $2N_c$ parameters by its value.

The variational energy for this digital optimal control problem is given by:

$$E_P^{\text{dQA}}(\mathbf{C}) = \langle \Psi_P^{\text{dQA}}(\boldsymbol{\theta}) | \hat{H}_z | \Psi_P^{\text{dQA}}(\boldsymbol{\theta}) \rangle , \quad (2.11)$$

where the $2P$ parameters $\boldsymbol{\theta} = (\boldsymbol{\theta}^x, \boldsymbol{\theta}^z)$ are regarded as functions of the $2N_c + 1$ coefficients $\mathbf{C} = (C_0, \mathbf{C}^x, \mathbf{C}^z)$. Besides the clear benefit of digitization, which is the lower computational cost of a time-evolution in a classical simulation, the digitized dynamics for a two-body $\hat{H}_z$ admits a standard decomposition into elementary one- and two-qubit gates, amenable to direct implementation on gate-based devices [73]. As an important side benefit, we can calculate *analytically* the gradients of the variational energy with respect to the variational parameters $\mathbf{C}$ [103, 169], see App. A.1 for details, hence allowing for faster gradient-based optimization routines. If $\mathbf{C}^*$ denotes the (local or global) optimal parameters found, the corresponding variational energy is given by:

$$E_P^{\text{fin}} = E_P^{\text{dQA}}(\mathbf{C}^*) . \quad (2.12)$$

Our practical implementation of the F-CRAB algorithm relies on an iterative procedure that is particularly effective in obtaining smooth solutions. In operational terms, these are characterized by a regular *smooth* pattern of each of the two optimal vectors $(\theta_1^x \cdots \theta_P^x)$ and $(\theta_1^z \cdots \theta_P^z)$, when plotted vs the layer index $m = 1 \cdots P$. The rationale behind the implementation of an iterative scheme is that a direct optimization of $2N_c + 1$ coefficients often gets trapped in a low-quality local minimum for large N_c . Even when optimal or nearly-optimal solutions are found, these are characterized by irregular non-smooth behavior, namely exhibiting sharp variations for a small increase of m .

To address this issue, we start by optimizing the $N_c = 2$ case. By following the CRAB prescription, the $N_c = 2$ frequencies are randomly generated for n_r times, a new optimization of the $2N_c + 1 = 5$ coefficients is performed for each realization, and the best result in terms of variational energy is selected. Then, we optimize the

cases $N'_c > N_c$ according to the following procedure. We use the optimal $2N_c + 1$ parameters obtained in the previous iteration as a warm start for the optimization of the new $2N'_c + 1$ coefficients, initializing the remaining $2(N'_c - N_c)$ coefficients to zero. Crucially, at each iterative step, the old N_c frequencies are kept fixed, whereas the new $N'_c - N_c$ frequencies are once again randomly generated for n_r times, selecting the optimal outcome with minimal variational energy.

We iterate this procedure, increasing the number of frequencies up to $N_c = P/2$. We choose this threshold as for $N_c > P/2$ we do not observe any significant performance improvement, compared with the increase of computational cost. This iterative optimization involves a computational overhead; however, for large P , it turns out to be sufficient to increase N_c in steps of 10 in order to obtain smooth optimal curves.

The occurrence of irregular solutions in a direct optimization for F-CRAB can be intuitively understood by considering that, for large N_c , the Fourier expansion involves high-frequency modes that may induce fast oscillations in the optimal parameters $\boldsymbol{\theta}$. Nevertheless, limiting the expansion to low-frequency modes does not empirically guarantee a good performance in terms of variational energy. The relation between irregular and smooth solutions is analyzed in greater detail in Sec. 2.4.4.

Such an iterative procedure, as it turns out, is not needed for C-CRAB. Indeed, this method provides high-quality smooth curves without any need for a warm-start: for each value of P and of N_c , we set a linear dQA initial condition for the optimization, i.e. $C_0 = 1$ and $\mathbf{C}^x = \mathbf{C}^z = 0$. This difference can be explained by noting that the Chebyshev polynomials of degree n oscillate slower compared to the n -th Fourier modes.

2.2.2. QAOA

The previous approach to the optimal control problem leads to a form of the variational state which is identical to that of the Quantum Approximate Optimization Algorithm (QAOA) by Farhi *et al.* [25]. The QAOA state is expressed as:

$$|\Psi_P^{\text{QAOA}}(\boldsymbol{\theta})\rangle = \widehat{U}(\theta_P^x, \theta_P^z) \cdots \widehat{U}(\theta_1^x, \theta_1^z) |\Psi_0\rangle, \quad (2.13)$$

where, again,

$$\widehat{U}(\theta_m^x, \theta_m^z) = e^{-i\theta_m^x \widehat{H}_x} e^{-i\theta_m^z \widehat{H}_z}. \quad (2.14)$$

The difference is that, here, $\boldsymbol{\theta}$ are interpreted as $2P$ free variational parameters, without any explicit reference to time-dependent smooth control fields. Although the form of the state in Eq. (2.13) is identical to that of the digitized optimal control state in Eq. (2.9), the idea behind the two approaches is, in principle, different: the first strategy is based on the optimization of the expansion coefficients over a

truncated basis of smooth driving fields, while the QAOA strategy is, in principle, agnostic of any smoothness property of optimal solutions. Nonetheless, smooth optimal solutions have been observed in the QAOA framework in several case studies [49, 87, 95, 153, 156–160, 170]. Here, the variational energy is defined as:

$$E_P^{\text{QAOA}}(\boldsymbol{\theta}) = \langle \Psi_P^{\text{QAOA}}(\boldsymbol{\theta}) | \hat{H}_z | \Psi_P^{\text{QAOA}}(\boldsymbol{\theta}) \rangle, \quad (2.15)$$

and if $\boldsymbol{\theta}^*$ denotes a (local or global) optimal solution, then the variational energy is given by:

$$E_P^{\text{fin}} = E_P^{\text{QAOA}}(\boldsymbol{\theta}^*). \quad (2.16)$$

The minimization of this cost function is, in general, a highly non-trivial task since random-start local optimization routines tend to get trapped into one of the many local minima of the $2P$ -dimensional search space [144]. Moreover, most of this landscape is plagued by the phenomenon of *barren plateaus* [145], whereby the gradients of the function to be minimized are exponentially small in the number of qubits.

Notice that digitized CRAB not only biases the search towards smooth solutions but also effectively explores a restricted portion of the QAOA energy landscape. Combined with an iterative procedure such as the one described for F-CRAB, digitized optimal control methods represent an effective way to optimize the QAOA variational parameters. Previous research on QAOA proposed alternative iterative methods to cope with the difficulties of an unfavorable optimization landscape, which are briefly reviewed below and compared to CRAB in Section 2.4.1.

Interpolation schemes

INTERP. One natural strategy is to proceed iteratively from a lower- P minimum to larger P' , for instance $P' = P + 1$, by appropriate interpolation of the known solution. More in detail, suppose that $\boldsymbol{\theta}^{(*,P)}$ is a previously determined (local or global) optimal schedule. Then, we can construct a candidate solution for $(P + 1)$ with the following procedure. Define an initial $2(P + 1)$ -dimensional vector $\boldsymbol{\theta}^{(0,P+1)}$ as follows (the same expression applies both to θ_m^x and θ_m^z , hence we omit indicating the label x/z):

$$\theta_m^{(0,P+1)} = \frac{m-1}{P} \theta_{m-1}^{(*,P)} + \frac{P-m+1}{P} \theta_m^{(*,P)}, \quad (2.17)$$

where $m = 1, \dots, P+1$. Notice that, although in principle referenced in the formula, $\theta_0^{(*,P)}$ and $\theta_{P+1}^{(*,P)}$ are always multiplied by zero, hence they are irrelevant. In particular, $\theta_1^{(0,P+1)} = \theta_1^{(*,P)}$ and $\theta_{P+1}^{(0,P+1)} = \theta_P^{(*,P)}$. The candidate solution $\boldsymbol{\theta}^{(0,P+1)}$ is utilized as a warm start for a local gradient-based optimization, yielding the optimal $\boldsymbol{\theta}^{(*,P+1)}$. This strategy was introduced in Ref. [95] and dubbed INTERP. In practice, it can be applied starting e.g. from $P = 2$, and iterating the algorithm for increasing $P = 3, 4, \dots$ up to the desired final value of P_{max} . Alternatively, it can be slightly modified to allow for $P' = P + \Delta P$, with $\Delta P > 1$.

LogINTERP. An alternative construction, introduced in Ref. [49], proceeds by *doubling* P , specifically by using an optimal solution at depth P to seed the optimization for depth $2P$. We apply it in two variants, each adapted to different boundary conditions for θ^z and θ^x .

Whenever the parameters are likely bound to vanish at the initial time-step, as for θ^z , we take as an initial guess:

$$\begin{cases} \theta_{2m}^{(0,2P)} = \theta_m^{(*,P)} & m = 1 \dots P \\ \theta_{2m-1}^{(0,2P)} = \frac{1}{2}(\theta_{m-1}^{(*,P)} + \theta_m^{(*,P)}) & m = 1 \dots P \end{cases}, \quad (2.18)$$

where the boundary condition $\theta_0^{(*,P)} = 0$ should be used. Alternatively, when the vanishing boundary condition is at the final time-step, as for θ^x , we take as initial guess:

$$\begin{cases} \theta_{2m-1}^{(0,2P)} = \theta_m^{(*,P)} & m = 1 \dots P \\ \theta_{2m}^{(0,2P)} = \frac{1}{2}(\theta_m^{(*,P)} + \theta_{m+1}^{(*,P)}) & m = 1 \dots P \end{cases}, \quad (2.19)$$

where the boundary condition $\theta_{P+1}^{(*,P)} = 0$ should be used. These qualitative features in the boundary conditions for the optimal angles have been observed in QAOA-like numerics for a breadth of different tasks [47, 87, 95, 159, 171–173]. The optimal $\theta^{(*,2P)}$ is found by applying a local gradient-based optimization starting from the guess $\theta^{(0,2P)}$.

Fourier-based schemes

FOURIER. A strategy based on expansion in Fourier modes was introduced in Ref. [95], denoted as FOURIER. It is based on the following parameterization:

$$\begin{cases} \theta_m^z = \sum_{n=1}^{N_c} C_n^z \sin\left((n - \frac{1}{2})(m - \frac{1}{2})\frac{\pi}{P}\right) \\ \theta_m^x = \sum_{n=1}^{N_c} C_n^x \cos\left((n - \frac{1}{2})(m - \frac{1}{2})\frac{\pi}{P}\right) \end{cases}. \quad (2.20)$$

Observe the similarities with the F-CRAB approach we discussed previously. Note, however, the absence of a linear dQA-like term multiplying the Fourier decomposition, and also the fact that the Fourier frequencies are fixed, and randomness is only used in searching for the optimal coefficients. Different variants of this strategy postulate how the final required $P_{\max}$ is reached starting from lower values, and how many Fourier modes are allowed, either $N_c = P$, or a fixed predetermined value of $N_c < P$. Details are given in the original Ref. [95], and summarized in App. A.2 for convenience.

2.3. The MaxCut problem

A general introduction to MaxCut and its Ising encoding was given in Section 1.1.1; here we specialize to the weighted instances studied in this Chapter.

The MaxCut problem is a classical combinatorial problem in the NP-complete complexity class. It is defined on a graph $\mathcal{G}$ of N vertices, with edges denoted as $\mathcal{E}$. We consider weighted graphs, i.e., each edge $(jj') \in \mathcal{E}$ is associated to a weight $J_{jj'}$ randomly drawn in the interval $[0, 1]$ with uniform distribution. The goal is to find a path which cuts the largest possible number of edges (more precisely, the largest sum of weights $J_{jj'}$) so that the graph $\mathcal{G}$ is partitioned into two subgraphs. Following the procedure described in Section 1.1.1, we reformulate the problem in terms of the ground state search of the antiferromagnetic spin Hamiltonian introduced in Eq. (1.5),

$$\hat{H}_{\text{targ}} = \hat{H}_z = \frac{1}{2} \sum_{(jj') \in \mathcal{E}} J_{jj'} (\hat{\sigma}_j^z \hat{\sigma}_{j'}^z - 1), \quad (2.21)$$

where each term in the sum has been rescaled and shifted to let the ground state energy coincide with (minus) the maximum number of cut edges (for the unweighted case with $J_{jj'} = 1$). The maximum energy is $E_{\text{max}} = 0$, associated with the configuration where all the spins are identical (ferromagnetic state).

Figure 2.1 illustrates a specific MaxCut instance, previously analyzed in Ref. [95]. This instance is particularly hard for a linear-schedule QA approach, as one quickly discovers by plotting the instantaneous spectrum of the annealing Hamiltonian $\hat{H}(s) = s\hat{H}_z + (1-s)\hat{H}_x$ with $s = t/\tau$, as shown in Fig. 2.2.

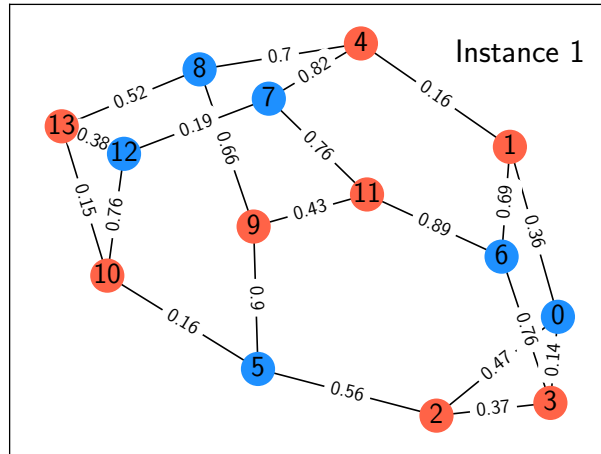


Figure 2.1.: A specific MaxCut instance with $N = 14$ vertices considered in our work. The red and the blue colors indicate the two optimal partitions that solve the MaxCut problem. This instance has been examined in full detail also in Ref. [95], proving particularly difficult for a linear-schedule QA approach due to a very small minimum gap.

In our work, we focus on a set of similarly hard instances of the MaxCut problem

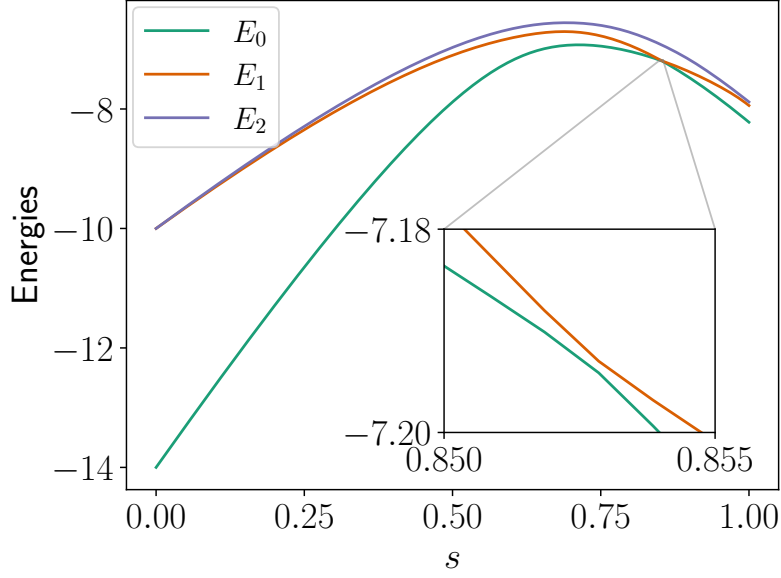


Figure 2.2.: Low-energy spectrum of the annealing Hamiltonian $s\hat{H}_z + (1-s)\hat{H}_x$ as a function of $s = t/\tau$, for the MaxCut instance 1. In the inset, we zoom on the closing gap, which is of order 10^{-3} .

with $N = 14$ vertices. To select hard instances, we first generate several random 3-regular weighted graphs and compute the instantaneous spectrum of the linear annealing Hamiltonian. In Fig. 2.3 we show the normalized cumulative distribution of the minimum gaps for 100 random instances. We define as hard the instances having a minimum gap $\Delta_{\min}$ of the order $\sim 10^{-3}$.

Since linear QA is not expected to solve these instances without resorting to very large values of the annealing time τ , a MaxCut optimal solution may be obtained more effectively by employing the quantum algorithms introduced in Section 2.2. We tested our methods on a total of 10 hard instances.

As a figure of merit for the effectiveness of our techniques, we will mostly use the residual energy, defined as a rescaled version of the variational energy in Eqs. (2.12) and (2.16):

$$\epsilon_P^{\text{res}} = \frac{E_P^{\text{fin}} - E_{\min}}{E_{\max} - E_{\min}}, \quad (2.22)$$

where $E_{\min}$ and $E_{\max}$ are the minimum (ground-state) and maximum energy of the problem Hamiltonian in Eq. (2.21). A traditional figure of merit used in the MaxCut literature [95] is $r = E_P^{\text{fin}}/E_{\min}$. Notice that $E_{\min} < 0$ and $0 < r < 1$, with r approaching 1 as the solution improves. The two quantities are related as $1 - r = \epsilon^{\text{res}}$ since $E_{\max} = 0$ when $J_{ij} > 0$ (ferromagnetic state).

An alternative figure of merit is the fidelity [4] between the final (optimal) state $|\Psi_P\rangle$, and the exact ground state of $\hat{H}_{\text{targ}}$:

$$\mathcal{F} = |\langle \Psi_{\text{gs}} | \Psi_P \rangle|^2. \quad (2.23)$$

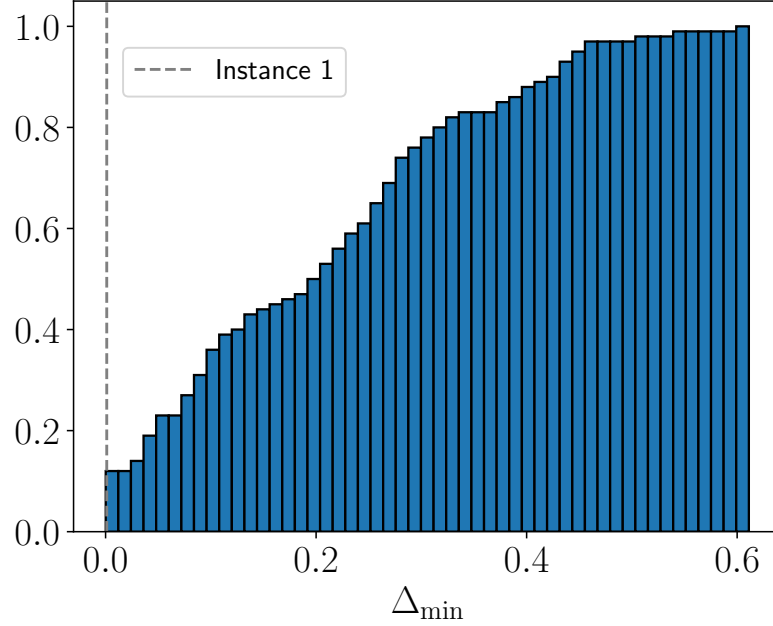


Figure 2.3.: Cumulative distribution of the minimum gaps of the annealing Hamiltonian for 100 weighted 3-regular graphs of $N = 14$ vertices. A vertical dashed line indicates the small minimum gap for instance 1: similarly hard instances are selected for our numerics. To compute the gap, we considered a linear annealing schedule.

In principle, very good solutions in terms of residual energy ϵ_p^{res} might have low fidelity, if the lowest excited states of the target Hamiltonian are very close in energy to the ground state. In the context of the MaxCut problems under consideration, we note that the infidelity, expressed as $1 - \mathcal{F}$, gives results that are qualitatively consistent with the residual energy ϵ_p^{res} for all the methods investigated.

2.4. Results

2.4.1. Comparison of different methods

In this section, we compare the performance of the different methods illustrated in Sec. 2.2. These include dQA with a linear schedule (dQA-LIN) or using Fourier- or Chebyshev-based CRAB (dQA-F-CRAB, dQA-C-CRAB) and QAOA-based techniques, in particular QAOA-INTERP, QAOA-LogINTERP, and QAOA-FOURIER. For ease of presentation, we show results focusing on the instance 1 depicted in Fig. 2.1, the main one considered also in Ref. [95]. We verified that our results are valid for all the other hard instances analyzed (See Appendix A.4 for further analysis).

In Fig. 2.4, we compare residual energy and infidelity values obtained using the

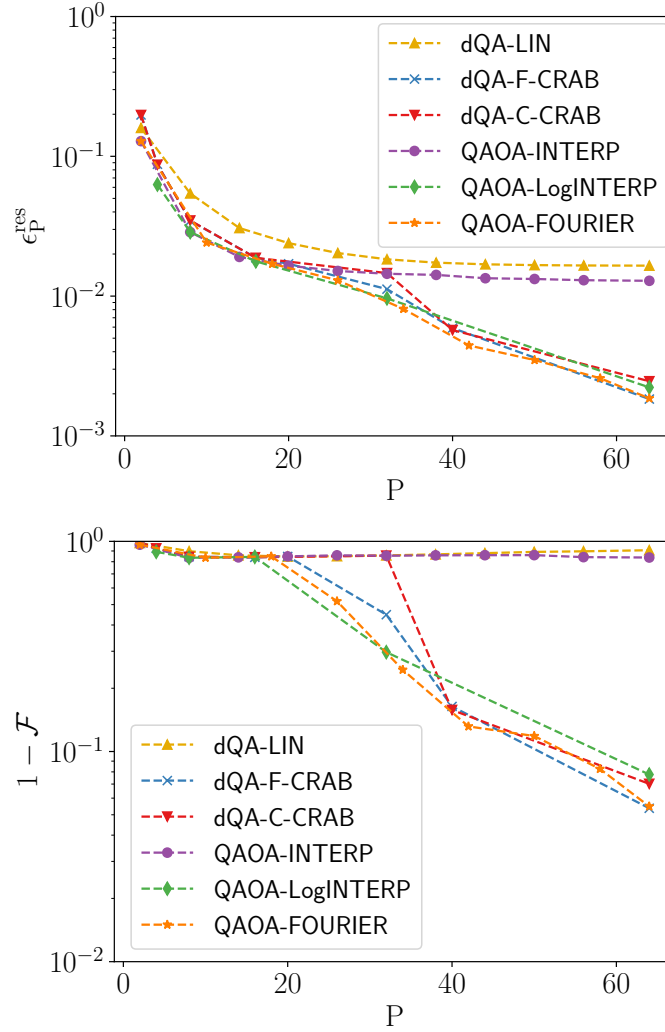


Figure 2.4.: Residual energies (top) and infidelity (bottom) as a function of P for the instance 1 and for different methods. Lines are guide to the eye. Infidelity and residual energy behave consistently. We see that digitized linear-schedule Quantum Annealing (dQA-LIN), and QAOA-INTERP get stuck in a local minimum for large P . The four methods that perform better are dQA-F-CRAB, dQA-C-CRAB, QAOA-FOURIER and QAOA-LogINTERP.

different methods. One immediately observes that, for shallow circuits, $P \leq 20$, all methods behave roughly in the same way. For larger values of P , two methods get trapped into some kind of plateau: dQA-LIN and QAOA-INTERP. Indeed, due to the small value of the minimum gap of the selected instance, the performance of a linear-schedule digitized QA (dQA-LIN) is rather poor. Perhaps surprisingly, a similar saturation occurs for QAOA-INTERP [95]. On the contrary, the two CRAB-based dQA, as well as QAOA-LogINTERP and QAOA-FOURIER perform much better, as witnessed by decreasing values of ϵ_P^{res} and infidelity $1 - \mathcal{F}$, for increasingly larger P . These methods show similar performance, and all provide smooth optimal schedules, as shown in Fig. 2.5.

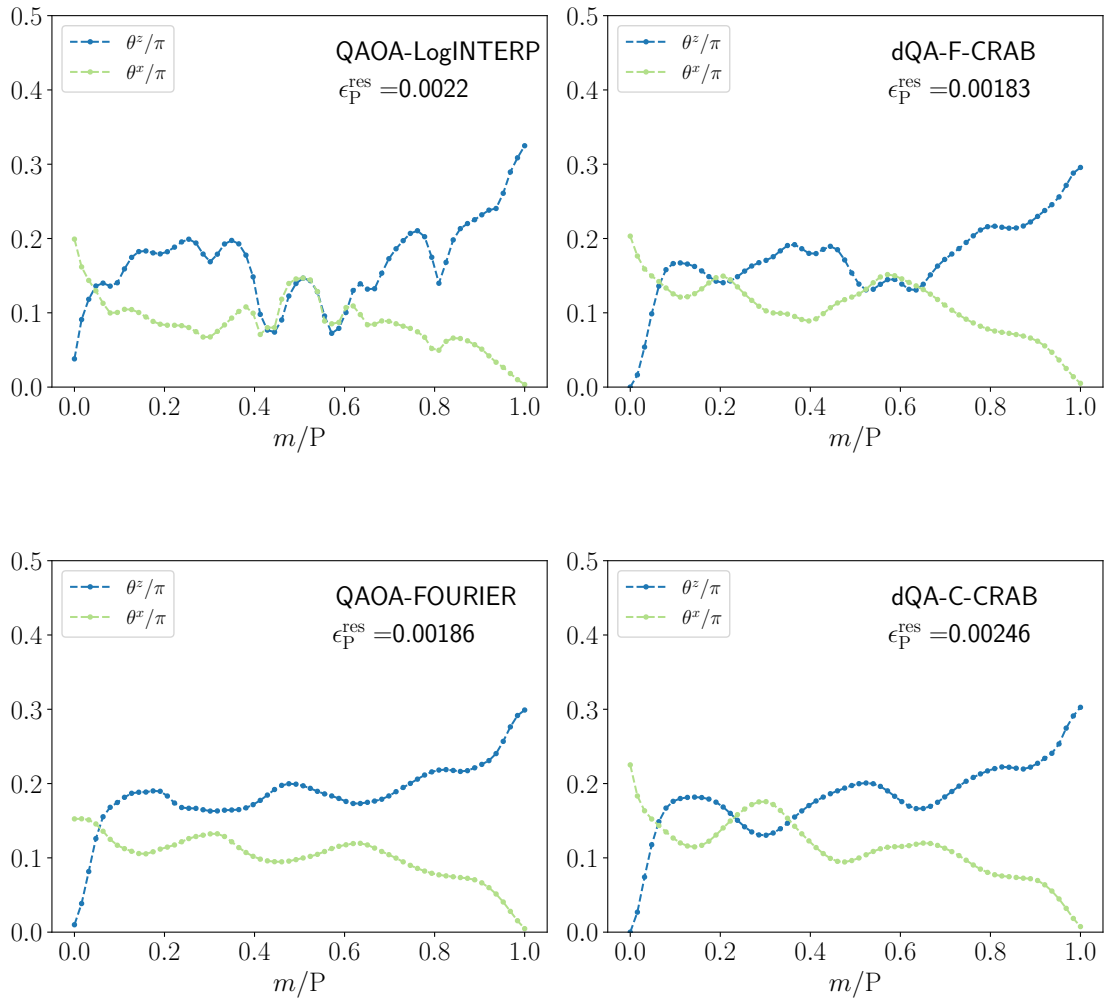


Figure 2.5.: Optimal parameters obtained with different methods for instance 1 and $P = 64$. For C-CRAB, F-CRAB, and FOURIER, the number of functional coefficients is $N_c = P/2 = 32$.

To conclude this section, we comment on the different performances of QAOA-INTERP and QAOA-LogINTERP. In both methods, the initial guess for the gradient-based optimization routine at depth $P' > P$ is close to the optimal pa-

rameters from the previous iteration at depth P . However, for QAOA-INTERP, since $P' = P + 1$, only two points are added to the previous optimal solution $\theta^{*,P}$: This does not allow to perturb the previous sub-optimal solution enough to escape from the local minima encountered at depth P . In contrast, for QAOA-LogINTERP, $P' = 2P$, and the number of parameters to optimize increases by a factor of 2. This enables the gradient descent routine to explore more directions in parameter space, making it easier to escape from local minima and to improve the quality of the solution.

2.4.2. Eigenstate population and shortcut to adiabaticity

Here, we analyze the details of the dynamics implemented by the different methods presented above. In order to describe the time evolution of the system, we introduce an effective Hamiltonian that generates the digital dynamics:

$$\hat{U}(\theta_m^x, \theta_m^z) = e^{-i\theta_m^x \hat{H}_x} e^{-i\theta_m^z \hat{H}_z} \equiv e^{-i\hat{H}_m^{\text{eff}}} . \quad (2.24)$$

The effective Hamiltonian $\hat{H}_m^{\text{eff}}$ can be expressed in terms of $\hat{H}_x$, $\hat{H}_z$ and their nested commutators, using the Baker–Campbell–Hausdorff (BCH) formula:

$$\begin{aligned} \hat{H}_m^{\text{eff}} &= \theta_m^x \hat{H}_x + \theta_m^z \hat{H}_z - \frac{i}{2} \theta_m^x \theta_m^z [\hat{H}_x, \hat{H}_z] \\ &\quad + \frac{1}{12} (\theta_m^x)^2 \theta_m^z [\hat{H}_x, [\hat{H}_x, \hat{H}_z]] \\ &\quad - \frac{1}{12} \theta_m^x (\theta_m^z)^2 [\hat{H}_z, [\hat{H}_x, \hat{H}_z]] + \dots . \end{aligned} \quad (2.25)$$

We will truncate the previous equation to the third order — i.e., to the terms shown — and perform exact diagonalization to obtain, at each time step m , the instantaneous eigenvectors of $\hat{H}_m^{\text{eff}}$. The choice of the truncation order is justified by the fact that including higher order in Eq. (2.25) does not significantly affect the values of the observables we computed.

Clearly, each protocol is characterized by a different effective Hamiltonian. The population of the j th instantaneous effective eigenstate $|\phi_m^j\rangle$ is defined as:

$$p_j(m) = |\langle \phi_m^j | \Psi_m \rangle|^2, \quad (2.26)$$

where, for each method, the state of the system at time step m reads:

$$|\Psi_m\rangle = \hat{U}(\theta_m^x, \theta_m^z) \cdots \hat{U}(\theta_1^x, \theta_1^z) |\Psi_0\rangle . \quad (2.27)$$

In Fig. 2.6, we focus on dQA-LIN, and we show the first three instantaneous eigenstate populations $p_j(m)$ for $j = 0, 1, 2$. Evidently, the system attempts to follow an adiabatic dynamics, maintaining a very high population $p_0(m)$ for the

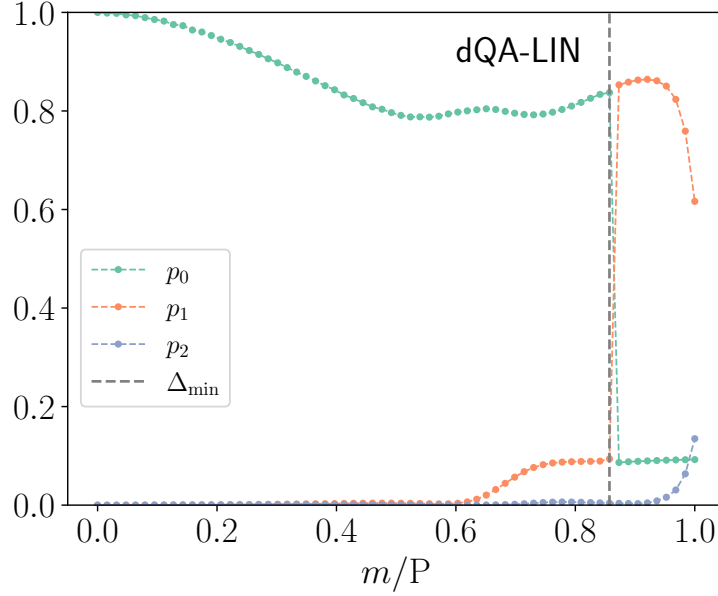


Figure 2.6.: Population of the instantaneous ground state and of the first two excited states as a function of the index m for the instance 1 as obtained using dQA-LIN. The vertical dashed line indicates where $\hat{H}_m^{\text{eff}}$ attains its minimum spectral gap. The optimal time step for dQA-LIN is $\Delta_t^* = 0.78$. The adiabatic protocol implemented by dQA-LIN fails to largely populate the ground state after the minimum gap.

instantaneous ground state until a small instantaneous gap is encountered: here, the system undergoes a population inversion between the instantaneous ground state and the first excited state, which turns out to be the most populated state at the final $m = P$. The four panels of Fig. 2.7 show the same quantities referring to the four methods that provide better solutions for this instance. In these cases, before the system reaches the minimum gap, there is a smooth population inversion between the instantaneous ground state and the first excited state. Afterward, a second population inversion occurs in correspondence with the minimum spectral gap of $\hat{H}_m^{\text{eff}}$, yielding a large final overlap between the state of the system and the target state. Remarkably, the system follows an STA path in order to achieve a large ground state population at the final time step $m = P$. These findings are consistent with the results of [95], obtained using the first-order BCH formula and the QAOA-FOURIER method.

We verified that these four methods implement similar but *quantitatively different* non-adiabatic protocols to solve the MaxCut problem. Due to the large dimension of the parameter space for $P = 64$, one cannot exclude the existence of a path linking the minima in Fig. 2.5 without encountering a barrier in the residual energy landscape. However, in App. A.3, we show that there is at least one direction along which the minima are not pairwise connected. This emphasizes the absence of a single broad minimum basin in the variational energy landscape to which all these optimal parameters belong.

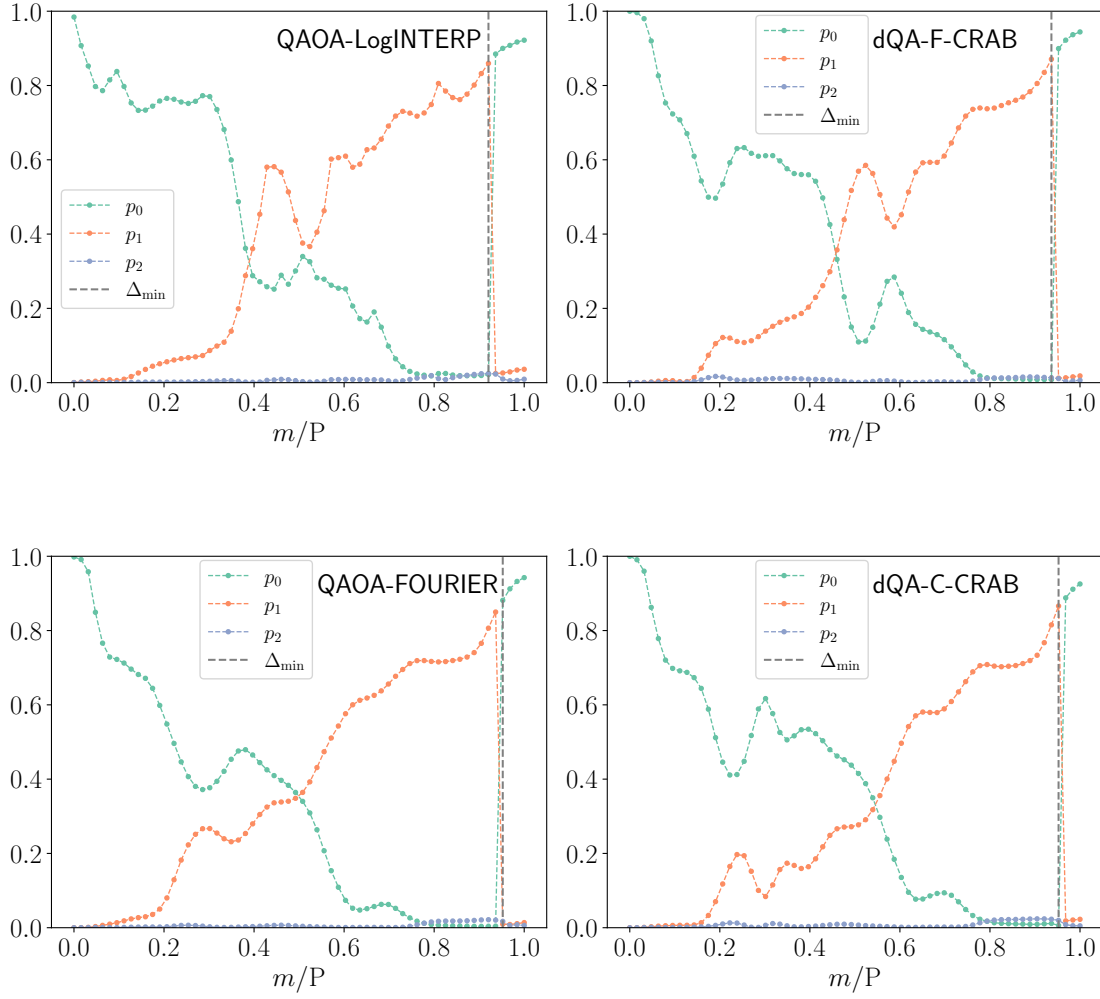


Figure 2.7.: Population of the instantaneous ground state and of the first two excited states as a function of the index m for the instance 1. We consider the same data as in Fig. 2.5. The vertical dashed lines indicate where the corresponding $\hat{H}_m^{\text{eff}}$ attain their minimum spectral gap. In the four panels, each solution implements a different nonadiabatic protocol to reach the target ground state at $m = P$.

We highlight that a similar non-adiabatic mechanism is observed for all the hard instances we examined. Specifically, the optimal parameters derived from methods outperforming dQA-LIN consistently exhibit an STA mechanism to avoid the lack of performance due to population inversion at the minimum gap. Interestingly, the gap of the effective Hamiltonian remains small regardless of the optimized digital dynamics. Rather than seeking an effective Hamiltonian with a large gap, the solution lies in inducing smooth population inversion to achieve a large final overlap with the target state. This trend holds across all the hard instances we considered.

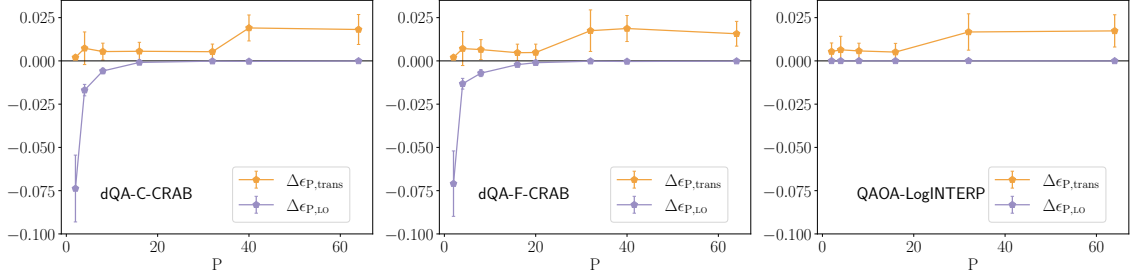


Figure 2.8.: Plot of the mean differences $\Delta\epsilon_{P,\text{trans}}$ and $\Delta\epsilon_{P,\text{LO}}$ as a function of P for three different methods, see the main text for details. We averaged over the 9 hard instances over which the parameters of instance 1 have been transferred. The error bars indicate the standard deviation of the mean.

2.4.3. Transferability among hard instances

In this section, we focus on the transferability of the optimal parameters for dQA-C-CRAB, dQA-F-CRAB, and QAOA-LogINTERP, which appear to be the most effective methods (QAOA-FOURIER turns out to be more computationally expensive).

Numerous studies have explored parameter concentration and transferability for both unweighted [174, 175] and weighted MaxCut [176, 177], often focusing on the regime of small P and large N . Special attention is devoted to typical instances drawn from a given distribution, also addressing problems beyond MaxCut or two-bodies interactions [158, 170, 178, 179]. Here, we focus on the transferability of optimal-control solutions at fixed circuit depth P among hard instances of 3-regular MaxCut with $N = 14$.

To evaluate it, we proceed as follows. First, we compute the residual energy for the 10 hard instances, for fixed P and for each method:

$$E_P^{\text{fin},k} = \langle \Psi_P^{(k)}(\boldsymbol{\theta}^{(k)}) | \hat{H}_z^{(k)} | \Psi_P^{(k)}(\boldsymbol{\theta}^{(k)}) \rangle, \quad (2.28)$$

where $k = 1, \dots, 10$ labels the instance.

Next, we take the optimal parameters $\boldsymbol{\theta}^{(1)}$ of instance $k = 1$, and we compute the variational energy of the instances $k = 2, \dots, 10$:

$$E_{P,\text{trans}}^{\text{fin},k} = \langle \Psi_P^{(k)}(\boldsymbol{\theta}^{(1)}) | \hat{H}_z^{(k)} | \Psi_P^{(k)}(\boldsymbol{\theta}^{(1)}) \rangle, \quad (2.29)$$

where $|\Psi_P^{(k)}(\boldsymbol{\theta}^{(1)})\rangle$ is the variational state for the instance k with the optimal parameters of instance $k = 1$, and $\hat{H}_z^{(k)}$ is the corresponding target Hamiltonian.

Finally, we refine these transferred solutions by local optimization (LO), by considering the parameters $\boldsymbol{\theta}^{(1)}$ as a warm-start for a local optimization of the other instances $k = 2, \dots, 10$, leading to final optimal parameters $\boldsymbol{\theta}_{\text{LO}}^{(k)}$. With these

transferred-and-optimized parameters, we compute the energy:

$$E_{P,LO}^{\text{fin},k} = \langle \Psi_P^{(k)}(\boldsymbol{\theta}_{LO}^{(k)}) | \hat{H}_z^{(k)} | \Psi_P^{(k)}(\boldsymbol{\theta}_{LO}^{(k)}) \rangle. \quad (2.30)$$

Equation (2.22) is then used to calculate the residual energies $\epsilon_P^{\text{res},k}$, $\epsilon_{P,\text{trans}}^{\text{res},k}$ and $\epsilon_{P,LO}^{\text{res},k}$.

We summarize these results in Fig. 2.8, where we show the average differences

$$\begin{aligned} \Delta\epsilon_{P,\text{trans}} &= [\epsilon_{P,\text{trans}}^{\text{res},k} - \epsilon_P^{\text{res},k}]_{\text{av}} \\ \Delta\epsilon_{P,LO} &= [\epsilon_{P,LO}^{\text{res},k} - \epsilon_P^{\text{res},k}]_{\text{av}} \end{aligned} \quad (2.31)$$

as a function of P for dQA-C-CRAB, dQA-F-CRAB, and QAOA-LogINTERP. The average is computed over the different hard instances we considered. We observe that, for small P , a direct transfer of the optimal parameters of instance 1 provides residual energies very close to the optima. On the other hand, at large P , a direct transfer of the optimal parameters of instance 1 does not guarantee a good residual energy $\epsilon_{P,\text{trans}}^{\text{res},k}$. Nonetheless, the transferred parameters turn out to be a good starting point, as witnessed by the fact that $\epsilon_{P,LO}^{\text{res},k}$ is, on average, comparable (or even better, for small P) to $\epsilon_P^{\text{res},k}$, with the extra advantage of being computationally cheaper.

2.4.4. Smoothness of the solutions and continuous-time annealing

In this section, we show the benefits of working with smooth optimal parameters. To draw a comparison, we start by building an irregular set of optimal parameters for instance 1. This is done by implementing dQA-F-CRAB for $P = 64$ and $N_c = P$, without any warm-start for the optimization, i.e., setting $C_0 = 1$ and $\mathbf{C}^x = \mathbf{C}^z = 0$ as initial conditions. The presence of high-frequency modes and the absence of a regularization induced by the iterative warm-start induce sudden variations, as a function of the layer index m , in the optimal parameters θ_m^x and θ_m^z , shown in the top panel of Fig. 2.9. Despite its irregularity, such solution provides a value of residual energy $\epsilon_{\text{IRR}}^{\text{res}} = 0.00190$, very close to that of the smooth solution for $P = 64$ and $N_c = P/2$, $\epsilon_P^{\text{res}} = 0.00183$, shown in the top-right panel of Fig. 2.5. In the digital framework, the smooth and the irregular solutions perform similarly, not only in terms of residual energy but also concerning the instantaneous populations of the first eigenstates. The bottom panel of Fig. 2.9 shows that, after a fast oscillatory phase, the population $p_1(t)$ of the first excited state becomes increasingly larger: the irregular solution performs similar STA dynamics, featuring the same population inversion described in the previous section for the smooth solution.

One of the key findings of our work is that only the smooth solutions can be used to define a valid continuous-time optimal control, which might be applied in

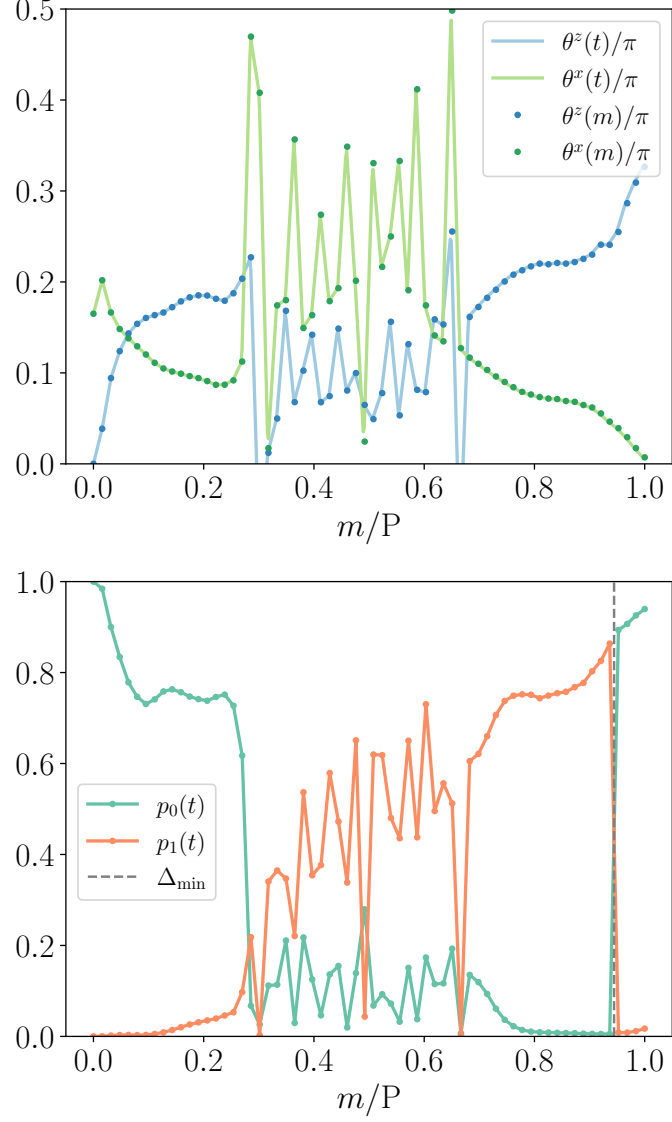


Figure 2.9.: (Top) The irregular optimal parameters for instance 1 obtained by implementing dQA-F-CRAB for $P = 64$ and $N_c = 64$ without any warm start. We also plot the linear interpolation $\theta^x(t)$ and $\theta^z(t)$ utilized in the continuous-time dynamics, for which the x-axis should be intended as t/τ . (Bottom) The discrete populations $p_j(m)$ for the digitized dynamics with irregular parameters.

a realistic experiment with real-time quantum dynamics. This procedure allows us to obtain an optimal QA schedule without the need for prior information on the minimum gap. To show this, we consider the continuous-time evolution driven by the annealing Hamiltonian $\hat{H}(t)$

$$\hat{H}(t) = (1 - s(t)) \hat{H}_x + s(t) \hat{H}_z, \quad (2.32)$$

where $s(t) = \theta^z(t)/(\theta^x(t) + \theta^z(t))$, with $\theta^x(t)$ and $\theta^z(t)$ obtained by interpolating the discrete optimal parameters θ_m^x and θ_m^z . This interpolation is shown in Fig. 2.9(top panel) for the irregular solution.

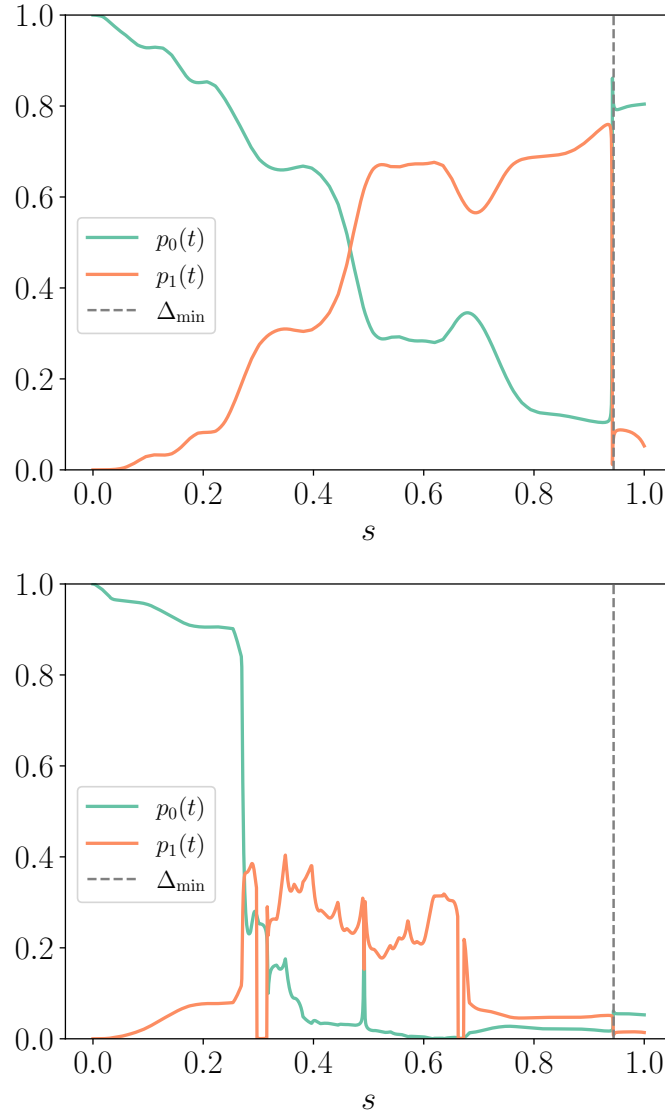


Figure 2.10.: (Top) Ground and first excited state populations $p_j(t)$ for the continuous-time evolution obtained by interpolation of the smooth digital parameters. (Bottom) The corresponding populations for the continuous-time evolution, as obtained by interpolation of the irregular digital parameters.

The state $|\Psi(t)\rangle$ is obtained by solving the Schrödinger equation (2.1). To do so, we efficiently simulate the action of the full evolution operator using a scaling and

squaring method, together with a truncated Taylor series approximation [180], which does not involve any Trotterization. In this continuous-time simulation, we used a sufficiently small integration time step ($dt_c = 0.1$ proved to be small enough) and we considered a total annealing time $\tau = \sum_{m=1}^P (\theta_m^x + \theta_m^z)$ [95]. Finally, we compute the average energy $E^{\text{fin}}(\tau) = \langle \Psi(\tau) | \hat{H}_z | \Psi(\tau) \rangle$, and the corresponding residual energy:

$$\epsilon^{\text{res}}(\tau) = \frac{E^{\text{fin}}(\tau) - E_{\min}}{E_{\max} - E_{\min}}. \quad (2.33)$$

Our simulations show that the continuous-time Schrödinger evolution with $\theta^x(t)$ and $\theta^z(t)$ obtained from the smooth optimal digital schedule θ_m^x and θ_m^z yields a low residual energy, $\epsilon^{\text{res}}(\tau) = 0.0078$. On the contrary, when $\theta^x(t)$ and $\theta^z(t)$ are obtained from the irregular optimal θ_m^x and θ_m^z shown in Fig. 2.9 (top panel), the result is very poor: $\epsilon^{\text{res}}(\tau) = 0.2382$.

This performance difference is also reflected in the population of the instantaneous eigenstates $p_j(t)$, shown in Fig. 2.10. We define the populations in the continuous-time framework as:

$$p_j(t) = |\langle \phi^j(t) | \psi(t) \rangle|^2, \quad (2.34)$$

where $|\phi^j(t)\rangle$ is the j -th eigenvector of $\hat{H}(t)$. For smooth parameters (top panel), despite some minor differences due to the time-discretization and Trotterization, the populations in the continuous-time evolution show qualitatively the same behavior as those in the digital evolution (see Fig. 2.7): we retrieve the STA protocol, which succeeds in largely populating the target state at the end of the time evolution.

On the other hand, the continuous-time evolution induced by the irregular parameters yields very different population results: this can be observed by comparing the bottom panel of Fig. 2.10 to the bottom panel of Fig. 2.9, showing the populations for discrete dynamics. Indeed, in the continuous-time evolution, the population of the low-energy spectrum drastically decreases after the irregular phase, and other excited states are significantly populated. As a consequence, the population inversion occurring at the minimum gap does not effectively populate the ground state.

We conclude that sharp variations in the time-dependent Hamiltonian controls affect the continuous-time dynamics and result in a drastic loss of performance compared to the digital evolution. On the other hand, optimal digital smooth solutions are suitable to approach the problem also in continuous-time, and can possibly be refined by local optimization in continuous time.

2.5. Conclusions

In this Chapter, we have presented a set of methods for constructing optimal Quantum Annealing (QA) schedules for hard MaxCut problems where a small spectral gap is a bottleneck to standard adiabatic dynamics.

The methods utilized fall into two distinct categories. The first consists of constructing a Trotter-digitized QA (dQA) with schedules parameterized in terms of Fourier modes or Chebyshev polynomials. These parameterizations are adapted from the Chopped Random Basis (CRAB) technique, which is employed for optimal control in continuous-time systems. The second category is based on the Quantum Approximate Optimization Algorithm (QAOA) [25], where solutions are found iteratively using linear interpolations (INTERP [95] or LogINTERP [49]) or expansions in Fourier modes (FOURIER [95]). In both cases, we have emphasized the importance of smooth optimal parameters and found that the same physical mechanism is behind the methods that perform better: STA-like dynamics, as discussed in Section 1.2.3, with a population inversion occurring well before the minimum spectral gap is encountered.

Summarizing the performance of the various methods on the hard instances we considered, we have found that QAOA-INTERP, based on iteratively increasing the layer depth at each step as $P \rightarrow P' = P + 1$, is not able to significantly outperform a linear-schedule dQA (dQA-LIN). On the contrary, its variant QAOA-LogINTERP, based on duplicating the layer depth at each step, $P \rightarrow P' = 2P$, is among the best-performing methods. QAOA-FOURIER [95] also shows comparable performance, but its implementation is rather time-consuming. Both digitized CRAB approaches, namely dQA-F-CRAB and dQA-C-CRAB, perform equally well. The latter, in particular, requires the least computational resources to attain smooth optimal solutions: for given P and N_c , dQA-C-CRAB finds optimal and smooth solutions using a discrete linear schedule as the initial condition without any iterative warm start.

A noteworthy feature of the smooth optimal solutions is their transferability among hard instances of MaxCut at fixed circuit depth P . Indeed, we find that the solutions obtained for a given instance perform rather well, even without any further local optimization, on other hard 3-regular instances with $N = 14$ vertices. If further local optimization is performed, the transferred solution proves even better than what would be obtained by directly optimizing the variational parameters for the target instance.

Finally, we find that smooth optimal protocols are a good starting point for a time-continuous schedule to be implemented in an analog device. Remarkably, this provides an optimized QA protocol that improves on the standard linear schedule, without requiring preliminary information on the instantaneous gap.

One natural continuation is to apply quantum optimal control to one of the simplest models featuring an exponentially small spectral gap: the frustrated Ising ring introduced in Ref. [80] and further explored in Ref. [140]. Motivated by this direction, Chapter 3 studies optimal-control protocols for this model.

The second direction is to further explore the connection between our digitized

solutions and the STA-like mechanism associated with continuous-time counterdiabatic driving, reviewed in Section 1.2.3 and Ref. [84]. Recent work [87] has pointed out that the first-order correction in the Baker–Campbell–Hausdorff formula has exactly the same structure and sign as an adiabatic-gauge-potential term realizing counterdiabatic driving; clarifying this connection between digitized and continuous-time dynamics remains an open direction. A subsequent study by Thengil *et al.* [181] explores this direction by systematically investigating local counterdiabatic terms in digitized protocols for the frustrated Ising ring that will be introduced in Chapter 3.

3. Optimal Control of a Frustrated Ising Ring Model

Having seen in Chapter 2 how small spectral gaps can bottleneck quantum annealing (QA) for hard MaxCut instances, we now examine this bottleneck more closely in a simple spin-glass setting. One of the simplest models exhibiting exponentially small spectral gaps is a ferromagnetic Ising ring with a single antiferromagnetic bond introducing frustration [80]. Previous studies of this model have explored continuous-time diabatic QA, where optimized non-adiabatic annealing schedules provided good solutions, avoiding exponentially large annealing times [140]. In our work, we move to a digital framework of Variational Quantum Algorithms, and present two main results:

- (1) We show that the model is *digitally controllable* with a scaling of resources that grows *quadratically* with the system size, achieving the exact solution using the Quantum Approximate Optimization Algorithm (QAOA), as shown in Section 3.4.1;
- (2) We combine a technique of quantum control — the Chopped RANdom Basis (CRAB) method — and digitized quantum annealing (dQA) to construct *smooth* digital schedules yielding optimal solutions with very high accuracy, as presented in Section 3.4.2.

3.1. Introduction

The QA framework and its small-gap bottleneck are reviewed in Sections 1.2.1 and 1.2.3, while dQA, QAOA, and their shared alternating-Hamiltonian Ansatz are introduced in Section 1.2.4 of Chapter 1.

In both of these digitized approaches, the threshold for the state controllability of the model [110, 182] — defined in this framework as the minimum number of unitary transformations required to reach the exact solution — plays a fundamental role. It is essential to evaluate algorithm efficiency and to analyze resource scaling with system size [183, 184]. For a generic model with an exponentially large Hilbert space, this quantity is very difficult to estimate. However, for specific models, symmetries

can be exploited to determine the critical depth required to reach the target state [49, 185].

In this work, we present results, using the alternating *Ansatz* in Eq. (1.41), on the frustrated-ring model proposed in [80], see also [140, 186], as the simplest problem exhibiting exponentially small spectral gaps as the number of spins increases. The model is a non-uniform Ising ring, where all the couplings except one are ferromagnetic, with a single antiferromagnetic coupling introducing frustration.

We first use QAOA to characterize the controllability properties of this system in terms of the number of layers of the quantum circuit needed to exactly reach the desired ground state. Operatively, we determine the critical depth P_1^{cr} as the minimum number of layers such that, among a large number N_{samples} of random initialization of the QAOA parameters, there is at least one set of optimized QAOA angles θ^{opt} that reaches the ground state. (In practice, $N_{\text{samples}} = 250$ in our work.)

Interestingly, we find that P_1^{cr} scales quadratically with the number of spins, $P_1^{\text{cr}} = (N^2 - 1)/4$, completely bypassing the exponential bottleneck associated with the standard QA solution of the model.

Next, we pursue the goal of finding smooth digital schedules for the parameters θ in Eq. (1.41), which optimize the average energy $E_P(\theta) = \langle \Psi_P(\theta) | \hat{H}_z | \Psi_P(\theta) \rangle$. A well-established quantum control technique, previously proposed and developed for continuous-time dynamics, is the Chopped RAndom Basis (CRAB) [105, 106, 108, 109]. In the context of QA, CRAB is used to decompose the annealing schedules $s(t)$ into a finite set of smooth basis functions, such as Fourier modes, turning the functional minimization problem into a finite-dimensional one. We adopted this technique — in its dressed CRAB (dCRAB) [108] flavour — in a digitized QA dynamics framework. We show how combining CRAB with dQA produces smooth digital schedules θ yielding results that outperform a randomly initialized QAOA in the regime $P < P_1^{\text{cr}}$; the approach provides high-accuracy results also in the controllable regime $P > P_1^{\text{cr}}$. We then characterize the digital dynamics generated by such smooth solutions in terms of ground state populations and shortcut-to-adiabaticity, unveiling the non-adiabatic nature of the optimal strategy implemented by such dQA-CRAB approach. Finally, we introduce an effective annealing time for the smooth digital dQA-CRAB protocol and compare our results with diabatic quantum annealing in continuous time [140]. We find that, for the same level of accuracy, our method exhibits better scaling with the system size.

The Chapter is organized as follows. Section 3.2 introduces the frustrated-ring model. Section 3.3 presents our methodology: the CRAB approach to digital schedule control in dQA (dubbed dQA-CRAB), 3.3.1, and QAOA from random starting initial points (dubbed QAOA-RAND), 3.3.2. Next, in Sec. 3.4, we present our results. We start with our evidence of digital controllability through QAOA-RAND, 3.4.1, and then move to present the digitally smooth schedules produced by dQA-

CRAB, 3.4.2. Next, we discuss the shortcut-to-adiabaticity [129, 187] mechanism behind the dQA-CRAB protocol, 3.4.3, and its superiority with respect to the competing continuous-time diabatic QA approach of Ref. [140], 3.4.4. Finally, Section 3.5 summarizes our key findings and discusses future developments.

3.2. The frustrated-ring model

Consider an *odd* number of spins N arranged in a ring, as shown in Fig. 3.1. This is the frustrated-ring model proposed in [80] and recently discussed in [140]. The problem Hamiltonian reads:

$$\hat{H}_{\text{targ}} = \hat{H}_z = - \sum_{j=1}^N J_j \hat{\sigma}_j^z \hat{\sigma}_{j+1}^z . \quad (3.1)$$

The couplings J_j between adjacent spins are:

$$J_j = \begin{cases} J_w & \text{if } j = (N \pm 1)/2 \\ -J_f & \text{if } j = N \\ J & \text{otherwise} \end{cases} . \quad (3.2)$$

Here, J is the ferromagnetic coupling between most sites, J_w denotes the weaker ferromagnetic coupling between the two central sites at $j = (N \pm 1)/2$, and J_f is the single frustrating (anti-ferromagnetic) coupling at site N . (The notation of Refs. [80] and [140] have $J_f = J_R$ and $J_w = J_L$.) We choose $0 < J_f < J_w < J$, specifically requiring that $JJ_f > J_w^2$, which leads to the spin-glass bottleneck regime studied in Ref. [80]. For the rest of this Chapter we take, following Ref. [140], $J = 1$, our unit of energy, $J_w = 0.5$ and $J_f = 0.45$. The ground state is the ferromagnetic state with energy $E_{\text{gs}} = -(N - 3)J - 2J_w + J_f$; the first excited state has two domain walls, at a central site $j = (N + 1)/2$ and at $j = N$, and is separated by an energy gap: $E_{\text{ex}} - E_{\text{gs}} \equiv \Delta_1 = 2(J_w - J_f)$.

We write the driving term in the annealing Hamiltonian of Eq. (1.6) as:

$$\hat{H}_{\text{drive}} = \hat{H}_x = -h \sum_{j=1}^N \hat{\sigma}_j^x . \quad (3.3)$$

We apply the standard Jordan–Wigner transformation [78] to express $\hat{H}_z$ and $\hat{H}_x$ as quadratic fermionic Hamiltonians:

$$\hat{H}_z = - \sum_{j=1}^N J_j \left(\hat{c}_j^\dagger \hat{c}_{j+1} + \hat{c}_j^\dagger \hat{c}_{j+1}^\dagger + \text{H.c.} \right) , \quad (3.4)$$

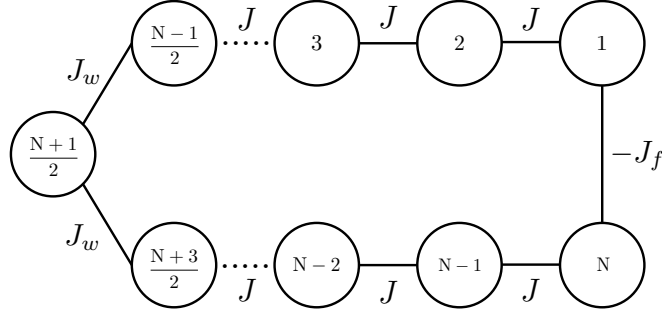


Figure 3.1.: A schematic diagram for the frustrated-ring model with an odd number of sites N . Most of the couplings between adjacent spins are uniform and ferromagnetic, J , except for the two central sites, coupled by weaker ferromagnetic bonds, J_w , and the first and last spins which are coupled by an antiferromagnetic (frustrating) bond $-J_f$.

with boundary condition $\hat{c}_{N+1} = (-1)^{p+1} \hat{c}_1$, and

$$\hat{H}_x = h \sum_{j=1}^N \left(\hat{c}_j^\dagger \hat{c}_j - \hat{c}_j \hat{c}_j^\dagger \right), \quad (3.5)$$

where $\hat{c}_j^\dagger$ and $\hat{c}_j$ are fermionic creation and annihilation operators. Here, $p = 0$ or 1 denotes even or odd fermion parity sectors of the Hilbert space, $(-1)^p = e^{i\pi \hat{N}_F}$, where $\hat{N}_F = \sum_j \hat{c}_j^\dagger \hat{c}_j$ is the total number of fermions. Notice that the ground state of $\hat{H}_x$ is the fermionic vacuum $|0\rangle$ (which has $p = 0$) for $h > 0$, and the fully occupied fermionic state $|F\rangle = \prod_j \hat{c}_j^\dagger |0\rangle$ for $h < 0$: this has $p = 1$ when N is odd. Throughout this Chapter we assume that $h < 0$, setting it to $h = -J$, and restrict ourselves to the sub-sector with $p = 1$, where the ground state of the interpolating Hamiltonian in Eq. (1.6) stays and the whole dynamics takes place.

Using the Nambu formalism [78], the fermionic Hamiltonians $\hat{H}_z$ and $\hat{H}_x$ are cast into matrix form:

$$\hat{H}_{x/z} = \hat{\Psi}^\dagger \mathbb{H}_{x/z} \hat{\Psi}, \quad (3.6)$$

where $\mathbb{H}_{x/z}$ are $2N \times 2N$ Hermitian matrices and $\hat{\Psi} = (\hat{c}_1, \dots, \hat{c}_N, \hat{c}_1^\dagger, \dots, \hat{c}_N^\dagger)^\top$, is a $2N$ -dimensional column vector combining fermionic annihilation operators $\hat{c}_1, \dots, \hat{c}_N$ and creation operators $\hat{c}_1^\dagger, \dots, \hat{c}_N^\dagger$. For the transverse field part we have:

$$\mathbb{H}_x = \begin{pmatrix} \mathbf{A}_x & \mathbf{0} \\ \mathbf{0} & -\mathbf{A}_x \end{pmatrix}, \quad (3.7)$$

where $\mathbf{A}_x$ is purely diagonal, here proportional to the identity since the transverse field is uniform:

$$\mathbf{A}_x = h \mathbf{1} = -\mathbf{1}, \quad (3.8)$$

where we set $h = -J = -1$.

For the target term, with $p = 1$, we have:

$$\mathbb{H}_z = \begin{pmatrix} \mathbf{A}_z & \mathbf{B}_z \\ -\mathbf{B}_z & -\mathbf{A}_z \end{pmatrix}, \quad (3.9)$$

where $\mathbf{A}_z$ and $\mathbf{B}_z$ are $N \times N$ real symmetric or anti-symmetric matrices, respectively, and their only non-zero elements are given by:

$$\begin{cases} (\mathbf{A}_z)_{j,j+1} = (\mathbf{A}_z)_{j+1,j} = -J_j/2 \\ (\mathbf{B}_z)_{j,j+1} = -(\mathbf{B}_z)_{j+1,j} = -J_j/2 \end{cases}, \quad (3.10)$$

with additional matrix elements, dependent on the fermionic parity, and given by:

$$(\mathbf{A}_z)_{N,1} = (\mathbf{A}_z)_{1,N} = (-1)^p J_N/2 = +J_f/2 \quad (3.11)$$

and

$$(\mathbf{B}_z)_{N,1} = -(\mathbf{B}_z)_{1,N} = (-1)^p J_N/2 = +J_f/2. \quad (3.12)$$

Figure 3.2 shows the spectral gap Δ_s between the ground state and 1st excited state of the interpolating Hamiltonian $s\hat{H}_z + (1-s)\hat{H}_x$ in the $p = 1$ subsector, for the frustrated-ring model with $N = 101$ sites. Notice that there are two small gaps. The first gap occurs at $s_c \approx 0.5$ and decays polynomially, as $1/N$: it is the critical gap separating the quantum paramagnetic phase ($s < s_c$) from the frustrated ferromagnetic phase ($s > s_c$). The second, at $s_b \approx 0.9$, is the exponentially small gap providing the spin-glass bottleneck discussed in Ref. [80]. It is predicted [80] to be located at:

$$s_b = \left(1 + \frac{(J^2 - J_w^2)(J_w^2 - J_f^2)}{J J_f (J^2 + J_f^2 - 2J_w^2)} \right)^{-1} \approx 0.89872068, \quad (3.13)$$

(the numerical value corresponds to $J_w/J = 0.5$ and $J_f/J = 0.45$), and to decay exponentially in N , as:

$$\Delta_{s_b} \underset{N \rightarrow \infty}{\propto} \left[\frac{J(J_w^2 - J_f^2)}{J_f(J^2 - J_w^2)} \right]^N. \quad (3.14)$$

3.3. Methodology

The continuous-time optimal-schedule problem was set up in Section 1.2.1 and, for the MaxCut case, in Section 2.2.1: the interpolating Hamiltonian $\hat{H}(t) = s(t)\hat{H}_z + (1-s(t))\hat{H}_x$ of Eq. (1.6), with $s(0) = 0$ and $s(\tau) = 1$, drives a unitary Schrödinger dynamics starting from the ground state of $\hat{H}_x$, and one seeks the schedule $s(t)$ that minimizes the final average energy $E^{\text{fin}}(\tau) = \langle \Psi(\tau) | \hat{H}_z | \Psi(\tau) \rangle$ at a given annealing time τ : a functional optimization problem.

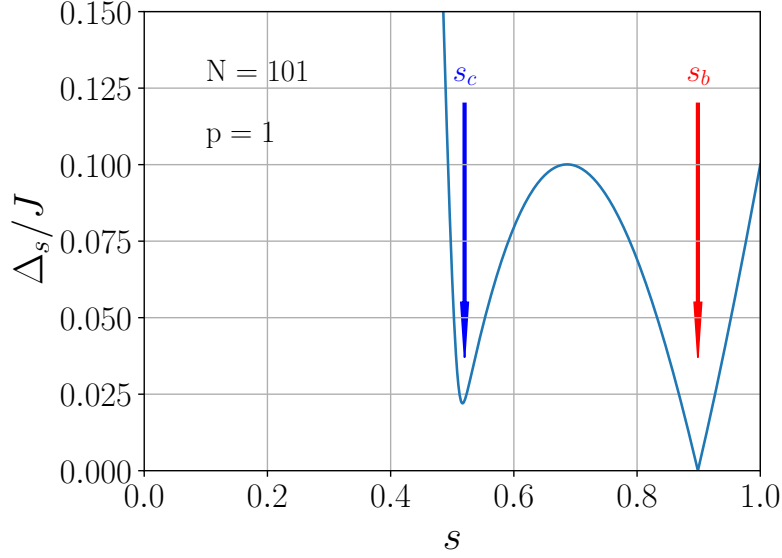


Figure 3.2.: The spectral gap Δ_s between ground state and first excited state of a frustrated-ring model of size $N = 101$ in the $p = 1$ fermionic sub-sector, as a function of the interpolation parameter s .

The continuous-time approach by Côté *et al.* [140] parametrizes the annealing schedule $s(t)$ as a piecewise linear decomposition with a finite number of parameters. The parametrization starts from randomly selecting $P - 1$ values s_p , in the interval $[0, 1)$, with $p = 1 \cdots P - 1$, which are assigned to $P - 1$ equally spaced times $t_p = p\Delta_t = p\tau/P$ in the time interval $[0, \tau]$: $s_p \equiv s(t_p)$. The initial schedule $s(t)$ is then the piecewise linear function $s^{\text{p-lin}}(t)$ obtained by connecting these $P - 1$ points (t_p, s_p) with the two endpoints $(t_0 = 0, s(0) = 0)$ and $(t_P = \tau, s(\tau) = 1)$. Different sets of random parameters $\{s_p\}$ are optimized with respect to the energy $E^{\text{fin}}(\tau)$ of the output state $|\Psi(\tau)\rangle$, and the set that produces the lowest energy is selected as a candidate schedule $s^{\text{p-lin}}(t)$. In subsequent iterations one doubles $P \rightarrow 2P$ by adding intermediate points in all linear segments of $s^{\text{p-lin}}(t)$, and then optimizing the $2P - 1$ internal points s_p to obtain a new piecewise-linear $s^{\text{p-lin}}(t)$. In practice, starting from $P = 4$ and trying $N_{\text{real}} = 10$ realizations of $\{s_p\}$, one takes the best $s^{\text{p-lin}}(t)$, and then moves to $2P = 8$ with $2P - 1 = 7$ internal points to be optimized; and so forth, until the final energy $E^{\text{fin}}(\tau)$ associated with $s^{\text{p-lin}}(t)$ differs from the actual ground state E_{gs} by less than a specific threshold Δ_E . For further details on this algorithm we refer the reader to the flow diagram (Figure 2) and the pseudocode (Algorithm 1) provided by Côté *et al.* [140]. (For the reader's convenience, we mention here some modifications we adopted in the notation: in Algorithm 1 of [140], $k_0 = P - 1$, $T = \tau$, $a = s$, and $A = s^{\text{p-lin}}$.)

3.3.1. Digitized QA dynamics with CRAB

Let us now move towards a digitized QA dynamics [46, 49, 73, 168], as introduced in Section 1.2.4. Suppose first that the annealing time τ is divided into a large number P of small time steps (not necessarily identical) Δ_p , with $p = 1 \cdots P$. The evolution operator associated to such time step is approximately given by (setting $\hbar = 1$):

$$e^{-i\Delta_p(s_p\hat{H}_z+(1-s_p)\hat{H}_x)} \xrightarrow{\text{Trotter}} e^{-i\theta_p^x\hat{H}_x}e^{-i\theta_p^z\hat{H}_z} ,$$

where $s_p = s(t_p)$ is the value of the schedule parameter at an appropriate point t_p inside the time-interval (for instance, the middle point), and the $\theta_p^{x,z}$ are given, in a lowest order Trotter decomposition, as:

$$\theta_p^x = \Delta_p(1 - s_p) , \quad \theta_p^z = \Delta_p s_p . \quad (3.15)$$

Inverting these relationships we get:

$$\Delta_p = \theta_p^x + \theta_p^z , \quad s_p = \frac{\theta_p^z}{\theta_p^x + \theta_p^z} . \quad (3.16)$$

The final state reached by applying the evolution operators for all time steps would then be given by:

$$|\Psi_P^{\text{dQA}}\rangle = \hat{U}(\theta_P^x, \theta_P^z) \cdots \hat{U}(\theta_1^x, \theta_1^z) |\Psi_0\rangle , \quad (3.17)$$

where

$$\hat{U}(\theta_p^x, \theta_p^z) = e^{-i\theta_p^x\hat{H}_x}e^{-i\theta_p^z\hat{H}_z} . \quad (3.18)$$

Details on how such digital dynamics is implemented in the Bogoliubov-de Gennes (BdG) formalism [78] for the frustrated-ring model after a Jordan-Wigner transformation are given in Appendix B.2.

The standard linear schedule $s(t) = t/\tau$ in a digitized-QA setting, first discussed in [49, 90, 168] and later considered in [156], would lead to

$$\theta_p^{x,0} = \Delta_t \left(1 - \frac{t_p}{\tau}\right) , \quad \theta_p^{z,0} = \Delta_t \frac{t_p}{\tau} , \quad (3.19)$$

where we set the time-grid $t_p = (p - \frac{1}{2})\Delta_t$, for $p = 1, \dots, P$, while $\tau = P\Delta_t$. By borrowing a well-established quantum control technique, the CRAB [105, 106, 108, 109] previously proposed and developed in continuous-time dynamics we might write an improved *Ansatz* for the parameters by adding a finite number N_c of Fourier modes over the previous linear schedule

$$\begin{aligned} \theta_p^{x,1} &= C_0^x \theta_p^{x,0} + \left(1 - \frac{t_p}{\tau}\right) \sum_{n=1}^{N_c} C_n^x \sin(\omega_n^{(1)} t_p) \\ \theta_p^{z,1} &= C_0^z \theta_p^{z,0} + \frac{t_p}{\tau} \sum_{n=1}^{N_c} C_n^z \sin(\omega_n^{(1)} t_p) \end{aligned} , \quad (3.20)$$

3. Optimal Control of a Frustrated Ising Ring Model

where $t_p/\tau = (p - \frac{1}{2})/P$, and we incorporate Δ_t in the free coefficients C_0^x and C_0^z . The $2N_c + 2$ variational parameters $\mathbf{C} = (C_0^x, C_0^z, \mathbf{C}^x, \mathbf{C}^z)$, with $\mathbf{C}^{x/z} = (C_1^{x/z}, \dots, C_{N_c}^{x/z})$, should be optimized to minimize the average energy

$$E_P(\mathbf{C}) = \langle \Psi_P^{\text{dQA}} | \hat{H}_z | \Psi_P^{\text{dQA}} \rangle. \quad (3.21)$$

The choice of how large N_c should be and how to choose the N_c frequencies $\omega_n^{(1)}$ is a matter that we addressed by numerical experiments. We came out with a good recipe, which is to set $N_c = P$, and to generate $x = \omega_n^{(1)}\tau/\pi$ from a Gamma probability distribution

$$\Gamma(x; \alpha, \beta) = \begin{cases} \frac{x^{\alpha-1} e^{-x/\beta}}{\beta^\alpha \Gamma(\alpha)} & x \geq 0 \\ 0 & x < 0 \end{cases}, \quad (3.22)$$

where $\alpha > 0$ is the shape parameter, $\beta > 0$ is the scale parameter, and $\Gamma(\alpha)$ the Gamma function. We took $\alpha = 3/2$ and $\beta = 4$, so that the probability distribution is peaked at $x_{\text{max}} = (\alpha - 1)\beta = 2$.

An important advantage of the digitized dynamics, over the continuous-time one, is that **gradients** of the energy $E_P(\mathbf{C})$ can be easily calculated analytically, see Appendix B.4, and given to the Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm [188] which we use to optimize the average energy.

The CRAB procedure can be iteratively improved by adopting a **dressed-CRAB** (dCRAB) approach [108], which adds new variational directions in the description of the control parameters:

$$\begin{aligned} \theta_p^{x,2} &= C_0^{x,2} \theta_p^{x,1} + \left(1 - \frac{t_p}{\tau}\right) \sum_{n=1}^{N_c} C_n^{x,2} \sin(\omega_n^{(2)} t_p) \\ \theta_p^{z,2} &= C_0^{z,2} \theta_p^{z,1} + \frac{t_p}{\tau} \sum_{n=1}^{N_c} C_n^{z,2} \sin(\omega_n^{(2)} t_p) \end{aligned}, \quad (3.23)$$

with $(\theta^{x,1}, \theta^{z,1})$ fixed from the previous CRAB optimization stage, and $2N_c + 2$ new variational parameters $\mathbf{C} = (C_0^{x,2}, C_0^{z,2}, \mathbf{C}^{x,2}, \mathbf{C}^{z,2})$, to be optimized so as to minimize $E_P(\mathbf{C})$ in equation (3.21). The new fixed frequencies $\omega_n^{(2)}$ appearing in equation (3.23), it turns out, are crucial in adding larger frequency components to the control fields: in this second dCRAB iteration we generate $x = \omega_n^{(2)}\tau/\pi$ from $\Gamma(x, \alpha = \frac{3}{2}, \beta = 20)$, which is peaked at $x_{\text{max}} = 10$. Figure 3.3 shows the two Gamma distributions from which we generate the frequencies for the two dCRAB iterations: $\Gamma(x, \frac{3}{2}, 4)$ for $\omega_n^{(1)}\tau/\pi$ in equation (3.20), and $\Gamma(x, \frac{3}{2}, 20)$ for $\omega_n^{(2)}\tau/\pi$ in equation (3.23). The algorithm described above will be referred to as dQA-CRAB.

For comparison, we also include the dQA-Linear protocol [49, 90, 156, 168], in which the linearity of the usual QA schedule is preserved in digitized QA. Specifically, we optimize the parameter Δ_t in the linear *Ansatz* described in equation (3.19), and adopt the resulting parameters for the digitized evolution.

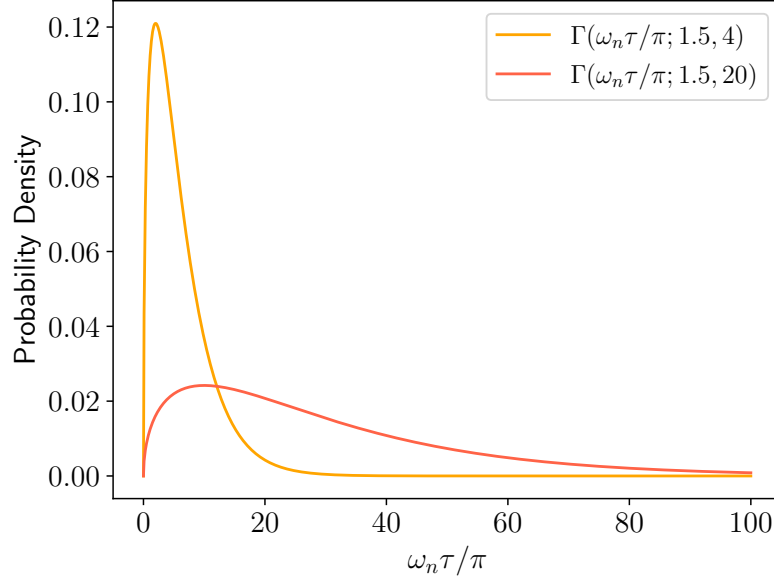


Figure 3.3.: The two gamma distributions from which we generate the frequencies for the two dCRAB iterations.

3.3.2. QAOA

The QAOA [25], introduced in Section 1.2.4, is a VQA designed to solve combinatorial optimization problems by approximating the ground state of the problem Hamiltonian. It prepares the final quantum state using a parametrized sequence of 2P alternating quantum gates which is identical to that appearing in equation (3.17), i.e.,

$$|\Psi_P^{\text{QAOA}}(\boldsymbol{\theta})\rangle = \hat{U}(\theta_P^x, \theta_P^z) \cdots \hat{U}(\theta_1^x, \theta_1^z) |\Psi_0\rangle, \quad (3.24)$$

where

$$\hat{U}(\theta_p^x, \theta_p^z) = e^{-i\theta_p^x \hat{H}_x} e^{-i\theta_p^z \hat{H}_z}, \quad (3.25)$$

with $p = 1, \dots, P$. The conceptual difference with respect to equation (3.17) is that here the $\boldsymbol{\theta} = (\boldsymbol{\theta}^x, \boldsymbol{\theta}^z)$, where $\boldsymbol{\theta}^x = (\theta_1^x \cdots \theta_P^x)$ and $\boldsymbol{\theta}^z = (\theta_1^z \cdots \theta_P^z)$ are P-dimensional vectors, are supposed to be *free* variational parameters with respect to which we should minimize the average energy:

$$E_P^{\text{QAOA}}(\boldsymbol{\theta}) = \langle \Psi_P^{\text{QAOA}}(\boldsymbol{\theta}) | \hat{H}_z | \Psi_P^{\text{QAOA}}(\boldsymbol{\theta}) \rangle. \quad (3.26)$$

The minimization of the cost function is, in general, non-trivial due to the complexity of the 2P-dimensional search space. Many strategies have been discussed in the literature: for a brief discussion of them, see [46]. Here, we present results obtained on the frustrated-ring model by initializing the optimization process with *randomly* generated starting points, employing the BFGS algorithm [188] with analytically calculated gradients to achieve efficient convergence toward a local or global minimum. We will call this algorithm QAOA-RAND.

3.4. Results

To assess the quality of the solutions obtained, we will use, as a figure of merit, a rescaled residual energy:

$$\epsilon_P^{\text{res}} = \frac{E_P - E_{\text{gs}}}{N}, \quad (3.27)$$

where E_P is the optimal solution obtained by either the dQA-CRAB/dQA-Linear, Eq. (3.21), or the QAOA-RAND, Eq. (1.42), approach, and $E_{\text{gs}} = -(N - 3)J - 2J_w + J_f$ is the ground state energy of the frustrated ring model, with N sites.

3.4.1. Digital controllability through QAOA-RAND

We start presenting evidence of a digital controllability of the ground state of the model, by employing QAOA-RAND. Figure (3.4) (a) - (c) shows histograms of the residual energy for a ring of $N = 13$ and $P = 41, 42, 45$: each histogram comprises $N_{\text{samples}} = 250$ QAOA optimization from random starting points (QAOA-RAND). We observe that $P = 42$ is the smallest P in which a (numerically) zero residual energy, that is, the exact ground state, could be reached. We denote as P_1^{cr} this critical P at which full control of the system's state is achieved. For $N = 13$, we find that $P_1^{\text{cr}} = 42$. As P increases, the percentage of simulations achieving zero residual energies increases. At $P = 45$, any random starting parameters in QAOA would ensure that the exact ground state is reached. We call P_2^{cr} the value of P at which the QAOA landscape appears to be simple enough that any random initial point leads to the solution.

The same pattern appears for larger system sizes N . Numerically, we find that P_1^{cr} grows **quadratically** with the system size N , in stark contrast with the exponentially closing gap Δ_{s_b} , which implies exponentially long annealing times for a linear schedule QA. As numerically shown in Fig. (3.4)(d) (see final discussion for a rationalization of this behavior) the expression for P_1^{cr} in terms of N appears to be:

$$P_1^{\text{cr}} = \frac{N^2 - 1}{4}. \quad (3.28)$$

A formal proof for the value of the critical depth P_1^{cr} in the frustrated-ring model will be presented in Ref. [125]. It suffices here to notice that the Bogoliubov-de Gennes theory [78] behind the exact Jordan-Wigner mapping of this and related quantum Ising chain models predicts that the dynamics (both in discrete and continuous time) is fully encoded into a $2N \times 2N$ unitary transformation: hence, a scaling with N^2 is to be expected. The precise value $P_1^{\text{cr}} = (N^2 - 1)/4$ follows from the additional reflection symmetry that the frustrated-ring model Hamiltonian possesses [125]. On the other hand, precisely determining the scaling of P_2^{cr} results to be more complex. When the circuit depth reaches P_2^{cr} , every local minima of the average energy becomes degenerate with the global minimum, leading to a much less complex optimization

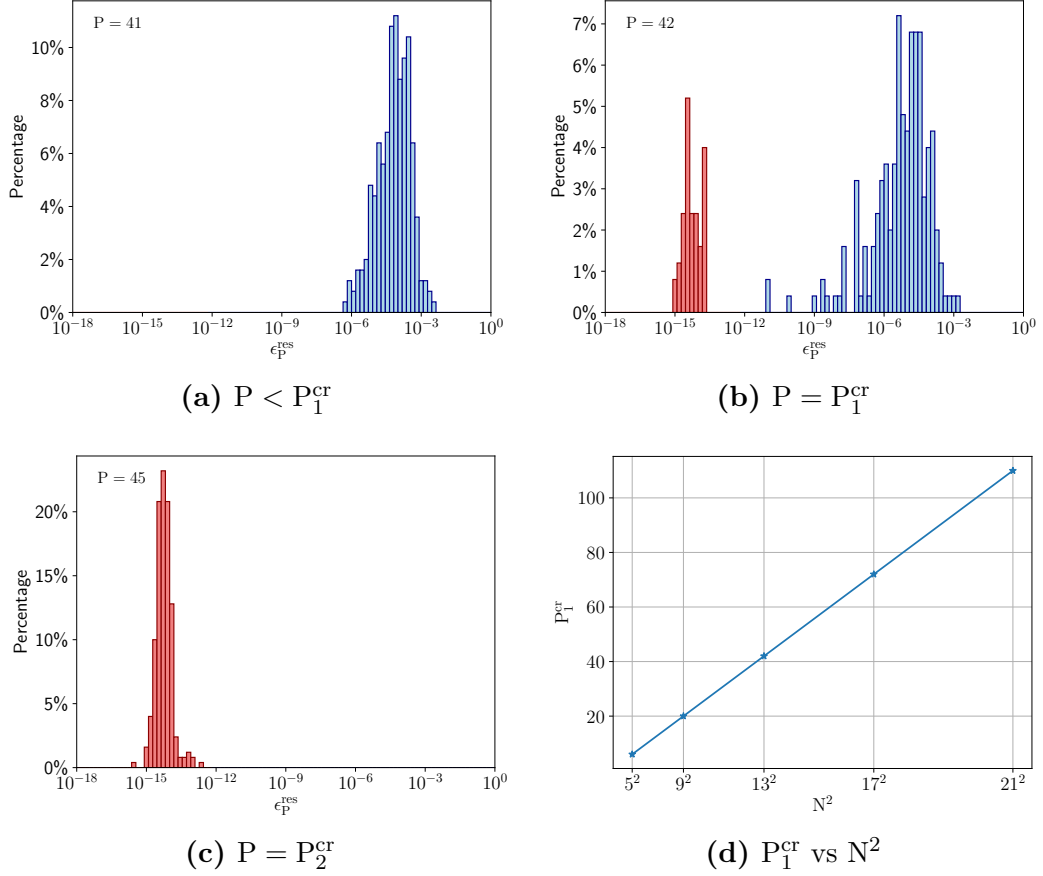


Figure 3.4.: Numerical results of QAOA following optimizations from random initial points, referred to as QAOA-RAND. (a) – (c): empirical distributions of optimized residual energies ϵ_P^{res} , as defined in Eq. (3.27), for $N = 13$ and different layers $P = 41, 42, 45$. Each distribution is constructed from $N_{\text{samples}} = 250$ independent QAOA runs with random initial points, leading to a spread of residual energy values after optimization. Residual energies $\epsilon_P^{\text{res}} < 10^{-12}$ — an introduced numerical threshold — are highlighted in red and are considered numerically zero. The critical layers P_1^{cr} and P_2^{cr} , defined in Section 3.4.1, are here met at $P = 42$ (panel (b)) and $P = 45$ (panel (c)), respectively. (d): Numerically determined quadratic scaling of P_1^{cr} with system size N (further analytical insights are discussed in the main text).

landscape. Characterizing this process is challenging because, unlike P_1^{cr} , it is not strictly related to the number of parameters required to describe the dynamics or to the state controllability [110, 112]. Furthermore, any numerical claim about the scaling of P_2^{cr} depends on the empirical, but inherently arbitrary, definition of a “zero residual energy”, which we set here to 10^{-12} .

We stress that the finite variance of the distributions of energies below the threshold we set as numerical zero stems from numerical errors of the optimization routines. This justifies our decision to consider protocols associated to residual energy below the threshold of 10^{-12} as leading to the exact ground state.

3.4.2. dQA-CRAB smooth optimal schedules

In analogy with the Trotter-discretized expressions in Eq. (1.40), we define, for a general digitized algorithm, even away from the Trotter limit, the total evolution time τ in terms of Δ_p as:

$$\tau = \sum_{p=1}^P \Delta_p = \sum_{p=1}^P (\theta_p^x + \theta_p^z), \quad (3.29)$$

where $p = 1, \dots, P$ labels the sequence of time steps.

We show in Fig. 3.5 the schedules $\{s_p\}$ defined in Eq. (1.40) and produced using QAOA-RAND and dQA-CRAB on the frustrated-ring model with $N = 13$ and $P = 56$. As expected dQA-CRAB produces smooth schedules, while QAOA-RAND is agnostic to any regularity. Observe that the dQA-CRAB schedule s_p has an ‘up-down-up’ shape qualitatively similar to the $s^{\text{p-lin}}(t)$ found in Ref. [140] (see their Figure 6).

To better understand the quality of the smooth solutions obtained with dQA-CRAB, we also show in Fig. 3.6 a direct comparison of the residual energies ϵ_p^{res} obtained from three methods, QAOA-RAND, dQA-CRAB, and dQA-Linear for fixed $N = 13$. The shaded region for the QAOA-RAND data spans from the minimum to the maximum of the residual energy values obtained from 100 different random realizations for each P . Notice how before P_1^{cr} , dQA-CRAB yields solutions notably better than even the best case of randomly initialized QAOA. This implies that for small P we achieve reasonably good solutions with a smooth schedule produced from dQA-CRAB, while QAOA-RAND should be avoided, as it leads to poor-quality local minima. When P increases and crosses P_1^{cr} , dQA-CRAB shows a less abrupt decrease in the residual energies obtained, compared to the steep drop obtained in QAOA-RAND. At and beyond $P = 56$, dQA-CRAB reaches residual energies around or below 10^{-12} , as shown by the lower horizontal dashed line. On the other hand, dQA-Linear [49, 90, 156, 168] performs poorly for any value of P and, within the number of layers we considered, it never overcomes the bottleneck of

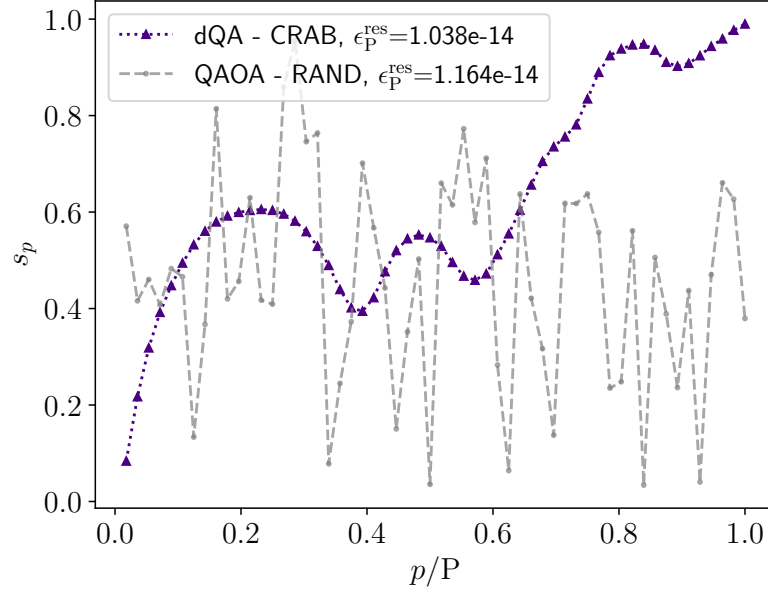


Figure 3.5.: Comparison of smooth and random schedules s_p (see Eq. (1.40)) obtained from dQA-CRAB and QAOA-RAND, respectively. Here we consider the frustrated-ring model for $N = 13$ and $P = 56 > P_1^{\text{cr}}$. Both approaches achieve a numerically zero residual energy ($\epsilon_P^{\text{res}} \approx 10^{-14}$), while dQA-CRAB exhibits smoothness and regularity in the resulting schedule.

a small spectral gap. In fact, based on the estimate of the spectral gap for $N = 13$, which is $\Delta_{\min} \approx 1.6 \times 10^{-5}$, one would anticipate Landau-Zener timescale of order $1/\Delta_{\min}^2 \approx 10^{10}$. Therefore, in a digitized QA with a linear protocol, we expect to need $P \approx 10^{10}$ before the adiabatic behavior sets in.

Figure 3.6 also shows, as the upper horizontal dashed line labeled $c = 0.1$, the performance criterion used by Côté *et al.* in [140] (their equation (9)):

$$E_P - E_{\min} \leq c \Delta_{s=1} , \quad (3.30)$$

where $\Delta_{s=1} = 2(J_w - J_f)$ is the energy gap between the ground state and the first excited state of the problem Hamiltonian [140]. When $c = 0.1$, $N = 13$ and $J_w - J_f = 0.05$, such threshold corresponds to $\epsilon_P^{\text{res}} \leq 7.6723 \times 10^{-4}$, and is reached by dQA-CRAB at $P = 28$. A more detailed comparison of the efficiency of our algorithm against the continuous-time strategy of Ref. [140] will be discussed in Sec. 3.4.4.

3.4.3. Populations of the instantaneous eigenstates in dQA-CRAB

To study what kind of dynamics, adiabatic or otherwise, the dQA-CRAB control parameters (θ_p^x, θ_p^z) generate at each discrete “time step” p , we would like to access

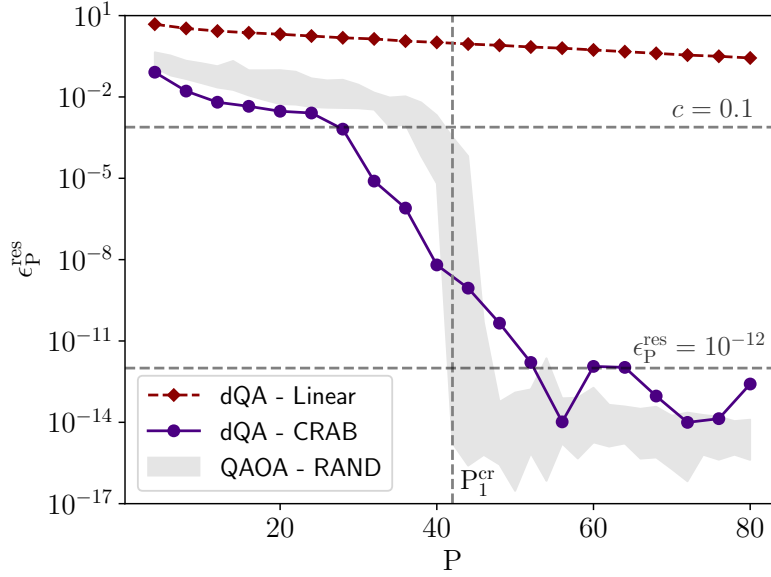


Figure 3.6.: Comparison of residual energy ϵ_P^{res} , defined in Eq. (3.27), versus number of layers P , for different protocols: dQA-Linear and dQA-CRAB from Section 3.3.1; QAOA-RAND from Section 3.3.2. The system size considered is $N = 13$. The dashed vertical line indicates the critical depth $P_1^{\text{cr}} = 42$, at which some realizations of QAOA-RAND start to produce numerically zero residual energies. The lower horizontal line $\epsilon_P^{\text{res}} = 10^{-12}$ is a threshold below which ϵ_P^{res} is considered a numerical zero. The upper horizontal line $c = 0.1$ is a performance criterion set by Côté *et al.* in [140], shown here for comparison.

the digitized equivalent of the *instantaneous* eigenstates of the evolving Hamiltonian. Unfortunately, at each step p we apply two unitary operators, $e^{-i\theta_p^x \hat{H}_x} e^{-i\theta_p^z \hat{H}_z}$, and this complicates the unambiguous identification of such “instantaneous eigenstates”. In this case, identifying a closed expression for the evolving Hamiltonian is not feasible, as it involves an infinite series of high-order products of $\hat{H}_x$ and $\hat{H}_z$. Moreover, the spectrum of the effective Hamiltonian at step p , defined as the logarithm of the evolution operator $\hat{U}(\theta_p^x, \theta_p^z)$, is determined only modulo- 2π . This leads to an inherent ambiguity in ordering the energy levels and identifying the instantaneous ground state. To bypass such difficulty, we consider, given the control parameters (θ_p^x, θ_p^z) , an effective Hamiltonian $\hat{H}_p^{\text{eff}}$ defined by the first-order term of the Baker-Campbell-Hausdorff (BCH) formula, i.e.:

$$\hat{H}_p^{\text{eff}} = \frac{1}{\Delta_p} \left(\theta_p^x \hat{H}_x + \theta_p^z \hat{H}_z \right) = s_p \hat{H}_z + (1 - s_p) \hat{H}_x, \quad (3.31)$$

where s_p and Δ_p are defined in Eq. (1.40). Let $|\phi_p^{j,\text{eff}}\rangle$ be the j th instantaneous eigenstate of the effective Hamiltonian $\hat{H}_p^{\text{eff}}$. We then define the population of the instantaneous eigenstates $|\phi_p^{j,\text{eff}}\rangle$ as:

$$P_j(p) = |\langle \phi_p^{j,\text{eff}} | \Psi_p \rangle|^2, \quad (3.32)$$

where $|\Psi_p\rangle = \hat{U}(\theta_p^x, \theta_p^z) \cdots \hat{U}(\theta_1^x, \theta_1^z) |\Psi_0\rangle$. The eigenstates $|\phi_p^{j,\text{eff}}\rangle$ are constructed, for small systems, by performing exact diagonalization; for larger systems, we rely on

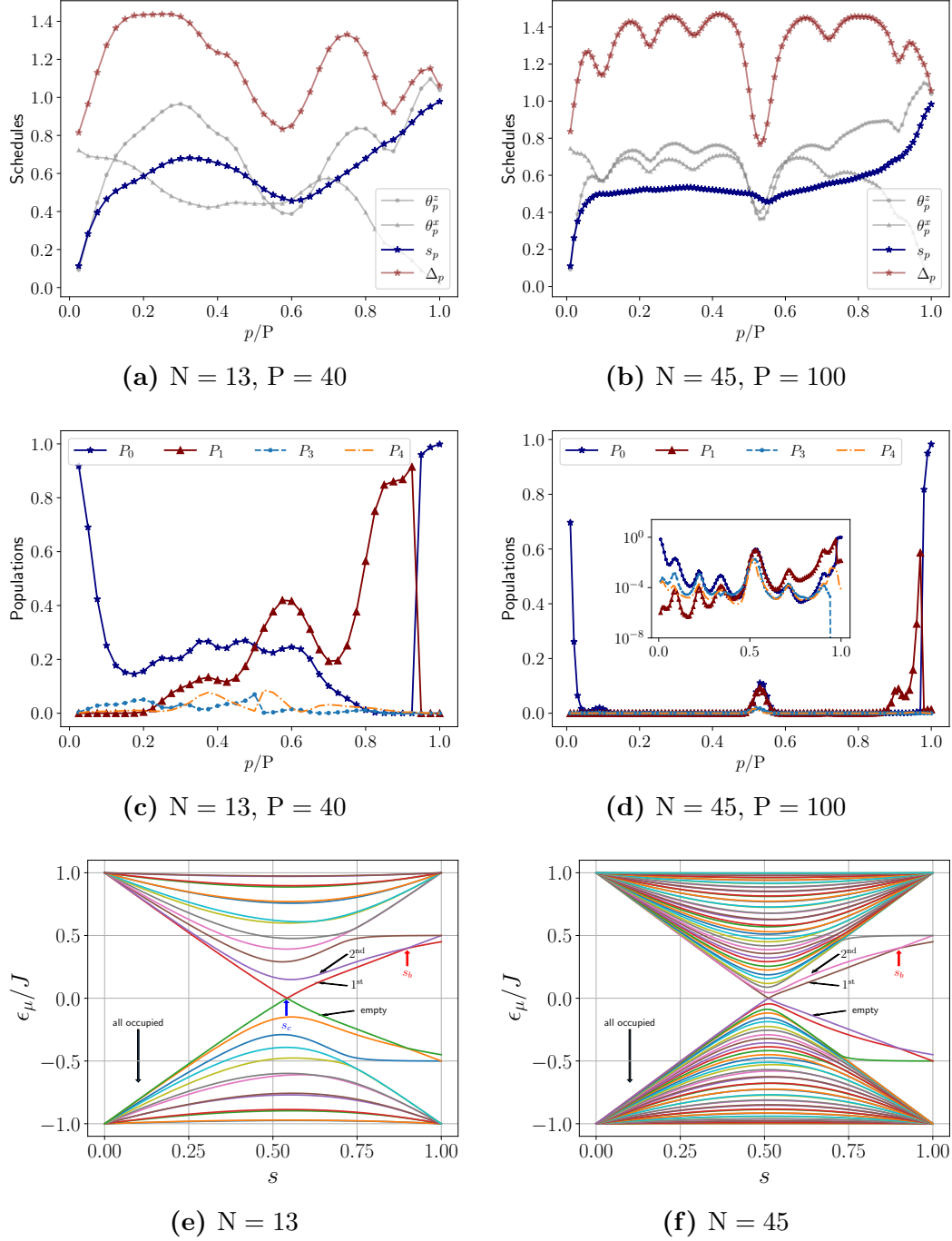


Figure 3.7.: (a,b) Optimal dQA-CRAB parameters for different N and P . The (dimensionless) protocol s_p and the time interval Δ_p are defined in Eq.(1.40). Δ_p , θ_p^x and θ_p^z have the dimension of a time in units of $\hbar/J$. (c,d) Populations of the instantaneous ground (P_0) and lowest excited states ($P_1 \dots P_4$). (e,f) Bogoliubov-de Gennes eigenvalue spectrum for the interpolating Hamiltonian $s\hat{H}_z + (1-s)\hat{H}_x$ versus s . The optimal schedules in (a,b) are generated using dQA-CRAB: two super-iterations of dressed-CRAB, where frequencies are sampled from gamma distributions, as described in Sec. 3.3.1. The residual energy for the case $N = 45, P = 100$ is $\epsilon_p^{\text{res}} = 1.2 \times 10^{-4}$.

the Jordan-Wigner mapping followed by a Bogoliubov transformation, and employ the techniques explained in Ref. [78][Appendix C] to calculate overlaps between states.

Figure 3.7(a,b) show the optimal dQA-CRAB schedules (θ_p^x , θ_p^z , and associated s_p and Δ_p) for $N = 13$ and $P = 40$ (a), and for $N = 45$ and $P = 100$ (b). Figure 3.7(c,d) show the corresponding discrete-time ground (P_0) and low excited states ($P_1 \cdots P_4$) instantaneous populations, for the same two values of N and P . An adiabatic algorithm would keep a large ground state population P_0 throughout the evolution. In contrast, we see that our dQA-CRAB dynamics implements, effectively, a shortcut-to-adiabaticity (STA) [129, 187] mechanism: the ground state population P_0 rapidly decreases, hence promoting excitations to higher levels, to increase sharply at the final stages of the dynamics, where the closing spectral gap is encountered.

To illustrate such a phenomenon in more detail, we refer the reader to the Bogoliubov-de Gennes spectrum shown in Fig. 3.7(e,f), again for the two values of N . We see that there is a clear singularity in the spectrum at $s_c \approx 0.5$, with a **vanishing gap** in the Nambu spectrum around the particle-hole symmetry point $E = 0$. This implies a true level crossing, which we will comment on shortly. For $s < s_c$, the system's state is obtained by occupying all the N negative energy Bogoliubov-de Gennes eigenvalues, with a large gap to the first excited state. However, as mentioned, at s_c a level crossing implies that the occupied state with energy $-\epsilon_1$ for $s < s_c$ is *adiabatically transformed* into the state at energy $+\epsilon_1$ when $s > s_c$, hence it remains *occupied*. Symmetrically, the (empty) state with energy $+\epsilon_1$ for $s < s_c$, is adiabatically transformed into the state at energy $-\epsilon_1$ for $s > s_c$, and remains empty. Such an adiabatic transformation is quite detrimental for the quality of the evolution: the occupied state at energy $+\epsilon_1$ when $s > s_c$ suffers from the closing gap bottleneck at s_b , where it would almost unavoidably become excited. The optimal strategy is that the system *preempts* such occurrence by exciting the system, around $s \approx s_c$, ideally to the 2nd level (see Fig. 3.7)(e,f): this will directly lead, when the system reaches the avoided crossing bottleneck at s_b , towards the final ground state. The populations of eigenstates calculated for the case with $N = 13$, see Fig. 3.7(a), follow quite closely such a two-state STA scenario, as witnessed by the small involvement of higher excited states. What is more surprising is the result for $N = 45$, see Fig. 3.7(b), where the vicinity of higher excited states in the Nambu spectrum at s_c , see Fig. 3.7(f), leads to appreciable populations of quite a number of excited states: indeed, both P_0 and P_1 remain quite small everywhere, with a final surge of P_1 right before the location of the bottleneck s_b , where the standard two-state STA scenario apparently takes place.

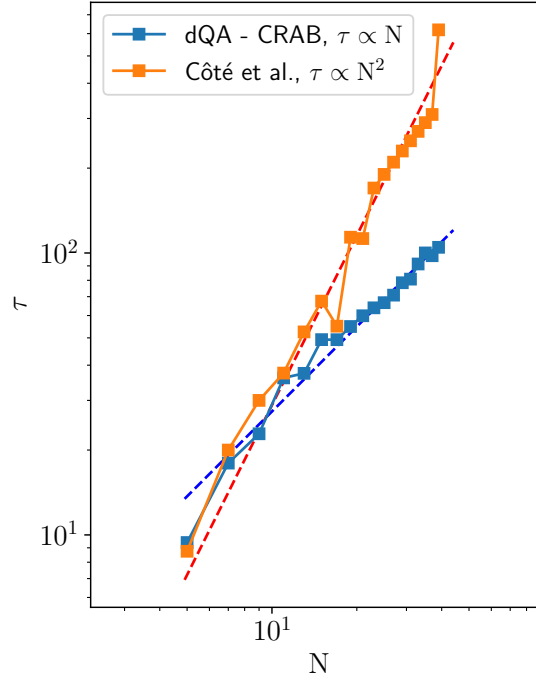


Figure 3.8.: Comparison of the annealing time τ needed to reach the threshold $c = 0.1$, see Eq. (3.30), for our dQA-CRAB method and the method of Côté *et al.* [140] (data taken from their Fig. 5), versus the system size N . Dashed lines show our best fit of the two approaches, compatible with a linear versus quadratic scaling in N .

3.4.4. Annealing time comparison

In the digitized setting, the total annealing time is defined, see Eq. (3.29), as:

$$\tau = \sum_{p=1}^P (\theta_p^x + \theta_p^z) . \quad (3.33)$$

We compare the efficiency of our dQA-CRAB algorithm against the continuous-time approach by Côté *et al.* [140] in terms of the annealing time required to reach the threshold $c = 0.1$ as defined in 3.30. Figure 3.8 shows such comparison. While a quadratic scaling $\tau \propto N^2$ is observed for the continuous-time approach of Ref. [140], we see that our digital algorithm using the smooth schedules obtained from dQA-CRAB shows a linear scaling $\tau \propto N$.

3.5. Conclusions and perspectives

We have explored digitized QA approaches and QAOA to overcome the challenges posed by exponentially small spectral gaps, a common bottleneck to adiabatic evolutions in traditional QA [128]. We have focused on the frustrated-ring model, proposed as the simplest class of Ising spin systems whose minimum gap decreases

3. Optimal Control of a Frustrated Ising Ring Model

exponentially with the system size N . By exploiting the Jordan-Wigner mapping, we were able to perform an exact simulation of the digital dynamics up to quite large values of N .

As a first result, we leveraged QAOA to provide a detailed characterization of the different controllability [110] regimes of the model. At any system size N , we identify two critical depths P_1^{cr} and $P_2^{\text{cr}} > P_1^{\text{cr}}$ for the associated QAOA quantum circuit, defining three distinct regimes. For depths with $P < P_1^{\text{cr}}$, QAOA with a random initialization of parameters (θ_p^x, θ_p^z) never achieves the exact ground state of the model. As the depth of the circuit equals P_1^{cr} , the probability of reaching the exact ground state with a random initialization of QAOA starts to be finite, and continues to grow until the second critical value P_2^{cr} is reached. Beyond this point, any random initialization of the angles (θ_p^x, θ_p^z) leads to an optimal protocol that converges to the exact ground state. Furthermore, we numerically determined an exact expression for $P_1^{\text{cr}} = (N^2 - 1)/4$, which scales as the square of the number of spins. This offers a significant advantage over the exponentially long evolution time required in standard QA.

The scaling of P_1^{cr} is supported by numerical evidences and analytical arguments, related to the number of parameters required to fully encode the dynamics [125], as mentioned in Sec. 3.4.1. Concerning P_2^{cr} , its dependence on N is less clear, partly due to the non-sharp definition of the threshold we consider for a “numerical zero”.

The second result of our work concerns the design of optimal smooth digital schedules (θ_p^x, θ_p^z) . To do so, we employed dressed-CRAB [108] quantum control techniques in our digitized QA approach. Beyond ensuring smoothness, this procedure, which we dubbed dQA-CRAB, outperforms a randomly-initialized QAOA for circuit depths below the critical threshold, $P < P_1^{\text{cr}}$. In the controllable regime $P > P_1^{\text{cr}}$, the smooth schedules produced with dQA-CRAB continue to perform very well, leading to final states very close to the actual ground state.

In order to characterize the dynamics implemented by smooth dQA-CRAB solutions, we focused on the populations of the instantaneous eigenstates of the effective Hamiltonian associated with the digital dynamics. Unlike in adiabatic protocols, we see clear signs of a shortcut-to-adiabaticity mechanism taking place in our protocol. Far before the closing of the spectral gap, the dQA-CRAB optimal schedules smoothly populate the excited states of the effective Hamiltonian spectrum. Subsequently, the system non-adiabatically crosses the small spectral gap, leading to population inversion and to a large occupation of the final ground state.

Finally, we compared our dQA-CRAB with the continuous-time diabatic annealing algorithm proposed in [140]. We show that, for a fixed predefined accuracy threshold, our method requires a total annealing time τ that scales linearly with the system-size, $\tau \propto N$, as opposed to quadratically, $\tau \propto N^2$, as in [140].

One of the most important messages of our study is that the spectral gap bottleneck, a nightmare of traditional QA approaches based on linear schedules [128], has nothing to do with the actual computational difficulty of the problem at hand or to how hard is the controllability issue: the frustrated-ring Ising model is an example of a polynomially controllable problem with a severe exponentially small spectral gap bottleneck.

An interesting avenue to pursue is that of a *quantum speed limit* [182, 189–192] in a digitized framework, i.e., the minimum circuit depth required to exactly reach a target state, starting from a fixed initial configuration. This quantity represents the ultimate limit for optimal digitized control protocols [104]. Previous work on specific models are known [49, 185, 193, 194], but a more systematic understanding is mandatory.

A detailed study of the optimal continuous-time evolution, and its associated quantum speed limit, would be an intriguing extension of this work, offering insights into the interplay between controllability in the digital setting and the fundamental limitations imposed by quantum speed limits.

Another promising avenue involves employing counterdiabatic driving [83, 84] into the current framework, see also [87, 186], which could potentially further improve the accuracy of the solution and computational efficiency. Exploring this integration in both continuous and digitized settings may reveal novel ways to enhance the performance of the proposed quantum algorithms.

Further analytical support for the digital-controllability threshold identified here is provided by Arezzo *et al.* [125] in our study of transverse-field Ising chains. By relating the critical QAOA depth for exact ground-state preparation to the number of independent parameters required to represent the free-fermion dynamics, our analysis recovers $P_1^{\text{cr}} = (N^2 - 1)/4$ for the reflection-symmetric frustrated ring, supporting the numerical scaling reported in this Chapter.

4. Parity Mapping for Quantum Optimization

Chapter 3 studied the frustrated Ising ring in its native encoding, where its exponentially closing gap creates a bottleneck for quantum annealing (QA). Here, we keep the same model but change its encoding, using the parity mapping, reviewed in Section 1.4, to investigate how this transformation affects the performance of both QA and the Quantum Approximate Optimization Algorithm (QAOA).

For QA, finite-size calculations show that the parity mapping increases the minimum spectral gap under the energy normalization used in this work, thereby enabling faster continuous-time ground state preparation protocols. An ideal implementation of Parity-QA, with a single global constraint, shows no evidence of exponential gap closing over the accessible system sizes, whereas a hardware-motivated decomposition into local constraints restores the exponential decrease, albeit with a smaller fitted exponent than conventional QA, as discussed in Section 4.5.1.

For the digitized protocol, we find that the number of Parity-QAOA layers required to prepare the exact ground state remains constant over the simulated sizes, improving upon the quadratic scaling required by conventional QAOA, as shown in Section 4.5.2.

To investigate the role of constraints in Parity-QAOA, we further consider a modified Ising ring instance in which the constraint term is essential for preparing the target ground state. We then compare the corresponding resource requirements with those of conventional QAOA, as discussed in Section 4.5.2.

4.1. Introduction

Previous studies of Parity-QA and Parity-QAOA focused on highly connected target Hamiltonians for which the parity mapping introduces a substantial qubit overhead [122, 195–197]. In this work, we investigate parity-based optimization for a sparse target Hamiltonian, namely the frustrated Ising ring model and its variants [80]. These models are of interest both for theoretical studies of QA and QAOA, and for efficient practical implementations within the parity architecture. The frustrated Ising ring, proposed in Ref. [80], is one of the simplest models exhibiting

exponentially closing spectral gaps, making it a compelling test case for continuous-time approaches that extend traditional QA [140, 186]. In Refs. [86, 125], digital QAOA approaches applied to this model demonstrated a remarkable exponential speedup with respect to standard QA, requiring a number of alternating unitaries that only grows quadratically with system size for exact ground-state preparation. Moving towards the parity framework, we find that these ring models differ from previously studied highly connected models because they incur zero qubit overhead and require fewer unitaries than standard QAOA. They therefore provide valuable test cases and serve as a natural starting point for evaluating parity-based methods on sparse graphs. Moreover, we observe an amplification of the spectral gaps when using Parity-QA, compared to standard linear-schedule QA, which suggests the possibility of mitigating the exponentially closing gap under ideal annealing hardware. The frustrated ring model also represents a special case for Parity-QAOA, as it is the first instance in which the unitary associated with the Parity-constraint is not required, and the problem can be solved using only single-qubit rotations. Finally, to further clarify the role of the constraint, we examine additional variants of the model.

The Chapter is structured as follows. Section 4.2 introduces the Parity transformation on the Hamiltonians of interest. Section 4.3 introduces the original QAOA and its parity-adapted version, which we call henceforth Parity-QAOA. Section 4.4 presents the frustrated Ising ring model in its original formulation, as well as a few variants that we will study. We further report numerical results on the scaling of the minimal spectral gap in Parity-QA as compared with standard QA. Section 4.5 presents the results obtained using Parity-QAOA. In Section 4.5.2, we analyze the case in which the protocol omits the constraint and achieves the exact solution with a constant resource scaling independent of system size. Finally, in Sec. 4.5.2, we examine the case where the constraint is included and compare the resulting CNOT depth and CNOT gate count with those of standard QAOA.

4.2. The Parity Transformation

In this section, we discuss the main technique in this work, namely the Parity or LHZ architecture [118]. We consider the problem of preparing the ground state of the following Ising ring target Hamiltonian [78]

$$\hat{H}_{\text{targ}} = \hat{H}_z = - \sum_{j=1}^N J_j \hat{\sigma}_j^z \hat{\sigma}_{j+1}^z, \quad (4.1)$$

with periodic boundary conditions $\hat{\sigma}_{N+1}^z = \hat{\sigma}_1^z$. An example of a ring with $N = 7$ is shown in Fig. 4.1 (a).

The parity mapping $\hat{\sigma}_j^z \hat{\sigma}_{j+1}^z \rightarrow \tilde{\sigma}_k^z$ maps pairs of interacting logical qubits to

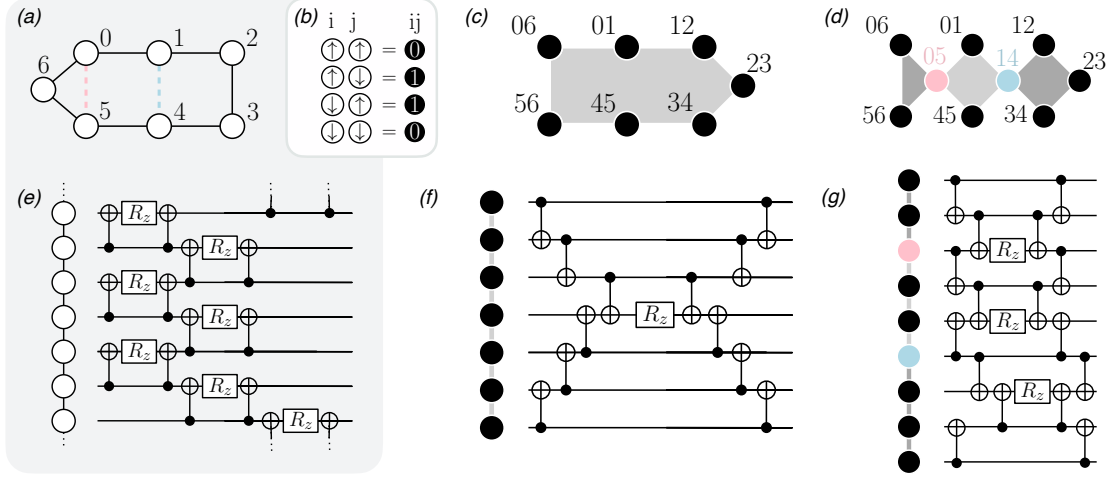


Figure 4.1.: The Ising ring with $N = 7$ in its logical representation (a) and parity representations (c)–(d). The parity transformation table is given in (b). Without ancillas the parity graph has an N -body constraint visualized by a grey polygon in (c). With ancilla interactions (red and blue dashed lines in (a)) the parity graph has three constraints of size three or four, shown in (d). The digital implementations of the multi-qubit gates are shown in (e), (f) and (g) for the logical graph, parity graph and the parity graph with ancillas, respectively.

individual physical (or parity) qubits. The translation table is shown in Fig. 4.1 (b). If two logical qubits $\hat{\sigma}_j^z$ and $\hat{\sigma}_{j+1}^z$ are parallel or anti-parallel the corresponding k -th parity qubit is 0 or 1, respectively. After the Parity mapping, the Hamiltonian $\hat{H}_z$ encoding the Ising ring model reads:

$$\hat{H}_z \rightarrow \tilde{H}_z = - \sum_{k=1}^K \tilde{J}_k \tilde{\sigma}_k^z, \quad (4.2)$$

with $\tilde{J}_k = J_k$ and $K = N$: The number of physical qubits K coincides with the number of pairwise interactions in the target Hamiltonian, and hence in this case equals N , the number of logical qubits in the ring structure.

The full 2^K -dimensional Hilbert space of the parity qubits contains configurations that cannot be decoded into logical spin assignments. For example, for odd N , the physical configuration $|1\rangle^{\otimes K}$ cannot be decoded to a logical one because no logical state has all neighboring qubits anti-parallel to each other. Therefore, constraints are required to guarantee the logical validity of physical configurations. In Ref. [118], the number of constraints needed is specified to be $K - N + 1$. Hence, a single constraint is required for the parity transformation of the Ising ring model at hand ¹.

¹Here $K = N$, but one parity configuration can be decoded into two logical ones and, therefore, 2^{N-1} physical states correspond to 2^N logical ones, the remainder are non-valid states. Hence, one constraint is needed.

In particular, a valid state $|\tilde{\psi}_{\text{valid}}\rangle$ satisfies the following identity:

$$\prod_{k=1}^K \tilde{\sigma}_k^z |\tilde{\psi}_{\text{valid}}\rangle = (\hat{\sigma}_1^z \hat{\sigma}_2^z)(\hat{\sigma}_2^z \hat{\sigma}_3^z) \cdots (\hat{\sigma}_{K-1}^z \hat{\sigma}_K^z)(\hat{\sigma}_K^z \hat{\sigma}_1^z) |\tilde{\psi}_{\text{valid}}\rangle = |\tilde{\psi}_{\text{valid}}\rangle. \quad (4.3)$$

In other words, a valid computational-basis configuration contains an even number of parity qubits occupying the state $|1\rangle$. To enforce this constraint, we add a non-local penalty term to the target Hamiltonian, denoted as $\tilde{H}_c$:

$$\tilde{H}_c = - \prod_{k=1}^K \tilde{\sigma}_k^z, \quad (4.4)$$

which is minimized by all valid states.

To solve the optimization problem within the parity picture, we must minimize the following target Parity Hamiltonian:

$$\tilde{H}_{\text{targ}} = \tilde{H}_z + C \tilde{H}_c. \quad (4.5)$$

Here, C is a constraint strength that should be chosen sufficiently large to ensure that the ground state of $\tilde{H}_{\text{targ}}$ is a valid state. Ref. [198] investigates the minimal required constraint strength and its scaling for various problem classes. For the problem we study in this work, the Ising ring, the constraint strength should be $C > \min_j |J_j|$. Finally, the driving Hamiltonian in Parity-QA and Parity-QAOA can be defined as the driving Hamiltonian defined in the Introduction, Sec. 4.1, namely $\tilde{H}_{\text{drive}} = \tilde{H}_x = - \sum_{k=1}^K \tilde{\sigma}_k^x$. Here, the $\tilde{\sigma}_k^x$ operators now act on K parity qubits.

Fig. 4.1 (c) illustrates the parity transformation of the Ising ring depicted in Fig. 4.1 (a). The constraint $\tilde{H}_c$, which is visualized as a gray area, involves a K -qubit term that is not practical for current annealing hardware, especially as the system size grows. To address this issue, ancilla qubits can be introduced to break the K -qubit term into 3- and 4-qubit local terms. These ancilla qubits usually correspond to pairs of non-interacting logical qubits in the original problem. An example is given in Fig. 4.1 (a), where the colored dashed lines (red and blue) indicate auxiliary interactions. The corresponding parity architecture is shown in Fig. 4.1 (d). The constraint Hamiltonian of Eq. (4.4) with ancilla qubits, as discussed in Ref. [119], changes to $\tilde{H}_c = - \sum_{l=1}^L \prod_{k \in C_l} \tilde{\sigma}_k^z$ where L is the total number of constraints and C_l is the set of qubit indices participating in the l -th constraint. The ancilla qubits play a crucial role in enabling implementations of quantum annealing on analog NISQ devices. Therefore, as we study the spectra in QA and Parity-QA in Section 4.5.1, we also explicitly investigate how these ancilla qubits affect the spectral gaps.

4.3. Parity-QAOA

The Quantum Approximate Optimization Algorithm (QAOA), later referred to as standard QAOA, adopts an alternating *Ansatz* in evolving the quantum states. The final state $|\psi_P(\boldsymbol{\theta}^x, \boldsymbol{\theta}^z)\rangle$ is obtained from

$$|\psi_P(\boldsymbol{\theta}^x, \boldsymbol{\theta}^z)\rangle = \hat{U}_x(\theta_P^x) \hat{U}_z(\theta_P^z) \cdots \hat{U}_x(\theta_1^x) \hat{U}_z(\theta_1^z) |\psi_0\rangle \quad (4.6)$$

where the initial state $|\psi_0\rangle = |+\rangle^{\otimes N} = \left(\frac{|1\rangle+|0\rangle}{\sqrt{2}}\right)^{\otimes N}$. The alternating unitary operators are defined using the driving Hamiltonian and the target Hamiltonian, respectively. At each layer $p = 1, \dots, P$, we have:

$$\hat{U}_x(\theta_p^x) = e^{-i\theta_p^x \hat{H}_x}, \quad \hat{U}_z(\theta_p^z) = e^{-i\theta_p^z \hat{H}_{\text{targ}}} = e^{-i\theta_p^z \hat{H}_z}. \quad (4.7)$$

The 2P parameters $(\boldsymbol{\theta}^x, \boldsymbol{\theta}^z)$ are optimized by minimizing the variational energy:

$$E_P(\boldsymbol{\theta}^x, \boldsymbol{\theta}^z) = \langle \psi_P(\boldsymbol{\theta}^x, \boldsymbol{\theta}^z) | \hat{H}_{\text{targ}} | \psi_P(\boldsymbol{\theta}^x, \boldsymbol{\theta}^z) \rangle. \quad (4.8)$$

A similar *Ansatz* can be written for the Parity-QAOA [120], following the parity mapping of Sec. 4.2:

$$|\tilde{\psi}_P(\boldsymbol{\theta}^x, \boldsymbol{\theta}^c, \boldsymbol{\theta}^z)\rangle = \tilde{U}_x(\theta_P^x) \tilde{U}_c(\theta_P^c) \tilde{U}_z(\theta_P^z) \cdots \tilde{U}_x(\theta_1^x) \tilde{U}_c(\theta_1^c) \tilde{U}_z(\theta_1^z) |\tilde{\psi}_0\rangle. \quad (4.9)$$

The initial state is $|\tilde{\psi}_0\rangle = |+\rangle^{\otimes K}$. Three types of unitaries are applied to the quantum state at each layer p , and this is due to the additional constraint term $\tilde{H}_c$ in the target Hamiltonian $\tilde{H}_{\text{targ}}$. These unitaries read:

$$\tilde{U}_x(\theta_p^x) = e^{-i\theta_p^x \tilde{H}_x}, \quad \tilde{U}_c(\theta_p^c) = e^{-i\theta_p^c \tilde{H}_c}, \quad \tilde{U}_z(\theta_p^z) = e^{-i\theta_p^z \tilde{H}_z}. \quad (4.10)$$

At each layer, there are three free parameters $(\theta_p^x, \theta_p^c, \theta_p^z)$ and hence the final state at layer P is determined by $3P$ parameters that are determined by an external classical optimization loop. The variational energy now reads:

$$E_P(\boldsymbol{\theta}^x, \boldsymbol{\theta}^c, \boldsymbol{\theta}^z) = \langle \tilde{\psi}_P(\boldsymbol{\theta}^x, \boldsymbol{\theta}^c, \boldsymbol{\theta}^z) | \tilde{H}_{\text{targ}} | \tilde{\psi}_P(\boldsymbol{\theta}^x, \boldsymbol{\theta}^c, \boldsymbol{\theta}^z) \rangle, \quad (4.11)$$

where $\tilde{H}_{\text{targ}}$ is as defined in Eq. (4.5).

In the standard QAOA, the interaction unitary $\hat{U}_z(\theta_p^z) = \prod_{j=1}^N e^{-i\theta_p^z J_j \hat{\sigma}_j^z \hat{\sigma}_{j+1}^z}$ can be compiled with a single-qubit rotation $R_z^{(j)}(\alpha) = e^{-i\alpha \hat{\sigma}_j^z/2}$ and two controlled-NOT (CNOT) gates for each edge $(j, j+1)$ of the ring, using the identity $e^{-i\alpha \hat{\sigma}_j^z \hat{\sigma}_{j+1}^z} = \text{CNOT}_{j,j+1} R_z^{(j+1)}(2\alpha) \text{CNOT}_{j,j+1}$. The interaction unitary $\hat{U}_z(\theta)$ therefore uses $2N$ CNOT gates per layer. Assuming a ring connectivity for the quantum device and scheduling the ring edges in parallel gives CNOT depth four for even N and six for odd N , as illustrated in Fig. 4.1(e).

In the Parity-QAOA implementation, the unitaries $\tilde{U}_x(\theta_p^x)$ and $\tilde{U}_z(\theta_p^z)$ require only single qubit rotations, while the constraint unitary $\tilde{U}_c(\theta_p^c)$ involves a K -qubit

entangling gate. It can be implemented with the circuit shown in Fig. 4.1 (f), comprising a chain sequence of CNOT gates connecting all physical qubits, followed by a single qubit z-rotation R_z , followed by the reversed preceding CNOT chain. The number of CNOT gates required to implement the constraint unitary $\tilde{U}_c$ is $2(N - 1)$, and the CNOT depth is N for even N and $N + 1$ for odd N (see Fig. 4.1, f).

Thus, per QAOA layer, standard QAOA has constant CNOT depth and $2N$ CNOTs, whereas the global-constraint Parity-QAOA implementation has linear CNOT depth and $2(N - 1)$ CNOTs. Whether this trade-off is advantageous depends on the number of layers required, the physical connectivity, and the relative importance of entangling-gate errors and circuit duration.

4.4. Frustrated Ising ring models

The standard frustrated Ising ring model is defined in Section 3.2, with its couplings given in Eq. (3.2). Figure 4.2 shows the bond labels used below to define its variants.

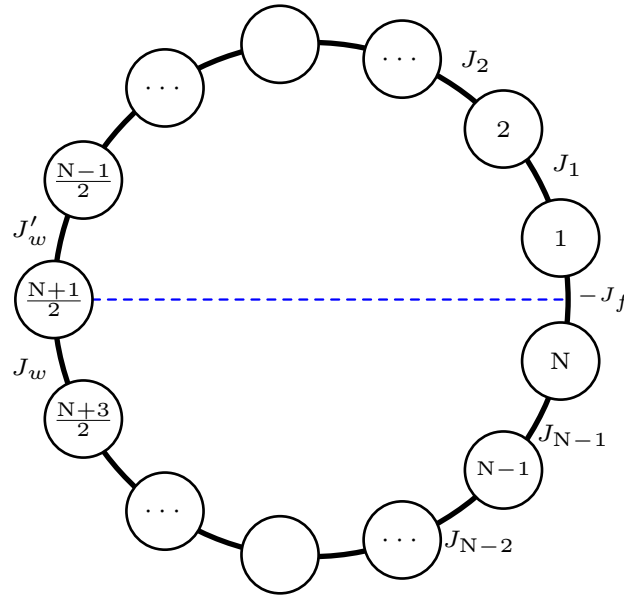


Figure 4.2.: The frustrated Ising ring with N spins. The antiferromagnetic bond $-J_f$ sits at site $j = N$; the two weak bonds J_w and J'_w flank the central spin at $j = (N + 1)/2$, with $J'_w = J_w$ in the standard symmetric case. All remaining bonds equal J in the standard model and may take distinct values in the modified instances (see text). The dashed blue line indicates the possible reflection symmetry axis.

As studied in [86, 125] (see Fig. 2 therein), the interpolating Hamiltonian, given in Eq. (1.6), with the frustrated Ising ring as $\hat{H}_{\text{targ}}$ exhibits two closing gaps between the ground state and the first excited state during the linear-schedule annealing

process: the first closes polynomially at $s \approx 0.5$ and the second closes exponentially at $s \approx 0.9$. We focus our analysis on the smallest gap because it is the dominant spectral bottleneck for linear schedules.

The minimum gap for the standard QA interpolating Hamiltonian and the Parity-QA interpolating Hamiltonian $s\tilde{H}_{\text{targ}} + (1-s)\tilde{H}_{\text{drive}}$ is given as

$$\Delta_{\min} = \min_s \Delta_s \quad \text{with } \Delta_s = E_d(s) - E_0(s), \quad (4.12)$$

where the integer d represents the ground state's degeneracy of the target Hamiltonian, while $E_0(s)$ and $E_d(s)$ represent the energies of the ground state and the d -th excited state of the interpolating Hamiltonian at a given point of the schedule s .

We further introduce variants of the original frustrated ring of Eq. (3.2) and Fig. 4.2 to support our subsequent analysis and provide richer results and comparisons. The first two variants break the reflection symmetry around the central site $j = (N+1)/2$ and increase the number of distinct coupling values J_j , while keeping the property of an exponentially closing gap: (i) we set $J_1/J = J_2/J = 0.8$ and keep the remaining couplings unchanged; (ii) Building on (i), we additionally set $J_{N-1}/J = 0.85$. Ref. [125] shows that these changes preserve the presence of an exponentially closing energy gap and hence the difficulty for linear-schedule QA. These variants are useful to investigate correlations between the performance of Parity-QAOA and the number of distinct coupling values appearing in the target Hamiltonian (see Sec. 4.5.2).

The third variant is an instance of the frustrated ring of Eq. (3.2), which is inspired by Refs. [199–201]: it consists of setting $J_w/J = 1$ and $-J_f/J = -1$, so that all couplings have unit magnitude and the ring carries a single antiferromagnetic bond of equal magnitude with respect to all other ferromagnetic couplings. The product of all J_j is $\prod_j \frac{J_j}{J} = -1$, so no spin configuration can satisfy all bonds simultaneously: exactly one bond must host a domain wall, which can sit on any of the N bonds. The classical ground space is therefore $2N$ -fold degenerate (N domain-wall positions times the global spin flip), in contrast to the two-fold degenerate ground state of the original frustrated ring of Eq. (3.2).

4.5. Results

First, we present the minimal energy gap $\Delta_{\min}$, Eq. (4.12), obtained from standard-QA and Parity-QA simulations. We then turn to QAOA simulations and examine whether the trends observed in QA carry over to QAOA. Simulation details of all simulations can be found in App. C.1.

4.5.1. Minimal Energy gap in QA

The frustrated Ising ring model in its original formulation, Eq. (3.2) (and the first two introduced variants in Sec. 4.4) has an exponential closing gap. This is verified in Fig. 4.3 (a) where the red line, labeled as ‘standard QA’, shows the results for $\Delta_{\min}$ for $N = 5, 7, \dots, 15$. The figure also shows the scaling for Parity-QA’s interpolating Hamiltonian, constraint strength set to $C = J$, for different numbers of ancilla qubits n_a .

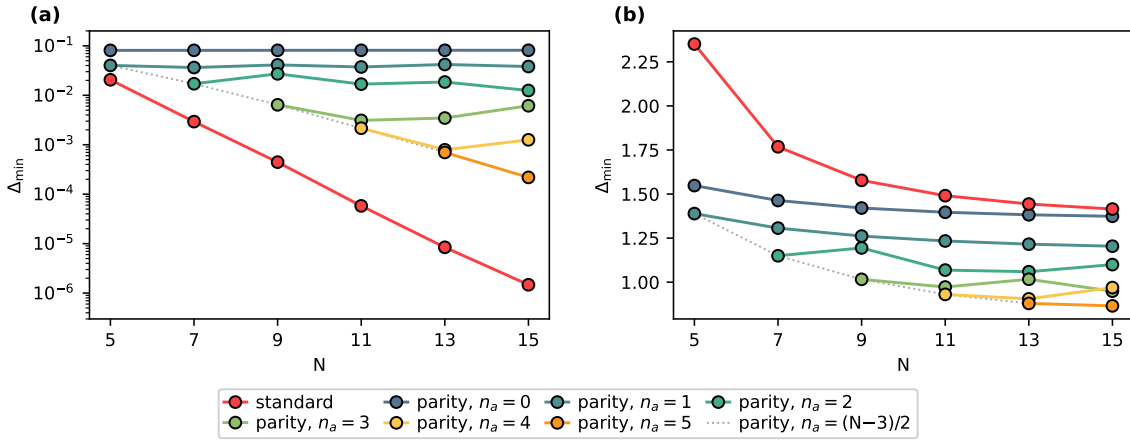


Figure 4.3.: Minimal energy gap $\Delta_{\min}$ versus system size N in the standard and parity implementations, for (a) the frustrated ring model (constraint strength $C = J$; note the logarithmic scale) and (b) the ring with uniform coupling magnitudes and a single antiferromagnetic bond (constraint strength $C = 2J$; linear scale). In both panels, solid coloured lines show the parity gap at fixed ancilla number $n_a = 0, 1, \dots, 5$, each starting at the smallest admissible ring $N = 2n_a + 3$; the grey dotted line is the realistic parity gap at the hardware-required $n_a = (N - 3)/2$. In (a), the standard and realistic-parity gaps close as $\sim e^{-0.96N}$ and $\sim e^{-0.51N}$, so the parity mapping roughly halves the decay exponent. In (b), no exponential closing occurs: all gaps saturate to finite values, with each additional ancilla lowering the gap only moderately (realistic parity: from 1.39 at $N = 5$ to 0.88 at $N = 13$).

As discussed in Sec. 4.2, ancilla qubits are needed to break the ‘unrealizable’ K-body constraint into smaller ones. Figure 4.1 presents an example for $N = 7$, where the dashed colored lines (red and blue) in the logical graph (a) represent auxiliary interactions between qubits that do not interact in the original model. These interactions are transformed into the red and blue ancilla physical qubits in panel (d). There are proposals for QA hardware based on superconducting qubits [202, 203] and rydberg atoms [122, 204] with the latter being capable, in principle, of realizing the configuration shown in Fig. 4.1 (d) on currently available hardware.

Figure 4.3(a) shows different scalings of the gap in the parity framework. When $n_a = 0$, there is a clear amplification of the minimal gap, visible even at a small system size $N = 5$. The gap then remains roughly unchanged with the system size

and no exponentially closing behavior is observed. Therefore, using Parity-QA to solve the frustrated ring model would in principle avoid the bottleneck associated with the exponentially closing gap. However, implementing the Hamiltonian without ancilla qubits remains beyond current hardware capabilities, as it requires a global constraint acting across all qubits. To be implementable on annealing hardware, $(N - 3)/2$ ancilla qubits are needed for N logical qubits. The black dotted line in Fig. 4.3(a) demonstrates the scaling of the minimal gap when the required number of ancilla qubits is used. With this more realistic consideration, the scaling of the spectral gap is back to an exponential decay, albeit with a smaller exponent: log-linear fits over $N = 5, \dots, 15$ give $\Delta_{\min}/J \sim e^{-0.96N}$ for standard QA and $\Delta_{\min}/J \sim e^{-0.51N}$ for the realistic parity implementation, so the parity mapping roughly halves the decay exponent.

The third variant introduced in Sec. 4.4 ($J_j/J \in \{-1, 1\}$) has no exponential closing gap. This is shown in Fig. 4.3(b). For both standard QA and ancilla-free Parity-QA ($n_a = 0$), the gap appears to saturate to a finite value. When ancillas are added ($n_a > 0$), the gap decreases but does not close exponentially and again appears to saturate. For this coupling choice, Parity-QA has a smaller gap than for the original frustrated ring; see Fig. 4.3(a). This model poses no small-gap bottleneck for standard QA or Parity QA. Nevertheless, it is instrumental for understanding the role of the constraint unitary in Parity-QAOA, as we elaborate in Sec. 4.5.2.

We now turn to analyzing the performance of Parity-QAOA on these models.

4.5.2. Parity-QAOA

We first show simulation results of the original frustrated Ising ring and its first two variants outlined in Sec. 4.4, before showing results of the third variant. Simulation details can be found in App. C.1.

Improved Controllability

Parity-QAOA simulations were performed on the original frustrated ring and on the modified cases obtained by (i) setting $J_1 = J_2 = 0.8$ and (ii) further setting $J_{N-1} = 0.85$; see Sec. 4.4. There are three distinct J_j values in the original case, four in case (i), and five in case (ii).

For $N = 7$, the ground-state probabilities are summarized in Fig. 4.4. Our simulation results show that the ground state is consistently obtained, with unit probability $P_{\text{GS}} = |\langle \tilde{\psi}_{\text{GS}} | \tilde{\psi}_{\text{P}}(\boldsymbol{\theta}^x, \boldsymbol{\theta}^c, \boldsymbol{\theta}^z) \rangle|^2 = 1$, at circuit depth P equal to the number of distinct J_j values appearing in the Hamiltonian.

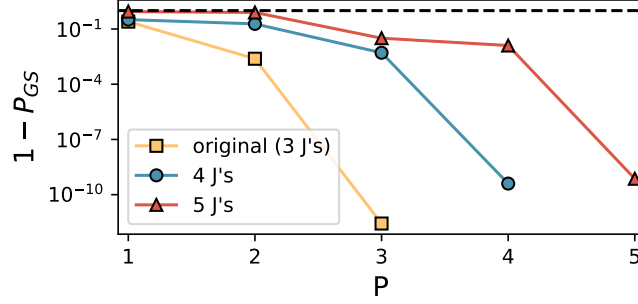


Figure 4.4.: Ground state infidelity $1 - P_{\text{GS}}$ vs circuit depth P (only single qubit z - and x -rotations) for the original formulation of the frustrated ring model and its adaptations for $N = 7$. The greater the number of distinct J_j values, the more layers P are needed to reach the ground state.

Furthermore, we observe that $\theta_p^c = 0$, for $p = 1, \dots, P$, which means that Parity-QAOA reaches the exact ground state of the frustrated ring and its variants *without* requiring the constraint unitary $\tilde{U}_c(\theta_p^c)$: It suffices to alternate applications of the local-field unitary $\tilde{U}_z(\theta_p^z)$, corresponding to single-qubit z -rotations, and the mixing unitary $\tilde{U}_x(\theta_p^x)$, corresponding to single-qubit x -rotations. This is a stark contrast to Parity-QA, where it is essential to include a non-zero constraint term $\tilde{H}_c$ in the target Hamiltonian $\tilde{H}_{\text{targ}}$ to ensure a valid outcome.

The fact that Parity-QAOA attains the ground state using only single-qubit rotations provides direct evidence that its performance is not governed by the spectral gap structure of the interpolating Hamiltonian. Rather, analogously to standard QAOA [125], the success of Parity-QAOA is determined by the controllability of the digitized dynamics and the structure of the reachable manifold, rather than by adiabatic considerations. Consistent with this, the same number of QAOA layers and the same set of optimized parameters can solve the problem across all system sizes N considered. The following sections outline this behavior and explain why Parity-QAOA can succeed without the constraint unitary U_c .

Optimal control parameters and transferability

We observe for small system sizes, $N = 7$, that Parity-QAOA consistently selects trivial constraint unitaries, with $\theta_p^c = 0$, for the frustrated Ising model. We can classically simulate the single-qubit rotations $\tilde{U}_x(\theta_p^x)$ and $\tilde{U}_z(\theta_p^z)$ to obtain a set of optimized parameters (θ^x, θ^z) that drive the system to the exact ground state, independent of system size. The resulting optimal parameters are listed in Table 4.1. The analytic, system-size-independent structure that makes this single-qubit optimization possible is derived in Appendix C.2.

Consider a single layer ($P = 1$) of the Parity-QAOA ansatz. Since the constraint

θ_1^z	θ_1^x	θ_2^z	θ_2^x	θ_3^z	θ_3^x	θ_4^z	θ_4^x	θ_5^z	θ_5^x
-9.60934341	0.20952926	-8.08672464	0.10182431	-5.23233597	0.0988746	-	-	-	-
-5.83795276	0.17247357	4.92515321	0.1598907	9.21350412	0.10889695	4.23400112	0.12517238	-	-
-0.66065018	0.18758133	-2.3049883	0.19599091	4.2849489	0.17213241	1.27555834	0.21802607	6.45839022	0.13112482

Table 4.1.: Parameter values to find the ground state for three different problem cases, in units of π . The first row shows the results for the original problem formulation in Eq. (3.2). The second row shows results for the modified ring with $J_1 = J_2 = 0.8$. For the third row, we further modified $J_{N-1} = 0.85$.

unitary is omitted here ($\theta^c = 0$), the layer consists only of $\tilde{U}_z(\theta^z)$ and $\tilde{U}_x(\theta^x)$, which act on each parity qubit separately and do not couple different qubits. Each parity qubit k therefore evolves on its own under $e^{i\theta_p^x \tilde{\sigma}_k^x}$ and $e^{i\tilde{J}_k \theta_p^z \tilde{\sigma}_k^z}$ from the initial state $|+\rangle$, and, as the ground state is the all-up configuration, $|\uparrow\rangle^{\otimes K}$, the ground-state probability factorizes into single-qubit up-probabilities,

$$P_{\text{GS}}(\theta^x, \theta^z) = \prod_{k=1}^K P_{\uparrow}(\theta^x, \tilde{J}_k \theta^z), \quad (4.13)$$

with $P_{\uparrow}(\theta^x, \tilde{J}_k \theta^z) = |\langle \uparrow | e^{i\theta_p^x \tilde{\sigma}_k^x} e^{i\tilde{J}_k \theta_p^z \tilde{\sigma}_k^z} | + \rangle|^2$ the probability that qubit k is in the $|\uparrow\rangle$ state after one layer; the derivation is given in Appendix C.2. Figure 4.5 evaluates these single-layer expressions with $\theta^x = \pi/4$ held fixed and the single parameter θ^z scanned (in units of π): the lower panel shows the single-qubit factor $P_{\uparrow}$ for each of the three distinct couplings $\tilde{J}_k = 1, 0.5, -0.45$ of the original ring, and the upper panel their product P_{GS} for $N = 5$ and $N = 51$.

Because the three couplings act as different effective rotation rates, their single-qubit factors peak at different values of θ^z , so a single layer cannot bring all qubits to the ground state simultaneously and the depth-one P_{GS} remains below one. The product is large only where the factors are simultaneously large, and is suppressed where they are not. We find that reaching $P_{\text{GS}} = 1$ requires a number of layers P equal to the number of distinct couplings, three here, which we report in Table 4.1. The product probability peaks of P_{GS} also sharpen markedly from $N = 5$ to $N = 51$, indicating that larger systems require more finely tuned parameters; this suggests that when transferring parameters optimized for a small N , additional refinement is needed. In particular, transferring parameters from smaller to larger N seems necessary to navigate barren plateaus, which become more pronounced for larger N [145].

This analysis shows that, to reach the ground state, qubits with the same $\tilde{J}_k$ should behave the same and there is no need to separate them, while individual control of qubits is needed when the $\tilde{J}_k$ are different. In fact, the whole problem can be reduced to qubits that equal the number of distinct couplings $\tilde{J}_k$ in the problem. Here, we have three distinct $\tilde{J}_k$ and each corresponding qubit accumulates a different phase under U_z , as shown in the bottom panel of Fig. 4.5. Given certain properties

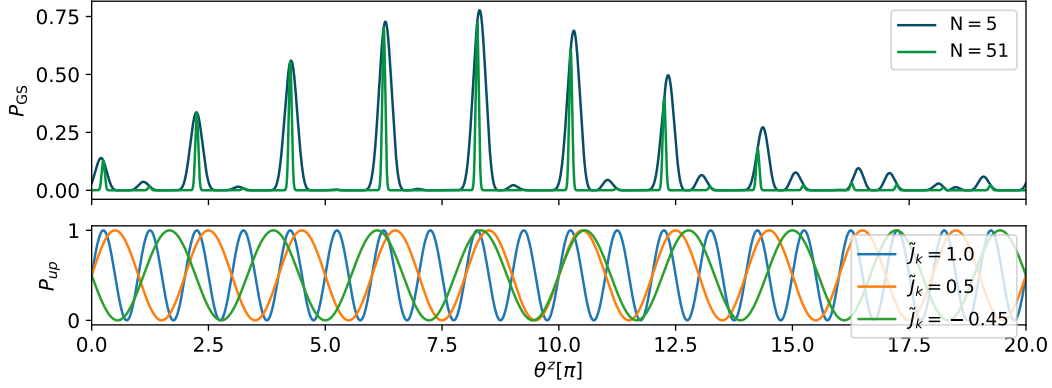


Figure 4.5.: The probability of being in the exact ground state P_{GS} (upper plot) and the probability of a single physical qubit being in the up state $P_{\uparrow}$ (lower plot) as a function of the single control parameter θ^z at fixed $\theta^x = \pi/4$ (a single QAOA layer, $P = 1$). In the upper plot we show comparisons between $N = 5$ and $N = 51$. In the lower plot we show comparisons of qubits associated with different parameters $\tilde{J}_k = 1.0, 0.5$ and -0.45 .

of the $\tilde{J}_k$ this frequency selectivity allows one to isolate actions on each qubit and, therefore, allow single qubit control. The proof in App. C.3 shows that the needed properties for arbitrary single qubit control are that the given $\tilde{J}_k$ and their absolute value are distinct, $|\tilde{J}_k| \neq |\tilde{J}_l|$ for any $k \neq l$. Hence, when $\tilde{J}_k \in \{-1, 1\}$ the qubits with $\tilde{J} = 1$ and $\tilde{J}' = -1$ cannot be separated with single-qubit rotations only and Parity-QAOA must include $\tilde{U}_c$ in this setting ($\theta_p^c \neq 0$).

Use of Constraints

We now turn to the regime in which the constraint unitary is indispensable. As described in Sec. 4.4, we set all coupling magnitudes to unity: $J_j/J = 1$ for all couplings except for a single coupling, $J_f/J = -1$. In the parity representation, the ground state is characterized by all qubits aligning with their local fields $\tilde{J}_k$, except for a single qubit, which must oppose its local field. As shown in App. C.3, when there is $J' = -J$ the qubit with the negative local field J' will always end up in the opposite state than the other one by only using a sequence of $\tilde{U}_z$ and $\tilde{U}_x$ alone. Therefore, the constraint-free mechanism identified in Sec. 4.5.2 cannot align all qubits with their ground-state orientations, making it necessary to actively employ $\tilde{U}_c$. Note also that frustration makes the ground space degenerate: the single unsatisfied bond can sit on any of the N bonds, yielding a $2N$ -fold degenerate logical ground manifold (N -fold in the parity representation). Throughout this section, we therefore define $P_{\text{GS}} = \langle \tilde{\psi}_P | \hat{P}_{\text{GS}} | \tilde{\psi}_P \rangle$, with $\hat{P}_{\text{GS}}$ the projector onto the ground manifold.

The numerical results, summarized in Fig. 4.6 (a), show P_{GS} displayed as a function of the accumulated CNOT depth (left) and CNOT gate count (right), with

successive markers corresponding to $P = 1, 2, \dots$. For standard QAOA, P_{GS} at fixed P decreases with increasing N , and the exact ground state ($P_{\text{GS}} = 1$) is first reached at $P^* = \lceil N/2 \rceil - 1$ (i.e., $P^* = 3, 5, 7, 10$ for $N = 7, 11, 15, 21$). For Parity-QAOA, by contrast, $P_{\text{GS}} > 0.99$ is reached at a constant $P^* = 4$ for all system sizes studied. The exact ground state is not attained at the depths we simulated (up to $P = 5$), but the residual error keeps decreasing with depth (e.g., $1 - P_{\text{GS}} \approx 2 \times 10^{-6}$ at $P = 5$ for $N = 11$), indicating that $P_{\text{GS}} \approx 0.99$ at $P = 4$ reflects the chosen depth rather than a limitation of the Ansatz. The optimal parameters transfer across system sizes: using the $N = 11$, $P = 4$ optimum as initial guess for a local re-optimization, we verified that $P_{\text{GS}} > 0.99$ is consistently achieved till $N = 21$. This can also be done for standard QAOA, see App. C.4.

We now combine P^* with the per-layer costs established in Sec. 4.3 (standard QAOA: $2N$ CNOTs and constant depth per layer; Parity-QAOA: $2(N - 1)$ CNOTs and depth $\mathcal{O}(N)$ per layer). Figure 4.6(b) compares the total resources required to reach $P_{\text{GS}} > 0.99$ at $N = 7, 11, 15, 21$.

For standard QAOA, the $\mathcal{O}(N)$ layers yield a total CNOT depth of $\mathcal{O}(N)$ and a total CNOT count of $\mathcal{O}(N^2)$. For Parity-QAOA, the constant number of layers yields a total depth that is also $\mathcal{O}(N)$, with a prefactor approximately twice that of standard QAOA, but a total CNOT count of only $\mathcal{O}(N)$. On NISQ hardware, where accumulated two-qubit gate errors typically dominate over coherence-time limitations, the total gate count is often the more relevant figure of merit, making this trade-off favorable for Parity-QAOA despite its depth overhead and the smaller residual $1 - P_{\text{GS}}$.

The optimal Parity-QAOA parameters at $N = 11$, $P = 4$ are shown in Fig. 4.7: the constraint unitary is actively engaged at the optimum, in direct contrast to the original frustrated ring of Sec. 4.5.2. Notably, an optimum with $\theta_1^c = 0$ exists, so that only three of the four constraint parameters are non-trivial.

4.6. Conclusions

The frustrated Ising ring is one of the simplest models exhibiting an exponentially closing spectral gap, making it a paradigmatic and stringent benchmark for quantum optimization. Its parity-compiled variant is particularly well-suited for this study because the ring topology incurs zero qubit overhead and requires fewer constraint terms than the previously studied dense graphs, placing it in a qualitatively different regime of the parity architecture. In this work, we have investigated the performance of the parity (LHZ) architecture on this model and its extensions, both in continuous-time quantum annealing and in digitized protocols based on QAOA.

In the continuous-time QA setting, we have shown that the parity mapping en-

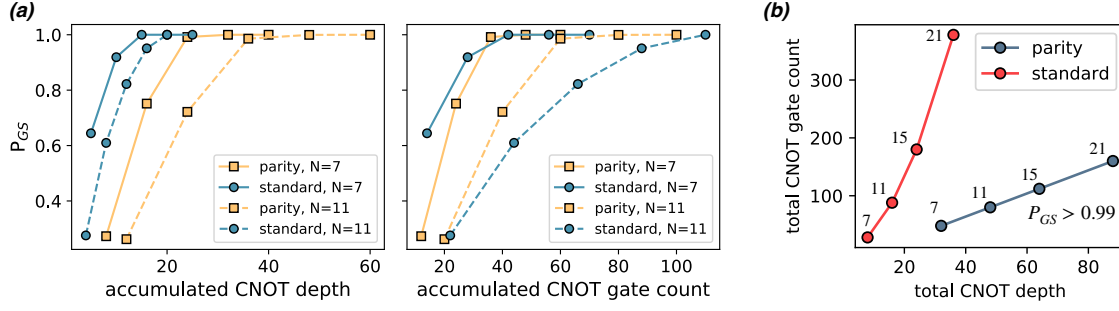


Figure 4.6.: Performance and resource scaling of standard and Parity-QAOA on the modified frustrated ring ($-J_f/J = -1$). (a) Ground-state probability P_{GS} as a function of the accumulated CNOT depth (left) and CNOT gate count (right) for $N = 7$ and 11 ; successive markers correspond to $P = 1, 2, \dots, 5$. At fixed P , P_{GS} decreases significantly with N for standard QAOA but not for Parity-QAOA. (b) Total CNOT depth and CNOT gate count required to reach $P_{GS} > 0.99$, shown for $N = 7, 11, 15, 21$. Standard QAOA needs a total depth of $\mathcal{O}(N)$ and a total gate count of $\mathcal{O}(N^2)$. Parity-QAOA needs a total depth and a total gate count that both scale as $\mathcal{O}(N)$.

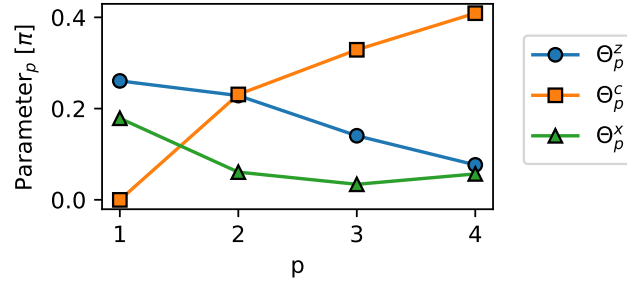


Figure 4.7.: Best found Parity-QAOA parameters ($\theta_p^z, \theta_p^c, \theta_p^x$) at each layer $p = 1, \dots, P$, for $N = 11$ and $P = 4$, for the ring with uniform coupling magnitudes and a single antiferromagnetic bond. The label [π] on the vertical axis indicates that all angles are expressed in units of π : a plotted value of 0.2 corresponds to a rotation angle of 0.2π . The constraint unitary is actively engaged, in contrast to the constraint-free regime of Sec. 4.5.2; the first constraint angle is trivial, $\theta_1^c = 0$.

larges the minimal spectral gap relative to standard linear-schedule QA. This advantage is quantifiably reduced when hardware-compatible ancilla qubits are introduced to decompose the global constraint into shorter loops: exponentially closing of the gap is restored, but with a smaller exponent. The result provides an honest, quantified picture of the spectral benefit versus hardware cost trade-off inherent to parity-based annealing on sparse graphs.

Parity-QAOA on the original frustrated ring and its symmetry-broken variants revealed an unexpected regime of *constraint-free* operation. The exact ground state is reached using only single-qubit rotations, with circuit depth equal to the number of distinct coupling values in the Hamiltonian and independent of system size. To

our knowledge, this is the first explicit instance in which the constraint unitary in Parity-QAOA is entirely unnecessary. The mechanism — qubits with distinct local fields in the parity basis have arbitrary qubit control — is concretely understood and is unlikely to work in more complex problem settings.

To probe a case in which Parity-QAOA’s constraint unitaries are necessary to prepare the ground state, we further considered a variant of the model where all coupling magnitudes are set to unity, with a single antiferromagnetic bond. In this regime, Parity-QAOA achieves $P_{\text{GS}} \approx 0.99$ with a number of layers that remains constant over the system sizes studied, while reducing the total CNOT count from $\mathcal{O}(N^2)$ to $\mathcal{O}(N)$ with respect to standard QAOA. The total circuit depth is $\mathcal{O}(N)$ in both approaches but Parity-QAOA exhibits a factor of two depth overhead. Interestingly, Ref. [124] found the opposite trade-off for complete graphs, where Parity-QAOA requires more two-qubit gates than standard QAOA but offers advantages in circuit depth. Taken together with the present results, this suggests that standard QAOA and Parity-QAOA are two complementary approaches, whose relative resource requirements strongly depend on the density of the underlying problem graph. This observation motivates extending the present analysis beyond rings to broader classes of sparse graphs, where the structure of the parity constraints and the associated ancilla requirements may also lead to qualitatively different scaling of circuit depth and two-qubit gate count.

On the experimental side, implementing Parity-QA and Parity-QAOA for these ring instances on current annealing and gate-based hardware would provide a direct test of whether the spectral gap amplification and gate-count reduction identified here survive realistic noise and connectivity constraints.

A Python code implementation of the QAOA circuits considered in this work is available online [205].

5. Conclusions

Progress in quantum computation continues to expand the range of problems for which quantum approaches are being explored. Yet current devices remain constrained, making careful algorithm design essential for using the available quantum resources effectively. This Thesis uses classical combinatorial optimization as a testbed and studies quantum annealing and hybrid quantum–classical variational algorithms. The three studies presented in Chapters 2–4 examine how the performance of the algorithms changes when varying, in turn, the annealing schedule, the independent parameters of a digitized evolution, and the problem encoding. Together, they show that these choices can open routes to ground-state preparation beyond a fixed linear annealing protocol.

Chapter 1 established the common framework for these studies. In standard QA, the adiabatic theorem relates the required evolution time to the spectral properties of the chosen Hamiltonian interpolation, so a small minimum gap can make a linear schedule ineffective. The Grover-search example shows, however, that replacing a linear schedule with one that moves quickly where the gap is large and slows near its minimum improves the runtime scaling from $O(N_s)$ to $O(\sqrt{N_s})$. Shortcuts to adiabaticity and quantum optimal control deliberately design nonadiabatic dynamics, while dQA and QAOA translate the evolution into alternating digital controls. The parity architecture further shows that an encoding changes the available generators, constraints, and hardware costs. These observations motivate assessing the complete protocol rather than only the gap of one prescribed path.

Chapter 2 first examined whether schedule optimization can overcome the poor performance of linear-schedule dQA on weighted, 3-regular MaxCut instances with $N = 14$ vertices and anomalously small annealing gaps. Smooth dQA schedules parameterized through Fourier or Chebyshev bases were compared with QAOA parameters obtained by iterative interpolation or Fourier-based procedures. At equal circuit depth, both dQA-CRAB variants improved on linear-schedule dQA. The LogINTERP and FOURIER variants of QAOA achieved comparable performance, whereas INTERP did not. Among the tested methods, dQA-C-CRAB required the least computational resources to obtain smooth, high-quality solutions.

The better solutions shared an STA-like mechanism: they did not eliminate the small gap or remain in the instantaneous ground state, but transferred population to an excited state before the minimum gap and returned it to the target ground state

at the end. Smooth solutions were also transferable among the ten hard instances studied, either directly or as initial conditions for local refinement. Interpolating them yielded continuous-time controls with low residual energy, whereas interpolating an equally competitive but irregular digital solution caused a substantial loss of performance. For these finite, noiseless instances, smoothness therefore enabled the optimized digital protocols to be transferred to continuous-time control. The resulting optimized QA protocol did not require prior knowledge of the minimum-gap location.

Chapter 3 then separated the controllability of a digitized protocol from the spectral gap of a particular annealing path, which was believed to be the bottleneck. The test case was an odd frustrated Ising ring whose minimum gap along the standard interpolation closes exponentially with N . Its free-fermion representation enabled exact simulation of the digital dynamics. With independently optimized QAOA angles, the first critical depth at which a randomly initialized run reached numerically zero residual energy was $P_1^{\text{cr}} = (N^2 - 1)/4$. Parameter counting for the accessible Bogoliubov–de Gennes dynamics explains the quadratic scaling, while the reflection symmetry gives the precise expression [125]. Thus, in this model, exact digital preparation requires a quadratic number of layers despite the exponential annealing-time bottleneck of the linear path.

As demonstrated in Chapter 2, smoothness is a valuable design principle for optimized digital protocols, and we pursued it here using dQA-CRAB. Below P_1^{cr} , dQA-CRAB outperformed randomly initialized QAOA; above this threshold, the smooth schedules continued to prepare states with high accuracy. Analysis of the instantaneous-state populations revealed an STA-like mechanism in which the protocols populated excited states before encountering the small-gap bottleneck and subsequently produced a population inversion into the final ground state. Finally, for the fixed accuracy criterion considered, the total annealing time of dQA-CRAB scaled linearly with N , compared with the quadratic scaling of the continuous-time diabatic protocol of Ref. [140]. Together with the preceding controllability analysis, these results show that an exponentially small gap along the standard interpolation does not determine the resources required by an optimized digital protocol.

Chapter 4 examined how changing the encoding alters the performance of QA and QAOA for the frustrated Ising ring. Because the ring contains as many interactions as logical spins, the parity mapping introduces no qubit overhead. In continuous-time QA, the ideal parity formulation with a single global constraint enlarges the minimum gap and shows no exponential closing over the accessible system sizes. Decomposing this constraint into hardware-compatible local terms requires ancillas and restores exponential gap closing, but reduces the fitted decay exponent from 0.96 for standard QA to 0.51 for Parity-QA. The spectral advantage therefore depends directly on how the parity constraint is implemented.

In the digital setting, Parity-QAOA reaches the exact ground state of the original ring and the variants with distinct coupling magnitudes without applying the constraint unitary: the dynamics factorizes into single-qubit rotations, and the required layer count equals the number of distinct couplings rather than growing with N . This constraint-free mechanism is model-specific; for equal-magnitude couplings with one antiferromagnetic bond, the constraint unitary is necessary. Nevertheless, Parity-QAOA attains $P_{\text{GS}} > 0.99$ in four layers for all tested sizes up to $N = 21$, reducing the total CNOT count from $\mathcal{O}(N^2)$ for standard QAOA to $\mathcal{O}(N)$, while both methods retain $\mathcal{O}(N)$ total CNOT depth and Parity-QAOA carries an approximately twofold depth overhead. So the advantage of the parity encoding depends on the constraint structure and on whether the target hardware is limited primarily by gate count, circuit depth, or qubit connectivity.

Taken together, the three studies show that, in these examples, a small gap along one annealing path does not exhaust the routes to a target state. Performance depends on the protocol as a whole: smooth schedules can exploit nonadiabatic transfer, independent digital controls reveal the reachable manifold, and alternative encodings reshape both controllability and compiled resources. This broader view suggests a natural next step: automating the joint design of these elements. Rather than optimizing parameters within a fixed protocol chosen in advance, learning-based methods could search jointly over control structures and their parameters subject to hardware constraints, then refine the resulting protocols through experimental feedback. Reinforcement learning can already refine conventional control solutions [206] and, experimentally, steer physical parameters continuously during quantum error correction [207]. Extending this closed-loop approach from calibration and error correction to many-body state preparation would integrate protocol discovery, compilation, and execution within a single design loop.

A second direction is to move beyond algorithms formulated as optimization problems. Quantum singular-value transformation (QSVT) applies polynomial transformations to the singular values of block-encoded operators and unifies primitives for Hamiltonian simulation, amplitude amplification, matrix inversion, and ground-state preparation [208]. It therefore recasts the questions studied here in terms of phase-sequence design, operator encoding, and compilation. Stochastic quantum signal processing has recently reduced query costs [209], and deep signal-processing sequences have been demonstrated experimentally, albeit with a degree–noise trade-off [210]. Hardware-aware QSVT could thus extend the control methods of this Thesis to quantum simulation and linear-algebra tasks, provided that block-encoding overhead and accumulated noise can be controlled.

A. Technical details for Chapter 2

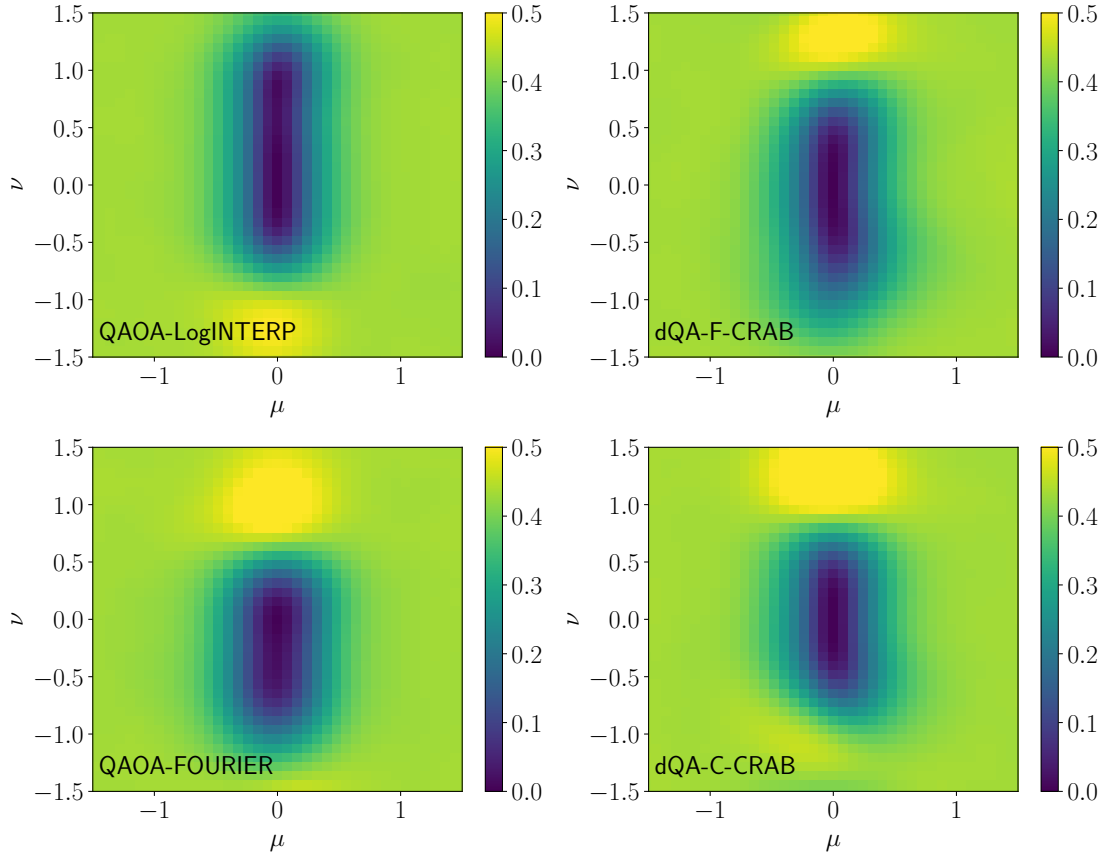


Figure A.1.: The residual energy $\epsilon_p^{\text{res}}(\theta)$ for different methods along two directions in parameter space, μ and ν . These correspond, respectively, to the direction of maximum and minimum curvature of $\epsilon_p^{\text{res}}(\theta)$ around the corresponding local minimum (see Fig. 2.5), set at $\mu = \nu = 0$.

A.1. Gradient-based optimization

Our classical simulations rely on BFGS as a gradient-based algorithm for local optimizations. To speed up the simulation runs and to avoid numerical approximations of the gradient, we calculate the analytic expression for the gradient of the variational energy in Eq. 2.11, exploiting the digitized form of the variational state.

Let C_j denote a generic parameter in the list $\mathbf{C} = (C_0, \mathbf{C}^x, \mathbf{C}^z)$. Hence:

$$\begin{aligned} \frac{\partial E_P^{\text{dQA}}}{\partial C_j} &= \langle \Psi_P^{\text{dQA}} | \hat{H}_z | \frac{\partial \Psi_P^{\text{dQA}}}{\partial C_j} \rangle + \langle \frac{\partial \Psi_P^{\text{dQA}}}{\partial C_j} | \hat{H}_z | \Psi_P^{\text{dQA}} \rangle \\ &= 2 \operatorname{Re} \langle \Psi_P^{\text{dQA}} | \hat{H}_z | \frac{\partial \Psi_P^{\text{dQA}}}{\partial C_j} \rangle . \end{aligned} \quad (\text{A.1})$$

The derivative of the state follows from the chain rule. Let us denote by θ_m^α the parameters, with $\alpha = x, z$. Then:

$$| \frac{\partial \Psi_P^{\text{dQA}}}{\partial C_j} \rangle = \sum_{\alpha=x,z} \sum_{m=1}^P | \frac{\partial \Psi_P^{\text{dQA}}}{\partial \theta_m^\alpha} \rangle \frac{\partial \theta_m^\alpha}{\partial C_j} . \quad (\text{A.2})$$

The two ingredients appearing here are simple to calculate from Eq. (2.8):

The other ingredient, the state derivative, depends on the Trotter splitting performed. For simplicity, let us focus on the lowest-order Trotter splitting. Define, for $m = 1 \cdots P$, the shorthands:

$$\hat{U}_m = \hat{U}(\theta_m^x, \theta_m^z) = e^{-i\theta_m^x \hat{H}_x} e^{-i\theta_m^z \hat{H}_z} ,$$

and

$$|\Psi_m\rangle = \hat{U}_m \cdots \hat{U}_1 |\Psi_0\rangle . \quad (\text{A.3})$$

Then:

$$i \langle \Psi_P^{\text{dQA}} | \hat{H}_z | \frac{\partial \Psi_P^{\text{dQA}}}{\partial \theta_m^x} \rangle = \langle \Psi_m | \hat{U}_{m+1}^\dagger \cdots \hat{U}_P^\dagger \hat{H}_z \hat{U}_P \cdots \hat{U}_{m+1} \hat{H}_x | \Psi_m \rangle , \quad (\text{A.4})$$

and

$$i \langle \Psi_P^{\text{dQA}} | \hat{H}_z | \frac{\partial \Psi_P^{\text{dQA}}}{\partial \theta_m^z} \rangle = \langle \Psi_{m-1} | \hat{U}_m^\dagger \cdots \hat{U}_P^\dagger \hat{H}_z \hat{U}_P \cdots \hat{U}_m \hat{H}_z | \Psi_{m-1} \rangle . \quad (\text{A.5})$$

The complete gradient can thus be calculated efficiently in a classical simulation by storing intermediate states $|\Psi_m\rangle$ and combining them to compute each partial derivative.

A.2. The FOURIER approach of Zhou et al.

We give here a few details on the implementation of the FOURIER method proposed in Ref. [95], reformulated in a notation consistent with the previous sections.

a) $N_c = P$, without randomness. In this case, one should keep the number of Fourier terms N_c equal to P and hence increase both by 1 at each step [95]. Suppose to have the optimal — most likely, *locally* optimal — Fourier coefficients $\mathbf{C}^{(*,P-1)}$ at level $P-1$. Then, take as an initial guess, for both $\mathbf{C}^z$ and $\mathbf{C}^x$:

$$\mathbf{C}^{(0,P)} = (\mathbf{C}^{(*,P-1)}, 0) , \quad (\text{A.6})$$

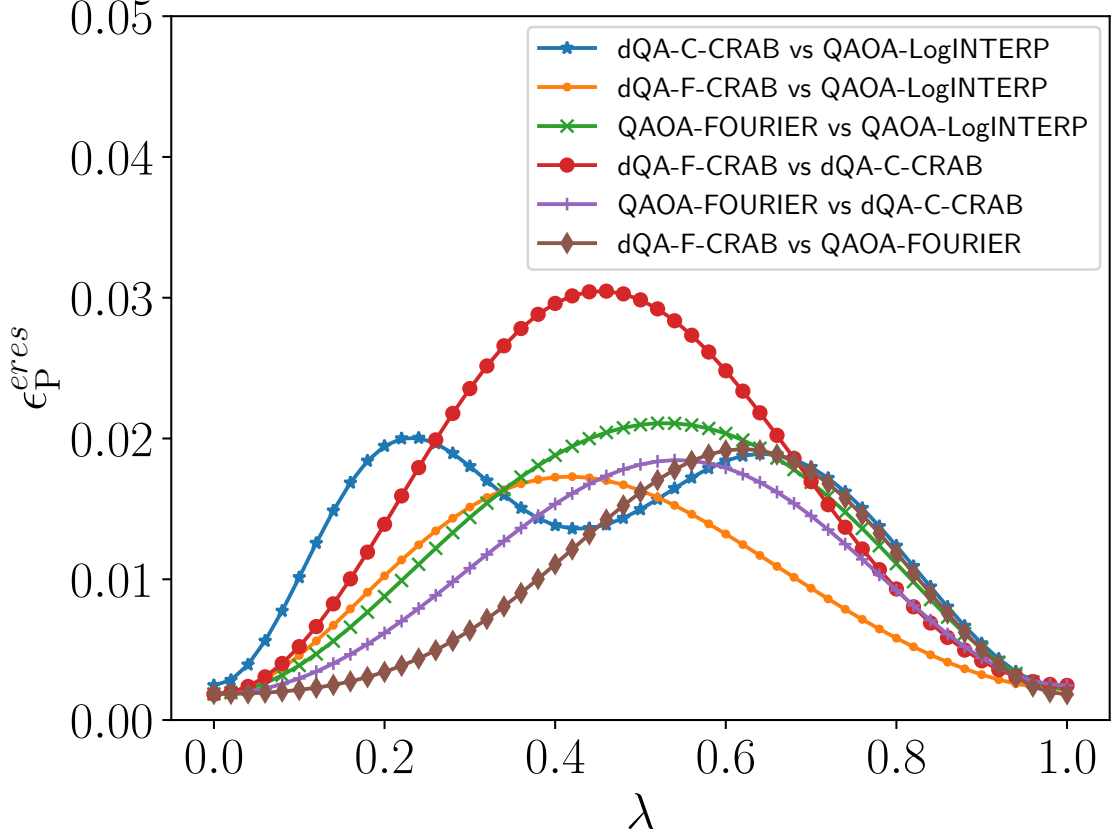


Figure A.2.: The residual energy $\epsilon_P^{\text{res}}(\theta^{12}(\lambda))$ along the direction λ linking two local minima in parameter space. The corresponding optimal parameters are shown in Fig.2.5

and determine the new minimum $\mathbf{C}^{(*,P)}$ by a local search strategy starting from $\mathbf{C}^{(0,P)}$.

b) $N_c = P$, with randomness. As in **a)** above, consider a (local) optimal solution $\mathbf{C}^{(*,P-1)}$ and construct a candidate initial point as in Eq. (A.6), with a corresponding new minimum $\mathbf{C}^{(*,P)}$. On top of that, multiple initial points are randomly generated (see below), starting from the *best* final solution $\mathbf{C}^{(\text{best},P-1)}$ kept at each P : notice that $\mathbf{C}^{(\text{best},P-1)}$ might differ from $\mathbf{C}^{(*,P-1)}$. More precisely, starting from $\mathbf{C}^{(\text{best},P-1)}$, one should generate $R > 0$ random candidate initial points, e.g., typically, $R = 10$, as follows:

$$\mathbf{C}^{(0,P)} = (\mathbf{C}^{(\text{best},P-1)} + \alpha \mathbf{r}, 0) . \quad (\text{A.7})$$

Here, α is a real coefficient, typically set to 0.6, and $\mathbf{r}$ is a $(P - 1)$ -dimensional vector of gaussian-distributed random numbers with zero mean and variance given by $r_n = (C_n^{(*,P-1)})^2$. An extra initial point is generated with $\mathbf{r} = 0$. Correspondingly, candidate final minima are found with a local optimization starting from each of these $R + 1$ initial points. The new $\mathbf{C}^{(\text{best},P)}$ at level P is selected as the best solution in a set of $R + 2$ candidates comprising the $R + 1$ minima just found, and the $\mathbf{C}^{(*,P)}$ generated as in **a)**. Note that, for the next iterative step, both $\mathbf{C}^{(*,P)}$ and

$\mathbf{C}^{(\text{best}, P)}$ need to be saved.

c) N_c fixed, with randomness. Everything is performed as in case **b)** above, only up to the value of $P = N_c$. Beyond that, N_c stays unchanged while P is further increased up to the desired P_{max} .

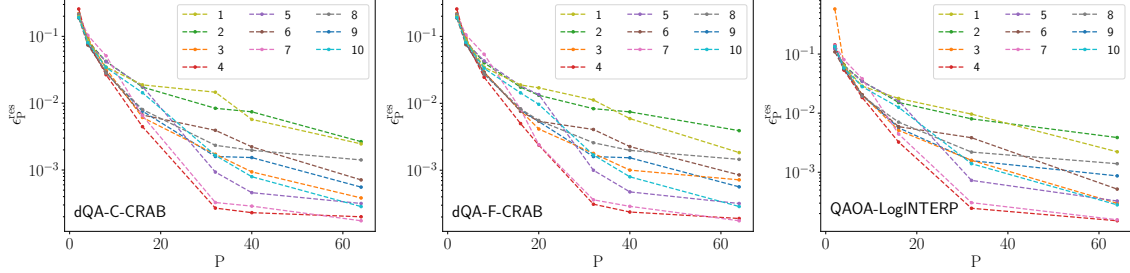


Figure A.3.: Residual energy as a function of P for the 10 hard instances we considered and for different methods. For each instance, these methods provide similar final results.

A.3. Local neighborhood of different optimal solutions

As the high complexity of the parameter landscape might hinder some connections among seemingly different local minima, in the following we further investigate this issue. Starting from the behavior of the residual energy in the neighborhood of these local minima, our goal is to determine whether the solutions we find are actually distinct in the $2P$ -dimensional parameter space. In particular, we focus on the Hessian of the residual energy $\epsilon^{\text{res}}(\boldsymbol{\theta})$ as a function of the variational parameters $\boldsymbol{\theta}$. For each method described in Section 2.2, we evaluate the Hessian of the energy in the minimum $\epsilon^{\text{res}}(\boldsymbol{\theta}^*)$ using finite differences on top of the analytical expression of the first-order gradient. The eigenvectors of the Hessian characterize the local curvature of the residual energy: the corresponding eigenvalues allow us to determine the directions where the curvature is locally maximum and minimum. In Fig. A.1, we plot $\epsilon^{\text{res}}(\boldsymbol{\theta})$ along these two directions. Specifically, we plot the residual energy as a function of the displacements μ and ν along the direction of maximum and minimum curvature, respectively. The displacements are evaluated from $\epsilon^{\text{res}}(\boldsymbol{\theta}^*)$, which is therefore located in the origin $\mu = \nu = 0$. We see that, even in the direction of minimum local curvature ν , a variation in the parameter $\boldsymbol{\theta}$ results in a sizable variation of the residual energy. Intuitively, the four solutions determined by the different methods appear to be isolated minima in the $2P$ -dimensional parameter space.

Furthermore, for each pair of local minima obtained through different methods, we focus on their convex linear combination. More in detail, given two points in

parameter space $\boldsymbol{\theta}^{(1)}$ and $\boldsymbol{\theta}^{(2)}$, we consider:

$$\boldsymbol{\theta}^{(12)}(\lambda) = \lambda \boldsymbol{\theta}^{(1)} + (1 - \lambda) \boldsymbol{\theta}^{(2)} \quad (\text{A.8})$$

and we compute the corresponding residual energy as a function of the interpolating parameter $\lambda \in [0, 1]$. We show the results in Fig. A.2: each pair of optimal solutions is separated by a barrier where the residual energy increases by an order of magnitude. This finding precludes the existence of a spherical broad minimum in the variational energy landscape where two of these optimal local minima rely. However, we stress that a complicated path in parameter space could connect two local minima without necessarily crossing a residual energy barrier.

A.4. Performance over different instances

In this section, we summarize the performance of the methods presented in the main text on the 10 hard instances we considered. In Fig. A.3, we show the residual energies for dQA-C-CRAB, dQA-F-CRAB, and QAOA-LogINTERP. Consistently with the results shown in the main text, these methods implement “shortcuts to adiabaticity” protocols and outperform the linear dQA-LIN method.

For each instance in exam, dQA-C-CRAB, dQA-F-CRAB, and QAOA-LogINTERP consistently succeed in decreasing residual energy as the circuit depth P increases.

Empirically, a linear interpolation method, QAOA-INTERP, often performs significantly worse. Finally, as expected, the residual energy of dQA-LIN reaches a plateau, stabilizing at a constant value as P increases.

B. Technical details for Chapter 3

B.1. Low-energy spectrum in the Nambu formalism

We consider here only the odd-parity case, $p = 1$. We denote by $\epsilon_\mu > 0$ the positive energy Bogoliubov-de Gennes eigenvalues [78], ordered from the smallest, ϵ_1 , to the largest. Recall that the Bogoliubov-de Gennes Hamiltonian has the diagonal form [78]:

$$\hat{H} = \sum_{\mu=1}^N \epsilon_\mu \left(\hat{\gamma}_\mu^\dagger \hat{\gamma}_\mu - \hat{\gamma}_\mu \hat{\gamma}_\mu^\dagger \right). \quad (\text{B.1})$$

Referring to Fig. 3.7(e,f), we see that there is a clear singularity in the spectrum at $s_c \approx 0.5$, with a vanishing gap in the Nambu spectrum around the particle-hole symmetry point $E = 0$, which implies a true level crossing, where $\epsilon_1(s = s_c) \equiv 0$. Hence, the (occupied) state with energy $-\epsilon_1$ for $s < s_c$ is transformed into the state at energy $+\epsilon_1 > 0$ when $s > s_c$, hence it remains occupied. Symmetrically, the (empty) state with energy $+\epsilon_1$ for $s < s_c$, is transformed into the state at energy $-\epsilon_1$ for $s > s_c$, hence it remains empty.

For $s < s_c$ the ground state is the Bogoliubov vacuum obtained by occupying all the negative energy states:

$$|\Psi_{gs}\rangle = |\emptyset\rangle = \left(\prod_{\mu=1}^N \hat{\gamma}_\mu \right) |0\rangle, \quad \implies \quad E_{gs}(s < s_c) = - \sum_{\mu=1}^N \epsilon_\mu. \quad (\text{B.2})$$

The first excited state (conserving the fermionic parity) is obtained by creating *two* fermions in the lowest two positive energy states, hence:

$$|\Psi_{ex}^{(21)}\rangle = \hat{\gamma}_2^\dagger \hat{\gamma}_1^\dagger |\emptyset\rangle, \quad \implies \quad E_{ex}^{(21)}(s < s_c) = - \sum_{\mu=1}^N \epsilon_\mu + 2(\epsilon_2 + \epsilon_1). \quad (\text{B.3})$$

Higher excited states are obtained similarly.

For $s > s_c$ the ground state is obtained by “occupying” the first positive eigenvalue. Formally, this amounts to performing a particle-hole transformation for the first level

$$\tilde{\gamma}_1 = \hat{\gamma}_1^\dagger, \quad \text{while} \quad \tilde{\gamma}_\mu = \hat{\gamma}_\mu \quad \text{for } \mu = 2, \dots, N, \quad (\text{B.4})$$

with energies transformed accordingly as $\tilde{\epsilon}_1 = -\epsilon_1 < 0$ and $\tilde{\epsilon}_\mu = \epsilon_\mu$ for $\mu = 2, \dots, N$. The ground state $|\tilde{\emptyset}\rangle$ is annihilated by the $\tilde{\gamma}_\mu$ and the expressions in Eq. (B.2)(B.3) can be directly used. For instance, the ground state energy is:

$$E_{\text{gs}}(s > s_c) = -\sum_{\mu=1}^N \tilde{\epsilon}_\mu = -\sum_{\mu=2}^N \epsilon_\mu + \epsilon_1 . \quad (\text{B.5})$$

The first excited state is:

$$|\Psi_{\text{ex}}^{(21)}\rangle = \tilde{\gamma}_2^\dagger \tilde{\gamma}_1^\dagger |\tilde{\emptyset}\rangle , \quad (\text{B.6})$$

with energy:

$$E_{\text{ex}}^{(21)}(s > s_c) = -\sum_{\mu=1}^N \tilde{\epsilon}_\mu + 2(\tilde{\epsilon}_1 + \tilde{\epsilon}_2) = -\sum_{\mu=2}^N \epsilon_\mu - \epsilon_1 + 2\epsilon_2 . \quad (\text{B.7})$$

Hence, the gap between the excited and ground state is:

$$\Delta_s^{(21)} = \begin{cases} 2(\epsilon_2 + \epsilon_1) & \text{for } s < s_c \\ 2(\epsilon_2 - \epsilon_1) & \text{for } s > s_c \end{cases} . \quad (\text{B.8})$$

This is the spectral gap shown in Fig. 3.2 for $N = 101$.

B.2. Digitized dynamics in the Nambu formalism

We derive here the expression for the digitized dynamics of the transverse field random Ising model, as implemented by the time-evolution operator:

$$\hat{U}_p^{\text{step}} = e^{-i\theta_p^x \hat{H}_x} e^{-i\theta_p^z \hat{H}_z} , \quad (\text{B.9})$$

The idea is that we move to the Heisenberg representation for the fermionic operators, in a way that closely follows the derivation of the time-dependent Bogoliubov-de Gennes (BdG) equations in Ref. [78]. Recall that for a system represented by a quadratic fermionic Hamiltonian $\hat{H}(t) = \hat{\Psi}^\dagger \mathbb{H}(t) \hat{\Psi}$, the Heisenberg operators for the fermions can be written as:

$$\hat{\Psi}_{\text{H}}(t) \equiv \begin{pmatrix} \hat{c}_{1\text{H}}(t) \\ \vdots \\ \hat{c}_{\text{NH}}(t) \\ \hat{c}_{1\text{H}}^\dagger(t) \\ \vdots \\ \hat{c}_{\text{NH}}^\dagger(t) \end{pmatrix} = \mathbb{U}(t) \begin{pmatrix} \hat{\gamma}_1 \\ \vdots \\ \hat{\gamma}_{\text{N}} \\ \hat{\gamma}_1^\dagger \\ \vdots \\ \hat{\gamma}_{\text{N}}^\dagger \end{pmatrix} , \quad (\text{B.10})$$

where the Bogoliubov operators $\hat{\gamma}_\mu$ annihilate the initial state $\hat{\gamma}_\mu|\Psi_0\rangle = 0$ and the $2N \times 2N$ unitary matrix $\mathbb{U}(t)$ satisfies the time-dependent BdG equations:

$$i\hbar \frac{d}{dt} \mathbb{U}(t) = 2 \mathbb{H}(t) \mathbb{U}(t) . \quad (\text{B.11})$$

From now on, we assume that we perform our digitized dynamics in the **odd fermion parity sector**, starting from the fully occupied fermionic state

$$|\Psi_0\rangle = |\text{F}\rangle = \prod_j \hat{c}_j^\dagger |0\rangle$$

corresponding to the ground state of $\hat{H}_x$ in Eq. (3.5) for $h = -J$. The Bogoliubov-de Gennes operators diagonalizing the initial Hamiltonian $\hat{H}_x$ are then given by $\hat{\gamma}_j = \hat{c}_j^\dagger$. With the standard [78] writing:

$$\hat{\Psi} \equiv \begin{pmatrix} \hat{c}_1 \\ \vdots \\ \hat{c}_N \\ \hat{c}_1^\dagger \\ \vdots \\ \hat{c}_N^\dagger \end{pmatrix} = \mathbb{U}_0 \begin{pmatrix} \hat{\gamma}_1 \\ \vdots \\ \hat{\gamma}_N \\ \hat{\gamma}_1^\dagger \\ \vdots \\ \hat{\gamma}_N^\dagger \end{pmatrix} ,$$

we deduce that the initial Bogoliubov transformation matrix appearing in Eq. (B.10) is:

$$\mathbb{U}_0 = \begin{pmatrix} \mathbf{0} & \mathbf{1} \\ \mathbf{1} & \mathbf{0} \end{pmatrix} . \quad (\text{B.12})$$

The unitary dynamics is such that at step $p = 1 \cdots P$ we have a state:

$$|\Psi_p\rangle = \hat{U}_p^{\text{step}} |\Psi_{p-1}\rangle , \quad (\text{B.13})$$

with the initial condition $|\Psi_0\rangle$, and $\hat{U}_p^{\text{step}}$ given by:

$$\hat{U}_p^{\text{step}} = e^{-i\theta_p^x \hat{H}_x} e^{-i\theta_p^z \hat{H}_z} . \quad (\text{B.14})$$

Such a dynamics can be seen as a series of *quantum quenches*, hence a special case of a general time-dependent $\hat{H}(t)$, for which the time-dependent Bogoliubov-de Gennes (BdG) equations (B.11) apply [78]. Consider now the single application of $e^{-i\theta \hat{H}}$ — with a *time-independent* $\hat{H}$, either $\hat{H}_z$ or $\hat{H}_x$ — to any Bogoliubov vacuum (Gaussian) state $|\Psi_0\rangle$. You can view it as the integration of a Schrödinger equation

$$i \frac{d}{d\theta} |\Psi(\theta)\rangle = \hat{H} |\Psi(\theta)\rangle , \quad (\text{B.15})$$

with initial condition $|\Psi(\theta = 0)\rangle$. The corresponding BdG equation reads:

$$i \frac{d}{d\theta} \mathbb{U}(\theta) = 2 \mathbb{H} \mathbb{U}(\theta), \quad (\text{B.16})$$

with boundary condition $\mathbb{U}(\theta = 0)$ set by $|\Psi(\theta = 0)\rangle$. Therefore, the explicit solution of the BdG equations reads:

$$\mathbb{U}(\theta) = e^{-2i\theta \mathbb{H}} \mathbb{U}(\theta = 0) . \quad (\text{B.17})$$

Let now $\mathbb{U}$ to be the unitary matrix formed by the diagonalization of $\mathbb{H}$:

$$\mathbb{H} = \mathbb{U} \mathbb{E}_{\text{diag}} \mathbb{U}^\dagger . \quad (\text{B.18})$$

Then, it is simple to show that

$$\mathbb{U}(\theta) = \mathbb{U} e^{-2i\theta \mathbb{E}_{\text{diag}}} \mathbb{U}^\dagger \mathbb{U}(\theta = 0) . \quad (\text{B.19})$$

The successive application of the two operators comprising $\hat{U}_p^{\text{step}}$ is therefore obtained from the diagonalizations, $\mathbb{H}_z = \mathbb{U}_z \mathbb{E}_{\text{diag}}^z \mathbb{U}_z^\dagger$, and $\mathbb{H}_x = \mathbb{U}_x \mathbb{E}_{\text{diag}}^x \mathbb{U}_x^\dagger$. The latter is particularly simple, as the Hamiltonian is diagonal in the original fermions,

$$\mathbb{E}_{\text{diag}}^x = \mathbb{H}_x = \begin{pmatrix} -\mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{pmatrix} .$$

Hence $\mathbb{U}_x = \mathbb{1}$, and therefore:

$$e^{-2i\theta_p^x \mathbb{H}_x} = \begin{pmatrix} e^{+2i\theta_p^x} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & e^{-2i\theta_p^x} \mathbf{1} \end{pmatrix} . \quad (\text{B.20})$$

The diagonalization of $\hat{H}_z$ in the $p = 1$ sector can be done by hand. Observe that in the original fermions, we can rewrite Eq. (3.4) as:

$$\hat{H}_z = - \sum_{j=1}^N J_j \left(\hat{c}_j^\dagger - \hat{c}_j \right) \left(\hat{c}_{j+1}^\dagger + \hat{c}_{j+1} \right) , \quad (\text{B.21})$$

where $\hat{c}_{N+1} = \hat{c}_1$. Next, observe that the BdG combinations

$$\begin{cases} \tilde{\gamma}_j &= \frac{1}{2} \left(\hat{c}_{j+1}^\dagger + \hat{c}_j^\dagger + \hat{c}_{j+1} - \hat{c}_j \right) \\ \tilde{\gamma}_j^\dagger &= \frac{1}{2} \left(\hat{c}_{j+1}^\dagger - \hat{c}_j^\dagger + \hat{c}_{j+1} + \hat{c}_j \right) \end{cases} , \quad (\text{B.22})$$

are such that

$$\begin{cases} \tilde{\gamma}_j^\dagger + \tilde{\gamma}_j &= \hat{c}_{j+1}^\dagger + \hat{c}_{j+1} \\ \tilde{\gamma}_j^\dagger - \tilde{\gamma}_j &= \hat{c}_j - \hat{c}_j^\dagger \end{cases} . \quad (\text{B.23})$$

Hence, we can write:

$$\hat{H}_z = \sum_{j=1}^N J_j \left(\tilde{\gamma}_j^\dagger - \tilde{\gamma}_j \right) \left(\tilde{\gamma}_j^\dagger + \tilde{\gamma}_j \right) = \sum_{j=1}^N J_j \left(\tilde{\gamma}_j^\dagger \tilde{\gamma}_j - \tilde{\gamma}_j \tilde{\gamma}_j^\dagger \right) = 2 \sum_{j=1}^N J_j \tilde{\gamma}_j^\dagger \tilde{\gamma}_j - \sum_{j=1}^N J_j . \quad (\text{B.24})$$

Since $J_N = -J_f < 0$, the ground state energy is correct: $E_{\text{gs}} = -\sum_{j=1}^N J_j = -(N-3)J - 2J_w + J_f$. These $\tilde{\gamma}_j$ are precisely the particle-hole transformed operators invoked in Eq. (B.4), and the ground state $|\tilde{\emptyset}\rangle$ of $\hat{H}_z$ is their vacuum state, $\tilde{\gamma}_j|\tilde{\emptyset}\rangle = 0$. The BdG eigenvectors are conveniently organized as follows [78]. If you define the $N \times N$ blocks $\mathbf{U}_z$ and $\mathbf{V}_z$ of the $2N \times 2N$ BdG matrix $\mathbb{U}_z$:

$$\mathbb{U}_z = \begin{pmatrix} \mathbf{U}_z & \mathbf{V}_z^* \\ \mathbf{V}_z & \mathbf{U}_z^* \end{pmatrix}, \quad (\text{B.25})$$

then, Eq. (B.22) tells us that:

$$\mathbf{U}_z = \begin{pmatrix} -\frac{1}{2} & 0 & 0 & \cdots & 0 & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & 0 & \cdots & \vdots & 0 \\ 0 & \frac{1}{2} & -\frac{1}{2} & \cdots & \vdots & 0 \\ 0 & 0 & \frac{1}{2} & \cdots & \vdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -\frac{1}{2} & 0 \\ 0 & 0 & 0 & \cdots & \frac{1}{2} & -\frac{1}{2} \end{pmatrix}, \quad \text{and} \quad (\text{B.26})$$

$$\mathbf{V}_z = \begin{pmatrix} \frac{1}{2} & 0 & 0 & \cdots & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 & \cdots & \vdots & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & \cdots & \vdots & 0 \\ 0 & 0 & \frac{1}{2} & \cdots & \vdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \frac{1}{2} & 0 \\ 0 & 0 & 0 & \cdots & \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

The application of $\hat{U}_p^{\text{step}} = e^{-i\theta_p^x \hat{H}_x} e^{-i\theta_p^z \hat{H}_z}$ is therefore carried out by:

$$\mathbb{U}_p^{\text{step}} = e^{-2i\theta_p^x \mathbb{H}_x} \mathbb{U}_z e^{-2i\theta_p^z \mathbb{E}_{\text{diag}}^z} \mathbb{U}_z^\dagger. \quad (\text{B.27})$$

By carrying out the algebra in Eq. (B.27), the final block form of $\mathbb{U}_p^{\text{step}}$ is found to be:

$$\mathbb{U}_p^{\text{step}} = \begin{pmatrix} \mathbf{U}(\theta_p^x, \theta_p^z) & \mathbf{V}^*(\theta_p^x, \theta_p^z) \\ \mathbf{V}(\theta_p^x, \theta_p^z) & \mathbf{U}^*(\theta_p^x, \theta_p^z) \end{pmatrix}, \quad (\text{B.28})$$

with:

$$\begin{aligned} \mathbf{U}(\theta_p^x, \theta_p^z) &= e^{+2i\theta_p^x} \left(\mathbf{U}_z \mathbf{D}_p \mathbf{U}_z^\dagger + \mathbf{V}_z^* \mathbf{D}_p^* \mathbf{V}_z^\dagger \right) \\ \mathbf{V}(\theta_p^x, \theta_p^z) &= e^{-2i\theta_p^x} \left(\mathbf{V}_z \mathbf{D}_p \mathbf{U}_z^\dagger + \mathbf{U}_z^* \mathbf{D}_p^* \mathbf{V}_z^\dagger \right), \end{aligned} \quad (\text{B.29})$$

and:

$$\mathbf{D}_p(\theta_p^z) = e^{-2i\theta_p^z \text{diag}(\epsilon_\mu^z)} = \begin{pmatrix} e^{-2i\theta_p^z J_1} & \dots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & e^{-2i\theta_p^z J_{N-1}} & 0 \\ 0 & \dots & 0 & e^{-2i\theta_p^z J_N} \end{pmatrix}, \quad (\text{B.30})$$

where we should keep in mind that $J_N = -J_f$, while all the other couplings are positive.

The final solution for the digitized dynamics can be summarized as follows:

$$\mathbb{U}_P = \mathbb{U}_P^{\text{step}} \mathbb{U}_{P-1}^{\text{step}} \dots \mathbb{U}_2^{\text{step}} \mathbb{U}_1^{\text{step}} \mathbb{U}_0, \quad (\text{B.31})$$

where

$$\mathbb{U}_0 = \begin{pmatrix} \mathbf{0} & \mathbf{1} \\ \mathbf{1} & \mathbf{0} \end{pmatrix}, \quad (\text{B.32})$$

as set by the initial state of the computation.

B.3. Residual energy

We can quantify the quality of the solution by the residual energy:

$$\epsilon_p^{\text{res}} = \frac{1}{N} \left(\langle \Psi_p | \hat{H}_z | \Psi_p \rangle - E_{\text{gs}} \right), \quad (\text{B.33})$$

where E_{gs} is the target ground state energy, which for the frustrated-ring model is given by

$$E_{\text{gs}} = - \sum_{j=1}^N J_j = - \sum_{j=1}^{N-1} J_j + J_f,$$

with $J_N = -J_f$. Hence, we can write:

$$\epsilon_p^{\text{res}} = \frac{1}{N} \sum_{j=1}^N J_j \langle \Psi_p | (1 - \hat{\sigma}_j^z \hat{\sigma}_{j+1}^z) | \Psi_p \rangle. \quad (\text{B.34})$$

In terms of Jordan-Wigner fermions we have that:

$$\langle \Psi_p | \hat{\sigma}_j^z \hat{\sigma}_{j+1}^z | \Psi_p \rangle \rightarrow \langle \Psi_p | (\hat{c}_j^\dagger \hat{c}_{j+1} + \hat{c}_j^\dagger \hat{c}_{j+1}^\dagger + \text{H.c.}) | \Psi_p \rangle.$$

One can define the elementary Green's functions:

$$(\mathbf{G}_p)_{j,j'} = \langle \Psi_p | \hat{c}_j \hat{c}_{j'}^\dagger | \Psi_p \rangle, \quad (\mathbf{F}_p)_{j,j'} = \langle \Psi_p | \hat{c}_j \hat{c}_{j'} | \Psi_p \rangle. \quad (\text{B.35})$$

From the t-dependent BdG equations [78], we can express both Green's functions in terms of the blocks of the matrix

$$\mathbb{U}_p = \mathbb{U}_p^{\text{step}} \mathbb{U}_{p-1}^{\text{step}} \dots \mathbb{U}_1^{\text{step}} \mathbb{U}_0 = \begin{pmatrix} \mathbf{U}_p & \mathbf{V}_p^* \\ \mathbf{V}_p & \mathbf{U}_p^* \end{pmatrix}.$$

Indeed:

$$\mathbf{G}_p = \mathbf{U}_p \mathbf{U}_p^\dagger, \quad \mathbf{F}_p = \mathbf{U}_p \mathbf{V}_p^\dagger. \quad (\text{B.36})$$

Hence, we can easily calculate

$$-\langle \Psi_p | \hat{\sigma}_j^z \hat{\sigma}_{j+1}^z | \Psi_p \rangle \rightarrow \left((\mathbf{F}_p + \mathbf{G}_p)_{j,j+1} + \text{c.c.} \right). \quad (\text{B.37})$$

From this expression, the residual energy follows by appropriate sums over j , as implied by Eq. (B.34).

B.4. Gradients of the residual energy

We now consider the residual energy ϵ_p^{res} , which is a function of the 2P variational parameters (θ^x, θ^z) . We would like to calculate the 2P gradients of such a quantity, so as to use gradient-based optimization algorithms. If θ_p^α , with $\alpha = x, z$ denotes the generic variable, we need to calculate:

$$\begin{aligned} -\frac{\partial}{\partial \theta_p^\alpha} \langle \Psi_P | \hat{\sigma}_j^z \hat{\sigma}_{j+1}^z | \Psi_P \rangle &\rightarrow \frac{\partial}{\partial \theta_p^\alpha} \left(\mathbf{U}_P (\mathbf{U}_P^\dagger + \mathbf{V}_P^\dagger) \right)_{j,j+1} + \text{c.c.} \\ &= \left(\left(\frac{\partial}{\partial \theta_p^\alpha} \mathbf{U}_P \right) (\mathbf{U}_P^\dagger + \mathbf{V}_P^\dagger) + \mathbf{U}_P \frac{\partial}{\partial \theta_p^\alpha} (\mathbf{U}_P^\dagger + \mathbf{V}_P^\dagger) \right)_{j,j+1} + \text{c.c.} . \end{aligned} \quad (\text{B.38})$$

The relevant derivative of $\mathbf{U}_P$ and $\mathbf{V}_P$ are obtained from the corresponding blocks of

$$\frac{\partial}{\partial \theta_p^\alpha} \mathbb{U}_P = \mathbb{U}_P^{\text{step}} \mathbb{U}_{P-1}^{\text{step}} \cdots \left(\frac{\partial}{\partial \theta_p^\alpha} \mathbb{U}_p^{\text{step}} \right) \cdots \mathbb{U}_1^{\text{step}} \mathbb{U}_0, \quad (\text{B.39})$$

which in turn follows from the blocks derivative:

$$\frac{\partial}{\partial \theta_p^\alpha} \mathbb{U}_p^{\text{step}} = \begin{pmatrix} \frac{\partial}{\partial \theta_p^\alpha} \mathbf{U}(\theta_p^x, \theta_p^z) & \frac{\partial}{\partial \theta_p^\alpha} \mathbf{V}^*(\theta_p^x, \theta_p^z) \\ \frac{\partial}{\partial \theta_p^\alpha} \mathbf{V}(\theta_p^x, \theta_p^z) & \frac{\partial}{\partial \theta_p^\alpha} \mathbf{U}^*(\theta_p^x, \theta_p^z) \end{pmatrix}. \quad (\text{B.40})$$

More precisely, one can perform the calculation explicitly. For the derivative with respect to θ_p^x the result is simple:

$$\frac{\partial}{\partial \theta_p^x} \mathbf{U}(\theta_p^x, \theta_p^z) = 2i \mathbf{U}(\theta_p^x, \theta_p^z), \quad \frac{\partial}{\partial \theta_p^x} \mathbf{V}(\theta_p^x, \theta_p^z) = -2i \mathbf{V}(\theta_p^x, \theta_p^z).$$

For the derivative with respect to θ_p^z we have:

$$\begin{aligned} \frac{\partial}{\partial \theta_p^z} \mathbf{U}(\theta_p^x, \theta_p^z) &= e^{+2i\theta_p^x} \left(\mathbf{U}_z \frac{\partial}{\partial \theta_p^z} \mathbf{D}_p \mathbf{U}_z^\dagger + \mathbf{V}_z^* \frac{\partial}{\partial \theta_p^z} \mathbf{D}_p^* \mathbf{V}_z^\dagger \right) \\ \frac{\partial}{\partial \theta_p^z} \mathbf{V}(\theta_p^x, \theta_p^z) &= e^{-2i\theta_p^x} \left(\mathbf{V}_z \frac{\partial}{\partial \theta_p^z} \mathbf{D}_p \mathbf{U}_z^\dagger + \mathbf{U}_z^* \frac{\partial}{\partial \theta_p^z} \mathbf{D}_p^* \mathbf{V}_z^\dagger \right), \end{aligned}$$

where the derivative $\frac{\partial}{\partial \theta_p^z} \mathbf{D}_p(\theta_p^z)$ follows directly from Eq. (B.30).

C. Technical details for Chapter 4

C.1. Simulation details

To calculate the minimal Energy gap $\Delta_{\min}$ of Eq. (4.12) we need to get the instantaneous energy spectra of the interpolation Hamiltonian. We can do this by simulating the time evolution of the time dependent Hamiltonian using QuTiP [211, 212]. However, this is only efficient for $N < 14$. For bigger systems, to get the low lying eigenenergies, we use the iterative eigensolver (*eigsh*) from SciPY [213]. Rather than constructing the Hamiltonian as a dense matrix, it was represented as a linear operator. This allows the Lanczos algorithm underlying function *eigsh* to determine the lowest eigenvalues and corresponding eigenvectors using only repeated matrix-vector products, avoiding the storage and manipulation of the full Hamiltonian matrix.

All QAOA simulations are exact state-vector simulations performed with QISKIT [214]. For all simulations the parameter domains were restricted to $\theta_p^x \in [0, \pi/4]$, and $\theta_p^c \in [-\pi/2, \pi/2]$. For the first variants, the parameter domain for θ_p^z was set to $\theta_p^z \in [-10\pi, 10\pi]$ and for the third variant (as $\tilde{J}_k = |1\rangle$) to $\theta_p^z \in [-\pi/2, \pi/2]$. The constraint strengths were set to $C = J$ for the first variants and $C = 2J$ for the third variant where $\tilde{J}_k \in \{-1, 1\}$. For each problem instance, we used 100 random restarts of the full parameter vectors (θ^x, θ^z) for standard QAOA and $(\theta^x, \theta^z, \theta^c)$ for Parity-QAOA and optimized the parameters using the BFGS [188] routine from python SciPY [213], retaining the best result.

C.2. Ground state probability for Parity-QAOA at arbitrary size

In the original formulation of the frustrated ring model, Parity-QAOA uses single-qubit rotations only to get to the ground state. The needed QAOA parameters can be found in Table 4.1. They are obtained as follows: Using the best found parameters $(\theta^{x*}, \theta^{z*})$ obtained from QAOA simulations $N = 7$ of QAOA simulations one can use a classical algorithm to refine these parameters (BFGS search) so that the ground state probability $P_{\text{GS}} = 1$ also for large $N > 10^6$. The classical algorithm works the

following way: The unitaries for single qubit rotations are

$$\tilde{U}_x(\theta_p^x) = \prod_{k=1}^K e^{i\theta_p^x \tilde{\sigma}_k^x} = \prod_{k=1}^K (\cos(\theta_p^x) \mathbf{1}^k + i \sin(\theta_p^x) \tilde{\sigma}_k^x), \quad (\text{C.1})$$

$$\tilde{U}_z(\theta_p^z) = \prod_{k=1}^K e^{i\tilde{J}_k \theta_p^z \tilde{\sigma}_k^z} = \prod_{k=1}^K (\cos(\tilde{J}_k \theta_p^z) \mathbf{1}^k + i \sin(\tilde{J}_k \theta_p^z) \tilde{\sigma}_k^z). \quad (\text{C.2})$$

While $\tilde{U}_x(\theta_p^x)$ represents a rotation around the x-axis which acts in the same way on every physical qubit k , the unitary $\tilde{U}_z(\theta_p^z)$ is a product of independent rotations around the z-axis which, however, depend on the local field $\tilde{J}_k$ of the physical qubit k . Since all physical qubits are independent, let us concentrate our attention on qubit k . The state for a qubit at site k starting in $|+\rangle_k = (|\uparrow\rangle + |\downarrow\rangle)/\sqrt{2}$ after P layers of single-qubit rotations is

$$|\tilde{\psi}_k(\boldsymbol{\theta}^x, \tilde{J}_k \boldsymbol{\theta}^z)\rangle = e^{i\theta_P^x \tilde{\sigma}_k^x} e^{i\tilde{J}_k \theta_P^z \tilde{\sigma}_k^z} \dots e^{i\theta_1^x \tilde{\sigma}_k^x} e^{i\tilde{J}_k \theta_1^z \tilde{\sigma}_k^z} |+\rangle_k. \quad (\text{C.3})$$

The ground state has all physical qubits in the up state $|\uparrow\rangle$. One can calculate the probability that qubit k is in the up state as $P_\uparrow(\boldsymbol{\theta}^x, \tilde{J}_k \boldsymbol{\theta}^z) = |\langle \uparrow | \tilde{\psi}_k(\boldsymbol{\theta}^x, \tilde{J}_k \boldsymbol{\theta}^z) \rangle|^2$. The total probability of being in the ground state can be obtained by multiplying $P_\uparrow(\boldsymbol{\theta}^x, \tilde{J}_k \boldsymbol{\theta}^z)$ for all K qubits:

$$P_{\text{GS}}(\boldsymbol{\theta}^x, \boldsymbol{\theta}^z) = \prod_{k=1}^K P_\uparrow(\boldsymbol{\theta}^x, \tilde{J}_k \boldsymbol{\theta}^z). \quad (\text{C.4})$$

In the case of the original formulation with three different J_j values ($N-3$ qubits with $\tilde{J}_k = J$, two qubits with $\tilde{J}_k = J_w$ and one qubit with $\tilde{J}_k = -J_f$), the simplified formula for the ground state probability is

$$P_{\text{GS}}(\boldsymbol{\theta}^x, \boldsymbol{\theta}^z) = P_\uparrow^{N-3}(\boldsymbol{\theta}^x, J \boldsymbol{\theta}^z) P_\uparrow^2(\boldsymbol{\theta}^x, J_w \boldsymbol{\theta}^z) P_\uparrow(\boldsymbol{\theta}^x, -J_f \boldsymbol{\theta}^z). \quad (\text{C.5})$$

The aim is to find the parameters $(\boldsymbol{\theta}^{x*}, \boldsymbol{\theta}^{z*})$ so that $P_{\text{GS}}(\boldsymbol{\theta}^{x*}, \boldsymbol{\theta}^{z*}) = 1$. The optimized parameters $(\boldsymbol{\theta}^{x*}, \boldsymbol{\theta}^{z*})$ found in this way are exactly those reported in Table 4.1 (first row, for the original three- J ring), and Eq. (C.5) is precisely what is plotted in Fig. 4.5: the lower panel shows the individual single-qubit probabilities $P_\uparrow(\boldsymbol{\theta}^x, \tilde{J}_k \boldsymbol{\theta}^z)$ for $\tilde{J}_k = 1, 0.5, -0.45$, and the upper panel their product P_{GS} , so that the exact ground state is reached at parameter values where all three factors are simultaneously maximal.

C.3. Local controllability

In the original formulation of the frustrated ring there are three different coupling strengths, $J_1 = 1$, $J_2 = 0.5$ and $J_{N-1} = -0.45$. In the Parity mapping the couplings

become local fields and, as shown in the main text, single qubit rotations are enough to find the ground state (all qubits up).

To be able to get the ground state one needs to align the qubits with the fields J_1 and J_2 to their local field, while the qubit with the negative field J_{N-1} has to oppose its field. Therefore, it is necessary to be able to address qubits with different local fields J_k individually. Obviously, it is not possible to address qubits with the same local field separately, and here it is also not necessary as we want qubits with the same field J_k to behave the same.

The following proof is based on the dynamical Lie algebra formalism widely used in optimal quantum control. The key step reduces the controllability condition to the invertibility of a Vandermonde matrix [215]. This proof does not say how many layers of single qubit rotations are needed to achieve individual control, just that it is possible.

Consider the Hamiltonian $H_x = \sum_k \sigma_k^x$ and $H_z = \sum_k J_k \sigma_k^z$, where we only use M distinct J_k (in the above example $M = 3$). To determine the set of Hamiltonians that can be synthesized from the available controls, we consider the Lie algebra generated by H_z and H_x , which is obtained by forming consecutive commutators. First $[H_z, H_x] = 2i \sum_k J_k \sigma_k^y$ and further $[H_z, [H_z, H_x]] = -4 \sum_k J_k^2 \sigma_k^x$. Continuing this yields Hamiltonians proportional to $J_k^m \sigma_k^P$ where $P = x$ when i is even and $P = y$ when m is odd. First, we focus on the even powers $H^{(m-1)} = \sum_k J_k^{2m} \sigma_k^x$ (omitting the unnecessary prefactors $(2i)^{2m}$ for simplicity) and write this equation as

$$\begin{pmatrix} H^{(0)} \\ H^{(1)} \\ \vdots \\ H^{(M-1)} \end{pmatrix} = V_{\text{even}} \begin{pmatrix} \sigma_1^x \\ \sigma_2^x \\ \vdots \\ \sigma_M^x \end{pmatrix}, \quad (\text{C.6})$$

where the coefficients form a Vandermonde matrix V_{even} given as

$$V_{\text{even}} = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ J_1^2 & J_2^2 & \cdots & J_K^2 \\ \vdots & \vdots & & \vdots \\ J_1^{(2M-2)} & J_2^{(2M-2)} & \cdots & J_K^{(2M-2)} \end{pmatrix}. \quad (\text{C.7})$$

Here, $\det(V_{\text{even}}) \neq 0$, as the J_k and their powers J_k^m are distinct, which means that V_{even} is invertible. Similarly, for odd powers we can write the equation $H^{(m-1)} = \sum_k J_k^{2m-1} \sigma_k^y$, where the resulting Vandermonde matrix is

$$V_{\text{odd}} = \begin{pmatrix} J_1 & J_2 & \cdots & J_K \\ J_1^3 & J_2^3 & \cdots & J_K^3 \\ \vdots & \vdots & & \vdots \\ J_1^{(2M-1)} & J_2^{(2M-1)} & \cdots & J_K^{(2M-1)} \end{pmatrix}. \quad (\text{C.8})$$

The matrix V_{odd} can be inverted if $|J_k| \neq |J_l|$ for $k \neq l$. The absolute value is important because when $J_k = -J_l$ the powers $J_k^{(m)}$ are distinct but the columns k and l only differ by a constant factor -1 which makes them dependent and, therefore, the matrix would not be invertible. Therefore, for both matrices V_{even} and V_{odd} the sufficient condition for invertibility is that $|J_k| \neq |J_l|$ for $k \neq l$.

Using the invertible Vandermonde matrix V_{even} in Eq. (C.6), we can rewrite it to

$$\begin{pmatrix} \sigma_1^x \\ \sigma_2^x \\ \vdots \\ \sigma_M^x \end{pmatrix} = V_{even}^{-1} \begin{pmatrix} H_x^{(0)} \\ H_x^{(1)} \\ \vdots \\ H_x^{(M-1)} \end{pmatrix} \quad (\text{C.9})$$

This equations shows that every individual operator σ_k^x can be expressed as a linear combination of the generated Hamiltonians. As $[H_z, \sigma_k^x] = 2iJ_k\sigma_k^y$ and $[\sigma_k^y, \sigma_k^x] = 2i\sigma_k^z$ the full set $\{\sigma_k^x, \sigma_k^y, \sigma_k^z\}$ can be generated for every qubit k , which implies arbitrary single-qubit control. This proof also works when considering $J_{N-1} = 0.8$ and $J_4 = 0.85$ in addition to the previous mentioned J_k as all $|J_k|$ are different.

As mentioned, in the case of opposite signs of the local fields $J_1 = 1$ and $J_2 = -J_1$ the above proof fails as V cannot be inverted. In fact, one can show that qubits with the negative local field J_2 will always be flipped compared to the ones with the positive field J_1 if only single qubit rotations are used. This can be proved the following way: Consider, for simplicity, only two qubit 1 and 2 and the following possible single qubit rotations: a x-rotation $R_x(\theta^x) = e^{-i\theta^x\sigma^x}$ that is the same for both qubits, and a z-rotation that is $R_z(\theta^z) = e^{-i\theta^z\sigma^z}$ for qubit 1 and $R_z(-\theta^z)$ for qubit 2 (because of the opposite signs of their local fields J_k).

Assume, qubit 1 experiences the following p sequences of single qubit rotations

$$|\psi\rangle_1 = R_x(\theta_p^x)R_z(\theta_p^z) \dots R_x(\theta_1^x)R_z(\theta_1^z)|+\rangle_1. \quad (\text{C.10})$$

For qubit 2 the same sequence leads to the following variational state

$$|\psi\rangle_2 = R_x(\theta_p^x)R_z(-\theta_p^z) \dots R_x(\theta_1^x)R_z(-\theta_1^z)|+\rangle_2. \quad (\text{C.11})$$

One can exploit the fact that $\sigma^x\sigma^z\sigma^x = -\sigma^z$ to make the transformation $R_z(-\theta^z) = \sigma^x R_z(\theta^z)\sigma^x$. Also, $R_x(\theta^x) = \sigma^x R_x(\theta^x)\sigma^x$. Substituting this into the previous equation gives

$$|\psi\rangle_2 = (\sigma^x R_x(\theta_p^x)\sigma^x)(\sigma^x R_z(\theta_p^z)\sigma^x) \dots (\sigma^x R_x(\theta_1^x)\sigma^x)(\sigma^x R_z(\theta_1^z)\sigma^x)|+\rangle_2. \quad (\text{C.12})$$

Neighboring σ^x cancel as $\sigma^x\sigma^x = \mathbf{1}$, which leaves σ^x only at the start and end of the sequence. Considering that $\sigma^x|+\rangle = |+\rangle$ also removes the σ^x at the start. Hence, the equation further simplifies to

$$|\psi\rangle_2 = \sigma^x R_x(\theta_p^x) R_z(\theta_p^z) \dots R_x(\theta_1^x) R_z(\theta_1^z) |+\rangle_2. \quad (\text{C.13})$$

Comparing this variational state, Eq. (C.13), with the variational state of qubit 1, Eq. (C.10), we see that they differ by one σ^x in the end of the single qubit rotation sequence. This means that qubit 2 is always flipped compared to qubit 1 - it is not possible to separate the qubits and address them individually. Hence, if the local fields are opposite, no sequence of single-qubit rotations can rotate both qubits into the same final state.

C.4. Optimal parameters for standard QAOA

In the main text, Sec. 4.5.2, simulation results for the ring with uniform couplings $-1, 1$ were presented. The results showed that standard QAOA can find the solution, $P_{GS} = 1$, when $P^* = \lceil N/2 \rceil - 1$. Further, we observed that the values of θ_p^z are related to those of θ_p^x . In particular, $\theta_i^x = \theta_j^z$ for $i = 1, 2, \dots, P^*$ and $j = P^*, P^* - 1, \dots, 1$ (θ_p^z follows the inverse order of θ_p^x). For simplicity, we show the best parameters θ_p^x found in the simulations ($N = 5 - 11$) in Fig. C.1. The pattern can be exploited to extrapolate to the parameters needed for $N = 15, 21$. Refining the results with local optimization (using QISKIT), as described in the main text Sec. 4.5.2, one can guarantee $P_{GS} > 0.99$. Those parameters are displayed in the Figure with dashed lines.

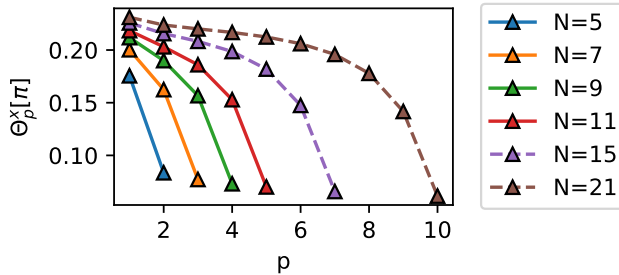


Figure C.1.: Best found QAOA parameters θ_p^x at each layer $p = 1, \dots, P$, for varying N to achieve $P_{GS} > 0.99$ for the ring with uniform coupling magnitudes and a single antiferromagnetic bond. Parameters for $N = 5 - 11$ (solid line) are obtained by taking the best of several optimization runs, parameters for $N = 15, 21$ (dashed lines) are obtained from extrapolation and local optimization. The corresponding parameters θ_p^z follow the inverse order of θ_p^x .

Bibliography

- [1] W. Heisenberg, *Über quantentheoretische Umdeutung kinematischer und mechanischer Beziehungen*, *Zeitschrift für Physik* **33**, 879 (1925).
- [2] E. Schrödinger, *Quantisierung als Eigenwertproblem*, *Annalen der Physik* **384**, 361 (1926).
- [3] M. Born, *Zur quantenmechanik der stoßvorgänge*, *Z. Phys.* **37**, 863 (1926).
- [4] M. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, 2000).
- [5] R. P. Feynman, *Simulating physics with computers*, *Int J Theor Phys* **21**, 467 (1982).
- [6] D. Deutsch, *Quantum theory, the church-turing principle and the universal quantum computer*, *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences* (1985).
- [7] P. W. Shor, *Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer*, *SIAM Journal on Computing* **26**, 1484 (1997).
- [8] J. I. Cirac and P. Zoller, *Quantum computations with cold trapped ions*, *Phys. Rev. Lett.* **74**, 4091 (1995).
- [9] C. D. Bruzewicz, J. Chiaverini, R. McConnell, and J. M. Sage, *Trapped-ion quantum computing: Progress and challenges*, *Appl. Phys. Rev.* **6**, 021314 (2019).
- [10] A. Ransford, M. S. Allman, J. Arkinstall, *et al.*, *A 98-qubit trapped-ion quantum computer with all-to-all connectivity*, *Nature* **655**, 81 (2026).
- [11] J. Koch, T. M. Yu, J. Gambetta, A. A. Houck, D. I. Schuster, J. Majer, A. Blais, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, *Charge-insensitive qubit design derived from the Cooper pair box*, *Phys. Rev. A* **76**, 042319 (2007).
- [12] M. Kjaergaard, M. E. Schwartz, J. Braumüller, P. Krantz, J. I.-J. Wang, S. Gustavsson, and W. D. Oliver, *Superconducting qubits: Current state of play*, *Annu. Rev. Condens. Matter Phys.* **11**, 369 (2020).

- [13] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. S. L. Brandao, D. A. Buell, B. Burkett, Y. Chen, Z. Chen, B. Chiaro, R. Collins, W. Courtney, A. Dunsworth, E. Farhi, B. Foxen, A. Fowler, C. Gidney, M. Giustina, R. Graff, K. Guerin, S. Habegger, M. P. Harrigan, M. J. Hartmann, A. Ho, M. Hoffmann, T. Huang, T. S. Humble, S. V. Isakov, E. Jeffrey, Z. Jiang, D. Kafri, K. Kechedzhi, J. Kelly, P. V. Klimov, S. Knysh, A. Korotkov, F. Kostritsa, D. Landhuis, M. Lindmark, E. Lucero, D. Lyakh, S. Mandrà, J. R. McClean, M. McEwen, A. Megrant, X. Mi, K. Michielsen, M. Mohseni, J. Mutus, O. Naaman, M. Neeley, C. Neill, M. Y. Niu, E. Ostby, A. Petukhov, J. C. Platt, C. Quintana, E. G. Rieffel, P. Roushan, N. C. Rubin, D. Sank, K. J. Satzinger, V. Smelyanskiy, K. J. Sung, M. D. Trevithick, A. Vainsencher, B. Villalonga, T. White, Z. J. Yao, P. Yeh, A. Zalcman, H. Neven, and J. M. Martinis, *Quantum supremacy using a programmable superconducting processor*, **Nature** **574**, 505 (2019).
- [14] Google Quantum AI, *Suppressing quantum errors by scaling a surface code logical qubit*, **Nature** **614**, 676 (2023).
- [15] Google Quantum AI and Collaborators, *Quantum error correction below the surface code threshold*, **Nature** **638**, 920 (2025).
- [16] E. Knill, R. Laflamme, and G. J. Milburn, *A scheme for efficient quantum computation with linear optics*, **Nature** **409**, 46 (2001).
- [17] H.-S. Zhong, H. Wang, Y.-H. Deng, M.-C. Chen, L.-C. Peng, Y.-H. Luo, J. Qin, D. Wu, X. Ding, Y. Hu, P. Hu, X.-Y. Yang, W.-J. Zhang, H. Li, Y. Li, X. Jiang, L. Gan, G. Yang, L. You, Z. Wang, L. Li, N.-L. Liu, C.-Y. Lu, and J.-W. Pan, *Quantum computational advantage using photons*, **Science** **370**, 1460 (2020).
- [18] H. J. Manetsch, G. Nomura, E. Bataille, *et al.*, *A tweezer array with 6,100 highly coherent atomic qubits*, **Nature** **647**, 60 (2025).
- [19] A. Browaeys and T. Lahaye, *Many-body physics with individually controlled Rydberg atoms*, **Nat. Phys.** **16**, 132 (2020).
- [20] S. Ebadi, T. T. Wang, H. Levine, A. Keesling, G. Semeghini, A. Omran, D. Bluvstein, R. Samajdar, H. Pichler, W. W. Ho, S. Choi, S. Sachdev, M. Greiner, V. Vuletić, and M. D. Lukin, *Quantum phases of matter on a 256-atom programmable quantum simulator*, **Nature** **595**, 227 (2021).
- [21] D. Bluvstein, S. J. Evered, A. A. Geim, S. H. Li, H. Zhou, T. Manovitz, S. Ebadi, M. Cain, M. Kalinowski, D. Hangleiter, J. P. Bonilla Ataides, N. Maskara, I. Cong, X. Gao, P. Sales Rodriguez, T. Karolyshyn, G. Semeghini, M. J. Gullans, M. Greiner, V. Vuletić, and M. D. Lukin, *Logical quantum processor based on reconfigurable atom arrays*, **Nature** **626**, 58 (2023).

-
- [22] D. Bluvstein *et al.*, *A fault-tolerant neutral-atom architecture for universal quantum computation*, *Nature* **649**, 39 (2026).
- [23] J. Preskill, *Quantum Computing in the NISQ era and beyond*, *Quantum* **2**, 79 (2018).
- [24] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O’Brien, *A variational eigenvalue solver on a photonic quantum processor*, *Nature Communications* **5**, 4213 (2014).
- [25] E. Farhi, J. Goldstone, and S. Gutmann, *A Quantum Approximate Optimization Algorithm*, (2014), arXiv preprint, [arXiv:1411.4028 \[quant-ph\]](#) .
- [26] B. M. Terhal, *Quantum error correction for quantum memories*, *Rev. Mod. Phys.* **87**, 307 (2015).
- [27] E. T. Campbell, B. M. Terhal, and C. Vuillot, *Roads towards fault-tolerant universal quantum computation*, *Nature* **549**, 172 (2017).
- [28] P. W. Shor, *Scheme for reducing decoherence in quantum computer memory*, *Phys. Rev. A* **52**, R2493 (1995).
- [29] A. R. Calderbank and P. W. Shor, *Good quantum error-correcting codes exist*, *Phys. Rev. A* **54**, 1098 (1996).
- [30] A. M. Steane, *Error Correcting Codes in Quantum Theory*, *Phys. Rev. Lett.* **77**, 793 (1996).
- [31] D. Gottesman, *Stabilizer Codes and Quantum Error Correction*, *Ph.D. thesis*, California Institute of Technology (1997), [quant-ph/9705052](#) .
- [32] H. Bombin and M. A. Martin-Delgado, *Topological quantum distillation*, *Phys. Rev. Lett.* **97**, 180501 (2006).
- [33] D. Bacon, *Operator quantum error-correcting subsystems for self-correcting quantum memories*, *Phys. Rev. A* **73**, 012340 (2006).
- [34] J.-P. Tillich and G. Zemor, *Quantum LDPC codes with positive rate and minimum distance proportional to the square root of the blocklength*, *IEEE Transactions on Information Theory* **60**, 1193 (2014).
- [35] P. Panteleev and G. Kalachev, *Asymptotically good quantum and locally testable classical LDPC codes*, in *Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing* (2022) pp. 375–388.
- [36] D. Gottesman, A. Kitaev, and J. Preskill, *Encoding a qubit in an oscillator*, *Phys. Rev. A* **64**, 012310 (2001).

- [37] E. Dennis, A. Kitaev, A. Landahl, and J. Preskill, *Topological quantum memory*, *J. Math. Phys.* **43**, 4452 (2002).
- [38] A. G. Fowler, M. Mariantoni, J. M. Martinis, and A. N. Cleland, *Surface codes: Towards practical large-scale quantum computation*, *Phys. Rev. A* **86**, 032324 (2012).
- [39] S. McArdle, S. Endo, A. Aspuru-Guzik, S. C. Benjamin, and X. Yuan, *Quantum computational chemistry*, *Rev. Mod. Phys.* **92**, 015003 (2020).
- [40] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd, *Quantum machine learning*, *Nature* **549**, 195 (2017).
- [41] A. Abbas *et al.*, *Challenges and opportunities in quantum optimization*, *Nature Reviews Physics* **6**, 718 (2024).
- [42] B. Korte and J. Vygen, *Combinatorial Optimization: Theory and Algorithms*, 6th ed., Algorithms and Combinatorics, Vol. 21 (Springer, Berlin, Heidelberg, 2018).
- [43] A. Schrijver, *Combinatorial Optimization: Polyhedra and Efficiency*, Algorithms and Combinatorics, Vol. 24 (Springer, Berlin, Heidelberg, 2003).
- [44] A. Lucas, *Ising formulations of many NP problems*, *Frontiers in Physics* **2**, 5 (2014).
- [45] F. Glover, G. Kochenberger, and Y. Du, A Tutorial on Formulating and Using QUBO Models, (2019), [arXiv:1811.11538 \[cs.DS\]](#) .
- [46] G. Pecci, R. Wang, P. Torta, G. B. Mbeng, and G. E. Santoro, *Beyond quantum annealing: optimal control solutions to maxcut problems*, *Quantum Sci. Technol.* **9**, 045013 (2024).
- [47] G. B. Mbeng, *Quantum annealing and digital quantum ground state preparation algorithms*, *Ph.D. thesis*, Available at <https://iris.sissa.it>, SISSA, Trieste (2019).
- [48] E. Boros and P. L. Hammer, *The max-cut problem and quadratic 0–1 optimization: Polyhedral aspects, relaxations and bounds*, *Annals of Operations Research* **33**, 151 (1991).
- [49] G. B. Mbeng, R. Fazio, and G. Santoro, *Quantum Annealing: a journey through Digitalization, Control, and hybrid Quantum Variational schemes*, [arXiv:1906.08948 \[quant-ph\]](#) (2019).
- [50] J. Hartmanis and R. E. Stearns, *On the computational complexity of algorithms*, *Transactions of the American Mathematical Society* **117**, 285 (1965).

-
- [51] S. A. Cook, *The P versus NP problem*, in *The Millennium Prize Problems*, edited by J. A. Carlson, A. Jaffe, and A. Wiles (American Mathematical Society, Providence, RI, 2006) pp. 87–104.
- [52] J. Watrous, *Quantum computational complexity*, in *Encyclopedia of Complexity and Systems Science*, edited by R. A. Meyers (Springer, New York, 2009) pp. 7174–7201, [arXiv:0804.3401](#).
- [53] R. M. Karp, *Reducibility among combinatorial problems*, in *Complexity of Computer Computations*, edited by R. E. Miller, J. W. Thatcher, and J. D. Bohlinger (Springer, Boston, MA, 1972) pp. 85–103.
- [54] M. R. Garey, D. S. Johnson, and L. Stockmeyer, *Some simplified NP-complete graph problems*, *Theoretical Computer Science* **1**, 237 (1976).
- [55] P. Alimonti and V. Kann, *Some APX-completeness results for cubic graphs*, *Theoretical Computer Science* **237**, 123 (2000).
- [56] S. A. Cook, *The complexity of theorem-proving procedures*, in *Proceedings of the Third Annual ACM Symposium on Theory of Computing* (Association for Computing Machinery, 1971) pp. 151–158.
- [57] E. Bernstein and U. Vazirani, *Quantum complexity theory*, *SIAM Journal on Computing* **26**, 1411 (1997).
- [58] A. C.-C. Yao, *Quantum circuit complexity*, in *Proceedings of the 34th Annual Symposium on Foundations of Computer Science* (IEEE, 1993) pp. 352–361.
- [59] J. Watrous, *Succinct quantum proofs for properties of finite groups*, in *Proceedings of the 41st Annual Symposium on Foundations of Computer Science* (IEEE, 2000) pp. 537–546, [arXiv:cs/0009002](#).
- [60] C. Marriott and J. Watrous, *Quantum Arthur–Merlin games*, *Computational Complexity* **14**, 122 (2005).
- [61] M. Agrawal, N. Kayal, and N. Saxena, *PRIMES is in P*, *Annals of Mathematics* **160**, 781 (2004).
- [62] A. B. Finnila, M. A. Gomez, C. Sebenik, C. Stenson, and J. D. Doll, *Quantum annealing: A new method for minimizing multidimensional functions*, *Chemical Physics Letters* **219**, 343 (1994).
- [63] T. Kadowaki and H. Nishimori, *Quantum annealing in the transverse Ising model*, *Phys. Rev. E* **58**, 5355 (1998).
- [64] G. E. Santoro and E. Tosatti, *Optimization using quantum mechanics: quantum annealing through adiabatic evolution*, *Journal of Physics A: Mathematical and General* **39**, R393 (2006).

- [65] T. Albash and D. A. Lidar, *Adiabatic quantum computation*, Rev. Mod. Phys. **90**, 015002 (2018).
- [66] UNSW Physics, [Proof of the adiabatic theorem](#), YouTube video (2020), published 28 July 2020; accessed 10 August 2026.
- [67] J. J. Sakurai and J. Napolitano, *Modern Quantum Mechanics*, 3rd ed. (Cambridge University Press, 2020).
- [68] E. Farhi, J. Goldstone, S. Gutmann, and M. Sipser, [Quantum computation by adiabatic evolution](#), (2000), [arXiv:quant-ph/0001106 \[quant-ph\]](#) .
- [69] P. Torta, *Quantum Approximate Optimization Algorithm and Variational Quantum Computing: from binary neural networks to ground state preparation*, [Ph.D. thesis](#), International School for Advanced Studies (SISSA), Trieste (2024).
- [70] D. K. Headley, *Angles and Devices for Quantum Approximate Optimization*, [Ph.D. thesis](#), Saarland University (2023).
- [71] M. W. Johnson *et al.*, *Quantum annealing with manufactured spins*, [Nature](#) **473**, 194 (2011).
- [72] A. D. King *et al.*, *Quantum critical dynamics in a 5,000-qubit programmable spin glass*, [Nature](#) **617**, 61 (2023).
- [73] R. Barends, A. Shabani, L. Lamata, J. Kelly, A. Mezzacapo, U. L. Heras, R. Babbush, A. G. Fowler, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, E. Jeffrey, E. Lucero, A. Megrant, J. Y. Mutus, M. Neeley, C. Neill, P. J. J. O’Malley, C. Quintana, P. Roushan, D. Sank, A. Vainsencher, J. Wenner, T. C. White, E. Solano, H. Neven, and J. M. Martinis, *Digitized adiabatic quantum computing with a superconducting circuit*, [Nature](#) **534**, 222 (2016).
- [74] D. Aharonov, W. van Dam, J. Kempe, Z. Landau, S. Lloyd, and O. Regev, *Adiabatic quantum computation is equivalent to standard quantum computation*, in [45th Annual IEEE Symposium on Foundations of Computer Science](#) (2004) pp. 42–51.
- [75] L. K. Grover, *Quantum mechanics helps in searching for a needle in a haystack*, [Phys. Rev. Lett.](#) **79**, 325 (1997).
- [76] C. Zalka, *Grover’s quantum searching algorithm is optimal*, [Phys. Rev. A](#) **60**, 2746 (1999).
- [77] J. Roland and N. J. Cerf, *Quantum search by local adiabatic evolution*, [Phys. Rev. A](#) **65**, 042308 (2002).

-
- [78] G. B. Mbeng, A. Russomanno, and G. E. Santoro, *The quantum ising chain for beginners*, [SciPost Physics Lecture Notes](#) , **82** (2024).
- [79] S. Knysh, *Zero-temperature quantum annealing bottlenecks in the spin-glass phase*, *Nature Communications* **7**, 12370 (2016).
- [80] D. Roberts, L. Cincio, A. Saxena, A. Petukhov, and S. Knysh, *Noise amplification at spin-glass bottlenecks of quantum annealing: A solvable model*, [Phys. Rev. A](#) **101**, 042317 (2020).
- [81] D. Guéry-Odelin, A. Ruschhaupt, A. Kiely, E. Torrontegui, S. Martínez-Garaot, and J. G. Muga, *Shortcuts to adiabaticity: Concepts, methods, and applications*, [Rev. Mod. Phys.](#) **91**, 045001 (2019).
- [82] M. Demirplak and S. A. Rice, *Adiabatic population transfer with control fields*, [J. Phys. Chem. A](#) **107**, 9937 (2003).
- [83] M. V. Berry, *Transitionless quantum driving*, [J. Phys. A: Math. Theor.](#) **42**, 365303 (2009).
- [84] M. Kolodrubetz, D. Sels, P. Mehta, and A. Polkovnikov, “Geometry” and non-adiabatic response in quantum and classical systems, [Physics Reports](#) **697**, 1 (2017).
- [85] D. Sels and A. Polkovnikov, *Minimizing irreversible losses in quantum systems by local counterdiabatic driving*, [Proc. Natl. Acad. Sci. USA](#) **114**, E3909 (2017).
- [86] R. Wang, V. R. Arezzo, K. Thengil, G. Pecci, and G. E. Santoro, *From exponential to quadratic: optimal control for a frustrated ising ring model*, [Quantum Sci. Technol.](#) **10**, 035052 (2025).
- [87] J. Wurtz and P. J. Love, *Counterdiabaticity and the quantum approximate optimization algorithm*, [Quantum](#) **6**, 635 (2022).
- [88] H. F. Trotter, *On the product of semi-groups of operators*, [Proc. Amer. Math. Soc.](#) **10**, 545 (1959).
- [89] M. Suzuki, *Generalized Trotter’s formula and systematic approximants of exponential operators and inner derivations with applications to many-body problems*, [Commun. Math. Phys.](#) **51**, 183 (1976).
- [90] G. B. Mbeng, R. Fazio, and G. E. Santoro, *Optimal quantum control with digitized quantum annealing*, (2019), [arXiv:1911.12259 \[quant-ph\]](#) .
- [91] M. Cerezo, A. Arrasmith, R. Babbush, S. Benjamin, S. Endo, K. Fujii, J. R. McClean, K. Mitarai, X. Yuan, L. Cincio, and P. Coles, *Variational Quantum Algorithms*, [Nature Reviews Physics](#) (2021).

- [92] S. Lloyd, *Quantum approximate optimization is computationally universal*, arXiv e-prints , arXiv:1812.11075 (2018), [arXiv:1812.11075 \[quant-ph\]](#) .
- [93] S. Hadfield, Z. Wang, B. O’Gorman, E. Rieffel, D. Venturelli, and R. Biswas, *From the Quantum Approximate Optimization Algorithm to a Quantum Alternating Operator Ansatz*, [Algorithms](#) **12**, 34 (2019).
- [94] A. Bäertschi and S. Eidenbenz, *Grover Mixers for QAOA: Shifting Complexity from Mixer Design to State Preparation*, in *2020 IEEE International Conference on Quantum Computing and Engineering (QCE)* (IEEE, 2020) pp. 72–82, [arXiv:2006.00354 \[quant-ph\]](#) .
- [95] 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*, [Physical Review X](#) **10**, 021067 (2020).
- [96] D. J. Tannor and S. A. Rice, *Control of selectivity of chemical reaction via control of wave packet evolution*, [J. Chem. Phys.](#) **83**, 5013 (1985).
- [97] A. P. Peirce, M. A. Dahleh, and H. Rabitz, *Optimal control of quantum-mechanical systems: Existence, numerical approximation, and applications*, [Phys. Rev. A](#) **37**, 4950 (1988).
- [98] D. J. Tannor, V. Kazakov, and V. Orlov, *Control of photochemical branching: Novel procedures for finding optimal pulses and global upper bounds*, in *Time-Dependent Quantum Molecular Dynamics*, NATO ASI Series B: Physics, Vol. 299, edited by J. Broeckhove and L. Lathouwers (Plenum Press, New York, 1992) pp. 347–360.
- [99] W. Zhu, J. Botina, and H. Rabitz, *Rapidly convergent iteration methods for quantum optimal control of population*, [J. Chem. Phys.](#) **108**, 1953 (1998).
- [100] R. S. Judson and H. Rabitz, *Teaching lasers to control molecules*, [Phys. Rev. Lett.](#) **68**, 1500 (1992).
- [101] A. Assion, T. Baumert, M. Bergt, T. Brixner, B. Kiefer, V. Seyfried, M. Strehle, and G. Gerber, *Control of chemical reactions by feedback-optimized phase-shaped femtosecond laser pulses*, [Science](#) **282**, 919 (1998).
- [102] J. P. Palao and R. Kosloff, *Quantum computing by an optimal control algorithm for unitary transformations*, [Phys. Rev. Lett.](#) **89**, 188301 (2002).
- [103] N. Khaneja, T. Reiss, C. Kehlet, T. Schulte-Herbrüggen, and S. J. Glaser, *Optimal control of coupled spin dynamics: design of nmr pulse sequences by gradient ascent algorithms*, [Journal of Magnetic Resonance](#) **172**, 296 (2005).

-
- [104] T. Caneva, M. Murphy, T. Calarco, R. Fazio, S. Montangero, V. Giovannetti, and G. E. Santoro, *Optimal control at the quantum speed limit*, *Phys. Rev. Lett.* **103**, 240501 (2009).
- [105] P. Doria, T. Calarco, and S. Montangero, *Optimal control technique for many-body quantum dynamics*, *Phys. Rev. Lett.* **106**, 190501 (2011).
- [106] T. Caneva, T. Calarco, and S. Montangero, *Chopped random-basis quantum optimization*, *Phys. Rev. A* **84**, 022326 (2011).
- [107] F. Motzoi, J. M. Gambetta, P. Rebentrost, and F. K. Wilhelm, *Simple pulses for elimination of leakage in weakly nonlinear qubits*, *Phys. Rev. Lett.* **103**, 110501 (2009).
- [108] N. Rach, M. M. Müller, T. Calarco, and S. Montangero, *Dressing the chopped-random-basis optimization: A bandwidth-limited access to the trap-free landscape*, *Phys. Rev. A* **92**, 062343 (2015).
- [109] C. P. Koch, U. Boscain, T. Calarco, G. Dirr, S. Filipp, S. J. Glaser, R. Kosloff, S. Montangero, T. Schulte-Herbrüggen, D. Sugny, and F. K. Wilhelm, *Quantum optimal control in quantum technologies. strategic report on current status, visions and goals for research in europe*, *EPJ Quantum Technology* **9**, 19 (2022).
- [110] D. D'Alessandro, *Introduction to quantum control and dynamics* (Chapman and Hall/CRC, 2007).
- [111] W. Zhu and H. Rabitz, *A rapid monotonically convergent iteration algorithm for quantum optimal control over the expectation value of a positive definite operator*, *J. Chem. Phys.* **109**, 385 (1998).
- [112] F. Albertini and D. D'Alessandro, *The lie algebra structure and controllability of spin systems*, *Linear Algebra and its Applications* **350**, 213 (2002).
- [113] J. Werschnik and E. K. U. Gross, *Quantum optimal control theory*, *J. Phys. B: At. Mol. Opt. Phys.* **40**, R175 (2007).
- [114] C. Altafini and F. Ticozzi, *Modeling and control of quantum systems: An introduction*, *IEEE Trans. Autom. Control* **57**, 1898 (2012).
- [115] U. Boscain, M. Sigalotti, and D. Sugny, *Introduction to the Pontryagin maximum principle for quantum optimal control*, *PRX Quantum* **2**, 030203 (2021).
- [116] Y. Lu, S. Joshi, V. S. Dinh, and J. Koch, *Optimal control of large quantum systems: assessing memory and runtime performance of GRAPE*, *J. Phys. Commun.* **8**, 025002 (2024).

- [117] G. C. Hegerfeldt, *Driving at the quantum speed limit: Optimal control of a two-level system*, *Phys. Rev. Lett.* **111**, 260501 (2013).
- [118] W. Lechner, P. Hauke, and P. Zoller, *A quantum annealing architecture with all-to-all connectivity from local interactions*, *Science advances* **1**, e1500838 (2015).
- [119] K. Ender, R. ter Hoeven, B. E. Niehoff, M. Drieb-Schön, and W. Lechner, *Parity Quantum Optimization: Compiler*, *Quantum* **7**, 950 (2023).
- [120] W. Lechner, *Quantum approximate optimization with parallelizable gates*, *IEEE Transactions on Quantum Engineering* **1**, 1 (2020).
- [121] Y. Susa and H. Nishimori, *Variational optimization of the quantum annealing schedule for the lechner–hauke–zoller scheme*, *Phys. Rev. A* **103**, 022619 (2021).
- [122] M. Lanthaler, C. Dłaska, K. Ender, and W. Lechner, *Rydberg-blockade-based parity quantum optimization*, *Phys. Rev. Lett.* **130**, 220601 (2023).
- [123] A. Weidinger, G. B. Mbeng, and W. Lechner, *Error mitigation for quantum approximate optimization*, *Physical Review A* **108**, 032408 (2023).
- [124] A. Weidinger, G. B. Mbeng, M. Fellner, D. Khachatryan, and W. Lechner, *Performance of parity qaoa for the signed max-cut problem*, *New Journal of Physics* **28**, 034514 (2026), [arXiv:2409.14786](#) .
- [125] V. R. Arezzo, R. Wang, K. Thengil, G. Pecci, and G. Santoro, *Digital controllability of transverse field ising chains*, *Physical Review A* **113**, 012610 (2026), [arXiv:2509.17754](#) .
- [126] R. Wang, A. Weidinger, G. B. Mbeng, W. Lechner, and G. E. Santoro, *Parity mapping for quantum optimization on frustrated ising rings*, (2026), [arXiv preprint, arXiv:2608.12213 \[quant-ph\]](#) .
- [127] T. Caneva, R. Fazio, and G. E. Santoro, *Adiabatic quantum dynamics of a random ising chain across its quantum critical point*, *Phys. Rev. B* **76**, 144427 (2007).
- [128] V. Bapst, L. Foini, F. Krzakala, G. Semerjian, and F. Zamponi, *The quantum adiabatic algorithm applied to random optimization problems: The quantum spin glass perspective*, *Physics Reports* **523**, 127 (2013).
- [129] D. Guéry-Odelin, A. Ruschhaupt, A. Kiely, E. Torrontegui, S. Martínez-Garaot, and J. G. Muga, *Shortcuts to adiabaticity: Concepts, methods, and applications*, *Rev. Mod. Phys.* **91**, 045001 (2019).

-
- [130] S. Matsuura, S. Buck, V. Senicourt, and A. Zaribafiyani, *Variationally scheduled quantum simulation*, *Phys. Rev. A* **103**, 052435 (2021).
 - [131] N. Barraza, G. A. Barrios, I. Montalban, E. Solano, and F. Albarrán-Arriagada, *Variational coherent quantum annealing*, (2023), [arXiv:2310.02248 \[quant-ph\]](#) .
 - [132] P. Chandarana, N. N. Hegade, K. Paul, F. Albarrán-Arriagada, E. Solano, A. del Campo, and X. Chen, *Digitized-counterdiabatic quantum approximate optimization algorithm*, *Phys. Rev. Research* **4**, 013141 (2022).
 - [133] C. M. Keever and M. Lubasch, *Towards adiabatic quantum computing using compressed quantum circuits*, (2023), [arXiv:2311.05544 \[quant-ph\]](#) .
 - [134] I. Čepaitė, A. Polkovnikov, A. J. Daley, and C. W. Duncan, *Counterdiabatic optimized local driving*, *PRX Quantum* **4**, 010312 (2023).
 - [135] J. R. Finžgar, M. J. A. Schuetz, J. K. Brubaker, H. Nishimori, and H. G. Katzgraber, *Designing quantum annealing schedules using bayesian optimization*, *Phys. Rev. Res.* **6**, 023063 (2024).
 - [136] G. Passarelli, V. Cataudella, and P. Lucignano, *Improving quantum annealing of the ferromagnetic p -spin model through pausing*, *Phys. Rev. B* **100**, 024302 (2019).
 - [137] G. Passarelli, K.-W. Yip, D. A. Lidar, H. Nishimori, and P. Lucignano, *Reverse quantum annealing of the p -spin model with relaxation*, *Phys. Rev. A* **101**, 022331 (2020).
 - [138] B. Seoane and H. Nishimori, *Many-body transverse interactions in the quantum annealing of the p -spin ferromagnet*, *Journal of Physics A: Mathematical and Theoretical* **45**, 435301 (2012).
 - [139] G. Quiroz, *Robust quantum control for adiabatic quantum computation*, *Phys. Rev. A* **99**, 062306 (2019).
 - [140] J. Côté, F. Sauvage, M. Larocca, M. Jonsson, L. Cincio, and T. Albash, *Diabatic quantum annealing for the frustrated ring model*, *Quantum Science and Technology* **8**, 045033 (2023).
 - [141] J. Tilly, H. Chen, S. Cao, D. Picozzi, K. Setia, Y. Li, E. Grant, L. Wossnig, I. Rungger, G. H. Booth, and J. Tennyson, *The Variational Quantum Eigensolver: a review of methods and best practices*, (2021), [arXiv:2111.05176 \[quant-ph\]](#) .

- [142] K. Bharti, A. Cervera-Lierta, T. H. Kyaw, T. Haug, S. Alperin-Lea, A. Anand, M. Degroote, H. Heimonen, J. S. Kottmann, T. Menke, W.-K. Mok, S. Sim, L.-C. Kwek, and A. Aspuru-Guzik, *Noisy intermediate-scale quantum algorithms*, *Rev. Mod. Phys.* **94**, 015004 (2022).
- [143] M. Cerezo, M. Larocca, D. García-Martín, N. L. Diaz, P. Braccia, E. Fontana, M. S. Rudolph, P. Bermejo, A. Ijaz, S. Thanasilp, E. R. Anschuetz, and Z. Holmes, Does provable absence of barren plateaus imply classical simulability? Or, why we need to rethink variational quantum computing, (2023), [arXiv:2312.09121 \[quant-ph\]](#).
- [144] L. Bittel and M. Kliesch, *Training variational quantum algorithms is np-hard*, *Phys. Rev. Lett.* **127**, 120502 (2021).
- [145] J. R. McClean, S. Boixo, V. N. Smelyanskiy, R. Babbush, and H. Neven, *Barren plateaus in quantum neural network training landscapes*, *Nature Communications* **9** (2018).
- [146] M. Ragone, B. N. Bakalov, F. Sauvage, A. F. Kemper, C. O. Marrero, M. Larocca, and M. Cerezo, A Unified Theory of Barren Plateaus for Deep Parametrized Quantum Circuits, (2023), [arXiv:2309.09342 \[quant-ph\]](#).
- [147] E. Fontana, D. Herman, S. Chakrabarti, N. Kumar, R. Yalovetzky, J. Heredge, S. H. Sureshbabu, and M. Pistoia, The Adjoint Is All You Need: Characterizing Barren Plateaus in Quantum Ansätze, (2023), [arXiv:2309.07902 \[quant-ph\]](#).
- [148] J. Dborin, F. Barratt, V. Wimalaweera, L. Wright, and A. G. Green, *Matrix product state pre-training for quantum machine learning*, *Quantum Science and Technology* **7**, 035014 (2022).
- [149] A. Skolik, J. R. McClean, M. Mohseni, P. van der Smagt, and M. Leib, *Layerwise learning for quantum neural networks*, *Quantum Machine Intelligence* **3**, 5 (2021).
- [150] E. Grant, L. Wossnig, M. Ostaszewski, and M. Benedetti, *An initialization strategy for addressing barren plateaus in parametrized quantum circuits*, *Quantum* **3**, 214 (2019).
- [151] M. Kliesch and I. Roth, *Theory of Quantum System Certification*, *PRX Quantum* **2** (2021).
- [152] S. H. Sack, R. A. Medina, A. A. Michailidis, R. Kueng, and M. Serbyn, *Avoiding Barren Plateaus Using Classical Shadows*, *PRX Quantum* **3**, 020365 (2022).

-
- [153] A. A. Mele, G. B. Mbeng, G. E. Santoro, M. Collura, and P. Torta, *Avoiding barren plateaus via transferability of smooth solutions in a hamiltonian variational ansatz*, *Phys. Rev. A* **106**, L060401 (2022).
 - [154] L. Broers and L. Mathey, *Mitigated barren plateaus in the time-nonlocal optimization of analog quantum-algorithm protocols*, *Phys. Rev. Res.* **6**, 013076 (2024).
 - [155] M. S. Rudolph, J. Miller, D. Motlagh, J. Chen, A. Acharya, and A. Perdomo-Ortiz, *Synergistic pretraining of parametrized quantum circuits via tensor networks*, *Nature Communications* **14**, 8367 (2023).
 - [156] S. H. Sack and M. Serbyn, *Quantum annealing initialization of the quantum approximate optimization algorithm*, *Quantum* **5**, 491 (2021).
 - [157] M. M. Wauters, E. Panizon, G. B. Mbeng, and G. E. Santoro, *Reinforcement-learning-assisted quantum optimization*, *Phys. Rev. Research* **2**, 033446 (2020).
 - [158] P. Torta, G. B. Mbeng, C. Baldassi, R. Zecchina, and G. E. Santoro, *Quantum approximate optimization algorithm applied to the binary perceptron*, *Phys. Rev. B* **107**, 094202 (2023).
 - [159] G. Pagano, A. Bapat, P. Becker, K. S. Collins, A. De, P. W. Hess, H. B. Kaplan, A. Kyprianidis, W. L. Tan, C. Baldwin, L. T. Brady, A. Deshpande, F. Liu, S. Jordan, A. V. Gorshkov, and C. Monroe, *Quantum approximate optimization of the long-range Ising model with a trapped-ion quantum simulator*, *Proceedings of the National Academy of Sciences* **117**, 25396 (2020), <https://www.pnas.org/content/117/41/25396.full.pdf>.
 - [160] L. Lumia, P. Torta, G. B. Mbeng, G. E. Santoro, E. Ercolessi, M. Burrello, and M. M. Wauters, *Two-Dimensional $\mathbb{Z}_2$ Lattice Gauge Theory on a Near-Term Quantum Simulator: Variational Quantum Optimization, Confinement, and Topological Order*, *PRX Quantum* **3**, 020320 (2022).
 - [161] T. S. Cubitt, D. Perez-Garcia, and M. M. Wolf, *Undecidability of the spectral gap*, *Nature* **528**, 207 (2015).
 - [162] G. Lami, P. Torta, G. E. Santoro, and M. Collura, *Quantum annealing for neural network optimization problems: A new approach via tensor network simulations*, *SciPost Physics* **14**, 117 (2023).
 - [163] A. Dranov, J. Kellendonk, and R. Seiler, *Discrete time adiabatic theorems for quantum mechanical systems*, *Journal of Mathematical Physics* **39**, 1340 (1998).
 - [164] M. M. M. Müller, R. S. Said, F. Jelezko, T. Calarco, and S. Montangero, *One decade of quantum optimal control in the chopped random basis*, *Reports on Progress in Physics* **85**, 076001 (2022).

- [165] S. Boyd and L. Vandenberghe, *Convex Optimization* (Cambridge University Press, 2004).
- [166] R. J. Duffin and E. L. Peterson, *Geometric programming with signomials*, *Journal of Optimization Theory and Applications* **11**, 3 (1973).
- [167] C. D. Maranas and C. A. Floudas, *Global optimization in generalized geometric programming*, *Computers & Chemical Engineering* **21**, 351 (1997).
- [168] G. B. Mbeng, L. Arceci, and G. E. Santoro, *Optimal working point in digitized quantum annealing*, *Phys. Rev. B* **100**, 224201 (2019).
- [169] P. de Fouquieres, S. Schirmer, S. Glaser, and I. Kuprov, *Second order gradient ascent pulse engineering*, *Journal of Magnetic Resonance* **212**, 412 (2011).
- [170] E. Farhi, J. Goldstone, S. Gutmann, and L. Zhou, *The Quantum Approximate Optimization Algorithm and the Sherrington-Kirkpatrick Model at Infinite Size*, *Quantum* **6**, 759 (2022).
- [171] E. Farhi, J. Goldstone, S. Gutmann, and L. Zhou, *The Quantum Approximate Optimization Algorithm and the Sherrington-Kirkpatrick model at infinite size*, (2020), arXiv preprint, [arXiv:1910.08187 \[quant-ph\]](https://arxiv.org/abs/1910.08187) .
- [172] G. E. Crooks, *Performance of the Quantum Approximate Optimization Algorithm on the Maximum Cut Problem*, (2018), arXiv preprint, [arXiv:1811.08419 \[quant-ph\]](https://arxiv.org/abs/1811.08419) .
- [173] L. T. Brady, L. Kocia, P. Bienias, A. Bapat, Y. Kharkov, and A. V. Gorshkov, *Behavior of analog quantum algorithms*, (2021), arXiv preprint, [arXiv:2107.01218 \[quant-ph\]](https://arxiv.org/abs/2107.01218) .
- [174] F. G. S. L. Brandao, M. Broughton, E. Farhi, S. Gutmann, and H. Neven, *For fixed control parameters the quantum approximate optimization algorithm’s objective function value concentrates for typical instances*, (2018), [arXiv:1812.04170 \[quant-ph\]](https://arxiv.org/abs/1812.04170) .
- [175] M. Streif and M. Leib, *Training the quantum approximate optimization algorithm without access to a quantum processing unit*, *Quantum Science and Technology* **5**, 034008 (2020).
- [176] S. H. Sureshbabu, D. Herman, R. Shaydulin, J. Basso, S. Chakrabarti, Y. Sun, and M. Pistoia, *Parameter Setting in Quantum Approximate Optimization of Weighted Problems*, *Quantum* **8**, 1231 (2024).
- [177] R. Shaydulin, P. C. Lotshaw, J. Larson, J. Ostrowski, and T. S. Humble, *Parameter transfer for quantum approximate optimization of weighted maxcut*, *ACM Transactions on Quantum Computing* **4** (2023).

- [178] J. Basso, D. Gamarnik, S. Mei, and L. Zhou, *Performance and limitations of the qaoa at constant levels on large sparse hypergraphs and spin glass models*, in *2022 IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS)* (2022) pp. 335–343.
- [179] J. Basso, E. Farhi, K. Marwaha, B. Villalonga, and L. Zhou, *The Quantum Approximate Optimization Algorithm at High Depth for MaxCut on Large-Girth Regular Graphs and the Sherrington-Kirkpatrick Model*, in *17th Conference on the Theory of Quantum Computation, Communication and Cryptography (TQC 2022)*, Leibniz International Proceedings in Informatics (LIPIcs), Vol. 232, edited by F. Le Gall and T. Morimae (Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl, Germany, 2022) pp. 7:1–7:21.
- [180] A. H. Al-Mohy and N. J. Higham, *Computing the action of the matrix exponential, with an application to exponential integrators*, *SIAM Journal on Scientific Computing* **33**, 488 (2011), <https://doi.org/10.1137/100788860>.
- [181] K. Thengil, V. R. Arezzo, and G. E. Santoro, *Digital techniques for the frustrated Ising ring: the role of counter-diabatic terms*, (2026), [arXiv:2607.23074 \[quant-ph\]](https://arxiv.org/abs/2607.23074).
- [182] N. Margolus and L. B. Levitin, *The maximum speed of dynamical evolution*, *Physica D: Nonlinear Phenomena* **120**, 188 (1998), proceedings of the Fourth Workshop on Physics and Consumption.
- [183] V. Akshay, H. Philathong, E. Campos, D. Rabinovich, I. Zacharov, X.-M. Zhang, and J. D. Biamonte, *Circuit depth scaling for quantum approximate optimization*, *Phys. Rev. A* **106**, 042438 (2022).
- [184] M. Larocca, P. Czarnik, K. Sharma, G. Muraleedharan, P. J. Coles, and M. Cerezo, *Diagnosing Barren Plateaus with Tools from Quantum Optimal Control*, *Quantum* **6**, 824 (2022).
- [185] M. Y. Niu, S. Lu, and I. L. Chuang, *Optimizing qaoa: Success probability and runtime dependence on circuit depth*, (2019), [arXiv:1905.12134 \[quant-ph\]](https://arxiv.org/abs/1905.12134).
- [186] F. Balducci, A. Grabarits, and A. del Campo, *Fighting exponentially small gaps by counterdiabatic driving*, *arXiv preprint arXiv:2410.02520* (2024).
- [187] E. Torrontegui, S. IbAAez, S. MartAnez-Garaot, M. Modugno, A. del Campo, D. GuAry-Odelin, A. Ruschhaupt, X. Chen, and J. G. Muga, *Chapter 2 - shortcuts to adiabaticity*, in *Advances in Atomic, Molecular, and Optical Physics*, Advances In Atomic, Molecular, and Optical Physics, Vol. 62, edited by E. Arimondo, P. R. Berman, and C. C. Lin (Academic Press, 2013) pp. 117–169.
- [188] J. Nocedal and S. J. Wright, *Numerical optimization* (Springer, 1999).

- [189] V. Giovannetti, S. Lloyd, and L. Maccone, *Quantum limits to dynamical evolution*, *Phys. Rev. A* **67**, 052109 (2003).
- [190] L. B. Levitin and T. Toffoli, *Fundamental limit on the rate of quantum dynamics: The unified bound is tight*, *Phys. Rev. Lett.* **103**, 160502 (2009).
- [191] B. Shanahan, A. Chenu, N. Margolus, and A. del Campo, *Quantum speed limits across the quantum-to-classical transition*, *Phys. Rev. Lett.* **120**, 070401 (2018).
- [192] M. Okuyama and M. Ohzeki, *Quantum speed limit is not quantum*, *Phys. Rev. Lett.* **120**, 070402 (2018).
- [193] M. M. Wauters, G. B. Mbeng, and G. E. Santoro, *Polynomial scaling of the quantum approximate optimization algorithm for ground-state preparation of the fully connected p -spin ferromagnet in a transverse field*, *Phys. Rev. A* **102**, 062404 (2020).
- [194] M. Christandl, N. Datta, T. C. Dorlas, A. Ekert, A. Kay, and A. J. Landahl, *Perfect transfer of arbitrary states in quantum spin networks*, *Phys. Rev. A* **71**, 032312 (2005).
- [195] M. Fellner, K. Ender, R. ter Hoeven, and W. Lechner, *Parity Quantum Optimization: Benchmarks*, *Quantum* **7**, 952 (2023).
- [196] E. Wybo and M. Leib, *Vanishing performance of the parity-encoded quantum approximate optimization algorithm applied to spin-glass models*, *Quantum* **8**, 1554 (2024).
- [197] J. Bennett, N. Chancellor, V. Kendon, and W. Lechner, *Improving success probability in lechner-hauke-zoller parity embedding by computing with quantum walks*, *Phys. Rev. A* **112**, 042616 (2025).
- [198] M. Lanthaler and W. Lechner, *Minimal constraints in the parity formulation of optimization problems*, *New Journal of Physics* **23**, 083039 (2021).
- [199] V. Marić, S. M. Giampaolo, D. Kuić, and F. Franchini, *The frustration of being odd: how boundary conditions can destroy local order*, *New Journal of Physics* **22**, 083024 (2020).
- [200] V. Marić, S. M. Giampaolo, and F. Franchini, *Quantum phase transition induced by topological frustration*, *Communications Physics* **3**, 220 (2020).
- [201] G. Torre, V. Marić, F. Franchini, and S. M. Giampaolo, *Effects of defects in the xy chain with frustrated boundary conditions*, *Phys. Rev. B* **103**, 014429 (2021).

-
- [202] R. Miyazaki and T. Yamamoto, *Four-body coupler for superconducting qubits based on josephson parametric oscillators*, *Phys. Rev. A* **111**, 062612 (2025).
- [203] L. Stenzel, R. ter Hoeven, R. Miyazaki, T. Yamaji, M. Shirane, and W. Lechner, *Spin model for quantum annealing with kerr parametric oscillators*, *Phys. Rev. A* **113**, 062621 (2026).
- [204] A. W. Glaetzle, R. M. W. van Bijnen, P. Zoller, and W. Lechner, *A coherent quantum annealer with rydberg atoms*, *Nature Communications* **8**, 15813 (2017).
- [205] Qaoa on frustrated ising rings, <https://gitlab.com/AnitaW/frustrated-ising-rings>.
- [206] S. Li, Y. Fan, X. Li, X. Ruan, Q. Zhao, Z. Peng, R.-B. Wu, J. Zhang, and P. Song, *Robust quantum control using reinforcement learning from demonstration*, *npj Quantum Information* **11**, 124 (2025).
- [207] V. Sivak, A. Morvan, M. Broughton, *et al.*, *Reinforcement learning control of quantum error correction*, *Nature* **655**, 879 (2026).
- [208] A. Gilyén, Y. Su, G. H. Low, and N. Wiebe, *Quantum singular value transformation and beyond: exponential improvements for quantum matrix arithmetics*, in *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing* (Association for Computing Machinery, 2019) pp. 193–204.
- [209] J. M. Martyn and P. Rall, *Halving the cost of quantum algorithms with randomization*, *npj Quantum Information* **11**, 47 (2025).
- [210] J.-T. Bu, L. Zhang, Z. Yu, J.-B. Wang, W.-Q. Ding, W.-F. Yuan, B. Wang, H.-J. Du, W.-J. Chen, L. Chen, J.-W. Zhang, J.-C. Li, F. Zhou, X. Wang, and M. Feng, *Exploring the experimental limit of deep quantum signal processing using a trapped-ion simulator*, *Phys. Rev. Applied* **23**, 034073 (2025).
- [211] J. R. Johansson, P. D. Nation, and F. Nori, *QuTiP: An open-source Python framework for the dynamics of open quantum systems*, *Computer Physics Communications* **183**, 1760 (2012).
- [212] J. R. Johansson, P. D. Nation, and F. Nori, *QuTiP 2: A Python framework for the dynamics of open quantum systems*, *Computer Physics Communications* **184**, 1234 (2013).
- [213] P. Virtanen, R. Gommers, T. E. Oliphant, *et al.*, *SciPy 1.0: fundamental algorithms for scientific computing in Python*, *Nature Methods* **17**, 261 (2020).
- [214] A. Javadi-Abhari, M. Treinish, K. Krsulich, C. J. Wood, J. Lishman, J. Gacon, S. Martiel, P. D. Nation, L. S. Bishop, A. W. Cross, B. R. Johnson, and

- J. M. Gambetta, [Quantum computing with Qiskit](#), (2024), [arXiv:2405.08810 \[quant-ph\]](#) .
- [215] F. Albertini and D. D'Alessandro, *The lie algebra structure and controllability of spin systems*, [Linear Algebra and its Applications](#) **350**, 213 (2002).