



UNIVERSITÀ
DEGLI STUDI
DI PADOVA



UNIVERSITY OF PADUA, DEPARTMENT OF PHYSICS AND ASTRONOMY "G. GALILEI"
MASTER THESIS IN PHYSICS OF DATA

Floquet counterdiabatic protocols for Quantum Annealing on Parity architecture

Francesco Pio Barone

ID 2058478

SUPERVISOR

Prof. Wolfgang Lechner

University of Innsbruck

INTERNAL SUPERVISOR

Prof. Simone Montangero

University of Padua

CO-SUPERVISOR

Dr. Glen Bigan Mbeng

University of Innsbruck

ACADEMIC YEAR 2023/2024

L^AT_EX document compiled on April 5, 2024.

Additional resources

This thesis work makes use of spin-system dynamics simulations.

The source code is publicly available online.



[baronefr/parity-floquet](https://github.com/baronefr/parity-floquet)

Contents

Abstract	v
Introduction	1
1 An introduction to Quantum Optimization	3
1.1 Combinatorial optimization problems	4
1.1.a Examples of combinatorial optimization problems	6
The Travelling Salesman Problem	7
The Maximum Cut problem	9
Quadratic Unconstrained Binary Optimization problems .	10
1.2 Quantum systems for optimization	11
1.2.a Spin glasses and the Ising model	12
1.2.b Optimization with spin systems	15
1.2.c Searching for the ground state: the connectivity mismatch	16
1.3 The Parity transformation	19
1.3.a The Parity/LHZ Hamiltonian	21
1.3.b Parity with arbitrarily connected graphs	22
2 Counterdiabatic Quantum Annealing	25
2.1 Quantum computation by adiabatic evolution	26
2.1.a The quantum adiabatic theorem	27
An approximate adiabaticity criterion	28
2.2 Counterdiabatic driving	31
2.2.a The Adiabatic Gauge Potential	31
2.2.b Approximations of the adiabatic gauge potential	34
Variational minimization	35
Nested commutator expansion	37
2.3 Quantum Annealing	39
2.4 Floquet engineered counterdiabatic protocols	41
3 Floquet engineering Parity annealing protocols	45
3.1 Floquet Parity protocols	46
3.1.a Additional controls of the counterdiabatic protocol	48
3.1.b Introducing the simulation instances	50

3.1.c	Single annealing sweep	51
3.1.d	Yield at longer annealing times	54
3.2	Tuning the FE protocol frequencies	56
3.2.a	Validation on other LHZ instances	59
3.2.b	Error sources	60
3.3	Cost analysis	62
3.3.a	The effective time compromise	64
3.3.b	Advantage over unassisted protocols	65
4	Can we improve Floquet protocols?	71
4.1	Alternative Floquet protocols	71
4.1.a	A family of protocols through variational minimization	72
4.1.b	Third-order Floquet protocols	76
	Results with second order counterdiabatic protocols	81
4.2	A quick look at the annealing of an Ising model	82
4.3	Alternative proposals within counterdiabaticity	86
4.3.a	Local Counterdiabatic Driving	86
4.3.b	Optimized Local Counterdiabatic Driving	87
	Advantage of protocols with different interactions	90
	Conclusions	93
A	AGP expressions for the Parity architecture	97
A.1	First order AGP coefficient	99
A.2	Second order AGP controls	100
B	Higher order Floquet protocols	103
C	Complement of simulations	105
C.1	LHZ6E and LHZ6H	108
C.2	LHZ10	112
C.3	Ising model	116
D	Further tests on models from reference paper	119
D.1	The three-level system	119
D.2	The magnetic trap	119
	Acronyms	125
	Bibliography	127
	Acknowledgments	137

Abstract

Combinatorial Optimization problems can be addressed with quantum computing techniques: the solution of an optimization problem can be encoded in the ground state of a quantum spin system, which in turn can be experimentally prepared through a suitable algorithm like Quantum Annealing.

Still, annealing applications have to face several challenges. Firstly, the limited spin connectivity of current hardware requires the use of clever encoding strategies. In this sense, the Parity architecture can encode fully connected optimization problems without requiring on-hardware long-range connectivity. Secondly, a quantum system naturally transits toward excited states during its evolution, resulting in a prepared state with a reduced overlap over the true ground state of the problem-encoding Hamiltonian. To overcome this limitation, one needs to find and implement non-trivial Quantum Annealing protocols.

In this work, we investigate Floquet counterdiabatic protocols for the Parity architecture. This combination of encoding and protocols constitutes a hardware-friendly recipe that does not require additional system controls.

On the one hand, we study the accuracy of the Floquet protocols, eventually finding a suitable criterion to tune the protocol hyperparameters. On the other hand, we investigate the efficiency of Floquet schedules. We identify a narrow driving frequency regime where the Floquet protocol provides an advantage over the standard Quantum Annealing approach.

We explore alternative protocols obtained by extending the derivation of Floquet schedules to higher orders of precision. Finally, we move away from the Floquet formulation and relax the requirement of preserving the original set of controls over the quantum system. Our work shows that numerically optimized counterdiabatic protocols using only extra local field controls can be advantageous over standard annealing approaches.



Introduction

The recent advancements in quantum technologies have sparked a lot of interest in the field of applied computing, due to the development of quantum algorithms that can solve those problems more efficiently than their classical counterparts. Combinatorial optimization problems [68] are paradigmatic examples of problems that are hard to solve with classical approaches; yet they are encountered across a wide range of fields, ranging from theoretical sciences to industrial applications. It has been shown that the solution of an optimization problem can be encoded in the ground state of quantum spin systems [60]. The ground states of spin systems can be experimentally prepared through suitable algorithms on dedicated hardware. One prominent quantum algorithm is Quantum Annealing [45, 41], in which a quantum spin system undergoes a time-dependent evolution that, under the most ideal assumptions, ends up in the desired ground state.

Still, Quantum Annealing implementations pose several challenges that limit their practical advantage. For instance, the quantum system naturally transits toward excited states during its evolution [43, 2], resulting in a prepared state with a small overlap over the true ground state of the problem-encoding Hamiltonian. To overcome this challenge, one needs to employ non-trivial Quantum Annealing protocols.

Counterdiabatic driving [11] is a ‘natural’ solution to the limitations that arise from the dynamics of an evolving quantum system. More precisely, Counterdiabatic driving can aid annealing protocols by actively suppressing the transitions towards excited eigenstates during the evolution. The advantage of speeding up the dynamics comes at the price of an extended set of controls [78], which is an overhead that significantly complicates the practical implementation of counterdiabatic schedules. Moreover, it has been shown that the counterdia-

batic evolution can be realized without extending the set of controls by using stroboscopic driving schedules: the Floquet counterdiabatic protocols [15].

Another limiting factor of Quantum Annealing is the hardness of engineering devices that permit interactions between non-neighboring spins. A generic optimization problem encoding may instead require spins that interact with all the other spins of the system (i.e., a *fully connected* spin model). This issue can be mitigated by using clever encoding strategies. As an example, the Parity architecture [55] can encode fully connected optimization problems without requiring on-hardware long-range connectivity.

In this thesis, we investigate Floquet counterdiabatic protocols for the Parity architecture. This combination of Parity encoding and Floquet protocols constitutes a hardware-friendly recipe that does not require additional controls over the system.

The first two chapters of this thesis provide a formal introduction to the topics that will be covered: Chapter 1 introduces Quantum Optimization and provides more details about the core features of the Parity architecture. In Chapter 2 we delve into the theory related to the preparation of a quantum state, focusing on counterdiabaticity. Eventually, we introduce Floquet Engineered protocols, which allow us to obtain counterdiabatic evolution without modifying the connectivity requirements.

Our original contributions are reported in the two remaining Chapters. Chapter 3 discusses Floquet Engineered protocols for the Parity architecture in the context of Quantum Annealing. Our work aims to study the accuracy and efficiency of Floquet protocols. Supported by numerical simulations, we report several examples of annealing schedules executed on paradigmatic Lechner-Hauke-Zoller (LHZ) models [55].

Chapter 4 explores alternative protocols obtained by extending the derivation of Floquet schedules to higher orders of precision. We also show that our conclusions are not limited to the specific study case of Parity models and are, in fact, of a more general validity. Finally, we move away from the Floquet formulation and relax the requirement of preserving the original set of controls over the quantum system, showing a numerically optimized counterdiabatic protocol that uses only extra local-field controls [78, 9, 65].

An introduction to Quantum Optimization

Scientific research has frequently shown that the dynamics of a variety of systems can be explained using a few basic principles, which are frequently formulated as the optimization of some quantity. For instance, we know that a molecule is stable if its configuration minimizes the system energy. Nonetheless, optimizing the bonds of a molecule to find a stable configuration is a notoriously difficult task, particularly when dealing with molecules of practical interest that involve numerous atoms. Optimization is also encountered through social and industrial problems: finding the shortest route when driving in the city, preferably avoiding traffic, or planning a train trip looking for cheaper or faster solutions. Indeed, one could claim that optimization problems are a common thread across numerous facets of our world.

Optimization problems are generally expressed as the minimization (or maximization) of a cost function $f(\underline{\sigma})$ with respect to a multidimensional variable $\underline{\sigma}$. Despite the existence of a common formulation, optimization problems can be solved using a variety of tools rather than a single, universal method.

Many optimization tasks are intrinsically formulated over a discrete set of input variables: namely, each optimizable variable $\sigma_i \in \underline{\sigma}$ assumes a discrete value. This class of optimization problems is labeled as **combinatorial** [68, 54]. Because of the different flavors of the problems, the literature about algorithms for solving combinatorial problems has diverged from continuous and multi-dimensional optimization methods.

The main goal of this Chapter is to explain how combinatorial problems can be encoded in the dynamics of quantum spin systems [33, 62], making it possible to employ more efficient algorithms. Intuitively, we will see that the cost function f can be mapped to the energy of a quantum spin system. Consequently, the

solution to the optimization problem is found by preparing the quantum system in its lowest energy state. Therefore, the paradigm is shifted to the research of quantum algorithms that can prepare a quantum system's ground state more efficiently than current algorithms can solve optimization problems [60].

On the one hand, if this thread of research proves to be successful, it would be possible to address significant problem instances that are untreatable with classic approaches. On the other hand, preparing the ground state of a spin system is itself a non-trivial task. In our discussion, we will focus on Quantum Annealing, a quantum algorithm that is executed on a specialized class of quantum computers (the quantum annealers).

By the end of this Chapter, we will see that there exist challenges on how efficiently the combinatorial problems can be encoded to quantum spin systems. Some encoding schemes come at the cost of either controlling numerous spins or implementing complex long-range interactions. The discussion will be functional to the introduction of the Parity architecture [55] which has the crucial advantage of achieving fully connected logical qubits by using only locally interacting physical qubits. In Chapter 2, we will extensively discuss Quantum Annealing, with a focus on counterdiabaticity as a promising method to improve the accuracy of the prepared state.

The first two Chapters of this work will cover all the theoretical requirements to address the main topic of this thesis: Floquet counterdiabatic protocols for Quantum Annealing on Parity architectures.

1.1 Combinatorial optimization problems

Consider a vector of N variables, $\underline{\sigma} = (\sigma_1, \dots, \sigma_N) \in X^N$. The finite set X is often referred to as the *alphabet* from which the variables are drawn. In a typical scenario, X could be a Boolean set ($X = \{0, 1\} = \{\text{false}, \text{true}\}$) or a subset of integer numbers ($X \subset \mathbb{Z}$). Let us add to the picture a suitable function $f : X^N \rightarrow \mathbb{R}$ that maps each $\underline{\sigma}$ to a real cost value that is intended to be minimized. The pair (X^N, f) formally determines a specific instance¹ of an unconstrained Combinatorial Optimization (CO) problem. The integer N is referred to as the *problem size* or the *input size*. The *optimal solution* of the problem is an input $\underline{\sigma}^*$, among the possible candidates, that globally minimizes the value of the cost function. Therefore, solving the optimization problem reads

¹The detail about the *instance* is a necessary formality. An instance is determined by some input data that provides enough information to obtain a solution. A *problem* is a collection of instances with different but similarly generated input data. [39]

as follows:

$$\text{find } \underline{\sigma}^* \in X^N \quad \text{s.t.} \quad f(\underline{\sigma}^*) \leq f(\underline{\sigma}) \quad \forall \underline{\sigma} \in X^N \quad .$$

Some clarifications are due:

- It might occur that the optimal solution is not uniquely determined. Indeed, there is no guarantee that the cost function has a single global minimum. In such a case, the problem simply allows for more than one optimal solution. In the following, we will neglect this possibility and refer to *the* globally optimal solution as it is uniquely determined, even though it might just be *one* of the optimal solutions.
- It is possible to consider additional constraints over the input variables, which are expressed in terms of additional equalities $h_j(\underline{\sigma}) = 0$ or inequalities $g_i(\underline{\sigma}) \geq 0$. The combinatorial problem is now said to be *constrained*, and all the previous definitions hold by substituting X^N with its subset $S \subset X^N$ of candidate solutions that satisfy the constraints.

Let us write plainly the most general definition of **combinatorial optimization problem**. A (instance of) combinatorial optimization problem is a pair (S, f) , with $f : S \rightarrow \mathbb{R}$ being the cost function, and S being a finite (or possibly countably infinite) set of candidate solutions [68].

Enumeration is the simplest algorithm for solving combinatorial optimization problems: the cost function is evaluated for each possible input and the one minimizing the cost function is chosen as the optimal solution. Yet, enumeration is an inefficient algorithm. It is trivial to observe that the number of candidate solutions of an unconstrained combinatorial problem scales as $|X|^N$. Meaningful instances of combinatorial problems typically fall in the regime $N \gg 1$, necessitating the search for a solution in a large, high-dimensional space. For such reason, enumeration becomes quickly infeasible with increasing N . Similar conclusions apply when the problem is constrained, even though the number of candidate solutions is reduced overall. One could argue that having fewer candidate solutions simplifies the search, but it depends mostly on the solving algorithm. There exist examples of problems that can be efficiently solved without constraints but that could become extremely complex when constraints have to be considered [39].

The **complexity of an algorithm** is determined by the functional dependence (in terms of the input size N) of the number of steps required to find a solution. It is of common use to indicate the asymptotic scaling of the complexity with the \mathcal{O} notation. For example, the number of steps required by the enumeration

algorithm is precisely equal to the cardinality of the search space. Thus the complexity of enumeration is $\mathcal{O}(|X|^N)$, which reads “exponential in N ”.

A problem is considered *easy* if there exists a solving algorithm whose complexity is polynomially bounded in the input size. An exponential complexity is indicative of either an intrinsically hard problem or a bad algorithm to solve it. Even if it is true that the computational power is increasing with time [79], the best chance to significantly increase the size of the problems that can be solved is either to wait for computers to catch up, or to find better algorithms now.

The study of problem complexity is itself an extensive topic. We say that an algorithm is *efficient* when the number of steps required to reach a solution grows polynomially in the input size. In Combinatorial Optimization, there are several examples of efficient algorithms, like matroid intersection, weighted matching, and linear programming. However, there are still problems for which no efficient algorithm is known yet. In modern computational theory, there exists a framework for classifying problems based on their complexity [46]. A pillar of this framework is the notion of **NP-complete problems** [34, 17]. Such class of problems is defined by these two properties [68]:

- Despite best efforts, no known polynomial algorithm can solve an NP-complete problem.
- If a problem \mathcal{A} is NP-complete, then there is a polynomial reduction that maps \mathcal{A} to any other NP problem \mathcal{B} .

The first property is just a reasonable assumption that no polynomial algorithm exists for any NP-complete problem. Nobody so far has been able to formally prove the statement, and it is a common belief that it cannot be proved with current mathematical techniques [68]. However, a consequence of the second property is that if one finds a polynomial algorithm to solve an NP-complete problem, then all NP problems are also solvable in polynomial time.

Finally, we observe that instead of focusing on the search for the exact algorithm, it is an option to develop algorithms that produce an arbitrary approximation of the optimal solution. Indeed, it is still very useful in practical applications to find solutions that are ‘fairly close’ to the global optimal, meaning that $|f(\underline{\sigma}^*) - f(\underline{\sigma})| \leq \delta$ for arbitrarily small δ . This task is accomplished by using heuristic algorithms, that trade the accuracy of the solution for significantly lower running time.

1.1.a Examples of combinatorial optimization problems

Let us introduce three well-known combinatorial optimization problems, following the footsteps of Ref. [28]. This is a good chance to review the concepts



Figure 1.1: Simple example of TSP instance with its optimal solution.

that have already been discussed in the previous section and to establish that optimization problems can be formulated in terms of graphs. The first problem has a prominent historical significance: in the early 1970s, its study played a crucial role in the development of the algorithmic complexity theory. The remaining two examples will be directly related to the topics covered in §1.2.

The Travelling Salesman Problem

Consider a set of N points $\{p_1, \dots, p_N\}$, each one representing the position of a city in a 2D plane. Introduce also a metric function $d(p_i, p_j)$ that measures the distance between the cities p_i and p_j . The Travelling Salesman Problem (TSP) [42] consists in finding the order $\underline{\sigma} \in \mathbb{Z}^N$ in which the cities have to be visited, minimizing the total travel distance

$$f(\underline{\sigma}) = \sum_{i=1}^{N+1} d(p_{\sigma_i}, p_{\sigma_{i+1}}) .$$

In addition, the salesman wants to visit each city only once, thus $\underline{\sigma}$ must be a permutation of $(1, \dots, N)$, and he has to return to the original city at the end of the tour, i.e. $\sigma_{N+1} = \sigma_1$ when evaluating the cost function.

The number of all possible paths is determined by the number of permutations of the sequence $(1, \dots, N)$ divided by $2N$, due to the cyclic invariance of a closed path; thus $(N - 1)!/2$ candidate solutions are available. The scaling of this problem w.r.t. the input size is super-exponential, making the enumeration utterly inefficient in solving the problem.

The relevance of TSP goes way beyond the mere happiness of salesmen: the same mathematical formulation is recurrent in optimal scheduling, computer wiring, vehicle routing, and much more. Let us define a matrix of elements $d_{i,j} \equiv d(p_i, p_j)$. In the TSP, $[d_{i,j}]$ contains all the possible distances between the cities and is naively symmetric since the distance between cities does not depend on the travel direction. The matrix $[d_{i,j}]$ is useful to generalize the TSP to a wide range of alternative situations, like in the following example.



Example: (from Ref. [7]) Consider a machine that performs several tasks, one at a time, which are labeled as t_1, \dots, t_N . During a working day, the machine has to perform once all the possible tasks. The machine can switch between two tasks with a time overhead, say from t_i to t_j , given by $d_{i,j}$. In general $d_{i,j} \neq d_{j,i}$. The goal of this problem is to find the order of execution of the tasks that minimizes the time overhead. This new problem is an asymmetric instance of the TSP.

TSP is proved to be an NP-complete problem [69]. A state-of-the-art algorithm managed to solve optimally a (nontrivial) 85900-city problem, taking almost 136 years of CPU time on a distributed network of computers [3, 74].

Combinatorial optimization is strongly connected to **graphs** [49]. In the specific case of the TSP, it is quite intuitive to transpose the problem into a graph $G = (V, E)$: each city is associated with a vertex V_i and the vertices V_i, V_j are connected by edges $E_{i,j}$ of weight $d_{i,j}$.

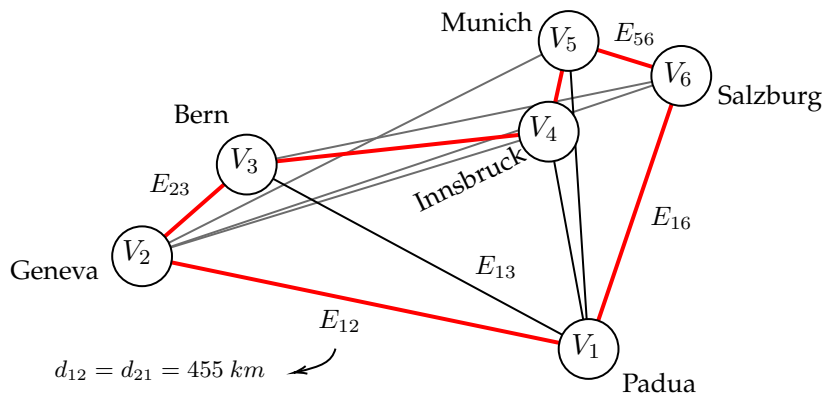


Figure 1.2: Example of TSP formulated on a graph. The red edges remark the solution seen in Figure 1.1.

The Maximum Cut problem

Consider an indirect graph $G = (V, E)$. A *cut* of G is a partition of its vertices in two complementary sets U and \bar{U} , which reads as $U \subset V$ and $\bar{U} = V \setminus U$. Given any cut² it is possible to find a set of edges $E(U)$ that have one vertex in U and one in \bar{U} . The **Maximum Cut** problem [16] consists of finding the cut U of G that maximizes $|E(U)|$ (the number of *cutted edges*).

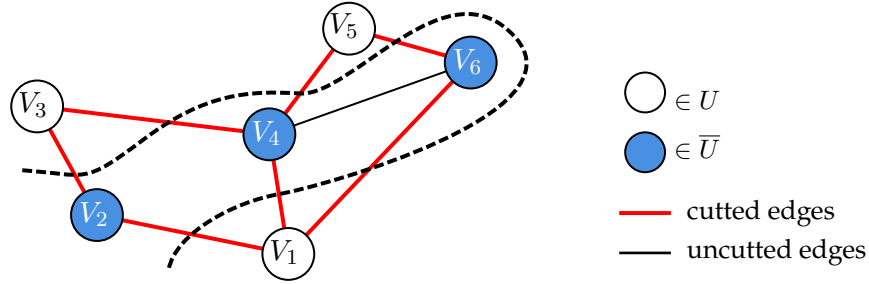


Figure 1.3: An example of Maximum Cut problem. The dashed line cuts the seven edges that are highlighted in red. Consequently, the vertices are marked with different colors to indicate their belonging to partition U or \bar{U} .

The Max-Cut problem can be modeled algebraically using binary variables. Assign to each vertex V_i a binary value that indicates if the vertex belongs to one of the two complementary partitions of a graph cut:

$$\sigma_i = \begin{cases} 1 & \text{if } V_i \in U \\ 0 & \text{if } V_i \in \bar{U} \end{cases}$$

The quantity $E_{ij} \equiv \sigma_i + \sigma_j - 2\sigma_i\sigma_j$ is conveniently defined to have the following property. If V_i and V_j belong to the same partition of a cut, it evaluates to $E_{ij} = 0$. Instead, if V_i and V_j belong to different partitions, it evaluates to $E_{ij} = 1$. The maximization of the number of cut edges $|E(U)|$, which was intuitively formulated on a graph, is now mapped to the minimization of

$$f(\underline{\sigma}) = - \sum_{i \neq j} E_{ij} = - \sum_{i \neq j} (\sigma_i + \sigma_j - 2\sigma_i\sigma_j) \quad (1.1)$$

Max Cut is also proved to be an NP-complete problem [56]. It is possible to add a further layer of complexity by considering weighted edges so that the goal of the problem is to maximize the sum of the weights of the cut $E(U)$. This varia-

²It is redundant to identify a cut with the symbol (U, \bar{U}) , so we prefer to simplify the notation by indicating only one of the two sets (U).

tion is included among Karp's 21 NP-complete problems [47]. In his work, Karp proved polynomial-time reductions between the boolean satisfiability problem (a well-known NP-complete problem) and 21 combinatorial problems, thus proving their NP-completeness.

Quadratic Unconstrained Binary Optimization problems

Consider a cost function that is quadratic in the inputs $\underline{\sigma} \in \{0, 1\}^N$. The expression can be conveniently written by storing the coefficients of each quadratic term in an upper triangular matrix $Q_{ij} \in \mathbb{R}^{N \times N}$:

$$f(\underline{\sigma}) = \sum_{i=1}^N \sum_{j>i}^N Q_{ij} \sigma_i \sigma_j$$

It is straightforward to include linear constraints as well, through a vector L_i .

$$f_{\text{QUBO}}(\underline{\sigma}) = \sum_i L_i \sigma_i + \sum_{i<j} Q_{ij} \sigma_i \sigma_j \quad (1.2)$$

The minimization of the previous equation is a typical example of Quadratic Unconstrained Binary Optimization (QUBO) problem [21, 38, 59]. The vector L_i and the matrix Q_{ij} determine an univocal QUBO instance.

QUBO problems are 'universal' since all combinatorial optimization problems can be recast to a QUBO formulation [52, 37]: set partitioning, graph coloring, warehouse location, and much more. Max-Cut is also a QUBO problem: indeed, Eq. (1.1) is easy to rewrite as Eq. (1.2) and vice-versa. The existence of a polynomial mapping from one problem to another indicates that they belong to the same complexity class.

As a final remark, the QUBO expression of Eq. (1.2) can be generalized to arbitrarily higher-order polynomials in the binary variables $\underline{\sigma}$.

$$f_{\text{PUBO}}(\underline{\sigma}) = \sum_i P_i^{(1)} \sigma_i + \sum_{i<j} P_{ij}^{(2)} \sigma_i \sigma_j + \sum_{i<j<k} P_{ijk}^{(3)} \sigma_i \sigma_j \sigma_k + \dots \quad (1.3)$$

The instances characterized by $(P^{(1)}, \dots, P^{(k)})$ fall under the name of Polynomial Unconstrained Binary Optimization (PUBO) problems. Despite being built on higher-order polynomials, the complexity class of a PUBO problem is the same as QUBO, because it is always possible to reduce the order of a polynomial

expression, at the price of increasing the number of binary variables. The order reduction is a polynomial mapping, hence the equivalence of the complexity class.

1.2 Quantum systems for optimization

Although quantum mechanics is not a recent physical theory, today it is still one of the most productive research areas. Quantum mechanics has been able to explain a wide range of microscopic phenomena since its foundations have been laid out. All of our technology, particularly semiconductors and lasers, is based on its advancements. Computers have of course benefited from better technology, resulting in a huge increment of computational power through the last decades. Even if singular computer components exploit quantum mechanical phenomena, the logic and computational processes that elaborate the information are still ‘classical’: no quantum property is used in the processing of information.

Quantum computers are devices that operate on quantum computational models, which explicitly take advantage of properties that do not have a classical analogous (like superposition, entanglement [14], ...). It is also known that those ‘super-classical’ properties are intrinsically difficult to simulate on a classical universal device [35]. To address this problem, Richard Feynman advanced in 1981 [31, 30] the idea of using quantum systems themselves to realize simulations of other quantum systems. By extension, he was suggesting the development of a quantum computational model, which would certainly be beneficial in other domains. The research line that originated from this brilliant intuition led to the development of quantum algorithms aiming to address classically intractable problems, i.e. with polynomial advantage or at least a significant speedup in their exponential scaling.

Even though the principles are promising, it is challenging to implement such devices. Numerous efforts have to be made to increase the number of available qubits, isolate the quantum system from the environment, and improve the control over its dynamics. Noisy Intermediate-Scale Quantum (NISQ) technology³ [73] is expected, in the next years, to advance to the point in which quantum computers can be used to solve problems that are intractable with classic devices.

³This term refers to the generation of non-fault-tolerant quantum devices that will be available for practical applications within the next few years. We expect those devices to employ from 50 to a few hundred qubits.

Combinatorial optimization problems are typical examples of problems that are notoriously difficult to solve with classical devices. The primary goal of the following sections is to establish the link between classical optimization and the dynamics of quantum systems [33, 62]. For the time being, we will not be discussing the specifics of how the quantum dynamics are implemented, as this topic will be covered in the next Chapter.

1.2.a Spin glasses and the Ising model

The term *spin glass* [27, 62, 32] has been introduced to describe materials that are characterized by random orientation of their magnetic moments, even though the atoms themselves are arranged in a periodic, ordered structure. The term *glass* has been borrowed from the common glass, which is a non-crystalline solid that shows an amorphous and random disposition of its atoms. In *spin glasses*, the “glassy randomness” refers to the disorder of the (frozen) spins rather than the lack of crystalline structure of the atoms. On the contrary, the atoms of a spin glass material are supposed to be distributed in a regular geometric pattern, which stands by the name of *lattice*.

Simple spin glasses can be obtained through alloys mixing non-magnetic metals with a few percent of magnetic atoms, such as the Au-Fe alloy of Ref. [13] which is sketched in Figure 1.4. Fundamentally, the atoms with a magnetic moment are randomly dispersed in the material. The strength of the interaction between magnetic moments depends on their random distance. Thus, a spin glass can be thought of as a system in which the couplings between magnetic moments are randomly distributed.⁴ Indeed, in 1975, S. Edward and P. Anderson formulated the spin glass problem in terms of an Ising model with couplings

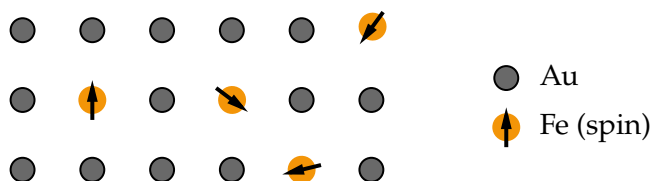


Figure 1.4: Sketch of a spin glass lattice. Au and Fe atoms are arranged in a regular structure (lattice). The Fe atoms’ magnetic moments, pictured with an arrow, are randomly oriented.

⁴The purpose of this work does not require a more formal treatment of spin glasses. But it is worth noting that more than 50 years of research into their properties resulted in Giorgio Parisi receiving the Nobel Prize in 2021 for his continuous replica symmetry breaking model [70]. Furthermore, spin glasses’ properties have proven to be more general and applicable to other systems that exhibit a similar type of disorder, despite different chemical compositions or different scales.

J_{ij} drawn from a real random distribution [26].

$$H_{\text{glass}} = - \sum_{\langle i,j \rangle} J_{ij} \sigma_i \sigma_j \quad (1.4)$$

To understand how spin glasses and Ising models are related to optimization problems, let us simplify the situation. Consider a one-dimensional lattice with N sites. The local magnetic moments are represented by Ising spins $\sigma_i = \pm 1$, which can be pictured as arrows pointing either up or down (± 1 , respectively). Suppose that each spin interacts only with its nearest neighbor, coupled through a constant $J_i = \pm 1$. The energy of the Ising chain is described by the Hamiltonian

$$H_{\text{chain}} = \sum_{i=1}^{N-1} J_i \sigma_i \sigma_{i+1} \quad (1.5)$$

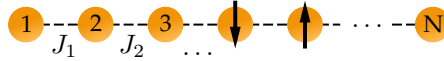


Figure 1.5: Chain of N spins described by the Hamiltonian of Eq. (1.5). The dashed lines connecting neighboring spins represent the couplings J_i .

It is in our interest to study the configuration of the spins that minimize the system energy: the ground state. We observe that the sign of J_i determines which type of interaction between spin pairs (σ_i, σ_{i+1}) is energetically favorable. $J_i = -1$ implies that parallel spins are energetically advantageous; vice-versa, a positive J_i favours an anti-parallel alignment of spins.

For instance, consider a chain of spins with uniform interactions J_i . If $J_i \equiv -1$, all the spins tend to align in the same direction, producing a non-null net magnetization (which is macroscopically measurable). This alignment produces a ferromagnetic phase. Instead, if all interactions are positively signed ($J_i \equiv +1$) the moments tend to counter-align, establishing two ferromagnetic sub-lattices that cancel each other, producing a net-zero magnetization. A sketch of the two phases is reported in Figure 1.6.

In general, there is no guarantee that there exists only one configuration of the system that minimizes the energy. If there are more, the ground state is said to be *degenerate*. Symmetries can be caused by a degeneracy. Consider, for instance, a square lattice of 4 spins with uniformly positive interactions J . There exist two configurations that minimize the energy (see Figure 1.7a), with no way to prefer one over the other, therefore, both states are degenerate ground states of the system. This behavior emerges due to symmetries in the geometry of the lattice that are not broken by external degrees of freedom (like an external field).



Figure 1.6: Two possible ground states of an Ising chain, in ferromagnetic (left) and antiferromagnetic (right) phases.

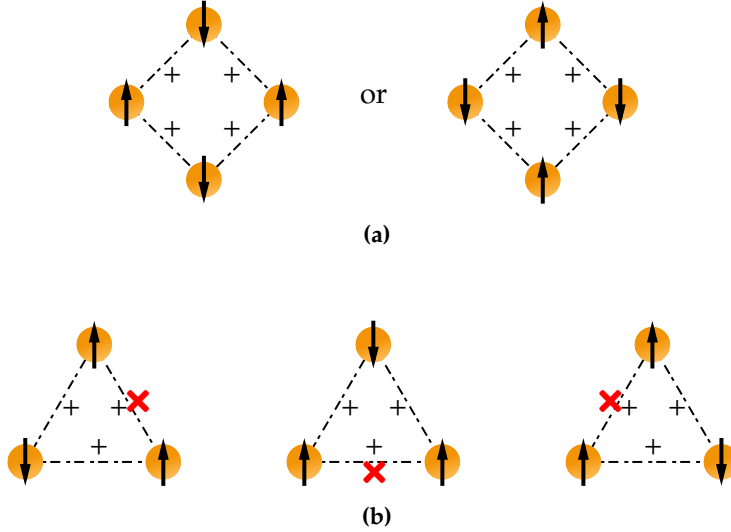


Figure 1.7: Examples of spin systems that present (a) degeneracy and (b) frustration. All the couplings of the spin chain, represented by a dash-dotted line, are intended to be positive: $J_i = +1$. The red crosses mark the couplings that are violated.

Frustration arises in spin systems when interactions cannot be satisfied due to some lattice geometries [63]. For instance, consider 3 spins in a triangular lattice, like in Figure 1.7b. The positive couplings drive the spins to be anti-aligned w.r.t. their neighbor. However, one can notice that one constraint is always violated. When the lattice geometry prevents any configuration from satisfying all the couplings, the system is said to show geometrical frustration.

The lattice geometry affects which pairs of spin interact. The most general way to express an Ising model is through the following Hamiltonian:

$$H_{\text{Ising}} = \sum_i^N h_i \sigma_i + \sum_{\langle i,j \rangle} J_{ij} \sigma_i \sigma_j \quad (1.6)$$

where h_i is a local field applied on each site and the bracket $\langle i, j \rangle$ denotes the couples of spins that are nearest-neighbor in a specific lattice geometry. The Ising model played a central role in the development of Statistical Mechanics, due to its behavior that retains most features of spin glasses and its numerous

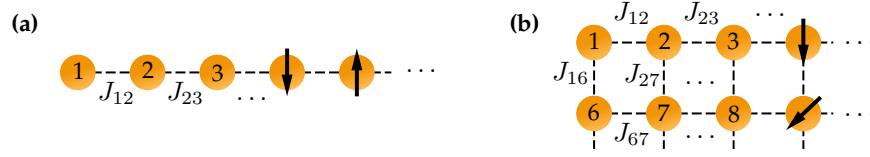
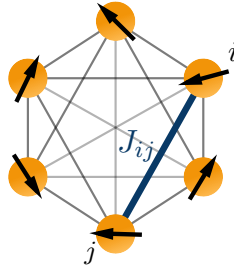


Figure 1.8: Ising model in one and two dimensions, respectively. The 1D Ising model was formulated by Wilhelm Lenz in 1920; the solution is due to Ernest Ising [61], from which the model took its name. The 2D case has been approached analytically by Lars Onsager [67], who published the solution in 1944. In contrast to the 1D case, the 2D Ising model exhibits a phase transition. It is proven that phase transitions occur also for higher dimensions, even though there is still no analytical solution.

applications, starting from the description of ferromagnetism. See Figure 1.8 for a sketch of the 1D and 2D Ising models.

There is no particular reason for which interactions have to be limited only to nearest-neighbor sites. Indeed, it is possible to generalize to all-to-all connected spin models, in which any pair of spins $i \neq j$ interacts through a constant J_{ij} . The coefficients of each interaction pair can be conveniently stored in an upper-triangular matrix $[J_{ij}]$ with $N(N - 1)/2$ entries. By setting some coefficients to zero, it is possible to formulate arbitrarily connected schemes of spins:



$$H_{\text{Ising}} = \sum_i^N h_i \sigma_i + \sum_{i < j} J_{ij} \sigma_i \sigma_j \quad (1.7)$$

1.2.b Optimization with spin systems

In the following, we wish to focus on the link between the Ising model and optimization problems. There are actually many ways to find connections with optimization problems. In broad terms, it is possible to map the loss function of the problem f to the Hamiltonian of a spin system. For instance, the expression of the spin glass Hamiltonian is isomorphic to the QUBO problem loss function

$$f_{\text{QUBO}}(\underline{\sigma}) = \sum_i P_i x_i + \sum_{i < j} Q_{ij} x_i x_j \quad \longleftrightarrow \quad H_{\text{Ising}} = \sum_i h_i \sigma_i + \sum_{i < j} J_{ij} \sigma_i \sigma_j$$

(Eq. 1.2) (Eq. 1.7)

through the transformation from binary to Ising variables $2x_i - 1 \equiv \sigma_i$. Consequently, $x_i = 1$ is mapped to $\sigma_i = 1$ and $x_i = 0$ to $\sigma_i = -1$. This argument is mathematical but can be interpreted intuitively: minimizing the cost function of the QUBO problem is formally equivalent to the minimization of the energy of the Ising spin system. This connection is quite acknowledged in the literature, to the extent that both Ising and QUBO problems are referred to under the common denomination of Binary Quadratic Models (BQM) [5].

From another point of view, it is possible to think of spin-glass interactions in terms of graphs. The spin system connectivity can be mapped to a graph G whose vertices V_i correspond to a site of the lattice; its edges E_{ij} , weighed by the coefficient J_{ij} , connect the interacting spins i and j . The key element consists again of formulating the problem of the graph in terms of a Hamiltonian.

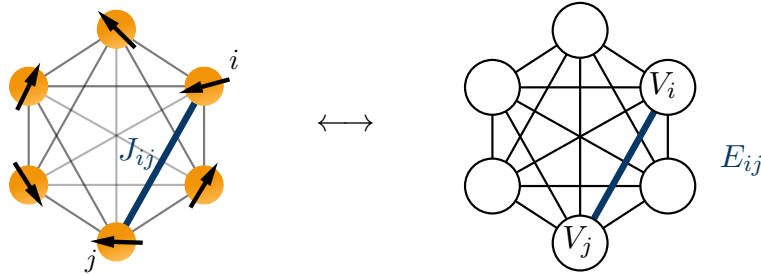


Figure 1.9: Conceptual mapping between spin systems and graphs. Each spin site i is associated with a node V_i and the non-null couplings J_{ij} are associated to an edge E_{ij} .

For instance, it can be proved [60] that the optimal solution of a Maximum Cut problem formulated on any graph $G = (V, E)$ is the ground state of the Ising Hamiltonian

$$H = \sum_{(i,j) \in E} J_{ij} \sigma_i \sigma_j \quad .$$

We wish to show a pictorial example of the connection between the Max Cut and the Ising model. Consider a 1D Ising chain with $J_i = +1$. We already know that in the absence of an external field, its ground state consists of anti-aligned spins. If we map the 1D spin chain to a graph, the Ising variables of the ground state suggest a partition of the graph vertices that maximizes the number of edges connecting nodes of different partitions. A graphic depiction of the previous statement is provided in Figure 1.10.

1.2.c Searching for the ground state: the connectivity mismatch

The previous observations established that combinatorial optimization problems can be mapped to Ising spin glasses and that their solution can be determined



Figure 1.10: Equivalence between the antiferromagnetic ground state and the solution of a specific Maximum Cut problem.

through the ground state of the spin system. In Ref. [60], the author proves Ising formulations for many NP-hard problems, including all of Karp’s 21 NP-complete problems. Remarkably, the required number of spins is at most cubic in the input size of the problem. In other words, the mapping between Ising spin glasses and specimen combinatorial optimization problems is polynomial in the input size.

The search for the ground state of an Ising spin glass through physical simulations is itself an NP problem [57]. Thus, it is not advantageous to use simulations of an Ising system to solve the optimization problem - unless we make use of some numerical approach that is not scaling exponentially in the input size. Say, instead, that we perform an experiment that physically prepares the ground state of any Ising glass system. A measure of the state would output the configuration of the spins, determining the solution of the optimization problem.

To perform such a task, we need an algorithm that can prepare the ground state of a spin system. It is crucial, though, that the quantum preparation algorithm scales more efficiently than any other classical algorithm.

The idea at the core of Adiabatic State Preparation (ASP) [4, 2] is that the target ground state can be prepared by slowly evolving a spin system from an initial state that is trivial to prepare. This objective is achieved by interpolating in the time interval $[0, \tau]$ two Hamiltonians, \mathcal{H}_i and \mathcal{H}_p :

$$\mathcal{H}(t) = (1 - \lambda(t)) \cdot \mathcal{H}_i + \lambda(t) \cdot \mathcal{H}_p \quad ,$$

imposing the boundary conditions $\lambda(0) = 0$ and $\lambda(\tau) = 1$. The total evolution time is marked with τ . If the system is prepared in the (trivial) ground state of \mathcal{H}_i , a result known as the adiabatic theorem [43] guarantees that the evolved state has a non-null overlap with the ground state of the “problem” Hamiltonian \mathcal{H}_p .

An adiabatic evolution (realized if $\tau \rightarrow \infty$) guarantees higher *probability of success*, which is customarily measured as the fidelity of the evolved state w.r.t. the true ground state of \mathcal{H}_p . It is inevitable however to settle on a finite τ and compromise on the resulting fidelity of the evolved state. In fact, there are significant

restrictions on the quantum system coherence time that make it difficult to arbitrarily extend τ . On the other extreme of the timescales, the main concern about ASP is the scaling of the minimum total evolution time τ to achieve satisfactory fidelities. The quantum adiabatic theorem (see §2.1.a) provides a more detailed insight into the dynamics that affect the minimum evolution time. For instance, in problems that involve first-order phase transitions, due to an exponentially closing energy gap, τ scales exponentially in the number of physical spins.

The challenges of ASP and Quantum Annealing (QA) [45, 64, 41, 87] - a protocol similar to ASP but that does not insist on adiabaticity - will be extensively discussed in Chapter 2. One of those challenging aspects, however, must be mentioned right away. Quantum Annealers do not allow for arbitrary physical connection of the qubits. The reason is prominently experimental and the details are strongly dependent on the type of hardware (superconducting qubits, trapped ions, etc...). In general, it is technologically challenging to connect qubits and the currently available connectivity topologies are far from being all-to-all. As a consequence, we have to deal with the fact that the required connectivity of the problem mismatches the actual host hardware connectivity.

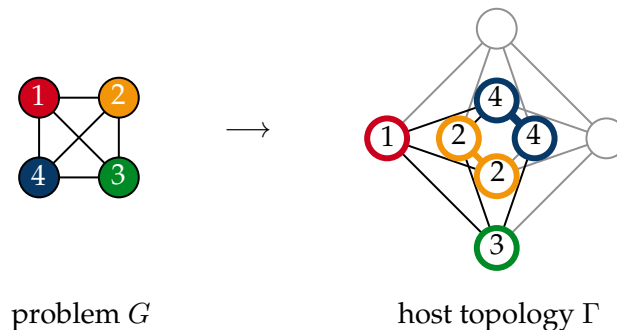


Figure 1.11: Example of minor embedding for the fully-connected graph G , on the left. Each of the 4 logical variables is laid out on the host topology graph Γ , on the right, which is a Chimera unit cell [66]. The connectivity of the variables in Γ is exactly the connectivity required by G , but the logical variables are stretched into multiple nodes. The contraction of the stretched logical variables returns the problem graph G .

The **connectivity mismatch** is quite an open challenge, with many possible solutions. One of them, the method of *minor embedding* [84, 6], consists of mapping the problem connectivity graph G over the target host topology graph Γ . This solution is routinely executed in NISQ quantum annealers, such as the D-Wave machines [66, 28].

Formally, a minor embedding of G in Γ is a sub-graph of Γ with the prop-

erty that the original graph G can be recovered by performing suitable edge contractions on the minor embedding. In simpler words, a minor-embedded graph maps a logical qubit to multiple (connected) physical qubits of the host machine. Moreover, by contracting the edges between physical qubits that encode the same logical qubit, it is possible to retrieve the problem connectivity graph. Figure 1.11 represents a simple example of embedding. To ensure that the physical qubits encode correctly the logical states, strong multi-qubit interactions are enforced on all the logically associated physical qubits. Finding an optimal minor embedding (i.e. the embedding that requires the smallest number of physical qubits) is itself an NP-hard problem [58], thus it is inevitable to settle for a sub-optimal embedding, which still serves the purpose of use in the current hardware. Nonetheless, the cost is a significant physical qubit overhead, which is currently the main restriction on problem sizes that can be solved.

Most of the current annealing hardware can perform quasi-local operations while maintaining a high level of accuracy, though complete fault tolerance remains a long-term goal [72, 10]. Achieving all-to-all interactions directly on hardware is still a challenging task. Nonetheless, ongoing efforts are trying to push further the connectivity of quantum devices, or at least to find topologies that allow more efficient embedding. Another branch of solutions, instead, would be to find an alternative encoding of a problem that can be realized within the available connectivity constraints, which is actually the topic addressed in the following section.

1.3 The Parity transformation

The Parity architecture [55] implements a logically fully connected model that can be implemented by sole local interactions. The formulation is not platform-specific, in the sense that it can be implemented on any hardware that offers a certain set of local controls. The price to pay is a $\mathcal{O}(N^2)$ redundancy in the number of physical qubits. The overhead, however, can be exploited to provide intrinsic fault tolerance. Remarkably, the idea at the core of Parity is inspired by the use of parity in error correction codes [77, 80, 85].

The first step to derive the architecture is to introduce the Parity variables $\tilde{\sigma}$, which encode the parity between two logical qubits. Two parallel interacting spins correspond to a +1 Parity variable; anti-parallel spins map instead to -1.



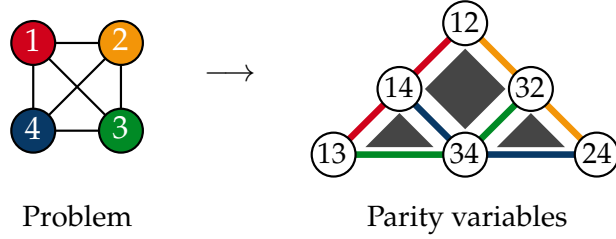
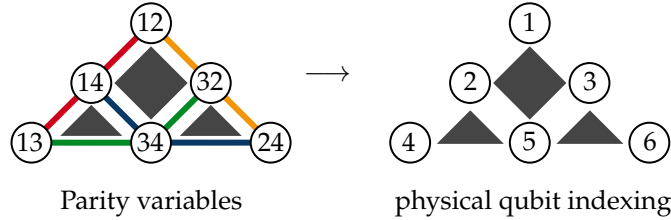


Figure 1.12: Example of Parity (LHZ) mapping for a fully connected problem with $N = 4$ logical variables, represented by the graph on the left. The Parity qubits are disposed on a 2D plane as in the right diagram. The Parity variables $\tilde{\sigma}_{ij}$ at the edge of a square (or a triangle) represent a closed cycle of logical qubits, and therefore a Parity constraint that has to be enforced. The colored lines connecting Parity variables remark the dependence on the original logical qubit: red for 1, yellow for 2, and so on.

which is the original name given to Parity diagrams that encode fully connected models.

1.3.a The Parity/LHZ Hamiltonian

From this point forward, let us neglect the step of the Parity transformation $(i, j) \rightarrow \tilde{\sigma}_{ij}$ and index each Parity variable with a single number: $\tilde{\sigma}^{(a)}, a \in \mathbb{N}$. The same number $a \in \{1, \dots, K\}$ will be used as the index of the physical qubit implementation.



Let us indicate each constraint with the index $l = 1, \dots, N_C$, where $N_C = K - N + 1$ is the number of required⁶ constraints. We greatly simplify the notation, denoting with $\tilde{\rho}^{[l]}$ the product of the Parity variables associated with the 4 (or 3) physical qubits of each constraint:

$$\begin{array}{c} \text{a} \quad \text{b} \\ \boxed{l} \\ \text{d} \quad \text{c} \end{array} = \tilde{\rho}^{[l]} \equiv \tilde{\sigma}^{(a)} \tilde{\sigma}^{(b)} \tilde{\sigma}^{(c)} \tilde{\sigma}^{(d)} > 0 \quad \text{for } (a, b, c, d) \in \text{constraint}[l]$$

Recall that the inequality > 0 is due to the observation that an even number

⁶ N_C constraints are sufficient to reduce the higher-dimension Parity space to the subspace of logically valid solutions [25].

of anti-parallel qubits is required to obtain a valid configuration in the space of Parity variables.

The constraints can be enforced by including in the “cost function” the terms $\tilde{\rho}^{[l]}$ multiplied by a negative constant, $C_l < 0$. The magnitude of the constant has to be sufficiently large to ensure that the constraint is satisfied. Formalizing the LHZ as a spin system, its “cost function” is written as the following spin Hamiltonian⁷, where we have chosen \hat{z} as the principal axis.

$$\mathcal{H}_{\text{LHZ}} = \sum_{k=1}^K J_k \hat{\sigma}_z^{(k)} + \sum_{l=1}^{K-N+1} C_l \hat{\rho}_z^{[l]} \quad (1.8)$$

1.3.b Parity with arbitrarily connected graphs

The LHZ model, introduced above, maps a fully connected graph of N logical variables to $K = N(N - 1)/2$ physical qubits representing the Parity variables. But what happens if the graph of the problem is not fully connected?

Clearly, the number of required Parity variables is reduced, because there is no need to map interactions that have null coefficients. However, the procedure used to lay down the LHZ scheme of Ref. [55] is not sensitive to such change and still maps the logical variables as the model is fully connected.

The immediate consequence is that LHZ schemes are not the most efficient when mapping arbitrarily connected graphs. The minimal number of required physical qubits K is exactly equal to the number of edges of the model graph, which is in general $K < N(N - 1)/2$ whenever the problem does not require full connectivity.

For the purpose of this work, there is no explicit need to investigate further this aspect of the Parity architecture. Nevertheless, we mention that it is possible to find a Parity mapping that uses exactly the minimum amount of required qubits. The method to build an optimal Parity mapping is referred to as *compilation* [29]. The Hamiltonian of a compiled Parity scheme is written by all means as the LHZ Hamiltonian of Eq. (1.8), so there is no practical change in all the work that will follow. An example of optimized mapping is provided in Figure 1.13.

⁷The operator $\hat{\rho}_z^{[l]}$ (which corresponds to the Parity variable $\tilde{\rho}^{[l]}$) stands for a tensor product of $\hat{\rho}_z$ Pauli operators acting on the physical qubits of each constraint:

$$\hat{\rho}_z^{[l]} = \begin{array}{cc} \text{a} & \text{b} \\ \blacksquare & \\ \text{d} & \text{c} \end{array} \equiv \hat{\sigma}_z^{(a)} \otimes \hat{\sigma}_z^{(b)} \otimes \hat{\sigma}_z^{(c)} \otimes \hat{\sigma}_z^{(d)} \quad \text{for qubits } (a, b, c, d) \in \text{constraint}[l]$$

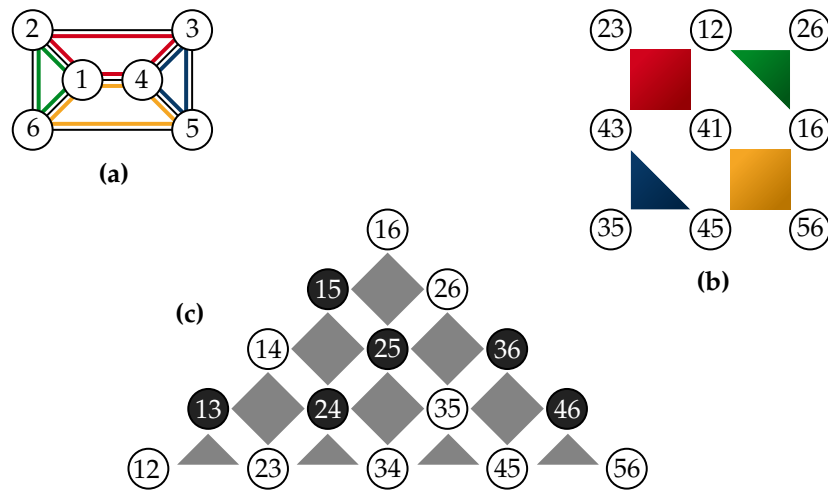


Figure 1.13: Example of compiled Parity mapping, taken from Figure 1 of Ref. [29]. **(a)** The problem connectivity scheme involves $N = 6$ not fully connected logical qubits. **(b)** is an example of compiled mapping that uses exactly 9 physical qubits (which is the number of edges in the problem graph). The constraints are color-coded to represent a cycle of four (or three) logical variables in the problem graph. **(c)** The LHZ mapping of a generic $N = 6$ problem uses instead $K = 15$ physical qubits. However, not all of them are required for the problem in (a): the Parity variables marked in black do not correspond to any edge of the problem graph and are therefore redundant in the Parity scheme.

Counterdiabatic Quantum Annealing

Despite the importance of quantum physics in the design of modern computer hardware, the underlying logic (the *computational model*) is based on the manipulation of ‘classical’ boolean variables, the bits. Therefore, the devices operating on such a model are called “classical computers”.

Lots of effort has been made to formulate a universal computational model relying on the principles of quantum mechanics. Deutsch made the first proposal in 1989, transposing the concept of classical gates to the quantum world. Summarizing, any evolution of the Hilbert space can be expressed as a discrete series of unitary quantum *gates*. The resulting framework has been called the circuit (or gate) model of quantum computing [22].

Today, circuit-based quantum computers are the standard and leading quantum technology. However, the current stage of quantum computation reveals several other quantum computing models that are promising and perhaps more indicated for certain classes of problems. We have seen, for instance, that optimization problems can be solved by searching for the ground state of a quantum spin system. Thus, it is necessary to find an algorithm to prepare the ground state of an arbitrary spin system. Even though one might try to accomplish this task in a circuit-based paradigm, there exists dedicated quantum hardware and algorithms that have been specifically designed for this task.

In this Chapter, we will introduce the Adiabatic State Preparation [4, 2] and Quantum Annealing [45, 41]. After discussing their limitations, we will focus on counterdiabatic driving as a promising solution. Finally, we will formally introduce the Floquet protocols [15].

2.1 Quantum computation by adiabatic evolution

Adiabatic Quantum Computing (AQC) is a computational model that enables universal quantum calculations [2, 1]. Such a framework is suitable for the preparation of the ground state of any Hamiltonian \mathcal{H}_p by controlling the evolution of the quantum system $\mathcal{H}(t)$ in a fixed time window. The system is initialized in the ground state of an initial Hamiltonian $\mathcal{H}_i = \mathcal{H}(0)$ and undergoes an evolution that ends up, at time τ , in the problem Hamiltonian $\mathcal{H}_p = \mathcal{H}(\tau)$.

$$\mathcal{H}_i \longrightarrow \mathcal{H}_p$$

This protocol is known as the Adiabatic State Preparation (ASP) algorithm [4, 2]. The ground state of \mathcal{H}_i is supposed to be unique and ‘easy’ to prepare - like a product state, that can be prepared through local operations. At the core of ASP lies the *Adiabatic Theorem*, which guarantees that the system stays in the instantaneous ground state of $\mathcal{H}(t)$ throughout the evolution, provided that it is carried out slowly enough [48, 43]. This condition certifies that the evolved state is exactly the ground state of the problem Hamiltonian.

It is customary to write the time-dependent Hamiltonian as an interpolation

$$\mathcal{H}(\lambda) = (1 - \lambda)\mathcal{H}_i + \lambda\mathcal{H}_p \quad (2.1)$$

weighted by a monotone function $\lambda(t) : [0, \tau] \rightarrow [0, 1]$ (the *schedule* function) satisfying the boundary conditions $\lambda(0) = 0$ and $\lambda(\tau) = 1$.

We generalize the concept of schedule by writing a generic Hamiltonian as a linear combination of time-independent Hamiltonians weighed by time-dependent coefficients.

$$\mathcal{H}(\lambda) = A(t)\mathcal{H}_A + B(t)\mathcal{H}_B + C(t)\mathcal{H}_C + \dots \quad (2.2)$$

We would then define the *schedule* as the set of functions $\{A(t), B(t), \dots\}$ that multiply any term of $\mathcal{H}(t)$ during its evolution. It does not matter what is the behavior of each schedule function within the time evolution, provided that the conditions $\mathcal{H}(0) = \mathcal{H}_i$ and $\mathcal{H}(\tau) = \mathcal{H}_p$ are still observed. The generalization is not a mere exercise of algebra, but has a physical relevance: it allows for more general “paths” between \mathcal{H}_i and \mathcal{H}_p , which might be advantageous w.r.t. the trivially interpolated protocol. For example, \mathcal{H}_p could be decomposed as a sum of numerous terms, each weighted by its own scheduling coefficient. Another idea could be to consider extra operators that act as a “catalyst” Hamiltonian [2], with coefficients being null at the start ($t = 0$) and end ($t = \tau$) of the protocol.

Right from its formulation, it is clear that ASP is particularly suitable for solving quantum optimization problems. To be advantageous, however, we would like this quantum algorithm to “scale better” than other classical algorithms. It is not straightforward to quantify the “cost” of the ASP, which is indeed a topic of active research [2]. Since the system evolves continuously, there is no notion of circuit depth, like in the gate-based model, or any number that quantifies the discrete operations to perform. In ASP, all we have is the protocol time-dependent Hamiltonian.

It is not even correct to consider the total evolution time τ as a meaningful measure of the cost. Say that the protocol Hamiltonian is multiplied by a constant, i.e. $\mathcal{H}(t) \rightarrow \kappa \cdot \mathcal{H}(t)$. The re-scaling operation changes its energy spectrum and alters the dynamics of the evolving system. For this reason, the total evolution time alone τ is not a reliable indicator of the algorithm complexity if the energy scales of the Hamiltonian are not appropriately considered [2]. To avoid this issue, it has been proposed in Ref. [1] to quantify the cost of the protocol through the dimensionless quantity

$$\mathcal{C} = \tau \cdot \max_t \|\mathcal{H}(t)\| \quad . \quad (2.3)$$

2.1.a The quantum adiabatic theorem

A key assumption at the heart of ASP is that the system evolves slowly. If not, the evolved state tends to get excited, resulting in a smaller probability of overlapping with the ground state of the problem Hamiltonian \mathcal{H}_p . This aspect of ASP goes in contrast with the definition of cost discussed above. On the one hand, it would be optimal to have a less expensive algorithm overall, reducing the total evolution time τ . On the other hand, the requirement of adiabatic evolution imposes that τ must increase to improve the probability of success, augmenting the cost. Furthermore, there are experimental considerations that make it impossible to increase the evolution time at will: current hardware has a very limited coherence time.

But what does ‘slow’ exactly mean? The literature proposes several instances of the quantum Adiabatic Theorem to answer this question. The theorems provide a lower bound of τ by estimating some quantities depending on the norms of \mathcal{H} and its time derivatives. For example, the following is a specific instance of the theorem, taken from Refs. [43, 2].

Theorem 1. *Suppose that the spectrum of $\mathcal{H}(\lambda)$ has the ground state eigenvalue separated by a gap $\Delta(\lambda) = \epsilon_1(\lambda) - \epsilon_0(\lambda) > 0$ from the rest of the spectrum, and that \mathcal{H} is twice continuously differentiable. Assume that \mathcal{H} , $\partial_\lambda \mathcal{H}$ and $\partial_\lambda^2 \mathcal{H}$ are bounded operators*

(an assumption that is always fulfilled in finite-dimensional spaces). Then

$$\tau \gg \max \left(\max_{\lambda} \frac{\|\partial_{\lambda} \mathcal{H}(\lambda)\|}{\Delta^2(\lambda)}, \max_{\lambda} \frac{\|\partial_{\lambda}^2 \mathcal{H}(\lambda)\|}{\Delta^2(\lambda)}, \max_{\lambda} \frac{\|\partial_{\lambda} \mathcal{H}(\lambda)\|^2}{\Delta^3(\lambda)} \right). \quad (2.4)$$

The theorem underlines a dependence of the minimum τ over the gap $\Delta(\lambda)$ between the ground state and the first excited state. To summarize, systems that experience smaller gaps during the evolution require larger time intervals to correctly evolve adiabatically. This is a limitation of the ASP, as there is no way to know a priori the gap without exact diagonalization. Besides, it is not hard to find examples of systems that violate the hypothesis of a non-vanishing spectral gap. Indeed, we know from statistical mechanics that spin systems exhibit closure of the gap in correspondence with phase transitions, which makes it harder to prepare their ground state.

An approximate adiabaticity criterion

Let us review a simplified adiabaticity criterion, adapted from Ref. [82]. Essentially, we will show that if specific terms of the dynamics are neglected, a quantum system prepared in any eigenstate at time $t = 0$ ends up in the correspondent eigenstate of the problem Hamiltonian. At the end of the derivation, we will provide a bound of τ that involves the off-diagonal matrix elements of $\partial_{\lambda} \mathcal{H}$ and the average gap between two distinct eigenstates. This form of the adiabaticity criterion will be useful to connect with the topics presented in §2.2.

Let $|\psi(t)\rangle$ be a quantum state that evolves in the time interval $[0, \tau]$. Its dynamics are described by the time-dependent Schrödinger equation.

$$i |\dot{\psi}(t)\rangle = \mathcal{H}(t) |\psi(t)\rangle$$

We use from now on the natural units $\hbar = 1$. Let us introduce the basis of instantaneous eigenstates $|n(t)\rangle$ with eigenvalues $\epsilon_n(t)$:

$$\mathcal{H}(t) |n(t)\rangle = \epsilon_n(t) |n(t)\rangle \quad (2.5)$$

Since $\{|n(t)\rangle\}$ is a complete orthonormal set, $|\psi(t)\rangle$ can be expressed as a linear combination in that basis:

$$|\psi(t)\rangle = \sum_n c_n(t) e^{i\theta_n(t)} |n(t)\rangle \quad (2.6)$$

We have explicitly factorized the phase $\theta_n(t) = \int_0^t \epsilon_n(t') dt'$ from $c_n(t)$. This

choice will be convenient in the following calculations, but it can be intuitively motivated straight away. It is known that an eigenstate picks up a phase $e^{-i\epsilon_n}$ when it evolves under a constant Hamiltonian. The $\theta_n(t)$ is a generalization of the picked-up phase, obtained by integrating it throughout the time interval in which the Hamiltonian evolves. It is known as the Berry phase.

Let us apply the time-dependent Schrödinger equation to $|\psi\rangle$:

$$i \sum_n e^{i\theta_n(t)} \left[\dot{c}_n |n(t)\rangle + ic_n \dot{\theta}_n |n(t)\rangle + c_n |\dot{n}(t)\rangle \right] = \sum_n c_n e^{i\theta_n(t)} \mathcal{H}(t) |n(t)\rangle \quad (2.7)$$

The derivative of the integral expression $\theta_n(t)$ is just $-\epsilon_n$. On the RHS, the Hamiltonian applied to the eigenstates $|n(t)\rangle$ (see Eq. (2.5)) brings a factor ϵ_n . As a result, the second term of LHS simplifies with RHS.

$$i \sum_n [\dot{c}_n |n(t)\rangle + c_n |\dot{n}(t)\rangle] e^{i\theta_n(t)} = 0$$

$$\sum_n \dot{c}_n |n(t)\rangle e^{i\theta_n(t)} = - \sum_n c_n |\dot{n}(t)\rangle e^{i\theta_n(t)}$$

The sum on the left side can be reduced to a single term by applying $\langle m(t)|$. The resulting equation,

$$\dot{c}_m = - \sum_n c_n e^{i(\theta_n - \theta_m)} \langle m(t)|\dot{n}(t)\rangle \quad , \quad (2.8)$$

can be integrated to determine the evolution of the quantum state. It is not difficult to evaluate $\langle m(t)|\dot{n}(t)\rangle$. Take the derivative of the instantaneous eigenbasis definition $|n(t)\rangle$ and project again with $\langle m(t)|$:

$$\begin{aligned} \dot{\mathcal{H}} |n\rangle + \mathcal{H} |\dot{n}\rangle &= \dot{\epsilon}_n |n\rangle + \epsilon_n |\dot{n}\rangle \\ \langle m|\dot{\mathcal{H}} |n\rangle + \langle m|\mathcal{H} |\dot{n}\rangle &= \dot{\epsilon}_n \langle m|n\rangle + \epsilon_n \langle m|\dot{n}\rangle \\ \langle m|\dot{\mathcal{H}} |n\rangle + \epsilon_m \langle m|\dot{n}\rangle &= \dot{\epsilon}_n \delta_{mn} + \epsilon_n \langle m|\dot{n}\rangle \end{aligned}$$

Thus, for $m \neq n$,

$$\langle m|\dot{n}\rangle = \frac{\langle m|\dot{\mathcal{H}} |n\rangle}{\epsilon_n - \epsilon_m} \quad (2.9)$$

We now observe that if terms $m \neq n$ are neglected in Eq. (2.8), the equality simplifies to $\dot{c}_m = -c_m \langle m|\dot{n}\rangle$, which has trivial solution:

$$c_m(t) = c_m(0) \cdot \exp \left[- \int_0^t \langle m(t')|\dot{n}(t')\rangle dt' \right] \quad (2.10)$$

Imposing the boundary condition $|\psi(0)\rangle = |n(0)\rangle$ - or equivalently $c_n(0) = 1$ and $c_m(0) = 0 \forall m \neq n$ - the evolved wave function integrates as

$$|\psi(t)\rangle = \exp \left[- \int_0^t \langle n(t') | \dot{n}(t') \rangle dt' \right] e^{i\theta_n(t)} |n(t)\rangle$$

Remarkably, the initial eigenstate $|n(0)\rangle$ is preserved throughout the evolution. For this reason, the off-diagonal terms ($m \neq n$) of $\dot{\mathcal{H}}$ have a prominent physical relevance in the adiabatic evolution of the quantum state: those terms are responsible for instantaneous transitions between eigenstates.

To formulate a better adiabaticity criterion, let us rewind to Eq. (2.8) and include the result of Eq. (2.9). Let us also relabel $n \rightarrow k$, $m \rightarrow n$ to make the notation more intuitive in the following computations.

$$\dot{c}_n(t) = - \sum_k c_k(t) \exp \left[-i \int_0^t dt' (\epsilon_k(t') - \epsilon_n(t')) \right] \frac{\langle n | \dot{\mathcal{H}} | k \rangle}{\epsilon_n - \epsilon_k} \quad (2.11)$$

We now impose that the m^{th} eigenstate is effectively preserved throughout the evolution, setting the conditions $c_{n \neq m}(t) = 0$ and $c_m(t) = 1$. The new equation

$$\dot{c}_{n \neq m} = \frac{\langle n | \dot{\mathcal{H}} | m \rangle}{\epsilon_n - \epsilon_m} \exp \left[-i \int_0^t dt' (\epsilon_m(t') - \epsilon_n(t')) \right]$$

integrates to

$$c_{n \neq m}(\tau) = \int_0^\tau \frac{\langle n | \dot{\mathcal{H}} | m \rangle}{\epsilon_n - \epsilon_m} \exp \left[-i \int_0^t dt' (\epsilon_m(t') - \epsilon_n(t')) \right] dt \quad .$$

Given the imposed conditions, we expect that $c_{n \neq m}(\tau) = 0$. Therefore, we can formulate the simplified adiabaticity criterion by imposing $|c_{n \neq m}(\tau)| \ll 1$. To quickly remove the integral out of the way, we substitute some quantities with their average value.

$$\begin{aligned} \epsilon_n(t) - \epsilon_m(t) &\longrightarrow \langle \epsilon_n - \epsilon_m \rangle \equiv \langle \Delta_{nm} \rangle \\ \langle n | \dot{\mathcal{H}} | m \rangle (t) &\longrightarrow \langle \partial_t \mathcal{H}_{nm} \rangle \end{aligned}$$

$$\begin{aligned} \implies c_{n \neq m}(\tau) &\simeq \int_0^\tau \frac{\langle \partial_t \mathcal{H}_{nm} \rangle}{\langle \Delta_{nm} \rangle} e^{i \int_0^t dt' \langle \Delta_{nm} \rangle} = \\ &= \frac{\langle \partial_t \mathcal{H}_{nm} \rangle}{\langle \Delta_{nm} \rangle} \cdot \frac{i}{\langle \Delta_{nm} \rangle} \left(e^{i \langle \Delta_{nm} \rangle \tau} - 1 \right) \end{aligned}$$

The upper bound $|c_{n \neq m}(\tau)| \ll 1$ eventually reads as

$$|c_{n \neq m}(\tau)| \leq \frac{\langle \partial_t \mathcal{H}_{nm} \rangle}{\langle \Delta_{nm} \rangle^2} \ll 1 \quad (2.12)$$

Once again, we observe that the criterion involves the off-diagonal elements of $\partial_t \mathcal{H}$ and the gap between eigenstates. The speed of the evolution is indirectly present through the time derivative of the off-diagonal elements of the Hamiltonian.

2.2 Counterdiabatic driving

So far, the Adiabatic Theorem motivates the need for longer evolution times τ . But in real implementations of ASP, much effort is instead made to reduce it. The reasons are essentially two. (1) Smaller evolution times tend to minimize the cost of ASP. Having smaller evolution times, we can at least execute many shots of the same schedule and retrieve more statistics to find the correct ground state. (2) Current hardware is subject to several limitations, one of them being the inability to perform long evolution times τ , due to incoherence arising in the system [23].

One of the possible solutions is to drop the assumptions of adiabaticity and modify the original protocol $\mathcal{H}(t)$ of Eq. (2.1). Counterdiabatic (CD) driving [11] is designed to instantly suppress the transitions towards excited eigenstates, by adding to the protocol a hermitian operator known as the Adiabatic Gauge Potential (AGP).

$$\mathcal{H}_{\text{CD}}(\lambda) = \mathcal{H}(\lambda) + \underbrace{\dot{\lambda} \mathcal{A}_\lambda(\lambda)}_{\text{AGP}}$$

The AGP is an analytical compensation of the transitions that arise when the system evolves in a non-adiabatic regime. In the previous discussions, we have established that the matrix elements $\langle m | \partial_\lambda \mathcal{H} | n \rangle$ are responsible for transitions between eigenstates. Remarkably, the AGP expression acts exactly on those terms.

2.2.a The Adiabatic Gauge Potential

Let us walk through the main steps that lead to the formal definition of the Adiabatic Gauge Potential, following the footsteps of Ref. [53].

Let $|\psi\rangle = \sum_n \psi_n |n\rangle_0$ be the wave function of a quantum state expanded in a generic basis $\{|n\rangle_0\}$. The subscript zero is just used to flag this basis. It is

always possible to perform a unitary transformation U and move to a new basis of choice $\{|m\rangle\}$. Suppose also that such unitary transformation is parameterized by the continuous parameter $\lambda(t)$. Therefore, if

$$|m(\lambda)\rangle = U_{nm}(\lambda) |n\rangle_0$$

the new basis set can be used to spell out the wave function as

$$|\psi\rangle = \sum_m \underbrace{\sum_n \psi_n U_{nm}^*}_{\equiv \tilde{\psi}_m} |m(\lambda)\rangle = \sum_m \tilde{\psi}_m(\lambda) |m(\lambda)\rangle \quad .$$

The same expression in vector notation is way more compact: $|\tilde{\psi}\rangle = U^\dagger(\lambda) |\psi\rangle$. Hereinafter we establish the notation that a tilde operator is to be understood as written in the transformed basis, e.g. $\tilde{\mathcal{O}} = U^\dagger \mathcal{O} U$.

Let us introduce a Hamiltonian $\mathcal{H}(t)$ that describes the dynamics of the system in the time interval $[0, \tau]$. In addition, we fix the unitary transformation such that the Hamiltonian in the new basis $\tilde{\mathcal{H}} = U^\dagger \mathcal{H} U$ is diagonalized, meaning that $\{|m(\lambda)\rangle\}$ is the instantaneous energy eigenbasis. The literature sometimes reports the instantaneous eigenbasis as the *adiabatic basis*.

We now write the Schrödinger equation for the transformed state

$$\begin{aligned} i \frac{d}{dt} |\tilde{\psi}\rangle &= i \left(\frac{dU^\dagger}{dt} |\psi\rangle + U^\dagger \frac{d|\psi\rangle}{dt} \right) = \\ &= i \dot{\lambda} \partial_\lambda U^\dagger \cdot U |\tilde{\psi}\rangle + U^\dagger \mathcal{H} \cdot U |\tilde{\psi}\rangle = \\ &\equiv \underbrace{\left(\tilde{\mathcal{H}} - \dot{\lambda} \tilde{\mathcal{A}}_\lambda \right)}_{\mathcal{H}_{rot}} |\tilde{\psi}\rangle \end{aligned} \quad (2.13)$$

and define, in the last passage, the *gauge potential* $\tilde{\mathcal{A}}_\lambda \equiv -i(\partial_\lambda U^\dagger)U$.

Gauge potentials are found in both classical and quantum mechanics. The Hamiltonian of a classical system is written as a function of canonical variables, which are arbitrary but must satisfy the canonical relations (in terms of Poisson brackets). The canonical transformations are a class of transformations that preserve the canonical relations. Eventually, Gauge potentials are the generators of said (continuous) canonical transformations. Moving to quantum systems, the equivalent of canonical transformations are unitary transformations. Indeed, by introducing U , we have taken advantage of the freedom of choosing the basis to express the quantum states, which is analogous to the freedom of choosing canonical variables. Therefore, we could have introduced the gauge potential as

the generator of the continuous unitary $U(\lambda)$:

$$i\partial_\lambda |\tilde{\psi}(\lambda)\rangle = i\partial_\lambda(U^\dagger |\psi\rangle) = i(\partial_\lambda U^\dagger)U |\tilde{\psi}\rangle = -\tilde{\mathcal{A}}_\lambda |\tilde{\psi}\rangle \quad .$$

Furthermore, gauge potentials are related to quantum geometry concepts. For instance, one can see them as connections and define covariant derivatives $D_\mu = \partial_\mu + i\mathcal{A}_\mu$ that are used to extend the parallel transport of vectors to wave functions: $D_\mu |\phi_n\rangle = 0, \forall$ eigenstates $|\phi_n\rangle$.

Let us discuss some properties. The gauge potential is hermitian

$$\tilde{\mathcal{A}}_\lambda \equiv -i(\partial_\lambda U^\dagger)U = -i\partial_\lambda(UU^\dagger) + iU^\dagger\partial_\lambda U = (\tilde{\mathcal{A}}_\lambda)^\dagger$$

and its matrix elements are

$${}_0\langle n|\tilde{\mathcal{A}}_\lambda|m\rangle_0 = i{}_0\langle n|U^\dagger\partial_\lambda U|m\rangle_0 = i\langle n(\lambda)|\partial_\lambda|m(\lambda)\rangle \quad .$$

The latter equation also implies that the gauge potential in the lab frame is $\mathcal{A}_\lambda = U\tilde{\mathcal{A}}_\lambda U^\dagger = i\partial_\lambda$, which is expected since the gauge is generator of the unitary transformation. We now wish to prove that

Proposition 1. For $m \neq n$,

$$\langle m|\mathcal{A}_\lambda|n\rangle = i\frac{\langle m|\partial_\lambda\mathcal{H}|n\rangle}{\epsilon_n - \epsilon_m} \quad . \quad (2.14)$$

Proof. Follows from the differentiation of the identity $\langle m|\mathcal{H}(\lambda)|n\rangle = 0$ for $m \neq n$:

$$\begin{aligned} \langle \partial_\lambda m|\mathcal{H}|n\rangle + \langle m|\partial_\lambda\mathcal{H}|n\rangle + \langle m|\mathcal{H}|\partial_\lambda n\rangle &= 0 \\ \epsilon_n \langle \partial_\lambda m|n\rangle + \langle m|\partial_\lambda\mathcal{H}|n\rangle + \epsilon_m \langle m|\partial_\lambda n\rangle &= 0 \end{aligned}$$

The underlined terms are related to the matrix elements of the gauge potential in the laboratory frame. $i\langle m|\partial_\lambda n\rangle = \langle n|\mathcal{A}_\lambda|m\rangle$, therefore

$$\langle m|\partial_\lambda\mathcal{H}|n\rangle \frac{-i}{\epsilon_m - \epsilon_n} = \langle m|\mathcal{A}_\lambda|n\rangle \quad .$$

Remark: we have omitted the explicit dependence on λ for simplicity. \square

The result of the proposition can be expressed in matrix form, being careful

to include the diagonal terms:

$$i\partial_\lambda \mathcal{H} = [\mathcal{A}_\lambda, \mathcal{H}] + i \underbrace{\sum_n \frac{\partial \epsilon_n}{\partial \lambda} |n\rangle \langle n|}_{\equiv -M_\lambda} \quad (2.15)$$

$$= [\mathcal{A}_\lambda, \mathcal{H}] - iM_\lambda \quad (2.16)$$

The newly defined operator M_λ represents a non-equilibrium generalized force corresponding to different eigenstates of the Hamiltonian. Because it is defined from the instantaneous eigenbasis, M_λ trivially commutes with the Hamiltonian, $[\mathcal{H}, M_\lambda] = 0$, which is actually a useful property of the gauge potential:

$$[\mathcal{H}, i\partial_\lambda \mathcal{H} - [\mathcal{A}_\lambda, \mathcal{H}]] = 0 \quad (2.17)$$

Let us go back to Eq. (2.13), where we have highlighted the effective Hamiltonian in the rotating frame $\mathcal{H}_{rot} \equiv \tilde{\mathcal{H}} - \dot{\lambda} \tilde{\mathcal{A}}_\lambda$. Looking at its terms in the instantaneous eigenbasis, $\tilde{\mathcal{H}}$ is by construction diagonal, therefore the off-diagonal terms of $\tilde{\mathcal{A}}_\lambda$ must be responsible for transitions between energy levels. This observation opens the way to a new understanding of the gauge potential: the off-diagonal elements generate non-adiabatic corrections of the Hamiltonian in the moving frame. For this reason, the \mathcal{A}_λ is called the Adiabatic Gauge Potential (AGP).

An immediate application of the AGP is transitionless driving. Intuitively, we wish to suppress the transitions between eigenstates to keep invariant the populations of the quantum system throughout its evolution. This condition is realized if the effective Hamiltonian is $\mathcal{H}_{rot} = \tilde{\mathcal{H}}$, which is obtained by adding $\dot{\lambda} \mathcal{A}_\lambda$ to the driving protocol in the laboratory frame:

$$\mathcal{H}_{CD}(\lambda) = \mathcal{H}(\lambda) + \dot{\lambda} \mathcal{A}_\lambda(\lambda) \quad (2.18)$$

2.2.b Approximations of the adiabatic gauge potential

The advantages of CD driving for state preparation are clear. The system excitation can be kept under control, and the evolved state instantly follows the Hamiltonian's ground state, even when the evolution time is far from the adiabatic limit. Nonetheless, several issues must be considered when applying this method to many-body systems.

The AGP is written, by definition, in the basis of instantaneous eigenstates. Even taking for granted the knowledge of the spectrum, it is still difficult to find

a suitable decomposition of the AGP in operators. It is the case in many-body systems that such decomposition demands the use of highly non-local interactions. Furthermore, there is evidence that the gauge potential cannot be local in chaotic systems “because no local operator is expected to be able to distinguish general many-body states with arbitrarily small energy differences” [15, 19].

Another argument of concern is that the gauge potential terms (Eq. (2.14)) can be ill-defined if the energy spectrum is dense, i.e. $\epsilon_{mn} = \epsilon_m - \epsilon_n \rightarrow 0$. This may occur in isolated points such as phase transitions, where the spectral gap becomes exponentially smaller in the system size. In other situations, due to chaotic dynamics, careful regularization is required [53, 19]. There are also proven issues in the thermodynamic limit. The widely accepted eigenstate thermalization hypothesis [19] suggests that excited states are exponentially susceptible to tiny perturbations. Furthermore,

$$\begin{aligned}\epsilon_{mn} &\sim \exp(-\mathcal{S}) \\ \langle m | \partial_\lambda \mathcal{H} | n \rangle &\sim \exp(-\mathcal{S}/2)\end{aligned}$$

where \mathcal{S} is the extensive thermodynamic entropy of the system. As a consequence, the off-diagonal matrix elements in Eq. (2.14) scale as $\exp(\mathcal{S}/2)$, so the AGP elements are exponentially divergent with the system size.

In light of the above considerations, there exists a solution that addresses most of the issues of AGP at once. Instead of seeking an exact gauge potential, let us settle for an approximation that still serves the purpose of avoiding transitions to excited states, with variable efficiency, during the state preparation sweep.

Two approximations are going to be discussed. We begin by presenting a method for calculating the AGP through a minimum action principle. This strategy allows the use of arbitrary ansatzes with arbitrarily restricted support, while still retaining most of the effectiveness of the CD driving in most applications. Secondly, we show an expansion of AGP in nested commutators of the original protocol Hamiltonian.

Variational minimization

Consider an hermitian operator \mathcal{X} and define

$$G_\lambda(\mathcal{X}) \equiv \partial_\lambda \mathcal{H} - i [\mathcal{H}, \mathcal{X}] \quad (2.19)$$

Because of Eq. (2.16), $G_\lambda(\mathcal{A}_\lambda) = -M_\lambda$. It can be difficult to find the AGP by solving the previous equation because it requires constructing the operator matrix explicitly in the full Hilbert space. Instead, the problem can be reformulated as a minimization, in terms of operator distance, between $G_\lambda(\mathcal{X})$ and $-M_\lambda$.

$$D(\mathcal{X}) \underset{\substack{\uparrow \\ \text{Frobenius norm}}}{=} \text{Tr} \left[(G_\lambda + M_\lambda)^2 \right] = \text{Tr} (G_\lambda^2) + 2 \text{Tr} (G_\lambda M_\lambda) + \text{Tr} (M_\lambda^2)$$

$$\begin{aligned} \text{where } \text{Tr} (G_\lambda M_\lambda) &= \text{Tr} (\partial_\lambda \mathcal{H} M_\lambda) + i \text{Tr} ([\mathcal{X}, \mathcal{H}] M_\lambda) = \\ &= - \text{Tr} (M_\lambda^2) + i \text{Tr} ([\mathcal{H}, M_\lambda] \mathcal{X}) \\ &\quad \uparrow \qquad \qquad \qquad \uparrow \\ &\quad \text{Eq. (2.16)} \qquad \text{cyclic property} \end{aligned}$$

$\Rightarrow D(\mathcal{X}) = \text{Tr} (G_\lambda^2) - \text{Tr} (M_\lambda^2)$. Since M_λ does not depend on \mathcal{X} , the second term is a constant that can be removed from the minimization procedure. The operation that ultimately leads to the AGP solution is to minimize G 's Frobenius norm. Furthermore, the distance over \mathcal{X} can be interpreted as the action associated to the gauge potential

$$S = \text{Tr} [G_\lambda^2(\mathcal{X})] \tag{2.20}$$

and the minimization procedure is formally equivalent to an Euler–Lagrange equation.

$$\delta S / \delta \mathcal{X} = 0$$

The variational minimization of Eq. (2.20) is a powerful tool to determine expressions of the AGP in numerous practical situations. Given a parameterized ansatz $\mathcal{X}(\underline{\alpha})$, its variational minimization returns the optimized parameters $\underline{\alpha}$ that approximate at best the AGP. It remains to determine which operators must be included in the ansatz. From Eq. (2.19), it is clear that the ansatz must not commute with the Hamiltonian, otherwise, the optimization procedure is not able to determine any solution. Local operators are the best options in practical applications because they are better suited to current hardware implementations. Nonetheless, including a finite set of more complex many-body interactions improves the approximation accuracy whilst avoiding the previously mentioned divergence issues of the AGP.

As a final remark, we observe that if \mathcal{X} is a good approximation of the AGP, then the property of Eq. (2.17) must hold: $[\mathcal{H}, G_\lambda] \simeq 0$. The quantity $[\mathcal{H}, G_\lambda]$ can be regarded as a good measure of the approximation error.

Nested commutator expansion

A general approximated ansatz for the AGP has been proposed in Ref. [15]:

$$\mathcal{A}_\lambda^{(\ell)} \equiv i \sum_{k=1}^{\ell} \alpha_k \underbrace{[\mathcal{H}, [\mathcal{H}, \dots [\mathcal{H}, \partial_\lambda \mathcal{H}]]]}_{2k-1} \quad (2.21)$$

where ℓ is the order of the expansion. Although the expression above can represent the exact potential at the limit $\ell \rightarrow \infty$, it is of practical interest to choose a finite value of ℓ . A finite order expansion is a well-defined gauge potential ansatz, that can be optimized variationally - through the machinery of Eq. (2.20) - to determine the optimal set of coefficients $\underline{\alpha}(\lambda) = (\alpha_1(\lambda), \dots, \alpha_\ell(\lambda))$. The calculation of the nested commutators returns a finite set of operators with restricted support. Even though there is still no guarantee that the operators are implementable in the experiment, the expansion provides a systematic method to identify the leading operator contributions.

There are several ways to motivate the expression of Eq. (2.21). The most simple way is to observe that the matrix elements of the ansatz can be expressed as a power series multiplying the state-dependent factor $\langle m | \partial_\lambda \mathcal{H} | n \rangle$:

$$\begin{aligned} \langle m | \mathcal{A}_\lambda | n \rangle &= i \sum_{k=1}^{\ell} \alpha_k \langle m | \underbrace{[\mathcal{H}, [\mathcal{H}, \dots [\mathcal{H}, \partial_\lambda \mathcal{H}]]]}_{2k-1} | n \rangle = \\ &= i \underbrace{\left[\sum_{k=1}^{\ell} \alpha_k (\omega_m - \omega_n)^{2k-1} \right]}_{a_\lambda^{(\ell)}(\omega_m - \omega_n)} \langle m | \partial_\lambda \mathcal{H} | n \rangle \end{aligned} \quad (2.22)$$

The latter expression matches the form of the AGP matrix elements¹

$$\langle m | \mathcal{A}_\lambda | n \rangle \stackrel{\text{Eq. (2.14)}}{=} -i \frac{\langle m | \partial_\lambda \mathcal{H} | n \rangle}{\omega_m - \omega_n}$$

if $a_\lambda^{(\ell)}(\omega_{mn}) \simeq -\omega_{mn}^{-1}$ for a good range of energy gaps $\omega_{mn} = \omega_m - \omega_n$. The approximation between the prefactor of the AGP elements and the power series is ill-defined at the limits $\omega_{mn} \rightarrow 0$ and $\omega_{mn} \rightarrow \infty$. However, none of those presents a serious limitation when the order of the ansatz is finite. On the first hand, the matrix elements of local operators at high ω_{mn} decay exponentially

¹We switch ϵ_n with the frequency ω_n to match the notation of the original Ref. [15]. This change of symbols is not relevant since we operate already in natural units: $\epsilon_n = \hbar\omega_n = \omega_n$.

[19], hence the contribution to the gauge potential is negligible. On the other hand, it is not crucial to suppress the transitions between eigenstates that occur in the small energy gap regime $\omega_{mn} \simeq 0$: intuitively, if the instantaneous energy gap through the schedule is Δ , only transitions with $\omega_{mn} \geq \Delta$ have to be suppressed. Even if $\Delta \sim 0$, the evolution of the system will result in excited states that are confined to a narrow energy shell of the spectrum.

Regularized gauge potentials [15] provide additional motivation for the ansatz:

$$\mathcal{A}_\lambda = \lim_{\epsilon \rightarrow 0^+} \int_0^\infty dt e^{-\epsilon t} \left(e^{-i\mathcal{H}(\lambda)t} \partial_\lambda \mathcal{H}(\lambda) e^{i\mathcal{H}(\lambda)t} - M_\lambda \right) \quad (2.23)$$

whose matrix elements resemble the ones of Eq. (2.14),

$$\begin{aligned} \langle m | \mathcal{A}_\lambda | n \rangle &= \lim_{\epsilon \rightarrow 0^+} \int_0^\infty dt e^{-\epsilon t} e^{-i\omega_{mn}t} \langle m | \partial_\lambda \mathcal{H} | n \rangle = \\ &= \lim_{\epsilon \rightarrow 0^+} \frac{\langle m | \partial_\lambda \mathcal{H} | n \rangle}{\epsilon + i\omega_{mn}}. \end{aligned}$$

The Baker-Campbell-Hausdorff formula expands the operators in a series of nested commutators:

$$e^{-i\mathcal{H}(\lambda)t} \partial_\lambda \mathcal{H}(\lambda) e^{i\mathcal{H}(\lambda)t} = \sum_{k=0}^{\infty} \frac{(-it)^k}{k!} \underbrace{[\mathcal{H}, [\mathcal{H}, \dots [\mathcal{H}, \partial_\lambda \mathcal{H}]]]}_k$$

Even order commutators are real (for real Hamiltonians) and contribute to M_λ , the odd commutators instead have an imaginary coefficient and contribute to \mathcal{A}_λ .

Even though the framework of AGP is suitable to determine optimal protocols analytically, their realization on hardware is still quite challenging, as CD protocols demand the control of additional interactions w.r.t. the reference (unassisted) protocol. This statement is easy to verify: the nested commutators of the power series expansion, in Eq. (2.21), would return highly non-local many-body interactions if the order of the ansatz ℓ is sufficiently high.

Nonetheless, we will see soon that the AGP power series expansion is useful for designing a family of protocols that address such limitation of CD protocols. Indeed, the CD protocols can be engineered as Floquet protocols [15] which do not require additional controls. Floquet protocols are central in the developments of this thesis and will be soon introduced in §2.4.

2.3 Quantum Annealing

One of the currently available technologies for solving on-hardware quantum optimization problems is Quantum Annealing (QA) [45, 41, 81]. The approach follows the same prescription of ASP, in which a quantum system is driven from an initial state towards the ground state of a final target Hamiltonian. But unlike ASP, QA does not require adiabaticity and it does not insist on universality. As a matter of fact, QA hardware is a compromise between idealized assumptions of AQC and experimental constraints [83].

The concept of (classical) *annealing* was developed to solve minimization problems that are affected by the presence of many local minima. Annealing is a heuristic algorithm, which is based on stochastic iterative improvements. Heuristic algorithms, in contrast to exact algorithms, do not guarantee the optimality of the solution but are effective in many problems.

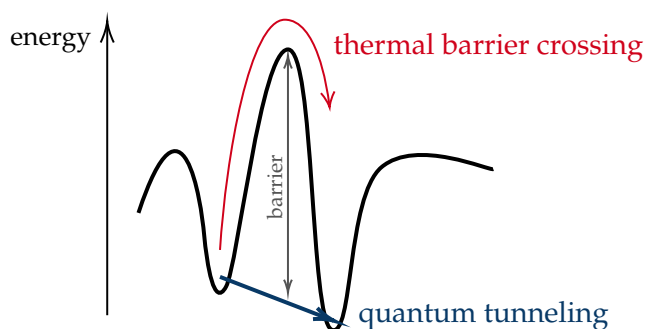
Excellent examples of heuristic optimization are stochastic gradient descent algorithms, which are extensively applied in machine learning. An issue that frequently arises in minimization problems is the presence of jagged loss landscapes with multiple local minima. Gradient descent algorithms might easily get stuck in one of these minima, returning a sub-optimal solution. A rather easy mitigation is to include fluctuations (noise) in each iteration [86], providing some randomness in the gradient descent updates that improve the optimizer capability to escape from local minima.

The concept of random noise can be formalized within statistical mechanics as a *thermal fluctuation*, with the temperature T playing the role of a noise factor [20]. Higher temperatures produce strong fluctuations. Simulated Annealing (SA) is a method for solving optimization problems that takes advantage of thermal fluctuations in a simulated environment [50]. The artificial temperature is progressively relaxed throughout the annealing schedule, approaching the zero-temperature limit. Eventually, the thermal noise is turned off, allowing the system to settle in one configuration.

Optimization algorithms that exploit spin-glass dynamics show a coarse (quantum) energy landscape, with many local minima. Introducing fluctuations is a reasonable mitigation method, even though there might be further complications. Indeed, in a glassy landscape, local minima might be surrounded by high energy barriers, making the optimization based on thermal noise less effective. Furthermore, the energetic barriers may be proportional to the system size and diverge in the thermodynamic limit, especially if the underlying interactions

have an infinite range [20, 76].

Luckily, the quantum nature of the system allows us to take advantage of another type of statistical fluctuations, which are due to *quantum tunneling*. Fluctuations due to quantum tunneling behave differently from (classical) thermal fluctuations. In the latter, the system has to gain sufficient energy in order to cross an energy barrier. Instead, in quantum tunneling, the crossing is understood in terms of probability and can occur even if the system does not match the energy requirement, provided the barrier is thin enough.



Quantum tunneling between states can be induced in spin systems through additional transverse fields. Therefore, the key element of QA is the use of transverse fields. We set the initial Hamiltonian of the schedule to

$$\mathcal{H}_i = - \sum_k^N \sigma_x^{(k)}$$

whose ground state is separable (i.e. can be prepared through local operations):

$$|\psi(0)\rangle = |+\rangle^{\otimes N} = \frac{1}{\sqrt{2}} (|0\rangle + |1\rangle)_1 \otimes \cdots \otimes \frac{1}{\sqrt{2}} (|0\rangle + |1\rangle)_N \quad .$$

Given the previous discussions, a schedule like $\mathcal{H}(t) = A(t)\mathcal{H}_i + B(t)\mathcal{H}_p$ mimics the same dynamics of SA, where the transverse field is being gradually turned off by the end of the sweep.

Even though we have introduced QA as a variant of ASP without the requirements of adiabaticity, the dynamics of the evolution are still subject to the consequences of the Adiabatic Theorem [43]. Particularly, shorter annealing times τ produce states with smaller fidelity w.r.t. the true ground state. However, it also means that QA can benefit from methods that mitigate the issues of ASP, such as Counterdiabatic driving.

2.4 Floquet engineered counterdiabatic protocols

The total evolution time is crucial in ASP algorithms. Unfortunately, the two extremes of the timescales present competitive limitations. From one point of view, it is compelling to relax the assumptions of adiabaticity: current hardware is inevitably subject to decoherence and noise, resulting in limited maximum evolution time. On another perspective, a faster driving speed produces diabatic excitations, undermining the original purpose of ASP and tampering with the preparation of a ground state.

A class of mitigations known as “shortcuts to adiabaticity” proposes protocols that approximate adiabatic dynamics when a system evolves in a finite time. We have focused our previous discussions on CD driving, which has been derived from the idea of instantaneously suppressing the dissipative losses of a quantum system that undergoes diabatic evolution. However, the construction of such gauge potential relies on the full knowledge of the spectrum, which forces in practical applications the use of approximations. Furthermore, a counterdiabatic protocol demands additional² controls w.r.t. the original adiabatic protocol. However, the implementation of those non-local multi-body interactions on hardware is not guaranteed.

The method [15] that we will now introduce stands as an ideal solution, achieving counterdiabaticity without requiring any additional controls over the quantum system. These novel protocols make use of stroboscopic driving that effectively reproduces the counterdiabatic dynamics in each driving cycle. In practice, a high-frequency oscillation is modulated in amplitude with coefficients that are determined from the power series expansion, Eq. (2.22), of the AGP.

A **Floquet Engineered (FE) protocol** is characterized by the Hamiltonian

$$\mathcal{H}_{\text{FE}} = \left[1 + \frac{\omega}{\omega_0} \cos \omega t \right] \mathcal{H} + \dot{\lambda} \underbrace{\left[\sum_k^{\infty} \beta_k \sin((2k-1)\omega t) \right]}_{\beta(t)} \partial_{\lambda} \mathcal{H} \quad (2.24)$$

The Fourier coefficients β_k modulate the driving frequency ω and its odd multiples, whereas ω_0 is just a parameter that arises from the protocol derivation, representing a reference frequency on the spectrum transition frequencies.

The leading order contributions in a regime of infinite-frequency driving are obtained by moving to the rotating frame w.r.t. $\mathcal{H}(\lambda) \cos(\omega t) \omega / \omega_0$. In the transformation, $\mathcal{H}(\lambda)$ can be treated as constant in each driving cycle, which leads to

²In this context, ‘additional’ refers to any other operator that is not present either in \mathcal{H}_i or \mathcal{H}_p .

the rotating frame Floquet Hamiltonian:

$$\begin{aligned}\tilde{\mathcal{H}}_{\text{FE}}(t) &= \exp\left(i\frac{\omega}{\omega_0}\int_0^t ds \cos(\omega s)\mathcal{H}(\lambda)\right) \mathcal{H}_{\text{FE}} \exp\left(-i\frac{\omega}{\omega_0}\int_0^t ds \cos(\omega s)\mathcal{H}(\lambda)\right) \\ &= e^{i\frac{\sin\omega t}{\omega_0}\mathcal{H}(\lambda)} \mathcal{H}_{\text{FE}} e^{-i\frac{\sin\omega t}{\omega_0}\mathcal{H}(\lambda)}\end{aligned}\quad (2.25)$$

The time evolution operator is written using the time-ordered exponential in each driving cycle of period $T = \omega/2\pi$:

$$e^{-i\tilde{\mathcal{H}}_F T} = \mathcal{T} \exp\left(-i\int_t^{t+T} \tilde{\mathcal{H}}_{\text{FE}}(t') dt'\right)$$

The Magnus expansion allows us to solve the equation for the effective Hamiltonian $\tilde{\mathcal{H}}_F$. We limit the expansion to the first order, obtaining

$$\begin{aligned}\tilde{\mathcal{H}}_F T &= \int_0^T \tilde{\mathcal{H}}_{\text{FE}}(t_1) dt_1 + \frac{i}{2} \int_0^T \left[\int_0^{t_1} \tilde{\mathcal{H}}_{\text{FE}}(t_2) dt_2, \tilde{\mathcal{H}}_{\text{FE}}(t_1) \right] + \dots = \\ &= \int_0^T e^{i\frac{\sin\omega t}{\omega_0}\mathcal{H}(\lambda)} \mathcal{H}_{\text{FE}} e^{-i\frac{\sin\omega t}{\omega_0}\mathcal{H}(\lambda)} dt \\ &\quad \uparrow \\ &\quad \text{1st order}\end{aligned}\quad (2.26)$$

It is convenient to write the matrix elements in the instantaneous eigenbasis. The off-diagonal elements are

$$\langle m | \tilde{\mathcal{H}}_F^{(1st)} | n \rangle = \frac{\dot{\lambda}}{T} \int_0^T dt e^{i\frac{\sin\omega t}{\omega_0}(\omega_m - \omega_n)} \beta(t) \langle m | \partial_\lambda \mathcal{H} | n \rangle$$

and using the Jacobi–Anger identity on the exponential we obtain

$$\begin{aligned}\langle m | \tilde{\mathcal{H}}_F^{(1st)} | n \rangle &= \dot{\lambda} \sum_{k=-\infty}^{\infty} \mathcal{J}_k\left(\frac{\epsilon_m - \epsilon_n}{\omega_0}\right) \langle m | \partial_\lambda \mathcal{H} | n \rangle \cdot \underbrace{\frac{1}{T} \int_0^T dt e^{ik\omega t} \beta(t)}_{k\text{th coefficient of } \beta(t)} \\ &= i\dot{\lambda} \sum_{k=1}^{\infty} \mathcal{J}_{2k-1}\left(\frac{\omega_m - \omega_n}{\omega_0}\right) \beta_k \langle m | \partial_\lambda \mathcal{H} | n \rangle \\ &= \langle m | \tilde{\mathcal{H}}_F^{(1st)} | n \rangle \\ &\quad \uparrow \\ &\quad \text{at multiples of } T\end{aligned}\quad (2.27)$$

The last equality is justified by observing that, by construction, the rotating frame overlaps with the laboratory frame at integer multiples of T . $\mathcal{J}_r(\cdot)$ denotes the Bessel functions of order r . Let $\omega_{mn} \equiv \omega_m - \omega_n$. Eventually, the effective

Hamiltonian is written as $\mathcal{H}_F = \mathcal{H} + \dot{\lambda}\mathcal{A}_F$, where

$$\langle m | \mathcal{A}_F | n \rangle = i \sum_{k=1}^{\infty} \beta_k \mathcal{J}_{2k-1} \left(\frac{\omega_{mn}}{\omega_0} \right) \langle m | \partial_{\lambda} \mathcal{H} | n \rangle .$$

After all these manipulations, the expression of the effective Floquet Hamiltonian \mathcal{H}_F is written in the same form as the CD protocol of Eq. (2.18). Furthermore, it can be matched to the off-diagonal terms of the AGP ansatz of Eq. (2.22). Indeed, the Bessel functions can be expanded as power series,

$$\langle m | \mathcal{A}_F | n \rangle = i \sum_{k=1}^{\infty} \beta_k \sum_{m=0}^{\infty} \frac{(-1)^m}{m!(2k+m-1)!} \left(\frac{\omega_{mn}}{2\omega_0} \right)^{2k+2m-1} \langle m | \partial_{\lambda} \mathcal{H} | n \rangle .$$

and comparison with the power series AGP implies the equality

$$\sum_{k=1}^{\ell} \alpha_k \omega_{mn}^{2k-1} = \sum_{k=1}^{\infty} \beta_k \sum_{m=0}^{\infty} \frac{(-1)^m}{m!(2k+m-1)!} \left(\frac{\omega_{mn}}{2\omega_0} \right)^{2k+2m-1} . \quad (2.28)$$

We observe that in order to make this match, it has been crucial to choose a $\beta(t)$ that selects only odd coefficients in Eq. (2.27). Indeed, the AGP ansatz shows only odd powers of ω_{mn} . This condition is realized by considering a generic $\beta(t)$ expanded in Fourier series, with non-null coefficients associated with sine waves of frequency $(2k-1)\omega$.

The condition that matches AGP ansatz with the effective Floquet dynamics can be developed up to any order ℓ . Let us start from the lowest order ($\ell = 1$):

$$\alpha_1 \omega_{mn} = \sum_{k=1}^{\infty} \beta_k \sum_{m=0}^{\infty} \frac{(-1)^m}{m!(2k+m-1)!} \left(\frac{\omega_{mn}}{2\omega_0} \right)^{2k+2m-1} \quad \forall \omega_{mn} .$$

The only power to match on LHS selects $2k+2m-1 = 1$ on RHS, thus the only values that are to be summed are $k = 1$ and $m = 0$, leading to

$$\alpha_1 = \beta_1 \frac{1}{0!1!} (2\omega_0)^{-1} \implies \beta_1 = 2\alpha_1 \omega_0 \quad (2.29)$$

Proceeding likewise to the second order, we have to match up to ω_{mn}^3 , so

$$\alpha_1 \omega_{mn} + \alpha_2 \omega_{mn}^3 = \beta_1 \frac{\omega_{mn}}{2\omega_0} - \beta_1 \frac{\omega_{mn}^3}{16\omega_0^3} + \beta_2 \frac{\omega_{mn}^3}{48\omega_0^3} \quad \forall \omega_{mn}$$

leading to the condition

$$\beta_2 = 6\omega_0\alpha_1 + 48\omega_0^3\alpha_2 \quad . \quad (2.30)$$

In conclusion, the Floquet protocol proposed in Ref. [15] consists of a stroboscopic driving that effectively reproduces the counterdiabatic dynamics in each driving cycle. The key advantage is that the set of required controls on the system is not modified, since only the interactions figuring in the original protocol (\mathcal{H}_i and \mathcal{H}_p) appear in the Floquet protocol.

Floquet engineering Parity annealing protocols

The preceding Chapters have reviewed all the topics at the heart of our work. Chapter 1 was a preface on Quantum Optimization, which culminated in the introduction of the LHZ/Parity architecture and its key advantages. Chapter 2 introduced the framework of Adiabatic State Preparation and argued the limitations posed by the quantum Adiabatic Theorem. We have then introduced Counterdiabatic driving as a feasible solution to such problems. The key element of a CD protocol is the addition of a gauge potential \mathcal{A}_λ that compensates for the losses of a quantum system in non-adiabatic regimes, effectively suppressing transitions towards excited eigenstates. We discussed the expansion of \mathcal{A}_λ (Eq. (2.22)) that determines a suitable operator basis to approximate the AGP. Despite the resulting operator basis having limited support, the set of required interactions is inevitably expanded, which makes more challenging the implementations on annealing hardware. Nevertheless, the same expansion of the AGP can be used to engineer a counterdiabatic Floquet protocol that does not require additional controls [15].

This chapter will cover the derivation and characterization of Floquet counterdiabatic protocols for the Parity architecture. At first, we determine the driving protocol at two leading orders ℓ of the CD ansatz. Our first milestone is to characterize how the yield of Floquet protocols depends on the choice of the driving frequency and the reference frequency. In doing so, we make extensive use of exact simulations on paradigmatic LHZ instances. Our work will point out a correlation between the yield of the Floquet protocol and errors affecting their mapping to the reference CD schedules. The analysis will proceed by defining a notion of “cost”, which will be used to assess the yield of Floquet protocols against the unoptimized annealing schedules. Eventually, in Chapter 4, we propose new Floquet protocols and repeat the same cost/yield analysis.

3.1 Floquet Parity protocols

The goal of this section is to derive explicitly the Floquet Engineered (FE) protocols for the Parity architecture. Therefore, we fix the problem Hamiltonian

$$\mathcal{H}_p = \sum_k^N J_k \hat{\sigma}_z^{(k)} + \sum_l^{N_c} C_l \hat{\rho}_z^{[l]} \quad , \quad (3.1)$$

where N denotes the number of physical qubits and N_c is the number of constraints. As in a typical Quantum Annealing setup, the initial Hamiltonian consists of transverse fields applied to each physical qubit.

$$\mathcal{H}_i = - \sum_k^N \hat{\sigma}_x^{(k)}$$

Because of this choice, the initial state of the annealing schedule is set to $|\psi(0)\rangle = |+\rangle^{\otimes N} = (|0\rangle + |1\rangle)^{\otimes N} / \sqrt{2}$. We stress that \mathcal{H}_i and \mathcal{H}_p are not time-dependent.

Floquet protocols provide effective counterdiabatic dynamics by modifying the Hamiltonian's driving terms. The original, *unassisted* protocol will be addressed henceforth as the Unassisted Annealing (UA) protocol. The UA Hamiltonian we consider is

$$\mathcal{H}_{\text{UA}}(t) = A(t)\mathcal{H}_i + B(t)\mathcal{H}_p \quad . \quad (3.2)$$

$A(t) = 1 - \lambda(t)$ and $B(t) = \lambda(t)$ are two parameterized interpolation functions. According to the reference work [15], the parameterization is a squared-sine function with null first derivatives at the boundaries:

$$\lambda(t) = \sin^2 \left(\frac{\pi}{2} \sin^2 \left(\frac{\pi t}{2\tau} \right) \right) \quad .$$

A plot of the schedule functions is provided in Figure 3.1.

Whereas CD driving expands the set of control operators, FE protocols retain the original protocol controls. It is rather simple to prove it explicitly. Consider the Floquet Hamiltonian of Eq. (2.24) and substitute \mathcal{H} from the reference UA protocol, then separate the coefficients from the time-independent Hamiltonian:

$$\begin{aligned} \mathcal{H}_{\text{FE}} &= \underbrace{\left[1 + \frac{\omega}{\omega_0} \cos \omega t \right]}_{\eta_1} \mathcal{H} + \beta(t) \partial_t \mathcal{H} = \\ &= \eta_1 [A(t)\mathcal{H}_i + B(t)\mathcal{H}_p] + \beta(t) [\partial_t A(t)\mathcal{H}_i + \partial_t B(t)\mathcal{H}_p] \end{aligned}$$

$$\begin{aligned}
&= \underbrace{\left[\eta_1 A(t) + \beta(t) \partial_t A \right]}_{A_{\text{FE}}(t)} \mathcal{H}_i + \underbrace{\left[\eta_1 B(t) + \beta(t) \partial_t B \right]}_{B_{\text{FE}}(t)} \mathcal{H}_p \\
&\implies \mathcal{H}_{\text{FE}} = A_{\text{FE}}(t) \mathcal{H}_i + B_{\text{FE}}(t) \mathcal{H}_p \quad . \quad (3.3)
\end{aligned}$$

The updated schedule functions are written as

$$\begin{cases} A_{\text{FE}}(t) = \left(1 + \frac{\omega}{\omega_0} \cos \omega t \right) (1 - \lambda(t)) - \dot{\lambda}(t) \sum_k \beta_k(t) \sin((2k-1)\omega t) \\ B_{\text{FE}}(t) = \left(1 + \frac{\omega}{\omega_0} \cos \omega t \right) \lambda(t) + \dot{\lambda}(t) \underbrace{\sum_k \beta_k(t) \sin((2k-1)\omega t)}_{\beta(t)} \end{cases} \quad (3.4)$$

The Fourier coefficients of $\beta(t)$ are determined through the equality in Eq. (2.28). Up to second order $\ell = 2$ of the counterdiabatic protocol, the mapping is

$$\ell = 1 : \left\{ \beta_1(t) = 2\omega_0 \alpha_1(t) \right. \quad \text{and} \quad \ell = 2 : \left\{ \begin{aligned} \beta_1(t) &= 2\omega_0 \alpha_1(t) \\ \beta_2(t) &= 6\omega_0 \alpha_1(t) + 48\omega_0^3 \alpha_2(t) \end{aligned} \right. \quad (3.5)$$

Now, in order to fully determine the driving terms, the AGP coefficients α_1 and α_2 must be calculated. The FE protocols have been derived by taking into consideration the nested commutator AGP ansatz of Eq. (2.21),

$$\mathcal{A}_\lambda^{(\ell)} = i \sum_{k=1}^{\ell} \alpha_k \underbrace{[\mathcal{H}, [\mathcal{H}, \dots [\mathcal{H}, \partial_\lambda \mathcal{H}]]]}_{2k-1} \quad ,$$

whose coefficients α are determined by the variational minimization of Eq. (2.20).

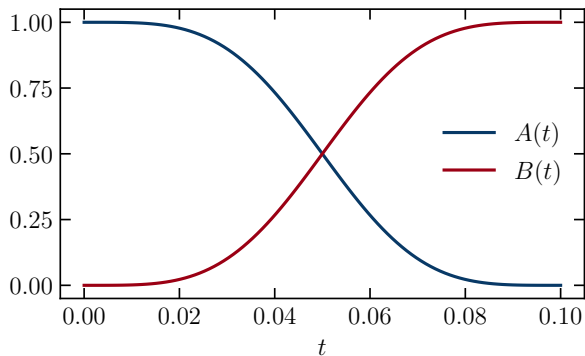


Figure 3.1: Example of the functions $A(t)$ and $B(t)$ that are used in the UA protocol of Eq. (3.2). The two functions interpolate between the \mathcal{H}_i and \mathcal{H}_p .

For this operation, the $2k - 1$ nested commutators must be developed explicitly; at each step, the commutators return numerous many-body interaction terms. Eventually, one must write G_λ (see Eq. (2.19)), square it, and compute its trace. Given the extent of the explicit calculations, we find it more convenient to postpone them in Appendix A. The results of interest are the optimized coefficients α_1 and α_2 for the driving at order $\ell = 1, 2$ respectively. Notably, the optimized CD driving coefficients are written as a function of the schedule functions - $A(t)$ and $B(t)$ - and the parameters of the specific Parity architecture instance (J_k, C_l), namely the local fields and the constraint coefficients.

For instance, the optimized α_1 expression is reported in Eq. (A.4):

$$\alpha_1(t) = -\frac{1}{4} \frac{\sum_l C_l^2 n_l + \sum_k J_k^2}{c_1 A^2(t) + c_2 B^2(t)} \quad , \quad (3.6)$$

where c_1, c_2 and n_l are suitably defined in the Appendix A. Instead, the expression of $\alpha_2(t)$ has been optimized numerically for convenience.

3.1.a Additional controls of the counterdiabatic protocol

It is insightful to look at the interactions that result from the variational CD ansatz. At higher orders of the ansatz ℓ , the resulting operators grant a better approximation of the adiabatic gauge, which in turn allows achieving higher fidelities on the prepared ground state. However, each order carries further control terms to be driven in the protocol, due to the $(2\ell - 1)$ -nested order commutators of Eq. (2.21). The support of the additional operators is still limited but might involve complex non-local interaction patterns, especially at higher ansatz orders. This aspect is particularly accentuated in the Parity architecture, which involves three- or four-body interactions already in the problem Hamiltonian.

At first order CD driving, $\ell = 1$, the number of additional controls is actually only two¹. The first additional term is a local control field on each site, $\sum_i J_i \hat{\sigma}_y^{(i)}$, which arises intuitively from the commutator $\sum_i [\hat{\sigma}_x^{(i)}, J_k \hat{\sigma}_z^{(i)}]$. The second set of additional terms is made of “non-homogeneous” 4-body constraints, acting with $\hat{\sigma}_y$ on one qubit and with $\hat{\sigma}_z$ on the others:

$$\begin{array}{c} y \\ \begin{array}{|c|c|} \hline a & b \\ \hline d & c \\ \hline \end{array} \\ z \end{array} \quad \equiv \quad \hat{\sigma}_y^{(a)} \otimes \hat{\sigma}_z^{(b)} \otimes \hat{\sigma}_z^{(c)} \otimes \hat{\sigma}_z^{(d)} \quad (3.7)$$

¹For the complete expression, refer to Eq. (A.2).

The previous interaction term comes from the commutator $[\sum_i \hat{\sigma}_x^{(i)}, C_l \hat{\sigma}_z^{[l]}]$, where qubit i is one of the qubits (a, b, c, d) in the constraint $[l]$. Because the $\hat{\sigma}_x$ is applied on each physical qubit, the commutator returns an additional 4-body constraint permuting the y control on each qubit:

$$\begin{array}{c} z \\ \blacksquare \\ z \end{array} \begin{array}{c} y \\ \\ \end{array} = \hat{\sigma}_z^{(a)} \hat{\sigma}_y^{(b)} \hat{\sigma}_z^{(c)} \hat{\sigma}_z^{(d)} \quad \begin{array}{c} z \\ \blacksquare \\ z \end{array} \begin{array}{c} z \\ \\ y \end{array} = \hat{\sigma}_z^{(a)} \hat{\sigma}_z^{(b)} \hat{\sigma}_y^{(c)} \hat{\sigma}_z^{(d)} \quad \dots$$

We underline that each of these new interaction patterns is applied on the same sets of qubits as the original constraints $\{C_l \hat{\rho}_z^{[l]}: l = 1, \dots, N_c\}$.

At the second order ($\ell = 2$) of the CD ansatz, the problem of the additional controls becomes substantially more complex. The new 4-body controls show more interaction patterns:

$$\begin{array}{c} x \\ \blacksquare \\ z \end{array} \begin{array}{c} y \\ \\ z \end{array} \quad \begin{array}{c} y \\ \blacksquare \\ z \end{array} \begin{array}{c} y \\ \\ z \end{array} \quad \begin{array}{c} x \\ \blacksquare \\ z \end{array} \begin{array}{c} x \\ \\ z \end{array}$$

As in the previous case, each new control pattern is present in all the unique permutations of x, y, z . Furthermore, new controls involve qubits belonging to neighbor constraints. Two examples are the following 5- and 6-body interactions:

$$\begin{array}{c} z \\ \blacksquare \\ z \end{array} \begin{array}{c} x \\ \\ z \end{array} \quad \begin{array}{c} z \\ \blacksquare \\ z \end{array} \begin{array}{c} z \\ \\ z \end{array} \quad \begin{array}{c} z \\ \blacksquare \\ z \end{array} \begin{array}{c} z \\ \\ y \end{array} \quad \begin{array}{c} z \\ \blacksquare \\ z \end{array} \begin{array}{c} z \\ \\ z \end{array}$$

This is due to the triple nested commutator that introduces terms that can involve up to three neighboring constraints.

The complex interaction patterns shown in the previous paragraphs motivate the need to find alternatives to high-order counterdiabatic protocols. Even though counterdiabaticity allows us to find optimal driving schedules analytically, improving the fidelity of the ground state, this approach requires an extensive pattern of interactions, which is unavailable for current quantum devices. After all, one of the primary benefits of the Parity architecture is to be hardware-friendly, reducing an all-to-all connection scheme to just quasi-local interactions, so introducing numerous extra non-local terms undermines its purpose.

label	protocol	ansatz order
UA	$\mathcal{H}_{\text{UA}} \stackrel{\text{Eq. (3.2)}}{=} A(t)\mathcal{H}_i + B(t)\mathcal{H}_p$	-
CD	$\mathcal{H}_{\text{CD}} \stackrel{\text{Eq. (2.18)}}{=} \mathcal{H}_{\text{UA}} + \dot{\lambda}\mathcal{A}^{(\ell)}(\lambda)$	$\ell = 1, 2$
FE	$\mathcal{H}_{\text{FE}} \stackrel{\text{Eq. (3.3)}}{=} A_{\text{FE}}(t)\mathcal{H}_i + B_{\text{FE}}(t)\mathcal{H}_p$	$\ell = 1, 2$

Table 3.1: Simulated protocol labels.

3.1.b Introducing the simulation instances

To characterize the performance of FE protocols, we run an extensive set of simulations on several LHZ instances. In particular, our problem demands standard solvers for the unitary time-dependent Schrödinger equation. The simulations have been realized in QuTip [44], which makes use of a Variable-coefficient Ordinary Differential Equation (ZVODE) integrator.

For quick reference, we conveniently label the annealing protocols as in Table 3.1: the “standard” unassisted protocol as UA, the counterdiabatic as CD and the related Floquet Engineered protocols as FE. The simulations must be run for a particular problem instance \mathcal{H}_p (Eq. (3.1)). We choose as the main study case a non-trivial LHZ instance with $N = 6$ physical qubits. To assess if our considerations are valid also for slightly larger system sizes, we extend the analysis to an LHZ instance with $N = 10$ physical qubits. Given the extent of the simulations, it is quite demanding to raise the number of qubits to higher values.

Another argument of concern is whether the simulation results are affected by the choice of a specific LHZ instance. To rule out this problem, we have selected three distinct paradigmatic instances with $N = 6$ physical qubits, based on a criterion to quantify the *hardness* of Parity instances.

The criterion consists of evaluating the energy of the ground state determined only by local interactions with respect to the constrained Hamiltonian. That is to say, if $|\Theta\rangle$ is the ground state of $\mathcal{H}_{\text{local}} = \sum_k J_k \hat{\sigma}_z^{(k)}$, then the hardness of the instance is evaluated as

$$\theta = \frac{\langle \Theta | \mathcal{H}_p | \Theta \rangle - \epsilon_{\min}}{\epsilon_{\max} - \epsilon_{\min}} \quad (3.8)$$

where ϵ_{\max} and ϵ_{\min} are the maximum and minimum eigenvalues of \mathcal{H}_p . An instance with $\theta \sim 0$ is said to be ‘easy’, since the energy of the ground state determined solely by local fields matches the energy of the problem Hamiltonian ground state. Higher values of θ correspond to increasingly non-trivial models.

In this work, we fix the constraints of the LHZ to $C_l = -2$ and the magnitude of the local fields to $J_k \in \{-1, +1\}$. The combination of J_k signs is fixing a specific

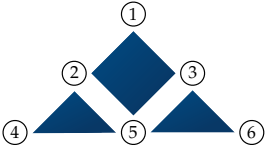
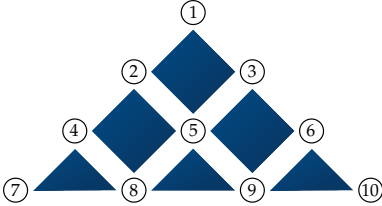
spin indexing	N	label	sign of J_k	θ
		LHZ6E	[+, +, -, -, +, -]	0.0
	6	LHZ6M	[-, -, -, +, -, +]	0.3
		LHZ6H	[+, +, +, -, -, +]	0.55
	10	LHZ10	[-, -, -, -, + +, +, -, -, -]	0.3

Table 3.2: Selected instances of LHZ models that are considered in the following analysis. The constraints are fixed to $C_i = -2$ and the magnitude of the local fields to $|J_k| = 1$. The signs of J_k that fully determine the instance are reported in the table, along with the label of the instance and its “hardness” θ . The reference study case is LHZ6M.

LHZ instance. Enumerating all the possible instances for $N = 6$, we observe that the values of θ range from 0.0 up to 0.6. Therefore, we settle on three instances of LHZ at $N = 6$: an “easy” instance with $\theta = 0$ (labelled LHZ6E), an “hard” instance with $\theta = 0.55$ (LHZ6H) and a “medium” instance with $\theta = 0.3$ (LHZ6M). The latter is considered to be the default study case, and the other two are checked for validation of the results. For validation purposes, we have selected only one instance with 10 physical qubits - LHZ10 - of intermediate hardness ($\theta = 0.3$). The detailed configurations of all four instances are reported in Table 3.2.

Finally, the default annealing time of the simulations is set to $\tau = 0.1$ for all the protocols. The scaling with respect to the annealing time is evaluated, if necessary, within the range of $\tau \in [0.1, 10]$. We do not study cases of smaller timescales, as the quenching limit is beyond the scope of this work.

3.1.c Single annealing sweep

Having described the general setup of the simulations, let us start the discussion by looking in detail at a simple set of simulations with optimally chosen parameters. The following simulations are executed with annealing time $\tau = 0.1$ and on the instance LHZ6M.

It is useful to keep track of two quantities during the evolution: the fidelity and energy of the current state. In other words, if $|\psi(t)\rangle$ is the state evolved at time t

and $|GS(t)\rangle$ is the instantaneous ground state of the annealing protocol Hamiltonian $\mathcal{H}(t)$, then the instantaneous fidelity is simply $f(t) = |\langle\psi(t)|GS(t)\rangle|^2$. Likewise, the instantaneous energy, normalized in the interval $[0, 1]$, is

$$\varepsilon(t) \equiv \frac{\langle\psi|\mathcal{H}|\psi\rangle(t) - \epsilon_{\min}(t)}{\epsilon_{\max}(t) - \epsilon_{\min}(t)} .$$

The normalization is operated with respect to the interval of minimum ($\epsilon_{\min}(t)$) and maximum ($\epsilon_{\max}(t)$) instantaneous eigenvalues of $\mathcal{H}(t)$. The reason for the normalization is easy to understand: it would be difficult to compare different protocols without it because each protocol would have a different spectral range. Furthermore, the use of a normalized quantity allows the definition of another universal metric, indicative of the ‘merit’ of the prepared state with respect to its energy. If the evolving state is closer to the ground state $\varepsilon \rightarrow 0$, so the quantity $1 - \varepsilon$ is maximized. Therefore, we define the *instantaneous energy merit* as

$$m(t) \equiv 1 - \varepsilon(t) .$$

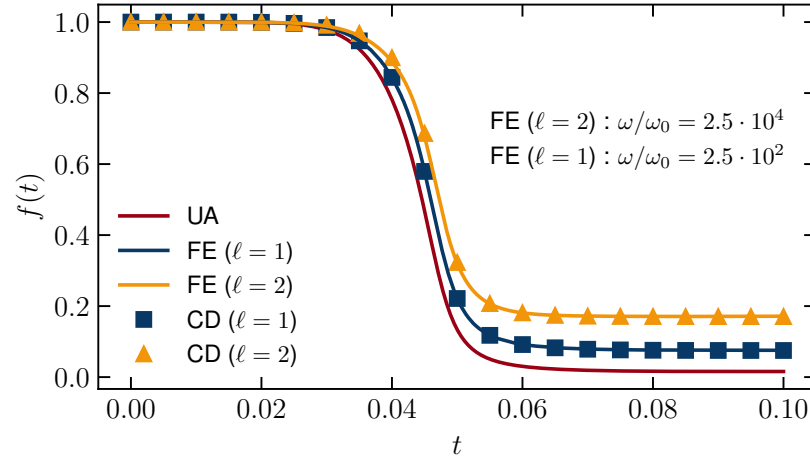
For each of the protocols under consideration, Figures 3.2a and 3.2b plot $f(t)$ and $m(t)$ in a paradigmatic annealing sequence. The parameters of FE protocols have been set in accordance with the suggestions of Ref. [15], with $\omega/\omega_0 = 2.5 \cdot 10^2$ for $\ell = 1$ and $\omega/\omega_0 = 2.5 \cdot 10^4$ for $\ell = 2$. The reference frequency is set to $\omega_0 = 4 \cdot 2\pi$.

$f(t)$ and $m(t)$ begin with a unitary value, indicating a complete overlap with the initial ground state. As the dynamic unfolds, $f(t)$ and $m(t)$ settle to lower values, with higher-order protocols producing better results. The FE protocols match the dynamics of the reference CD schedules; we underline once again that the dynamics of FE protocols are obtained without introducing numerous additional control terms, but rather by modifying the coefficients of \mathcal{H}_i and \mathcal{H}_p .

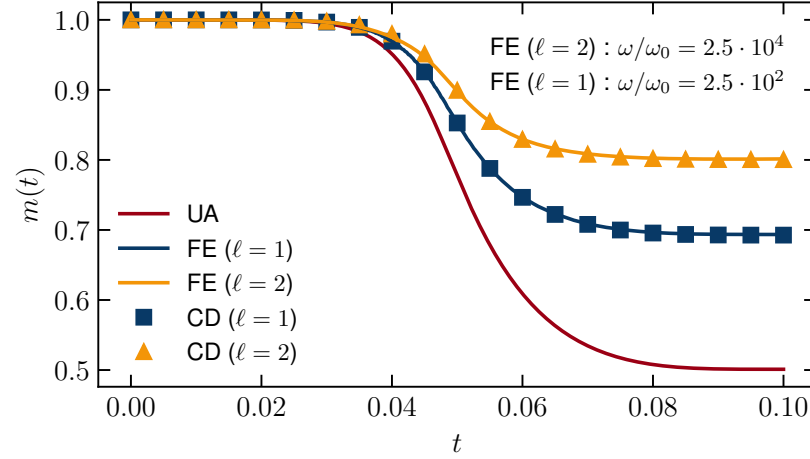
A crucial clarification is due. The FE curves plotted in Figure 3.2 have been sampled at stroboscopic time intervals $t = n \cdot T = n \cdot 2\pi/\omega, n \in \mathbb{N}$, i.e. at the end of each driving oscillation. The reason is that FE protocols have been derived by matching the time evolution operator over a driving cycle to a CD protocol. In general, the values for which $t \neq nT$ exhibit different values w.r.t. the CD protocol. Indeed, if one could zoom within one driving period, the Floquet schedules would show high-frequency oscillations due to the periodic driving. An example of this behavior is shown in Figure 3.3.

	UA	CD		FE	
		$\ell = 1$	$\ell = 2$	$\ell = 1$	$\ell = 2$
$f(\tau) [10^{-2}]$	1.5793	7.4970	17.039	7.5554	17.162
$m(\tau)$	0.50107	0.69264	0.80038	0.69354	0.80143

Table 3.3: Fidelity and energy merit resulting from an annealing simulation of the LHZ6M instance (see Table 3.2) at $t = \tau = 0.1$.



(a) Instantaneous fidelity.



(b) Instantaneous energy merit.

Figure 3.2: (a) Fidelity and (b) energy merit during an annealing simulation on the LHZ6M instance (see Table 3.2). The values of FE protocol have been sampled at stroboscopic times $t = n \cdot 2\pi/\omega, n \in \mathbb{N}$. The choice of the driving frequency ω is done according to the values used in Ref. [15], but ω_0 is changed to $4 \cdot 2\pi$.

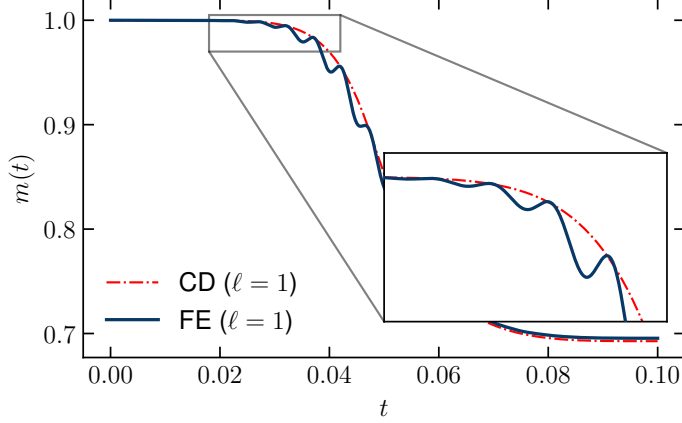


Figure 3.3: Floquet oscillations within a driving cycle for a FE schedule at order $\ell = 1$. LHZ6M instance (see Table 3.2). $\omega_0 = 2 \cdot 2\pi$. To make the oscillations more visible, the driving frequency has been lowered to $\omega = 100 \cdot 2\pi$.

3.1.d Yield at longer annealing times

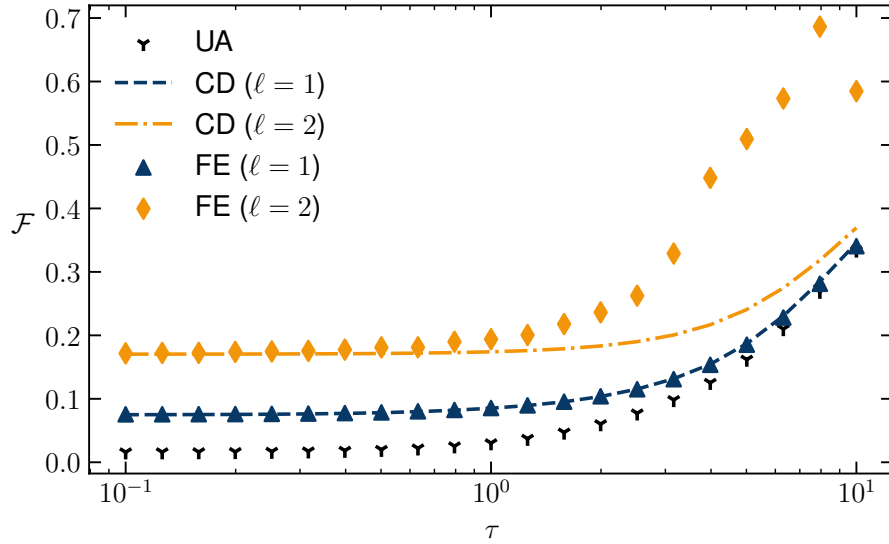
In this section, we study the performance of the various protocols at larger annealing timescales. To do so, we look at the results achieved at the end of the protocols. Therefore, we define two new symbols that indicate the fidelity and the energy merit obtained at the end of a protocol with total evolution time τ .

$$\mathcal{F}(\tau) \equiv f(\tau) = |\langle \psi(\tau) | GS(\tau) \rangle|^2 \quad (3.9)$$

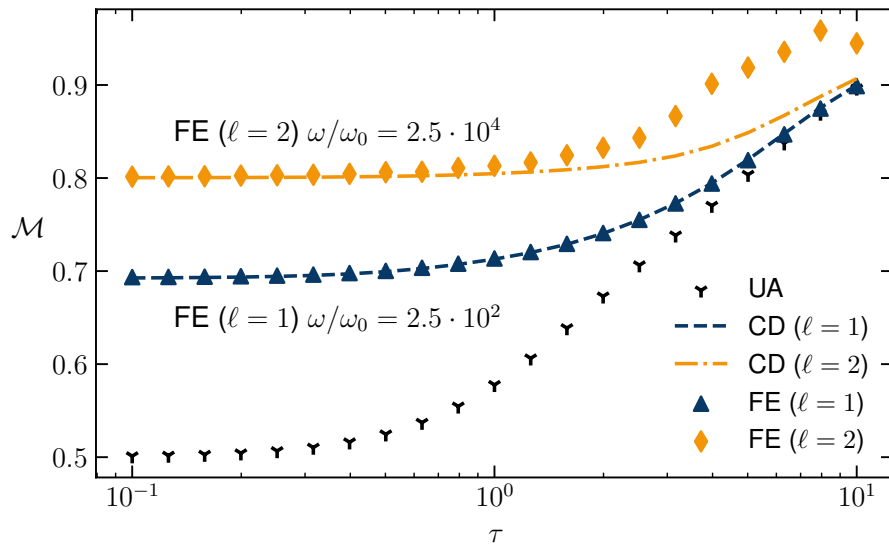
$$\mathcal{M}(\tau) \equiv 1 - \varepsilon(\tau) = m(\tau) \quad (3.10)$$

The behavior of the two functions for different values of τ matches the expectations (see Figure 3.4): longer annealing times improve the fidelity and produce an evolved state closer to the problem ground state. The improvement of CD protocols is considerable for $\tau \in [0.1, 1]$ and becomes negligible at higher timescales. Indeed, as the protocol approaches the $\tau \rightarrow \infty$ limit the action of CD driving becomes marginal, since the AGP is multiplied by $\dot{\lambda}$ and the schedule follows the hypothesis of the Adiabatic Theorem. Eventually, the results of all the various protocols flatten to the UA values.

The agreement between CD and FE dynamics is perfect for the $\ell = 1$ ansatz but fails for high values of τ at second order $\ell = 2$. Fortunately, the discrepancy appears to result in higher energy merit values (and fidelities) in favor of the Floquet protocol. The reason for such disagreement is not clear at first hand and requires further investigation. From a theoretical standpoint, the FE protocols should not improve or worsen the result w.r.t. their reference CD protocols, as



(a) Final fidelity.



(b) Final energy merit.

Figure 3.4: (a) Fidelity and (b) energy merit at the end of a quantum annealing simulation on the LHZ6M instance (see Table 3.2). The annealing time is changed in the range $\tau \in [0.1, 10]$. The choice of the driving frequency ω is done according to the values used in Ref. [15], but we fix $\omega_0 = 4 \cdot 2\pi$.

they should exactly replicate the counterdiabatic dynamics.

As we will see in the following section, this discrepancy occurs also at $\tau = 0.1$ for specific choices of ω and ω_0 . The goal of §3.3 will be to motivate this discrepancy by looking at the sources of errors introduced in the derivation of the FE protocol.

3.2 Tuning the FE protocol frequencies

We now wish to discuss the choice of ω , the stroboscopic driving frequency, and ω_0 , the reference frequency of the spectrum. Following the suggestions from the Ref. [15], we have adopted the same criterion of $\omega/\omega_0 \sim 10^{2\ell}$, retrieving good results in all our tests at $\tau = 0.1$. This paragraph aims to explore different regimes of such parameters.

The values of ω and ω_0 appear in the derivation of the Floquet protocols and are directly involved in several approximations. Therefore, one can try to infer their magnitude from the underlying theory.

- Bigger values of ω realize FE protocols with finer stroboscopic driving. Given a protocol of length τ and frequency ω , the total number of driving oscillations is

$$n_d = \omega\tau/2\pi \quad . \quad (3.11)$$

To have $n_d \gg 1$, one should choose $\omega \gg 2\pi/\tau$. Nonetheless, it may be worthwhile to investigate regimes in which the driving cycles are not dense and see how far the protocol may be pushed into lower-frequency regimes.

- The reference frequency ω_0 is involved in the expansion of the Bessel functions with the Jacobi-Anger formula (see Eq. (2.28)). For the expansion to hold, its argument should be $\rightarrow 0$, i.e. $\omega_{mn} \ll \omega_0$. This consideration suggests that it is necessary to push ω_0 towards higher frequency scales. However, the Bessel function expansion does not need to be very accurate at higher excitation frequencies ω_{mn} . Indeed, during evolution, it is more likely² for the instantaneous state to get excited towards “closer” eigenstates ($\omega_{mn} \simeq 0$) of the spectrum. Therefore, the expansion does not need to be accurate for a large window of ω_{mn} , but only for a limited range $\omega_{mn} \leq \Delta$. This rationale suggests it could be worth examining how far the Floquet protocols can be stretched to lower reference frequencies ω_0 .

We benchmark FE protocols in a wide range of ω and ω_0 , considering also low-frequencies regimes. The results of these benchmarks will be shown in

²Recall that the off-diagonal matrix elements of the AGP - responsible for instantaneous transitions between eigenstates - are proportional to ω_{mn}^{-1} (see Eq. (2.14)).

color-mapped plots. On the vertical axis, we have fixed the number of driving cycles in the schedule (i.e. $\omega\tau/2\pi$). The horizontal axis denotes the value of the reference spectrum frequency in units of 2π .

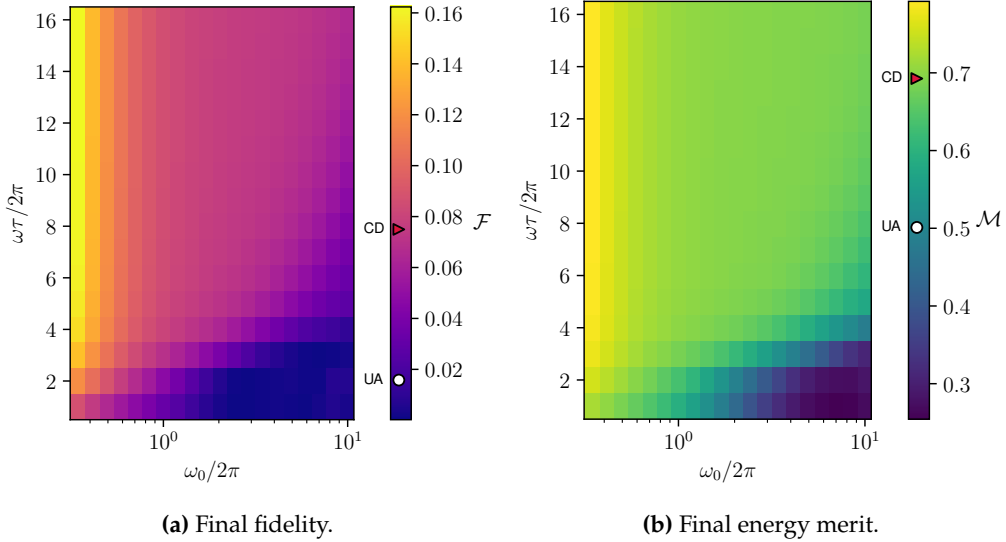


Figure 3.5: FE $\ell = 1$ simulations on the LHZ6M instance (see Table 3.2), in various regimes of driving & reference frequencies. Figure (a) shows the final fidelity \mathcal{F} , whereas Figure (b) the energy merit \mathcal{M} . The total evolution time of the annealing protocol is set to $\tau = 0.1$. The vertical axis fixes the total number of driving cycles ($\omega\tau/2\pi$); the horizontal axis denotes the spectral reference frequency $\omega_0/2\pi$. In addition, the white circle on the color bar indicates the value - either fidelity or energy merit - achieved by the UA protocol. Likewise, the red triangle marks the same quantities for the $\ell = 1$ CD protocol.

The results of the FE simulations, at order $\ell = 1$ on the LHZ6M instance, are shown in Figure 3.5, where we plot the fidelity \mathcal{F} and the energy merit \mathcal{M} at the end of the schedule. To aid the interpretation, we have marked on the color bar the values achieved by UA and CD schedules. Looking at the color scale range of Figure 3.5, it is clear that the FE protocol achieves fidelities (or energy merit) below the UA protocol and, surprisingly, above the reference CD protocol. In other words, depending on the choice of (ω_0, ω) , the FE protocols do not reproduce correctly the CD evolution. This behavior has already been seen in Figure 3.4, where the $\ell = 2$ FE protocol exceeded the fidelity of the CD protocol at larger total evolution times.

Having settled that the FE protocols are supposed to replicate the dynamic of a CD schedule, we wish to determine which combinations of (ω_0, ω) the

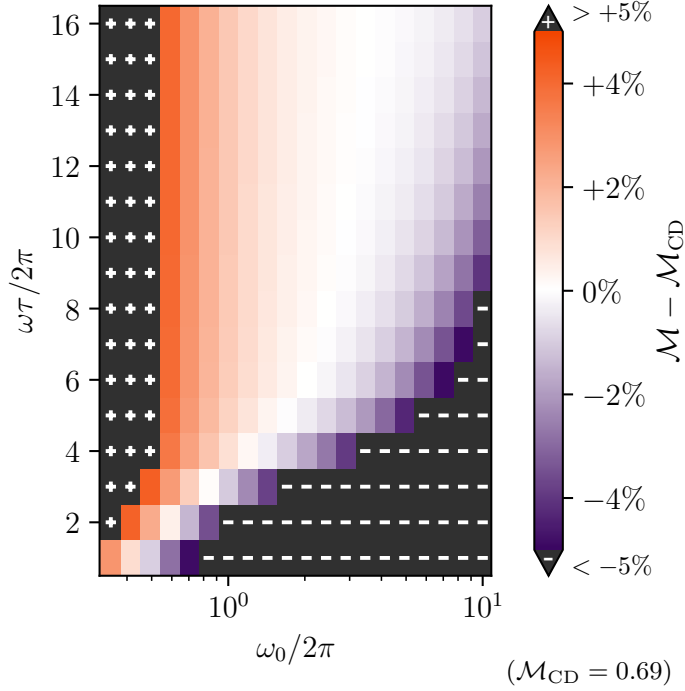


Figure 3.6: Difference in energy merit for the FE protocol w.r.t. the CD protocol at first order ($\ell = 1$). The annealing is executed on the LHZ6M instance (see Table 3.2). The points that exceed the $\pm 5\%$ interval around zero are cropped to a grey color and marked with a $+/-$ symbol to indicate the sign of the extreme. The vertical axis fixes the total number of driving cycles of the protocol ($\omega\tau/2\pi$); the horizontal axis denotes the spectral reference frequency $\omega_0/2\pi$.

FE protocol are most successful. To facilitate the visualization, we consider $\mathcal{M} - \mathcal{M}_{\text{CD}}$, which is the energy merit of the FE protocol diminished by the counterdiabatic protocol energy merit \mathcal{M}_{CD} . The extremes of the new colormap are bound to a $\pm 5\%$ interval around 0. The area of best “accuracy of the protocol” ($\mathcal{M} - \mathcal{M}_{\text{CD}} \simeq 0$) is therefore identified with a neutral white color. All the values that exceed the $\pm 5\%$ interval have been marked with a gray color and a $+$ or $-$ symbol, depending on whether the value exceeds positively or negatively the color scale.

Looking at Figure 3.6, we observe that below 4 driving cycles, only a small strip of parameters (ω_0, ω) shows optimal accuracy. Above 6 driving cycles the area of agreement is significantly larger, improving even more when using higher driving frequencies. Furthermore, it appears that the minimum value of ω_0 in order to get the best accuracy has to be greater than 2π .

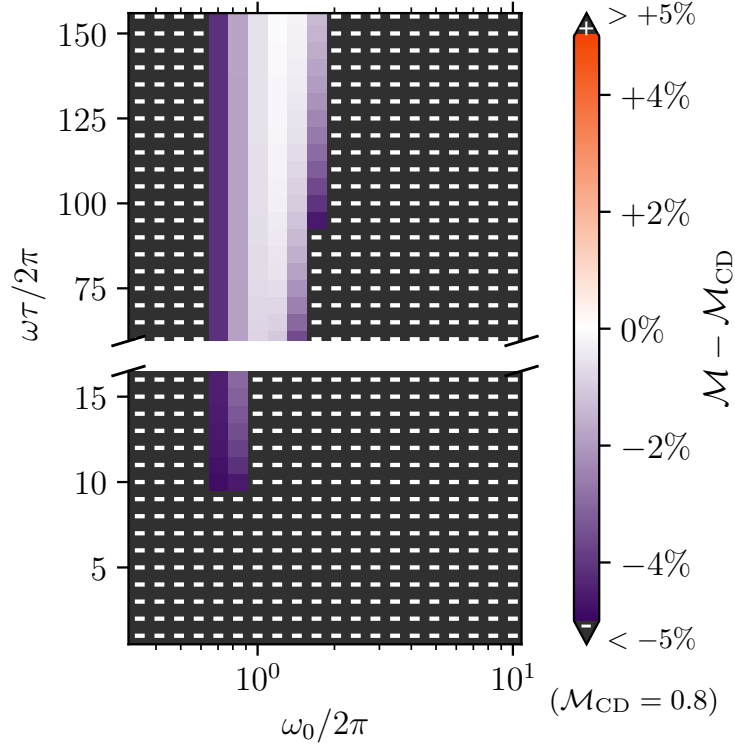


Figure 3.7: Difference in energy merit for the FE protocol w.r.t. the CD protocol at second order ($\ell = 2$). The annealing is executed on the LHZ6M instance (see Table 3.2). The points that exceed the $\pm 5\%$ interval around zero are cropped to a grey color and marked with a \pm symbol to indicate the sign of the extreme. The vertical axis fixes the total number of driving cycles of the protocol ($\omega\tau/2\pi$); the horizontal axis denotes the spectral reference frequency $\omega_0/2\pi$.

Figure 3.7 shows the energy merit difference of the $\ell = 2$ FE protocol. The window of driving oscillations between 1 and 16 does not reproduce correctly the target CD energy merit. We have benchmarked a second window of driving cycles between 60 and 160. The accuracy of the energy merit ($\mathcal{M} - \mathcal{M}_{\text{CD}} \simeq 0$) improves significantly around $\omega_0 \sim 2\pi$, even though the feasible interval of ω_0 is tighter w.r.t. the first order FE protocol. Moreover, the difference is always negative, in contrast to the first-order protocol (i.e. no energy merit of FE protocols is above the CD value).

3.2.a Validation on other LHZ instances

The results plotted in Figures 3.6 and 3.7 have considered only one of the three LHZ instances with $N = 6$ physical qubits. The same type of simulations has

been conducted on the LHZ6E and LHZ6H instances (see Table 3.2), showing no significant differences with respect to the results discussed above. At larger system sizes, using the instance LHZ10, the energy merit accuracy seems unaltered. To avoid unnecessary redundancy, we will refer to Appendix C for the plots on LHZ6E, LHZ6H, and LHZ10.

3.2.b Error sources

As we have previously observed, it is not expected that the evolution under FE protocols differs significantly from the reference CD protocols, as they are explicitly constructed to reproduce their dynamics. Thus, we would like to find a suitable criterion to motivate the difference in terms of energy merit. The most plausible explanation is to be found among the approximations that have been employed to derive the FE protocols.

Our analysis has found a plausible source of errors in the Magnus expansion of Eq. (2.26). The error associated with the Magnus expansion truncated at first-order is upper bounded [8] by the integral of the driving Hamiltonian norm in each driving cycle of period $T = 2\pi/\omega$:

$$\delta = 2 \int_0^T \|\mathcal{H}_{\text{FE}}\|_2 dt \quad . \quad (3.12)$$

The average error bound per driving cycle has been plotted in Figure 3.8a. A comparison with the nearby Figure 3.8b shows that the areas of best accuracy ($\mathcal{M} - \mathcal{M}_{\text{CD}} \simeq 0$) correspond to the ones in which δ is smaller.

If the error of the Magnus expansion is responsible for the difference of the FE energy merit, then minimizing the error bound should be a suitable criterion to select the best parameters ω and ω_0 . Furthermore, δ can be minimized without solving the dynamics, which is instead mandatory if one wishes to maximize either the fidelity or the energy merit of the protocol.

For instance, suppose to fix the number of driving oscillations to 10; what is the value of ω_0 that allows for the best accuracy of the protocol? To answer this question we have computed the norm of the FE driving protocol for many values of τ , which allows us to see whether the optimal value of ω_0 would change for longer annealing times; the bound δ has been reported in Figure 3.9. One observes that the error bound is constant in a wide range of evolution times. In particular, $\omega_0 = 4 \cdot 2\pi$ minimizes δ , and by extension the error of the protocol.

In conclusion, the criteria of minimizing the error bound δ is effective in determining the combination of parameters (ω_0, ω) of the FE protocol that returns the best accuracy, i.e. $\mathcal{M} - \mathcal{M}_{\text{CD}} \simeq 0$. This region of parameters corresponds approximately to $\omega_0/2\pi \simeq 4$ if the driving frequency is $\omega/2\pi > 5/\tau$. In the

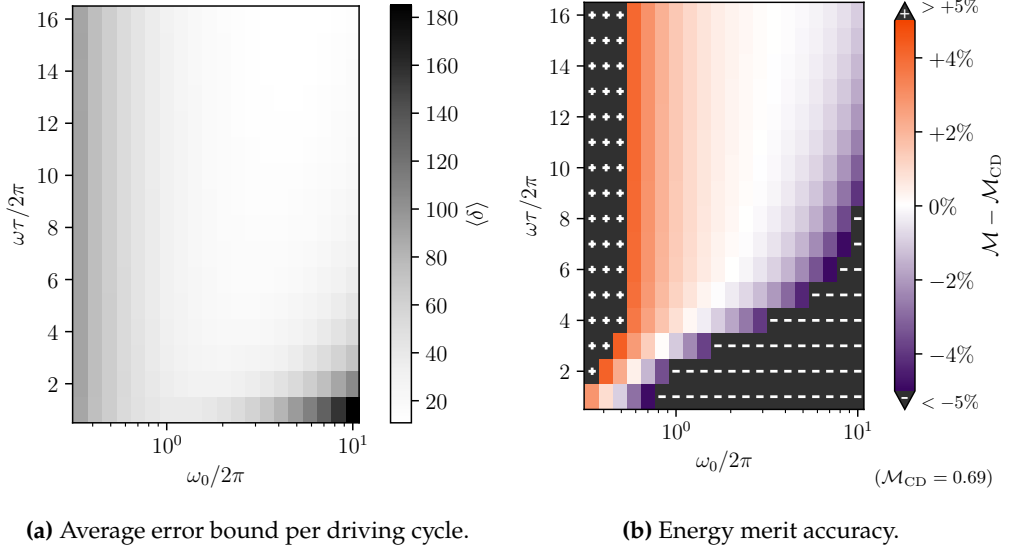


Figure 3.8: Comparison between error per driving cycle and the energy merit difference for a FE protocol at $\ell = 1$, executed on the LHZ6M instance (see Table 3.2). The best accordance of \mathcal{M} with \mathcal{M}_{CD} occurs where the average error bound per driving cycle $\langle \delta \rangle = 2\langle \int_0^T \|\mathcal{H}\| dt \rangle$ is small. The vertical axis fixes the total number of driving cycles of the protocol ($\omega\tau/2\pi$); the horizontal axis denotes the spectral reference frequency $\omega_0/2\pi$.

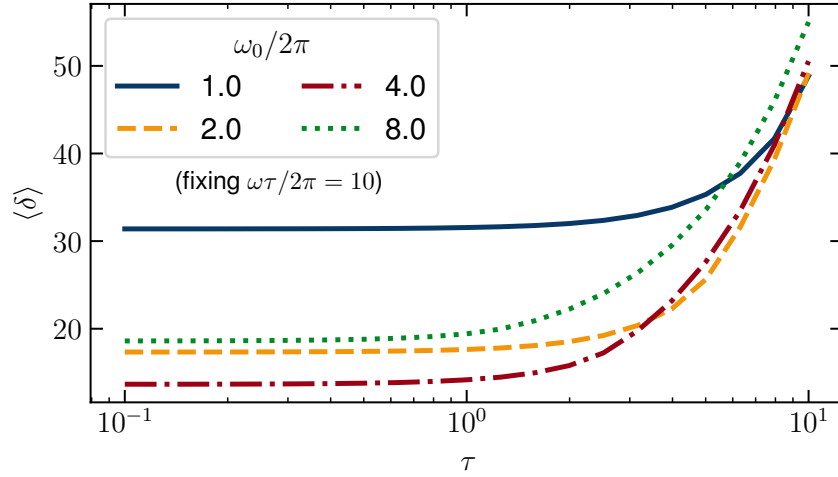


Figure 3.9: Error per driving cycle on the LHZ6M instance (see Table 3.2), obtained by fixing the number of stroboscopic oscillations to 10 and changing the total annealing time τ . As remarked in the main text, $\omega_0 = 4 \cdot 2\pi$ minimizes the error in the wide range of τ s from 0.1 to 3.0.

low driving frequency regime $\omega/2\pi < 5\tau$, the optimal reference frequency ω_0 is gradually reduced to $\omega_0/2\pi \simeq 0.4$.

3.3 Cost analysis

So far, we have compared the FE protocols to the reference CD dynamics, with the intent of finding a criterion to determine the optimal set of parameters (ω_0, ω) that assures the accuracy of the protocol. We must now raise some questions about the *efficiency* of the Floquet protocols themselves when compared to the unassisted annealing protocols (UA). In other words, is it convenient to execute Floquet protocols?

By looking at the outcome of individual simulations, like in Figure 3.2, it seems that FE protocols are advantageous, in the sense that they successfully reproduce CD evolutions at arbitrary order ℓ without requiring additional interaction terms. However, we must take into account the magnitude of the coefficients that are used in Floquet schedules. Indeed, as hinted in §2.1, a proper ‘cost’ metric of the protocol does not depend only on the total evolution time, but also on the norm of the driving Hamiltonian [2]. We observe that the driving coefficients A_{FE} and B_{FE} - in Eq. (3.4) - scale asymptotically as $\sim \omega/\omega_0 + \omega_0^{2\ell-1}$, so the magnitude of the coefficients can grow quickly in the optimal region of frequency parameters. In turn, the norm of the driving Hamiltonian depends on its coefficients, leading to a larger cost of the FE protocol.

The role of the coefficient magnitude is not as trivial as it seems. One could think that quantum hardware can realize arbitrarily large driving terms, but this is incorrect. Say that the hardware can handle coefficients up to a limit magnitude Λ . Then, the protocol Hamiltonian must be proportionally rescaled such that $\max_t (A(t), B(t)) \leq \Lambda$. However, rescaling the driving coefficients is an operation that has a physical relevance.

Let us consider the Schrödinger equation to solve the dynamics of a quantum system described by a generic Hamiltonian H :

$$i\partial_t |\phi\rangle = H |\phi\rangle$$

Suppose that Hamiltonian is rescaled by a coefficient $\kappa > 1$: $H \rightarrow H' = \kappa H$. The Schrödinger equation on the new Hamiltonian is identical to the former, but its RHS is multiplied by the Hamiltonian scaling coefficient.

$$i\partial_t |\phi\rangle = H' |\phi\rangle = \kappa \cdot H |\phi\rangle$$

It is trivial to get rid of the additional constant by redefining the time variable $t \rightarrow t'$ such that $\partial_{t'} = \kappa^{-1} \partial_t$,

$$i \frac{1}{\kappa} \partial_t |\phi\rangle = i \partial_{t'} |\phi\rangle = H |\phi\rangle \quad ,$$

which is satisfied by any linear transformation $t = \kappa t' + \text{const.}$ The Schrödinger equation for the new time variable yields the same solutions as the Schrödinger equation for the original variable, but t is substituted by t' , of course. To compare the dynamics between the two solutions it is necessary to revert the transformation of the time variable, i.e. returning to the original time coordinate. We now understand that the solution coming from the rescaled Hamiltonian is effectively representing a dynamics having a dilated time variable $t = \kappa t' > t'$. This demonstrates that a linear rescaling of the Hamiltonian is equivalent to a linear rescaling of the time variable.

This argument can be applied to a Quantum Annealing schedule, where we operate a protocol $\mathcal{H}(t)$ in a total annealing time τ . If the protocol is rescaled by a coefficient $\kappa > 1$, then the “effective time” of the dynamics would be $\tau' = \kappa \tau > \tau$. As a consequence, it would be possible to tamper with the dynamics of the system and obtain arbitrarily large virtual annealing times τ' in a fixed hardware time τ . Unfortunately, the hardware is not able to handle coefficient rescaling to an arbitrary extent. For this very same reason, the total evolution time τ alone is not a reliable measure of the ‘cost’ of the ASP algorithm: it does not take into account the ‘magnitude’ of the Hamiltonian terms and it would be possible to improve the prepared state by rescaling the Hamiltonian coefficients.

It is crucial to establish a concept of cost that takes into consideration the freedom to scale the Hamiltonian. Generalizing the Eq. (2.3) from Ref. [2], a valid estimation of the cost would be to integrate the norm of the Hamiltonian in the evolution time interval:

$$\mathcal{C}(\mathcal{H}) \equiv \int_0^\tau \|\mathcal{H}\| dt \quad . \quad (3.13)$$

The new cost function fixes the problem of arbitrary rescaling, taking into account the magnitude of the driving terms. Indeed, a rescaling of the Hamiltonian $\mathcal{H}' = \kappa \mathcal{H}$ implies that $\mathcal{C}(\mathcal{H}') = \kappa \mathcal{C}(\mathcal{H})$.

We point out, however, that in real applications, the norm of the driving Hamiltonian of Eq. (3.13) could not be an optimal metric, since a realistic cost estimation must depend on the experimental annealing setup. But if two protocols are written as linear interpolation, like the UA and FE protocols in Eqs.(3.2) and (3.3), $\mathcal{C}(\mathcal{H})$ is useful to compare their cost.

3.3.a The effective time compromise

The cost metric for a Hamiltonian in the form $\mathcal{H}(t) = A(t)\mathcal{H}_i + B(t)\mathcal{H}_p$ - with \mathcal{H}_i and \mathcal{H}_p being time-independent - can be simplified by considering only the coefficient functions. We introduce the *effective annealing time* as

$$\tau_{\text{eff}} \equiv \int_0^\tau |A(t)| + |B(t)| dt \quad . \quad (3.14)$$

The effective time is related to the cost \mathcal{C} . From the properties of the norm, it is

$$\|A(t)\mathcal{H}_i + B(t)\mathcal{H}_p\| \leq |A(t)|\|\mathcal{H}_i\| + |B(t)|\|\mathcal{H}_p\|$$

and if one assumes that $H_{\text{max}} = \max_t (\|\mathcal{H}_i\|, \|\mathcal{H}_p\|)$, then

$$\int_0^\tau \|\mathcal{H}\| dt \leq H_{\text{max}} \int_0^\tau |A(t)| + |B(t)| dt = H_{\text{max}} \tau_{\text{eff}} \quad . \quad (3.15)$$

We also observe that if the schedule is a simple interpolation like the UA protocol - that is $A(t) = 1 - \lambda(t)$ and $B(t) = \lambda(t)$, with $\lambda : [0, \tau] \rightarrow [0, 1]$ - then the effective annealing time is exactly the (real) annealing time.

$$\tau_{\text{eff}}^{\text{UA}} = \int_0^\tau dt (1 - \lambda(t) + \lambda(t)) = \tau$$

The same property does not hold, of course, for FE schedules. So, what is a typical value of τ_{eff} for Floquet protocols? To answer, we have plotted in Figure 3.10a the effective time for many values of ω_0 , fixing the total number of stroboscopic oscillations to 10. The adjacent Figure 3.10b shows the difference in energy merit between the FE protocol and its reference CD value. We observe that the best accuracy of the FE protocol (i.e. $\mathcal{M} - \mathcal{M}_{\text{CD}} \simeq 0$) occurs for those values of ω_0 where the effective time is globally smaller.

In the previous section, we have seen that the accuracy of the FE protocols is optimal when the average error bound in each driving cycle is minimized (see Figure 3.8). It turns out that the error bound is related to τ_{eff} . The average error bound per driving cycle is proportional to the integral of the norm in the annealing time interval, which is upper-bounded by the effective time:

$$\frac{\tau}{T} \underbrace{2 \left\langle \int_0^T \|\mathcal{H}\| dt \right\rangle}_{\text{avg error bound per cycle}} = 2 \int_0^\tau \|\mathcal{H}\| dt \stackrel{\text{Eq. (3.15)}}{\leq} 2H_{\text{max}} \tau_{\text{eff}} \quad .$$

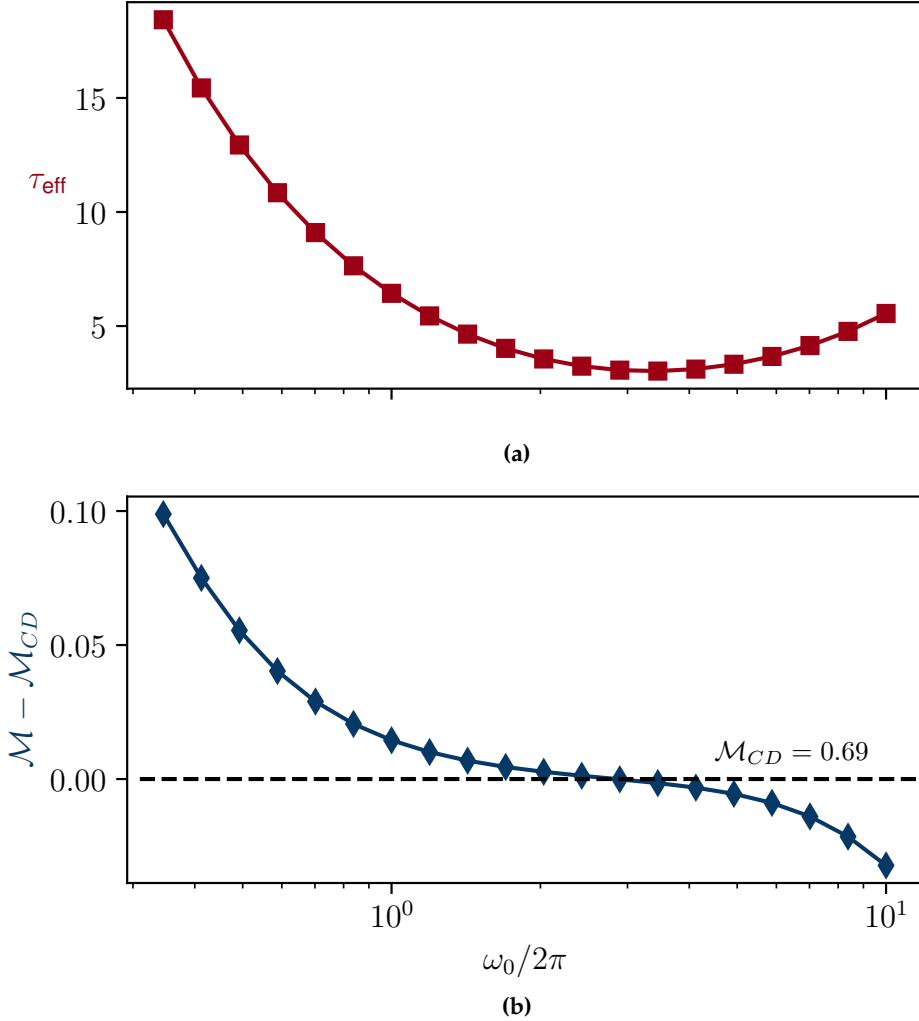


Figure 3.10: (a) Effective time and (b) energy merit accuracy for FE protocol at first order ($\ell = 1$). The data is simulated on the LHZ6M instance (see Table 3.2), fixing the number of stroboscopic oscillations to 10. The values of the effective time are $\tau_{\text{eff}} \geq 3.5$, much bigger than the (real) total annealing time $\tau = 0.1$. Furthermore, the protocol accuracy (i.e. $\mathcal{M} - \mathcal{M}_{CD} \simeq 0$) is optimal when the effective annealing time is minimized.

3.3.b Advantage over unassisted protocols

The crucial question is now whether the Floquet protocols bring some advantage with respect to UA protocols. Actually, the typical effective times shown in Figure 3.10a are much bigger than the real annealing time ($\tau = 0.1$). Thus, it is quite intuitive to check what energy merit would an UA protocol achieve at an equivalent effective time. If the energy merit of the FE schedule was greater than the UA's, it would imply that executing a FE protocol is advantageous.

To formalize this idea, let us define the **gain** as the difference in energy merit between FE and UA protocols at equivalent effective time.

$$\mathcal{G} \equiv \mathcal{M}_{\text{FE}}(\tau_{\text{eff}}) - \mathcal{M}_{\text{UA}}(\tau_{\text{eff}}) \quad (3.16)$$

Intuitively, the gain is visualized in Figure 3.11 as the vertical distance from the energy merit of the FE protocol and the energy merit of UA. The latter can be obtained by benchmarking \mathcal{M}_{UA} in a wide interval of τ . A simple geometrical consideration implies that all the points $(\tau_{\text{eff}}, \mathcal{M}_{\text{FE}})$ above the UA line will have a positive gain.

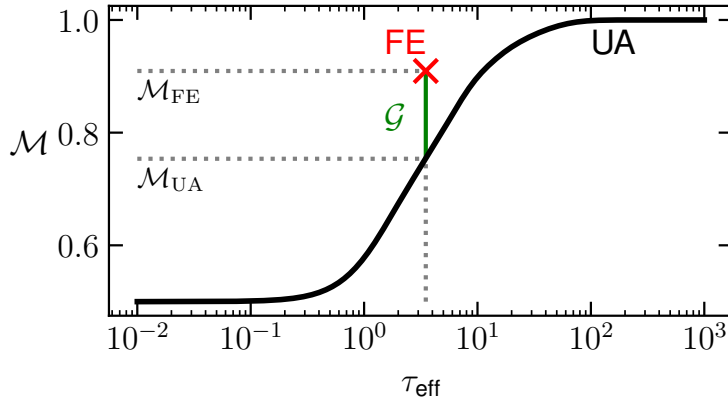


Figure 3.11: Energy merit vs τ_{eff} . The solid black line represents the energy merit of UA schedules on the LHZ6M instance (see Table 3.2). The red marker is a dummy indicator used to represent an ideal FE simulation. The vertical distance between the FE red marker and the UA line, highlighted in green, is the gain.

We have calculated the gain using the same approach of the simulations shown in the previous analysis, i.e. changing the FE protocol driving frequency ω and the spectral reference frequency ω_0 in broad intervals. The resulting gain values are plotted in Figure 3.12, which testify to the dominance of negative values of \mathcal{G} for both counterdiabatic protocol orders. At $\ell = 1$, there exists a wide region of (ω_0, ω) that achieves a null gain, with some values being slightly positive. The area of gain positivity ($\mathcal{G} \geq 0$) has been surrounded by a dashed line. Just a few simulations with small driving frequencies achieve positive gains, and all the other values in the parameter space are negative. The highest value of the gain is 7.15%, which is obtained by choosing a driving frequency $\omega/2\pi = 1/\tau$.

In the $\ell = 2$ FE protocol, the gain values are generally worse, with only one simulation at $\omega\tau/2\pi = 1$ achieving gain neutrality (see Figure 3.12b).

An alternative visualization of the gain is provided in Figure 3.13, where the

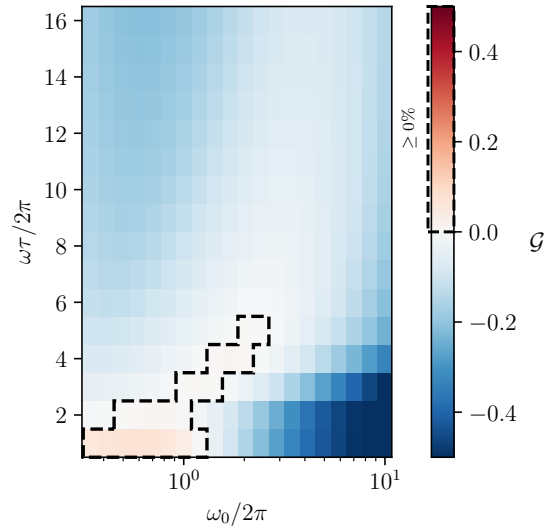
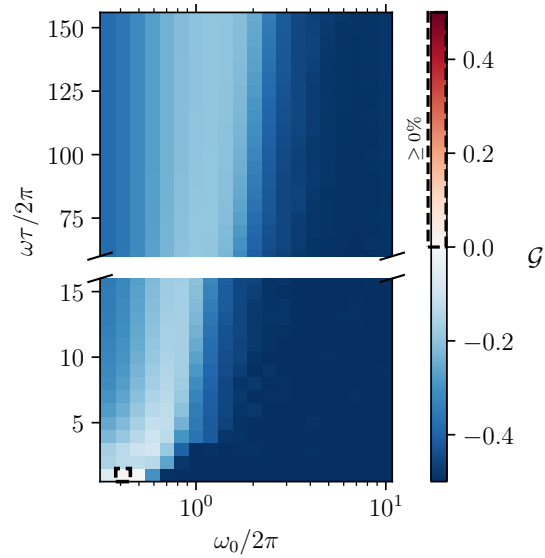
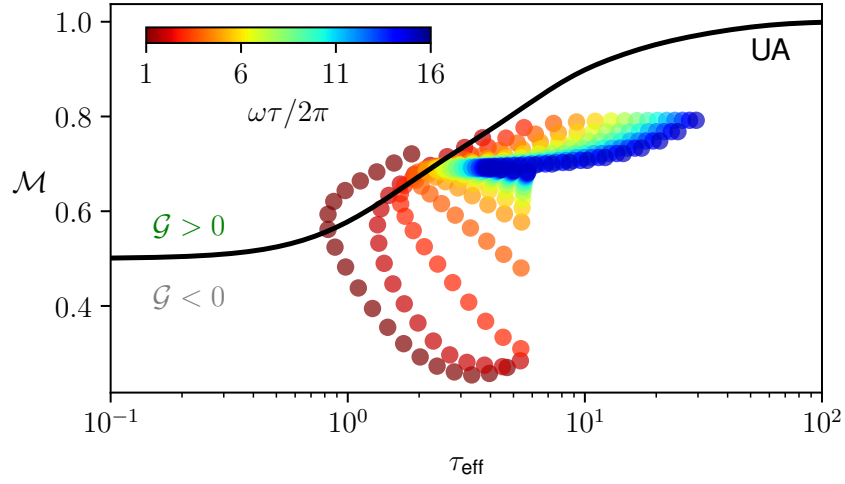
(a) $\ell = 1$ FE protocol.(b) $\ell = 2$ FE protocol.

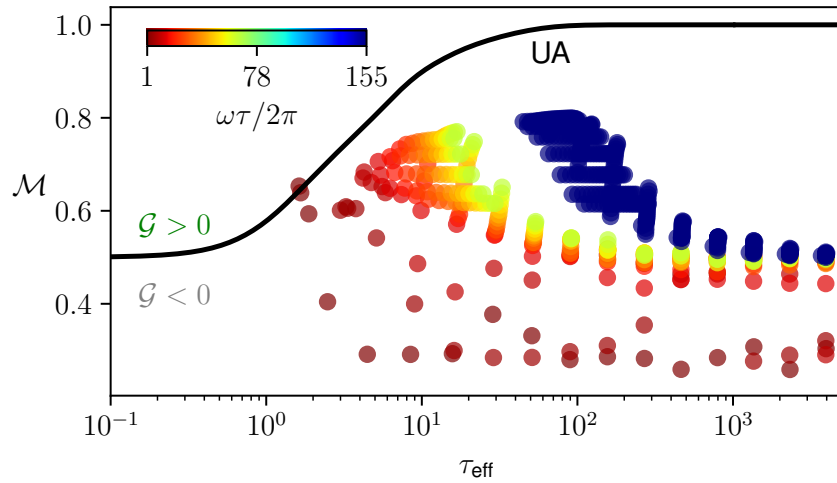
Figure 3.12: Gain for Floquet simulations of the LHZ6M instance (see Table 3.2). The vertical axis fixes the total number of driving cycles of the protocol ($\omega\tau/2\pi$); the horizontal axis denotes the spectral reference frequency $\omega_0/2\pi$. The region of $\mathcal{G} \geq 0$ is surrounded by a dashed line. The subfigures show results at different CD approximation order: (a) for the $\ell = 1$ protocol and (b) for the $\ell = 2$ protocol.

energy merit is plotted against the effective annealing time. The points fall mostly in the region below the UA energy merit curve, implying overall negative values of the gain. The trend is particularly accentuated for the $\ell = 2$ protocol.

Based on the results of the gain analysis, we conclude that FE protocols are generally unadvantageous w.r.t. UA protocols. The few positive values of \mathcal{G} are obtained in the regime of low driving frequencies $\omega\tau/2\pi \simeq 1$, which is tricky for two reasons. (1) The FE protocol realizes the CD evolution in only one driving cycle, which is not optimal for the hypothesis under which the protocols have been derived, in particular the one that $\mathcal{H}(t)$ is constant in a driving cycle (see Eq. (2.25)). (2) In the regime $\omega\tau/2\pi \simeq 1$, the protocol accuracy, $\mathcal{M} - \mathcal{M}_{\text{CD}} \simeq 0$, is guaranteed only in a very narrow window around $\omega_0 \simeq 0.4 \cdot 2\pi$.



(a) First order ($\ell = 1$) FE protocol. Only some points are located in the region of positive gain $\mathcal{G} > 0$. However, they correspond to a regime of small driving frequency that is not optimal to reach the target CD energy merit.



(b) Second order ($\ell = 2$) FE protocol. The points corresponding to high-frequency driving show very large values of τ_{eff} , which is due to the large coefficients in the driving protocol.

Figure 3.13: Alternative visualizations of the gain, for the LHZ6M instance (see Table 3.2) at $\tau = 0.1$. The subfigures show results at different CD approximation order: (a) for the $\ell = 1$ protocol and (b) for the $\ell = 2$ protocol. Each point represents the outcome of a FE protocol simulation for a different combination of ω and $\omega_0 \in [0.3, 10] \cdot 2\pi$. The points have been color-coded by the number of driving cycles $\omega\tau/2\pi$ (see the color map). The vertical axis shows the final energy merit \mathcal{M} ; the horizontal the effective annealing time τ_{eff} of the protocol. Overall, the simulation results are located below the UA line, indicating negative values of the gain.

Can we improve Floquet protocols?

In the previous Chapter, we evaluated the efficiency of Floquet protocols for quantum annealing. We found a narrow region of parameters (ω_0, ω) with limited advantage in the regime of small driving frequencies. In this work, we quantify the efficiency of Floquet protocols by comparing the energy merit of the evolved state between the FE and the UA schedules at equal ‘cost’ of the protocol. However, we were not able to go beyond 7% of efficiency. Furthermore, the regime that presents such an advantage is not optimal due to the hypothesis underlying the protocol derivation. In regimes of high driving frequency, Floquet protocols typically achieve a lower energy merit w.r.t. UA protocols at equal cost.

Before formulating any further conclusion, we would like to address three additional questions. Can we improve Floquet protocols? Is this analysis valid also for the annealing of models that are not the Parity architecture? Are there any other ways to exploit counterdiabaticity that are more advantageous than unassisted annealing schedules?

4.1 Alternative Floquet protocols

Floquet protocols implement the counterdiabatic evolution through a stroboscopic driving of the system. The coefficients of the oscillating terms have been determined by matching the time evolution operator in each driving cycle to the CD Hamiltonian. However, the Floquet protocol proposed by Claeys et al. in Ref. [15] - discussed in this text with the Hamiltonian in Eq. (2.24) - is just an instance of the possible Floquet protocols that realizes counterdiabatic driving. In other words, it is possible to derive new FE protocols through different approaches, or even just by changing the approximations made during the original

derivation.

Throughout this section, we will introduce several new stroboscopic protocols that can implement the desired counterdiabatic evolution. To assess their effectiveness, we will rely on the analysis framework established in the previous Chapter. The *accuracy* of the protocol is quantified by looking at $\mathcal{M} - \mathcal{M}_{\text{CD}}$ (see §3.2), where we compare the energy merit (Eq. (3.10)) of FE schedules with respect to the one obtained with the target CD protocol. We quantify the *efficiency* of the protocol through the gain, defined in §3.3.b as $\mathcal{G} = \mathcal{M}(\tau_{\text{eff}}) - \mathcal{M}_{\text{UA}}(\tau_{\text{eff}})$.

4.1.a A family of protocols through variational minimization

To get started, we show that the protocols of Ref. [15] can be derived from slightly more general conditions. Let us start with a periodic driving Hamiltonian in a suitable rotating frame. Consider a generic Hamiltonian $\mathcal{H}(\lambda)$ and move to a rotating frame with the transformation $R(t) = e^{-i \int^t f(t') \mathcal{H} dt'}$:

$$\begin{aligned} \mathcal{H} &\longrightarrow \mathcal{H}_{\text{rot}} = e^{-i \int^t f \mathcal{H}} \mathcal{H} e^{i \int^t f \mathcal{H}} - i e^{-i \int^t f \mathcal{H}} \left(i \dot{f} \mathcal{H} + i f \partial_\lambda \mathcal{H} \dot{\lambda} \right) e^{i \int^t f \mathcal{H}} = \\ &= \left(1 + \dot{f} \right) \mathcal{H} + f \dot{\lambda} \partial_\lambda \mathcal{H} \quad . \end{aligned}$$

If the rotating Hamiltonian is evaluated in its eigenbasis, the only term that can lead to off-diagonal elements is $\partial_\lambda \mathcal{H}$. Therefore, we can write a ‘tunable’ Floquet Hamiltonian by replacing the coefficient of $\dot{\lambda} \partial_\lambda \mathcal{H}$ with a generic function $\beta(t)$,

$$\mathcal{H}_{\text{FE}} = \left(1 + \dot{f} \right) \mathcal{H} + \beta(t) \dot{\lambda} \partial_\lambda \mathcal{H} \quad . \quad (4.1)$$

To realize the stroboscopic driving, $f(t)$ should be an oscillating function of period $T = \omega/2\pi$. We observe that the case of $f = \sin(\omega t)/\omega_0$ has been used in the Floquet Hamiltonian (see Eq. (2.24)) of Ref. [15].

In the following passages, we will determine the conditions on $\beta(t)$ that allow us to replicate the counterdiabatic dynamics. The effective dynamic of the general Floquet Hamiltonian in a driving cycle is given by the time evolution operator, which can be expanded at first order with the Magnus expansion:

$$\begin{aligned} \mathcal{H}_{\text{eff}} &= \frac{1}{T} \int_0^T dt e^{-i \int^t f \mathcal{H}} \mathcal{H}_{\text{FE}} e^{i \int^t f \mathcal{H}} = \\ &= \frac{1}{T} \int_0^T dt \left[\left(1 + \dot{f} \right) \mathcal{H} + \beta(t) \dot{\lambda} e^{-i \int^t f \mathcal{H}} \partial_\lambda \mathcal{H} e^{i \int^t f \mathcal{H}} \right] \quad . \end{aligned}$$

In the instantaneous eigenbasis of \mathcal{H} , the matrix elements of \mathcal{H}_{eff} are

$$\langle m | \mathcal{H}_{\text{eff}} | n \rangle = \frac{1}{T} \int_0^T dt (1 + \dot{f}) \langle m | \mathcal{H} | n \rangle + \frac{1}{T} \int_0^T dt \dot{\lambda} e^{-if(t)\omega_{mn}} \beta(t) \langle m | \partial_\lambda \mathcal{H} | n \rangle .$$

The second integral is responsible for off-diagonal terms in the eigenbasis, thus it can be used to control transitions towards excited eigenstates. For this purpose, we can fix $\beta(t)$ to match the off-diagonal term to the power series expansion of the AGP in Eq. (2.22).

Assuming that $\langle m | \partial_\lambda \mathcal{H} | n \rangle$ and $\dot{\lambda}$ are constant during a driving period, they can be taken out of the integral. The generic condition that matches the Floquet driving with CD driving reads as

$$i \left[\sum_{k=1}^{\ell} \alpha_k (\omega_m - \omega_n)^{2k-1} \right] = \frac{1}{T} \int_0^T dt e^{i(\omega_m - \omega_n)f(t)} \beta(t) .$$

It is more convenient to change variables such that the period T does not explicitly figure in the equivalence. Substituting $\omega dt = ds$, we get

$$i \left[\sum_{k=1}^{\ell} \alpha_k \omega_{mn}^{2k-1} \right] = \frac{1}{2\pi} \int_0^{2\pi} ds e^{i\omega_{mn}f(s)} \beta(s) \quad \forall \omega_{mn} . \quad (4.2)$$

To retrieve an expression in power series of ω_{mn} in RHS, the exponential $e^{i\omega_{mn}f}$ can be expanded in Taylor series for small values of $\omega_{mn}f$:

$$\int_0^{2\pi} e^{i\omega_{mn}f} \beta ds = \underbrace{\int_0^{2\pi} \beta(s) ds}_{=0} + i\omega_{mn} \underbrace{\int_0^{2\pi} \beta(s)f(s) ds}_{=2\pi\alpha_1} - \frac{\omega_{mn}^2}{2!} \underbrace{\int_0^{2\pi} \beta f^2 ds}_{=0} + \dots$$

Matching every power ω_{mn} with the AGP is equivalent to impose that

$$\int_0^{2\pi} ds f^k \beta = 0 \quad \forall \text{ even } k ,$$

$$\frac{1}{2\pi} \int_0^{2\pi} ds f^{2k-1} \beta \frac{(-1)^{k+1}}{(2k-1)!} = \alpha_k \quad \forall \text{ odd } k .$$

The two previous conditions are slightly more general than the ones imposed in Ref. [15], where the matching between AGP and the off-diagonal elements of \mathcal{H}_{eff} was made directly through the Bessel function expansion.

Let us proceed by developing the condition up to the first order in the CD

driving. We express $\beta(s)$ as a linear combination in $\underline{\alpha}$,

$$\beta(s) = \sum_k \alpha_k \beta_k(s)$$

which at first order ($\ell = 1$) simply reads as $\beta = \alpha_1 \beta_1$. The conditions required to match the counterdiabatic evolution at first order are

$$\begin{cases} \int_0^{2\pi} \beta_1(s) ds = 0 \\ \int_0^{2\pi} \beta_1(s) f(s) ds = 2\pi \end{cases} \quad (4.3)$$

To derive some new Floquet protocols, we choose to fix the driving functions to

$$\begin{cases} f(t) = \omega_0^{-1} \cdot (a \sin(\omega t) + b \cos(\omega t)) \\ \beta(t) = A \sin(\omega t) + B \cos(\omega t) + C \sin(2\omega t) + D \cos(2\omega t) \end{cases}, \quad (4.4)$$

requiring also $a^2 + b^2 = 1$, to limit the amplitude of $f(t)$. We observe that β includes only oscillating terms with frequency up to 2ω , but in principle, we could include any integer multiple $n\omega, n \in \mathbb{N}$. Following our derivation, the final result would not change, since we will see that all the frequencies multiple of ω will not show up in the protocol.

The two integrals in Eq. (4.3) are easy to solve. The first is already satisfied by our choice of $\beta(t)$; the second integral instead returns the condition

$$\pi Aa + \pi Bb = 2\pi\omega_0 \quad .$$

In the derivation of these new protocols, we also wish to minimize the error bound (see Eq. (3.12)) of the first-order Magnus expansion. In our previous discussions, the minimization of the error bound has proved to be a suitable criterion to improve the accuracy of the FE protocol (§3.2.b). This condition is imposed by minimizing the L2 norm of the Hamiltonian written in the rotating frame w.r.t. $\dot{f}\mathcal{H}$, i.e. $\tilde{\mathcal{H}} = \mathcal{H} + \beta \dot{\lambda} \partial_\lambda \mathcal{H}$. Proceed by writing explicitly \mathcal{V} :

$$\begin{aligned} \mathcal{V} &= \int_0^{2\pi} \|\tilde{\mathcal{H}}\|^2 ds = \\ &= \int_0^{2\pi} ds \underbrace{\|\mathcal{H}\|^2}_{\zeta_1} + \underbrace{2\alpha_1 \dot{\lambda} \langle \mathcal{H}, \partial_\lambda \mathcal{H} \rangle \beta(s)}_{\zeta_2} + \underbrace{\alpha_1^2 \|\partial_\lambda \mathcal{H}\|^2 \dot{\lambda}^2 \beta^2(s)}_{\zeta_3} = \\ &= \zeta_1 2\pi + \zeta_2 \cdot 0 + \zeta_3 \left(\int_0^{2\pi} ds A^2 \sin^2(s) + \int_0^{2\pi} ds B^2 \cos^2(s) + \dots \right) = \\ &= \zeta_1 2\pi + \zeta_3 (A^2 + B^2 + C^2 + D^2) \pi \end{aligned}$$

We now solve a system of equations that includes the minimization of \mathcal{V} with respect to A, B, C, D :

$$\begin{cases} \frac{\partial \mathcal{V}}{\partial A} = \frac{\partial \mathcal{V}}{\partial B} = \frac{\partial \mathcal{V}}{\partial C} = \frac{\partial \mathcal{V}}{\partial D} = 0 \\ \pi Aa + \pi Bb = 2\pi\omega_0 \\ a^2 + b^2 = 1 \end{cases}$$

The system is solved by the following conditions, expressed as a function of b :

$$\begin{cases} a = \pm\sqrt{1-b^2} \\ A = \pm 2\omega_0\sqrt{1-b^2} \\ B = 2\omega_0 b \\ C = D = 0 \end{cases} \quad (4.5)$$

In conclusion, we find a family of FE protocols parameterized by $b \in [-1, 1]$, whose driving Hamiltonian is

$$\begin{aligned} \mathcal{H}_{\text{FE}}^{(b)} = & \left(1 + \frac{\omega}{\omega_0} \left(\pm\sqrt{1-b^2} \cos(\omega t) - b \sin(\omega t) \right) \right) \mathcal{H} + \\ & + 2\omega_0\alpha_1\dot{\lambda} \left(\pm\sqrt{1-b^2} \sin(\omega t) + b \cos(\omega t) \right) \partial_\lambda \mathcal{H} \end{aligned} \quad (4.6)$$

We observe that the case for $b = 0$ leads back to the protocol of Ref. [15], whose Hamiltonian has already been discussed in Eq. (2.24). Instead, the other values of $b \in [-1, 1]$ identify a new $\ell = 1$ FE protocol.

We benchmark the protocol fixed by choosing $b = -1$:

$$\mathcal{H}_{\text{FE}}^{(b=-1)} = \left(1 + \frac{\omega}{\omega_0} \sin(\omega t) \right) \mathcal{H} + 2\omega_0\dot{\lambda}\alpha_1 \cos(\omega t) \partial_\lambda \mathcal{H} \quad (4.7)$$

The results in terms of the energy merit accuracy and gain are reported in Figure 4.1. Those results, to put it briefly, are comparable to the ones obtained with the protocol of Ref. [15]. This statement is supported by comparing bespoke Figure 4.1 to Figures 3.8b and 3.12a. Because of such similarity, our conclusions about the protocol of Eq. (4.7) are identical to the ones discussed at the end of Chapter 3: there exists a narrow region of parameters (ω_0, ω) with limited advantage in the regime of small driving frequencies, whereas there is no advantage with higher driving frequencies.

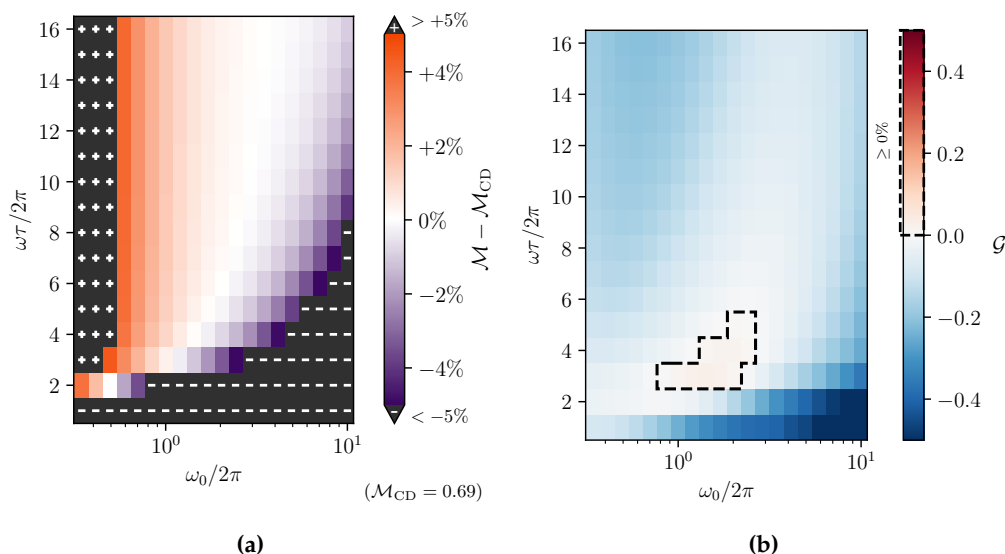


Figure 4.1: (a) Energy merit difference and (b) gain for the protocol of Eq. (4.7), executed on the LHZ6M instance (see Table 3.2) with $\tau = 0.1$. The vertical axis fixes the total number of driving cycles of the protocol ($\omega\tau/2\pi$); the horizontal axis denotes the spectral reference frequency $\omega_0/2\pi$.

To summarize, we have derived a family of Floquet protocols at first counterdiabatic order. The family includes also the original protocol for $\ell = 1$ of Ref. [15]. However, the evidence collected through simulations of a specific instance of the new protocol family does not point to any substantial improvement in the gain, nor in the optimal driving frequency area.

4.1.b Third-order Floquet protocols

The FE protocols of Ref. [15] have been derived under two fundamental approximations, as we have already established in section §2.4. (1) The Magnus expansion at first order, to find the effective Hamiltonian from the time evolution operator. (2) The Bessel function expansion, to match the power series in ω_{mn} between the FE effective Hamiltonian and the AGP expression.

We now inquire if the protocol yield improves by pushing these approximations to higher orders of accuracy. In detail, we have focused our efforts on finding new protocols that push the Magnus expansion up to third order. The Hamiltonians of the new protocols are reported in Appendix B. Now, we will analyze the accuracy and efficiency of the new protocols using the framework introduced in Chapter 3.

The new protocols at higher order have a driving Hamiltonian in the form of

$$\mathcal{H}_{\text{FE}} = \left(1 + \frac{1}{\omega_0} \dot{f}\right) \mathcal{H} + \beta(t) \dot{\lambda} \cdot \partial_{\lambda} \mathcal{H} \quad (4.8)$$

with f and β specifically defined for each protocol. The features of the new protocols have been summarized in Table 4.1. Essentially, we have found new protocols at the second $\mathcal{O}(\delta^2)$ and third order $\mathcal{O}(\delta^3)$ of the Magnus expansion¹ for both $\ell = 1$ and $\ell = 2$. The protocols will be conveniently labeled $\mathcal{O}2$ and $\mathcal{O}3$ during the discussion. Additionally, we consider two more third-order protocols that have been formulated by removing the highest driving frequency multiple from the $\mathcal{O}3$ protocols. Those ‘reduced frequency’ protocols are instead labeled as $\mathcal{O}3\text{L}$ (‘L’ standing for *Lower-frequency*). From now on, the protocols of Ref. [15] will be labeled as the ‘Claeys’ protocols, for simplicity.

Magnus order label		$\mathcal{O}(\delta^2)$	$\mathcal{O}(\delta^3)$	
		$\mathcal{O}2$	$\mathcal{O}3$	$\mathcal{O}3\text{L}$
$\ell = 1$	frequency	ω	$\omega, 3\omega$	ω
	protocol	Eq. (B.1)	Eq. (B.2)	Eq. (B.3)
$\ell = 2$	frequency	$\omega, 3\omega$	$\omega, 3\omega, 5\omega$	$\omega, 3\omega$
	protocol	Eq. (B.4)	Eq. (B.5)	Eq. (B.6)

Table 4.1: Additional FE protocols, pushing to higher order the Magnus expansion. The Table references the equation of the protocols from Appendix B, highlighting their driving frequencies. The $\mathcal{O}3\text{L}$ protocols have been obtained from the $\mathcal{O}3$ protocols by removing the highest multiple of the driving frequency.

Before proceeding with the analysis, we have noticed that the protocol for $\ell = 1$ and expansion order $\mathcal{O}(\delta^2)$ matches with the FE protocol derived in the previous section (see Eq. (4.7)). We also observe that the protocols at second order $\mathcal{O}(\delta^2)$ are very similar to the Claeys protocols, for both $\ell = 1, 2$, which are obtained instead through a Magnus expansion at first order. The only noticeable difference is that the oscillating terms in f and β have been shifted in phase by $\pi/2$. We stress that even though our derivation of $\mathcal{O}(\delta^2)$ protocols has been carried out explicitly to find protocols at a higher order of accuracy, it is not guaranteed that the resulting schedules cannot be derived within a lower order of approximation too.

Nevertheless, this fortunate coincidence allows us to skip the analysis of the

¹The notation $\mathcal{O}(\delta^n)$ denotes the error bound of the Magnus expansion at order n . It is shown in Ref. [8] that the order- n terms of the Magnus expansion are upper bounded by δ^n , where δ is defined in Eq. (3.12).

$\mathcal{O}2$ protocols, since the $\ell = 1$ instance has already been examined in the preceding section, finding no appreciable improvement over the Claeys protocol.

It is more interesting to study the new protocols at order $\mathcal{O}(\delta^3)$. Indeed, the new protocol schedules differ clearly from the previous ones, due to the addition of new oscillating terms that are divided by the factor $\left(2 + \frac{2\pi}{\omega} \frac{\dot{\lambda}}{\lambda}\right) \cdot \omega_0$. For instance, the driving functions $f(t)$ and $\beta(t)$ of the $\ell = 2$ $\mathcal{O}3L$ protocol are defined by (B.6), which we report hereby:

$$\begin{cases} \beta(t) = 2\omega_0 [\cos(\omega t) - 3 \cos(3\omega t)] \alpha_1(t) - 48\omega_0^3 [\cos(3\omega t)] \alpha_2(t) \\ \quad + \omega_0^{-1} \left(2 + \frac{2\pi}{\omega} \frac{\dot{\lambda}(t)}{\lambda(t)}\right)^{-1} [\cos(\omega t) - 3 \cos(3\omega t)] \\ f(t) = \cos(\omega t) \end{cases} .$$

Let us discuss, first, the yield of the $\ell = 1$ $\mathcal{O}(\delta^3)$ protocols. To evaluate the accuracy of the protocol, as done in Chapter 3, we look at the energy merit \mathcal{M} obtained in many configurations (ω_0, ω) of the FE protocol and we compare it to the energy merit \mathcal{M}_{CD} obtained with the reference CD protocol. We execute the benchmarks using the reference setup, which simulates the annealing schedules of the LHZ6M instance (see Table 3.2) with total annealing time $\tau = 0.1$. The energy merit difference $(\mathcal{M} - \mathcal{M}_{\text{CD}})$ of the $\mathcal{O}3$ and $\mathcal{O}3L$ protocols is reported in Figures 4.2b and 4.2c, respectively. The main comparison case is the Claeys protocol of Ref. [15], whose energy merit difference has been plotted again in Figure 4.2a for quick comparison. We observe that the region of optimal accuracy (i.e. $\mathcal{M} - \mathcal{M}_{\text{CD}} = 0 \pm 5\%$) is improved for both the new protocols, with the low-frequency variant $\mathcal{O}3L$ achieving better results. More in detail, the region of spectral reference frequency $\omega_0/2\pi \leq 1$ is now accessible with good convergence over the target CD energy merit.

What about the efficiency of the new protocols? In the previous Chapter, we have introduced the gain \mathcal{G} as a measure of the advantage in the energy merit which takes into account the cost of the protocol w.r.t. unassisted annealing schedules. Following the same approach, we plot the gain achieved by the $\mathcal{O}3$ and $\mathcal{O}3L$ protocols in Figures 4.3b and 4.3c. Comparing the results with the Claeys protocol - which is plotted again in Figure 4.3a to aid the comparison - it is possible to observe a slightly larger region of gain neutrality in the $\mathcal{O}3L$ protocol at higher driving frequencies. However, the maximum gain value, 0.0365, is reached with only 2 driving cycles per schedule, which is a regime of driving frequencies that should be avoided for the reasons discussed at the end of Chapter 3: the FE protocol realizes the CD evolution in only one driving cycle, which is not optimal for the hypothesis under which the protocols have been derived.

We observe that the third-order protocol featuring all the additional frequen-

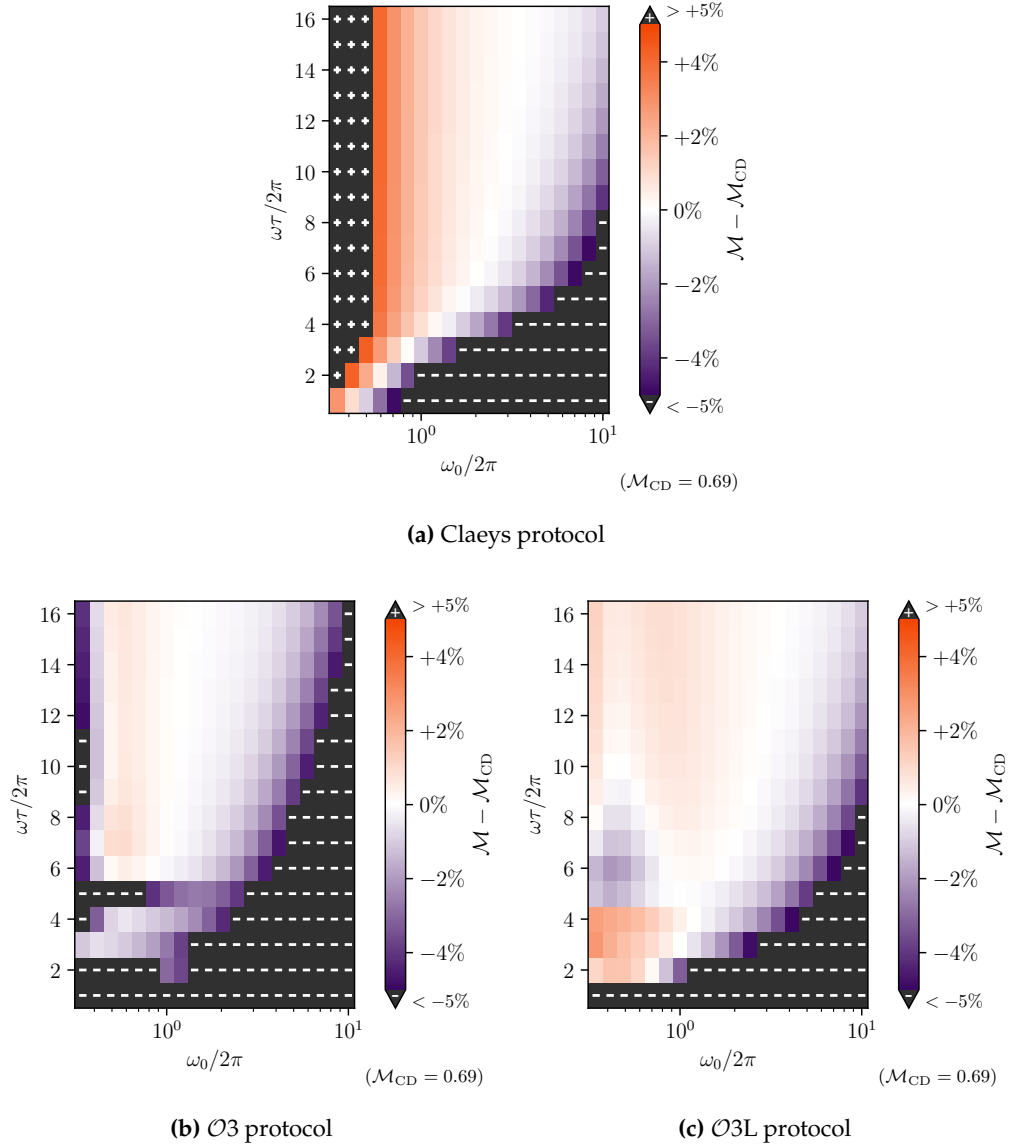


Figure 4.2: Difference in energy merit for the new $\mathcal{O}(\delta^3)$ FE protocols w.r.t. the reference $\ell = 1$ CD protocol. Figure (b) is realized by using the $\mathcal{O}3$ protocol, whereas the $\mathcal{O}3L$ protocol (i.e. its version with lower frequency oscillations) is used in Figure (c). For comparison, Figure (a) shows the results of the Claeys protocol. The annealing is performed on the LHZ6M instance (see Table 3.2) at $\tau = 0.1$. The points that exceed the $\pm 5\%$ interval around zero are cropped to a grey color and marked with a $+/-$ symbol to indicate the sign of the extreme. The vertical axis fixes the total number of driving cycles of the protocol ($\omega\tau/2\pi$); the horizontal axis denotes the spectral reference frequency $\omega_0/2\pi$.

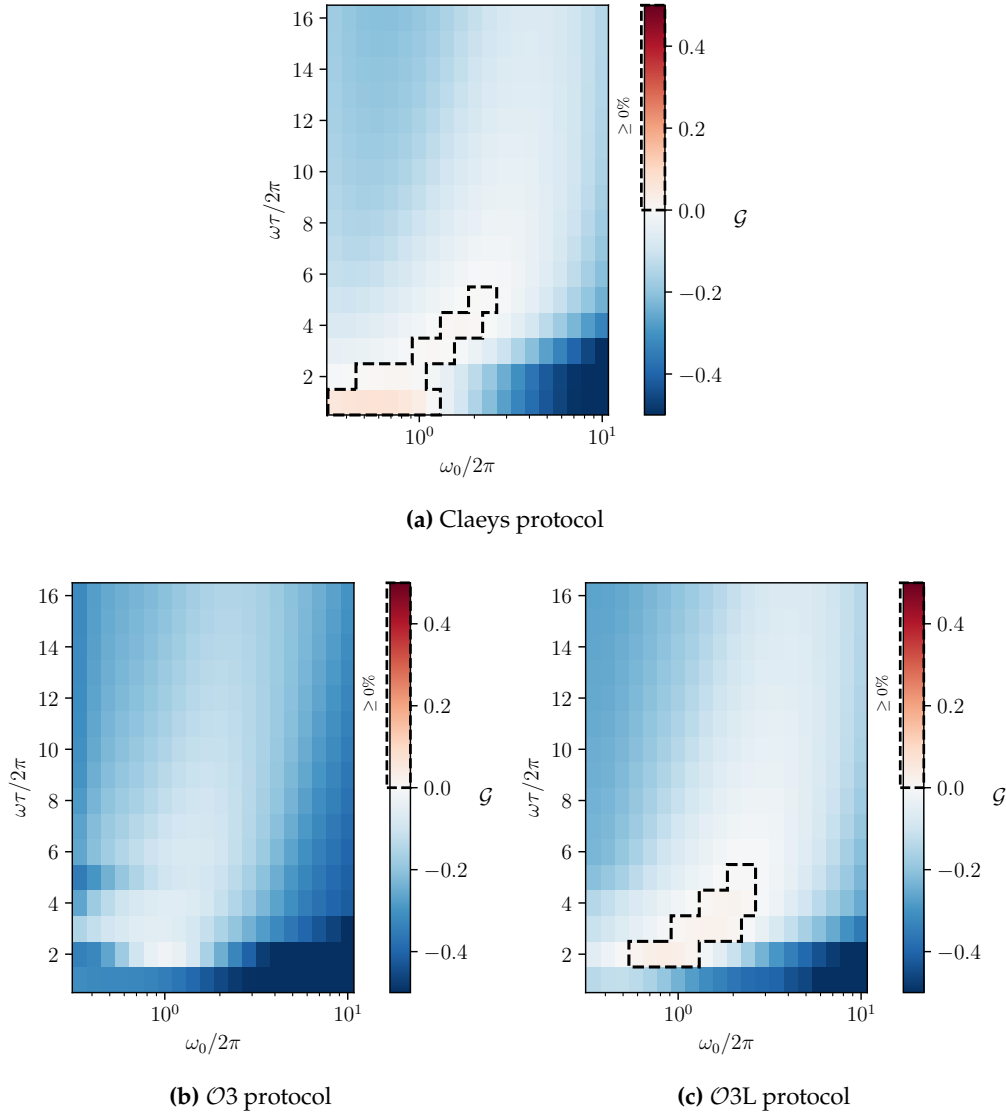


Figure 4.3: Gain for the new $\ell = 1$, $\mathcal{O}(\delta^3)$ FE protocols. Figure (b) is realized by using the $\mathcal{O}3$ protocol, whereas the $\mathcal{O}3L$ protocol (i.e. its version with lower frequency oscillations) is used in Figure (c). For comparison, Figure (a) shows the results of the Claeys protocol. The annealing is performed on the LHZ6M instance (see Table 3.2) at $\tau = 0.1$. The vertical axis fixes the total number of driving cycles of the protocol ($\omega\tau/2\pi$); the horizontal axis denotes the spectral reference frequency $\omega_0/2\pi$.

cies, $\mathcal{O}3$, is not able to achieve gain neutrality with any value of the driving frequencies ω . We conjecture that this behavior is due to the larger coefficients that multiply the higher frequency terms (3ω , in this case), which increase the effective time τ_{eff} of the schedule and negatively impact the gain. To support this claim, we show in Figure 4.4 that the effective time τ_{eff} is sensitively higher for the $\mathcal{O}3$ protocol in the window of reference frequencies $\omega_0 > 2\pi$. Instead, the $\mathcal{O}3\text{L}$ protocol shows effective times similar to the Claeys protocol of Ref. [15], with absolute differences being in the order of 10^{-2} .

Since our analysis has shown that the most interesting region of driving frequencies is within $2 \leq \omega\tau/2\pi \leq 5$, we also show the corresponding data in 2D plots, separating each driving frequency ω in its own figure. Figure 4.5 shows the gain \mathcal{G} for a few, small driving frequencies. Positive gain values are achieved by the $\mathcal{O}3\text{L}$ protocol, on the contrary of the $\mathcal{O}3$ protocol. The gain of Claeys protocols is lower than the one obtained by the $\mathcal{O}3\text{L}$.

In conclusion, we see that the $\mathcal{O}3\text{L}$ protocol leads to some advantage w.r.t. the Claeys protocol:

- The $\mathcal{O}3\text{L}$ protocol can substantially improve the convergence to the target CD energy merit in the regime of low spectral reference frequencies $\omega_0 < 2\pi$, which was previously inaccessible with the protocols of Ref. [15].
- The gain of $\mathcal{O}3\text{L}$ is slightly positive only in the regime of small driving frequencies. However, this regime is not optimal due to the hypothesis underlying the derivation of the FE protocols. Nevertheless, we observe that the optimal driving frequency has increased by one driving cycle w.r.t. the protocol of Ref. [15].

Results with second order counterdiabatic protocols

The $\ell = 2$ Claeys protocols of Ref. [15] showed a limitation in the feasible window of driving frequencies ω . Indeed, we have been forced to benchmark a window of driving oscillations in the range $60 \leq \omega\tau/2\pi \leq 160$ to find a suitable region of energy merit accuracy $\mathcal{M} - \mathcal{M}_{\text{CD}} \simeq 0 \pm 5\%$ (see Figure 3.7). The feasible region of parameters, however, is narrow in the reference spectrum frequency ω_0 , showing suboptimal results even with $\omega_0 = 1 \cdot 2\pi$. Furthermore, the analysis based on the gain concluded that $\ell = 2$ FE protocols were never advantageous w.r.t. UA schedules (see Figure 3.12b). What about the new $\mathcal{O}(\delta^3)$ protocols?

It is more interesting, once again, to look at the results on the $\mathcal{O}3\text{L}$ protocol. For convenience, the plots are reported in Appendix C. The region of energy merit

accuracy (Figure C.1c) is significantly amplified. In particular, the region of low-frequency driving $6 \leq \omega\tau/2\pi \leq 16$ benefits from the new protocol, showing optimal accuracy with small spectral reference frequencies $\omega_0 < 1 \cdot 2\pi$. The gain \mathcal{G} is still overall negative, as can be seen in Figure C.2c, even though the improvement of $\mathcal{O}3L$ w.r.t. Claeys is evident.

4.2 A quick look at the annealing of an Ising model

It is interesting to determine whether our results apply generally to annealing with FE protocols or are restricted to the particular case of Parity architecture annealing. Unfortunately, the evaluation of the gain metric requires knowledge of the energy merit of the protocol, which can be computed only through numerical simulations. This aspect of our analysis framework makes it difficult to formulate predictions about larger system sizes, even for a fixed problem. For the same reason, it is impossible to generalize analytically our conclusions to any possible combination of \mathcal{H}_i and \mathcal{H}_p .

Nevertheless, to probe the general validity of our conclusion, we repeat our analysis on a paradigmatic model that is traditionally used as a test bench in Quantum Annealing: the Ising model. More in detail, we benchmark the annealing of a 1D chain with homogeneous fields and interactions, fixing the number of spins $N = 6$ to match the same number of qubits used until now with

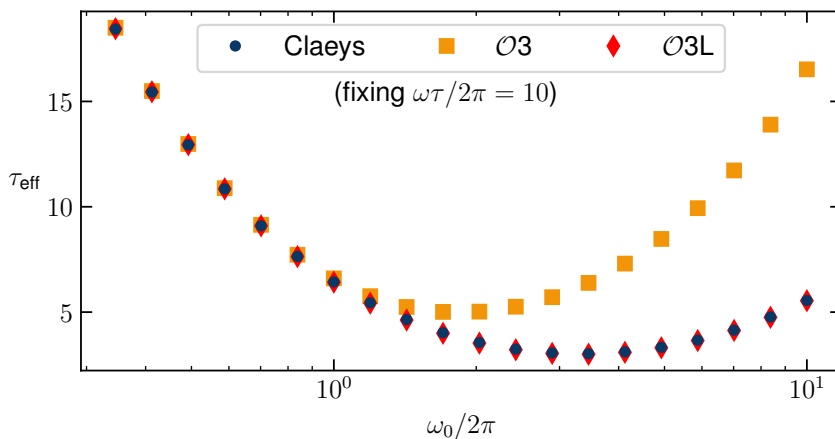


Figure 4.4: Effective time across protocols at first counterdiabatic order ($\ell = 1$), obtained by fixing $\omega\tau/2\pi = 10$ driving cycles and changing the spectral reference frequency ω_0 . Simulation for the LHZ6M instance (see Table 3.2) with $\tau = 0.1$.

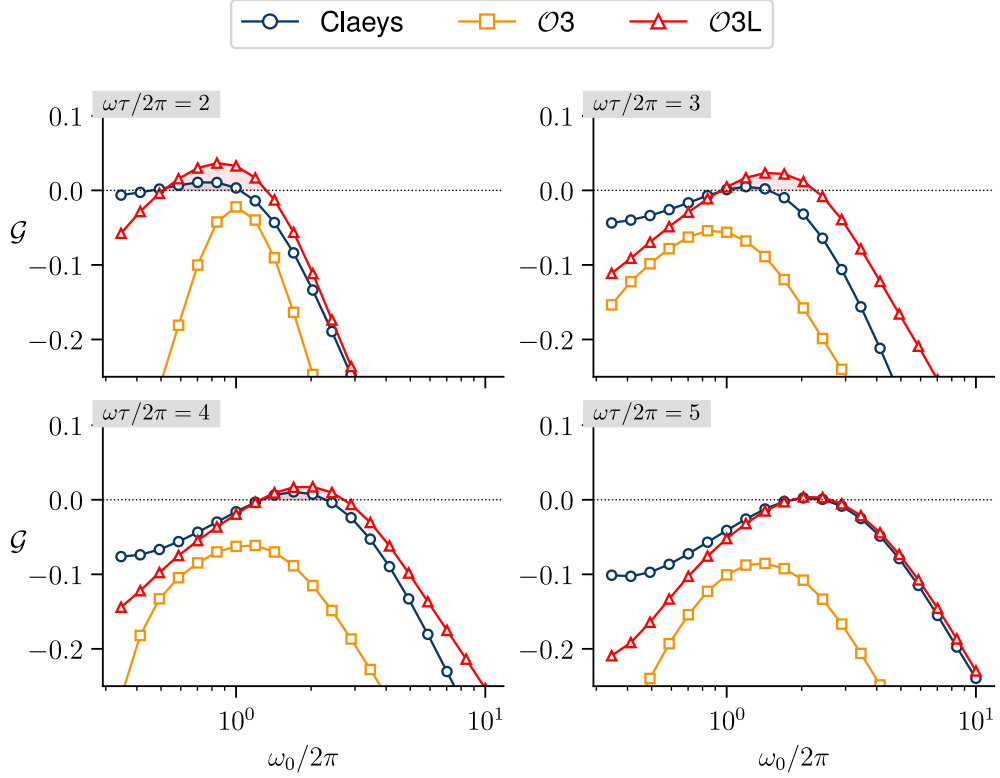


Figure 4.5: Gain comparison across the Claeys, $\mathcal{O}3$, and $\mathcal{O}3L$ protocols at $\ell = 1$. Each subplot fixes the number of driving oscillations in $2 \leq \omega\tau/2\pi \leq 5$. The simulations are executed on the LHZ6M instance (see Table 3.2) with annealing time $\tau = 0.1$. We look after positive gain values $\mathcal{G} > 0$, highlighting the area under the curve if it is positive. No protocol achieves positive gain for $\omega\tau/2\pi \geq 6$. In the plots with fewer driving oscillations, the $\mathcal{O}3L$ protocol stands out and achieves a marginally positive gain of $\mathcal{G}_{\max} = 3.65\%$. The Claeys protocol gain is positive but much smaller. The $\mathcal{O}3$ protocol gain is strongly negative through all the data.

the LHZ model. The problem Hamiltonian is

$$\mathcal{H}_p^{\text{Ising}} = \sum_i^N \hat{\sigma}_z^{(i)} - \sum_i^N \hat{\sigma}_z^{(i)} \hat{\sigma}_z^{(i+1)} \quad , \quad (4.9)$$

and the rest of the simulation setup is unchanged: the total annealing time is set to $\tau = 0.1$ and the initial Hamiltonian is $\mathcal{H}_i = -\sum_i^N \hat{\sigma}_x^{(i)}$.

Comparing the Claeys and $\mathcal{O}3\text{L}$ protocols with CD order $\ell = 1$, we observe the same improvements that have been previously observed in the LHZ instances:

- As shown in Figure 4.6, the new $\mathcal{O}(\delta^3)$ protocol can drive the system correctly at lower reference spectrum frequencies ω_0 .
- The gain is slightly improving with the higher-order protocol - see Figure 4.7. However, the positive gain remains in a region of driving frequencies that is infeasible given the hypotheses that have been used to derive the Floquet driving protocols.

The same conclusions apply for the $\ell = 2$, $\mathcal{O}3\text{L}$ protocol, whose plots are reported in Appendix C.

It is hard to draw general conclusions only from a single batch of simulations. Still, the higher-order protocols do not bring a noticeable improvement in efficiency, which is understood in terms of gain. However, the new $\mathcal{O}3\text{L}$ protocol improves the accuracy in the region of small spectral reference frequencies ω_0 .

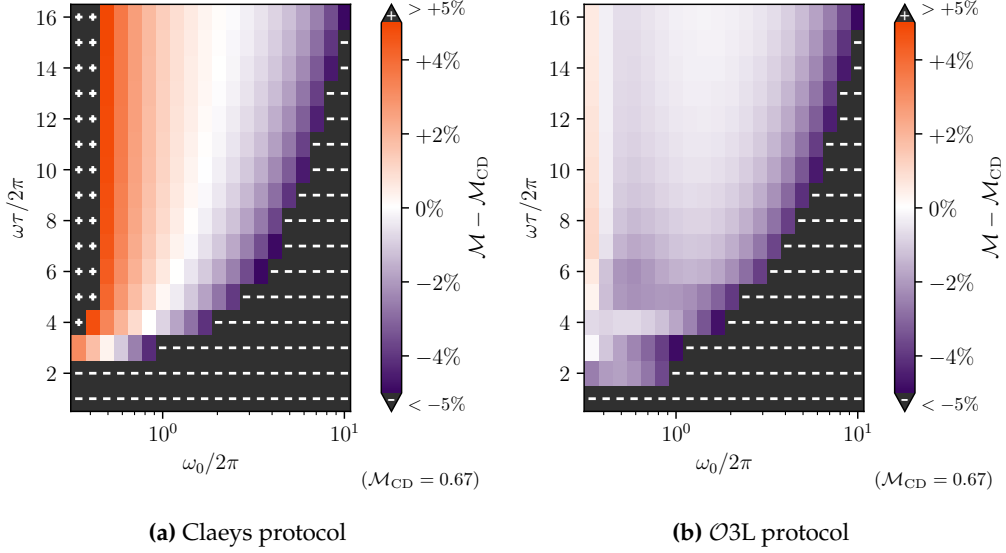


Figure 4.6: Ising model ($N = 6$) annealing. Difference in energy merit of the FE protocols w.r.t. the reference $\ell = 1$ CD protocol. For comparison, Figure (a) shows the results of the Claey's protocol. Figure (b) is realized by using the $\mathcal{O}3\text{L}$ protocol. The points that exceed the $\pm 5\%$ interval around zero are cropped to a grey color and marked with a $+/-$ symbol to indicate the sign of the extreme. The annealing is simulated with $\tau = 0.1$. The vertical axis fixes the total number of driving cycles of the protocol ($\omega\tau/2\pi$); the horizontal axis denotes the spectral reference frequency $\omega_0/2\pi$.

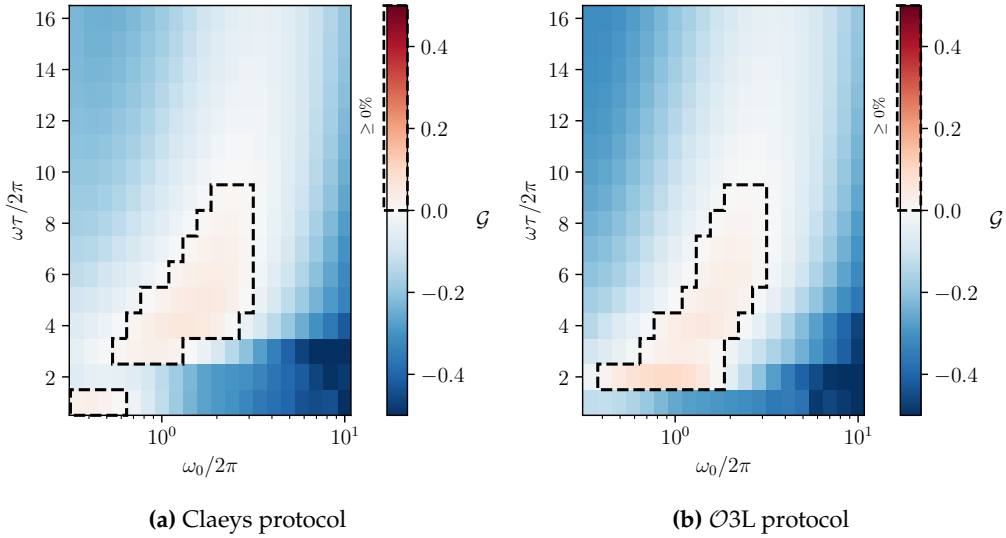


Figure 4.7: Ising model ($N = 6$) annealing. Gain for the new $\ell = 1$, $\mathcal{O}(\delta^3)$ FE protocols. For comparison, Figure (a) shows the results of the Claey's protocol. Figure (b) is realized by using the $\mathcal{O}3\text{L}$ protocol. The annealing is simulated with $\tau = 0.1$. The vertical axis fixes the total number of driving cycles of the protocol ($\omega\tau/2\pi$); the horizontal axis denotes the spectral reference frequency $\omega_0/2\pi$.

4.3 Alternative proposals within counterdiabaticity

We have seen that Floquet protocols retain the advantage of effectively implementing a counterdiabatic evolution without requiring additional system controls. However, our analysis has determined a limited advantage in using Floquet protocols w.r.t. unassisted annealing schedules. This leads us to a final question: are there other ways to exploit counterdiabaticity that show instead some extent of advantage?

This last section of the thesis explores alternative counterdiabatic protocols which exempt from the Floquet formulation. Specifically, we wish to drop the requirement of not extending the set of protocol controls: perhaps, it might be advantageous to allow a very limited set of extra control operators, without impacting too much on the experimental realization. Therefore, our goal is now to identify a ‘compromise’ protocol that allows for improved ground-state preparation while keeping the experimental realization as simple as possible.

4.3.a Local Counterdiabatic Driving

The machinery required to perform this task has already been introduced in §2.2.b. In practice, we need to find an ansatz of the AGP that has very limited operator support. Once the ansatz is determined, it can be optimized via the variational minimization of Eq. (2.20).

In Local Counterdiabatic Driving (LCD) [78], the approximation ansatz of the AGP is formed by local operators. The most simple example in our case scenario would be a local $\hat{\sigma}_y$ control on each qubit². In a LCD protocol, we can also allow the use of quasi-local interactions, i.e. terms involving a small number of qubits.

Ref. [40] has already implemented this kind of protocol on the LHZ model. The authors make use of an ansatz with time-dependent local y-magnetic fields,

$$\mathcal{A}_{\text{LCD}} = \sum_i^N \alpha_i(t) \hat{\sigma}_y^{(i)} \quad ,$$

and optimize variationally $\underline{\alpha}(t)$. The LCD protocol effectively improves the fidelity of the evolved state w.r.t. the true ground state, without posing a significant overhead in the practical implementation due to the local nature of the extra controls. The authors show also that a suitable *quasi-local ansatz* for the

²Recall that the AGP ansatz must not commute with the system Hamiltonian. Therefore, local fields like $\hat{\sigma}_x$ and $\hat{\sigma}_z$ would not suffice, since they are already present in the protocol Hamiltonian (\mathcal{H}_i and \mathcal{H}_p).

LHZ architecture would include the following non-homogeneous constraints:

$$\mathcal{A}_{\text{qLCD}} = \sum_i^N \alpha_i(t) \hat{\sigma}_y^{(i)} + \mu_0(t) \sum_l \hat{\rho}_{[yzzz]}^{[l]} + \mu_1(t) \sum_l \hat{\rho}_{[zyyy]}^{[l]} \quad (4.10)$$

where we have used the notation introduced in Appendix A. In simple terms, a symbol like $\hat{\rho}_{[yzzz]}^{[l]}$ expands as a sum of 4-body constraints that permute the operators ($\hat{\sigma}_y, \hat{\sigma}_z, \hat{\sigma}_z, \hat{\sigma}_z$) on each qubit of the constraint $[l]$:

$$\hat{\rho}_{[yzzz]}^{[l]} = \begin{array}{c} y \quad z \\ \blacksquare \\ z \quad z \end{array} + \begin{array}{c} z \quad y \\ \blacksquare \\ z \quad z \end{array} + \begin{array}{c} z \quad z \\ \blacksquare \\ z \quad y \end{array} + \begin{array}{c} z \quad z \\ \blacksquare \\ y \quad z \end{array}$$

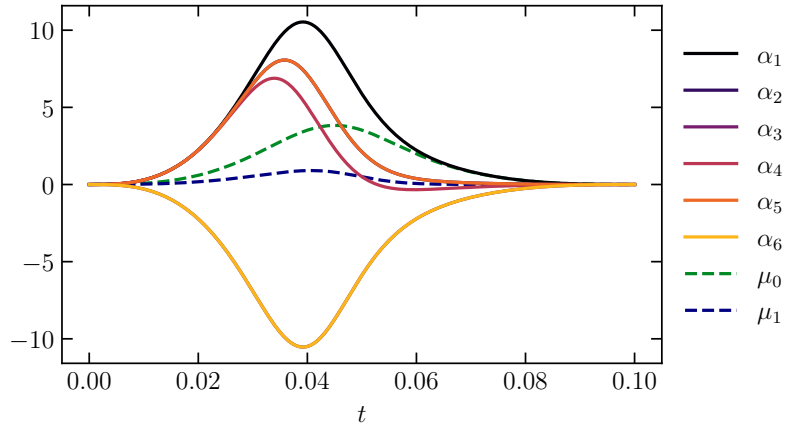
The ‘qLCD’ in Eq. (4.10) stands for *quasi*-LCD. We report in Figure 4.8 the optimized coefficients (α, μ_0, μ_1) and the energy merit for the LCD and the qLCD protocols executed on the LHZ6M instance (see Table 3.2). Not surprisingly, the qLCD protocol benefits from increased operator support and reaches higher energy merits. However, the constraints controlled through the coefficients μ_i are hard to implement and might pose some limitations in the protocol engineering. We refer to Ref. [40] for more details about LCD protocols applied to the Parity architecture.

4.3.b Optimized Local Counterdiabatic Driving

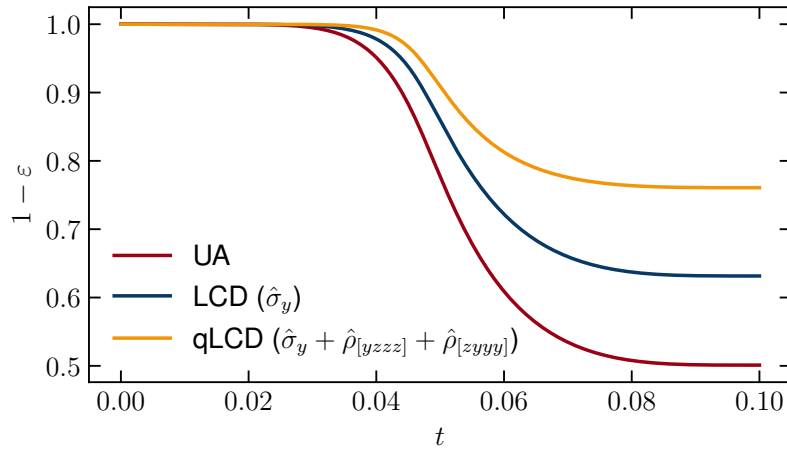
In this section, we will adapt an approach that has been proposed in Ref. [9], where the authors combine LCD with a method belonging to the domain of Quantum Optimal Control (QOC) [18, 36, 51, 71] theory.

The Chopped RANdom Basis (CRAB) algorithm [65, 75] is a paradigmatic QOC method, in which the controls of a quantum system are modified through additional optimizable pulses. The key element of this algorithm is to optimize the shape of the pulses to maximize a metric, which in the context of state preparation is typically the fidelity of the prepared state. To this extent, the control pulse is expanded in a suitable finite basis of functions: $f_{\text{OC}}(t) = \sum_k^{N_k} \gamma_k f_k(t)$, where the γ_k are the coefficients that should be optimized. CRAB has already been applied to many-body systems [24, 12], showing consistent results even by using a small number of tunable parameters N_k .

In Ref. [9], the schedule coefficients of the trivial annealing approach ($A(t), B(t), \dots$) are modified with tunable pulses while operating LCD. As in CRAB, the authors optimize the pulses by maximizing the fidelity of the prepared state. A limitation of such a method is that it requires the execution of annealing sim-



(a)



(b)

Figure 4.8: LCD protocols on LHZ6M instance, with total annealing time $\tau = 0.1$. Figure (a) shows the optimized AGP coefficients α_i , μ_0 and μ_1 of the ansatz in Eq. (4.10). Figure (b) reports the energy merit of the LCD protocols across a single annealing simulation.

ulations to optimize the parameters through the fidelity of the prepared state, which is an operation that might be expensive if the target quantum system involves a large number of qubits. However, the authors also show that typically the optimization ends up in maximizing the contribution on the AGP from local operators while minimizing the amplitude of non-local terms of the ansatz - i.e. the μ_s of Eq. (4.10). This observation suggests that the optimization of the fidelity via simulations could be replaced by a much cheaper optimization of the non-local terms amplitude. In our approach, we would like to take advantage of this property.

In the following, we propose an Optimized Local Counterdiabatic Driving (OLCD) protocol. Let us start by rewriting the UA protocol,

$$\mathcal{H}_{\text{UA}}(t) = A(t) \left(- \sum_k \hat{\sigma}_x^{(k)} \right) + B(t) \sum_k J_k \hat{\sigma}_z^{(k)} + C(t) \sum_l^{N_c} \hat{\rho}_z^{[l]} \quad (4.11)$$

where $A(t) = 1 - \lambda(t)$ and $B(t) = C(t) = \lambda(t)$. The only variation from Eq. (3.2) is that the coefficient of the constraints is now independent from the coefficient of the local fields in the problem Hamiltonian.

We now introduce a parameterized pulse in the schedule functions,

$$A(t) \longrightarrow A(t, \underline{\gamma}^{(A)}) = A(t) + f_{\text{OC}}(\underline{\gamma}^{(A)})$$

and likewise for $B(t) \rightarrow B(t, \underline{\gamma}^{(B)})$ and $C(t) \rightarrow C(t, \underline{\gamma}^{(C)})$. The CRAB parameterized pulse is expanded in a basis made of sinusoidal waveforms:

$$f_{\text{OC}}(t, \underline{\gamma}) = \sum_{k=1}^{N_k} \gamma_k \sin(\pi k t / \tau) \quad ,$$

where γ_k is a coefficient associated with each frequency $k\pi/\tau$ of the control pulse. The number of frequencies N_k is a hyperparameter of our protocol. Larger values of N_k enable the creation of pulses with more complex shapes, while $N_k \rightarrow 1$ represents a straightforward sinusoidal parabola in the limit case. We stress that each schedule function (A, B, C) has its own parameter vector $\underline{\gamma}^{(A)}$, $\underline{\gamma}^{(B)}$, and $\underline{\gamma}^{(C)}$. All the pulse parameters can be collected, for simplicity of notation, under a single vector of parameters $\underline{\gamma} = (\underline{\gamma}^{(A)}, \underline{\gamma}^{(B)}, \underline{\gamma}^{(C)})$ with $3 \cdot N_k$ entries.

After choosing the annealing Hamiltonian of Eq. (4.11), we perform the variational optimization of the counterdiabatic coefficients ($\underline{\alpha}, \mu_0, \mu_1$) using the $\mathcal{A}_{\text{qLCD}}$ ansatz in Eq. (4.10). In our framework, the optimized expressions of the CD coef-

ficients will depend on $\underline{\gamma}$. We use this dependence to our advantage, effectively minimizing the contributions coming from hard many-body LCD terms, i.e. the μ_0 and μ_1 . Indeed, the vector of parameters $\underline{\gamma}$ is updated using a classical Powell optimizer, minimizing the following loss function.

$$\mathcal{L} = \int_0^\tau (|\mu_0(t)| + |\mu_1(t)|) dt + r \|\underline{\gamma} - 1\| \quad (4.12)$$

The first term of the loss function is the integral of non-local counterdiabatic coefficients $\mu_0(t)$ and $\mu_1(t)$. The second term acts as a regularization to avoid large values of the control vector $\underline{\gamma}$. The factor r is suitably small: $r = 0.005$.

After executing the numerical optimization, we end up with an optimized vector of pulse parameters $\underline{\gamma}'$ and the optimized control terms $\underline{\alpha}'$. Eventually, we drive our system with the new OLCD protocol by adding only the local control fields:

$$\mathcal{H}_{\text{OLCD}}(t, \underline{\gamma}') = \mathcal{H}_{\text{UA}}(t, \underline{\gamma}') + \sum_k \alpha'_k \hat{\sigma}_y^{(k)} \quad (4.13)$$

We provide an example of the optimized OLCD schedule coefficients in Figure 4.9, along with the energy merit of a single annealing schedule, comparing it across all the LCD protocols that we have discussed.

Advantage of protocols with different interactions

Eventually, our goal is to assess if the execution of a new protocol is advantageous. This essentially means comparing the prepared state energy of the new protocol with the one obtained through unoptimized annealing schedules at equal cost, which is exactly how the gain \mathcal{G} has been defined in §3.3.b. However, we cannot use anymore the effective time τ_{eff} to play the role of the cost since the protocol $\mathcal{H}_{\text{OLCD}}(t)$ is not expressed anymore in the form $A(t)\mathcal{H}_i + B(t)\mathcal{H}_p$. We need instead to use a cost definition that considers arbitrary control terms in \mathcal{H} . We have already introduced a suitable definition of cost in Eq. (3.13):

$$\mathcal{C} = \int_0^\tau \|\mathcal{H}\| dt \quad .$$

Therefore, we generalize the gain as the difference of energy merit of the OLCD protocol w.r.t. UA protocol at an equal cost \mathcal{C} .

$$\mathcal{G}' = \mathcal{M}_{\text{OLCD}}(\mathcal{C}) - \mathcal{M}_{\text{UA}}(\mathcal{C}) \quad (4.14)$$

To compute the gain under the new definition, we have integrated numerically

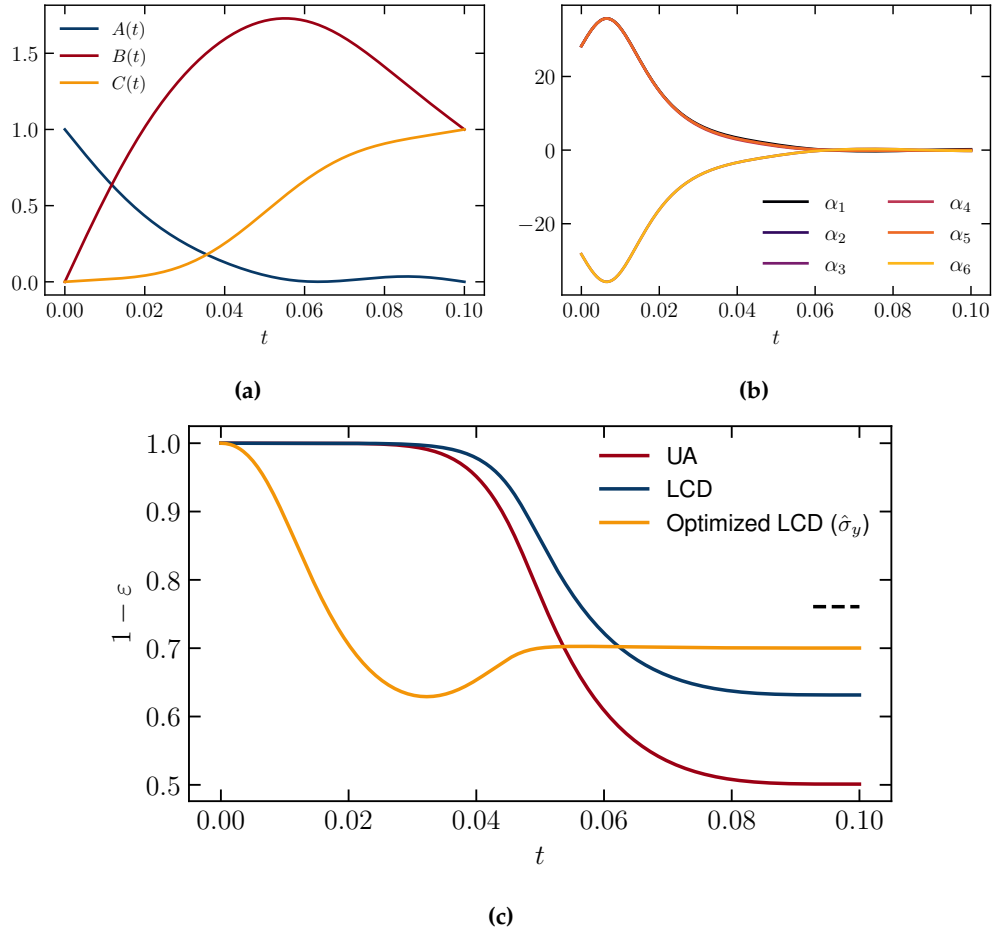


Figure 4.9: OLCD protocol executed on LHZ6M instance (see Table 3.2), with $\tau = 0.1$. The Figures (a) and (b) show the optimized driving coefficients. Figure (c) compares the energy merit across LCD protocols discussed in this section. The dashed black line in Figure (c) indicates the value of $1 - \epsilon(\tau)$ obtained via the qLCD protocol in equivalent settings.

the Hamiltonian of LCD, qLCD, and OLCD protocols. The resulting gain - shown in Table 4.2 - is positive for all three protocols, with our OLCD protocol improving by 33% the gain over the LCD protocol that uses the same extra set of controls (the locally tuned $\hat{\sigma}_y$). The OLCD gain is still lower than the one obtained with the qLCD protocol, which is not expected since qLCD can take advantage of a larger set of support operators for the AGP ansatz.

protocol	control terms	\mathcal{G}'	\mathcal{M}
LCD	$\hat{\sigma}_y$	0.12	0.6315
Optimized LCD	$\hat{\sigma}_y$	0.16	0.7002
qLCD	$\hat{\sigma}_y, \hat{\rho}_{[yzzz]}, \hat{\rho}_{[zyyy]}$	0.25	0.7607
UA	-	0	0.5011

Table 4.2: Statistics about the LCD protocols applied to the quantum annealing on the LHZ6 instance (see Table 3.2) with $\tau = 0.1$. We show the gain, computed with the definition of Eq. (4.14), and the energy merit achieved by the protocols. We also show the corresponding values of the UA protocol and the additional control operators used in the LCD protocols.

In conclusion, our OLCD protocol is derived from a numerical optimization of LCD schedules, using only local field controls. The optimization is performed by modifying the schedule functions of the protocol with parametrized pulses, like in a CRAB approach. Contrary to several other QOC approaches, this OLCD protocol has the advantage that the optimization does not require the feedback of annealing simulations, making this method suitable also for larger system instances. The positive gains of LCD protocols in Table 4.2 can be seen as proof of the advantage. However, we stress that our definition of the cost \mathcal{C} in Eq. (3.13) does not take into account the experimental limitations on the control fields. However, if one can pay the price of an extended set of control operators, LCD protocols seem to be a viable and advantageous option.

Conclusions

In this thesis, we analyze the effectiveness of Floquet counterdiabatic protocols in the context of Quantum Annealing. This study has focused specifically on Parity architecture models. However, we explicitly verified that our conclusions also hold for the annealing of a typical Ising chain.

Floquet-engineered protocols are designed to reproduce the counterdiabatic evolution through a stroboscopic driving schedule, which is derived analytically from an ansatz expression of the Adiabatic Gauge Potential. The variational ansatz is considered at the two leading orders $\ell = 1, 2$. Our analysis aimed to evaluate the yield of Floquet protocols in light of two complementary factors: the ability of the protocol to reproduce correctly the counterdiabatic dynamics, and the advantage of using the Floquet protocol over an unoptimized annealing schedule. Our investigation is supported by extensive benchmarks of the protocols, realized through numerical annealing simulations on paradigmatic instances of the Lechner-Hauke-Zoller model.

Firstly, we have performed a fine-tuning of the Floquet protocols hyperparameters: the driving frequency ω and the spectral reference frequency ω_0 . We have successfully identified the configurations of (ω_0, ω) that allow us to reproduce accurately the counterdiabatic evolution. We have linked the pattern of optimal parameters (ω_0, ω) to theoretical estimates for the error of the approximations involved in the Floquet protocols' construction. In particular, we have discovered that the error of the Magnus expansion - proportional to the integral norm in each driving cycle of the stroboscopic protocol - is minimized in correspondence with the optimal (ω_0, ω) . This criterion is suitable for practical applications since the integral of the norm can be minimized analytically without needing access to (expensive) simulations of the annealing schedule.

Secondly, we have used a notion of ‘cost’ of the annealing protocol based on the magnitude of its schedule coefficients: the *effective annealing time* τ_{eff} . We have explained why the (real) annealing time τ of the protocol is not a faithful indicator of the algorithm cost. The effective annealing time was functional to the definition of a new metric, the gain, that quantifies the advantage of the Floquet protocol. The *gain* compares the energy of the evolved state between unoptimized schedules and Floquet protocols at equal cost τ_{eff} . We found that the gain points overall towards a negative result. In a broad window of parameters, the unoptimized annealing schedules approximate better the ground state at an equivalent cost. There exists a narrow regime of driving frequencies ω which shows a slightly positive gain. That hyper-parameter regime, however, is outside the scope of approximations that have been used to derive the Floquet protocols.

In the final part of this work, we have tried to amend the criticalities of the reference Floquet protocols of Ref. [15]. For this purpose, we have derived a family of protocols that pushes the Magnus expansion up to the third order. One instance of these new protocols has shown some noticeable improvements over the reference work. The region of optimal convergence (ω_0, ω) is extended towards lower spectral reference frequencies ω_0 . In the case of $\ell = 2$ protocols, this region is significantly extended also towards lower driving frequencies ω , which was a strong limiting factor of the original protocols. Nevertheless, the negative verdict reached in the gain-based analysis is unaltered.

In conclusion, we have shown that Floquet protocols achieve a limited practical advantage w.r.t. unoptimized annealing schedules. Our results strongly rely on numerical simulations, since our analysis framework, based on the gain metric, requires access to the energy of the evolved state. To avoid selection bias, we have validated the conclusion mentioned above on two more instances of the LHZ model (with different ‘hardness’), on a new LHZ instance with 10 physical qubits, and an Ising model as well.

Overall, our conclusions point to a more realistic perspective on counterdiabatic annealing protocols. The compelling reason to investigate Floquet protocols was that the set of required system controls is not expanded, which is indeed a strong limiting factor in the implementation of counterdiabatic methods. Nonetheless, it seems reasonable to relax this assumption and allow the usage of a small set of additional controls. Under these considerations, in the final part of this work, we have explored an approximation of the Adiabatic Gauge Potential that uses only quasi-local operators. These methods are known

in the literature as Local Counterdiabatic Driving.

We have discussed a numerically optimized LCD schedule that employs only local field controls. The optimized protocol is developed upon the method introduced in Ref. [9], where the authors combine LCD with a simplified version of the Chopped RAndom-Basis algorithm [24]. In particular, we modify the LCD schedule with parametrized pulse functions; but instead of optimizing the pulses to maximize the fidelity of the prepared state, we minimize the contributions from high-order contributions of the Adiabatic Gauge Potential. Therefore, our method does not require access to numerical simulation of the quantum system, like in some other Quantum Optimal Control approaches.

We show that the minimization of the non-homogeneous 4-body contributions of the gauge potential favors the control by local transverse fields and improves the prepared ground state in the Parity architecture. The analysis of the efficiency of this protocol (based on a similar formulation of cost as the Floquet protocols) leads to a positive gain.

We have concluded that our optimized LCD schedule can be advantageous w.r.t. unoptimized protocols, provided that one can pay the price of a slightly extended set of controls. For this reason, we recommend further exploration of counterdiabatic-aided approaches.

A

AGP expressions for the Parity architecture

In this Appendix, we optimize the coefficients $\underline{\alpha}$ of the nested-commutator AGP ansatz applied to the annealing of the Parity model.

The derivation consists of the following steps.

- Consider an annealing schedule for the Parity model of Eq. (1.8):

$$\mathcal{H}(\lambda) = A(\lambda) \sum_{k=1}^K \hat{\sigma}_x^{(k)} + B(\lambda) \left(\sum_{k=1}^K J_k \hat{\sigma}_z^{(k)} + \sum_{l=1}^{K-N+1} C_l \hat{\rho}_z^{[l]} \right)$$

- Compute the ansatz $\mathcal{A}^{(\ell)}$ at counterdiabatic order ℓ ,

$$\mathcal{A}^{(\ell)} \equiv i \sum_{k=1}^{\ell} \alpha_k \underbrace{[\mathcal{H}, [\mathcal{H}, \dots [\mathcal{H}, \partial_\lambda \mathcal{H}]]]}_{2k-1} \quad ,$$

which is usually the most time-consuming step of the computation because of the $2\ell - 1$ nested commutators that have to be developed.

- Compute the Hermitian operator G defined in Eq. (2.19),

$$G = \partial_\lambda \mathcal{H} - i[\mathcal{H}, \mathcal{A}^{(\ell)}] \quad ,$$

and minimize its trace square w.r.t. $\underline{\alpha}$, i.e. solve the equation

$$\frac{\partial \text{Tr}[G^2]}{\partial \alpha_i} = 0 \quad \forall i = 1, \dots, \ell \quad .$$

Throughout this Appendix, the dotted variables \dot{A} and \dot{B} denote the deriva-

tives of the functions w.r.t. λ , for instance $\dot{A} = \partial_\lambda A(\lambda)$. Furthermore, we adopt the notation of Ref. [40] to indicate the many-body interactions on each constraint of the Parity model:

$${}^1_4 \square_l {}^2_3 = \hat{\sigma}_1^{(a)} \otimes \hat{\sigma}_2^{(b)} \otimes \hat{\sigma}_3^{(c)} \otimes \hat{\sigma}_4^{(d)} \quad (\text{A.1})$$

where a, b, c, d are the qubits on which the constraint l is acting, and the pedices $1, 2, 3, 4 \in \{x, y, z\}$. We extend the notation introducing several new elements:

- The variable n_l counts the number of physical qubits of each constraint l .

$$\begin{array}{c} a & & b \\ & \blacksquare & \\ d & & c \end{array} : n_l = 4 \qquad \begin{array}{c} & a & \\ & \blacktriangle & \\ c & & b \end{array} : n_l = 3$$

- When l is referenced in a sum symbol, i.e. like in $\sum_{k \in \{l\}}$, the index k runs on the qubits $\{a, b, c, d\}$ affected by the constraint l .
- The scripture $\mathcal{N}(l)$ returns the indices of the “neighboring” constraints that share at least a qubit with the constraint l . For example, consider the Parity instance with 6 constraints shown on the left. We report the values of \mathcal{N} for three of those constraints.

$$\begin{array}{ccc} \circ & & \circ \\ \circ & \blacksquare_1 & \circ \\ \circ & & \circ \\ \circ & & \circ \end{array} \quad \begin{array}{ccc} \circ & & \circ \\ \circ & \blacktriangle_2 & \circ \\ \circ & & \circ \\ \circ & & \circ \end{array} \quad \begin{array}{ccc} \circ & & \circ \\ \circ & \blacktriangle_3 & \circ \\ \circ & & \circ \\ \circ & & \circ \end{array} \quad \begin{array}{c} \mathcal{N}(1) = \{2, 5, 6\} \\ \mathcal{N}(3) = \{2, 4\} \\ \mathcal{N}(5) = \{1, 2, 4, 6\} \end{array}$$

- The quantity $I_{n,l}$ counts the number of physical qubits shared between the constraints l and n . For instance, consider again the Parity instance of the above figure. Then, $I_{1,2} = I_{2,6} = 1$, and $I_{1,6} = I_{2,5} = I_{4,5} = 2$.
- The quantity χ_k is defined by taking into consideration the constraints that act on the k^{th} qubit, summing the squared constraints coefficients. For instance:

$$\begin{array}{ccc} & \textcircled{1} & \\ & \blacktriangle_{C_1} & \\ \textcircled{2} & & \textcircled{3} \\ \textcircled{4} & \blacktriangle_{C_2} & \textcircled{5} & \blacktriangle_{C_3} & \textcircled{6} \end{array} \quad \begin{array}{c} \chi_1 = C_1^2 \\ \chi_2 = C_1^2 + C_2^2 \\ \chi_5 = C_1^2 + C_2^2 + C_3^2 \end{array}$$

- We introduce new constraint patterns involving more than 4 qubits. The following pattern involves the qubits of two neighboring constraints l and m :

$${}^2_1 \square_l {}^3_6 \square_m {}^4_5 = \hat{\sigma}_1^{(a)} \otimes \hat{\sigma}_2^{(b)} \otimes \hat{\sigma}_3^{(c)} \otimes \hat{\sigma}_4^{(d)} \otimes \hat{\sigma}_5^{(e)} \otimes \hat{\sigma}_6^{(f)}$$

We observe that the two constraints above share the qubits 3 and 6. If two neighboring constants share only one qubit, the qubit 6 is to be neglected in the formula: ${}^2_1\boxed{l}^3_3\boxed{m}^4_5 = \hat{\sigma}_1^{(a)} \otimes \hat{\sigma}_2^{(b)} \otimes \hat{\sigma}_3^{(c)} \otimes \hat{\sigma}_4^{(d)} \otimes \hat{\sigma}_5^{(e)}$.

- The letters a, b, c, d, e, f are reserved to index the qubits in each constraint. For example, $J_{l[a]}^x \boxed{l}_z^z$ means that the constraint ${}^x_z\boxed{l}_z^z$ is multiplied by the element of the vector \underline{J} associated to the qubit (a) of constraint $[l]$.

A.1 First order AGP coefficient

Consider the AGP ansatz at order $\ell = 1$. The expression of $\mathcal{A}^{(1)}$ is therefore

$$\begin{aligned} \mathcal{A}^{(1)} &= i\alpha_1[\mathcal{H}, \partial_\lambda \mathcal{H}] = \\ &= 2\alpha_1(A\dot{B} - B\dot{A}) \left(\sum_i J_i \hat{\sigma}_y^{(i)} + \sum_l C_l \left({}^y_z\boxed{l}_z^z + {}^z_l\boxed{l}_z^y + {}^z_l\boxed{l}_y^z + {}^z_y\boxed{l}_z^z \right) \right) \end{aligned} \quad (\text{A.2})$$

To compute G , we need to commute the ansatz of Eq. (A.2) with \mathcal{H} :

$$\begin{aligned} [\mathcal{H}, \mathcal{A}^{(1)}] &= 4i\alpha_1(A\dot{B} - B\dot{A}) \left(A \sum_k J_k \hat{\sigma}_z^{(k)} + A \sum_l C_l n_l {}^z_l\boxed{l}_z^z - B \sum_k J_k^2 \hat{\sigma}_x^{(k)} \right. \\ &\quad - A \sum_l C_l \left[\left({}^y_z\boxed{l}_z^y + {}^y_l\boxed{l}_y^z + {}^y_y\boxed{l}_z^z \right) + \left({}^y_z\boxed{l}_z^y + {}^z_l\boxed{l}_y^y + {}^z_y\boxed{l}_z^y \right) + \dots \right] \\ &\quad - 2B \sum_l C_l \left(J_{l[a]}^x {}^x_z\boxed{l}_z^z + J_{l[b]}^z {}^z_l\boxed{l}_z^x + J_{l[c]}^z {}^z_l\boxed{l}_x^z + J_{l[d]}^z {}^z_x\boxed{l}_z^z \right) \\ &\quad \left. - B \sum_l C_l^2 \sum_{k \in \{l\}} \hat{\sigma}_x^{(k)} - B \sum_l \sum_{m \in \mathcal{N}(l)} C_l C_m \left({}^z_l\boxed{l}^x \boxed{m}_z^z + {}^z_l\boxed{l}_x \boxed{m}_z^z \right) \right) \end{aligned} \quad (\text{A.3})$$

The trace squared of G reads as follows:

$$\begin{aligned} \text{Tr}[G^2] &= \sum_k \left(\dot{A} - 2B\tilde{\alpha}_1 \chi_k - 2B\tilde{\alpha}_1 J_k^2 \right)^2 + \sum_k (\dot{B} + 2A\tilde{\alpha}_1)^2 J_k^2 + \\ &\quad + \sum_l C_l^2 (\dot{B} + 2A\tilde{\alpha}_1 n_l)^2 + 4\tilde{\alpha}_1^2 \left(2A^2 \sum_l C_l^2 n_l (n_l - 1) + \right. \\ &\quad \left. + 4B^2 \sum_l C_l^2 \sum_{i \in [l]} J_i^2 + 2B^2 \sum_l \sum_{m \in \mathcal{N}(l)} (C_l C_m)^2 I_{l,m} \right) \end{aligned}$$

where $\tilde{\alpha}_1 = 2\alpha_1(A\dot{B} - B\dot{A})$ for convenience. Eventually, the minimization w.r.t. α_1 returns

$$\alpha_1 = -\frac{1}{4} \frac{\sum_l C_l^2 n_l + \sum_k J_k^2}{c_1 A^2 + c_2 B^2} \quad (\text{A.4})$$

where

$$c_1 = \sum_l C_l^2 n_l^2 + \sum_k J_k^2 + 2 \sum_l C_l^2 n_l (n_l - 1) \quad ,$$

$$c_2 = \sum_k (\chi_k^2 + J_k^2)^2 + \sum_l \sum_{i \in [l]} C_l^2 J_i^2 + 2 \sum_l \sum_{m \in \mathcal{N}(l)} (C_l C_m)^2 I_{l,m} \quad .$$

We note down that if two constraints l and m share only one qubit, the sum $\left(\begin{smallmatrix} z \\ z \end{smallmatrix} \overline{l}^x \overline{m}^z + \begin{smallmatrix} z \\ z \end{smallmatrix} \overline{l}^x \overline{m}^z \right)$ of Eq. (A.3) simplifies to $\begin{smallmatrix} z \\ z \end{smallmatrix} \overline{l}^x \overline{m}^z$. A similar convention is used in all the following calculations.

A.2 Second order AGP controls

Consider the AGP ansatz at order $\ell = 2$. Given the complexity of the calculation, we will not derive analytically the expression of α_2 . Instead, we have optimized it numerically, since the instances that we simulate involve a small number of qubits. In this section, however, we wish to report the calculations up to the AGP expression, without optimizing the coefficients α_1 and α_2 . Our goal is to show which additional interaction patterns have to be controlled in a $\ell = 2$ CD protocol.

$$\begin{aligned} \mathcal{A}^{(2)} &= \underbrace{i\alpha_1 [\mathcal{H}, \partial_\lambda \mathcal{H}] + i\alpha_2 [\mathcal{H}, [\mathcal{H}, [\mathcal{H}, \partial_\lambda \mathcal{H}]]]}_{\text{see } \mathcal{A}^{(1)}} = \\ &= i\alpha_1 [\mathcal{H}, \partial_\lambda \mathcal{H}] + 8\alpha_2 (A\dot{B} - \dot{A}B) \left(\textcircled{1} + \textcircled{2} + \textcircled{3} \right) \end{aligned} \quad (\text{A.5})$$

$$\begin{aligned}
\textcircled{1} &= \frac{i}{2} \left[A \sum_k \hat{\sigma}_x^{(k)}, [\mathcal{H}, [\mathcal{H}, \partial_\lambda \mathcal{H}]] \right] = \\
&= -A^2 \sum_k J_k \hat{\sigma}_y^{(k)} - A^2 \sum_l C_l (3n_l - 2) \left({}^y \square_l^z + {}^z \square_l^y + {}^z \square_l^z + {}^z \square_y^z \right) \\
&\quad + 2AB \sum_l C_l \left[J_{l[a]} \left({}^x \square_l^y + {}^x \square_l^z + {}^x \square_y^z \right) + \right. \\
&\quad \quad \quad \left. + J_{l[b]} \left({}^y \square_l^x + {}^z \square_l^x + {}^z \square_l^y \right) + \dots \right] \\
&\quad + AB \sum_l \sum_{m \in \mathcal{N}(l)} C_l C_m \left({}^y \square_l^x \square_m^z + {}^z \square_l^x \square_m^y + {}^z \square_l^x \square_m^z + {}^z \square_y^x \square_m^z \right)
\end{aligned}$$

$$\begin{aligned}
\textcircled{2} &= \frac{i}{2} \left[B \sum_k J_k \hat{\sigma}_x^{(k)}, [\mathcal{H}, [\mathcal{H}, \partial_\lambda \mathcal{H}]] \right] = \\
&= -B^2 \sum_k J_k^3 \hat{\sigma}_y^{(k)} - B^2 \sum_l C_l^2 \sum_{i \in \{l\}} J_i \hat{\sigma}_y^{(i)} \\
&\quad - 2B^2 \sum_l C_l \left(J_a^2 {}^y \square_l^z + J_{l[b]}^2 {}^z \square_l^y + J_{l[c]}^2 {}^z \square_l^z + J_{l[d]}^2 {}^z \square_y^z \right) \\
&\quad + 2AB \sum_l C_l \left[J_{l[a]} \left({}^x \square_l^y + {}^x \square_l^z + {}^x \square_y^z \right) + \right. \\
&\quad \quad \quad \left. + J_{l[b]} \left({}^y \square_l^x + {}^z \square_l^x + {}^z \square_l^y \right) + \dots \right] \\
&\quad - B^2 \sum_l \sum_{m \in \mathcal{N}(l)} C_l C_m \left({}^z \square_l^y \square_m^z + {}^z \square_l^y \square_m^z \right)
\end{aligned}$$

$$\begin{aligned}
\textcircled{3} &= \frac{i}{2} \left[B \sum_l C_l \hat{\rho}_z^{[l]}, [\mathcal{H}, [\mathcal{H}, \partial_\lambda \mathcal{H}]] \right] = \\
&= -B^2 \sum_l C_l^2 \sum_{i \in \{l\}} J_i \hat{\sigma}_y^{(i)} - B^2 \sum_l C_l \left(J_{l[a]}^2 {}^y \square_l^z + J_{l[b]}^2 {}^z \square_l^y + \dots \right) \\
&\quad - B^2 \sum_l \sum_{m \in \mathcal{N}(l)} C_l C_m \left(J_{l[b]} {}^z \square_l^y \square_m^z + J_{l[e]} {}^z \square_l^y \square_m^z \right) \\
&\quad - AB \sum_l \sum_{m \in \mathcal{N}(l)} C_l C_m \left({}^z \square_l^x \square_m^y + {}^z \square_l^x \square_m^z + {}^z \square_l^x \square_m^y + {}^z \square_l^x \square_m^z \right)
\end{aligned}$$

B

Higher order Floquet protocols

Using the general driving Hamiltonian

$$\mathcal{H}_{\text{FE}} = \left(1 + \frac{1}{\omega_0} \dot{f}\right) \mathcal{H} + \dot{\lambda} \beta(t) \cdot \partial_{\lambda} \mathcal{H}$$

we define the following protocols. These have been determined through a higher-order $\mathcal{O}(\delta^n)$ Magnus expansion of the effective driving in an oscillation period. The derivation is omitted.

At first order $\ell = 1$ of CD, we define the protocols:

- $\mathcal{O}2$, i.e. up to $\mathcal{O}(\delta^2)$:

$$\begin{cases} \beta(t) = 2\omega_0 \cos(\omega t) \alpha_1(t) \\ f(t) = \cos(\omega t) \end{cases} \quad (\text{B.1})$$

- $\mathcal{O}3$, i.e. up to $\mathcal{O}(\delta^3)$:

$$\begin{cases} \beta(t) = 2\omega_0 [\cos(\omega t) - 3 \cos(3\omega t)] \alpha_1(t) \\ \quad + \omega_0^{-1} \left(2 + \frac{2\pi}{\omega} \frac{\ddot{\lambda}(t)}{\dot{\lambda}(t)}\right)^{-1} [\cos(\omega t) - 3 \cos(3\omega t)] \\ f(t) = \cos(\omega t) \end{cases} \quad (\text{B.2})$$

- $\mathcal{O}3\text{L}$, i.e. up to $\mathcal{O}(\delta^3)$ without the high-frequency terms:

$$\begin{cases} \beta(t) = 2\omega_0 \cos(\omega t) \alpha_1(t) + \omega_0^{-1} \left(2 + \frac{2\pi}{\omega} \frac{\ddot{\lambda}(t)}{\dot{\lambda}(t)}\right)^{-1} \cos(\omega t) \\ f(t) = \cos(\omega t) \end{cases} \quad (\text{B.3})$$

At second order $\ell = 2$ of CD, we define the protocols:

- $\mathcal{O}2$, i.e. up to $\mathcal{O}(\delta^2)$:

$$\begin{cases} \beta(t) = 2\omega_0 [\cos(\omega t) - 3 \cos(3\omega t)] \alpha_1(t) - 48\omega_0^3 \cos(3\omega t) \alpha_2(t) \\ f(t) = \cos(\omega t) \end{cases} \quad (\text{B.4})$$

- $\mathcal{O}3$, i.e. up to $\mathcal{O}(\delta^3)$:

$$\begin{cases} \beta(t) = 2\omega_0 [\cos(\omega t) - 3 \cos(3\omega t)] \alpha_1(t) \\ \quad - 48\omega_0^3 [\cos(3\omega t) - 5 \cos(5\omega t)] \alpha_2(t) \\ \quad + \omega_0^{-1} \left(2 + \frac{2\pi}{\omega} \frac{\ddot{\lambda}(t)}{\dot{\lambda}(t)} \right)^{-1} [\cos(\omega t) - 3 \cos(3\omega t) + 5 \cos(5\omega t)] \\ f(t) = \cos(\omega t) \end{cases} \quad (\text{B.5})$$

- $\mathcal{O}3L$, i.e. up to $\mathcal{O}(\delta^3)$ without the high-frequency terms:

$$\begin{cases} \beta(t) = 2\omega_0 [\cos(\omega t) - 3 \cos(3\omega t)] \alpha_1(t) - 48\omega_0^3 [\cos(3\omega t)] \alpha_2(t) \\ \quad + \omega_0^{-1} \left(2 + \frac{2\pi}{\omega} \frac{\ddot{\lambda}(t)}{\dot{\lambda}(t)} \right)^{-1} [\cos(\omega t) - 3 \cos(3\omega t)] \\ f(t) = \cos(\omega t) \end{cases} \quad (\text{B.6})$$

Complement of simulations

Throughout the main text, we have frequently shown plots of $\mathcal{M} - \mathcal{M}_{\text{CD}}$ that evaluate the accuracy of the protocol; the plots of the gain \mathcal{G} assess the efficiency of the protocol. This Appendix will report several additional plots that have been obtained by changing the annealing problem instance for various protocols: the ‘Claeys’ protocol of Ref. [15], the $\mathcal{O}3$ and $\mathcal{O}3\text{L}$ protocols of Chapter 4.

In the following sections, we consider the LHZ models of Table 3.2 and the Ising model of Eq. (4.9). A summary of the Figures is provided in Table C.1.

	$\ell = 1$		$\ell = 2$	
	$\mathcal{M} - \mathcal{M}_{\text{CD}}$	\mathcal{G}	$\mathcal{M} - \mathcal{M}_{\text{CD}}$	\mathcal{G}
LHZ6M	shown in main text		Figure C.1	Figure C.2
LHZ6E	Figure C.3	Figure C.4	-	-
LHZ6H	Figure C.5	Figure C.6	-	-
LHZ10	Figure C.7	Figure C.8	Figure C.9	Figure C.10
Ising ($N = 6$)	shown in main text		Figure C.11	Figure C.12

Table C.1: Recap of the Figures shown in this Appendix for various LHZ and Ising models.

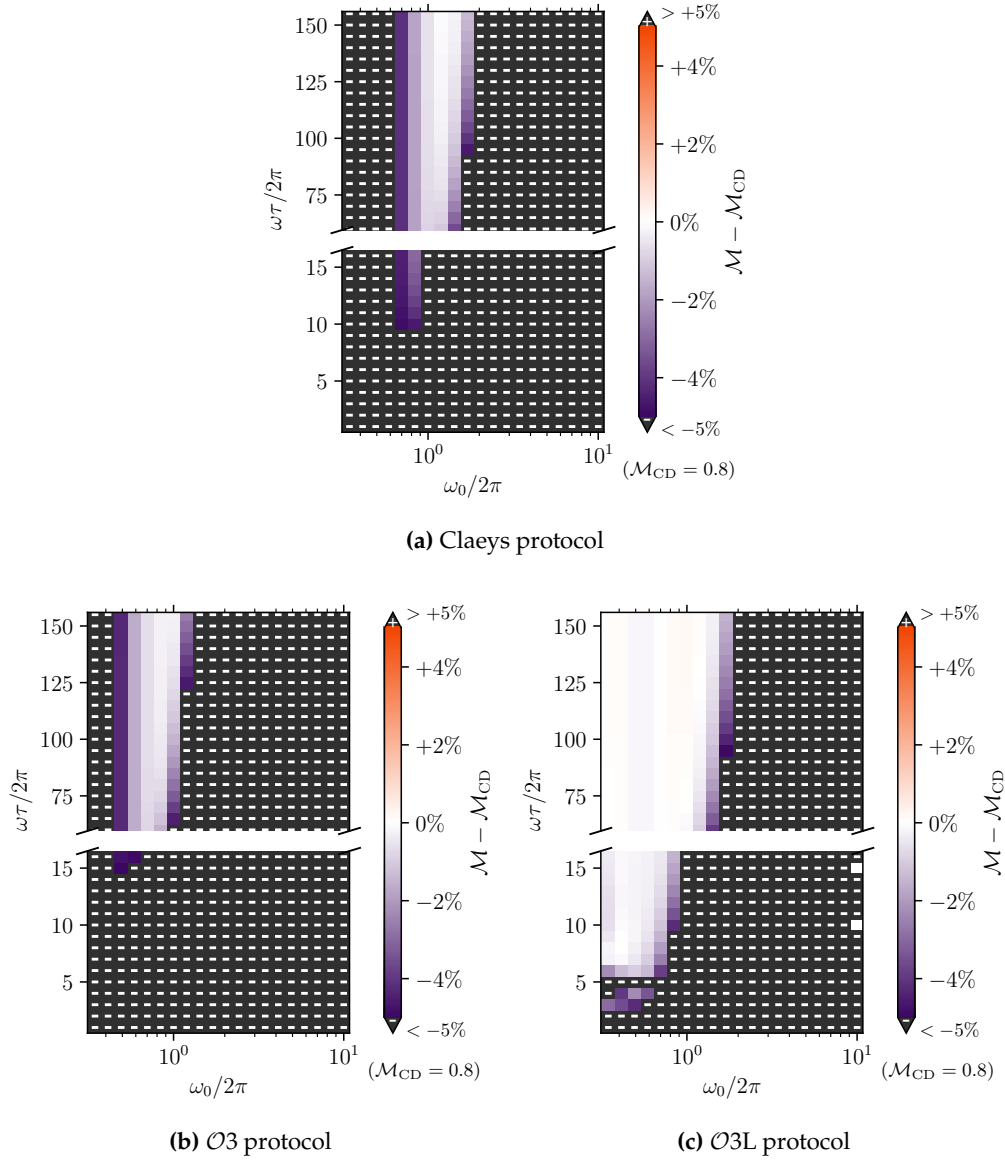
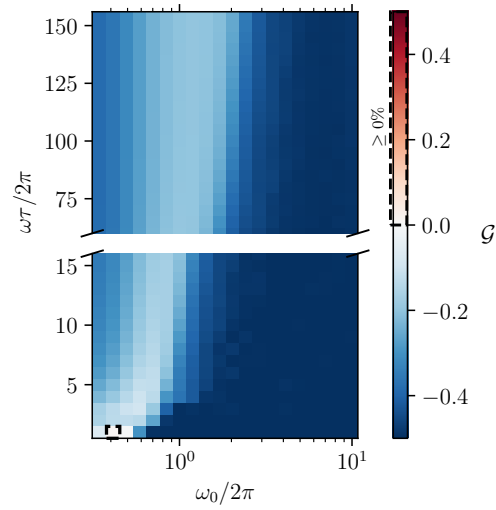


Figure C.1: LHZ6M instance: difference in energy merit for the new $\mathcal{O}(\delta^3)$ FE protocols w.r.t. the reference $\ell = 2$ CD protocol. Figure (b) is realized by using the $\mathcal{O}3$ protocol, whereas the $\mathcal{O}3L$ protocol is used in Figure (c). For comparison, Figure (a) shows the results of the Claeys protocol. The annealing schedule is simulated with $\tau = 0.1$. The points that exceed the $\pm 5\%$ interval around zero are cropped to a grey color and marked with a $+/-$ symbol to indicate the sign of the extreme. The vertical axis fixes the total number of driving cycles of the protocol ($\omega\tau/2\pi$); the horizontal axis denotes the spectral reference frequency $\omega_0/2\pi$.



(a) Claeys protocol

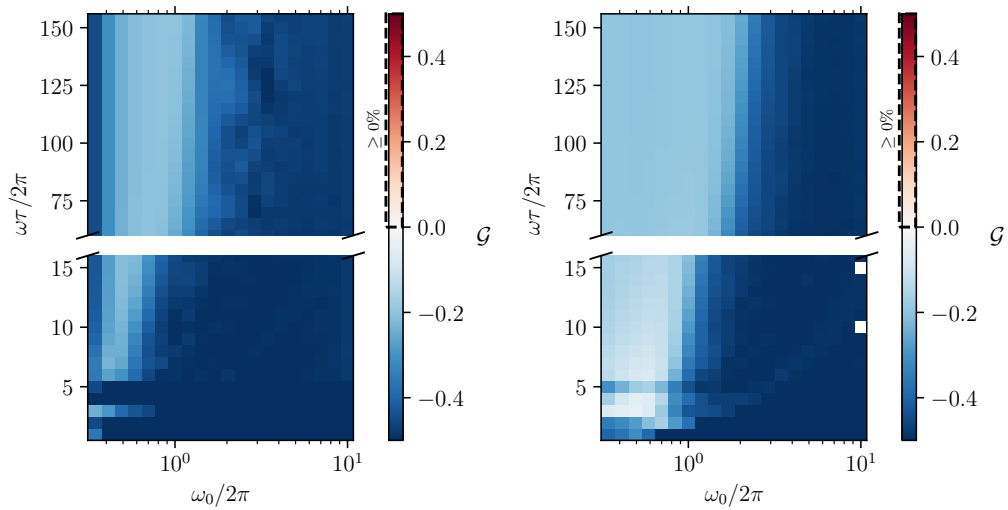
(b) $\mathcal{O}3$ protocol(c) $\mathcal{O}3L$ protocol

Figure C.2: LHZ6M instance: gain for the new $\ell = 2$, $\mathcal{O}(\delta^3)$ FE protocols. Figure (b) is realized by using the $\mathcal{O}3$ protocol, whereas the $\mathcal{O}3L$ protocol is used in Figure (c). For comparison, Figure (a) shows the results of the Claeys protocol. The annealing is simulated with $\tau = 0.1$. The vertical axis fixes the total number of driving cycles of the protocol ($\omega\tau/2\pi$); the horizontal axis denotes the spectral reference frequency $\omega_0/2\pi$.

C.1 LHZ6E and LHZ6H

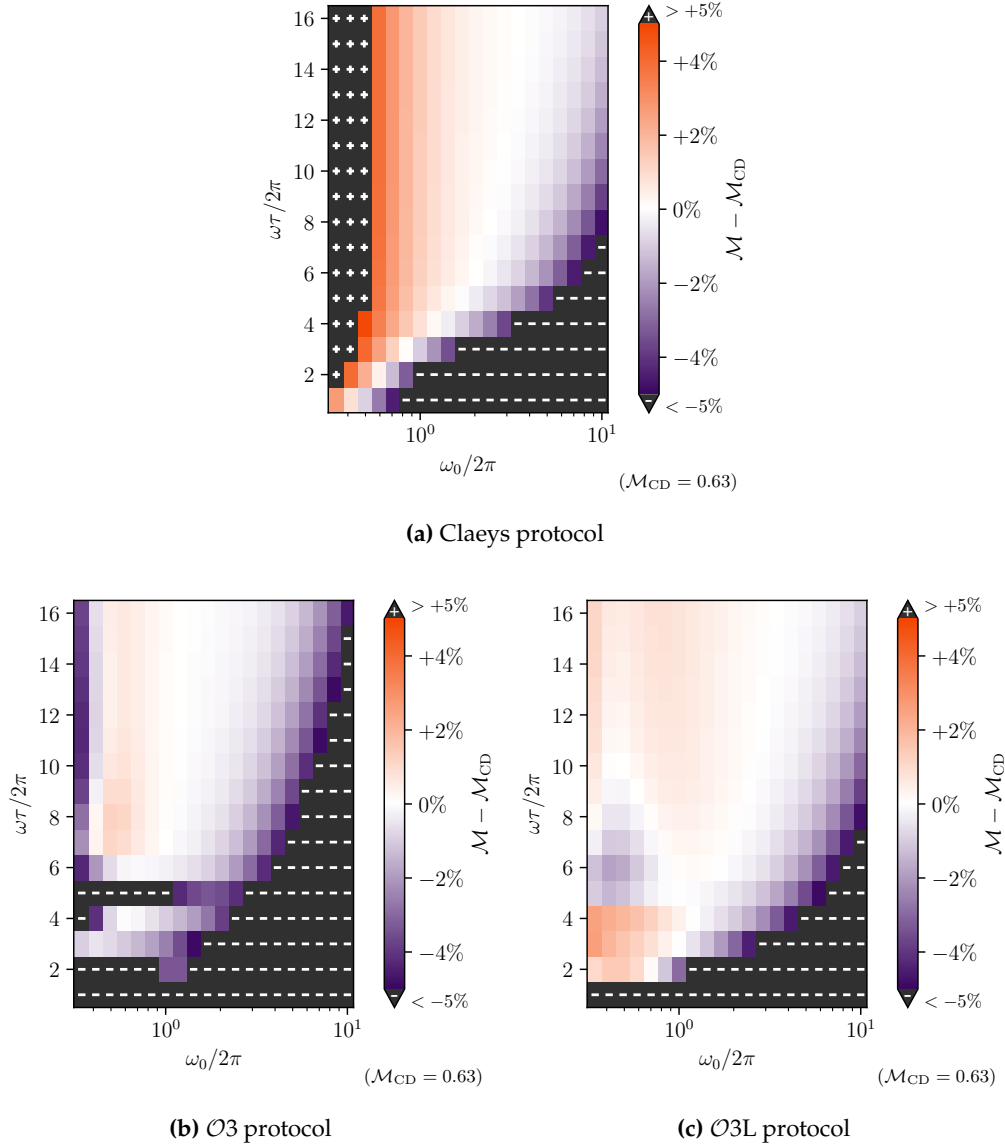
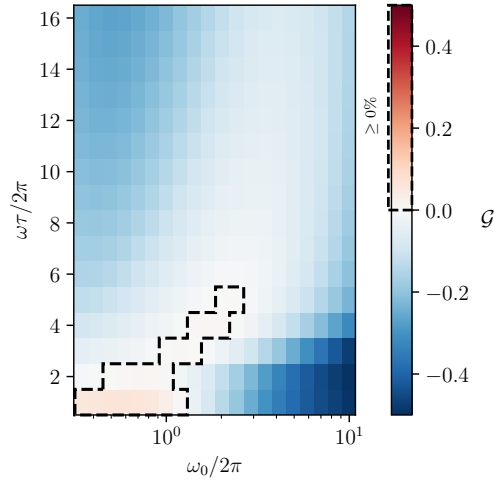


Figure C.3: LHZ6E instance: difference in energy merit for the new $\mathcal{O}(\delta^3)$ FE protocols w.r.t. the reference $\ell = 1$ CD protocol. Figure (b) is realized by using the $\mathcal{O}3$ protocol, whereas the $\mathcal{O}3L$ protocol is used in Figure (c). For comparison, Figure (a) shows the results of the Claeys protocol. The annealing schedule is simulated with $\tau = 0.1$. The points that exceed the $\pm 5\%$ interval around zero are cropped to a grey color and marked with a $+/-$ symbol to indicate the sign of the extreme. The vertical axis fixes the total number of driving cycles of the protocol ($\omega\tau/2\pi$); the horizontal axis denotes the spectral reference frequency $\omega_0/2\pi$.



(a) Claeys protocol

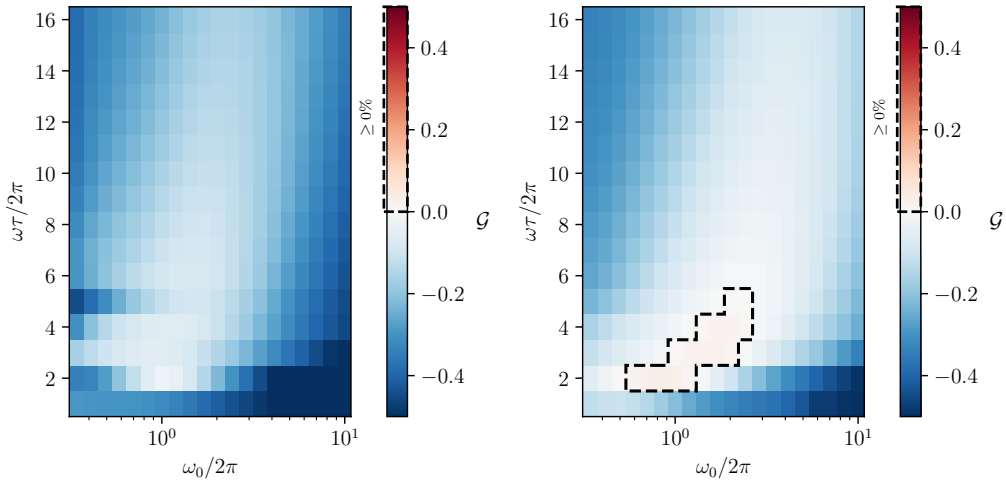
(b) $\mathcal{O}3$ protocol(c) $\mathcal{O}3L$ protocol

Figure C.4: LHZ6E instance: gain for the new $\ell = 1$, $\mathcal{O}(\delta^3)$ FE protocols. Figure (b) is realized by using the $\mathcal{O}3$ protocol, whereas the $\mathcal{O}3L$ protocol is used in Figure (c). For comparison, Figure (a) shows the results of the Claeys protocol. The annealing is simulated with $\tau = 0.1$. The vertical axis fixes the total number of driving cycles of the protocol ($\omega\tau/2\pi$); the horizontal axis denotes the spectral reference frequency $\omega_0/2\pi$.

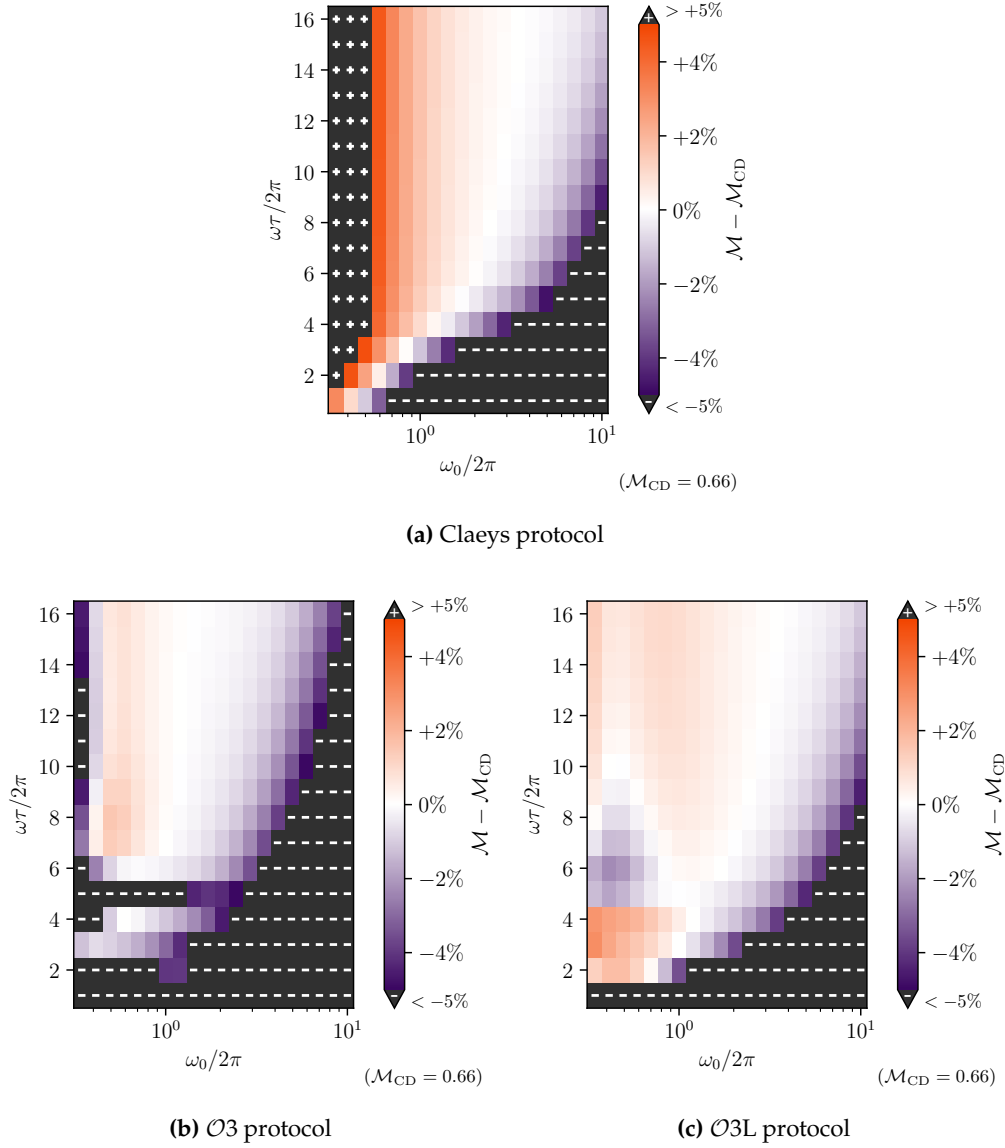
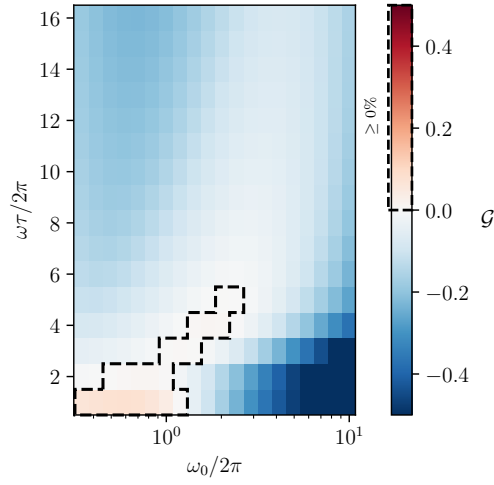


Figure C.5: LHZ6H instance: difference in energy merit for the new $\mathcal{O}(\delta^3)$ FE protocols w.r.t. the reference $\ell = 1$ CD protocol. Figure (b) is realized by using the $\mathcal{O}3$ protocol, whereas the $\mathcal{O}3L$ protocol is used in Figure (c). For comparison, Figure (a) shows the results of the Claeys protocol. The annealing schedule is simulated with $\tau = 0.1$. The points that exceed the $\pm 5\%$ interval around zero are cropped to a grey color and marked with a $+/-$ symbol to indicate the sign of the extreme. The vertical axis fixes the total number of driving cycles of the protocol ($\omega\tau/2\pi$); the horizontal axis denotes the spectral reference frequency $\omega_0/2\pi$.



(a) Claeys protocol

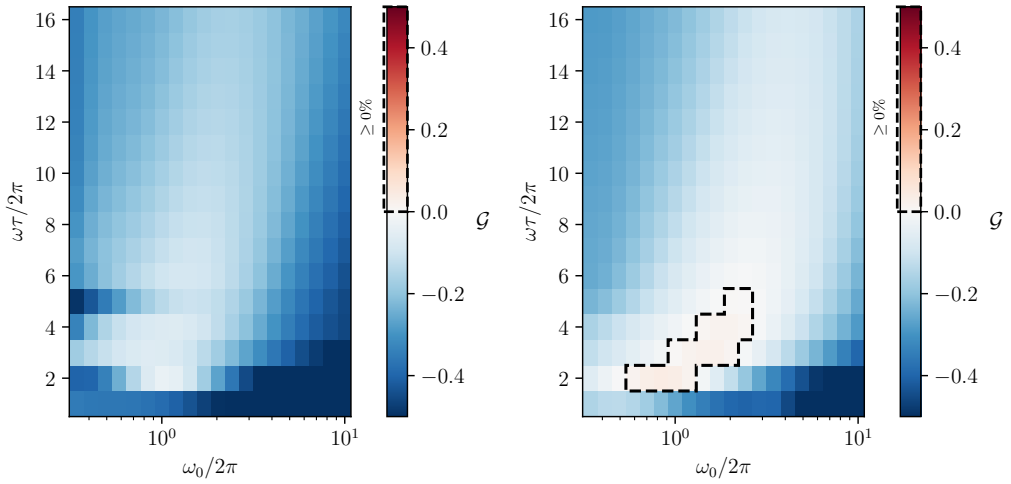
(b) $\mathcal{O}3$ protocol(c) $\mathcal{O}3L$ protocol

Figure C.6: LHZ6H instance: gain for the new $\ell = 1$, $\mathcal{O}(\delta^3)$ FE protocols. Figure (b) is realized by using the $\mathcal{O}3$ protocol, whereas the $\mathcal{O}3L$ protocol is used in Figure (c). For comparison, Figure (a) shows the results of the Claeys protocol. The annealing is simulated with $\tau = 0.1$. The vertical axis fixes the total number of driving cycles of the protocol ($\omega\tau/2\pi$); the horizontal axis denotes the spectral reference frequency $\omega_0/2\pi$.

C.2 LHZ10

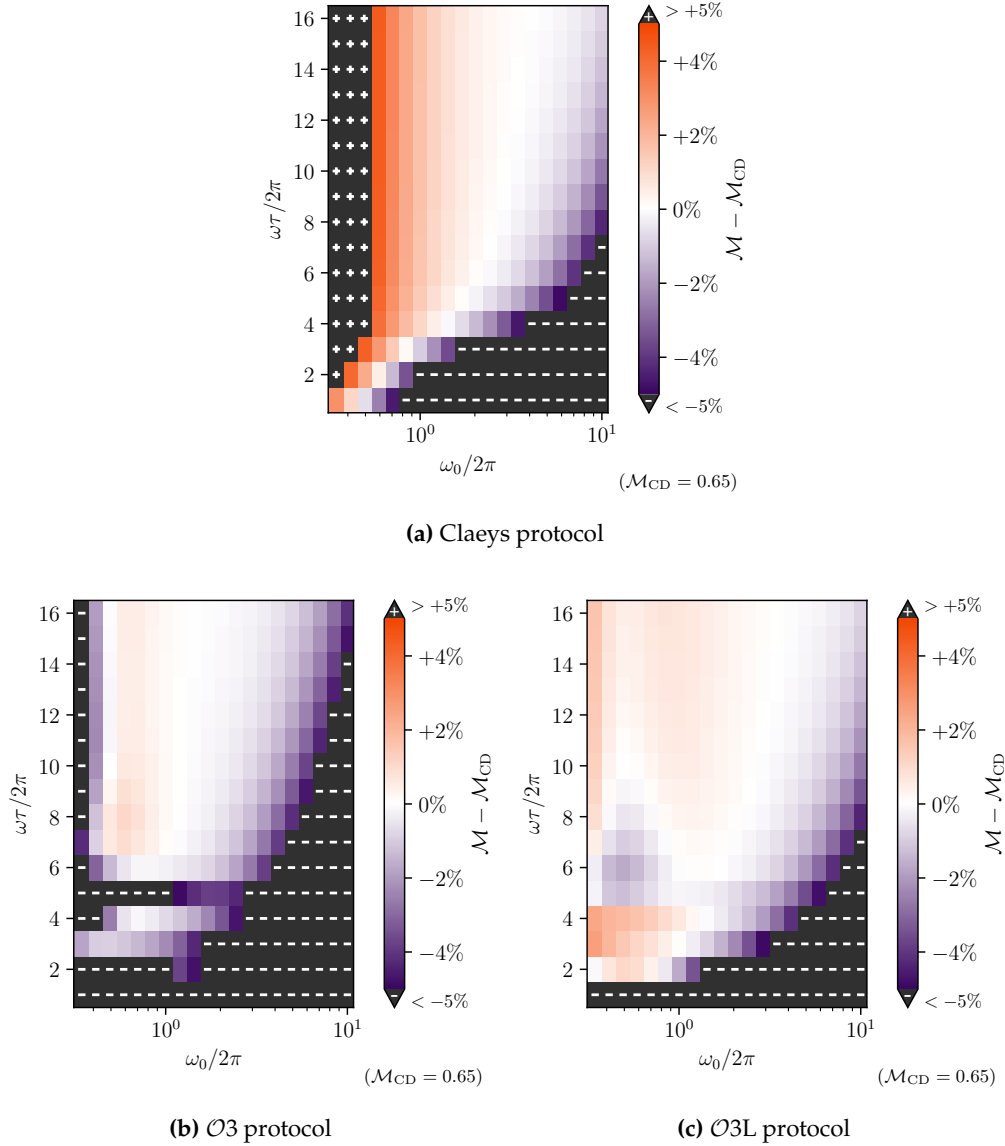
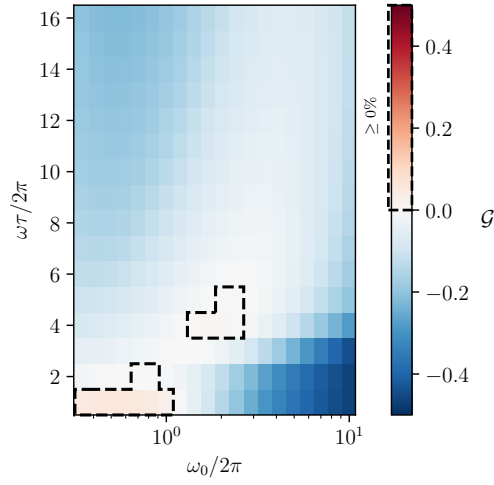


Figure C.7: LHZ10 instance: difference in energy merit for the new $\mathcal{O}(\delta^3)$ FE protocols w.r.t. the reference $\ell = 1$ CD protocol. Figure (b) is realized by using the $\mathcal{O}3$ protocol, whereas the $\mathcal{O}3L$ protocol is used in Figure (c). For comparison, Figure (a) shows the results of the Claeys protocol. The annealing schedule is simulated with $\tau = 0.1$. The points that exceed the $\pm 5\%$ interval around zero are cropped to a grey color and marked with a $+/-$ symbol to indicate the sign of the extreme. The vertical axis fixes the total number of driving cycles of the protocol ($\omega\tau/2\pi$); the horizontal axis denotes the spectral reference frequency $\omega_0/2\pi$.



(a) Claeys protocol

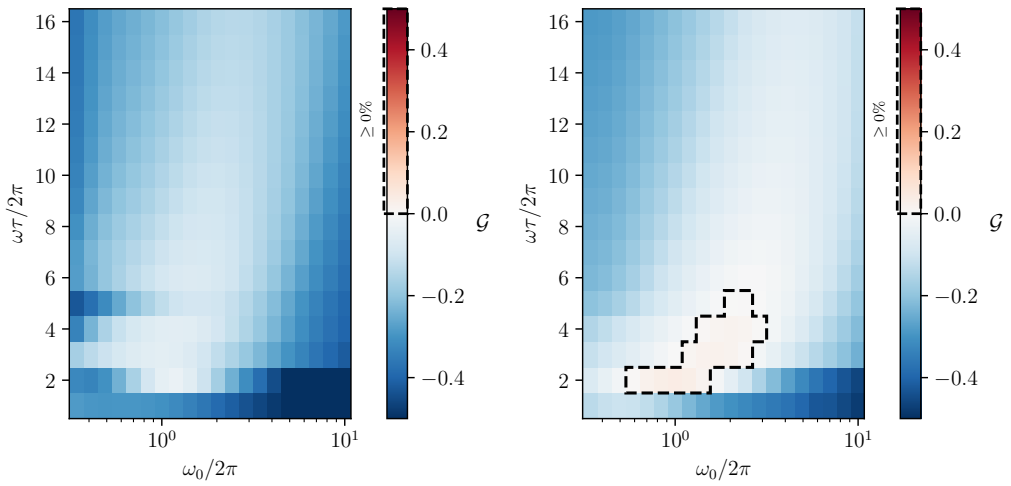
(b) $\mathcal{O}3$ protocol(c) $\mathcal{O}3L$ protocol

Figure C.8: LHZ10 instance: gain for the new $\ell = 1$, $\mathcal{O}(\delta^3)$ FE protocols. Figure (b) is realized by using the $\mathcal{O}3$ protocol, whereas the $\mathcal{O}3L$ protocol is used in Figure (c). For comparison, Figure (a) shows the results of the Claeys protocol. The annealing is simulated with $\tau = 0.1$. The vertical axis fixes the total number of driving cycles of the protocol ($\omega\tau/2\pi$); the horizontal axis denotes the spectral reference frequency $\omega_0/2\pi$.

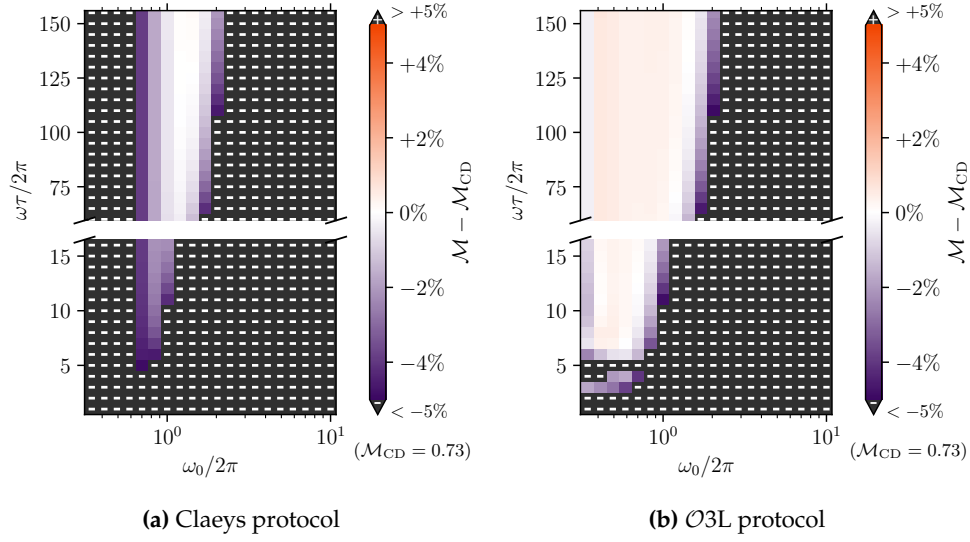


Figure C.9: LHZ10 instance: difference in energy merit for the new $\mathcal{O}(\delta^3)$ FE protocols w.r.t. the reference $\ell = 2$ CD protocol. For comparison, Figure (a) shows the results of the Claey's protocol. Figure (b) is realized by using the $\mathcal{O}3\text{L}$ protocol. The annealing schedule is simulated with $\tau = 0.1$. The points that exceed the $\pm 5\%$ interval around zero are cropped to a grey color and marked with a $+/-$ symbol to indicate the sign of the extreme. The vertical axis fixes the total number of driving cycles of the protocol ($\omega\tau/2\pi$); the horizontal axis denotes the spectral reference frequency $\omega_0/2\pi$.

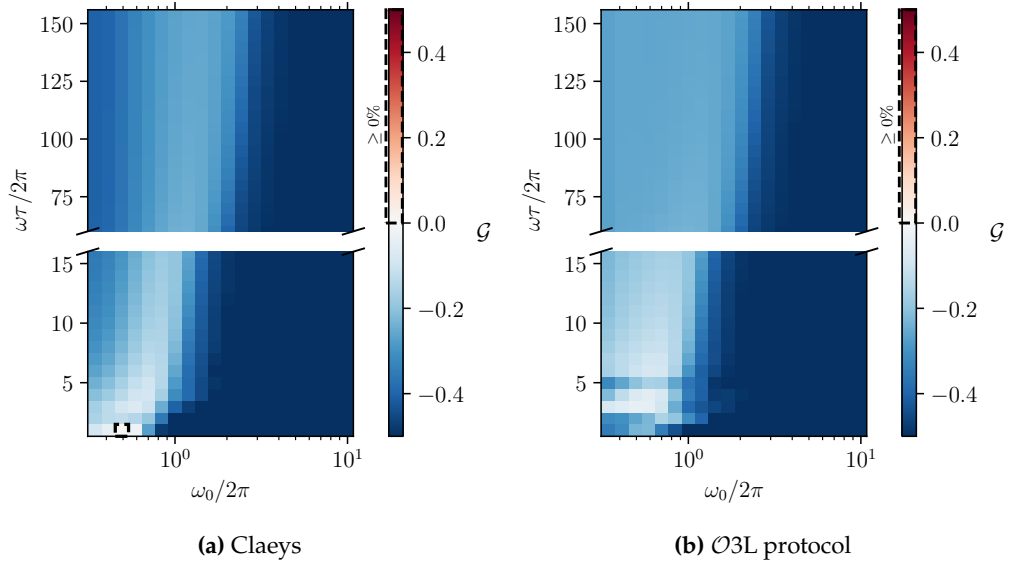


Figure C.10: LHZ10 instance: gain for the new $\ell = 2$, $\mathcal{O}(\delta^3)$ FE protocols. For comparison, Figure (a) shows the results of the Claeys protocol. Figure (b) is realized by using the O3L protocol. The annealing is simulated with $\tau = 0.1$. The vertical axis fixes the total number of driving cycles of the protocol ($\omega\tau/2\pi$); the horizontal axis denotes the spectral reference frequency $\omega_0/2\pi$.

C.3 Ising model

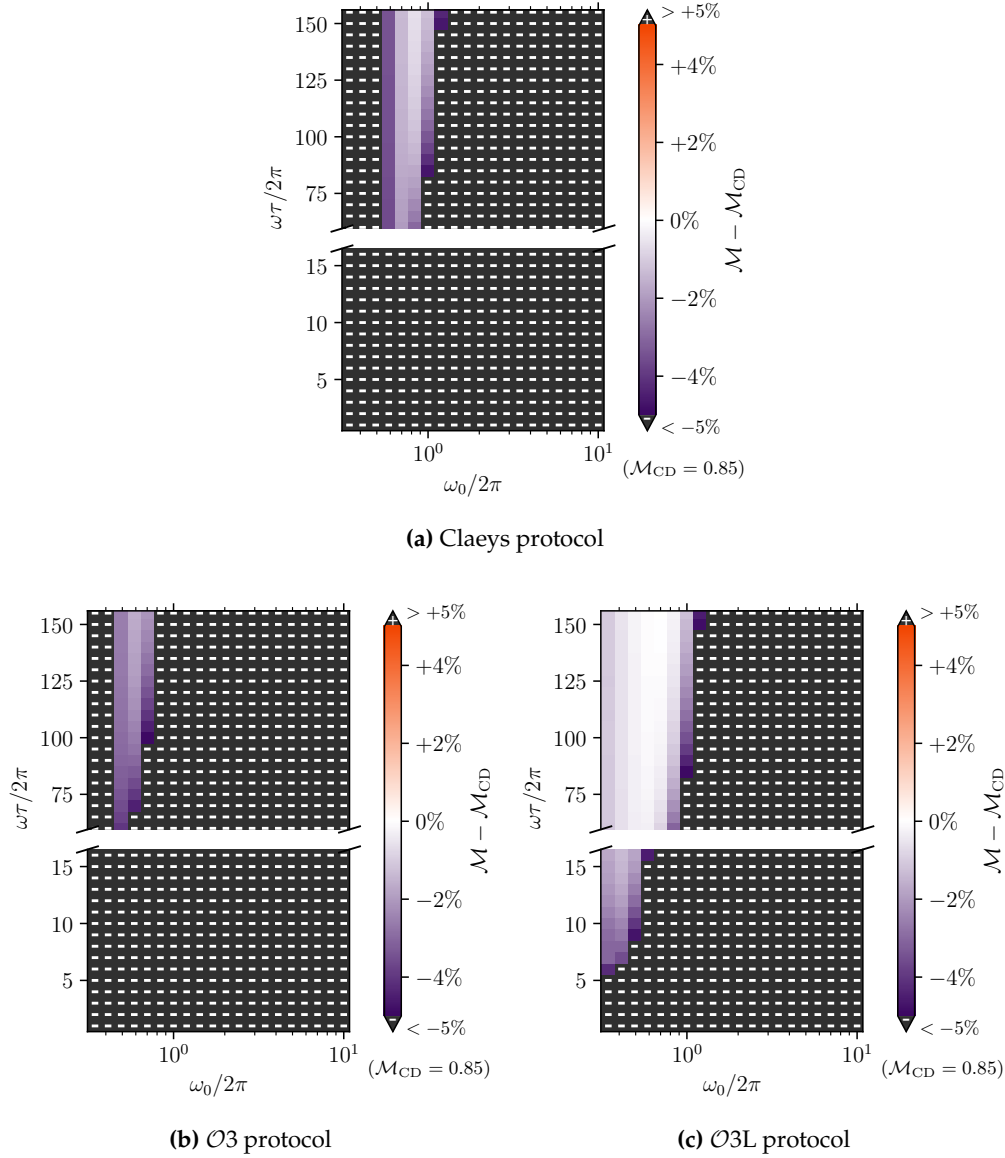
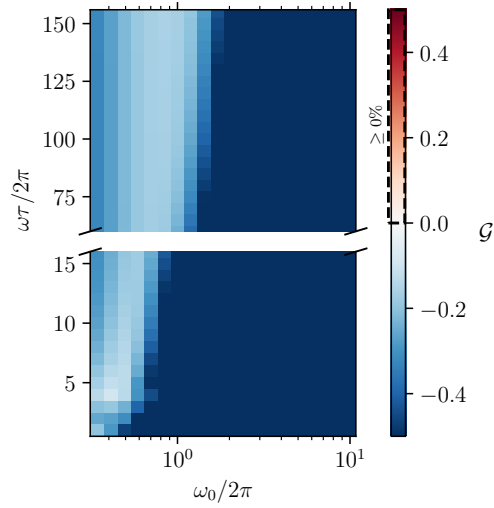


Figure C.11: Ising model instance with $N = 6$ qubits: difference in energy merit for the new $\mathcal{O}(\delta^3)$ FE protocols w.r.t. the reference $\ell = 2$ CD protocol. Figure (b) is realized by using the $\mathcal{O}3$ protocol, whereas the $\mathcal{O}3L$ protocol is used in Figure (c). For comparison, Figure (a) shows the results of the Claeys protocol. The annealing schedule is simulated with $\tau = 0.1$. The points that exceed the $\pm 5\%$ interval around zero are cropped to a grey color and marked with a $+/-$ symbol to indicate the sign of the extreme. The vertical axis fixes the total number of driving cycles of the protocol ($\omega\tau/2\pi$); the horizontal axis denotes the spectral reference frequency $\omega_0/2\pi$.



(a) Claey's protocol

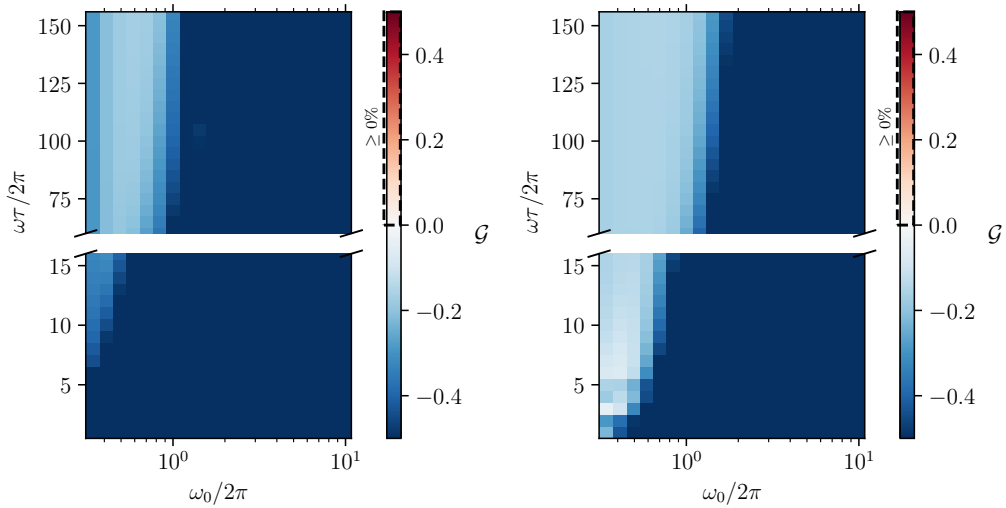
(b) $\mathcal{O}3$ protocol(c) $\mathcal{O}3L$ protocol

Figure C.12: Ising model instance with $N = 6$ qubits: gain for the new $\ell = 2$, $\mathcal{O}(\delta^3)$ FE protocols. Figure (b) is realized by using the $\mathcal{O}3$ protocol, whereas the $\mathcal{O}3L$ protocol is used in Figure (c). For comparison, Figure (a) shows the results of the Claey's protocol. The annealing is simulated with $\tau = 0.1$. The vertical axis fixes the total number of driving cycles of the protocol ($\omega\tau/2\pi$); the horizontal axis denotes the spectral reference frequency $\omega_0/2\pi$.

D

Further tests on models from reference paper

In this Appendix, we test the $\mathcal{O}(\delta^3)$ Floquet protocols on two systems that have been used in Ref. [15]. Since the Hamiltonians of these systems are not explicitly written in the form $\mathcal{H}(\lambda) = A(\lambda)\mathcal{H}_i + B(\lambda)\mathcal{H}_p$, we compute the gain as done in §4.3.b - see Eq. (4.14) - where the cost of the algorithm has been defined as the integral of the driving Hamiltonian norm.

D.1 The three-level system

This system consists of two qubits that behave as a three-level system,

$$\mathcal{H}(\lambda) = -2J\sigma_z^{(1)}\hat{\sigma}_z^{(2)} - h\left(\hat{\sigma}_z^{(1)} + \hat{\sigma}_z^{(2)}\right) + 2h\lambda\left(\hat{\sigma}_x^{(1)} + \hat{\sigma}_x^{(2)}\right) \quad (\text{D.1})$$

where the parameters are fixed to $J = 1$, $h = 2$, as in Ref. [15]. The results are shown in Figure D.1 and D.2.

D.2 The magnetic trap

This system consists of a non-integrable Ising chain with a magnetic trap whose center is shifted through the chain. The driving Hamiltonian is

$$\mathcal{H}(\lambda) = J \sum_{i=1}^{L-1} \hat{\sigma}_z^{(i)} \hat{\sigma}_z^{(i+1)} + h_z \sum_{i=1}^L \hat{\sigma}_z^{(i)} - h_t \sum_{i=1}^L \exp\left[-\frac{(i - c_t(\lambda))^2}{w_t^2}\right] \hat{\sigma}_z^{(i)} + h_x \sum_{i=1}^L \hat{\sigma}_x^{(i)}$$

with $c_t(\lambda) = (1 - \lambda)i_0 + \lambda i_f$ representing the center of the trap. The parameter $\lambda(t)$ is increasing from 0 to 1 during the annealing schedule of duration τ , shifting

the center of the trap from site i_0 to i_f . The results are shown in Figure D.3 and D.4.

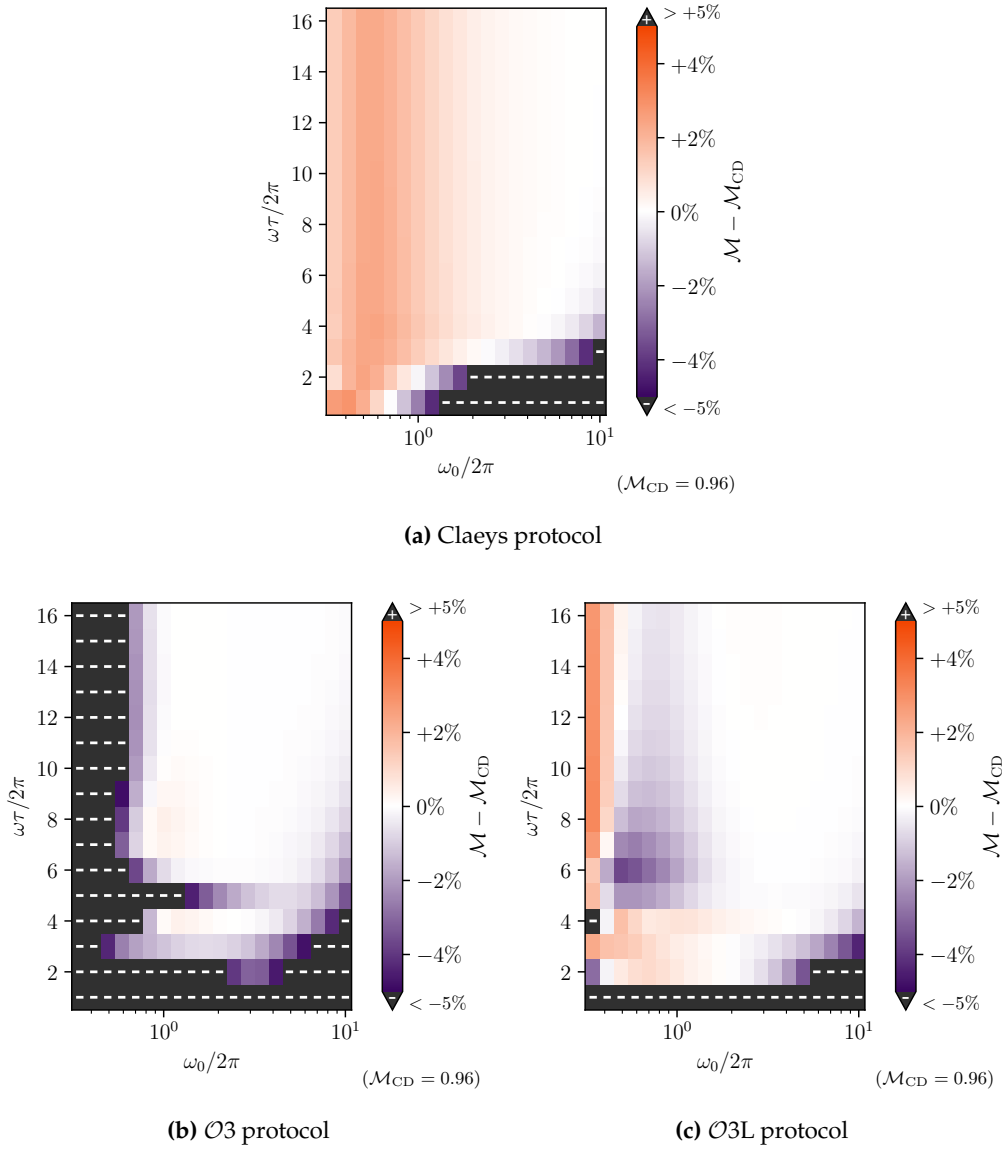
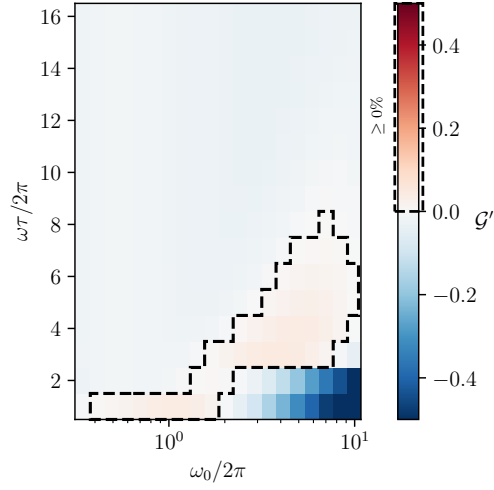


Figure D.1: 3-level system: difference in energy merit for the new $\mathcal{O}(\delta^3)$ FE protocols w.r.t. the reference $\ell = 1$ CD protocol. Figure (b) is realized by using the $\mathcal{O}3$ protocol, whereas the $\mathcal{O}3L$ protocol is used in Figure (c). For comparison, Figure (a) shows the results of the Claey's protocol. The annealing schedule is simulated with $\tau = 0.1$. The points that exceed the $\pm 5\%$ interval around zero are cropped to a grey color and marked with a $+/-$ symbol to indicate the sign of the extreme. The vertical axis fixes the total number of driving cycles of the protocol ($\omega\tau/2\pi$); the horizontal axis denotes the spectral reference frequency $\omega_0/2\pi$.



(a) Claeys protocol

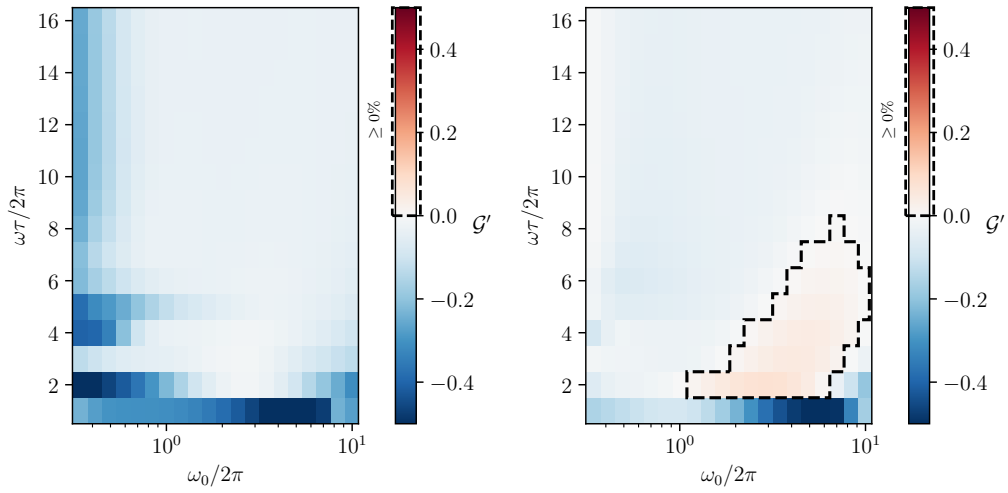
(b) $\mathcal{O}3$ protocol(c) $\mathcal{O}3L$ protocol

Figure D.2: 3-level system: gain for the new $\ell = 1$, $\mathcal{O}(\delta^3)$ FE protocols. Figure (b) is realized by using the $\mathcal{O}3$ protocol, whereas the $\mathcal{O}3L$ protocol is used in Figure (c). For comparison, Figure (a) shows the results of the Claeys protocol. The annealing is simulated with $\tau = 0.1$. The vertical axis fixes the total number of driving cycles of the protocol ($\omega\tau/2\pi$); the horizontal axis denotes the spectral reference frequency $\omega_0/2\pi$.

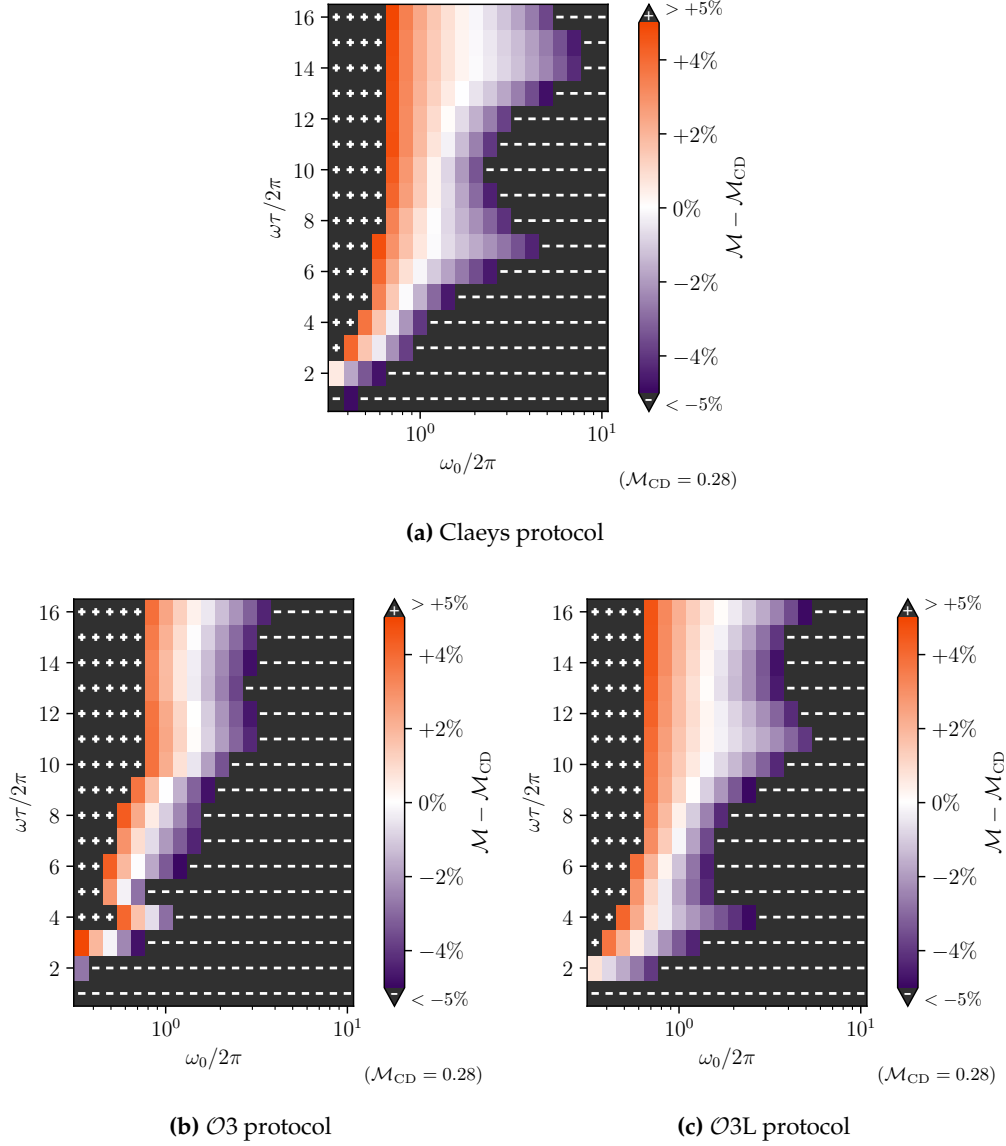


Figure D.3: Magnetic trap, $N = 6$ qubits: difference in energy merit for the new $\mathcal{O}(\delta^3)$ FE protocols w.r.t. the reference $\ell = 1$ CD protocol. Figure (b) is realized by using the $\mathcal{O}3$ protocol, whereas the $\mathcal{O}3L$ protocol is used in Figure (c). For comparison, Figure (a) shows the results of the Claeys protocol. The annealing schedule is simulated with $\tau = 0.1$. The points that exceed the $\pm 5\%$ interval around zero are cropped to a grey color and marked with a $+/-$ symbol to indicate the sign of the extreme. The vertical axis fixes the total number of driving cycles of the protocol ($\omega\tau/2\pi$); the horizontal axis denotes the spectral reference frequency $\omega_0/2\pi$.

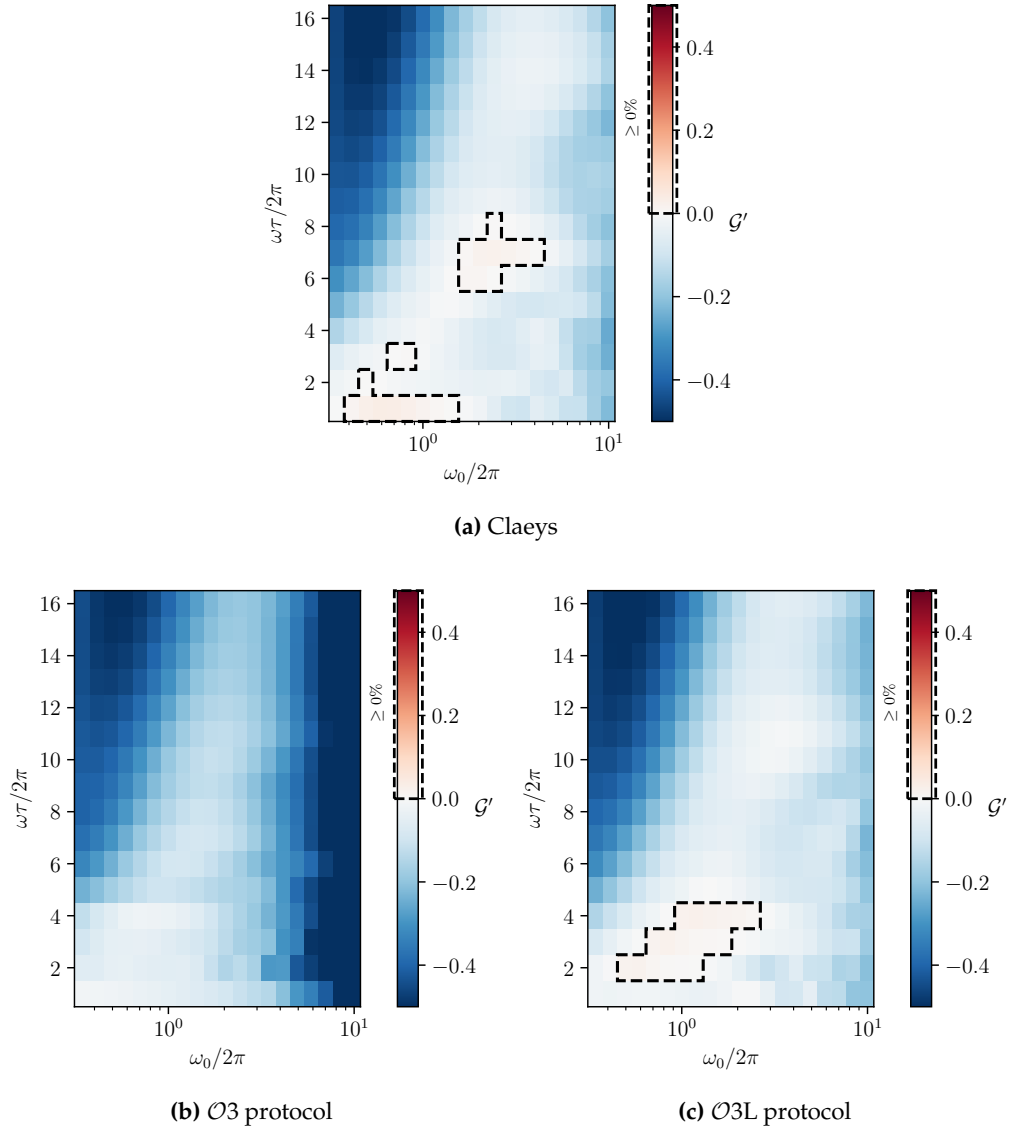


Figure D.4: Magnetic trap, $N = 6$ qubits: gain for the new $\ell = 1$, $\mathcal{O}(\delta^3)$ FE protocols. Figure (b) is realized by using the $\mathcal{O}3$ protocol, whereas the $\mathcal{O}3L$ protocol is used in Figure (c). For comparison, Figure (a) shows the results of the Claeys protocol. The annealing is simulated with $\tau = 0.1$. The vertical axis fixes the total number of driving cycles of the protocol ($\omega\tau/2\pi$); the horizontal axis denotes the spectral reference frequency $\omega_0/2\pi$.

Acronyms

- AGP** Adiabatic Gauge Potential. 31, 34–38, 41, 43, 45, 47, 54, 56, 73, 76, 86, 88, 89, 92, 97, 99, 100
- AQC** Adiabatic Quantum Computing. 26, 39
- ASP** Adiabatic State Preparation. 17, 18, 26–28, 31, 39–41, 63
- BQM** Binary Quadratic Models. 16
- CD** Counterdiabatic. 31, 34, 35, 38, 41, 43, 45, 46, 48–50, 52, 54, 57–60, 62, 64, 67–69, 71–73, 78, 79, 81, 84, 85, 89, 100, 103, 104, 106, 108, 110, 112, 114, 116, 120, 122
- CO** Combinatorial Optimization. 4
- CRAB** Chopped RAndom Basis. 87, 89, 92
- FE** Floquet Engineered. 41, 46, 47, 50, 52–54, 56–69, 71, 72, 74–82, 85, 106–117, 120–123
- LCD** Local Counterdiabatic Driving. 86–88, 90–92
- LHZ** Lechner-Hauke-Zoller. 2, 20, 22, 45, 50, 51, 59, 84, 86, 87, 105
- NISQ** Noisy Intermediate-Scale Quantum. 11, 18
- OLCD** Optimized Local Counterdiabatic Driving. 89–92
- PUBO** Polynomial Unconstrained Binary Optimization. 10

QA Quantum Annealing. 18, 39, 40

QOC Quantum Optimal Control. 87, 92

QUBO Quadratic Unconstrained Binary Optimization. 10, 15, 16

SA Simulated Annealing. 39, 40

TSP Travelling Salesman Problem. 7, 8

UA Unassisted Annealing. 46, 47, 50, 54, 57, 62–66, 68, 69, 71, 81, 89, 90, 92

Bibliography

- [1] Dorit Aharonov et al. “Adiabatic Quantum Computation is Equivalent to Standard Quantum Computation”. In: *SIAM Journal on Computing* 37.1 (2007), pp. 166–194. DOI: [10.1137/S0097539705447323](https://doi.org/10.1137/S0097539705447323). eprint: <https://doi.org/10.1137/S0097539705447323>. URL: <https://doi.org/10.1137/S0097539705447323>.
- [2] Tameem Albash and Daniel A. Lidar. “Adiabatic quantum computation”. In: *Reviews of Modern Physics* 90.1 (Jan. 2018). ISSN: 1539-0756. DOI: [10.1103/revmodphys.90.015002](https://doi.org/10.1103/revmodphys.90.015002). URL: <http://dx.doi.org/10.1103/RevModPhys.90.015002>.
- [3] David L. Applegate et al. 2007.
- [4] Ryan Babbush, Peter J Love, and Alán Aspuru-Guzik. “Adiabatic quantum simulation of quantum chemistry”. In: *Scientific reports* 4.1 (2014), p. 6603.
- [5] John E Beasley. *Heuristic algorithms for the unconstrained binary quadratic programming problem*. Tech. rep. Working Paper, The Management School, Imperial College, London, England, 1998.
- [6] David E Bernal et al. “Integer programming techniques for minor-embedding in quantum annealers”. In: *Integration of Constraint Programming, Artificial Intelligence, and Operations Research: 17th International Conference, CPAIOR 2020, Vienna, Austria, September 21–24, 2020, Proceedings 17*. Springer. 2020, pp. 112–129.
- [7] Louis-Philippe Bigras, Michel Gamache, and Gilles Savard. “The time-dependent traveling salesman problem and single machine scheduling problems with sequence dependent setup times”. In: *Discrete Optimization* 5.4 (2008), pp. 685–699.

- [8] S. Blanes et al. “The Magnus expansion and some of its applications”. In: *Physics Reports* 470.5–6 (Jan. 2009), pp. 151–238. ISSN: 0370-1573. DOI: [10.1016/j.physrep.2008.11.001](https://doi.org/10.1016/j.physrep.2008.11.001). URL: <http://dx.doi.org/10.1016/j.physrep.2008.11.001>.
- [9] Ieva Čepaitė et al. “Counterdiabatic Optimized Local Driving”. In: *PRX Quantum* 4 (1 Jan. 2023), p. 010312. DOI: [10.1103/PRXQuantum.4.010312](https://doi.org/10.1103/PRXQuantum.4.010312). URL: <https://link.aps.org/doi/10.1103/PRXQuantum.4.010312>.
- [10] Earl T Campbell, Barbara M Terhal, and Christophe Vuillot. “Roads towards fault-tolerant universal quantum computation”. In: *Nature* 549.7671 (2017), pp. 172–179.
- [11] Adolfo del Campo. “Shortcuts to Adiabaticity by Counterdiabatic Driving”. In: *Phys. Rev. Lett.* 111 (10 Sept. 2013), p. 100502. DOI: [10.1103/PhysRevLett.111.100502](https://doi.org/10.1103/PhysRevLett.111.100502). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.111.100502>.
- [12] Tommaso Caneva, Tommaso Calarco, and Simone Montangero. “Chopped random-basis quantum optimization”. In: *Phys. Rev. A* 84 (2 Aug. 2011), p. 022326. DOI: [10.1103/PhysRevA.84.022326](https://doi.org/10.1103/PhysRevA.84.022326). URL: <https://link.aps.org/doi/10.1103/PhysRevA.84.022326>.
- [13] V. Cannella and J. A. Mydosh. “Magnetic Ordering in Gold-Iron Alloys”. In: *Phys. Rev. B* 6 (11 Dec. 1972), pp. 4220–4237. DOI: [10.1103/PhysRevB.6.4220](https://doi.org/10.1103/PhysRevB.6.4220). URL: <https://link.aps.org/doi/10.1103/PhysRevB.6.4220>.
- [14] Nikolai M. Chtchelkatchev et al. “Bell inequalities and entanglement in solid-state devices”. In: *Phys. Rev. B* 66 (16 Oct. 2002), p. 161320. DOI: [10.1103/PhysRevB.66.161320](https://doi.org/10.1103/PhysRevB.66.161320). URL: <https://link.aps.org/doi/10.1103/PhysRevB.66.161320>.
- [15] Pieter W. Claeys et al. “Floquet-Engineering Counterdiabatic Protocols in Quantum Many-Body Systems”. In: *Phys. Rev. Lett.* 123 (9 Aug. 2019), p. 090602. DOI: [10.1103/PhysRevLett.123.090602](https://doi.org/10.1103/PhysRevLett.123.090602). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.123.090602>.
- [16] Clayton W Commander. “Maximum cut problem, MAX-cut.” In: *Encyclopedia of Optimization* 2 (2009).
- [17] Stephen A Cook. “The complexity of theorem-proving procedures”. In: *Logic, Automata, and Computational Complexity: The Works of Stephen A. Cook*. 2023, pp. 143–152.
- [18] Domenico d’Alessandro. *Introduction to quantum control and dynamics*. Chapman and hall/CRC, 2021.

- [19] Luca D'Alessio et al. "From quantum chaos and eigenstate thermalization to statistical mechanics and thermodynamics". In: *Advances in Physics* 65.3 (May 2016), pp. 239–362. ISSN: 1460-6976. DOI: [10.1080/00018732.2016.1198134](https://doi.org/10.1080/00018732.2016.1198134). URL: <http://dx.doi.org/10.1080/00018732.2016.1198134>.
- [20] Arnab Das and Bikas K. Chakrabarti. "Colloquium: Quantum annealing and analog quantum computation". In: *Rev. Mod. Phys.* 80 (3 Sept. 2008), pp. 1061–1081. DOI: [10.1103/RevModPhys.80.1061](https://doi.org/10.1103/RevModPhys.80.1061). URL: <https://link.aps.org/doi/10.1103/RevModPhys.80.1061>.
- [21] Prasanna Date et al. "Efficiently embedding QUBO problems on adiabatic quantum computers". In: *Quantum Information Processing* 18 (2019), pp. 1–31.
- [22] David Elieser Deutsch and Roger Penrose. "Quantum computational networks". In: *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences* 425.1868 (1989), pp. 73–90. DOI: [10.1098/rspa.1989.0099](https://doi.org/10.1098/rspa.1989.0099). eprint: <https://royalsocietypublishing.org/doi/pdf/10.1098/rspa.1989.0099>. URL: <https://royalsocietypublishing.org/doi/abs/10.1098/rspa.1989.0099>.
- [23] David P DiVincenzo and Daniel Loss. "Quantum computers and quantum coherence". In: *Journal of Magnetism and Magnetic Materials* 200.1 (1999), pp. 202–218. ISSN: 0304-8853. DOI: [https://doi.org/10.1016/S0304-8853\(99\)00315-7](https://doi.org/10.1016/S0304-8853(99)00315-7). URL: <https://www.sciencedirect.com/science/article/pii/S0304885399003157>.
- [24] Patrick Doria, Tommaso Calarco, and Simone Montangero. "Optimal control technique for many-body quantum dynamics". In: *Physical review letters* 106.19 (2011), p. 190501.
- [25] Maike Drieb-Schön et al. "Parity quantum optimization: Encoding constraints". In: *Quantum* 7 (2023), p. 951.
- [26] S F Edwards and P W Anderson. "Theory of spin glasses". In: *Journal of Physics F: Metal Physics* 5.5 (May 1975), p. 965. DOI: [10.1088/0305-4608/5/5/017](https://doi.org/10.1088/0305-4608/5/5/017). URL: <https://dx.doi.org/10.1088/0305-4608/5/5/017>.
- [27] Samuel Frederick Edwards and Phil W Anderson. "Theory of spin glasses". In: *Journal of Physics F: Metal Physics* 5.5 (1975), p. 965.
- [28] Kilian Ender. "Advances in parity quantum computing: compilation, implementation, and application". In: (Mar. 2024).
- [29] Kilian Ender et al. "Parity Quantum Optimization: Compiler". In: *Quantum* 7 (Mar. 2023), p. 950. ISSN: 2521-327X. DOI: [10.22331/q-2023-03-17-950](https://doi.org/10.22331/q-2023-03-17-950). URL: <https://doi.org/10.22331/q-2023-03-17-950>.

- [30] Richard P Feynman. “Quantum mechanical computers”. In: *Optics news* 11.2 (1985), pp. 11–20.
- [31] Richard P Feynman et al. “Simulating physics with computers”. In: *Int. j. Theor. phys* 21.6/7 (2018).
- [32] Konrad H Fischer and John A Hertz. *Spin glasses*. 1. Cambridge university press, 1993.
- [33] Yaotian Fu and Philip W Anderson. “Application of statistical mechanics to NP-complete problems in combinatorial optimisation”. In: *Journal of Physics A: Mathematical and General* 19.9 (1986), p. 1605.
- [34] Michael R Garey and David S Johnson. *Computers and intractability*. Vol. 174. freeman San Francisco, 1979.
- [35] I. M. Georgescu, S. Ashhab, and Franco Nori. “Quantum simulation”. In: *Rev. Mod. Phys.* 86 (1 Mar. 2014), pp. 153–185. doi: [10.1103/RevModPhys.86.153](https://doi.org/10.1103/RevModPhys.86.153). URL: <https://link.aps.org/doi/10.1103/RevModPhys.86.153>.
- [36] Steffen J Glaser et al. “Training Schrödinger’s cat: Quantum optimal control: Strategic report on current status, visions and goals for research in Europe”. In: *The European Physical Journal D* 69 (2015), pp. 1–24.
- [37] Fred Glover, Gary Kochenberger, and Yu Du. *A Tutorial on Formulating and Using QUBO Models*. 2019. arXiv: [1811.11538](https://arxiv.org/abs/1811.11538) [cs.DS].
- [38] Fred Glover, Mark Lewis, and Gary Kochenberger. *Logical and Inequality Implications for Reducing the Size and Complexity of Quadratic Unconstrained Binary Optimization Problems*. 2017. arXiv: [1705.09545](https://arxiv.org/abs/1705.09545) [cs.AI].
- [39] Alexander K Hartmann and Martin Weigt. *Phase transitions in combinatorial optimization problems: basics, algorithms and statistical mechanics*. John Wiley & Sons, 2006.
- [40] Andreas Hartmann and Wolfgang Lechner. “Rapid counter-diabatic sweeps in lattice gauge adiabatic quantum computing”. In: *New Journal of Physics* 21.4 (Apr. 2019), p. 043025. ISSN: 1367-2630. doi: [10.1088/1367-2630/ab14a0](https://doi.org/10.1088/1367-2630/ab14a0). URL: <http://dx.doi.org/10.1088/1367-2630/ab14a0>.
- [41] Philipp Hauke et al. “Perspectives of quantum annealing: methods and implementations”. In: *Reports on Progress in Physics* 83.5 (May 2020), p. 054401. ISSN: 1361-6633. doi: [10.1088/1361-6633/ab85b8](https://doi.org/10.1088/1361-6633/ab85b8). URL: <http://dx.doi.org/10.1088/1361-6633/ab85b8>.
- [42] Karla L Hoffman, Manfred Padberg, Giovanni Rinaldi, et al. “Traveling salesman problem”. In: *Encyclopedia of operations research and management science* 1 (2013), pp. 1573–1578.

- [43] Sabine Jansen, Mary-Beth Ruskai, and Ruedi Seiler. “Bounds for the adiabatic approximation with applications to quantum computation”. In: *Journal of Mathematical Physics* 48.10 (Oct. 2007), p. 102111. ISSN: 0022-2488. DOI: [10.1063/1.2798382](https://doi.org/10.1063/1.2798382). eprint: https://pubs.aip.org/aip/jmp/article-pdf/doi/10.1063/1.2798382/16055030/102111_1_online.pdf. URL: <https://doi.org/10.1063/1.2798382>.
- [44] J.R. Johansson, P.D. Nation, and Franco Nori. “QuTiP 2: A Python framework for the dynamics of open quantum systems”. In: *Computer Physics Communications* 184.4 (2013), pp. 1234–1240. ISSN: 0010-4655. DOI: <https://doi.org/10.1016/j.cpc.2012.11.019>. URL: <https://www.sciencedirect.com/science/article/pii/S0010465512003955>.
- [45] Mark W Johnson et al. “Quantum annealing with manufactured spins”. In: *Nature* 473.7346 (2011), pp. 194–198.
- [46] Richard M Karp. “Combinatorics, complexity, and randomness”. In: *Communications of the ACM* 29.2 (1986), pp. 98–109.
- [47] Richard M Karp. *Reducibility among combinatorial problems*. Springer, 2010.
- [48] Tosio Kato. “On the adiabatic theorem of quantum mechanics”. In: *Journal of the Physical Society of Japan* 5.6 (1950), pp. 435–439.
- [49] Elias Khalil et al. “Learning combinatorial optimization algorithms over graphs”. In: *Advances in neural information processing systems* 30 (2017).
- [50] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. “Optimization by Simulated Annealing”. In: *Science* 220.4598 (1983), pp. 671–680. DOI: [10.1126/science.220.4598.671](https://doi.org/10.1126/science.220.4598.671). eprint: <https://www.science.org/doi/pdf/10.1126/science.220.4598.671>. URL: <https://www.science.org/doi/abs/10.1126/science.220.4598.671>.
- [51] Christiane P Koch et al. “Quantum optimal control in quantum technologies. Strategic report on current status, visions and goals for research in Europe”. In: *EPJ Quantum Technology* 9.1 (2022), p. 19.
- [52] Gary A. Kochenberger and Fred Glover. “A Unified Framework for Modeling and Solving Combinatorial Optimization Problems: A Tutorial”. In: *Multiscale Optimization Methods and Applications*. Ed. by William W. Hager et al. Boston, MA: Springer US, 2006, pp. 101–124. ISBN: 978-0-387-29550-3. DOI: [10.1007/0-387-29550-X_4](https://doi.org/10.1007/0-387-29550-X_4). URL: https://doi.org/10.1007/0-387-29550-X_4.

- [53] Michael Kolodrubetz et al. "Geometry and non-adiabatic response in quantum and classical systems". In: *Physics Reports* 697 (June 2017), pp. 1–87. ISSN: 0370-1573. DOI: [10.1016/j.physrep.2017.07.001](https://doi.org/10.1016/j.physrep.2017.07.001). URL: <http://dx.doi.org/10.1016/j.physrep.2017.07.001>.
- [54] Bernhard H Korte et al. *Combinatorial optimization*. Vol. 1. Springer, 2011.
- [55] Wolfgang Lechner, Philipp Hauke, and Peter Zoller. "A quantum annealing architecture with all-to-all connectivity from local interactions". In: *Science Advances* 1.9 (2015), e1500838. DOI: [10.1126/sciadv.1500838](https://doi.org/10.1126/sciadv.1500838). eprint: <https://www.science.org/doi/pdf/10.1126/sciadv.1500838>. URL: <https://www.science.org/doi/abs/10.1126/sciadv.1500838>.
- [56] Harry R Lewis. "Michael R. Garey and David S. Johnson. Computers and intractability. A guide to the theory of NP-completeness. WH Freeman and Company, San Francisco 1979, x+ 338 pp." In: *The Journal of Symbolic Logic* 48.2 (1983), pp. 498–500.
- [57] Daniel A Lidar. "On the quantum computational complexity of the Ising spin glass partition function and of knot invariants". In: *New Journal of Physics* 6 (Nov. 2004), pp. 167–167. ISSN: 1367-2630. DOI: [10.1088/1367-2630/6/1/167](https://doi.org/10.1088/1367-2630/6/1/167). URL: <http://dx.doi.org/10.1088/1367-2630/6/1/167>.
- [58] Elisabeth Lobe and Annette Lutz. "Minor embedding in broken chimera and derived graphs is np-complete". In: *Theoretical Computer Science* 989 (2024), p. 114369.
- [59] Bas Lodewijks. *Mapping NP-hard and NP-complete optimisation problems to Quadratic Unconstrained Binary Optimisation problems*. 2020. arXiv: [1911.08043](https://arxiv.org/abs/1911.08043) [cs.DS].
- [60] Andrew Lucas. "Ising formulations of many NP problems". In: *Frontiers in physics* 2 (2014), p. 5.
- [61] LW McKeehan. "A contribution to the theory of ferromagnetism". In: *Physical Review* 26.2 (1925), p. 274.
- [62] Marc Mézard, Giorgio Parisi, and Miguel Angel Virasoro. *Spin glass theory and beyond: An Introduction to the Replica Method and Its Applications*. Vol. 9. World Scientific Publishing Company, 1987.
- [63] Frederic Mila. "Frustrated spin systems". In: *Many-Body Physics: From Kondo to Hubbard* 5 (2015).
- [64] Satoshi Morita and Hidetoshi Nishimori. "Mathematical foundation of quantum annealing". In: *Journal of Mathematical Physics* 49.12 (2008).

- [65] Matthias M Müller et al. “One decade of quantum optimal control in the chopped random basis”. In: *Reports on progress in physics* 85.7 (2022), p. 076001.
- [66] Shuntaro Okada et al. “Improving solutions by embedding larger sub-problems in a D-Wave quantum annealer”. In: *Scientific reports* 9.1 (2019), p. 2098.
- [67] Lars Onsager. “Crystal Statistics. I. A Two-Dimensional Model with an Order-Disorder Transition”. In: *Phys. Rev.* 65 (3-4 Feb. 1944), pp. 117–149. DOI: [10.1103/PhysRev.65.117](https://doi.org/10.1103/PhysRev.65.117). URL: <https://link.aps.org/doi/10.1103/PhysRev.65.117>.
- [68] Christos H Papadimitriou and Kenneth Steiglitz. *Combinatorial optimization: algorithms and complexity*. Courier Corporation, 1998.
- [69] Christos H. Papadimitriou. “The Euclidean travelling salesman problem is NP-complete”. In: *Theoretical Computer Science* 4.3 (1977), pp. 237–244. ISSN: 0304-3975. DOI: [https://doi.org/10.1016/0304-3975\(77\)90012-3](https://doi.org/10.1016/0304-3975(77)90012-3). URL: <https://www.sciencedirect.com/science/article/pii/0304397577900123>.
- [70] Giorgio Parisi. “The physical meaning of replica symmetry breaking”. In: *arXiv preprint cond-mat/0205387* (2002).
- [71] Francesco Pio Barone et al. “Counterdiabatic optimized driving in quantum phase sensitive models”. In: *New Journal of Physics* (2024). URL: <http://iopscience.iop.org/article/10.1088/1367-2630/ad313e>.
- [72] John Preskill. “Fault-tolerant quantum computation”. In: *Introduction to quantum computation and information*. World Scientific, 1998, pp. 213–269.
- [73] John Preskill. “Quantum Computing in the NISQ era and beyond”. In: *Quantum* 2 (Aug. 2018), p. 79. ISSN: 2521-327X. DOI: [10.22331/q-2018-08-06-79](https://doi.org/10.22331/q-2018-08-06-79). URL: <https://doi.org/10.22331/q-2018-08-06-79>.
- [74] S.N. Parragh R.F. Hartl. *TSP history and milestones*. Slide 8/18. URL: <https://prolog.univie.ac.at/teaching/LVAs/KFK-TL/WS%2012/TL-Part4-TSP-History.pdf> (visited on 01/10/2024).
- [75] Niklas Rach et al. “Dressing the chopped-random-basis optimization: A bandwidth-limited access to the trap-free landscape”. In: *Physical Review A* 92.6 (2015), p. 062343.

- [76] P. Ray, B. K. Chakrabarti, and Arunava Chakrabarti. “Sherrington-Kirkpatrick model in a transverse field: Absence of replica symmetry breaking due to quantum fluctuations”. In: *Phys. Rev. B* 39 (16 June 1989), pp. 11828–11832. DOI: [10.1103/PhysRevB.39.11828](https://doi.org/10.1103/PhysRevB.39.11828). URL: <https://link.aps.org/doi/10.1103/PhysRevB.39.11828>.
- [77] Gerald E Sacks. “Multiple error correction by means of parity checks”. In: *IRE transactions on information theory* 4.4 (1958), pp. 145–147.
- [78] Dries Sels and Anatoli Polkovnikov. “Minimizing irreversible losses in quantum systems by local counterdiabatic driving”. In: *Proceedings of the National Academy of Sciences* 114.20 (2017), E3909–E3916.
- [79] John Shalf. “The future of computing beyond Moore’s Law”. In: *Philosophical Transactions of the Royal Society A* 378.2166 (2020), p. 20190061.
- [80] Isaac D. Smith, Hendrik Poulsen Nautrup, and Hans J. Briegel. *Parity Quantum Computing as YZ-Plane Measurement-Based Quantum Computing*. 2024. arXiv: [2401.10079](https://arxiv.org/abs/2401.10079) [quant-ph].
- [81] Byron Tasseff et al. *On the Emerging Potential of Quantum Annealing Hardware for Combinatorial Optimization*. 2022. arXiv: [2210.04291](https://arxiv.org/abs/2210.04291) [math.OC].
- [82] Delft University of Technology. *Applications of Quantum Mechanics: lecture notes at TU Delft*. Chapter 6, section 2. 2019. URL: https://appquantmech.quantumtinkerer.tudelft.nl/adiabatic_proof/ (visited on 01/10/2024).
- [83] Walter Vinci and Daniel A. Lidar. “Non-stoquastic Hamiltonians in quantum annealing via geometric phases”. In: *npj Quantum Information* 3.1 (Sept. 2017). ISSN: 2056-6387. DOI: [10.1038/s41534-017-0037-z](https://doi.org/10.1038/s41534-017-0037-z). URL: <http://dx.doi.org/10.1038/s41534-017-0037-z>.
- [84] Walter Vinci et al. “Quantum annealing correction with minor embedding”. In: *Phys. Rev. A* 92 (4 Oct. 2015), p. 042310. DOI: [10.1103/PhysRevA.92.042310](https://doi.org/10.1103/PhysRevA.92.042310). URL: <https://link.aps.org/doi/10.1103/PhysRevA.92.042310>.
- [85] Anita Weidinger, Glen Bigan Mbeng, and Wolfgang Lechner. “Error mitigation for quantum approximate optimization”. In: *Phys. Rev. A* 108 (3 Sept. 2023), p. 032408. DOI: [10.1103/PhysRevA.108.032408](https://doi.org/10.1103/PhysRevA.108.032408). URL: <https://link.aps.org/doi/10.1103/PhysRevA.108.032408>.
- [86] Stephan Wojtowysch. “Stochastic gradient descent with noise of machine learning type Part I: Discrete time analysis”. In: *Journal of Nonlinear Science* 33.3 (2023), p. 45.

- [87] Sheir Yarkoni et al. "Quantum annealing for industry applications: Introduction and review". In: *Reports on Progress in Physics* 85.10 (2022), p. 104001.



Acknowledgments

Thinking about the last months, I realized how much I have learned, how many *quantum people*¹ I have met, and how beautiful are the places I have visited. My hope for the future is that this path, which I began five years ago by enrolling in a Physics degree, will lead to many more of these experiences. In the meantime, the following acknowledgments are for all those who contributed to these wonderful years of MSc studies.

Let me start by acknowledging my supervisors at the University of Innsbruck. To Dr. Glen Mbeng, for looking after me through this thesis work and respecting the sacredness of coffee breaks. I am thankful to Prof. Wolfgang Lechner, for the guidance and support that I have received during my stay in Innsbruck. Of course, I still remember the good time I had with the members of the Quantum Optimization group: thank you for being part of this experience! Cheers to Pedro, who wanted so desperately to be mentioned here.

Speaking of my time in foreign lands, I can't forget the internship at CERN Quantum Technology Initiative. I couldn't have asked for better supervisors than Michele Grossi and Oriel Kiss, so I hope to cross paths with you again, possibly far away from the ways of counterdiabaticity.

Turning the sight back to my home institution, there is one figure that has been crucial to my education and growth. I am thankful to Prof. Simone Montangero for his unconditional support and for his lectures, which brought me into the quantum realm in the first place.

Family matters for many reasons, and a paragraph wouldn't be enough to

¹Students, researchers, and professors who work in the vast multiverse of quantum physics.

explain why. However, over five years of studies, I've seen examples of how unpredictable a family environment can be and how much it can affect a person's serenity. Therefore, I'm grateful to my family for always supporting my decisions and for the sacrifices that they have made to provide me with the opportunities I have had.

Finally, there is another 'family' that must be acknowledged. In the last two years, I've shared a roof and bread with approximately thirty students. This group of pariahs, the VPMG, has boosted my caffeine addiction, filled my afternoons with nonsense activities, and provided me with the best games of Cards Against Humanity. By all means, the family of VPMG includes the phenomenal girls that, due to a hostile heritage, dwell in another residence. The coffee room wouldn't be as cheerful without you. So, VPMG: thanks to you all, for being the best university friendships I could ever have asked for during these two years in Padua.