UNIVERSITÀ DEGLI STUDI DI TRENTO

Department of Physics

Quantum computing for biophysical and optimization problems

SUPERVISORS:

Prof. Pietro Faccioli
Prof. Philipp Hauke
Prof. Davide Pastorello
Prof. Enrico Blanzieri

REFEREES:

Prof. Roberto Covino
Prof. Mikel Sanz

PH.D. CANDIDATE:

Veronica Panizza

Academic Year 2024/2025

Ad Enrico, per l'affetto che ci lega,
a mia sorella Anna, per essere da sempre la mia ancora,
al nonno Marco, per aver sempre incoraggiato le sue "bambine"

Acknowledgements

First and foremost, I would like to express my deepest gratitude to my supervisors, Pietro Faccioli, Philipp Hauke, Davide Pastorello, and Enrico Blanzieri. I owe special thanks to Pietro Faccioli, with whom I had the pleasure of working in strict collaboration and quite intensively throughout this PhD journey. Looking back, I imagine I will once again see Pietro at the blackboard — the eyes twinkling with enthusiasm — after long hours of discussion on a stifling summer day. I don't just thank Pietro for the physics he taught me but especially for sparking a genuine love for curiosity and research, for knowing *exactly* when to push me, and for offering unwavering encouragement when I felt "lost in research". I would also like to express my deepest gratitude to Philipp, who encouraged and supported me in pursuing an interdisciplinary research line. While I had the freedom to explore the applications of quantum computing in biophysics — what I consider one of the "Far West" of science, both in terms of its vast potential and the risk of getting lost in it — I always knew I could count on Philipp's solid scientific guidance.

I am sincerely grateful to Roberto Covino and Mikel Sanz for their time and dedication in carefully revising this manuscript.

I would also like to extend my sincere thanks to my collaborators, Cristian Micheletti, Alessandro Roggero, and Ricardo Costa de Almeida. Working with them has been an enriching experience, and I greatly appreciate their contributions, expertise, and stimulating discussions, which have helped shape the direction of my work.

A special mention goes to my colleagues (in Trento, Milano, and Trieste), whose support and good company have made this journey all the more enjoyable. The stimulating conversations we had during our fruit breaks and shared challenges — comprising quests for chairs in Palermo, fence jumping, burocratical via crucis at the Commissariato del Governo, and other adventures — have been as valuable as the research itself. It was a pleasure to be part of a group where birthdays and happy moments are celebrated with such enthusiasm — and possibly with delicious indian home-cooked food for everyone. Special thanks to Matteo and Danial, who generously dedicated some of their time to reviewing this dissertation. Their insightful feedback significantly enhanced the quality of this manuscript.

Lastly, I am immensely grateful to my loved ones, especially my sister,

Anna, my grandfather, Marco, and my dear Enrico. Marco is not “just” a grandfather; he is a teacher, a passionate geologist, and a great lover of mountains. Above all, he has always placed immense value on knowledge and continuously encouraged my sister and me to seek it. My little Anna is my colorful and chaotic sunshine. Ever since she stormed into my life like a hurricane, she has continued to astonish me with her intelligence, humanity, and courage. Now, as a teacher, she cultivates curiosity in her students with passion and diligence. I profoundly admire her! As for Enrico, with whom I have chosen to share my life, I have always found him to be much more than just a partner. He is a compassionate listener, always there to hear my thoughts and concerns, no matter how big or small. He has stood by my side through difficulties, offering his unwavering love and care.

These years have had their share of challenges, and only a fraction of them were related to physics. We helped one another and stuck together — and we did great!

To all of you — THANK YOU!

Summary

The remarkable progress of quantum technologies over recent years has driven significant efforts toward developing algorithms with applications to a wide range of research fields. Beyond fully quantum algorithms — whose efficacy remains constrained by technological limitations — hybrid quantum-classical algorithms and quantum-inspired methods have emerged as promising avenues for tackling real-world problems. In this study, we focus on two particularly challenging biophysics problems: protein design and polymer sampling.

Protein design involves engineering the primary sequence of a protein to ensure that it folds into a specific target conformation of biological interest. Our approach employs a physics-based machine learning model that incorporates a QUBO (Quadratic Unconstrained Binary Optimization) encoding of the design problem that is amenable to adiabatic quantum platforms such as the D-Wave device.

For the polymer sampling problem — where the objective is to sample both the sequence and the conformation of polymers according to a thermal distribution — we establish a deep connection with an Abelian lattice gauge theory populated with fermions. Building on this theoretical framework, we develop a quantum-inspired Monte Carlo protocol that not only eliminates the sign problem but also features a decorrelation time that scales linearly with the system size in the dense-melt polymer regime, providing a novel approach to computational polymer physics.

Within the framework of lattice gauge theories, where physically realizable measurements are heavily constrained by local symmetries, we analyze from the quantum-information perspective the problem of pinpointing entangled states by resorting to entanglement witnesses. Furthermore, we develop a numerical optimization protocol that enhances the effectiveness

of entanglement witnesses while ensuring their physical implementation within the lattice gauge theory framework.

Contents

PART I

INTRODUCTION

Chapter 1

Quantum computing

The first theoretical investigations concerning the possibility of devising quantum mechanical systems reproducing arbitrary computational processes can be traced back to 1980 [1]. These first investigations laid the basis for the idea — introduced in the famous Feynman's speech "Simulating physics with computers" [2] — of simulating physics using a well-controlled quantum system. A key motivation for this approach is the potential to efficiently simulate strongly correlated quantum systems, a task that would otherwise be exponentially hard on a classical computer. Further theoretical developments showed that a quantum generalization of a universal Turing machine exists and is compatible with the foundations of quantum mechanics [3]; in addition, Deutsch intuited the potential advantages of quantum machines with respect to classical ones, especially in terms of speed, owing to 'quantum parallelism'. A couple of decades later, the minimal requirements featuring a quantum computer were formulated [4]. According to these criteria, any quantum platform has to be a *scalable* quantum system equipped with well-characterized qubits, whose idle state can be initialized by applying a suitable protocol. Moreover, the system has to be fully tunable, so that *any unitary evolution* operating on the qubits can be established. To support the entire computation, the qubits' decoherence time must be sufficiently long. Finally, the availability of a measurement protocol is crucial, as it allows reading-off the outcomes of the quantum computation.

The approach adopted in manipulating these quantum platforms to perform quantum computing tasks is not unique. Indeed, several paradigms exist for approaching quantum computing, including the gate-based model, adiabatic quantum computing, and the measurement-based approach.

The growing interest in quantum computers resulted in many efforts toward developing applications extending beyond the simulation of physical systems. Indeed, significant milestones in quantum computing include the formulation of a quantum algorithm to factorize integer composite numbers [5] and search in a large and unstructured database [6]. On the one hand, these works corroborate previous theoretical studies concerning the capability of quantum computers to reproduce arbitrary classic processes; on the other hand, they prove that quantum algorithms, due to inherited quantum mechanical features, achieve substantial speed-up. Specifically, Shor's algorithm offers exponential speed-up with respect to the best-known classical algorithm, whereas Grover's search gives a quadratic gain in speed. With the passing of years, the applications of quantum computing extended to other relevant problems, such as Knapsack, traveling salesman, max-cut, bin-packing, and minimum k-cut problems.

Since the foundations of quantum computing were established, a primary goal has been to identify the potential advantages offered by this fundamentally different approach to computation. Initially, the advantage was mostly intended as a gain in speed, but in recent years, other factors, such as energy efficiency and reduced memory requirements, have also been considered. Stronger claims regarding quantum computers' capabilities — known as *quantum supremacy* claims — involve the possibility that for certain problems, the time required by quantum computers to solve them scales much more favorably with respect to the system's size compared to classical computers. In such cases, solving large problem instances classically would become practically infeasible, whereas quantum computers could provide solutions within reasonable timeframes. Reference [7] addresses the theoretical limitations of classic computers by relying on complexity theory results and states the minimal requirements that experiments need to meet to prove quantum supremacy. One of the first attempts to prove quantum supremacy involved an experiment conducted on the Sycamore 53-qubit quantum processor produced by Google. This allowed successful sampling of the output of pseudo-random quantum circuits with a time rate of an instance every 200 s [8]; it was argued that the same task would have required state-of-the-art supercomputers 10,000 years. This claim was soon disproved, as the novel tensor-network algorithm named "big-batch method" succeeded in predicting the outcome probability distribution deriving from random quantum circuits within a reasonable timescale [9]. Another attempt to prove quantum supremacy consisted in an experiment that employed a squeezed source of light and threshold detectors to perform boson sampling [10]. Indeed, the outcome distribution of such a setup can be predicted classically through repeated evaluations of Torontians, leading to a total computational complexity comparable to that of evaluating the Hafnian, a generalization of the permanent [11]. Several years later, the trotterized dynamics for the Kicked Ising Model was implemented on the IBM's 127-qubit processor [12]. By pushing the simulation into a strongly entangled regime, where brute-force classical computing techniques break down, the solutions retrieved using the quantum computer were proven to be accurate. However, more recently, a novel tensor network approach was employed to efficiently derive even more accurate results than those obtained using the IBM's quantum platform [13]. Another recent claim [14] of quantum supremacy is based on experiments carried out on the D-Wave quantum annealing superconducting platform, where it was possible to simulate the non-equilibrium dynamics of large ferromagnetic systems through a quantum phase transition, a result that is expected to take to the Frontier supercomputer order of millions of years to be reproduced. Despite these claims, there is still considerable debate regarding quantum advantage and supremacy, particularly in the Noisy Intermediate-Scale Quantum (NISQ) era of quantum technologies, where fault-tolerant systems are out of reach due to current technological limitations.

In the following sections, we specialize the discussion on digital and analog quantum computers and review relevant notions of information theory.

1.1 Digital quantum computers

As we have anticipated, the fundamental degrees of freedom that are employed in quantum computing are *qubits*, which are elements of a Hilbert space — the so-called *qubit Hilber space* — that is spanned by two states, namely $|0\rangle$ and $|1\rangle$. Quantum computation can be achieved by manipulating these elementary units, a task that can be approached in several ways. One of these is known as the *gate-based* paradigm of digital quantum computing and requires to decompose generic unitaries acting on a system of qubits — initialized in an idle state, usually chosen to be $|\mathbf{0}\rangle = \prod_k |0_k\rangle$ — as a concatenation and/or tensor product of simpler unitaries that typically act on either single qubits or pairs.

To allow the representation of *any* unitary evolution, it is necessary to dispose of a so-called *universal set of gates*. In general, proving a set of gates to be universal is not an easy task, but once one such a set has been singled out, it is possible to combine elements therein to produce other universal sets. An example of a universal set of gates is embodied by $\{\hat{R}_Z(\frac{\pi}{4}), \hat{H}, \hat{C}_{\mathrm{not}}\}$, where $\hat{R}_Z(\frac{\pi}{4})$ acts on a single qubit shifting the phase of $|1\rangle$ by $\frac{\pi}{4}$ with respect to $|0\rangle$, $\hat{H}$ is a single-qubit gate capable of creating and destroying balanced superposition among $|0\rangle$ and $|1\rangle$ states, and $\hat{C}_{\mathrm{not}}$ — known as *controlled-not* gate — is a gate operating on two qubits that unlocks the possibility of generating quantum correlation (i.e., entanglement). This universal set is often referred to as *elementary set* and usually serves as a reference to build different sets of universal gates that better suit the technological requirements of individual quantum platforms.

To design a quantum circuit implementing an arbitrary unitary of interest, $\hat{U}$, it is necessary to decompose $\hat{U}$ in terms of gates belonging to a universal set. This operation, known as *quantum compilation*, has no unique outcome, as potentially many quantum circuits can feature the same unitary evolution. Usually the aim of quantum compilation is not to reproduce the unitary of interest *exactly*, but rather to reproduce it within a *fixed tolerance*. This is so especially when devising algorithms addressing current NISQ technologies. Indeed, given the limitations of these devices, a viable strategy for choosing among different circuit designs is to minimize both the circuit depth and the number of gates employed. The number of gates necessary to approximate with accuracy δ a generic unitary $\hat{U}$ acting on N qubits scales as $O(e^N)\mathrm{poly}\big(\log(\frac{1}{\delta})\big)$ [15]. However, by restricting $\hat{U}$ to establish up to two-body interaction terms among qubits, i.e., avoiding higher-order interactions, $O(N \mathrm{poly}\log\frac{N}{\delta})$ elementary gates are sufficient for approximating $\hat{U}$.

In designing the quantum circuit, it is also necessary to plan how to extract the relevant information from the outcome wavefunction. To this end, projective measurements of all qubits or a subset of the register are taken and the measured outcomes are stored as classic bit strings. In most applications, the whole procedure has to be repeated several times to ensure robust statistics associated with the output wavefunction.

There are several kinds of quantum platforms that support a universal set of operations, including trapped-ions and neutral-atoms platforms, superconducting chips, semiconductor systems, and photonic devices. In what follows, we outline the main features of the most common available technologies.

In devices based on trapped ions [16–18], the two-level system is realized using two hyperfine atomic levels, which can be manipulated through laser controls to perform single-qubit rotations. In this device, two-qubit gates are realized by coupling the atomic degree of freedom of two qubits with the phononic degree of freedom of the ions in the trap. The main advantage of this platform stems from the long coherence time, on the order of 1 minute [19], which compares well with the typical gate's speed, as gates operate within the timescale of microseconds. A drawback associated with the design of this technology is the difficulty in scaling up the system, as implementing 1D traps containing more than a few dozen qubits remains challenging [20]. Overcoming this technical challenge potentially amounts to implementing smaller clusters of trapped ions that are allowed to exchange qubits by resorting to dynamic electric fields [21].

Another technology, instead of involving ions, employs neutral atoms. These are arranged on a lattice by means of optical tweezers (Rydberg atoms) or optical lattices (ultracold atom platforms). When optical tweezers are involved, the two-level system featuring a qubit stems from the two states that an atom can assume: groundstate or Rydberg state, where the valence electron is excited to a very high energy orbit. Relevantly, digital quantum computation is enabled when using optical tweezers, as a phenomenon known as Rydberg-blockade offers the possibility of establishing a native two-qubit gate [22, 23].

Regarding superconducting facilities [24–29], which can be directly controlled with electronic devices, the two-level system adopted as a qubit is obtained by shifting the equally-spaced energetic levels of an LC oscillator with the aid of a non-linear device (i.e., a Josephson junction). Despite this platform lacking the possibility of having all-to-all connectivity, it allows parallel implementation of several gates. A drawback associated with this technology lies in the short coherence time of qubits, which is typically in the order of hundreds of microseconds. However, with gate operation times around $10ns$, many operations can be executed within the coherence window [20].

Another relevant class of solid-state platforms consists of semiconductor-based systems, where a lattice of semiconductor islands on a silicon substrate is used to isolate electrons. In this case, the two electronic spin states form a two-level system that represents a qubit. In this framework, single-qubit gates are induced by magnetic fields, while two-qubit gates naturally arise from direct magnetic coupling between electrons lying in neighbouring dots. To achieve a *tunable* coupling, it is necessary to employ an additional quantum dot to mediate the interaction [30]. With respect to superconducting facilities, a system of quantum dots operates at higher temperatures, in the order of 1K.

In the next section, we introduce another relevant paradigm of quantum computing, on which we resorted to extrapolate results in Chapter 3.

1.2 Analog quantum computers

In the analog paradigm of quantum computing, instead of manipulating an initial state to produce the desired output, $|\Psi\rangle$, the Hamiltonian governing the system is chosen to encode $|\Psi\rangle$ in its ground-state (i.e., $\hat{H}\,|\Psi\rangle = e_{\text{gnd}}\,|\Psi\rangle$). Importantly, it was proven that any quantum circuit in the gate-based model can be efficiently mapped to a Hamiltonian in the analog quantum paradigm reproducing the same outcome [31]. It is worth noticing that this result proves that adiabatic quantum computing, being equivalent to circuit-based models, allows universal quantum computation. However, no quantum adiabatic processor capable of implementing universal quantum computation has been developed yet, so current technologies remain purpose-specific.

Turning our focus to the implementation of analog quantum computation, once a suitable Hamiltonian, $\hat{H}_P$, for the application at hand has been singled out, it is necessary to lead the system toward its ground-state. To do so, an alternative to decreasing the temperature to reach the absolute zero is establishing a *quantum anneal*. This protocol finds its theoretical foundation in the adiabatic theorem [32], which ensures that a system initially in the ground state of a given Hamiltonian, $\hat{H}_D$, will remain in the instantaneous ground state of the time-dependent Hamiltonian

$$\hat{H}(t) = \left(1 - \frac{t}{\tau}\right)\hat{H}_D + \frac{t}{\tau}\hat{H}_P \tag{1.1}$$

as long as the evolution is sufficiently slow. In the previous, $\hat{H}_P$ is the Hamiltonian governing the system at the end of the annealing process and it encodes in its ground state the solution of a given problem. For this reason, it is often referred to as *problem Hamiltonian*, while $\hat{H}_D$ goes under the name of *driver Hamiltonian* and, if chosen to be non-commuting with

$\hat{H}_P$, it stimulates the quantum fluctuations that are necessary for quantum tunneling to take place. A crucial request on $\hat{H}_D$ prescribes that its ground state has to be easily reached by the quantum platform. With this premise, for the quantum annealing to be successful, the annealing time τ has to be sufficiently long to guarantee the applicability of the adiabatic theorem. As a general rule, the annealing time τ is inversely proportional to the minimum energy gap separating the instantaneous ground- and first-excited states.

Biamonte and Love [33] proved that the simplest problem Hamiltonians, $\hat{H}_P$, supporting quantum universal computation are given by

$$\begin{aligned} \hat{H}_P =& \sum_{\mathrm{i}} \mu_{\mathrm{i}}\, \hat{\sigma}_{\mathrm{i}}^{(z)} + \sum_{\mathrm{i}} \Delta_{\mathrm{i}}\, \hat{\sigma}_{\mathrm{i}}^{(x)} \\ &+ \sum_{\mathrm{i}>\mathrm{j}} J_{\mathrm{ij}}\, \hat{\sigma}_{\mathrm{i}}^{(z)} \hat{\sigma}_{\mathrm{j}}^{(z)} + \sum_{\mathrm{i}>\mathrm{j}} K_{\mathrm{ij}}\, \hat{\sigma}_{\mathrm{i}}^{(x)} \hat{\sigma}_{\mathrm{j}}^{(x)} \,, \end{aligned} \tag{1.2}$$

and

$$\begin{aligned} \hat{H}_P =& \sum_{\mathrm{i}} \mu_{\mathrm{i}}\, \hat{\sigma}_{\mathrm{i}}^{(z)} + \sum_{\mathrm{i}} \Delta_{\mathrm{i}}\, \hat{\sigma}_{\mathrm{i}}^{(x)} \\ &+ \sum_{\mathrm{i}>\mathrm{j}} J_{\mathrm{ij}} \hat{\sigma}_{\mathrm{i}}^{(x)} \hat{\sigma}_{\mathrm{j}}^{(z)} + \sum_{\mathrm{i}>\mathrm{j}} K_{\mathrm{ij}} \hat{\sigma}_{\mathrm{i}}^{(z)} \hat{\sigma}_{\mathrm{j}}^{(x)} \,. \end{aligned} \tag{1.3}$$

Even if current adiabatic technologies do not support universal quantum computation, their applications proved successful in approximating solutions to integer optimization problems including the Traveling Salesman problem, Knapsack problems, portfolio optimization, max-cut problems, and other relevant NP-hard problems [34–36].

In modern technologies, spurious interactions between the quantum processor and the environment cause decoherence of qubits and the adiabatic theorem is hardly satisfied so that the system behaves as if operating at a finite low-temperature regime. At first sight, this could be interpreted as a heavy short-coming of the whole approach, but it can be turned into a useful tool, as quantum annealers can be used to sample thermal distributions as an alternative to classic Markov-chain algorithms.

One appealing application of adiabatic machines, even if not universal, is the solution of integer optimization problems by encoding them into the problem Hamiltonian, $\hat{H}_P$. In the following section, we discuss the encoding of integer optimization problems on quantum adiabatic machines and propose an illustrative example based on the Knapsack problem. In Chapter 3, we will make use of these strategies to encode a particular integer problem — connected to the optimization of an amino-acid sequence — as a QUBO.

1.2.1 QUBO problems

The problem Hamiltonians that most common quantum analog devices allow to implement take the form of Ising Hamiltonians, namely

$$\hat{H}_P = \sum_{\mathrm{i}} \mu_{\mathrm{i}}\, \hat{\sigma}_{\mathrm{i}}^{(z)} + \sum_{\mathrm{i}>\mathrm{j}} J_{\mathrm{ij}}\, \hat{\sigma}_{\mathrm{i}}^{(z)}\, \hat{\sigma}_{\mathrm{j}}^{(z)} \,, \tag{1.4}$$

where μ_{i} and J_{ij} are real parameters defining, respectively, local fields and couplings acting on the qubits. The driving Hamiltonian is given by a uniform transverse field, i.e., $\hat{H}_D = -\sum_{\mathrm{i}} \hat{\sigma}_{\mathrm{i}}^{(x)}$. As mentioned in the previous section, these systems are not capable of reproducing arbitrary Hamiltonians and consequently do not support universal quantum computing. Nonetheless, the applications of these quantum solvers range from finding solutions to relevant optimization problems [37–40], simulating phase transitions [41–43], and sampling from thermal distributions [44–46].

To illustrate how these machines can be harnessed to solve optimization problems, let us consider the minimization of the quadratic expression defined by

$$f(\mathbf{q}) = \sum_{i\leq j=1}^{n} q_i\, Q_{ij}\, q_j \,, \tag{1.5}$$

where $\mathbf{q} = \{q_1, \ldots, q_n\}$ is an array of binary variables that are unconstrained and take values in $\{0, 1\}$ and Q_{ij} are arbitrary real parameters. The previous expression can be regarded as the eigenvalue equation corresponding to the operator identity

$$\hat{H}_Q = \sum_{\mathrm{i}\leq\mathrm{j}=1}^{n} \hat{q}_{\mathrm{i}}\, Q_{\mathrm{ij}}\, \hat{q}_{\mathrm{j}} \,, \tag{1.6}$$

where $\hat{q}_{\mathrm{i}}$ are operators having eigenvalues 0 and 1, corresponding, respectively, to the eigenstates $|0\rangle$ and $|1\rangle$. Then, the binary array $\bar{\mathbf{q}}$ minimizing $f(\mathbf{q})$ is related to the ground-state, $|\Psi\rangle$, of $\hat{H}_Q$ though

$$|\Psi\rangle = \otimes_{\mathrm{i}=1}^{n} |\bar{q}_{\mathrm{i}}\rangle \;. \tag{1.7}$$

Notably, $\hat{q}_{\mathrm{i}}$ operators can be recast as $\frac{\hat{\mathbb{1}}-\hat{\sigma}_{\mathrm{i}}^{(z)}}{2}$, causing Eq. (1.6), up to an irrelevant additive constant, to be equivalent to Eq. (1.4). Hence, by performing adiabatically a quantum anneal having as problem Hamiltonian $\hat{H}_Q$, the final state will be its ground-state, $|\Psi\rangle$. Consequently, to read off the sought solution to the integer binary problem, it will be sufficient to measure all qubits along the z direction. This concludes that the solution to a Quadratic Unconstrained Binary Optimization problem, such as minimizing $f(\mathbf{q})$, can be successfully retrieved by resorting to adiabatic computing protocols leading to Ising-like problem Hamiltonians, as the one in Eq. (1.4). Naturally, not all problems are natively formulated as QUBOs and these cases require some additional manipulation, as outlined in Appendix A.

To better illustrate the process of deriving a QUBO formulation for an integer programming optimization task, the following section provides an example based on the Knapsack problem.

1.2.1.1 QUBO formulation of the Knapsack problem

Suppose we have a bag with a maximum weight capacity, W, and a set of N objects, each characterized by a specific value, v_i, and a positive integer weight w_i. The Knapsack problem consists in determining how to select the objects to be packed into the bag such that (i) the total weight of the objects does not exceed the bag's weight limit, W, and (ii) the total value of the packed objects is maximized. Concretely, by describing the content of the bag using a binary string $\mathbf{x} = (x_1, x_2, \dots, x_N)$, where $x_{\mathrm{i}} = 1$ ($x_{\mathrm{i}} = 0$) if the i-th object is packed, the Knapsack problem can be expressed as

$$\mathbf{x}_{\mathrm{i}}^{*} = \underset{\mathbf{x}}{\mathrm{argmax}} \sum_{\mathrm{i}=1}^{N} x_{\mathrm{i}}\, v_i \quad \text{subject to} \quad \sum_{\mathrm{i}=1}^{N} x_{\mathrm{i}}^{*}\, w_i \leq W\,. \tag{1.8}$$

The inequality constraint appearing in Eq. (1.8) can be manipulated into an equality constraint at the cost of introducing W additional binary variables — known as *slack* variables — $\mathbf{s} = (s_1, \dots, s_W)$, where $s_n = 1$ signals that the total weight of the bag corresponds to n. In this setting, it is necessary to impose that *exactly* one slack variable is active and that the total weight packed in the bag is in agreement with the active slack variable, leading to a reformulation of the Knapsack problem in form of

$$\mathbf{x}^*, \mathbf{s}^* = \underset{\mathbf{x},\mathbf{s}}{\mathrm{argmax}} \sum_{\mathrm{i}=1}^{N} x_\mathrm{i}\, v_\mathrm{i} \quad \text{subject to} \quad \begin{cases} \sum_{n=1}^{W} s_n^* = 1 \\ \\ \sum_{\mathrm{i}=1}^{N} x_\mathrm{i}^*\, w_\mathrm{i} = \sum_n n\, s_n^* \end{cases} \tag{1.9}$$

Yet, Eq. (1.9) does not describe a QUBO problem, as the binary arrays $\mathbf{x}$ and $\mathbf{s}$ are subject to constraints. However, it is possible to show that, finding the solution to Eq. (1.9) is equivalent to minimizing

$$\mathbf{x}^*, \mathbf{s}^* = \underset{\mathbf{x},\mathbf{s}}{\mathrm{argmin}} - \mathcal{H}_v[\mathbf{x}] + \mathcal{H}_{C_1}[\mathbf{s}] + \mathcal{H}_{C_2}[\mathbf{x}, \mathbf{s}]\,, \tag{1.10}$$

with

$$\begin{aligned} \mathcal{H}_v[\mathbf{x}] &= A_v \sum_{\mathrm{i}=1}^{N} x_\mathrm{i}\, v_\mathrm{i} \\ \mathcal{H}_{C_1}[\mathbf{s}] &= A_1 \left(\sum_{n=1} s_n - 1 \right)^2 \\ \mathcal{H}_{C_2}[\mathbf{x}, \mathbf{s}] &= A_2 \left(\sum_{\mathrm{i}=1}^{N} x_\mathrm{i}\, w_i - \sum_{n=1}^{W} n\, s_n \right)^2 \,. \end{aligned} \tag{1.11}$$

In the limit of $A_1, A_2 \gg A_v > 0$, the binary sequence minimizing (1.10) satisfies both constraints in Eq. (1.9) and, by *minimizing* $-\sum_{\mathrm{i}=1}^{N} x_\mathrm{i}\, v_\mathrm{i}$, adapts the packed content so to maximize the overall value carried in the bag. Further, Eq. (1.10) is a QUBO problem, as the binary variables are not subject to a *hard constraints.* Indeed, configurations $(\mathbf{x}, \mathbf{s})$ not meeting the requirements correspond to $\mathcal{H}_{C_1}[\mathbf{s}] > 0$ or $\mathcal{H}_{C_2}[\mathbf{x}, \mathbf{s}] > 0$ and thus do not minimize the objective function.

As discussed in the previous section, the solution to QUBO problems can be encoded in the ground-space of the Ising Model. As an instance, the problem Hamiltonian required to solve Eq. (1.10) operates on a $N \times W$-dimensinal Hilbert space and is given by

$$\hat{H}_P = -\hat{H}_v + \hat{H}_{C_1} + \hat{H}_{C_2}\,, \tag{1.12}$$

with

$$\begin{aligned}
\hat{H}_v &= A_v \sum_{\mathrm{i}=1}^{N} \hat{q}_{\mathrm{i}}\, v_i \\
\hat{H}_{C_1} &= A_1 \left(\sum_{n=1}^{W} \hat{s}_n - 1 \right)^2 \\
\hat{H}_{C_2} &= A_2 \left(\sum_{\mathrm{i}=1}^{2} \hat{q}_{\mathrm{i}} - \sum_{n=1}^{W} n\, \hat{s}_n \,. \right)^2
\end{aligned} \tag{1.13}$$

In the previous, the operators $\hat{q}_{\mathrm{i}}$ and $\hat{s}_n$ are, respectively, projectors on active x_{i} and s_n states and can be linearly mapped to standard Pauli operators.

In the next section, we briefly present the main features of the D-Wave superconducting analog device we adopted to derive results in Chapter 3.

1.2.2 Embedding of QUBOs on the D-Wave chip

In this section, we will outline how QUBOs can be implemented on the D-Wave chip (see Appendix B for a detailed description of the platform).

The quantum anneal implemented on this device consists in a slow transition from the driver Hamiltonian $\hat{H}_D = -\sum_{\mathrm{i}} \hat{\sigma}_{\mathrm{i}}^{(x)}$, having non-degenerate ground-state given by $\otimes_{\mathrm{i}} \frac{|0\rangle_{\mathrm{i}} + |1\rangle_{\mathrm{i}}}{\sqrt{2}}$, to the problem Hamiltonian, taking the form

$$\hat{H}_P = \sum_{\mathrm{i}} \mu_{\mathrm{i}} \hat{\sigma}_{\mathrm{i}}^{(z)} + \sum_{(\mathrm{i},\mathrm{j}) \in \mathcal{G}} \hat{\sigma}_{\mathrm{i}}^{(z)}\, \hat{\sigma}_{\mathrm{j}}^{(z)} \,. \tag{1.14}$$

Due to the technological implementation of qubits, bringing two of them to interact requires their topological vicinity. More formally, only couplings between pairs of qubits belonging to a well-defined graph, $\mathcal{G}$, can be established. Fig. 1.1 illustrates one of the possible architectures that can be implemented on the D-Wave platform (i.e., the so-called Chimera topology). The specific structure of $\mathcal{G}$ depends solely on fabrication details. Equation (1.14), while seemingly equivalent to Eq. (1.4), exhibits a key difference: due to the limited connectivity of physical qubits, directly embedding arbitrarily dense problems is usually not feasible and thus it requires additional manipulations. In what follows, we will refer to problem

Hamiltonians in the form of Eq. (1.4) as *logical Hamiltonians*, involving *logical qubits* that can be arbitrarily connected.

To effectively reproduce a logical Hamiltonian on a device with limited connectivity, there are several viable options. The first, consists in employing the so-called LHZ architecture [48], where the *relative orientation* of pairs of logical qubits are regarded as the relevant degrees of freedom and as such they are represented on the quantum device. This comes at the expense of introducing additional consistency constraints, coming as 4-body interaction terms that are prone to be ancillarized to match the form of Eq. (1.14).

Yet another possibility consists in mapping each logical qubit, $|Q_a\rangle$, to a chain of ferromagnetically coupled physical qubits, $\{q_1^{(a)}, \ldots, q_n^{(a)}\}$. Exploiting this overhead, individual couplings involving the logical qubits $|Q_a\rangle$ and $|Q_b\rangle$ are realized by coupling two arbitrary physical qubits rep-

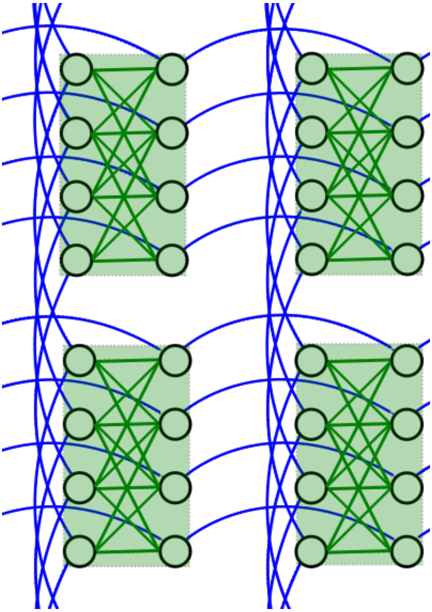


Figure 1.1: A cropped view of two unit cells of a Chimera graph. Qubits are arranged in 4 unit cells (translucent green squares) interconnected by external couplers (blue lines). Reproduced from Ref. [47].

resenting them, e.g., $|q_{\mathrm{i}}^{(a)}\rangle$ and $|q_{\mathrm{j}}^{(b)}\rangle$. Implementing this strategy requires careful consideration of two aspects: the mapping of logical qubits to physical ones, known as *minor embedding*, and the selection of the strength of the ferromagnetic couplings among physical qubits representing the same logical one, referred to as *chain strength*. As for the first aspect, the problem of finding an embedding to a problem while minimizing the overhead in quantum resources employed is on its own account an NP-complete problem [49]. Nonetheless, various protocols and algorithms exist to determine a feasible minor embedding [48, 50], even when the problem Hamiltonian cannot be fully represented on the quantum chip [51, 52].

Regarding the second aspect, since the chain strength dictates the tendency of each physical qubit to align with others representing the same logical qubit, an infinitely strong chain would in principle be ideal. Indeed, this would allow interpreting unambiguously the outcome, i.e., without resorting to major-vote criteria that would otherwise introduce biases in the outcome distribution [53]. Unfortunately, the limit of infinite chain strength incurs the risk of losing the resolution of low-energy levels, as the intensity of couplings that can be established on the D-Wave platform is limited. For this reason, the D-Wave software implements an auto-scaling feature that adapts all couplings (including chain strengths) to fall within the range $[-1, +1]$ [54].

Once a suitable minor embedding and corresponding chain strengths have been selected, the logical Hamiltonian is effectively reformulated in the form of Eq. (1.14), making the system ready for the annealing protocol to initiate. Ideally, the annealing protocol should proceed slowly enough for the dynamics to be adiabatic, but in practice, the annealing time is limited by the coherence time of qubits. Indeed, the annealing time, τ, on the D-Wave Advantage platform is comprehended between $1\mu s$ and $2ms$, while the time-dependent Hamiltonian governing the system is given by

$$\hat{H}(t = s\,\tau) = A(s)\,\hat{H}_0 + B(s)\,\hat{H}_P\,. \tag{1.15}$$

The scheduling functions $A(s)$ and $B(s)$ in the previous equation satisfy $A(s = 0) \gg B(s = 0)$ and $A(s = 1) \ll B(s = 1)$, as depicted in Fig. 1.2. It is possible to customize the annealing schedule by combining the native schedules, $A(s)$ and $B(s)$, with a user-defined function $f(s)$ satisfying $f(s = 0) = 0$ and $f(s = 1) = 1$. This approach modifies the annealing process by adopting the new schedules $A(f(s))$ and $B(f(s))$. Such customization is particularly beneficial when the approximate size and po-

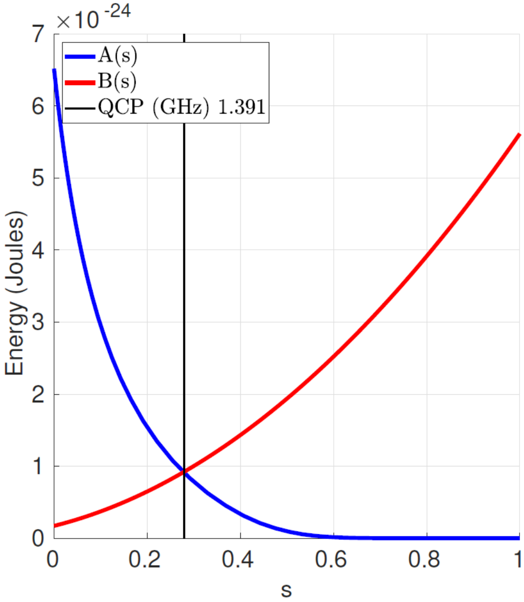


Figure 1.2: Annealing functions $A(s)$, $B(s)$. Annealing begins at $s = 0$ with $A(s) \gg B(s)$ and ends at $s = 1$ with $A(s) \ll B(s)$. Data shown are representative of D-Wave 2X systems. Reproduced from Ref. [55].

sition of avoided crossings along the annealing protocol are known. In these cases, slowing down the annealing process at avoided crossings lowers the probability of having diabatic transitions taking place.
This concludes the introduction of the main features of the quantum platform we used to derive the results in Chapter 3. In the next section, we review the quantum information theory concepts that underpin the contents of Chapter 5.

1.3 Entanglement

In a famous “gedanken Experiment” proposed by Einstein, Podolsky, and Rosen [56], it was proposed to consider a wave function defined on two spatially separated domains A and B. Then, supposing to perform a measurement in A, it was shown that a *spooky action at large distance* was also forcing the collapse of the wave function localized in B, leading to the breaking of the local-realism principle. This opened the door to alternative theories to quantum mechanics, such as the Local Hidden Variables Model (LHVM), which aimed at completing quantum mechanics with additional variables in a way that locality and realism were restored. For the LHVM to be consistent, it was necessary to satisfy a set of constraints known as Bell inequalities [57], which curiously happen to fail when specific quantum states — that were successully experimentally prepared — are considered. This invalidates the LHVM and points to the conclusion that the spooky action at large distace is a consequence of a fully legitimate feature that only quantum systems possess. This feature was later on baptized as *quantum entanglement* and opened new fields of research, both theoretical and technological. It is worth stressing that a violation of the Bell inequalities irrefutably signals that the considered quantum state is entangled; however, the inverse does not hold. Indeed, there are entangled quantum states for which the Bell inequalities are verified.

Entangled states among domains A and B are, by definition, not separable between these domains. Concerning the definition of separability, a pure state $|\Phi\rangle$ belonging to $\mathcal{H}_\mathrm{A} \otimes \mathcal{H}_\mathrm{B}$ [1]is separable if there exist two states, $|\phi_\mathrm{A}\rangle$ and $|\phi_\mathrm{B}\rangle$, such that $|\Phi\rangle = |\phi_\mathrm{A}\rangle \otimes |\phi_\mathrm{B}\rangle$.

Alternatively, if $|\Phi\rangle$ is separable, it factorizes the expectation value of any observable taking the form $\hat{T}_\mathrm{A} \otimes \hat{T}_\mathrm{B}$ into $\langle \hat{T}_\mathrm{A} \otimes \mathbb{1}_\mathrm{B} \rangle_\Phi \times \langle \mathbb{1}_\mathrm{A} \otimes \hat{T}_\mathrm{B} \rangle_\Phi$, where $\hat{T}_\mathrm{A}$ and $\hat{T}_\mathrm{B}$ are arbitrary hermitian operators acting, respectively, on $\mathcal{H}_\mathrm{A}$ and $\mathcal{H}_\mathrm{B}$. Differently, a mixed state $\hat{\rho}$, is separable if it can written as a convex sum of pure separable states, namely if

$$\hat{\rho} = \sum_n p_n \, \hat{\rho}_n \quad \text{with} \quad \forall\, n \begin{cases} \hat{\rho}_n = |\phi_n\rangle\langle\phi_n| \,, \\ |\phi_n\rangle \;:\; \text{is separable} \,. \end{cases} \tag{1.16}$$

In the following sections, we elaborate more on the technological applications of entanglement and the techniques that are capable of discriminating entangled states from separable ones.

---

[1] The tensor product of two spaces $\mathcal{H}_\mathrm{A}$ and $\mathcal{H}_\mathrm{B}$, respectively having $\left\{ \left|\phi_1^{(\mathrm{A})}\right\rangle, \left|\phi_2^{(\mathrm{A})}\right\rangle, \dots \right\}$ and $\left\{ \left|\phi_1^{(\mathrm{B})}\right\rangle, \left|\phi_2^{(\mathrm{B})}\right\rangle, \dots \right\}$ as base, is the Hilber space spanned by the states $\left|\phi_n^{(\mathrm{A})} \phi_m^{(\mathrm{B})}\right\rangle$, with $n = 1, \dots, \dim(\mathrm{A})$ and $m = 1, \dots, \dim(\mathrm{B})$.

1.3.1 Entanglement as a resource

In this section, we introduce some of the most relevant applications that use entanglement as their main ingredient.

In quantum teleportation [58–60], a two-qubit system is initialized as a maximally entangled state (i.e., a Bell state). After the initialization, these two qubits are shared between the two parties, Alice and Bob, establishing a communication channel between them. Suppose that Alice has in her possession an additional single-qubit state $|\psi\rangle$ and wants Bob to have it. It is sufficient for Alice to measure her two qubits against the orthogonal base defined by the four Bell states and to communicate *classically* the outcome of her measurement to Bob, who, depending on it, will adjust his state by applying a specific qubit rotation. At the end of the protocol, Alice has a collapsed state with no memory of the initial state she sent to Bob, while Bob's state is exactly the one Alice wanted him to have.

Entanglement finds a significant application in cryptography, where many protocols are initiated with two parties, Alice and Bob, sharing a key for encrypting and decrypting messages over a potentially insecure channel. Ensuring the security of this shared key is essential. One approach, introduced in [61], involves initializing a two-qubit state into one of the four Bell states and sending a qubit to each party. In this setup, measurements performed by Alice are perfectly correlated (or anticorrelated) with those of Bob. If the shared state is partially measured by an eavesdropper, it would collapse to a separable state. Thus, a certification test that would grant Alice and Bob the integrity of their shared key consists in verifying whether Bell inequalities are violated. If they are, the shared state is entangled, certifying that no eavesdropper tampered with it. Then, the key is trusted and can be safely used.

Entanglement enables quantum-dense coding [62], a protocol that allows the communication of two classical bits using a single qubit, apparently challenging the Holevo bound [63], which constrains the *accessible* classical information encoded in a n-qubit system to be at most n bits. The process begins with the establishment of a communication channel between Alice and Bob, who share a maximally entangled two-qubit state. Next, Alice selects one of four possible messages — 00, 01, 10, 11 — and applies the corresponding rotation — $\hat{\mathbb{1}}$, $\hat{\sigma}_x$, $\hat{\sigma}_y$, $\hat{\sigma}_z$ — to the qubit in her possession. She then sends this qubit to Bob, who decodes the message by identifying the resulting two-qubit state as one of the four Bell states.

This protocol demonstrates that Alice can communicate two classical bits to Bob by transmitting a single qubit. However, it is worth noticing that this does not violate the Holevo bound, as the shared state was *a priori* entangled. Indeed, a deeper analysis reveals that the algorithm requires 1 ebit — 1 unit of bipartite entanglement — and the transmission of 1 qubit to communicate 2 classical bits.

While not exhaustive, these examples highlight the wide range of applications involving entanglement as a key element. In the following section, we consider a technique that can be easily implemented in experimental setups to qualify a state as entangled.

1.3.2 Entanglement witnessing

In principle the quantity of bipartite entanglement a state possesses is related to its Von Neumann entropy, logarithmic negativity, entanglement cost, and other measures [64,65]. Unfortunately, prior to quantifying these measures, it is usually necessary to perform full or partial tomography [66] of the state under consideration, which is known to be a costly operation, even if more efficient and sophisticated tomography techniques are a current field of research [67]. Hence, quantifying entanglement can be prohibitively expensive.

Similarly, even with full availability of the system's state, determining whether a state is separable based on Eq. (1.16) is generally hard. However, as we shall discuss later on, entanglement witnessing [68–72] is an approach allowing to detect *some* entangled states through measurements of purposely chosen observables.

To fix the notation, the tested states belong to $\mathcal{H}_{\mathrm{A}} \otimes \mathcal{H}_{\mathrm{B}}$ and can be either separable or entangled between regions A and B. An entanglement witness is an observable $\hat{W}$ acting both on A and B regions such that (i) it acquires a non-negative expectation value on any separable state,

$$\forall \ |\Phi\rangle_{\mathrm{sep}} \in \mathcal{H}_{\mathrm{A}} \otimes \mathcal{H}_{\mathrm{B}} \Rightarrow \langle \hat{W} \rangle_{\Phi_{\mathrm{sep}}} \geq 0 \,, \tag{1.17}$$

and (ii) there exists at least one entangled state $|\Phi\rangle$ featuring $\langle \Phi | \hat{W} | \Phi \rangle < 0$. In mathematical terms, the first condition is equivalent to asking $\hat{W}$ to be block positive semi-definite when considering the bipartition between domains A and B, while the second excludes completely positive observables from serving as entanglement witnesses. There are two crucial observations concerning Eq. (1.17). First, it is not sufficient to have $\langle \hat{W} \rangle_\Phi \geq 0$ to guarantee the separability of $|\Phi\rangle$. Second, it is sufficient to measure $\langle \hat{W} \rangle_\Phi < 0$ to qualify $|\Phi\rangle$ as entangled. Under this circumstance, we will say that $\hat{W}$ detected $|\Phi\rangle$ as being an entangled state.

To exemplify this, consider the entanglement witness defined on a two-qubit system

$$\hat{W}_2 = \hat{\mathbb{1}} - \hat{\sigma}_1^{(x)}\, \hat{\sigma}_2^{(x)} - \hat{\sigma}_1^{(z)}\, \hat{\sigma}_2^{(z)} \,. \tag{1.18}$$

To prove that $\langle \hat{W}_2 \rangle_{\Phi_{\text{sep}}} \geq 0$, we build $|\Phi\rangle_{\text{sep}}$ as

$$|\Phi\rangle_{\text{sep}} = |\Phi_1\rangle \otimes |\Phi_2\rangle \quad \text{with} \quad \begin{cases} |\Phi_1\rangle = \sin\theta_1\,|0\rangle + \cos\theta_1 e^{i\phi_1}\,|1\rangle \\ |\Phi_2\rangle = \sin\theta_2\,|0\rangle + \cos\theta_2 e^{i\phi_2}\,|1\rangle \end{cases} \tag{1.19}$$

and we evaluate the associated expectation value of $\hat{W}_2$ as

$$\begin{aligned} \langle \hat{W}_2 \rangle_{\Phi_{\text{sep}}} &= 1 - \left(\cos^2\theta_1 - \sin^2\theta_1\right)\left(\cos^2\theta_2 - \sin^2\theta_2\right) + \\ &\quad - 2\sin\theta_1\cos\theta_1\sin\theta_2\cos\theta_2\big(\cos(\phi_1+\phi_2) + \cos(\phi_1-\phi_2)\big) \\ &\geq 1 - \left(\cos^2\theta_1 - \sin^2\theta_1\right)\left(\cos^2\theta_2 - \sin^2\theta_2\right) \\ &\quad - 4|\sin\theta_1\cos\theta_1\sin\theta_2\cos\theta_2| \end{aligned} \tag{1.20}$$

and consequently

$$\langle \hat{W}_2 \rangle_{\Phi_{\text{sep}}} \geq 1 - \cos(2\theta_a \pm 2\theta_b) \geq 0 \,. \tag{1.21}$$

Then, considering the Bell state given by $|\Psi\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}}$, we recover $\langle \Psi | \hat{W}_2 | \Psi \rangle = -1$, and this corroborates the fact that $|\Psi\rangle$ is an entangled state. On the contrary, when considering another Bell state, say $|\Psi'\rangle = \frac{|01\rangle + |10\rangle}{\sqrt{2}}$, the entanglement witness would give as expectation value $\langle \Psi' | \hat{W}_2 | \Psi' \rangle = 1$, and thus it would fail in detecting $|\Psi'\rangle$ as being entangled, even if it is a maximally entangled state!

As we have seen, a witness typically can pinpoint only a fraction of entangled states. Then, it is reasonable to wonder whether some entanglement witnesses may be "better" than others. To formalize this discussion, let us define $\mathcal{S}_W$ the set of states that are detected as entangled by the witness $\hat{W}$. Then, we will say that another entanglement witness $\hat{W}'$ is *finer* than $\hat{W}$ if $\mathcal{S}_{W'} \supset \mathcal{S}_W$. In other words, if $\hat{W}'$ detects all the entangled states that $\hat{W}$ does, plus additional ones. In the case where is not possible to find a finer witness that $\hat{W}'$, this is regarded as *optimal*.

In a more rigorous language, $\hat{W}'$ is optimal if it does not exist any positive semi-definite operator $\hat{P}$ such that $\hat{W}'' = \hat{W}' - \lambda\hat{P}$ is a valid entanglement witness for some strictly positive λ [73, 74]. In Chapter 5, we will address the problem of optimizing entanglement witnesses in the special context of lattice gauge theories where local symmetries pose additional challenges.

Chapter 2

Soft-matter physics applications

Proteins stand as a prototypical example of soft matter objects whose study is computationally challenging. As such, it offers the opportunity of devising relevant applications of quantum technologies to real-world problems.

A protein's sequence typically comprises 50 to 2,000 amino-acid units, each of which can be chosen in a set of 20 possible chemical identities [75]. Despite the remarkable variety of sequences available a priori, only a fraction of it is found in Nature. This obervation already poses a concern regarding the circumstances allowing an amino-acid sequence to fold into a stable configuration. We will refer to *protein-like* sequences as those that fold into a well-defined and stable conformation — *native state* — within timescales ranging from milliseconds (ms) to microseconds (μs). Predicting the native state given the amino acid sequence, known as the problem of protein folding, is highly challenging. Indeed, this problem has no deterministic solution in polynomial time (i.e., it falls in the NP class of computational complexity).

Another important challenge arise from reverse engineering the folding process to predict which protein-like sequences — if they exist — give rise to specific conformations. This challenge, referred to as the *inverse* or *protein design* problem, holds significant promise for real-world applications and falls within the category of NP-complex problems. In Section 2.1 we provide a detailed introduction to the topic of protein design.

Another intriguing research line involves sampling simultaneously amino-acid sequences and backbone conformations based on a given distribution. Indeed, this offers the opportunity of gaining insight into the statistical mechanical features of proteins. Unfortunately, from a computational perspective, sampling compact backbone configurations that satisfy topological constraints (e.g., continuity and self-avoidance requirements) is already a computationally demanding endeavor, as Markov chain techniques typically lead to high rejection rates and consequent long autocorrelation times. Additional detail concerning polymer sampling is provided in Section 2.2.

2.1 Protein folding and design

Before introducing the protein design problem, it is propedeutic to outline first the common models that have been adopted to study statistical mechanical features of proteins (see Section 2.1.1) and to consider the conditions under which an amino-acid sequence is protein-like (see Section 2.1.2). In Section 2.1.3 we formalize in mathematical terms the problem of protein design, and in Section 2.1.4 we report relevant related work.

2.1.1 Protein models

In this section, we overview the most relevant models that have been adopted to study proteins and polypeptides. The main features distinguishing different protein models are (i) the choice of the spatial degree of freedom, which can be discrete or continuous, (ii) the solvent degrees of freedom, which can be implicit or explicit, (iii) the chosen resolution on the degrees of freedom associated with the polymeric chain, which can have atomistic detail or can be coarse-grained [76, 77].

Although there are research lines considering full-atomistic detail and explicit solvent degrees of freedom [78–80], many applications allow for avoiding the computational effort required to simulate such models in full detail. For this reason, coarse-grained units are employed to describe the polymer, while the solvent degrees of freedom are integrated out. This results in an effective potential governing the remaining coarse-grained degrees of freedom. The coarse-graining technique, while substantially reducing the computational cost of simulating proteins, has to be handled carefully. Indeed, unwise choices of the coarse-grained degrees of freedom can lead to models that fail to reproduce relevant physical phenomena, leading to incorrect results. In what follows, we reference typical models adopted in the study of polymers.

Examples of coarse-grained models realized in the spatial continuum are given by those employed in [81–83], where "patchy monomers" are shaped as hard-spheres having specific sites on their surface — called patches — that interact anisotropically. Covalent bonds are realized as spring-like interactions among adjacent monomers; depending on the model, these binding forces are applied to anchors placed on the surface of the monomer (freely-rotating chain) or to their center of mass (freely-joint chain).

Another class of models employs coarse-grained units constrained to lie on the vertices of a lattice. The most simplified model involves only two kinds of elements: those that are attracted by water — hydrophilic (P) — and those that are repelled by it — hydrophobic (H) [84–88]. The rationale underlying this choice is guided by the observation that hydrophobic-hydrophilic forces are mainly responsible for protein's collapse into a folded configuration. Indeed, HP heteropolymer sequences typically fold into conformations having H residues buried in the core, and P chemical entities lying on the surface of the globule, exposed to the solvent. Thus, by integrating out the solvent degrees of freedom, the effective potential provides an attractive force among H-type monomers. For this reason, a typical approach consists in accounting only for the H-H non-covalent interactions to contribute to the effective potential, though alternative parametrizations have also been explored [89].

Another coarse-grained model relies on 20 types of amino-acids, where residues interact through a potential that is statistically inferred from large protein data banks [90]. This stands as the so-called quasichemical approximation and is employed in many studies concerning proteins (e.g., [85, 91–99]).

Both the HP and the 20-letter lattice models introduced so far associate to the oriented sequence $S = \{s_1, \ldots, s_N\}$ disposed on the structure Γ the energy

$$E(\Gamma, S) = \sum_{\mathrm{i}>\mathrm{j}=1}^{N} C_{\mathrm{ij}}\,\Delta_{\mathrm{ij}}\,\varepsilon(s_{\mathrm{i}}, s_{\mathrm{j}})\,, \tag{2.1}$$

Here, s_{i} is the chemical identity of the i-th amino-acid of the sequence, $\varepsilon(s_{\mathrm{i}}, s_{\mathrm{j}})$ are the entries of a suitable $D \times D$ symmetric matrix — D being the size of the amino-acids chemical alphabet employed— and $\boldsymbol{C}(\Gamma)$ is the contact map of the conformational state Γ, with entries $C_{\mathrm{ij}}(\Gamma)$ equal to 1 if the amino-acids i and j establish a non-covalent contact and zero otherwise. The Δ_{ij} term serves to weight the contribution to the potential energy given by monomers establishing a non-covalent interaction depending on their distance. In models accounting only for nearest-neighbor interactions, Δ_{ij} is 1 if i and j are first-neighbors on the lattice, and it is zero otherwise.

2.1.2 Funnel theory and protein folding

In this section, we introduce the concept of protein-like sequences, along with the features characterizing the cooperative folding process they undergo.

First: do polymers fold into their native structures following a definite pathway towards it? To answer this question, we start by considering how the dimensionality of phase space scales with the number of degrees of freedom associated with the polymer chain. Consider, for example, a polypeptide chain comprising L amino-acids. To define the spatial arrangement of this amino-acid chain it is sufficient to specify the relative orientation of consecutive covalent bonds. This amounts to specifying the angles between adjacent bonds ($L-2$ real-valued parameters) and the dihedral angles formed by planes comprising consecutive adjacent covalent bonds ($L-3$ parameters). Assuming that each parameter can assume $n > 1$ stable configurations, the lower bound on the number of conformations available to the protein is n^{2L-5} and thus grows exponentially with the length of the polypeptide chain. Now, assuming that the folding process consists in uniformly sampling this vast space of configurations to find the native structure, the time required for a protein to fold correctly would exceed the age of the Universe (even if one could sample a configuration on a timescale of picoseconds). This contrasts with the behavior of real proteins, whose folding time ranges from milliseconds to microseconds. These considerations are at the basis of the paradox posed in 1969 by Levinthal [100,101], which stimulated research efforts towards revealing the processes that allow the folding of polymers.

In principle, to find a solution to the Levinthal paradox, it would be necessary to address directly the folding dynamics of polymeric chains. Luckily, there is experimental evidence [103, 104] suggesting that the native state associated with a polymer is the one minimizing the Helmholtz free energy, thus circumventing the necessity of simulating its dynamics. This observation was formalized as the "thermodynamic hypothesis" and was also successfully reproduced in Monte Carlo simulations involving a simplified lattice model of proteins [105].

Furthermore, there was evidence indicating that cooperative mechanisms guide the polymeric chains toward their native state, thereby avoiding uniform sampling of configurations and significantly reducing the time required to achieve the native state compared to Levinthal's prediction. Qualitatively, when the correct folding of a protein's segment has a beneficial effect on the folding process of the rest of the amino acid chain, the protein is said to fold cooperatively.

To better define these cooperative mechanisms, Wolynes and coworkers [106] proposed that the energy landscape associated with the folding process of protein-like amino-acid sequences is a *funnel*, namely it is a

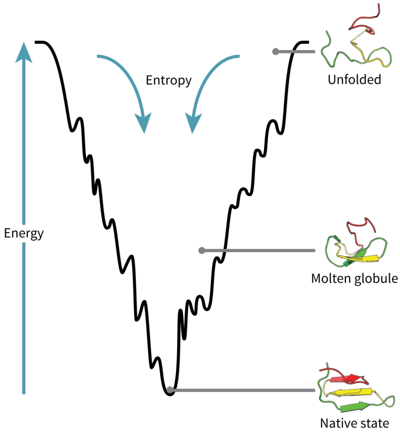


Figure 2.1: Funnel energy landscape (reproduced from Ref. [102]).

compromise between a smooth and a rough surface (as the one depicted in Fig. 2.1). The ripples in the free energy landscape are associated with rapid and local conformational changes, while the smooth component takes care of more complex and slower rearrangements of the backbone.

Let $n(\Gamma)$ be a similarity measure between the backbone configuration Γ and the native one, so that $n(\Gamma) = 1$ if Γ is the native configuration and $0 < n(\Gamma) < 1$ otherwise. Then, the energy of a n-similar state is assumed to be distributed as

$$P_n(E) = \mathcal{N} e^{-\frac{[E-\bar{E}(n)]^2}{2[\Delta E(n)]^2}} , \tag{2.2}$$

where $\bar{E}(n)$ is modeled to decrease as n approaches 1, while $\Delta E(n)$ sets the order of magnitude of energy fluctuations arising from local rearrangements of the backbone. With the number of n-similar conformational states being $\Omega(n) = e^{\frac{S_0(n)}{k_B}}$, the joint probability of having a configuration that is n-similar and has energy E is

$$P(n, E) = \mathcal{N} e^{\frac{S_0(n)}{k_B}} e^{-\frac{[E-\bar{E}(n)]^2}{2[\Delta E(n)]^2}} e^{-\beta E} . \tag{2.3}$$

The most probable energy of a n-similar conformation can be found by solving $\partial_E P(n, E)\big|_{E_{\text{m.p.}}} = 0$, leading to

$$E_{\text{m.p.}} = \bar{E}(n) - \beta \left[\Delta E(n)\right]^2 , \tag{2.4}$$

so that the saddle-point approximation of Eq. (2.3) is

$$P(n) \simeq \mathcal{N}\, e^{\frac{S_0(n)}{k_B} + \frac{1}{2}\beta^2[\Delta E(n)]^2 - \beta\bar{E}(n)} , \tag{2.5}$$

which corresponds, up to an additive offset value, to the free energy profile given by

$$F(n) \simeq \bar{E}(n) - \frac{1}{2}\beta\Delta E(n) - TS_0(n) . \tag{2.6}$$

At high temperatures, when the entropic term dominates, the system will preferably populate conformations that are largely dissimilar to the native state (i.e., the minimum of F is located at $n \simeq 0$). On the contrary, at low temperatures, the energetic term favors the population of conformations that are similar to the native state (i.e., the minimum of F is located at $n \simeq 1$). An intermediate value of T corresponds to a free energy landscape with two local minima: one close to $n = 0$ and one to $n = 1$. The basic thermodynamic assumption for a protein to fold into its native state is having as global minima the one located close to $n = 1$.

This seminal work shed some light on understanding the statistical and dynamical properties of protein folding; more importantly, it clarified the requirements on the free energy landscape that allow a protein to reach its native state. Later investigations [76, 107, 108] aimed at finding patterns that underlie protein-like amino-acid sequences based on simplified protein models. Up to this point, we have discussed properties that characterize protein-like sequences for which the folding problem is well-posed. In what follows, we formulate the protein folding problem in a manner that well-suits the application we develop in Chapter 3.

Protein folding: Let us suppose that S is known to be a protein-like sequence with respect to a given energy model $E(S, \Gamma)$. Then, owing to the thermodynamic hypothesis, solving the direct protein folding problem involves finding the conformer(s) Γ with the largest occupation probability in canonical equilibrium,

$$\begin{aligned} P_{\mathrm{opt}}(\Gamma|S) &= \max_{\Gamma} \frac{e^{-\beta E(\Gamma,S)}}{\sum_{\Gamma'} e^{-\beta E(\Gamma',S)}} \\ &\equiv \max_{\Gamma} e^{-\beta(E(\Gamma,S)-F(S))} \geq p_{\mathrm{fold}} \end{aligned} \tag{2.7}$$

where β is the inverse thermal energy in physiological conditions and $F(S) = -\frac{1}{\beta} \ln \sum_{\Gamma'} e^{-\beta E(\Gamma',S)}$ is the free energy of sequence S, which involves the sum over the possible conformational states. As S is assumed to be a foldable polypeptide chain, such as naturally-occurring proteins, it is characterized by the thermodynamic stability of the state Γ maximizing Eq. (2.7), i.e., $P_{\mathrm{opt}}(\Gamma|S) > p_{\mathrm{fold}}$, where p_{fold} is a suitable threshold, typically larger than 0.5. In this case, Γ is the native state of S.

In the following section, we turn our attention to the inverse problem of protein folding, which consists in finding protein-like sequences that have a target conformation as their native state.

2.1.3 Engineering protein-like sequences folding in a target configuration

As we have seen in introducing the problem of protein folding, only a small fraction of sequences have the right properties allowing cooperative folding, which means that the majority of sequences do not identify as protein-like.

Analogously, it is tempting to ask whether all structures allow for a designing sequence (i.e., a protein-like sequence that quickly reaches the desired structure, resulting in a thermally stable conformation), and, if not the case, to investigate the features of designable structures.
Results derived by considering simplified lattice models indicate that only a small fraction of structures are designable [109–112] and that, among these designable structures, a minority is supported by a large number of designing sequences, whereas most are associated with only a few possible designs.

Given these premises, how is it possible to associate a designable structure with its designing sequences? Despite the huge progress in the field of protein design — culminated in 2024 with the awarding of the Nobel Prize in Chemistry to Baker for his endeavors on the subject [113] — the techniques employed to solve this problem remain computationally expensive. To understand the origin of these computational costs, let us consider the most intuitive attempt to find designing sequences for a target structure. That consists in searching the sequences that, once installed on a target structure, minimize the energy functional. Yue and coworkers tested this hypothesis using an HP on-lattice model [114], obtaining that, contrary to expectations, the native states associated with such sequences do not correspond to the target structure. These results indicate that solving the inverse folding problem requires more than simply swapping monomers on a fixed, presumed native structure. Indeed, solving the inverse folding problem for a given target state Γ_T amounts to finding a sequence S, if any exists, such that

$$P_{\text{opt}}(S|\Gamma_T) = \max_S \; e^{-\beta(E(\Gamma_T,S)-F(S))} \geq p_{\text{fold}} \, . \tag{2.8}$$

Sequences that satisfy inequality (2.8) are said to design the target state Γ_T [111,115].

Thus, solving the design problem is equivalent to finding sequences that minimize the function

$$G(S) = E(\Gamma_T, S) - F(S) \; , \tag{2.9}$$

and satisfy the inequality in Eq. (2.8).

A key point is that the computational demands for solving the above rigorous physics-based formulations of the direct and inverse folding problems differ greatly. Solving the former involves, in principle, computing the energy of the sequence of interest over the entire set of physically viable conformational states. Conversely, to solve the design problem, it is not sufficient to compute the energy of all viable sequences on the given target structure, Γ_T. Indeed, the sequences that minimize the energy when mounted on Γ_T may fold into a different structure, Γ', with even lower effective energy, i.e., $E(\Gamma', S) < E(\Gamma_T, S)$ [111, 115]. Hence, solving the design problem involves two nested searches: one over the sequences and one over the structures [85, 111, 115, 116]. For this reason, much effort has been spent on finding practical schemes and approximations to curb the computational expenditure entailed by this problem [85, 111, 115–119].

2.1.4 Related work on protein design

A novel twist in this field of research has come from recent advancements in machine learning [120–122]. Relevant examples include the development of Bayesian learning design strategies [123], and the experimental validation of deep-learning models [124], including generative ones [125, 126]. On the one hand, these approaches provide elegant and valuable demonstrations of the striking extent to which sequence-structure correlations present in protein databases might be harnessed by empirical scoring functions for protein design. In perspective, such endeavors could emulate the breakthrough in the empirical solution to the direct folding problem, where deep neural networks now yield remarkably reliable predictions [127].

On the other hand, an inherent limitation of all such empirical methods is the lack of interpretability. In contrast, physics-based approaches, based on an explicit definition of the energy function $E(\Gamma, S)$, would enable abstracting principles applicable to more general contexts [128]. For this reason, the quest for computationally-amenable physics-based approaches to the protein design and related problems remains an active research avenue as well as a natural testbed for new computing hardware paradigms, including quantum computing [87, 88, 97, 99, 129], as we shall discuss in Chapter 3.

In physics-based schemes, the cornerstone notion is that the energy function $E(\Gamma, S)$ in Eq. (2.9) is the only theoretical ingredient needed to solve both the direct and the inverse folding problems. However, in detailed atomistic approaches, even a single computation of $E(\Gamma, S)$ would require extensive calculations, e.g., to integrate out the solvent degrees of freedom. Customarily, this prohibitively expensive computation is alleviated by resorting to coarse-grained models and implicit-solvent energy functions (see Section 2.1.1). At the same time, coarse-graining also tames the complexity of the sampling problem by smoothing the energy landscape and drastically reducing the number of conformational degrees of freedom [130–135]. Yet, reliably estimating the functional form and the parametrization of the effective energy $E(\Gamma, S)$ remains challenging.

In recent years, several quantum-inspired algorithms have been proposed to unveil protein sequence-structure relationship [87,88,97–99,136], compute protein folding pathways [44, 137], and more generally address equilibrium properties polymeric systems [138, 139]. By contrast, the protein design problem has been tackled by comparatively fewer attempts using algorithms designed for quantum hardware. Such pioneering efforts have relied on lattice protein models [84,105,109,140] because their discrete nature enables a straightforward mapping onto the quantum simulation hardware. In Ref. [129], the authors used a gate-based quantum algorithm to reshuffle a sequence to minimize its energy on a reference structure. In contrast, in Ref. [141] a quantum-annealing platform was used for an analogous objective. Both studies employed a simplified two-letter amino-acid alphabet and postulated the effective amino-acid interactions.

2.2 Polymer sampling

In the previous section, we introduced two relevant problems in biophysics, namely the protein folding and design problems. In protein folding the primary sequence is held fixed, while the polymer's conformation is adapted to minimize its free energy. In protein design, the search scans primary sequences to find the ones that exhibit a free energy minima in correspondence to the target configuration. This results in a nested search both in sequence and configuration spaces, which can be addressed asynchronously.

A completely different standpoint can be taken by allowing synchronous sampling of both primary sequence and backbone conformation according to an established statistical mechanical ensemble. Despite this approach is not expected to capture the intricate evolutionary processes likely to occur in Nature, it offers the opportunity to gain insight into the statistical properties of heteropolymers, such as average length, relative abundance of chemical species, average end-to-end point distance, and other relevant quantities.

In Section 2.2.1 we outline the close connection that exists between statistical polymer models and quantum field theories, which can be addressed by leveraging path-integral Monte Carlo techniques. Nonetheless, there are polymer models whose statistical properties are associated with path integrals that exhibit frequent fluctuations in sign, leading to poor convergence of stochastic methods such as Monte Carlo chains. Crucially, self-avoiding polymer models fall into this category. This stimulated new research efforts, including those harnessing quantum technologies. In Section 2.2.2 we discuss relevant details of a recently established mapping of the self-avoiding polymer model to a QUBO function that finds implementation on existing quantum platforms. Building on this formalism, our research aims to circumvent the limitations arising when incorporating chemical identities and their interaction into the model and when further restricting the ensemble of polymers to account only for structures displaying no isolated rings.

2.2.1 Polymer field theory

The pioneering work of Edwards [142] and De Gennes [143] established deep theoretical connections between the statistical mechanical models adopted to describe polymers and several quantum field theories.

As an illustrative example, we consider the *freely joint model* of polymers, where each covalent bond has a fixed length, b, and an orientation that is uniformly distributed and is statistically independent of the orientation of other covalent bonds. The partition function of a canonical ensemble populated by N freely joint covalent bonds, each characterized by its displacement vector $\mathbf{r}_i$, in three dimensions writes

$$\mathcal{Z} = \sum_{\mathbf{r}_1,\ldots,\mathbf{r}_N} \delta(|\mathbf{r}_1| - b) \ldots \delta(|\mathbf{r}_N| - b)\,. \tag{2.10}$$

This causes the end-to-end point vector distribution of the polymer, $\mathbf{R} = \sum_{\mathrm{i}} \mathbf{r}_i$, to be approximated by

$$\mathcal{P}(N, \mathbf{R}) = \left(\frac{3}{2\pi N b^2}\right)^{\frac{3}{2}} e^{-\frac{3\mathbf{R}^2}{2Nb^2}}\,. \tag{2.11}$$

Crucially, Eq. (2.11) is the solution of the diffusion equation given by

$$\left(\partial_N - \frac{b^2}{6}\nabla^2_{\mathbf{R}}\right) \mathcal{P}(N, \mathbf{R}) = 0\,. \tag{2.12}$$

Considering the Shrödinger equation of a free quantum particle, namely

$$\left(i\hbar\,\partial_t + \frac{\hbar^2}{2m}\nabla^2_{\mathbf{x}}\right)\,\Psi(t,\mathbf{x}) = 0\,, \tag{2.13}$$

it becomes apparent that applying the Wick rotation $N \to -it$ to Eq. (2.12) yields an equation that is equivalent to Eq. (2.13). Then, the statistical properties a freely joint chain can be encoded in the wave-function of a free quantum particle having mass $m = \frac{3\hbar}{b^2}$, where b is the length of covalent bonds. The potential of this mapping lies in the possibility of leveraging the powerful path integral formalism when dealing with classical models of polymers. Indeed, the probability of having a random walk joining two points, $\mathbf{x}_{\mathrm{i}}$ and $\mathbf{x}_{\mathrm{f}}$, results from considering all polymer paths joining $\mathbf{x}_{\mathrm{i}}$ and $\mathbf{x}_{\mathrm{f}}$ according to their statistical weight. Namely, the polymer "propagator" of a freely joint chain composed of N covalent bonds between points $\mathbf{x}_{\mathrm{i}}$ and $\mathbf{x}_{\mathrm{f}}$ is

$$\mathcal{P}(\mathbf{x}_{\mathrm{i}},\,\mathbf{x}_{\mathrm{f}},N) = \int_{\mathbf{x}_{\mathrm{i}}}^{\mathbf{x}_{\mathrm{f}}} \mathcal{D}\mathbf{r}\, e^{-\frac{3}{2b^2}\int_0^N \mathrm{d}s\, H[\mathbf{r}(s)]}\,, \tag{2.14}$$

where $H[r(s)] = \left(\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t}\right)^2$ and $\mathbf{r}(s)$ represents the polymer's conformation.

Deepening the connection between polymer physics and quantum mechanics, we now focus on deriving a *field* theory and the associated path integral representation associated with polymers. To do so, we write Eq. (2.12) as

$$\partial_N \left|P_N\right\rangle = -\hat{H}\left|P_N\right\rangle\,, \tag{2.15}$$

where $\hat{H}$ is the Hamiltonian that is represented in spatial coordinates by the operator $-\frac{b^2}{6}\nabla^2_{\mathbf{R}}$ and the state $\left|P_s\right\rangle$ encodes the probability distribution associated with the position of the s–th residue constituting the polymer chain. Building on this, it is possible to *evolve* a polymer state, $\left|P_0\right\rangle$, up to "time" N by resorting to

$$\left|P_N\right\rangle = e^{-\hat{H}N}\left|P_0\right\rangle\,, \tag{2.16}$$

which enables us to rewrite the polymer propagator in the form of

$$\mathcal{P}(\mathbf{x}_{\mathrm{i}},\mathbf{x}_{\mathrm{f}},N) = \left\langle P_N\right|e^{-\hat{H}N}\left|P_0\right\rangle\,. \tag{2.17}$$

A mathematical expedient to evaluate Eq. (2.17) involves introducing coherent fields of bosons, $\left|\psi\right\rangle$, that satisfy

$$\hat{a}_{\mathbf{x}}\left|\psi\right\rangle = \psi(\mathbf{x})\left|\psi\right\rangle \quad \text{and} \quad \left\langle\psi\right|\hat{a}^\dagger_{\mathbf{x}} = \left\langle\psi\right|\psi^*(\mathbf{x})\,, \tag{2.18}$$

where $\hat{a}^\dagger_{\mathbf{x}}$ and $\hat{a}_{\mathbf{x}}$ are, respectively, bosonic creation and annihilation operators. Adopting coherent states, $|\psi\rangle$, as a base for the Fock space, the resolution of the identity writes

$$\hat{\mathbb{1}} = \int \mathcal{D}\psi^* \mathcal{D}\psi \; |\psi\rangle\langle\psi| \, e^{-\psi^* \cdot \psi} \, , \tag{2.19}$$

where $\psi^* \cdot \psi$ is a shorthand for $\int \mathrm{d}\mathbf{x} \, \psi^*(\mathbf{x})\psi(\mathbf{x})$. The usefulness of these states is apparent when evaluating matrix elements associated with normally ordered operators, i.e., operators having creation operators on the left of annihilation operators (e.g., $\hat{a}^\dagger_{\mathbf{x}}\hat{a}^\dagger_{\mathbf{y}}\hat{a}_{\mathbf{x}}\hat{a}_{\mathbf{y}}$). Indeed, considering such a quadratic operator, $\hat{A}$, we have

$$\begin{aligned} \langle\psi|\hat{A}|\psi'\rangle &= \langle\psi|\sum_{\mathbf{xy}} \alpha_{\mathbf{xy}} \, \hat{a}^\dagger_{\mathbf{x}}\hat{a}_{\mathbf{y}}|\psi'\rangle \\ &= \sum_{\mathbf{xy}} \alpha_{\mathbf{xy}} \, \psi^*(\mathbf{x})\psi'(\mathbf{y}) e^{\psi^* \cdot \psi'} \\ &= A[\psi^*, \psi'] \, e^{\psi^* \cdot \psi} \, . \end{aligned} \tag{2.20}$$

This result extends to the matrix elements of any normal-ordered operator, $\langle\psi|\hat{X}|\psi'\rangle$, which is recast as $X[\psi^*, \psi'] \, e^{\psi^* \cdot \psi'}$, where $X[\psi^*, \psi']$ is a function involving the fields amplitudes.

With this, we return to the evaluation of Eq. (2.17). Upon expressing $e^{-\hat{H}N}$ as $\left(e^{-\frac{N}{M}\hat{H}}\right)^M$ and resolving the identity as in Eq. (2.19), we obtain

$$\begin{aligned} \mathcal{P}(\mathbf{x}_\mathrm{i}, \mathbf{x}_\mathrm{f}, N) = \int \mathcal{D}\psi\mathcal{D}\psi^* \; & \langle\mathbf{x}_\mathrm{f}|\psi_M\rangle \, \langle\psi_0|\mathbf{x}_\mathrm{i}\rangle \times \\ & \prod_{n=1}^{M} \langle\psi_n|e^{-\frac{N}{M}\hat{H}}|\psi_{n-1}\rangle \, \prod_{n=0}^{M} e^{-\psi_n^* \cdot \psi_n} \, , \end{aligned} \tag{2.21}$$

where $\int \mathcal{D}\psi^*\mathcal{D}\psi$ collectively addresses the integration on all the $(M+1)$ coherent fields. In the limit where a large number of Trotter steps, M, is employed, we can rewrite Eq. (2.21) as

$$\begin{aligned}\mathcal{P}(\mathbf{x}_\mathrm{i}, \mathbf{x}_\mathrm{f}, N) = &\int \mathcal{D}\psi\mathcal{D}\psi^* \, \psi_M(\mathbf{x}_\mathrm{f})\psi_0^*(\mathbf{x}_\mathrm{i})\times \\ &\prod_{n=1}^{M} e^{-\frac{N}{M}\hat{H}[\psi_n,\psi_{n-1}]} e^{-\psi_n^*\cdot\,\psi_{n-1}} \prod_{n=0}^{M} e^{-\psi_n^*\psi_n} \\ = &\int \mathcal{D}\psi\mathcal{D}\psi^* \, \psi_M(\mathbf{x}_\mathrm{f})\psi_0^*(\mathbf{x}_\mathrm{i})\times \\ &\prod_{n=1}^{M} e^{-\frac{N}{M}\hat{H}[\psi_n,\psi_{n-1}]-\psi_n^*\cdot\,(\psi_n-\psi_{n-1})} \; e^{-\psi_0^*\cdot\,\psi_0} \,,\end{aligned} \tag{2.22}$$

where we leveraged the properties of coherent fields defined in Eq. (2.20). In the same limit, we can apply the mapping

$$\begin{aligned}&\psi_n \longrightarrow \psi(s) \quad \text{with} \quad s = \frac{Nn}{M} \in [0, N] \in \mathbb{R} \\ &\psi_n - \psi_{n-1} \longrightarrow \frac{N}{M}\,\partial_s\psi(s) \\ &\sum_n \frac{N}{M} \longrightarrow \int_0^N \mathrm{d}s \,,\end{aligned} \tag{2.23}$$

so that Eq. (2.22) is expressed as

$$\begin{aligned}\mathcal{P}(\mathbf{x}_\mathrm{i}, \mathbf{x}_\mathrm{f}, N) = &\int \mathcal{D}\psi\mathcal{D}\psi^* \, \psi(N, \mathbf{x}_\mathrm{f})\psi^*(0, \mathbf{x}_\mathrm{i})\times \\ &e^{-\int_0^N \mathrm{d}\mathbf{s}\,(H[\psi^*(s),\psi(s)]-\psi^*(s)\cdot\,\partial_s\psi(s))} \quad e^{-\psi^*(0)\cdot\,\psi(0)} \,.\end{aligned} \tag{2.24}$$

Until now, we have not specified the modes that are created and destroyed by $\hat{a}_\mathbf{x}^\dagger$ and $\hat{a}_\mathbf{x}$. If we choose these to be real space modes, then we have

$$\begin{aligned}\mathcal{P}(\mathbf{x}_\mathrm{i}, \mathbf{x}_\mathrm{f}, N) = &\int \mathcal{D}\psi\mathcal{D}\psi^* \, \psi(N, \mathbf{x}_\mathrm{f})\psi^*(0, \mathbf{x}_\mathrm{i})\times \\ &e^{-\int_0^N \mathrm{d}s \int \mathrm{d}\mathbf{x}\, \psi^*(s,\mathbf{x})\left(\partial_s - \frac{b^2}{6}\nabla_\mathbf{x}^2\right)\psi(s,\mathbf{x})} \times \\ &e^{-\int \mathrm{d}\mathbf{x}\, \psi^*(0,\mathbf{x})\psi(0,\mathbf{x})} \,.\end{aligned} \tag{2.25}$$

Due to the prescription of causality, the previous path integral is equivalent to

$$\begin{aligned}\mathcal{P}(\mathbf{x}_{\mathrm{i}}, \mathbf{x}_{\mathrm{f}}, N) = \int \mathcal{D}\psi\mathcal{D}\psi^* \, \psi(N, \mathbf{x}_{\mathrm{f}})\psi^*(0, \mathbf{x}_{\mathrm{i}}) \times \\ e^{-\int_0^N \mathrm{d}s \int \mathrm{d}\mathbf{x}\, \psi^*(s,\mathbf{x})\left(\partial_s - \frac{b^2}{6}\nabla^2_{\mathbf{x}}\right)\psi(s,\mathbf{x})} \, .\end{aligned} \tag{2.26}$$

Crucially, when the action is at most quadratic in the field amplitudes, the path-integral is Gaussian and can be analytically evaluated. Otherwise, it is possible to take a perturbative approach, leading to Feynmann diagrams and related rules.

While several polymer models, including the random self-avoiding walk model, can be reformulated in terms of quantum field theories, the latter features a path integral with a complex term contributing to $H[r(s)]$. As a consequence, computational methods based on Markov Chains aiming at sampling from the distribution defined through Eq. (2.14) suffer the so-called *sign problem.* This is encountered when integrating a highly oscillating function, causing poor convergence of the Markov Chain, and can be *mitigated* by employing sophisticated, model-specific techniques, such as the one described in [144].

With the emergence of quantum computing as a powerful tool for sampling from established distributions [145], significant efforts have been devoted to exploring its applications to the polymer sampling problem. In the following section, we introduce a recent QUBO formulation of the homopolymer sampling problem that inspired the binary encoding underlying the sampling algorithm outlined in Chapter 4.

2.2.2 QUBO formulation of the homopolymer sampling problem

In Ref. [138], the problem of sampling self-avoiding configurations on a lattice was recast as a QUBO problem whose minimization is amenable to quantum analog devices such as the one provided by the D-Wave Inc. company.

In particular, homopolymer configurations are encoded into an array of binary variables, $\mathbf{\Gamma} = \{\Gamma_{\mathrm{i}}, \Gamma_{\mathrm{ij}}\}$, where $\Gamma_{\mathrm{i}} = 1$ if the lattice site i is occupied by a polymer's bead, and $\Gamma_{\mathrm{ij}} = 1$ if the lattice edge ij is populated with a covalent bond. In this work, the simplest ensemble that is addressed consists of polymers characterized by a fixed number of beads, N, and covalent bonds, L. Configurations satisfying the constraints introduced so far are the degenerate minima of the functional

$$\mathcal{H}[\mathbf{\Gamma}] = \mathcal{H}_{\rm beads} + \mathcal{H}_{\rm length} + \mathcal{H}_{\rm consistency}\ , \tag{2.27}$$

where $\mathcal{H}_{\rm beads} = A_{\rm beads}\ \left(\sum_{\rm i}\Gamma_{\rm i} - N\right)^2$ and $\mathcal{H}_{\rm length} = A_{\rm length}\ \left(\sum_{\langle {\rm i,\,j}\rangle}\Gamma_{\rm ij} - L\right)^2$. When $A_{\rm beads} > 0$ and $A_{\rm length} > 0$ the sum of these two terms is minimized when the binary field describes configurations having the proper number of beads and covalent bonds. The consistency between the binary fields $\Gamma_{\rm i}$ and $\Gamma_{\rm ij}$, is favored by incorporating into the Hamiltonian functional the term

$$\mathcal{H}_{\rm consistency} = A_{\rm consistency} \sum_{\langle {\rm i,\,j}\rangle} \Gamma_{\rm ij}\left(1-\Gamma_{\rm i}\right)\ , \tag{2.28}$$

where i and j are bound to be first-neighboring lattice sites. It is worth noticing that, by choosing $A_{\rm consistency} > 0$, Eq. (2.28) penalizes energetically all those configurations where lattice sites not belonging to the polymer chain (i.e., $\Gamma_{\rm i} = 0$) establish covalent interactions with neighboring sites, which corresponds to the situation where at least one edge connected to i is active (i.e., $\Gamma_{\rm ij} = 1$).

To address an ensemble of polymer configurations that have no self-intersections nor branches, it is necessary to introduce an additional set of binary variables, $\Gamma_{\rm ijk}$, to describe *trimer* states associated with consecutive lattice sites i, j, and k. More specifically, only when both the edges ij and jk are populated with a covalent bond, the corresponding trimer variable, $\Gamma_{\rm ijk}$, is activated. It is worth noticing that the new set of variables $\Gamma_{\rm ijk}$ is fully determined by the binary field $\Gamma_{\rm ij}$ and it is thus necessary to introduce suitable quadratic energy penalties to favor binary field states that locally satisfy $\Gamma_{\rm ijk} = \Gamma_{\rm ij}\,\Gamma_{\rm jk}$. Indeed, the Hamiltonian functional chosen to address this polymer ensemble was

$$\mathcal{H}'[\Gamma] = \mathcal{H} + \mathcal{H}_{\rm trimers} + \mathcal{H}_{\rm SA}\ , \tag{2.29}$$

where $\Gamma = \{\Gamma_i, \Gamma_{\rm ij}, \Gamma_{\rm ijk}\}$ collectively represents all the binary fields describing polymer configurations, $\mathcal{H}_{\rm trimers}$ implements consistency relations between binary fields associated with trimers and edges, and $\mathcal{H}_{\rm SA}$ exploits trimer variables to penalize the formation of branches and self-crossings of the polymer's backbone. As a result, configurations minimizing Eq. (2.29) are valid configurations that display no branching points

nor self-intersections.
Concretely, H_{SA} writes

$$\mathcal{H}_{SA} = A_{\mathrm{SA}} \sum_{\mathrm{i,j,k,l,m}}' \Gamma_{\mathrm{ijk}} \Gamma_{\mathrm{ljm}} \, , \tag{2.30}$$

where A_{SA} is a positive multiplier and the ($'$) in the summation imposes $\{\mathrm{i,j,k}\}$ to be a different set of vertices with respect to $\{\mathrm{l,j,m}\}$. In light of this, Eq. (2.30) provides a positive energy shift to binary vector states where two distinct trimers of edges sharing the central site are active. Indeed, this circumstance qualifies the polymer configuration to be either branched or self-intersecting. Further, consistency relations between the binary fields encoding edges and trimers were enforced with the term

$$\mathcal{H}_{\mathrm{trimers}} = A_{\mathrm{trimers}} \sum_{\mathrm{ijk}} 3\Gamma_{\mathrm{ijk}} + \Gamma_{\mathrm{ij}} \Gamma_{\mathrm{jk}} - 2\Gamma_{\mathrm{ijk}} \left(\Gamma_{\mathrm{ij}} + \Gamma_{\mathrm{jk}}\right) \, , \tag{2.31}$$

with A_{trimers} being a positive multiplicative factor. As it can be easily checked, the minima of Eq. (2.31) are binary fields that locally satisfy $\Gamma_{\mathrm{ijk}} = \Gamma_{\mathrm{ij}} \Gamma_{\mathrm{jk}}$.

Notably, the binary field configurations that minimize Eq. (2.29) include configurations representing polymeric backbones comprising isolated rings. To avoid such states, it was proposed to introduce additional terms contributing to the Hamiltonian functional in Eq. (2.29) involving ancillary binary fields. Indeed, to prevent the formation of $2n$–long rings, it was necessary to employ binary fields describing the state of all n–long strings composed of consecutive edges. While this mapping holds the potential to harness NISQ devices in the context of homopolymer sampling, the large overhead in quantum resources required to penalize configurations containing rings up to a given size motivated us to explore the alternative algorithm discussed in Chapter 4. Although our algorithm does not entail a QUBO problem, it highlights deep connections between lattice gauge theory (LGT) and polymer sampling, which can be exploited to design Monte Carlo trial moves that naturally preserve the continuity of the backbone chain. In addition, we introduce fermionic auxiliary degrees of freedom coupled to the electric fields of the LGT, which, once integrated out, yield an ensemble where loopy confomations are assigned with vanishing canonical weight.

PART II

SOFT-MATTER PHYSICS APPLICATIONS

Chapter 3

Protein design by integrating machine learning and quantum-encoded optimization

As anticipated in Sec. 2.1.3, to design the structure Γ_{T}, it is necessary to find a sequence, S, such that

$$P_{\mathrm{opt}}(S|\Gamma_{\mathrm{T}}) = \max_{S} \; e^{-\beta(E(\Gamma_{\mathrm{T}},S)-F(S))} \geq p_{\mathrm{fold}} \, , \tag{3.1}$$

where p_{fold} is a suitably chosen constant, typically greater than 0.5, and $F(S)$ is the free energy associated with the sequence S, namely

$$F(S) = -\frac{1}{\beta} \ln \left(\sum_{\Gamma} e^{-\beta \, E(\Gamma,S)} \right) \, . \tag{3.2}$$

Solving Eq. (3.1) turns into minimizing

$$G(S) = E(\Gamma_{\mathrm{T}}, S) - F(S) \tag{3.3}$$

with respect to S and check if the constraint in Eq. (3.1) holds. The first challenge lies in $E(\Gamma, S)$ not being known *a priori*; in addition, the estimation of $F(S)$ hinders a search in structure space that, if addressed exhaustively, makes the overall computation untractable even for moderately large systems.

The main goal of the present study is to demonstrate that it is possible to integrate advancements in both machine learning and quantum computing technologies to tackle the design problem without abandoning the physics-based standpoint of Eqs. (3.1) and (3.3). In this context, the research in quantum computing may also drive the development of radically new physics-based formulations that are advantageous even when implemented on classical machines [139].

In our first illustrative application, we resort to minimalistic lattice protein models. This choice is particularly suited to assessing the accuracy of our scheme, allowing us to better control the sources of errors. Indeed, it eliminates the uncertainties associated with heuristic machine-learning algorithms for protein folding, as they can be replaced by an exhaustive search of the conformational space. Furthermore, it enables to assess the accuracy through which our iterative learning scheme is able to learn the underlying physics-based energy function.

Even after this major simplification, the combinatorial search over the sequence space can be computationally demanding. Quantum annealing machines are ideally suited to solve this kind of discrete combinatorial optimizations after a suitable mathematical reformulation, or encoding, of the original problem. A very relevant question to address is if such reformulation can lead to performance improvement even when adopted on classical machines [139].

We answer this question in the affirmative. Indeed, in our proof-of-concept study, quantum-encoded approaches implemented on both classical and quantum computers outperform a well-established scheme based on simulated annealing. At the same time, our iterative machine learning scheme enables us to reach solutions to the design problem with a high success rate. Our algorithm's main merit is its ability to simultaneously harness the new possibilities offered by Machine Learning applications and promised by the advancements in quantum computing hardware, while remaining rooted into the physics-based modelling paradigm. Furthermore, its portability to off-lattice all-atom molecular representations paves the ways to perspective realistic applications. Collectively, our results suggest that, if the size and performance of quantum simulators continue to improve over the next several years, the integration of quantum annealers and classical machine learning may represent a transformative new paradigm with broad implications in several areas of Life Sciences and Pharmacology.

In Section 3.1, we introduce the approximation adopted to address the term $F(S)$ in Eq. (3.3). Despite this approximation, the problem remains NP-hard and is tackled using the iterative scheme extensively discussed in Section 3.2. This protocol is applied to the simple lattice model in Section 3.3, leading to the results presented in Section 3.4. The details concerning the numerical simulations presented are enclosed in Section 3.5, while in Section 3.6 we discuss how the present study can be extended to realistic off-lattice protein models.

3.1 Approximate scoring function $G(S)$

We simplify the complexity of the design problem by introducing two approximations to circumvent the nested sequence-structure search implied by the minimization of the design scoring function in Eq. (3.3).

First, we resort to a customary linear Ansatz for $E(\Gamma, S)$:

$$E(\Gamma, S) \simeq \sum_{\mathrm{i}>\mathrm{j}=1}^{N} C_{\mathrm{ij}}(\Gamma)\, \varepsilon(s_{\mathrm{i}}, s_{\mathrm{j}})\,. \tag{3.4}$$

Importantly, this equation, despite being formally identical to Eq. (2.1) introduced in Section 2.1.1, does not directly address a model with discretized space. Indeed, the entries of the contact map $\mathbf{C}(\Gamma)$ can span the entire $[0:1]$ interval and thus weigh the interaction between residues based on their distance. When choosing a sharp cutoff to pinpoint interacting residues, the entries of $\mathbf{C}(\Gamma)$ take the discrete value 1, if an interaction between residues takes place, and 0 otherwise. We recall that in Eq. (3.4) s_{i} is the chemical identity of the i-th amino acid of the sequence S, which has length n, and $\varepsilon(s_{\mathrm{i}}, s_{\mathrm{j}})$ are the entries of a suitable $D \times D$ energy matrix — D being the size of the amino acids chemical alphabet.

In Eq. (3.3), the energy of the sequence S mounted over the target structure is computed relative to the sequence free energy $F(S)$. Evaluating the latter implies computing the energy of S mounted over all possible states, a computationally prohibitive task. A key approximation of our approach consists in replacing this reference with the average energy evaluated over a database of known native structures,

$$F(S) \simeq \sum_{\mathrm{i}>\mathrm{j}=1}^{N} \varepsilon(s_{\mathrm{i}}, s_{\mathrm{j}}) \, \langle C_{\mathrm{ij}} \rangle \ . \tag{3.5}$$

In the expression above, $\langle C_{\mathrm{ij}} \rangle$ is the average contact map of those structures in the database that have the same length as the target one. In more general contexts, this restriction can be relaxed by setting $\langle C_{\mathrm{ij}} \rangle$ equal to the contact probability of amino acids at chemical distance $|\mathrm{i}-\mathrm{j}|$ computed over all database entries that are sufficiently longer than $|\mathrm{i}-\mathrm{j}|$, to avoid end effects.

The gist of the approximation in Eq. (3.5) is to yield a free energy estimate that, while remaining computationally amenable, is still informed by the structural properties of viable states. The average pairwise contact probabilities appear to be the most natural and effective choice in this respect, considering that the approach could be systematically generalized to include three-body and higher-order contact probabilities.

With this proviso, our approximation to the design scoring function $G(S)$ becomes

$$G(S) \simeq \sum_{\mathrm{i}>\mathrm{j}=1}^{N} \varepsilon(s_{\mathrm{i}}, s_{\mathrm{j}}) \, \left(C_{\mathrm{ij}}(\Gamma_{\mathrm{T}}) - \langle C_{\mathrm{ij}} \rangle\right) \ . \tag{3.6}$$

Minimizing this function thus selects sequences whose native energy is as low as possible compared to the average state in the database. Importantly, the energy matrix $\varepsilon(s_{\mathrm{i}}, s_{\mathrm{j}})$ does not need to be known *a priori*. Instead, we propose an iterative approach, discussed in the next section, through which an initial guess is refined until consistency between the direct and inverse folding problems is reached.

3.2 Iterative scheme to tackle the design problem

Key parameters of our design scheme are the entries of the $D \times D$ symmetric energy matrix $\boldsymbol{\varepsilon}$ of Eq. (3.6). Inspired by earlier work on the extraction of effective potentials for protein folding or design [146–149], we adopted an iterative scheme based on enforcing consistency between the solutions of the direct and inverse folding problems. Our choice is motivated by the increasing availability of reliable and fast algorithms for predicting protein folds even in realistic contexts [127]. In principle, this opens the possibility of harnessing these efficient methods for tuning $\boldsymbol{\varepsilon}$ to design a specific type or family of target structures. Such optimized schemes would also limit adverse effects inherent to structural coarse-graining, which inevitably impacts the transferability of potential energies obtained by thermodynamic integration. For the same reason, it is not apparent *a priori* that database-wide schemes for extracting interaction potentials, including the powerful quasichemical approximation [90, 91, 94–96] for capturing amino acids interaction propensities, are ideally suited to the design task at hand based on the minimization of Eq. (3.6).

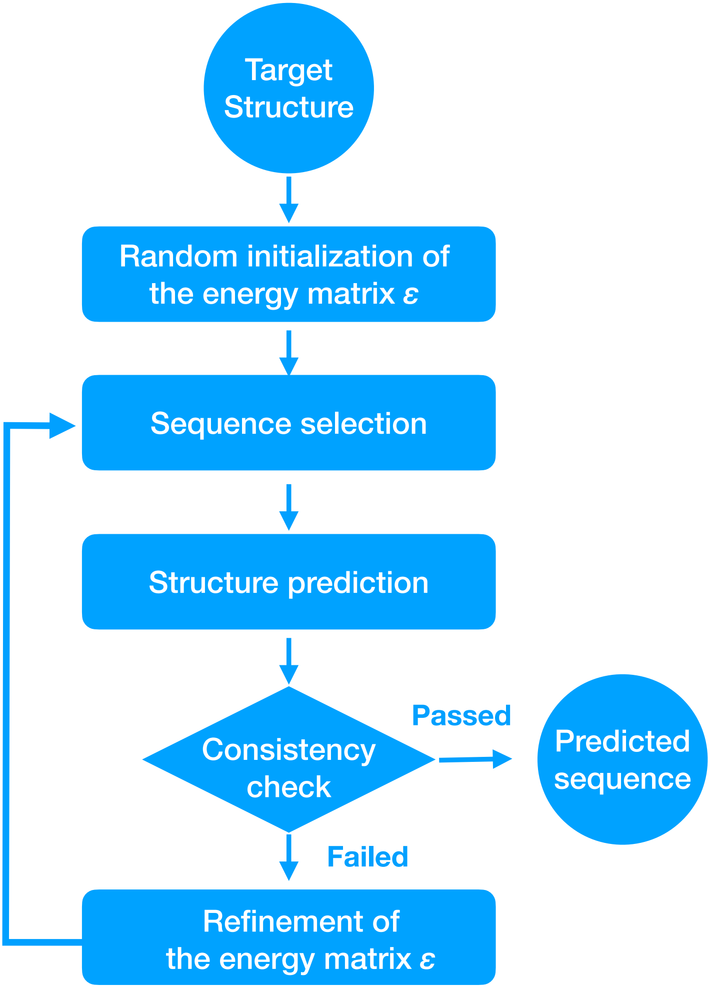


Figure 3.1: Schematic representation of our protein-design algorithm. Alternating steps of sequence selection and structure prediction are repeated. The energy map from an initial guess is updated until the algorithm passes consistency checks. The result is a sequence that folds into the desired target structure. Importantly, sequence selection and structure prediction can be integrated from different algorithmic or hardware platforms, permitting leveraging on rapid and complementary progress in technologies such as AI and quantum computing.

The key steps of our iterative scheme are sketched in Fig. 3.1. Given the target structure to be designed, Γ_{T}, and a random initialization of the energy matrix $\boldsymbol{\varepsilon}$, the scheme proceeds by iterating at each cycle the following steps:

1. *Sequence selection.* Explore the combinatorial space of sequences with fixed chemical composition $\mathbf{N} = \{N_1, \ldots, N_D\}$ — where N_m specifies the abundancy of the chemical identity m in the sequence — to find the set $\mathcal{S} \equiv \{S_1, S_2, ...\}$ corresponding to the lowest values of $G(S)$.

2. *Structure prediction.* For each sequence in $\mathcal{S}$ obtain a reliable prediction of the native state. We shall indicate such corresponding native set as $\mathcal{N} \equiv \{\tilde{\Gamma}_1, \tilde{\Gamma}_2, ...\}$.

3. *Energy function refinement.* Assess whether the states in $\mathcal{N}$ match the target structure Γ_{T} within a specified tolerance. If so, the design problem of Γ_{T} is solved, and the procedure ends. Otherwise, the symmetric energy matrix $\boldsymbol{\varepsilon}$ is refined to impose a consistency with the structure prediction results, i.e., to account for the fact that the native states of $\mathcal{S}$ do not include Γ_{T}.

A fixed point in this iterative scheme embodies the highest achievable consistency between our heuristic scoring function and the chosen protein structure prediction algorithm. In the following sections, we individually examine the three steps mentioned above, emphasizing how quantum technologies and advanced deep-learning tools can contribute.

3.2.1 Step 1: Sequence selection.

The first step of our iterative scheme involves solving a combinatorial optimization problem over the space of amino acid sequences. We will perform this step by constraining the overall amino acid composition, i.e., the abundances of the different types of amino acids. Thus, the first step involves minimizing Eq. (3.6) over the possible reshuffling of a given initial sequence with the desired composition. This task is an integer programming problem that can be carried out on conventional computers. However, considerable speed-ups for the same NP-complete class of problems may be achieved with quantum annealers. [150–153].

In the following, we illustrate two viable QUBO encodings of the constrained integer optimization problem set by Eq. (3.6), which can be either solved by resorting to the use of the quantum D-Wave facility or on the heuristic industry-grade software GUROBI. Further, we discuss the encoding employed in the Simulated Annealing protocol adopted to benchmark results obtained relying on the QUBO encoding.

In Sec. 3.2.1.1 we introduce the encoding of the constrained optimization of Eq. (3.6) on the Advantage quantum platform that is chosen to derive results in 3.4, while in Section 3.2.1.2 we report an alternative encoding to the one adopted in the present work. In Section 3.2.1.3, we discuss the technique employed to derive classical results.

3.2.1.1 QUBO linear encoding

The quantum solver to whom we address the present discussion is the superconducting quantum annealer provided by the D-Wave Inc., whose features we discuss in greater detail in Appendix B. We recall that, the problem Hamiltonians are limited to have the form

$$\hat{H}_{\mathrm{QUBO}} = \sum_{\mathrm{i}} \mu_{\mathrm{i}}\, \hat{\sigma}_{\mathrm{i}}^{(z)} + \sum_{\mathrm{i}>\mathrm{j}} J_{\mathrm{ij}}\, \hat{\sigma}_{\mathrm{i}}^{(z)}\, \hat{\sigma}_{\mathrm{j}}^{(z)}\,, \tag{3.7}$$

where $\hat{\sigma}_{\mathrm{i}}^{(z)}$ single qubit operators have eigenvalues ±1 and both local fields, μ_{i}, and couplings, J_{ij}, are tunable in sign and intensity. To shift the values of the binary variables involved to be in $\{0,1\}$, it is customary to set $\hat{\sigma}_{\mathrm{i}}^{(z)} = \hat{\mathbb{1}} - 2\hat{q}_{\mathrm{i}}$ in Eq. (3.7), where $\hat{q}_{\mathrm{i}}$ operators have eigenvalues 0 and 1.

To solve the optimization problem set by Eq. (3.6) using this quantum solver, it is necessary to (i) introduce suitable binary variables to describe the system undergoing optimization, (ii) manipulate the constraints into energy penalties targeting binary strings that are constraints-violating, and (iii) express both penalty and objective functions in a form that can be reproduced by plugging suitable local fields, μ_{i}, and couplings, J_{ij}, in Eq. (3.7).

Encoding To recast the minimization of $G(S)$ in Eq. (3.6) as a QUBO problem, we introduce binary variables to describe the chemical type of each amino acid in the sequence. Specifically, to the i-th element of the sequence we associate an array of D binary variables, $q_m^{(\mathrm{i})}$, with $m =$

$\{1, \ldots, D\}$. With this choice, if the i-th element of the primary sequence corresponds to the m-th chemical identity in the dictionary, then $q_m^{(\mathrm{i})} = 1$; otherwise, if the m-th identity does not appear at position i, then $q_m^{(\mathrm{i})} = 0$.

To prevent two or more monomers from occupying the same site and to impose the global chemical composition given by the occupation array $\mathbf{N} = \{N_1, \ldots, N_D\}^1$, we enforce, respectively, the excluded volume condition

$$\forall\, \mathrm{i} = 1, \ldots, N\,, \quad \sum_{m=1}^{D} q_m^{(\mathrm{i})} = 1 \tag{3.8}$$

and the composition constraint

$$\forall\, m = 1, \ldots, D\,, \quad \sum_{\mathrm{i}=1}^{N} q_m^{(\mathrm{i})} = N_m\,. \tag{3.9}$$

The constraints in Eqs. (3.8) and (3.9) are not linearly independent; to see this, let us consider $\sum_{m=1}^{D} \sum_{\mathrm{i}=1}^{N} q_m^{(\mathrm{i})}$ and, using Eq. (3.8), express it as $\sum_{i=1}^{N} 1 = N$. The same result can be derived using Eq. (3.9) combined with the fact that the composition array is satisfies $\sum_{m=1}^{D} N_m = N$. This allows us to remove at least one of the constraints appearing in Eqs. (3.8) and (3.9). In particular, we retain the constraints given by

$$\forall\, i = 1, \ldots, N, \quad \sum_{m=1}^{D} q_m^{(\mathrm{i})} = 1 \quad \text{and} \tag{3.10}$$

$$\forall\, m = 2, \ldots, D, \quad \sum_{i=1}^{N} q_m^{(\mathrm{i})} = N_m\,. \tag{3.11}$$

It is worth noticing that the variables $q_1^{(\mathrm{i})}$ no longer appear in the constraints in Eq. (3.11). As a next step, we leverage Eq. (3.10) to reduce the number of binary variables required to describe the sequence. Namely, we set $q_1^{(\mathrm{i})} = 1 - \sum_{m=2}^{D} q_m^{(\mathrm{i})}$ and we subject the remaining $N(D-1)$ variables

---

[1] As the protein sequence has length N, the compostion array $\mathbf{N} = \{N_1, \ldots, N_D\}$ is bound to satisfy $\sum_{m=1}^{D} N_m = N$.

to the new set of constraints given by

$$\begin{aligned} &\forall\, \mathrm{i} = 1, \ldots, N\,, \quad \sum_{m=2}^{D} q_m^{(\mathrm{i})} \leq 1 \quad \text{and} \\ &\forall\, m = 2, \ldots, D\,, \quad \sum_{\mathrm{i}=1}^{N} q_m^{(\mathrm{i})} = N_m\,. \end{aligned} \tag{3.12}$$

With this, we conclude that $(D-1)$ binary variables are sufficient to specify the chemical identity of a residue along the polymer chain and that $N(D-1)$ logical qubits satisfying Eq. (3.12) describe valid sequences having the right composition.

Penalty terms featuring constraints To embed the conditions set through Eq. (3.12) on the quantum platform developed by D-Wave Inc., we need to express them in terms of quadratic penalties.

To manipulate the first constraint, it is worth noticing that, since $q_m^{(\mathrm{i})} \in \{0, 1\}$, the excluded volume condition is satisfied if and only if $\sum_{m=2}^{D} q_m^{(\mathrm{i})}$ is 0 or 1, which is equivalent to enforce $\left(\sum_{m=2}^{D} q_m^{(\mathrm{i})}\right)^2 = \sum_{m=2}^{D} q_m^{(\mathrm{i})}$ and leads to the quadratic penalty term

$$\mathcal{H}_{\mathrm{occ}} = A_1 \sum_{\mathrm{i}=1}^{N} \sum_{m \neq n = 2}^{D} q_m^{(\mathrm{i})} q_n^{(\mathrm{i})}\,, \tag{3.13}$$

which, for $A_1 > 0$, penalizes cases where more than one entry of the $q_m^{(\mathrm{i})}$ array is equal to 1.

The second constraint is recast as the quadratic term

$$\mathcal{H}_{\mathrm{comp}} = A_2 \sum_{m=2}^{D} \left(\sum_{\mathrm{i}=1}^{N} q_m^{(\mathrm{i})} - N_m \right)^2, \tag{3.14}$$

where $A_2 > 0$ and N_m is the assigned number of amino acids of type m. Minimizing this term ensures that the non-zero entries of the $\mathbf{q}^{(\mathrm{i})}$ arrays are consistent with the prescribed composition.

Eqs. (3.13) and (3.14) are quadratic expressions involving the binary variables $q_m^{(\mathrm{i})} \in \{0,1\}$ and are reproduced, respectively, in the eigenvalues associated with the operators

$$\hat{H}_{\mathrm{occ}} = A_1 \sum_{\mathrm{i}=1}^{N} \sum_{m \neq n=2}^{D} \hat{q}_m^{(\mathrm{i})}\, \hat{q}_n^{(\mathrm{i})} \quad \text{and} \tag{3.15}$$

$$\hat{H}_{\mathrm{comp}} = A_2 \sum_{m=2}^{D} \left(\sum_{\mathrm{i}=1}^{N} \hat{q}_m^{(\mathrm{i})} - n_m \right)^2 , \tag{3.16}$$

which, by applying the substitution $\hat{q}_m^{\mathrm{i}} = \frac{\hat{\mathbb{1}} - \sigma_m^{(z)}}{2}$, match exactly the form of Eq. (3.7).

Objective function Let us now show how Eq. (3.6) can be expressed as a quadratic form involving the binary variables $q_m^{(\mathrm{i})}$, where $m = 2, \ldots, D$ and $\mathrm{i} = 1, \ldots, N$.

With our definition of $q_m^{(\mathrm{i})}$, $\varepsilon(s_{\mathrm{i}}, s_{\mathrm{j}}) = \sum_{m,n=1}^{D} q_n^{(\mathrm{i})} q_m^{(\mathrm{j})}\, \varepsilon(n,m)$, so that

$$\begin{aligned} G(\Gamma, S) = &\sum_{i>j=1}^{N} \sum_{m,n=2}^{D} \tilde{C}_{\mathrm{ij}}\, q_m^{(\mathrm{i})} q_n^{(\mathrm{j})}\, \varepsilon(m,n) + \\ &\sum_{i>j=1}^{N} \sum_{m=2}^{D} \tilde{C}_{\mathrm{ij}}\, q_m^{(\mathrm{i})} q_1^{(\mathrm{j})}\, \varepsilon(m,1) + \\ &\sum_{i>j=1}^{N} \sum_{n=2}^{D} \tilde{C}_{\mathrm{ij}}\, q_1^{(\mathrm{i})} q_n^{(\mathrm{j})}\, \varepsilon(1,n) + \\ &\sum_{i>j}^{N} \tilde{C}_{\mathrm{ij}}\, q_1^{(\mathrm{i})} q_1^{(\mathrm{j})}\, \varepsilon(1,1)\,, \end{aligned} \tag{3.17}$$

where $\tilde{\mathbf{C}} = \mathbf{C}(\Gamma_{\mathrm{T}}) - \langle \mathbf{C} \rangle$, $\mathbf{C}(\Gamma_{\mathrm{T}})$ is the contact map assosciated with the target structure (Γ_{T}), and $\langle \mathbf{C} \rangle$ is the average contact map evaluated on a database of representative native structure. Using the $q_1^{(\mathrm{i})} = 1 - \sum_{m=2}^{D} q_m^{(\mathrm{i})}$

identity yields

$$
\begin{aligned}
G(\Gamma, S) = &\sum_{\mathrm{i}>\mathrm{j}=1}^{N} \sum_{m,n=2}^{D} \tilde{C}_{\mathrm{ij}}\, q_m^{(\mathrm{i})} q_n^{(\mathrm{j})}\, \varepsilon(m,n) + \\
&\sum_{\mathrm{i}>\mathrm{j}=1}^{N} \sum_{m=2}^{D} \tilde{C}_{\mathrm{ij}}\, q_m^{(\mathrm{i})} \left(1 - \sum_{n=2}^{D} q_n^{(\mathrm{j})}\right) \varepsilon(m,1) + \\
&\sum_{\mathrm{i}>\mathrm{j}=1}^{N} \sum_{n=2}^{D} \tilde{C}_{\mathrm{ij}} \left(1 - \sum_{m=2}^{D} q_m^{(\mathrm{i})}\right) q_n^{(\mathrm{j})}\, \varepsilon(1,n) + \\
&\sum_{\mathrm{i}>\mathrm{j}=1}^{N} \tilde{C}_{\mathrm{ij}} \left(1 - \sum_{m=2}^{D} q_m^{(\mathrm{i})}\right) \left(1 - \sum_{n=2}^{D} q_n^{(\mathrm{j})}\right) \varepsilon(1,1)\,,
\end{aligned}
\tag{3.18}
$$

and finally

$$
\begin{aligned}
G(\Gamma, S) = &\sum_{\mathrm{i}>\mathrm{j}=1}^{N} \sum_{m,n=2}^{D} \tilde{C}_{\mathrm{ij}}\, q_m^{(\mathrm{i})} q_n^{(\mathrm{j})} [\varepsilon(m,n) - \varepsilon(m,1) - \varepsilon(1,n) + \varepsilon(1,1)] + \\
&\sum_{\mathrm{i}>\mathrm{j}=1}^{N} \sum_{m=2}^{D} \tilde{C}_{\mathrm{ij}} \left(q_m^{(\mathrm{i})} + q_m^{(\mathrm{j})}\right) [\varepsilon(1,m) - \varepsilon(1,1)] + \sum_{\mathrm{i}>\mathrm{j}=1}^{N} \tilde{C}_{\mathrm{ij}}\, \varepsilon(1,1)\,.
\end{aligned}
\tag{3.19}
$$

Building on these premises, the quadratic objective function associated with a given target structure, Γ_{T}, is

$$
\begin{aligned}
\mathcal{H}_{\mathrm{contact}} = B\Bigg(&\sum_{\mathrm{i}>\mathrm{j}=1}^{N} \sum_{m,n=2}^{D} q_m^{(\mathrm{i})} q_n^{(\mathrm{j})} \tilde{C}_{\mathrm{ij}} \alpha_{mn} + \\
&\sum_{\mathrm{i}>\mathrm{j}=1}^{N} \sum_{m=2}^{D} \left(q_m^{(\mathrm{i})} + q_m^{(\mathrm{j})}\right) \tilde{C}_{\mathrm{ij}} \gamma_m + \sum_{\mathrm{i}>\mathrm{j}=1}^{N} \tilde{C}_{\mathrm{ij}} \varepsilon_{11} \Bigg)\,,
\end{aligned}
\tag{3.20}
$$

where, $B > 0$, $\alpha_{mn} = \varepsilon_{mn} - \varepsilon_{m1} - \varepsilon_{n1} + \varepsilon_{11}$, and $\gamma_m = \varepsilon_{m1} - \varepsilon_{11}$. Consequently, the problem Hamiltonian that reproduces in its eigenvalues the behavior of Eq. (3.20) is

$$\begin{aligned}\hat{H}_{\text{contact}} = B\Bigg(&\sum_{\mathrm{i,j}=1}^{N}\sum_{m,n=2}^{D} \hat{q}_m^{(\mathrm{i})}\,\hat{q}_n^{(\mathrm{j})}\tilde{C}_{\mathrm{ij}} + \\ &\sum_{\mathrm{i}>\mathrm{j}=1}^{N}\sum_{m=2}^{D}\left(\hat{q}_m^{(\mathrm{i})} + \hat{q}_m^{(\mathrm{j})}\right)\tilde{C}_{\mathrm{ij}}\gamma_m + \sum_{\mathrm{i}>\mathrm{j}=1}^{N}\tilde{C}_{\mathrm{ij}}\varepsilon_{11}\Bigg) .\end{aligned} \tag{3.21}$$

Final problem Hamiltonian The overall problem Hamiltonian results from combining the constraint terms in Eqs. (3.15) and (3.16) with the objective in Eq. (3.21), resulting in

$$\hat{H} = \hat{H}_{\text{comp}} + \hat{H}_{\text{occ}} + \hat{H}_{\text{contact}} \ . \tag{3.22}$$

We emphasize that $\hat{H}_{\text{occ}}$ and $\hat{H}_{\text{comp}}$ encode the strong constraints, while $\hat{H}_{\text{contact}}$ represents the molecular energy. As long as $A_1, A_2 \gg B$, the ground state solutions of Eq. (3.22) simultaneously satisfy all the hard constraints and correspond to sequences that minimize $G(s)$ for the given chemical composition.

The sought ground states of the QUBO Hamiltonian can be found with various methods, including the purely classical heuristic search implemented in GUROBI and the hybrid-quantum scheme provided by the D-Wave OCEAN library. The latter combines classical taboo search heuristic optimization with quantum annealing steps [154].

3.2.1.2 QUBO Logarithmic encoding

In this section, we introduce an alternative encoding that reformulates the optimization problem defined in Eq. (3.6) as a QUBO. While this approach is theoretically valid, it is less intuitive than the encoding presented in Section 3.2.1.1 and does not offer any advantage in terms of quantum resource requirements. For these reasons, this alternative encoding is not used to derive the results discussed in Section 3.4. Instead, this section serves to illustrate with an example that QUBO formulations of a given optimization problem are generally not unique.

Considering an alphabet of D monomers, this encoding makes use of an array of $d = \log_2 D^2$qubits to identify each monomer of the sequence. In particular, if the n-th monomer of the alphabet occupies the i-th

---

[2] For the sake of simplicity, let us consider the case where $D = 2^d$.

position in the sequence, then $\mathbf{q}^{(\mathrm{i})} = \mathbf{n}$, with $\mathbf{n} = (n_1, \ldots, n_d)$ being the binary representation of n, with the convention that the first element of the alphabet is associated with $n = 0$.

To exemplify the use of this encoding, let us suppose that the alphabet consists of four letters — A, B, C, and D. According to the encoding described sofar, the binary representation of the sequence "ACBD" is $\mathbf{q}^{(1)} = (0, 0)$, $\mathbf{q}^{(2)} = (1, 0)$, $\mathbf{q}^{(3)} = (0, 1)$, and $\mathbf{q}^{(4)} = (1, 1)$.

In this setting, there is no necessity to impose excluded volume conditions, as any configuration of the d–binary array $\mathbf{q}^{(\mathrm{i})}$ uniquely identifies a chemical species. Besides, the sequence must be compatible with the composition constraint, which is specified through the occupation array $\mathbf{N} = \{N_1, \ldots, N_D\}$. Before moving to the description of the associated quadratic functional, it is worth noticing that to identify the chemical species corresponding to the i-th residue along the sequence, it is necessary to rely on functions $\delta(\mathbf{n}, \mathbf{q}^{(\mathrm{i})})$ — with $n = 1, \ldots, D$ — such that $\delta(\mathbf{n}, \mathbf{q}^{(\mathrm{i})}) = 1$ if and only if $\mathbf{q}^{(\mathrm{i})} = \mathbf{n}$, so that $\mathbf{q}^{(\mathrm{i})}$ accounts for the presence of the n-th monomer in site i.

With the aid of such functions, we rewrite the composition constraint as

$$\mathcal{H}_{\mathrm{comp}} = \sum_{m=1}^{D} \left(\sum_{\mathrm{i}=1}^{N} \delta(\mathbf{m}, \mathbf{q}^{(\mathrm{i})}) - N_m \right)^2 , \tag{3.23}$$

and the $G(S)$ functional as

$$G(\Gamma, S) = \sum_{\mathrm{i}>\mathrm{j}}^{N} \sum_{n,m} \tilde{C}_{\mathrm{ij}}(\Gamma)\, \delta(\mathbf{n}, \mathbf{q}^{(\mathrm{i})})\, \delta(\mathbf{m}, \mathbf{q}^{(\mathrm{j})})\, \varepsilon(n, m) . \tag{3.24}$$

Relevantly, Eqs. (3.23) and (3.24) are bound to contain terms like $\delta(\mathbf{m}, \mathbf{q}^{(\mathrm{i})})\, \delta(\mathbf{n}, \mathbf{q}^{(\mathrm{j})})$ that need to be manipulated into quadratic expressions involving the binary variables $q_s^{(\mathrm{i})}$ with $\mathrm{i} = 1, \ldots, N$ and $s = 1, \ldots, d$. This implies expressing each function $\delta(\mathbf{m}, \mathbf{q}^{(\mathrm{i})})$, whose expansion is

$$\delta(\mathbf{m}, \mathbf{q}^{(\mathrm{i})}) = \prod_{s=1}^{d} \delta(m_s, q_s^{(\mathrm{i})}) , \tag{3.25}$$

as a linear term in the $q_s^{(\mathrm{i})}$ variables. It is worth noticing that the terms $\delta(m_s, q_s^{(\mathrm{i})})$ are linear. Indeed, when $m_s = 1$, it is sufficient to set $\delta(m_s, q_s^{(\mathrm{i})}) = q_s^{(\mathrm{i})}$, otherwise if $m_s = 0$, the delta function is faithfully reproduced by the term $(1-q_s^{(\mathrm{i})})$. Then, it is clear that Eq. (3.25) is in general a d-body term that needs to be recast in a linear term by adding ancillary variables. To evaluate the number of such ancillary variables it is convenient to proceed by induction.

First, notice that anicillarizing $\delta(\mathbf{m},\mathbf{q}^{(\mathrm{i})})$ for any choice of binary array $\mathbf{m}$ involves linearizing 2-, 3-, …, d-body intraction terms involving any subset of the d variables. Further, observe that any k–body interaction can be obtained by "gluing" two $(k-1)$–body terms by means of an additional ancillary variable (e.g., $(q_1q_0) = (q_1)(q_0), (q_2q_1q_0) = (q_2q_1)(q_1q_0)$, $(q_3q_2q_1q_0) = (q_3q_2q_1)(q_2q_1q_0)$, etc.). Considering all possible combinations of k variables belonging to a set of d elements giving rise to a k–body interaction term, we deduce that $\begin{pmatrix} k \\ d \end{pmatrix} = \frac{d!}{k!\,(d-k)!}$ ancillary variables are necessary to degrade any k-body interaction term. Finally, summing the required additional variables to represent 2-, 3-, …, d-interaction terms, we find that the total amount of ancillas employed for each bead is

$$\sum_{k=2}^{d} \begin{pmatrix} k \\ d \end{pmatrix} = \sum_{k=0}^{d} \begin{pmatrix} k \\ d \end{pmatrix} - 1 - d = 2^d - 1 - d\,. \tag{3.26}$$

Adding the d qubits initially employed to encode the chemical identity of residues through the $\mathbf{q}^{\mathrm{i}} = (q_1^{(\mathrm{i})}, \ldots, q_d^{(\mathrm{i})})$ array, the total number of necessary qubits required *per bead* amounts to $2^d - 1 = D - 1$. This concludes that the total amount of qubits that are needed to reproduce Eqs. (3.23) and (3.24) in the eigenstates of QUBO Hamiltonian is $N(D-1)$.

The encoding introduced in this section is physically less transparent than the one discussed previously, in Section 3.2.1.1, and does not imply a saving in terms of quantum resources. For these reasons, the logarithmic encoding is not employed to derive results in Section 3.4.

3.2.1.3 Simulated annealing on classical machine

In this section, we provide information regarding the classical combinatorial optimization algorithms employed in the sequence optimization step of our iterative algorithm. In our approach, we employed an improved version of the simulated annealing scheme to retain only trial moves that preserve the global relative abundance of chemical identities in the chain, thereby ensuring a much larger acceptance rate. In this approach, the chain sequence was not specified by means of binary variables (as in the QUBO formulation), but rather directly as a sequence $(X_1, X_2, \dots)$ of alphabet elements ($X_i \in \mathcal{D}$). Trial moves were proposed by swapping randomly chosen pairs of monomers in the sequence. For example, given a sequence $S = (X_1, X_2, \dots)$ of alphabet elements, selecting i-th and j-th monomers, proposes the sequence

$$S' = (\dots, X_i', \dots, X_j', \dots) = (\dots, X_j, \dots, X_i, \dots)\,. \tag{3.27}$$

We stress that this prescription automatically avoids configurations where several monomers occupy the same site; in addition to this, since moves conserve the composition of a given sequence, it is sufficient to select as initial state a sequence that meets the composition constraint.

In all annealing simulations, we relied on the `Simanneal` package for python [155], with the following choice of parameters: $T_{\text{max}} = 100$, $T_{\text{min}} = 10^{-4}$, and $N_{\text{steps}} = 10^4$.

3.2.2 Step 2: Structure prediction.

The second step of the iterative scheme involves the application of structure prediction methods to the sequences identified from the minimization of $G(S)$ at the previous step. The key point is that the native states of such sequences are obtained with an independent structure prediction method. In particular, the scheme used to predict native structures given a sequence is not informed by the energy matrix defining the design scoring function $G(S)$.

In realistic off-lattice applications, the go-to structure prediction methods would naturally be those based on heuristic machine-learning algorithms, which have proved to be reliable and efficient. Considering the minimalistic nature of the protein model that we adopt as an illustrative example in this work (see Section 3.3), we opted for the most transparent and feasible method: an exhaustive search of conformational space to identify the lowest energy state(s) of a sequence based on a ground-truth energy matrix. This ground-truth energy matrix is used solely in this step and for selecting viable target structures for the design problem, as detailed in Section 3.3, and is never used in the $G(S)$ definition.

3.2.3 Step 3: Energy function refinement.

The third step in the proposed iterative scheme involves updating the energy matrix entering $G(S)$ to improve consistency with the chosen ground-truth structure prediction algorithm.

To this end, we have devised the following scheme: At the k-th step of the iterative procedure, let S be a putative designing sequence obtained by minimizing $G(S)$ based on the current energy matrix, $\boldsymbol{\varepsilon}^{(k)}$. The external protein structure prediction algorithm may find several structures for the sequence S that are better faring than Γ_{T} as native states. Let $\{\Gamma_0(S), \Gamma_1(S), \ldots, \Gamma_n(S)\}$ be a ranked set of such competing structures, ordered by increasing ground-truth energy, i.e., decreasing confidence score.

Since the structure prediction step is assumed to be reliable (and it certainly is in our minimalistic context where it entails an exhaustive search in structure space), the observation that the competing structures $\Gamma_0, \ldots, \Gamma_n$ have a higher confidence score than Γ_{T} signals the imperfect parametrization of the $\boldsymbol{\varepsilon}^{(k)}$ matrix.

To compensate for this, we move to a new iteration where the energy matrix $\boldsymbol{\varepsilon}^{(k+1)}$ is updated over the k–th one by requiring that $\Gamma_0, \ldots, \Gamma_n$ have a lower energy than Γ_{T}, consistent with the outcome of the ground-truth predictor:

$$E^{(\mathrm{k+1})}(\Gamma_i, S) \leq E^{(\mathrm{k+1})}(\Gamma_{\mathrm{T}}, S), \qquad i \leq n, \tag{3.28}$$

where $E^{(k+1)}$ is the energy function of Eq. (3.4) informed by the interaction matrix $\boldsymbol{\varepsilon}^{(k+1)}$.

Let us consider the case where the canonical weight associated with S disposed along the best ranking structure, Γ_0, exceeds a given threshold, p_{fold}, indicating that, according to Eq. (3.1), Γ_0 is the native structure of S. In this case, we impose that the energy of Γ_0 should be significantly lower than that of all competing structures:

$$E^{(\mathrm{k+1})}(S, \Gamma_i) \geq E^{(\mathrm{k+1})}(S, \Gamma_0) + \Delta(p_{\mathrm{fold}}, \beta)\,, \qquad \forall i > 0\,, \tag{3.29}$$

where $\Delta(p_{\mathrm{fold}}, \beta) = \frac{1}{\beta} \ln \frac{1-p_{\mathrm{fold}}}{p_{\mathrm{fold}}}$ is the minimum energy gap separating ground– and first–excited state energies of a protein-like sequence S — that would fold into its native state with probability $\geq p_{\mathrm{fold}}$ at inverse temperature β.

To derive the term $\Delta(p_{\text{fold}}, \beta)$ contributing to Eq. (3.29), let S be a protein-like sequence having Γ_0 as its native conformation, i.e., satisfying $\frac{-\beta E(S,\Gamma_0)}{\sum_{\Gamma'} e^{-\beta E(S,\Gamma')}} \geq p_{\text{fold}}$. Considering the ordered spectrum of S, $\gamma = \{\Gamma_0, \Gamma_1, \dots, \Gamma_M\}$, we rewrite this expression as

$$1 \geq p_{\text{fold}}\left(1 + e^{-\beta\Delta_{01}} + e^{-\beta\Delta_{02}} + \cdots + e^{-\beta\Delta_{0M}}\right), \tag{3.30}$$

where $\Delta_{0n} = E(S,\Gamma_n) - E(S,\Gamma_0)$. If the previous inequality holds, then

$$\begin{gathered} \frac{1}{p_{\text{fold}}} - 1 \geq e^{-\beta\Delta_{01}} \\ \Updownarrow \\ E(S,\Gamma_1) - E(S,\Gamma_0) = \Delta_{01} \geq \frac{1}{\beta} \ln \frac{1 - p_{\text{fold}}}{p_{\text{fold}}} = \Delta(p_{\text{fold}}, \beta), \end{gathered} \tag{3.31}$$

setting a necessary but not sufficient condition for S to be protein-like at inverse temperature β. As a consequence, the constraint in which $\Delta(p_{\text{fold}}, \beta)$ is involved is quite a loose constraint; nonetheless, it can be employed without requiring to know the details of the ground-truth energy spectrum of S.

Settled this point, we go back to the task of optimizing the coefficients entering Eq. (3.6) to satisfy conditions in Eqs. (3.28) and (3.29). Given that the coefficients to be learned enter linearly in the approximated scoring function $G(S)$, fulfilling the set of inequalities is equivalent to solving a linear separability problem. This task can be conveniently tackled using established algorithms [156–159]. In what follows, we discuss our implementation based on the perceptron technique [156, 160–162], which has been previously used in different protein folding and design contexts to iteratively learn interaction potentials between amino acids [92, 148].

In our algorithm, $\boldsymbol{\varepsilon}^{(k+1)}$ is required to satisfy a heuristic set of conditions that depend on the set $\mathcal{S}^{(k)}_{\text{cum}}$ formed by all the sequences that have been generated by minimizing $G(S)$ up to k–th iteration: $\mathcal{S}^{(k)}_{\text{cum}} = \mathcal{S}^{(k)} \cup \mathcal{S}^{(k-1)} \cup \cdots \cup \mathcal{S}^{(0)}$.

We adopt the notation $\Gamma_i \prec_S \Gamma_{\text{T}}$ if the ground-truth energy of sequence S mounted on Γ_i is lower than the one of S mounted on Γ_{T}. With this, the set of constraints is expressed as

$$\underbrace{\forall \text{ S} \in \mathcal{S}^{(k)}_{\text{cum}}, \quad \Gamma_i \preceq_S \Gamma_{\text{T}}}$$
$$\Downarrow$$
$$E^{(k+1)}(\Gamma_i, S) \leq E^{(k+1)}(\Gamma_{\text{T}}, S),$$

$$\underbrace{\forall \text{ foldable } S \in \mathcal{S}^{(k)}_{\text{cum}}, \quad \forall i > 0}$$
$$\Downarrow$$
$$E^{(k+1)}(\Gamma_i, S) \geq E^{(k+1)}(\Gamma_0, S) + \Delta(p_{\text{fold}}, \beta)\,. \tag{3.32}$$

Using the notation

$$E^{(k+1)}(\Gamma, S) = \sum_{n,m=1}^{D} \left(\sum_{\text{i}>\text{j}=1}^{N} C_{\text{ij}}(\Gamma) q_n^{(\text{i})} q_m^{(\text{j})} \right) \varepsilon^{(k+1)}(n, m) =$$
$$= \mathbf{n}(\Gamma, S) \cdot \boldsymbol{\varepsilon}^{(k+1)}, \tag{3.33}$$

we compactly rewrite the previous constraints as

$$\boldsymbol{\varepsilon}^{(k+1)} \cdot \mathbf{x}_i + c_i \geq 0\,, \tag{3.34}$$

with suitably chosen arrays $\mathbf{x}_i$ and constants c_i. To refine the parameters entering in the designing scoring function so that they minimally violate constraints in Eq. (3.34), we resort to a procedure inspired by Ref. [156].

In particular, the refinement protocol sketched in Fig. 3.2 is initiated by setting the temporary parameters $\boldsymbol{\varepsilon}^*$ to the last available coefficients, i.e., $\boldsymbol{\varepsilon}^{(k)}$. At each refinement iteration, the temporary parameters are used to evaluate the quantities in Eq. (3.34) and to find the most violated inequality, which is used to update the temporary parameters, as follows

$$\boldsymbol{\varepsilon}^* \to \boldsymbol{\varepsilon}^* + \eta\, \mathbf{x}_{i^*}\,, \quad \text{with} \quad i^* = \underset{i}{\operatorname{argmin}}\, \boldsymbol{\varepsilon}^* \cdot \mathbf{x}_i + c_i\,. \tag{3.35}$$

In the previous, η is a regularizing parameter and can be regarded as a “learning rate” in the context of machine learning.

This iterative update proceeds until one of the following conditions occurs: (i) the maximum number of iterations is reached or (ii) all constraints are simultaneously satisfied.

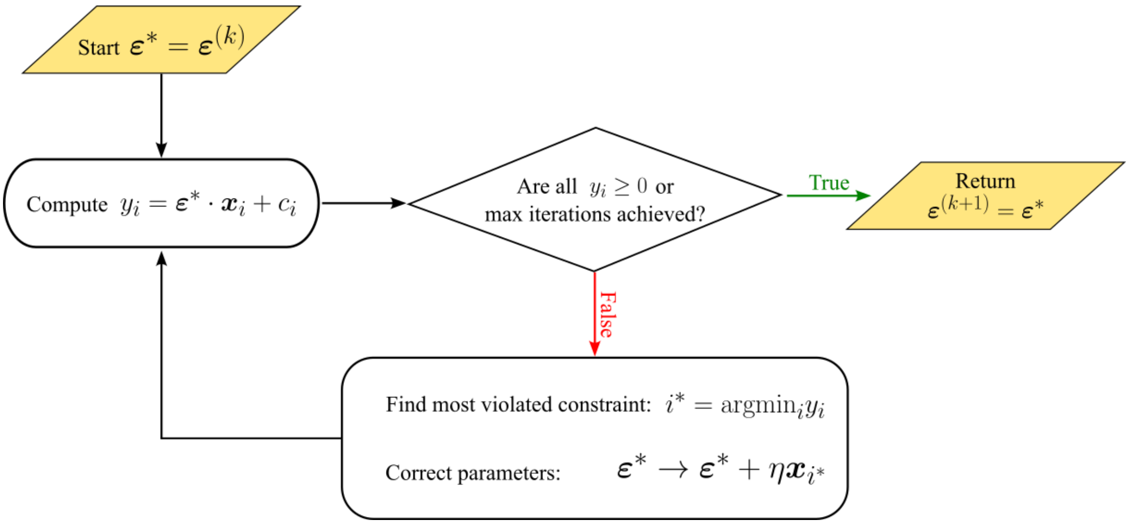


Figure 3.2: This algorithm takes as inputs the parameters entering in the designing function $G(S)$ at k–th cycle and refines them to minimally violate the set of constraints in Eqs. (3.32) and (3.33).

The number of constraints that are simultaneously applied to the parameters of the energy model underlying the definition of $G(S)$ increases at each refinement iteration k. Hence, it is convenient to adapt the parameter η to decrease at each iteration, leading to finer and finer updates of the energy matrix entries. In particular, we set $\eta = \frac{\eta_0}{1+3k}$. The numerical values of η_0 adopted in our simulations are specified in Section 3.5.

3.3 Lattice protein model

We test the algorithm discussed so far on a protein lattice model defined on a two-dimensional lattice (square) and involving coarse-grained chemical identities. The two-dimensional lattice is customarily preferred over the cubic lattice because it offers a more realistic surface-to-volume ratio of compact structures of small length, $\lesssim$ 100 amino acids. We consider sequence alphabets of $D = 3, 4$, and 5 letters and target structures filling 4×4, 5×5, and 6×6 lattices.

In our proof of concept application, we choose to define the ground-truth energy functional as

$$E_{\mathrm{gnd}}(\Gamma, S) = \sum_{\mathrm{i}>\mathrm{j}=1}^{N} C_{\mathrm{ij}}(\Gamma)\, \varepsilon_{\mathrm{gnd}}(s_{\mathrm{i}}, s_{\mathrm{j}})\,, \tag{3.36}$$

where $S = \{s_1, \ldots, s_N\}$ is the primary sequence, which is composed of chemical entities $s_{\mathrm{i}} \in \{1, \ldots, D\}$, $\mathbf{C}(\Gamma)$ is the contact map of the conformation Γ, and $\boldsymbol{\varepsilon}$ is the ground-truth symmetric $D \times D$ matrix.

The ground-truth interaction potential — employed only in the structure prediction step — has to be chosen in a way that as many structures as possible are designable. To achieve this, we developed a systematic approach for selecting the $\boldsymbol{\varepsilon}_{\mathrm{gnd}}$ matrix that defines the ground-truth potential. Specifically, we generate a set of 1,000 candidate matrices, each representing a viable interaction potential. These potentials are then tested by attempting to fold a sample set of sequences. By analyzing the folding outcomes, we assess each candidate's ability to support designable structures. Based on this evaluation, we select the energy matrix that maximizes the number of designable structures. The numerical choices of the ground-truth energy matrices for the different alphabets are detailed in Section 3.5.

3.4 Results

In our iterative design strategy, the optimal parameters of the scoring function $G(S)$ are obtained by comparing the results of direct vs. inverse folding predictions. Several factors may determine the quality of the predictions: (i) the feasibility of minimizing $G(S)$ in the combinatorial space of sequences, (ii) the accuracy of the "external" structure prediction method, (iii) the viability of the functional form of $G(S)$ for yielding accurate design predictions when suitably parametrized, (iv) the feasibility of identifying such optimal parametrizations of $G(S)$ using the iterative scheme.

In this proof-of-concept study, the uncertainties associated with points (ii) and (iii) are ruled out from the outset. Indeed, modeling proteins as compact structures on square lattices makes it possible to perform exhaustive searches in structure space, thus enabling the exact determination of the lowest energy state(s) of any given sequence. In addition, the functional form of the scoring function $G(S)$, namely a pairwise-contacts Hamiltonian, was purposely chosen to match that of the ground-truth Hamiltonian used to pick designable structures as viable targets in Eq. (3.36), thus guaranteeing that suitable parametrizations of $G(S)$ exist and are, in principle, learnable.

In our context, where we shall use alphabets of limited size, $D = 3$, 4, 5, point (i) could be addressed by exhaustive enumeration, similarly to the structure prediction step. However, in realistic contexts the exhaustive search of sequence or structure spaces would be unfeasible. While the structure space search can today be circumvented using the now available rapid and accurate structure prediction methods based on machine learning, the challenge of minimizing $G(S)$ in sequence space still persists. For this reason, we address point (i) by recasting the minimization of $G(S)$ as a QUBO problem, which can be tackled with classical and quantum combinatorial algorithms. As we demonstrate by considering various values of D, such an approach is straightforwardly adapted to amino acid alphabets of any size. In connection to points (iii) and (iv) outlined above, we first assess the viability of the approximate functional form of $G(S)$ as a scoring function to tackle the sequence optimization step. We considered two different contexts. In the first, the ε matrix is not learned but is set equal to that used in the ground truth protein folding predictor. In the second, ε is learned through our iterative procedure.

To carry out this assessment, we take the compact structure Γ_{T} of

Fig. 3.3a as the design target. We use an alphabet of $D = 3$ letters and set the composition to $[N_1 = 5, N_2 = 5, N_3 = 6]$, a choice that combines a sizeable combinatorial space of sequences with the existence of numerous solutions to the design problem.

For both ε choices, we computed $G(S)$ for all sequences with the above composition. We then obtained the receiver-operating curve (ROC), $y(x)$, where x is the rank index for increasing $G(S)$ and y indicates what fraction of the exhaustive set of design solutions are found up to that value of the scoring function.

A perfect design performance would yield the steepest ramping ROC curve, where all the design solutions are exclusively at the highest-ranking positions (lowest values of the scoring function). Accordingly, a customary measure of ROC performance is the so-called *normalized area under the curve*, Q, i.e., the area between the curve and the diagonal divided by the area of the upper triangle. The aforementioned perfect performance would correspond to $Q \sim 1$. In contrast, in a baseline performance — where solutions are discovered with uniform probability independent of their $G(S)$ ranking — Q would be close to 0.

The results of our ROC analysis are shown in Fig. 3.4. As anticipated, we preliminarily tested our $G(S)$ approximation by plugging the ground-truth potentials in place of the ε matrix. The corresponding ROC curve, shown with a dashed line in Fig. 3.4, shows a near-ideal performance, $Q > 0.99$. This demonstrates that the heuristic scoring function $G(S)$ of Eq. (3.6), which is based on average contact probabilities, is indeed viable

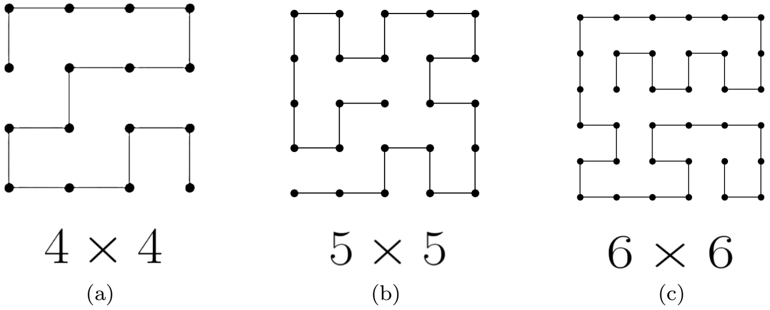


Figure 3.3: Selected target structures for different lattice sizes.

for design purposes, as it can lead to a near-perfect scoring when informed by suitable potentials.

We then moved to the second assessment, aimed at ascertaining if suitable parametrizations of $G(S)$ can be learned by our iterative design procedure starting from arbitrary initializations of the $\boldsymbol{\varepsilon}$ matrix. A further question is how many iterations are required for convergence.

For these tests, we applied the iterative procedure to the same designable target structure starting from 50 different random choices of the initial matrix. The results are summarized in Fig. 3.4, which shows the ROC curves at different iteration stages, averaged over the different initializations — see Fig. 3.5 for the individual ROC curves.

The blue curve in Fig. 3.4 shows the average performance at the beginning of the iterative procedure (labeled cycle 0) when the energy matrix is yet to be learned. The curve is near-diagonal in this case, demonstrating the expected baseline performance. The performance steadily improves at each iteration, converging to a nearly perfect parametrization, corresponding to $Q > 0.99$, in as few as three iterations.

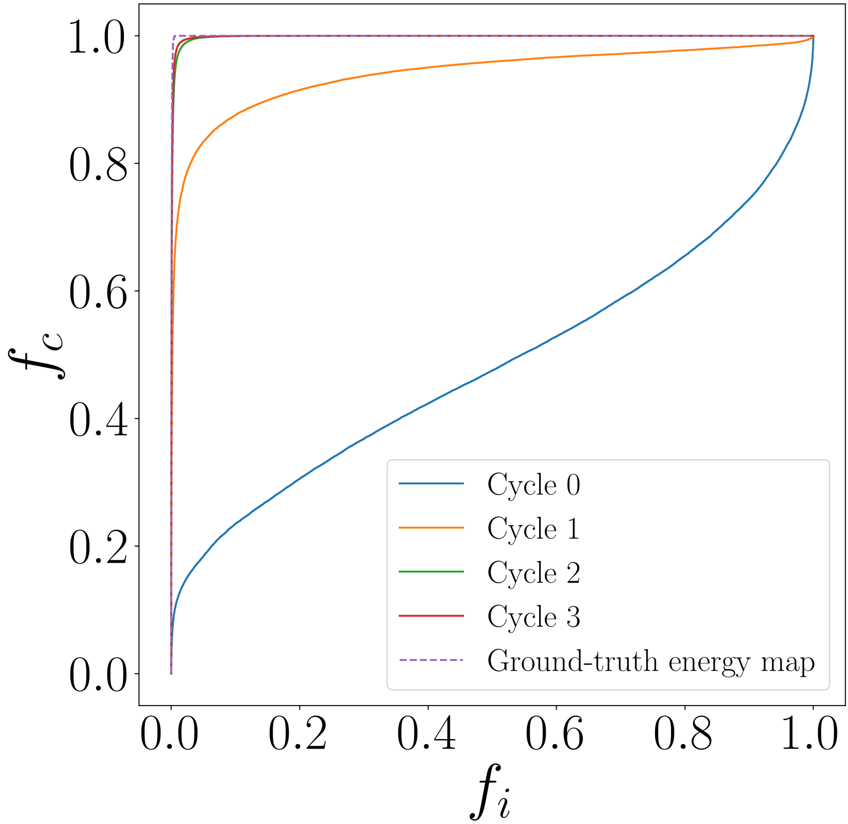
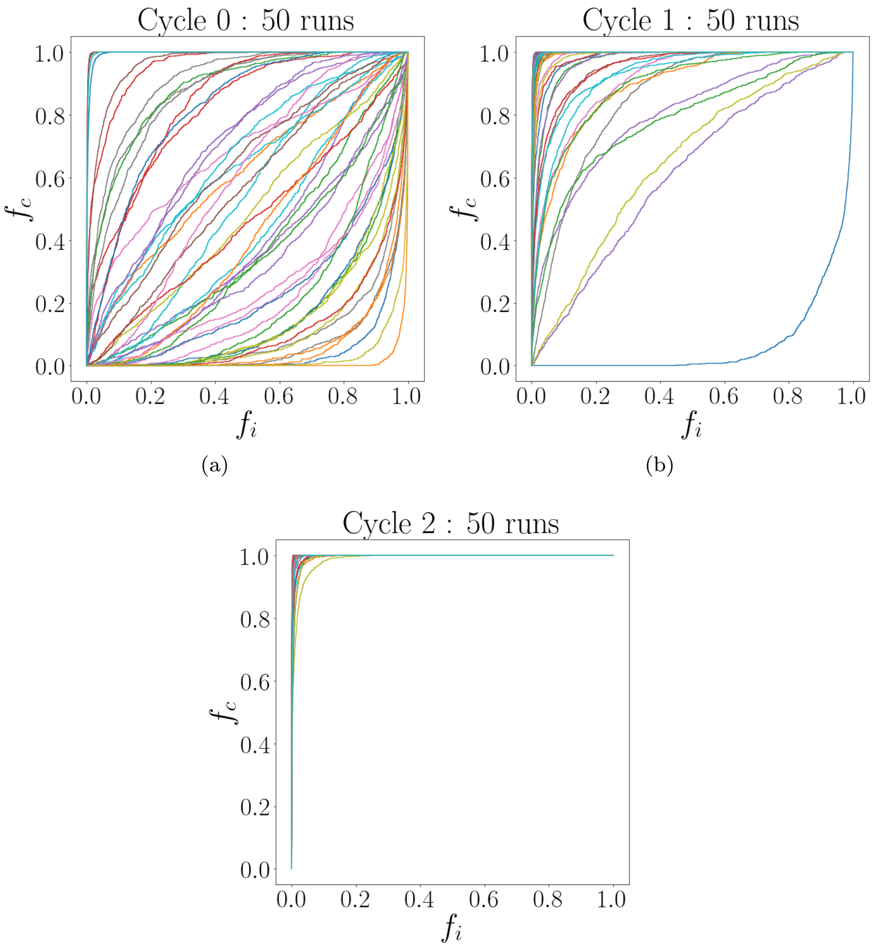


Figure 3.4: **Solid lines:** ROC curve at different iterations tests the goodness of $G(S)$ (see Eq. (3.6)) as a threshold binary classifier distinguishing designs for Γ_T from arbitrary sequences. These results show that classifiers based on $G(S)$ evolve from being nearly random (at cycle 0) to nearly optimal (at cycle 3). **Dashed line:** Plugging the ground-truth energy map in $G(S)$ produces a nearly-optimal classifier.



(a)

(b)

(c)

Figure 3.5: Individual ROC curves of threshold binary classifiers based on the scoring function $G(S)$ at different iteration steps. Each plot accounts for the evolution of 50 parametrizations of $G(S)$ that are randomly initialized at cycle 0. The averaged results are enclosed in Fig. 3.4.

3.4.1 Performance scaling with lattice and alphabet size.

We next turned to larger lattice sizes and amino acids alphabets; see Section 3.5 for numerical details. In such cases, ROC curves are not the best way to assess the design performance as they require exhaustive coverage of sequence space, which becomes rapidly impractical with growing protein length and alphabet size.

Instead, we estimate the design success rate using a sampling scheme. Specifically, at each iteration, we select the 30 best-scoring sequences according to $G(S)$ and compute which fraction of them, f_c, admits the target structure as the unique ground state and satisfies Eq. (3.1).

The results are given in Fig. 3.6a and show that, for all three alphabet sizes considered, $D = \{3, 4, 5\}$, our algorithm reaches a success rate of about 80% after just a few iterations. Notably, the highest performance is achieved with the largest alphabet size, corresponding to 5 amino-amino acid types. Importantly, this trend is robust over different choices of the target structure (see Figs. 3.7a and 3.7b).

In Fig. 3.6b, we report the results of a similar analysis for a fixed alphabet of $D = 3$ letters but for three compact structures filling lattices of increasing sizes. Again, in all cases, the algorithm reaches a plateau after a few iterations. For the systems we have considered, the overall success rate ranges from about 65% up to nearly 100%. While these specific instances do not show a clearly identifiable trend with lattice size, when the analysis is extended to an ensemble of structures, we observe that the success rate decreases with increasing chain size (see Figs. 3.7c and 3.7d).

3.4.2 Comparison of conventional and QUBO-based minimizers.

A key feature of our approach, is that the combinatorial search underpinning the sequence selection step is formulated as a QUBO problem, which is, in principle, amenable to quantum annealers. This poses two questions: (i) Does the QUBO encoding boost the design performance compared to working directly in sequence space? (ii) How do currently available quantum annealers fare at the design task compared to state-of-the-art classical QUBO solvers?

To address these questions, in Fig. 3.8, we report, for compact structures of different sizes, the lowest values of $G(S)$ obtained after many runs

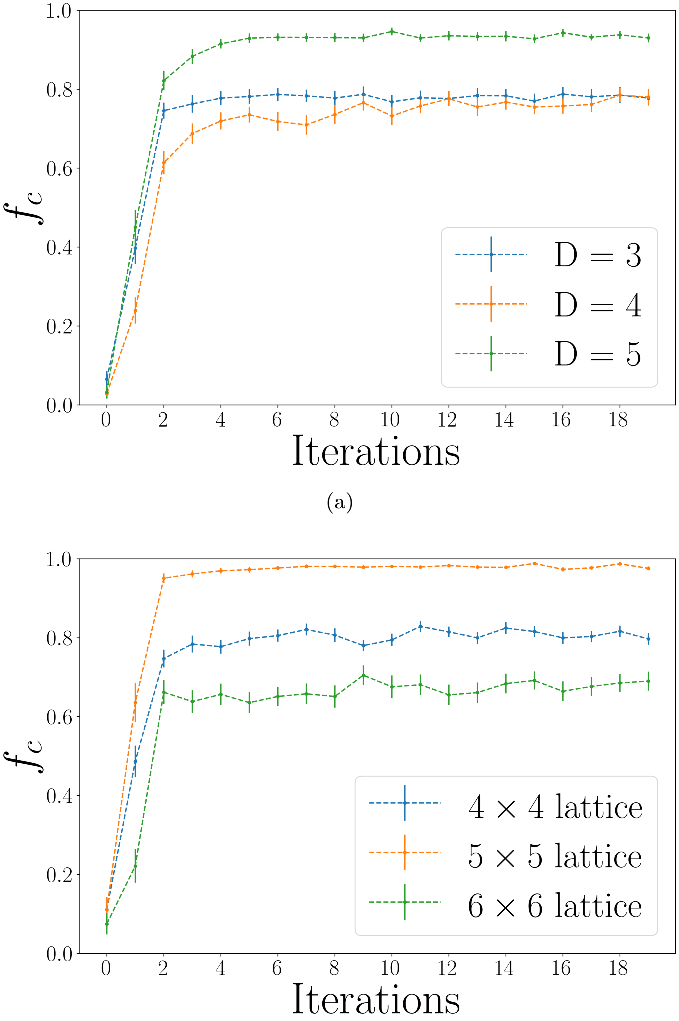
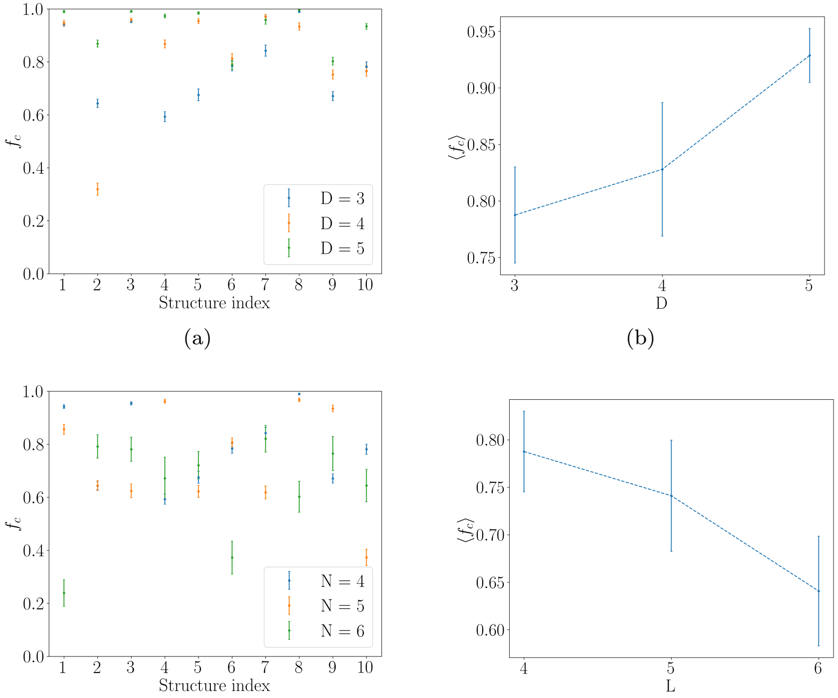


Figure 3.6: Fraction f_c of correctly identified sequences as a function of refinement iterations for different alphabet and lattice sizes. In **(a)**, we consider a fixed structure on a 4×4 lattice, see Fig. 3.3a, and vary the size D of the alphabet. In **(b)**, we fix the alphabet size $D = 3$ and span over different lattice sizes. The corresponding target structures are in Figs. 3.3a, 3.3b, and 3.3c.



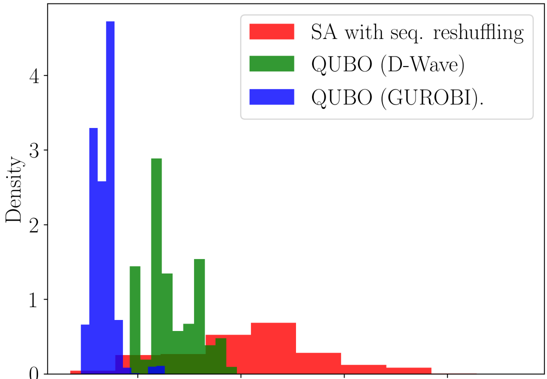
(a)

(b)

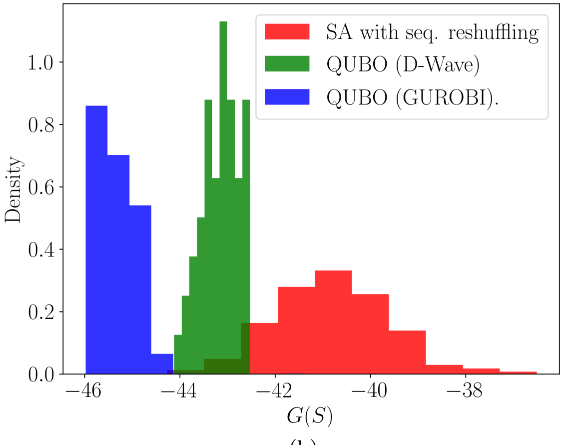
(c)

(d)

Figure 3.7: Fraction of sequences correctly folding to the target structure, f_c, scaling alphabet size and chain size. **(a)** We test our algorithm using 10 target structures lying on a 4×4 lattice with a variable alphabet size, D. **(b)** Average values associated with data in (a), indicating that our algorithm improves when considering larger alphabets. **(c)** Maintaining the alphabet size fixed to 3, we test our algorithm on different lattice sizes, $L \times L$, where for each lattice format, we select 10 target structures. **(d)** Average values associated with data in (c), showing that, as the protein size increases, the performance of our algorithm decreases.



(a)



(b)

Figure 3.8: Statistics of the $G(S)$ values resulting from the use of different optimization approaches (simulated annealing, hybrid annealing on D-Wave, and GUROBI optimizer). In particular, we represent 1000 samples obtained by using simulated annealing, 1000 using the GUROBI optimizer, and 100 using the D-Wave hybrid optimizer. In **(a)** we consider a structure on a 9×9 lattice, while in **(b)** we consider a 13×13 lattice.

of classical optimizations of the scoring functions, parametrized with the ground-truth potentials, with different encodings and hardware at equal duration (3s). The red histogram corresponds to the results of simulated annealing directly formulated in sequence space, where the moves correspond to composition-preserving reshufflings of the sequence. Instead, the blue curve was obtained using GUROBI, an industry-grade QUBO solver. Finally, the green histogram reports the results of 3s runs on D-Wave using the hybrid classical-quantum solver, given that the complexity of the problem at hand exceeds the size currently addressable with fully-quantum annealing algorithms.

The most striking feature of these results is that the $G(S)$ distributions of the minimizers based on the QUBO formulation (green and blue) extend well below the lower tail of the distribution generated with a conventional optimization based on the combination of sequence reshuffling and simulated annealing (red). Hence, the QUBO reformulation required to harness quantum computing technologies has generated a major improvement in the sequence optimization step even when adopted on classical machines. Notably, this difference in performance is enhanced for the largest lattice size.

Focusing on QUBO solvers, we note that the best performance is achieved by GUROBI, an entirely classical scheme based on heuristic searches. This result highlights the maturity reached by classical optimizers following from decades of hardware and software development. At the same time, we emphasize that the hybrid scheme implemented in D-Wave interleaves classical and quantum steps with internal criteria that are not easily controllable by the user. Thus, the results of the hybrid algorithm arguably represent a lower bound of the performance achievable by optimal combinations of classical and quantum steps.

3.5 Numerical details

In this section, we provide the parameters that are necessary to reproduce the results reported in Section 3.4. The model used in our application is intended to serve as a comprehensive benchmark for our approach, rather than to investigate the protein design problem within a specific physical regime. As a consequence, the choice of parameters has to be regarded as purely illustrative.

The inverse thermal energy parameter β was arbitrarily fixed to 3 (in appropriate units), the probability threshold defining foldable sequences p_{fold} was set to 0.8, and the ground truth energy maps, for dictionaries $\mathcal{D}_3, \mathcal{D}_4$, and $\mathcal{D}_5$ are, respectively,

$$\boldsymbol{\varepsilon}_3 = \begin{pmatrix} -0.35346 & 0.30399 & 0.42582 \\ 0.30399 & 0.17115 & -0.30167 \\ 0.42582 & -0.30167 & 0.34102 \end{pmatrix} \tag{3.37}$$

$$\boldsymbol{\varepsilon}_4 = \begin{pmatrix} 0.05375 & 0.21861 & 0.00656 & 0.14191 \\ 0.21861 & 0.43261 & -0.50441 & -0.5146 \\ 0.00656 & -0.50441 & 0.23041 & 0.34485 \\ 0.14191 & -0.5146 & 0.34485 & 0.34976 \end{pmatrix} \tag{3.38}$$

$$\boldsymbol{\varepsilon}_5 = \begin{pmatrix} -0.05777 & 0.26095 & -0.00228 & 0.26162 & 0.0197 \\ 0.26095 & 0.14214 & -0.37257 & 0.13965 & 0.18096 \\ -0.00228 & -0.37257 & 0.04771 & 0.12568 & 0.11891 \\ 0.26162 & 0.13965 & 0.12568 & -0.38521 & 0.02284 \\ 0.0197 & 0.18096 & 0.11891 & 0.02284 & -0.32999 \end{pmatrix}. \tag{3.39}$$

Results in Fig. 3.6a of the main text are derived by designing the target structure in Fig. 3.3a employing sequences with compositions $\{5, 5, 6\}$, $\{5, 4, 2, 5\}$, and $\{3, 3, 2, 4, 4\}$. Fig. 3.6b of the main text results when designing targets in Figs. 3.3a to 3.3c considering, respectively, sequences with composition $\{5, 5, 6\}$, $\{7, 9, 9\}$, and $\{12, 18, 6\}$. The parameter η_0 employed to refine the energy map in Eq. (3.35) is set to 0.325 for $D = 3$, to 0.288 for $D = 4$, and to 0.263 for $D = 5$.

3.6 Pathway towards realistic applications

In this section, we outline one possible route for transferring the method discussed so far to realistic off-lattice design contexts. With reference to the flowchart of Fig. 3.1, this endeavor would involve three main challenges: (i) the generalization of the design scoring function of Eq. (3.6) to off-lattice models (ii) the use of tools such as AlphaFold for the structure prediction and energy refinement steps, and (iii) the generalization of the energy and free-energy terms of Eqs. (3.4) and (3.5) to account for the amino acid structure and cooperative interactions. In addition to these conceptual points, one should additionally consider on how the extension from a minimalistic to a realistic setting can impact the required computational resources.

Generalization to an off-lattice protein representation: Our protein design strategy is based on the scoring function of Eq. (3.6), where the structural information is encoded through contact maps. The discretized nature of lattice configurations is only reflected by the binary nature of the first matrix C_{ij} in Eq. (3.6), i.e., the contact map of the target structure. However, the binary nature of C_{ij} is not a requirement of our method. This is manifest by the fact that the second matrix in Eq. (3.6) is an average contact map, and hence real-valued even in lattice contexts. In fact, the structural encoding based on contact maps is inherently robust and general because it does not hinge on explicit Cartesian coordinates representations nor the discreteness of the embedding space, and not even its dimensionality. Therefore, our formulation of the design scoring function is manifestly transferrable to off-lattice models.

Use of AlphaFold to perform protein structure predictions: The minimalistic lattice model used in our proof-of-concept application enabled us to obtain protein folding predictions from a pre-determined (ground-truth) energy function, resorting to the exhaustive exploration of all compact protein structures.

The generalization from minimalistic to realistic models requires a reliable tool to perform protein structure predictions. AlphaFold provides the most accurate option for this task, to date. In this scheme, the native structures are predicted by a deep neural network trained on a databank, not by a physics-based model. This structure prediction tool can be seamlessly integrated in our energy-refinement step. In fact, Alphafold calls do not return a single structure, but rather a set of ranked structure predictions. The ranking is based on a physical measure (pLDDT) that reflects the propensity of amino acids to adopt local structures different from the target one. Thus, the ranked set of structures and their pLDDt scores can be seamlessly used to select the competing structures to use in the energy-refinement step based on Eq. (3.28).

Improvement of the energy model in the scoring functional: Since the goal of the protein design problem is to identify sequences that yield a given *fold*, it is natural to resort to a scoring function based on a coarse-grained representation of the chain. Our minimalistic energy model is already equipped to account for pairwise interactions of the coarse-grained amino acids. However, in realistic applications, two-body interactions would not typically suffice, and effective many-body interactions would be needed to implicitly account for, e.g., internal degrees of freedom of the amino acids, such as rotameric states of the sidechains. While this multi-body approach is entirely general, its viability in the protein design context is ensured by Anfisen's principle, which guarantees that it is possible to define a scoring function, depending on an effective energy E, that discriminates between native and competing folds on the basis of the sole sequence information. Indeed, such expansions have been exploited before in accurate coarse-grained protein contexts, such as the UNited RESidue models developed by Scheraga and coworkers [163, 164]. Thus, both the energy and free-energy expressions of Eqs. (3.4) and (3.5) may be extended through a many-body expansion,

$$E = E_1 + E_2 + E_3 + \dots, \tag{3.40}$$

where E_k represents the k-body interaction.
As noted above, the minimalistic model discussed so far was based on retaining only the E_2 contribution. The inclusion in Eqs. (3.4) and (3.5) of unary terms, such as those needed to account for amino acid hydrophathies or Ramanchandran angle constraints, would be straightforward. Including three-body terms would additionally enable accounting, at least in part, for effects arising from rotameric degrees of freedom, as well as *bona fide* cooperative interactions. To this end, a natural form of three-body terms to include in Eqs. (3.4) and (3.5) would be

$$E_3 = \sum_{\mathrm{ijk}} \Gamma_{\mathrm{ijk}}(\Gamma)\, \epsilon_{q(\mathrm{i})q(\mathrm{j})q(\mathrm{k})} \tag{3.41}$$

In this equation, $\Gamma_{\mathrm{ijk}}(\Gamma) \in [0,1]$ is a rank-3 tensor determined entirely by the structure upon which a sequence $q(1), \dots, q(N)$ is mounted. This tensor is defined to approach 1 only when residues i, j, and k are simultaneously within a cutoff distance. Similar expressions can be provided for the order N terms and involve rank-N tensors.

Computational challenges for realistic protein design: Upgrading from a minimalistic to a realistic calculation will impact the computational cost of our algorithm. Most importantly, including many-body interactions

greatly enlarges the space of parameters to be learned iteratively, thus requiring a larger training data set.
A possible strategy to balance the accuracy of the energy map and the computational cost of learning it consists of using physico-chemical insight to restrict the structure of the many-body interactions, leaving only a few phenomenological parameters to be learned.
For example, one may harness three-body interactions to penalize the colocalization of triplets of amino acids that are seldom observed to be mutually proximal in protein data bank entries. This way, all such interactions would be weighted by the same coupling constant, i.e., an energy penalty. This practical and feasible scheme would greatly reduce the combinatorial search space due to the incorporation of phenomenological information, including steric constraints depending on rotameric degrees of freedom and cooperative amino acid interactions.

The inclusion of many-body terms would also affect the quantum encoding required to perform the sequence optimization step on a quantum annealing machine. Indeed, the current QUBO formulation is natively predisposed to deal with one- and two-body interactions.
However, higher-order interactions can still be introduced by resorting to ancillary variables, as explicitly shown, e.g., in [139]. We expect that such demand of increased number of qubits will be met by the rapid growth in size and performance of quantum annealing machines. Our basic (two-body interactions only) QUBO formulation would require no more than about 2000 qubits for representing sequences of 100 amino acids with the full 20-letter chemical alphabet. For reference, the recent ground-breaking applications of D-Wave to a physics-based problem [14] have employed 5000 qubits. We expect the next generations of quantum annealers to leave adequate room for accounting for a sizable number of many-body interactions in realistic design problems.

Finally, regarding the computational cost of the structure prediction step, we note that a single call of AlphaFold-2 using the dedicated web server takes about 10 minutes to return the structure of a globular protein of about 100 amino acids. This time is shorter than the one required in this work to perform the exhaustive search in the largest lattices considered, thus supporting the viability of integrating neural-network-based structure prediction tools.

3.7 Conclusion and perspectives

In this proof-of-concept study, we have shown that the availability of reliable algorithms for protein structure predictions can be capitalized to envision efficient strategies for tackling the protein design problem. Our iterative method has two distinctive features. First, the structure prediction algorithm is used to learn an optimal scoring function for the design problem instead of using postulated models and interaction parameters. Second, having mapped the sequence selection step to a combinatorial QUBO problem allows for addressing the design problem by harnessing existing powerful classical optimizers and promising quantum technologies.

Strikingly, we found that the QUBO encoding brings *per se* a significant improvement, to the point that matching the performance of classical or quantum QUBO optimizers with conventional schemes becomes computationally impractical even for modest protein lengths.

In our first illustrative benchmarks, we resorted to exhaustive enumeration to remove the uncertainty associated with heuristic protein folding predictors. This choice had the downside of considering simplified lattice models and relatively small chains and alphabets. However, we note that the key theoretical ingredient of our method is the scoring function $G(S)$ that is entirely specified in terms of contact maps, regardless of the specific representation of protein conformations. As such, our method is primed to be generalized to off-lattice contexts.

State-of-the-art machine-learning predictors such as AlphaFold [127] provide the key tool to upgrade our approach to realistic applications. Notably, these schemes return accurate folding solutions in a time much shorter than what is required by our exhaustive enumeration in lattice models.

Furthermore, the trajectory in addressing the technological limitations of the existing quantum hardware has been impressive [150, 165]. This gives hope that significantly more complex problems will become solvable in the near future, opening the possibility of using more accurate energy models and optimizing $G(S)$ in the realistic sequence alphabet space.

Succeeding at this task will be transformative not only for protein design. However, it would also pave the way towards a broad range of related applications, e.g., protein origami, drug screening, and *de novo* drug design.

Chapter 4

Statistical mechanics of heteropolymers from lattice gauge theory

Lattice polymers have been extensively studied to gain qualitative insight into the statistical physics of a wide class of soft matter systems [106]. In particular, compact structures of lattice heteropolymers have emerged as paradigmatic simplified models for protein native states. While efficient algorithms have been developed for dilute polymer ensembles [166–168], compact structures of heteropolymers have proven significantly more challenging to sample [169, 170]. Indeed, growth-based methods become exponentially inefficient, and the low acceptance rate of pivot moves hinders Monte Carlo (MC) schemes. Combinatorial algorithms derived from graph theory can be more efficient in generating compact structures, particularly Hamiltonian paths [171,172], but do not satisfy ergodicity. Recently developed advanced MC schemes overcome this limitation [173, 174]. However, despite these advances, generating uncorrelated samples of large compact heteropolymer lattice structures remains challenging.

In recent years, the development of quantum computing hardware has inspired alternative approaches that rely on a binary encoding of lattice polymer configurations [138,175–177]. The upgrade from classical to quantum encoding involves replacing each binary variable with a two-level quantum system (qubit). In the classical and the quantum encoding, the task of sampling the ergodic surface can be formulated as a Quadratic Unconstrained Binary Optimization (QUBO) problem, enabling the characterization of remarkably complex polymer ensembles, such as self-assembled ring melts [177]. Unfortunately, as discussed in Section 2.2.2, the computational advantage of the QUBO encoding rapidly degrades when it is applied to ensembles that do not contain ring polymers. In these cases, many new hard constraints and ancillary variables must be introduced to remove rings up to a given size. The number of these interactions and related ancillary variables diverges in the thermodynamic limit, where rings of any size can occur. In addition, the QUBO approach cannot be directly applied to sample heteropolymer ensembles, because the viable conformational states are not confined to a ground space. In principle, this problem may be overcome by resorting to MC schemes. However, naive local trial moves based on randomly flipping binary variables would not preserve the chain's topology, leading to exceedingly high rejection rates.

In this work, we overcome all these limitations by introducing a new quantum encoding, in which the partition function is mapped into a vacuum expectation value of a quantum field theory with both fermionic and bosonic degrees of freedom. The dynamics of this theory is shaped by its $\mathbb{Z}_2$ gauge symmetry, which expresses the chain's continuity condition and inspires MC moves that preserve the chain's topology. By tuning the coupling constants of our LGT, we can vary the ring density or even set it to zero. As we show below, this scheme is particularly apt to sample compact polymer configurations. Furthermore, this LGT can be naturally extended to study heteropolymers by adding chemistry-specific soft interactions. Finally, this encoding relies on qubits, so it provides a suitable framework to develop future quantum computing algorithms.

4.1 From polymers to LGT

To establish our mapping, we begin by assigning a qubit state $|\Gamma_{\mathrm{ij}}\rangle$ to each edge connecting *neighboring* lattice sites i and j (see Fig. 4.1) and a qubit state $|\Gamma_{\mathrm{i}}\rangle$ to each lattice site i.

The presence (absence) of a covalent bond between residues located at neighboring sites i and j is signaled by $|\Gamma_{\mathrm{ij}}\rangle = |1\rangle$ ($|0\rangle$), which we call *active* (*inactive*) bond state. Conversely, a chain terminal occupies site i if $|\Gamma_{\mathrm{i}}\rangle = |1\rangle$. We also define the bond operator $\hat{B}_{\mathrm{ij}} \equiv |1\rangle\langle 1|$ (projector on active bonds) and the site operator $\hat{g}_{\mathrm{i}} = |1\rangle\langle 1|$ (projector on active sites).

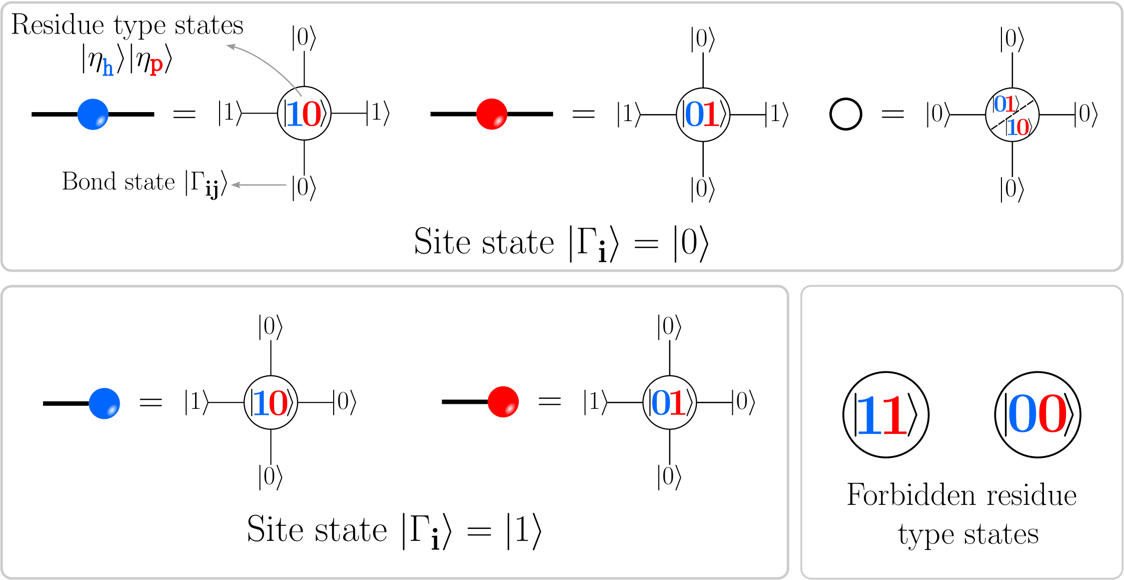


Figure 4.1: In the upper panel, the central site is not a chain's endpoint, i.e., it corresponds to the inactive site state $|0\rangle$. On the contrary, in the lower panel, the central site is a chain's endpoint, i.e., it is related to the active site state $|1\rangle$. Active links are in the $|1\rangle$ state and correspond to covalent bonds, marked with bold lines. In our quantum encoding, illustrated for the so-called HP model outlined in Section 4.5, residues are grouped into hydrophobic (h), depicted in blue, and polar (p), marked in red. The colored circles correspond to sites connected to active bonds (i.e., belonging to the heteropolymer chain). The chemical identity is specified through the *residue type state* $|\eta_h\rangle\,|\eta_p\rangle$, which is allowed to take only two states, namely $|1_h\rangle\,|0_p\rangle$, which identifies a h-type residue, or $|0_h\rangle\,|1_p\rangle$, which corresponds to a p-type residue.

We now consider the tensor product states of all the qubits, $|\Psi\rangle = \left(\prod_{\otimes\langle \mathrm{i,j}\rangle} |\Gamma_{\mathrm{ij}}\rangle\right) \otimes \prod_{\mathrm{i}} |\Gamma_{\mathrm{i}}\rangle$, constituting our computational basis. In a subset of these states, the bond qubits in the $|1\rangle$ state align to form continuous paths in the d-dimensional lattice. We now show that these so-called *polymer states* are solutions of

$$\prod_{k=1}^{d} (-1)^{\hat{B}_{\mathrm{i\,i}+\mathbf{e}_k}} (-1)^{\hat{B}_{\mathrm{i\,i}-\mathbf{e}_k}} |\Psi\rangle = (-1)^{\hat{g}_{\mathrm{i}}} |\Psi\rangle, \tag{4.1}$$

where $\{\mathbf{e}_k\}_{k=1,\ldots,d}$ is the set of lattice basis vectors.

To prove Eq. (4.1), we first note that the computational basis states are eigenstates of the operators at the two sides of the equation. The corresponding eigenvalues obey

$$\prod_{k=1}^{d} (-1)^{\Gamma_{\mathrm{i\,i}+\mathbf{e}_k} + \Gamma_{\mathrm{i\,i}-\mathbf{e}_k}} = (-1)^{\Gamma_{\mathrm{i}}}. \tag{4.2}$$

Equation (4.2) expresses a flux conservation condition associated with a binary field Γ_{ij} that is "emitted" and "absorbed" at the chain endpoints (see also Fig. 4.2). To see this, consider first a generic site i where $\Gamma_{\mathrm{i}} = 0$. There, Eq. (4.2) is fulfilled only if the number of active bonds connected to i, given by the sum

$$\rho_{\mathrm{i}} = \sum_{k=1}^{d} \left(\Gamma_{\mathrm{i\,i}+\mathbf{e}_k} + \Gamma_{\mathrm{i\,i}-\mathbf{e}_k}\right) , \tag{4.3}$$

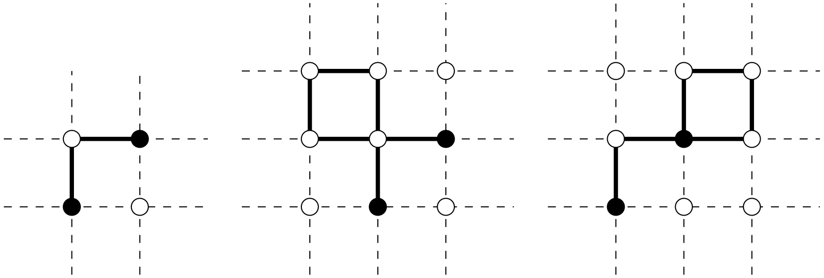


Figure 4.2: Example of continuous polymer configurations satisfying Eq. (4.2) at every vertex. Indeed, the configuration of each vertex is locally among the allowed ones illustrated in Fig. 4.3.

is even. In particular, for 3-dimensional cubic lattice it is equal to 2 along the chain's backbone and 4 or 6 at the intersection point between chains. Conversely, on the same lattice, the number of active bonds attached to the sites with eigenvalue $\Gamma_{\mathrm{i}} = 1$ (chain's endpoints) is equal to 1, or to 3 and 5 if chain intersections occur at the endpoint. A graphical representation of the analog constraints for a 2-d square lattice is reported in Fig. 4.3.

The continuity Eq. (4.1) is in fact the Gauss' law of the $\mathbb{Z}_2$ LGT introduced by Wegner and Kogout in Ref. [178], with $\hat{B}_{\mathrm{ij}}$ playing the role of the electric operator and $\hat{g}_{\mathrm{i}}$ that of the topological charge operator. Inspired by this connection, we also introduce a so-called *chain deformation operator* $\hat{C}_\Box$, which plays the role of the magnetic operator of the $\mathbb{Z}_2$ LGT, and is defined on each plaquette on the lattice as

$$\hat{C}_\Box = \prod_k \hat{X}_k \qquad \text{with} \quad \hat{X} = |1\rangle\langle 0| + |0\rangle\langle 1| \ , \tag{4.4}$$

where k labels the edges defining the given plaquette $\Box$. Realizing the existence of an underlying gauge symmetry is the key to devise a new type of MC algorithms to sample polymer configurations. Indeed, we note that $\hat{C}_\Box$ commutes with the operator on the left-hand side of Eq. (4.1). This implies that if $|\Psi\rangle$ is a polymer state, then $\hat{C}_\Box|\Psi\rangle$ is a different (i.e.,

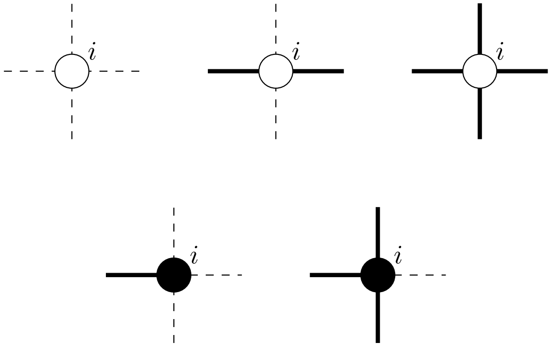


Figure 4.3: Configurations satisfying the eigenvalue Eq. (4.2) at vertex i. Black circles represent end-points with a topological charge present, described by $\Gamma_i = 1$, while white circles represent regular points along the chain backbone, described by $\Gamma_i = 0$. Solid lines represent active bonds ($\Gamma_\ell = 1$) and dashed lines denote inactive bonds ($\Gamma_\ell = 0$). These configurations can be seen as those permitted by a local symmetry (Gauss' law) encoding flux conservation.

locally deformed) polymer state. In practice, $\hat{C}_\square$ flips the state of all qubits assigned to the edges of a plaquette (see Fig. 4.4). Another class of operators that preserve Eq. (4.1) is the chain terminal displacement operator $\hat{D}_{\mathrm{ij}}$ that acts on the qubits at the neighboring sites i and j and on the bond qubit connecting them,

$$\hat{D}_{\mathrm{ij}} = \hat{X}_{\mathrm{ij}}\, \hat{a}_{\mathrm{i}}^{\dagger}\, \hat{a}_{\mathrm{j}} + \mathrm{h.c.}\,, \tag{4.5}$$

where $\hat{a}_{\mathrm{i}} = |0\rangle\langle 1|$. In a LGT, this term represents the motion of charged matter with a corresponding instantaneous adjustment of electric field lines. Two arbitrary polymer configurations with the same number of open chains can be deformed one into the other by a combination of chain deformations and terminal displacements.

Let us further restrict the set of viable tensor product states to retain only *self-avoiding* polymer states. These can be identified with the ground space of the Hamiltonian $\hat{H} = \hat{H}_{\mathrm{Gauss}} + \hat{H}_{\mathrm{SA}}$, with

$$\begin{aligned} \hat{H}_{\mathrm{Gauss}} &= e_0 \sum_{\mathrm{i}} \left((-1)^{\hat{\rho}_{\mathrm{i}}} - (-1)^{\hat{g}_{\mathrm{i}}} \right)^2 , \\ \hat{H}_{\mathrm{SA}} &= e_1 \sum_{\mathrm{i}} \left(\hat{\rho}_{\mathrm{i}} + \hat{g}_{\mathrm{i}} \right)\left(\hat{\rho}_{\mathrm{i}} + \hat{g}_{\mathrm{i}} - 2 \right) , \end{aligned} \tag{4.6}$$

where $\hat{\rho}_{\mathrm{i}} \equiv \sum_{k=1}^{d} \left(\hat{B}_{\mathrm{i}\,\mathrm{i}+\mathbf{e}_k} + \hat{B}_{\mathrm{i}\,\mathrm{i}-\mathbf{e}_k} \right)$ counts the number of active links at

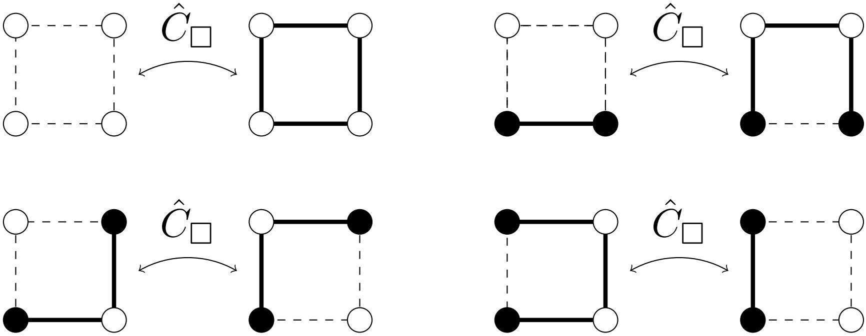


Figure 4.4: Action of the chain deformation operator in Eq. (4.4). Active and inactive bonds around a plaquette are inverted, while the topological charges remain unaltered.

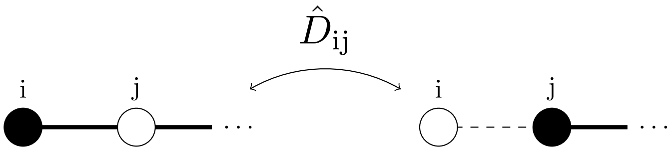


Figure 4.5: Action of the diffusion operator for the chain's endpoints, as given in Eq. (4.5). A topological charge is moved along one link, whose active/inactive state is adjusted such that the vertices assume only configurations (illustrated in Fig. 4.3) that are allowed by the Gauss' law.

the site i. If the couplings e_0 and e_1 are large and $e_0 > e_1$, $\hat{H}$ implements two hard constraints: $\hat{H}_{\mathrm{Gauss}}$ enforces the Gauss' law in Eq. (4.1), while $\hat{H}_{\mathrm{SA}}$ the self-avoidance condition. Indeed, lattice sites where $|\Gamma_{\mathrm{i}}\rangle = |0\rangle$ yield vanishing eigenvalues of $\hat{H}_{\mathrm{SA}}$ and $\hat{H}_{\mathrm{Gauss}}$ as long as they are either not attached to any active bond (empty sites) or they are linked to exactly two active bonds (backbone sites). Lattice sites where $|\Gamma_{\mathrm{i}}\rangle = |1\rangle$ avoid an energetic penalty if they are attached to exactly 1 active bond, thus forming one of the chain's endpoints.

Without loss of generality, in the following we specialize to the case of a single open chain, i.e., we require that only two $|\Gamma_{\mathrm{i}}\rangle$ qubits must be in the $|1\rangle$ state (note, at this stage this restriction still allows for an arbitrary number of additional ring polymers). We denote with $\{|\Psi_s\rangle\}_{s=1,2,\ldots}$ the numerable set of all self-avoiding polymer states (i.e., the subset of degenerate ground states of $\hat{H}$) and introduce the linear sum $|\Omega\rangle \equiv \sum_s |\Psi_s\rangle$. The matrix element $\mathcal{Z} \equiv \langle\Omega|\Omega\rangle$ counts the total number of distinct self-avoiding configurations and, thus, is identified with the classical (grand-canonical) partition function of a system consisting of a single chain and arbitrary rings. Upon inserting the resolutions of the identity in their respective spaces, $\hat{\mathbb{1}} = \sum_{\Gamma_{\mathrm{ij}}} |\Gamma_{\mathrm{ij}}\rangle\langle\Gamma_{\mathrm{ij}}|$ and $\hat{\mathbb{1}}' = \sum_{\Gamma_{\mathrm{i}}} |\Gamma_{\mathrm{i}}\rangle\langle\Gamma_{\mathrm{i}}|$, we obtain an explicit representation:

$$\mathcal{Z} = \sum_{\Gamma} \left(\prod_{\mathrm{i}} \delta\left(\mathcal{G}_{\mathrm{i}}[\Gamma]\right) \right) e^{-\beta H_{\mathrm{SA}}[\Gamma]}, \quad \beta = (k_B T)^{-1}. \tag{4.7}$$

In this expression, we have collectively denoted with Γ the fields defined by the eigenstates of the bond and chain terminal operators. In Eq. (4.7), we have imposed the Gauss' law via a set of delta-functions for the functions $\mathcal{G}_{\mathrm{i}}[\Gamma] \equiv \prod_{k=1}^{d} (-1)^{\Gamma_{\mathrm{i}\,\mathrm{i}+\mathbf{e}_k} + \Gamma_{\mathrm{i}\,\mathrm{i}-\mathbf{e}_k}} - (-1)^{\Gamma_{\mathrm{i}}}$, while the self-avoiding condition follows from choosing a large coupling e_1 in the Hamiltonian H_{SA}, ensuring $\beta e_1 \gg 1$ for all feasible temperatures. The motivation for using two different notations to express these hard constraints is that we can define ergodic trial moves that preserve Gauss' law while we resort to a Metropolis criterion to enforce self-avoidance.

In the ensemble defined by Eq. (4.7), the density of ring polymers and the average length of the open chain is not fixed. We now show that this limitation is overcome by coupling the qubits $|\Gamma_{\mathrm{ij}}\rangle$ and $|\Gamma_{\mathrm{i}}\rangle$ to "spinless fermions" with M degenerate flavors, via the tight-binding Hamiltonian

$$\hat{H}_{\mathrm{F}} = \sum_{\langle \mathrm{i\,j} \rangle} \sum_{m=1}^{M} \hat{\psi}_{\mathrm{i}}^{(m)\dagger} \hat{\psi}_{\mathrm{j}}^{(m)} \hat{T}_{\mathrm{ij}} \,, \tag{4.8}$$

$$\hat{T}_{\mathrm{ij}} = \Big[\delta_{\mathrm{ij}} \big(m_{\mathrm{f}} - \bar{g}^2 \hat{\rho}_{\mathrm{i}} \big) - \lambda^2 \hat{B}_{\mathrm{ij}} \Big] \,, \tag{4.9}$$

where $\hat{\psi}_{\mathrm{i}}^{(m)}$ and $\hat{\psi}_{\mathrm{i}}^{(m)\dagger}$ are anti-communting fields. The motivation for introducing these fermions into the theory is that the density of rings in the ensemble can be tuned by varying the parameters $m_{\mathrm{f}}, \bar{g}$, and λ.

Again, we consider the matrix element $\mathscr{Z}' = \langle \Omega' | \Omega' \rangle$, where now $|\Omega'\rangle$ denotes the linear sum of all the degenerate ground states of $\hat{H}' = \hat{H} + \hat{H}_{\mathrm{F}}$. The explicit expression for $\mathscr{Z}'$ involves a Gaussian path integral over M Grassmann fields, which we carry out analytically to obtain

$$\mathscr{Z}' = \sum_{\Gamma} \left(\det T[\Gamma] \right)^M \prod_{\mathrm{i}} \delta(\mathcal{G}_{\mathrm{i}}[\Gamma]) \, e^{-\beta H_{\mathrm{SA}}[\Gamma]}, \tag{4.10}$$

where $T_{\mathrm{ij}}(\Gamma) = \delta_{\mathrm{ij}} \big(m_{\mathrm{f}} - \bar{g}^2 \rho_{\mathrm{i}} \big) - \lambda^2 \Gamma_{\mathrm{ij}}$.

4.2 Setting the density of ring polymers to zero

In what follows, we explicitly prove that, for a suitable choice of parameters, the determinant arising from integrating out Grassmann fields in Eq. (4.10) completely removes the undesired polymer rings from the statistical ensemble. Let us consider a configuration of the binary field Γ associated with an arbitrary polymer configuration containing an ℓ–long isolated ring (see Fig. 4.6). To build the Fermion bilinear $T(\Gamma)$ defined in Eq. (4.9), we adopt the convenient labeling of sites where the first ℓ nodes are those belonging to the ring, while the following $(N-\ell)$ identify the remaining vertices left on the lattice, i.e., the ones that are not connected to the ring. Due to the fact that the ring is *isolated*, the $T(\Gamma)$ matrix in the chosen basis acquires the block diagonal form

$$T(\Gamma) = \left(\begin{array}{c|c} A_{(\ell)} & 0 \\ \hline 0 & R_{(N-\ell)} \end{array}\right)_{N\times N}, \tag{4.11}$$

where $A_{(\ell)}$ and $R_{(N-\ell)}$ represent, respectively, the ring and an arbitrary configuration supported on the $(N-\ell)$ remaining vertices. The sub-matrix $A_{(\ell)}$ is a matrix having the structure

$$A_{(\ell)} = \begin{pmatrix} m_{\mathrm{f}}-2\bar{g}^2 & -\lambda^2 & 0 & 0 & \boldsymbol{-\lambda^2} \\ -\lambda^2 & m_{\mathrm{f}}-2\bar{g}^2 & -\lambda^2 & 0 & 0 \\ 0 & -\lambda^2 & \ddots & \ddots & 0 \\ 0 & \cdot & \ddots & m_{\mathrm{f}}-2\bar{g}^2 & -\lambda^2 \\ \boldsymbol{-\lambda^2} & 0 & 0 & -\lambda^2 & m_{\mathrm{f}}-2\bar{g}^2 \end{pmatrix}_{\ell\times\ell},$$

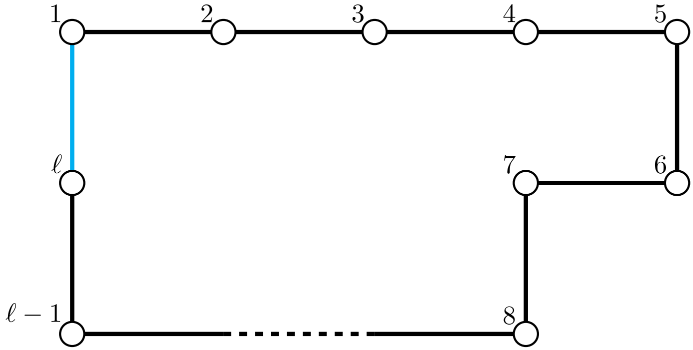


Figure 4.6: Ring configuration consisting of ℓ nodes and ℓ edges. In this picture, we omit the representation of the remaining $(N-\ell)$ nodes that do not belong to the loop, whose contribution to the $T(\Gamma)$ matrix in Eq. (4.11) is embodied by the square sub-matrix $R_{(N-\ell)}$. The contribution of the first ℓ sites belonging to the loop is reflected by the square sub-matrix $A_{(\ell)}$ specified in Eq. (4.12), where the entries marked in cyan are associated with the link joining the first node to the ℓ–th.

where the entries marked in cyan are related to the connection of the first and the ℓ–th node, which is highlighted in Fig. 4.6 with the same color. Crucially, these contributions qualify $A_{(\ell)}$ as a circulant matrix having well-studied properties [179].

In what follows, we will establish the conditions granting $\det T(\Gamma) = 0$ for all those matrices with the above-mentioned form, i.e., those that describe a configuration corresponding to an arbitrarily long isolated ring. The condition $\det T(\Gamma) = 0$ is equivalent to requiring $T(\Gamma)$ to have at least one *vanishing eigenvalue.* Since $T(\Gamma)$ is a block diagonal matrix, a subset of its eigenvalues are those associated with the circulant matrix $A_{(\ell)}$, which can be evaluated exactly (see, e.g., [180]) to be

$$\begin{aligned} &\lambda_k = m_{\mathrm{f}} - 2\bar{g}^2 - 2\lambda^2 \cos\left(\frac{2\pi k}{\ell}\right) \\ &\text{for } k = 1, \dots, \ell\,. \end{aligned}$$

The ℓ-th eigenvalue of $T(\Gamma)$ vanishes by setting

$$m_{\mathrm{f}} = 2\bar{g}^2 + 2\lambda^2\,. \tag{4.12}$$

Among the many possible choices, this condition can be fulfilled by setting

$$m_f = 4\bar{g}^2 \quad \text{and} \quad \bar{g}^2 = \lambda^2\,. \tag{4.13}$$

yielding $\det T(\Gamma) = 0$, for any possible ring number, size and shape.

4.3 Tuning the chain length

In Section 4.2, we explicitly showed that a proper choice of the parameters m_{f}, $\bar{g}^2$, and λ^2, completely removes configurations containing isolated rings from the ensemble, as their canonical weight is zero due to the vanishing of $detT(\Gamma)$.

Once ring polymers have been suppressed, the length of the remaining single chain is fixed by the total number of active bonds and, as we show in the following sections, any given average length ℓ_0 can be set by introducing a *soft* constraint via the additional Hamiltonian term $\hat{H}_{\mathrm{L}} = c\sum_{\langle \mathrm{ij}\rangle}(\hat{B}_{\mathrm{ij}} - \overline{\Gamma})^2$, corresponding to the new matrix element $\mathcal{Z}'' = \langle\Omega'|e^{-\beta\hat{H}_{\mathrm{L}}}|\Omega'\rangle$. The parameter $\overline{\Gamma} \in (0,1)$ controls the average propensity of bond qubits to be in the $|1\rangle$ state. Using the idempotence of $\hat{B}_{\mathrm{ij}}$, one can recast the operator $e^{-\beta \hat{H}_{\mathrm{L}}}$ into the form of a chemical potential term, yielding $\mathcal{Z}'' \propto \langle \Omega' | e^{\beta \mu \hat{L}} | \Omega' \rangle$, where $\hat{L} = \frac{1}{2} \sum_{\mathrm{i}} \hat{\rho}_{\mathrm{i}}$ counts the total number of active bonds and $\mu = 2c(2\bar{\Gamma} - 1)$.

For the sake of clarity, we divide the discussion into two parts; the first delves into the calculations that lead to a general expression for $\det T(\Gamma)$ when Γ describes an open-chain configuration. The second part makes use of this result to fine-tune μ in a way that the ensemble average chain length is fixed to an arbitrary value, ℓ_0.

4.3.1 Determinant of chain configurations

Let us now assume m_{f}, $\bar{g}^2$, and λ^2 are chosen to obey Eq. (4.13), such that we are ensured the absence of any ring. To evaluate Eq. (4.10) we require the numerical value of $\det T(\Gamma)$ for the case we are interested in, i.e., binary field configurations Γ corresponding to an open chain of length ℓ. Remarkably, the structure of the $T(\Gamma)$ matrix does not depend on the specific self-avoiding chain conformation, but only on its length ℓ (see Fig. 4.7):

$$T(\Gamma) = \left(\begin{array}{c|c} B_{(\ell+1)} & 0 \\ \hline 0 & D_{(N-\ell-1)} \end{array} \right) . \tag{4.14}$$

Here, $D_{(N-\ell-1)}$ is a square diagonal matrix with all entries set to $4\bar{g}^2$ and $B_{(\ell+1)}$ takes the form

$$B_{(\ell+1)} = \bar{g}^2 \begin{pmatrix} \mathbf{3} & -1 & \cdot & \cdot & \cdot \\ -1 & 2 & -1 & \cdot & \cdot \\ \cdot & -1 & \ddots & \ddots & \cdot \\ \cdot & \cdot & \ddots & 2 & -1 \\ \cdot & \cdot & \cdot & -1 & \mathbf{3} \end{pmatrix}_{(\ell+1)\times(\ell+1)} . \tag{4.15}$$

With this,

$$\det T(\Gamma) = (4\bar{g}^2)^{N-(\ell+1)} \det B_{(\ell+1)} . \tag{4.16}$$

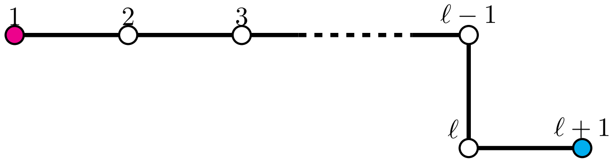


Figure 4.7: Open chain configuration consisting of ℓ edges and $\ell+1$ nodes. In this picture, we omit the representation of the remaining $(N-\ell-1)$ nodes that do not belong to the chain, whose contribution to the $T(\Gamma)$ matrix in Eq. (4.14) is embodied by the diagonal square sub-matrix $D_{(N-\ell-1)}$. The contribution of the first $\ell+1$ sites belonging to the chain is reflected by the square sub-matrix $B_{(\ell+1)}$ specified in Eq. (4.15), where the entries marked in magenta and blue are associated to the extremes of the chain (marked with the same colors in the present figure).

To compute the determinant of $B_{(\ell+1)}$, we recall a general result concerning the determinant of block matrices [181]:

$$\det\left(\begin{array}{c|c} A & C \\ \hline D & B \end{array}\right) = \det A \times \det\left(A - C\,B^{-1}\,D\right)\ . \tag{4.17}$$

To exploit this result, it is convenient to adopt a change of basis such that the transformed $B'_{(\ell+1)}$ acquires the block-diagonal form

$$B'_{(\ell+1)} = \bar{g}^2 \left(\begin{array}{cc|c} 3 & 0 & 0\cdots\cdots 0 \quad -1 \\ 0 & 3 & -1 \quad 0\cdots\cdots 0 \\ \hline 0 & -1 & \\ \vdots & & \\ \vdots & 0 & X_{(\ell-1)} \\ \vdots & \vdots & \\ 0 & \vdots & \\ -1 & 0 & \end{array}\right)_{(\ell+1)\times(\ell+1)}, \tag{4.18}$$

where the matrix $X_{(\ell-1)}$ is a Toepliz matrix having the structure

$$X_{(\ell-1)} = \begin{pmatrix} 2 & -1 & 0 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 & 0 \\ 0 & -1 & \ddots & \ddots & 0 & 0 \\ 0 & 0 & \ddots & 2 & -1 & 0 \\ 0 & 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & 0 & -1 & 2 \end{pmatrix}_{(\ell-1)\times(\ell-1)}. \tag{4.19}$$

The determinant of $X_{(\ell-1)}$ can be computed through successive applications of Eq. (4.17) and is found to have the simple expression

$$\det X_{(\ell-1)} = \ell\,, \tag{4.20}$$

so that, by applying once more Eq. (4.17), the determinant of $B'_{(\ell+1)}$ becomes

$$\begin{aligned}\det B'_{(\ell+1)} &= (\bar{g}^2)^{\ell+1} \det X_{(\ell-1)} \times \\ &\times \det \begin{pmatrix} 3-(X^{-1}_{(\ell-1)})_{\ell-1,\ell-1} & -(X^{-1}_{(\ell-1)})_{\ell-1,1} \\ -(X^{-1}_{(\ell-1)})_{1,\ell-1} & 3-(X^{-1}_{(\ell-1)})_{1,1} \end{pmatrix}.\end{aligned} \tag{4.21}$$

The column vectors of the inverse matrix needed to evaluate Eq. (4.21) have the simple form

$$\left(X^{-1}_{(\ell-1)}\right)_{k,1} = \frac{\ell-k}{\ell} \tag{4.22}$$

$$\left(X^{-1}_{(\ell-1)}\right)_{k,\ell-1} = \frac{k}{\ell}, \quad k = 1,\ldots,\ell-1\,. \tag{4.23}$$

To check Eqs. (4.22,4.23), it is sufficient to express the j-th row of $X_{(\ell-1)}$ as

$$\left(X_{(\ell-1)}\right)_{j\cdot} = \begin{cases} 2\delta_{1,k} - \delta_{2,k} & \text{if} \quad j=1\,, \\ 2\delta_{j,k} - \delta_{j,k+1} - \delta_{j+1,k} & \text{if} \quad 1<j<\ell-1\,, \\ 2\delta_{\ell-1,k} - \delta_{\ell-2,k} & \text{if} \quad j=\ell-1\,, \end{cases} \tag{4.24}$$

so that, plugging Eq. (4.22) in the expression $\left(X_{(\ell-1)}\right)_{j,k}\left(X^{-1}_{(\ell-1)}\right)_{k,1}$, we obtain

$$\begin{aligned}&\sum_k \left(X_{(\ell-1)}\right)_{jk} \frac{\ell-k}{\ell} = \\ &= \begin{cases} \frac{2(\ell-1)}{\ell} - \frac{\ell-2}{\ell} = 1 & \text{if} \quad j=1 \\ \frac{2(\ell-j)}{\ell} - \frac{\ell-j-1}{\ell} - \frac{\ell-j+1}{\ell} = 0 & \text{if} \quad 1<j<\ell-1 \\ \frac{2(\ell-(\ell-1))}{\ell} - \frac{\ell-(\ell-2)}{\ell} = 0 & \text{if} \quad j=\ell-1\,, \end{cases} \\ &= \delta_{j,1}\end{aligned} \tag{4.25}$$

confirming that Eq. (4.22) is correct. To verify that Eq. (4.23) correctly reproduces $\left(X_{(\ell-1)}\right)_{j,k}\left(X^{-1}_{(\ell-1)}\right)_{k,\ell-1} = \delta_{j,\ell-1}$, we evaluate

$$\begin{aligned}
&\sum_k \left(X_{(\ell-1)}\right)_{j,k}\frac{k}{\ell} = \\
&= \begin{cases} \frac{2}{\ell} - \frac{2}{\ell} = 0 & \text{if} \quad j = 1 \\ \frac{2j}{\ell} - \frac{j+1}{\ell} - \frac{j-1}{\ell} = 0 & \text{if} \quad 1 < j < \ell - 1 \\ \frac{2(\ell-1)}{\ell} - \frac{\ell-2}{\ell} = 1 & \text{if} \quad j = \ell - 1 \end{cases} \\
&= \delta_{j,\ell-1}\,.
\end{aligned} \tag{4.26}$$

Since $B_{(\ell+1)}$ and $B'_{(\ell+1)}$ are related by a change of basis, they share the same determinant, and we finally obtain

$$\begin{aligned}
\det B_{(\ell+1)} &= (\bar{g}^2)^{\ell+1}\,\ell\,\det\begin{pmatrix} 2-\frac{1}{\ell} & -\frac{1}{\ell} \\ -\frac{1}{\ell} & 2-\frac{1}{\ell}\end{pmatrix} \\
&= 4(\bar{g}^2)^{\ell+1}(\ell+1)\,.
\end{aligned} \tag{4.27}$$

Thus,

$$\det T(\Gamma) = 4^{N-\ell}\,\bar{g}^{2N}\,(\ell+1). \tag{4.28}$$

4.3.2 Tuning of average chain length

The result in Eq. (4.28) enables us to tune the chemical potential μ acting on link variables to yield any chosen average chain length. To this end, we recall that the partition function in Eq. (4.10)

$$\mathcal{Z}' = \sum_{\Gamma}(\det T(\Gamma))^M \prod_i \delta[\mathcal{G}_i] e^{-\beta H_{\text{SA}}}\,, \tag{4.29}$$

in the limit of $\beta e_1 \gg 1$, can we rewritten as

$$\mathcal{Z}' = \sum_{\Gamma_{\text{SA}}}(\det T(\Gamma_{\text{SA}}))^M \prod_i \delta[\mathcal{G}_i]\,, \tag{4.30}$$

where the summation is over the ensemble of self-avoiding lattice states Γ_{SA}.

By introducing the contribution given by the chemical potential operating on the link variables, we obtain

$$\mathcal{Z}'' = \sum_{\Gamma_{\text{SA}}}(\det T(\Gamma_{\text{SA}}))^M \prod_i \delta[\mathcal{G}_i]\, e^{\beta\mu\ell}\,, \tag{4.31}$$

with ℓ being the total number of active links. If we consider the specific case of a summation over ensembles featuring a single chain $\Gamma_{\mathrm{SA}}(\mathrm{i_i}, \mathrm{i_j})$ with endpoints $\mathrm{i_i}$ and $\mathrm{i_f}$ (i.e., lattice sites occupied by topological charges and thus associated to the states $|\Gamma_{\mathrm{i_i}}\rangle = |1\rangle$ and $|\Gamma_{\mathrm{i_f}}\rangle = |1\rangle$), Eq. (4.31) reads

$$\mathcal{Z}''_{(1)} = \sum_{\mathrm{i_i},\mathrm{i_f}} \sum_{\Gamma_{\mathrm{SA}}(\mathrm{i_i},\mathrm{i_f})} (\det T(\Gamma_{\mathrm{SA}}))^M \prod_i \delta[\mathcal{G}_i] e^{\beta\mu\ell} \,. \tag{4.32}$$

In light of Eq. (4.28), $\det T(\Gamma)$ in Eq. (4.32) depends solely on the total length of the chain, making it possible to rewrite Eq. (4.32) as

$$\mathcal{Z}''_{(1)} = \sum_{\mathrm{i_i},\mathrm{i_f}} \sum_{\ell} n(\ell, \mathrm{i_i}, \mathrm{i_f}) \left(\det T(\Gamma)\right)^M e^{\beta\mu\ell} \,, \tag{4.33}$$

where $n(\ell, \mathrm{i_i}, \mathrm{i_f})$ is the number of binary field configurations describing a chain of length ℓ having one end-point at the site $\mathrm{i_i}$ and the other at $\mathrm{i_f}$.

Using Eq. (4.28) we obtain:

$$\begin{aligned} \mathcal{Z}''_{(1)} &= \sum_{\mathbf{i_i},\mathbf{i_f}} \sum_{\ell} n(\ell, \mathbf{i_i}, \mathrm{i_f})\, e^{M\left[2(N-\ell)\ln 2 + N \ln \bar{g}^2 + \ln(\ell+1)\right] + \beta\mu\ell} \\ &= \sum_{\ell} n(\ell)\, e^{M\left[2(N-\ell)\ln 2 + N \ln \bar{g}^2 + \ln(\ell+1)\right] + \beta\mu\ell} \,, \end{aligned} \tag{4.34}$$

where $n(\ell) = \sum_{\mathbf{i_i},\mathrm{i_f}} n(\ell, \mathbf{i_i}, \mathrm{i_f})$ is the total number of chains having length ℓ, irrespectively of the position of the chain's end-points. To identify the most probable value of the chain length ℓ_0 we consider the saddle-point solution of Eq. (4.34), i.e.:

$$\partial_\ell \ln n(\ell)|_{\ell_0} - M\left(2\ln 2 - \frac{1}{\ell_0 + 1}\right) + \beta\mu = 0 \,. \tag{4.35}$$

Our goal is to solve this equation for μ, for an arbitrarily chosen value of ℓ_0. Unfortunately, computing $\partial_\ell \log n(\ell)$ involves the challenge of estimating the number $n(\ell)$ of independent lattice field states describing polymer configurations of length ℓ. Fortunately, this problem can be overcome by considering the regime in which the chemical potential μ and the number of degenerate fermion fields M are simultaneously large.

With this choice, Eq. (4.35) simplifies to

$$\lim_{M,\mu\to\infty} \frac{\mu}{M} \simeq \frac{1}{\beta}\left(2\ln 2 - \frac{1}{\ell_0 + 1}\right) . \tag{4.36}$$

For any (large) choice of M, Eq. (4.36) allows us to set a value of the chemical potential μ such that the most frequently sampled chains have length ℓ_0.

In the large M limit, the distribution of ℓ is exponentially peaked around the mean-value ℓ_0. To see this, we first re-write the partition function of Eq. (4.34) as

$$\mathcal{Z}''_{(1)} = \sum_{\ell} e^{-f(\ell)}\,, \tag{4.37}$$

with $f(\ell) = -\ln n(\ell) - M\big[2(N-\ell)\ln 2 + N\ln \bar{g}^2 + \ln(\ell+1)\big] - \beta\mu\ell$. Expanding $f(\ell)$ to second order around the saddle-point ℓ_0, where $\partial_\ell f(\ell)|_{\ell_0} = 0$, we obtain

$$\begin{aligned} f(\ell) \simeq & f(\ell_0) + \frac{1}{2}(\ell-\ell_0)^2\bigg(-\partial_\ell^2 \ln n(\ell)\bigg|_{\ell_0} + \frac{M}{(\ell_0+1)^2}\bigg) \\ \underset{M,\mu\to\infty}{\simeq} & \, f(\ell_0) + \frac{M}{2}\frac{(\ell-\ell_0)^2}{(\ell_0+1)^2}\,. \end{aligned}$$

Hence, chain length fluctuations around ℓ_0 are exponentially suppressed in the large M limit.

4.4 Suppressing self-intersections

In this section, we analyze the problem of fixing the order of magnitude of the parameter e_1 appearing in Eq. (4.6) in a way that self-crossing configurations are reasonably suppressed. For the sake of clarity, we divide the discussion into two parts. In Section 4.4.1 we derive an analytical expression of $\det T(\Gamma)$ when Γ describes a generic self-intersecting chain. In Section 4.4.2 we use this result to derive a suitable order of magnitude for e_1.

4.4.1 Determinant of self-crossing configurations

In this section, we derive the analytical expression of $\det T(\Gamma)$ when the state $|\Gamma\rangle$ describes a self-intersecting chain. Assuming that the self-crossing configuration is obtained by joining two chains, with length ℓ_1 and ℓ_2, and a ℓ_R-long ring, as depicted in Fig. 4.8, the resulting $T(\Gamma)$ is

$$T(\Gamma)=\left(\begin{array}{c|c} C(\ell_1,\ell_2,\ell_R) & 0 \\ \hline 0 & D_{(N-\ell_1-\ell_2-\ell_R)}\end{array}\right), \tag{4.38}$$

where $D_{(N-\ell_1-\ell_2-\ell_R)}$ is a square matrix with all diagonal entries set to $4\bar{g}^2$ and $C(\ell_1,\ell_2,\ell_R)$ has the form

$$C(\ell_1,\ell_2,\ell_R)=\bar{g}^2\begin{array}{cc}
\left(\begin{array}{ccc|ccccc|ccccc|ccccc}
3&\cdot&\cdot&-1&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot\\
\cdot&0&\cdot&\cdot&\cdot&\cdot&\cdot&-1&\cdot&\cdot&\cdot&\cdot&-1&-1&\cdot&\cdot&\cdot&-1\\
\cdot&\cdot&3&\cdot&\cdot&\cdot&\cdot&\cdot&-1&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot\\
\hline
-1&\cdot&\cdot&&&&&&&&&&&&&&&\\
\cdot&\cdot&\cdot&\multicolumn{5}{c|}{\boldsymbol{X_{(\ell_1-1)}}}&\multicolumn{5}{c|}{\mathbf{0}}&\multicolumn{5}{c}{\mathbf{0}}\\
\cdot&\cdot&\cdot&&&&&&&&&&&&&&&\\
\cdot&-1&\cdot&&&&&&&&&&&&&&&\\
\hline
\cdot&\cdot&-1&&&&&&&&&&&&&&&\\
\cdot&\cdot&\cdot&\multicolumn{5}{c|}{\mathbf{0}}&\multicolumn{5}{c|}{\boldsymbol{X_{(\ell_2-1)}}}&\multicolumn{5}{c}{\mathbf{0}}\\
\cdot&\cdot&\cdot&&&&&&&&&&&&&&&\\
\cdot&-1&\cdot&&&&&&&&&&&&&&&\\
\hline
\cdot&-1&\cdot&&&&&&&&&&&&&&&\\
\cdot&\cdot&\cdot&\multicolumn{5}{c|}{\mathbf{0}}&\multicolumn{5}{c|}{\mathbf{0}}&\multicolumn{5}{c}{\boldsymbol{X_{(\ell_R-1)}}}\\
\cdot&\cdot&\cdot&&&&&&&&&&&&&&&\\
\cdot&-1&\cdot&&&&&&&&&&&&&&&
\end{array}\right)
&
\begin{array}{c}
\spadesuit\\ \bigstar\\ \clubsuit\\ \mathbf{1}\\ \vdots\\ \vdots\\ \boldsymbol{\ell_1-1}\\ \mathbf{1}\\ \vdots\\ \vdots\\ \boldsymbol{\ell_2-1}\\ \mathbf{1}\\ \vdots\\ \vdots\\ \boldsymbol{\ell_R-1}
\end{array}
\\
\begin{array}{cccccccccccc}
\spadesuit&\bigstar&\clubsuit&\mathbf{1}&\cdots&\boldsymbol{\ell_1-1}&\mathbf{1}&\cdots&\boldsymbol{\ell_2-1}&\mathbf{1}&\cdots&\boldsymbol{\ell_R-1}
\end{array}
&
\end{array}. \tag{4.39}$$

The matrices $X_{(\ell_1-1)}$, $X_{(\ell_2-1)}$, and $X_{(\ell_R-1)}$ appearing in Eq. (4.39) are Toepliz matrices defined through Eq. (4.19). The 1-s appearing in the first rows and columns realize the connections between the points ♠, ★, and ♣ and the rest of the configuration (see Fig. 4.8).

Straightforwardly, we have

$$\det T(\Gamma)=(4\bar{g}^2)^{N-\ell_1-\ell_2-\ell_R}\,\det C(\ell_1,\ell_2,\ell_R)\,. \tag{4.40}$$

The first step towards evaluating $\det C(\ell_1,\ell_2,\ell_R)$ consist in applying Eq. (4.17) to integrate out the block matrix $X_{(\ell_R-1)}$. With this,

$$\begin{aligned}\det T(\Gamma)&=(4\bar{g}^2)^{N-\ell_1-\ell_2-\ell_R}\,\det X_{(\ell_R-1)}\,\det C(\ell_1,\ell_2)\\ &=(4\bar{g}^2)^{N-\ell_1-\ell_2-\ell_R}\,\left(\bar{g}^2\right)^{\ell_R-1}\,\ell_R\det C(\ell_1,\ell_2)\,,\end{aligned} \tag{4.41}$$

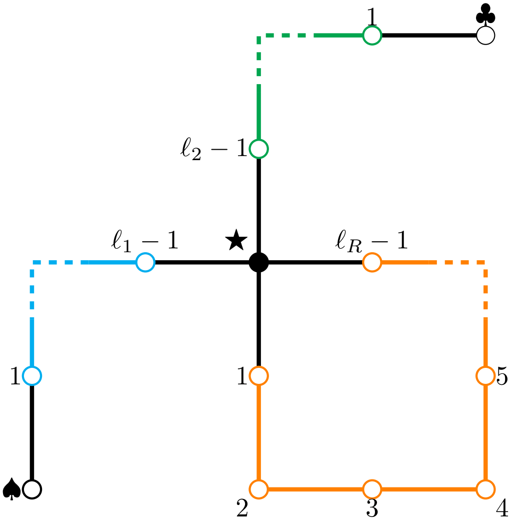


Figure 4.8: A configuration having a self-intersection in ★ can be broken into three pieces: two chains of length ℓ_1 and ℓ_2, with end-points, respectively, in ♠ and ♣; and a ring involving ℓ_R edges. The colors and symbols are used to clarify the origin of the terms in Eq. (4.39).

where

$$C(\ell_1,\ell_2)=\bar{g}^2\begin{array}{cl}\left(\begin{array}{c|c|c}
\boxed{\mathbf{S_1}} & \begin{array}{cccc} -1&\cdot&\cdot&\cdot\\ \cdot&\cdot&\cdot&-1\\ \cdot&\cdot&\cdot&\cdot\end{array} & \begin{array}{cccc}\cdot&\cdot&\cdot&\cdot\\ \cdot&\cdot&\cdot&-1\\ -1&\cdot&\cdot&\cdot\end{array} \\ \hline
\begin{array}{ccc}-1&\cdot&\cdot\\ \cdot&\cdot&\cdot\\ \cdot&\cdot&\cdot\\ \cdot&-1&\cdot\end{array} & \boxed{\boldsymbol{X}_{(\ell_1-1)}} & \mathbf{0} \\ \hline
\begin{array}{ccc}\cdot&\cdot&-1\\ \cdot&\cdot&\cdot\\ \cdot&\cdot&\cdot\\ \cdot&-1&\cdot\end{array} & \mathbf{0} & \boxed{\boldsymbol{X}_{(\ell_2-1)}}
\end{array}\right) & \begin{array}{l}\spadesuit\\ \star\\ \clubsuit\\ \mathbf{1}\\ \vdots\\ \boldsymbol{\ell_1-1}\\ \mathbf{1}\\ \vdots\\ \boldsymbol{\ell_2-1}\end{array} \\
\begin{array}{ccc}\spadesuit\ \star\ \clubsuit & \mathbf{1}\cdots\cdots\boldsymbol{\ell_1-1} & \mathbf{1}\cdots\cdots\boldsymbol{\ell_2-1}\end{array} &
\end{array}\quad . \tag{4.42}$$

In the previous,

$$S_1 = \begin{pmatrix} 3 & \cdot & \cdot \\ \cdot & -\chi_1 & \cdot \\ \cdot & \cdot & 3 \end{pmatrix} \tag{4.43}$$

and χ_1 is the sum of four elements of $X^{-1}_{(\ell_R-1)}$, namely the enties $\left(X^{-1}_{(\ell_R-1)}\right)_{1,1}$, $\left(X^{-1}_{(\ell_R-1)}\right)_{\ell_R-1,1}$, $\left(X^{-1}_{(\ell_R-1)}\right)_{1,\ell_R-1}$, and $\left(X^{-1}_{(\ell_R-1)}\right)_{\ell_R-1,\ell_R-1}$. A direct evaluation follows from Eqs. (4.22) and (4.23), yielding $\chi_1 = 2$.

Proceeding our calculation, we iterate the application of Eq. (4.17) to integrate out the diagonal block $X_{(\ell_2-1)}$, which results in

$$\begin{aligned} \det T(\Gamma) &= (4\bar g^2)^{N-\ell_1-\ell_2-\ell_R}\,(\bar g^2)^{\ell_R-1}\,\ell_R\,\det X_{(\ell_2-1)}\,\det C(\ell_1) \\ &= (4\bar g^2)^{N-\ell_1-\ell_2-\ell_R}\,(\bar g^2)^{\ell_R+\ell_2-2}\,\ell_R\,\ell_2\,\det C(\ell_1)\,. \end{aligned} \tag{4.44}$$

In the previous $C(\ell_1)$ is

$$C(\ell_1) = \bar g^2 \begin{array}{l} \left(\begin{array}{c|c} \boxed{\mathbf{S_2}} & \begin{array}{cccc} -1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & -1 \\ \cdot & \cdot & \cdot & \cdot \end{array} \\ \hline \begin{array}{ccc} -1 & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & -1 & \cdot \end{array} & \boxed{\boldsymbol{X_{(\ell_1-1)}}} \end{array}\right) \begin{array}{c} \spadesuit \\ \bigstar \\ \clubsuit \\ \mathbf{1} \\ \vdots \\ \boldsymbol{\ell_1 - 1} \end{array} \\ \quad \begin{array}{cccc} \spadesuit & \bigstar & \clubsuit & \mathbf{1}\cdots\cdots\boldsymbol{\ell_1-1} \end{array} \end{array} \tag{4.45}$$

with

$$\begin{aligned} S_2 = S_1 - &\begin{pmatrix} \cdot & \cdot & \cdot \\ \cdot & \left(X^{-1}_{\ell_2-1}\right)_{\ell_2-1,\ell_2-1} & \left(X^{-1}_{\ell_2-1}\right)_{\ell_2-1,1} \\ \cdot & \left(X^{-1}_{\ell_2-1}\right)_{1,\ell_2-1} & \left(X^{-1}_{\ell_2-1}\right)_{1,1} \end{pmatrix} = \\ = &\begin{pmatrix} 3 & \cdot & \cdot \\ \cdot & -2 - \frac{\ell_2-1}{\ell_2} & -\frac{1}{\ell_2} \\ \cdot & -\frac{1}{\ell_2} & 3 - \frac{\ell_2-1}{\ell_2} \end{pmatrix}. \end{aligned} \tag{4.46}$$

Applying once more Eq. (4.17) to integrate out the sub-matrix $X_{(\ell_1-1)}$ gives

$$\begin{aligned}\det T(\Gamma) &= (4\bar{g}^2)^{N-\ell_1-\ell_2-\ell_R}\,(\bar{g}^2)^{\ell_R+\ell_2-2}\,\ell_R\,\ell_2\,\det X_{(\ell_1-1)}\,\det R_3\\ &= (4\bar{g}^2)^{N-\ell_1-\ell_2-\ell_R}\,(\bar{g}^2)^{\ell_R+\ell_2+\ell_1-3}\,\ell_R\,\ell_2\,\ell_1\det R_3\end{aligned} \tag{4.47}$$

with

$$\begin{aligned}R_3 &= \bar{g}^2\left[S_2 - \begin{pmatrix} \left(X^{-1}_{(\ell_1-1)}\right)_{1,1} & \left(X^{-1}_{(\ell_1-1)}\right)_{1,\ell_1-1} & \cdot \\ \left(X^{-1}_{(\ell_1-1)}\right)_{\ell_1-1,1} & \left(X^{-1}_{(\ell_1-1)}\right)_{\ell_1-1,\ell_1-1} & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix}\right] =\\ &= \bar{g}^2\begin{pmatrix} 2+\frac{1}{\ell_1} & -\frac{1}{\ell_1} & \cdot \\ -\frac{1}{\ell_1} & -4+\frac{1}{\ell_1}+\frac{1}{\ell_2} & -\frac{1}{\ell_2} \\ \cdot & -\frac{1}{\ell_2} & 2+\frac{1}{\ell_2}\end{pmatrix}.\end{aligned} \tag{4.48}$$

Then, when the total amount of active edges is ℓ, i.e. $\ell_1+\ell_2+\ell_R=\ell$, the expression for $\det T(\Gamma)$ is

$$\det T(\Gamma) = -4\left(4+\frac{1}{\ell_1}+\frac{1}{\ell_2}\right)(4)^{N-\ell}\,(\bar{g}^2)^N\,\ell_1\,\ell_2\,\ell_R\,. \tag{4.49}$$

In the next section, it will be relevant to have an estimate of the upper-bound to the absolute value of $\det T(\Gamma)$ based on the total number of active bonds. Once ℓ is fixed, $|\mathrm{det}T(\Gamma)|$ attains its maximum when $\ell_1=\ell_2=\frac{1}{6}\left(\ell-1+\sqrt{\ell^2+\ell+1}\right)$ and its value scales as

$$\left|\mathrm{det}T(\Gamma)\right| = 4^{N-\ell}\,\bar{g}^{2N}\left(\frac{16}{27}\,\ell^3+O(\ell^2)\right). \tag{4.50}$$

4.4.2 Setting the order of magnitude of η

Accounting for self-avoiding configurations and those having a single point of self-intersection as the one in Fig. 4.8, the partition function in Eq. (4.10) can be approximated with

$$\mathcal{Z}'' = \sum_{\ell}\left(n_0(\ell)\det T(\Gamma_0)^M\,e^{\beta\mu\ell} + n_1(\ell)\det T(\Gamma_1)^M\,e^{\beta\mu\ell}\,e^{-8\beta e_1}\right), \tag{4.51}$$

where $n_0(\ell)$ denotes the number of self-avoiding configurations (Γ_0) having length ℓ and $n_1(\ell)$ represents the number of configurations (Γ_1) with a single self-intersection and a total length fixed at ℓ. The term $e^{-8\beta\nu}$ comes from evaluating $e^{-\beta \hat{H}_{\mathrm{SA}}}$ (see Eq. (4.6)) on states describing self-intersecting configurations as the one in Fig. 4.8. It is worth noticing that the contribution to Eq. (4.51) of configurations displaying n self-intersections is expected to be negligible, as the canonical weight associated with such configurations is exponentially supprssed by the factor $e^{-8\beta n e_1}$. For this reason, their contribution can be neglected in first approximation. To suppress the sampling of self-crossing configurations, it is sufficient to choose e_1 so that

$$n_0(\ell)\,\det T(\Gamma_0)^M \gg n_1(\ell)\,\det T(\Gamma_1)^M\, e^{-8\beta e_1}\,. \tag{4.52}$$

Recalling that $\det T(\Gamma_0)$ is given by Eq. (4.28) and that the upper-bound of $|\det T(\Gamma)|$ scales as in Eq. (4.50), the condition for suppressing at first order the self-crossing configurations is satisfied if

$$n_0(\ell)\,(\ell+1)^M \gg n_1(\ell) \left(-\frac{16}{27}\,\ell^3 + O(\ell^2)\right)^M e^{-8\beta e_1}\,. \tag{4.53}$$

Further assuming that M is even and taking the limit of large M and e_1 values, we can neglect the terms $n_0(\ell)$ and $n_1(\ell)$. Moreover, since our focus is not on fine-tuning e_1 but rather obtaining its order of magnitude, we can also discard sub-leading terms in ℓ, which results in $e_1 \gg \frac{M}{4\beta}\ln \ell$. Since ℓ is bound to be $O(N)$, we conclude that a safe choice relies on setting e_1 to be a multiple of $\frac{M}{4\beta}\ln N$.

4.5 Heteropolymers

Up to here, we have shown how the *backbone structure* of polymer chains with the desired length can be expressed through a LGT. We now further generalize our theory to include soft interactions that depend on the chemical properties of the residues. To account for chemical variability, we introduce additional qubits at each lattice site — one for each type of residue — and a Hamiltonian that encodes non-covalent interactions:

$$\hat{H}_{\mathrm{chem}} = \sum_{\mathrm{ij}} \sum_{a,b=1}^{D} \varepsilon_{ab}\,\Delta_{\mathrm{ij}}\,\hat{\chi}_{\mathrm{i}}\,\hat{\chi}_{\mathrm{j}}\,(1-\hat{B}_{\mathrm{i>j}})\,\hat{n}_{\mathrm{i}}^{a}\,\hat{n}_{\mathrm{j}}^{b}\,, \tag{4.54}$$

where D is the size of the chemical alphabet, $\hat{n}_{\mathrm{i}}^{a} = |1_a\rangle\langle 1_a|$ is the lattice site occupation operator acting on the qubit of type a located at site i, and $\hat{\chi}_{\mathrm{i}} \equiv \frac{1}{2}(\hat{\rho}_{\mathrm{i}} + \hat{g}_{\mathrm{i}})$ is an operator that identifies the sites belonging to the chain. With the self-avoiding constraint, we have $\rho_{\mathrm{i}} = 0$ (site i not occupied by the polymer), $\rho_{\mathrm{i}} = 1$ (site occupied by the endpoint), and $\rho_{\mathrm{i}} = 2$ (site along the backbone), so $\hat{\chi}_{\mathrm{i}} = \hat{\mathbb{1}}$ at the endpoints and along the backbone. The factor $\hat{\chi}_{\mathrm{i}}\hat{\chi}_{\mathrm{j}}$ in Eq. (4.54) ensures that the non-covalent interaction Hamiltonian only couples qubits located at lattice sites that are occupied by the chain. Furthermore, the term $(1 - \hat{B}_{\mathrm{ij}})$ excludes non-covalent interactions between covalently bound residues. Δ_{ij} is a decaying function of the distance $|\mathrm{i} - \mathrm{j}|$. In a nearest-neighbor model, $\Delta_{\mathrm{ij}} = 1$ if $|\mathrm{i} - \mathrm{j}| = 1$ and 0 otherwise. ε_{ab} is a $D \times D$ symmetric matrix that defines the interaction between the residue types.

Finally, to set the relative abundance of the different residue types along the chain, we introduce the term

$$\hat{H}_{\mathrm{rel}} = g' \sum_{a=1}^{D} \sum_{\mathrm{i}} \hat{\chi}_{\mathrm{i}} \left(\hat{n}_{\mathrm{i}}^{a} - \nu_a \right)^2 . \tag{4.55}$$

This term can also be written as $\hat{H}_{\mathrm{rel}} = -\sum_a \mu_a \hat{N}^a + g' \sum_a \nu_a^2 \sum_{\mathrm{i}} \hat{\chi}_{\mathrm{i}}$, where $\hat{N}^a = \sum_{\mathrm{i}} \hat{\chi}_{\mathrm{i}}\, \hat{n}_{\mathrm{i}}^{a}$ counts the number of residues of type a in the polymer backbone and $\mu_a \equiv g'(2\nu_a - 1)$ are the associated chemical potentials. Since for a single chain $\sum_i \hat{\chi}_{\mathrm{i}} = \hat{L} + 1$, the second term contributing to $\hat{H}_{\mathrm{rel}}$ can be absorbed into the chemical potential term $\sim \mu$ acting on the bond qubits. After combining all terms, we obtain the final expression for the partition function of a single heteropolymer:

$$\mathcal{Z}_{\mathrm{het}} = \sum_{\Gamma,\eta} \prod_{\mathrm{i}} \delta[\mathcal{G}_{\mathrm{i}}(\Gamma)]\delta[\Phi_{\mathrm{i}}(\eta)]\, e^{\beta(\mathcal{L}[\eta,\Gamma] - \mathcal{H}[\eta,\Gamma])}. \tag{4.56}$$

The summation $\sum_\eta$ spans the configurations of the D fields η_{i}^{a}, constructed from the eigenvalues of $\hat{n}_{\mathrm{i}}^{a}$, while

$$\mathcal{L}[\Gamma, \eta] \equiv \mu L + \sum_a \mu_a N_a + M \mathrm{Tr} \log T \tag{4.57}$$

$$\mathcal{H}[\Gamma, \eta] \equiv H_{\mathrm{SA}} + H_{\mathrm{chem}}. \tag{4.58}$$

H_{chem} is obtained by replacing $\hat{\Gamma}_{\mathrm{ij}}$ and $\hat{n}_{\mathrm{i}}^{a}$ with the binary fields Γ_{ij} and η_{i}^{a} in Eq. (4.54). Setting $\Phi_{\mathrm{i}} = (\sum_{a=1}^{D} \eta_{\mathrm{i}}^{a} - 1)^2$, so $\prod_{\mathrm{i}} \delta[\Phi_{\mathrm{i}}]$ ensures that at any lattice site i there is exactly one active residue-type state (see bottom-right panel in Fig. 4.1).

Importantly, the theory defined by Eq. (4.56) does not allow us, in general, to choose a specific primary sequence. Nevertheless, it can be used to characterize the thermodynamically stable sequences along with their structures. Specific patterns in the sequence may be imposed by additional interactions. For example, co-polymers can be sampled by adding an "anti-ferromagnetic" interaction, $\hat{H}_{\text{co-pol}} = \sum_{a,b} \sum_{\langle \mathrm{ij} \rangle} \hat{n}_{\mathrm{i}}^{a} \hat{n}_{\mathrm{j}}^{b} \hat{B}_{\mathrm{ij}}$, while block co-polymers are obtained from the analog "ferromagnetic" term.

4.6 Illustrative application

To illustrate our LGT approach to polymer sampling, we choose a minimalistic setting based on the so-called HP lattice model [140, 182]. In this model, the residues are grouped into hydrophobic (h) and polar (p), with interaction matrix elements $\varepsilon_{hh} = -1$, $\varepsilon_{pp} = \varepsilon_{hp} = 0$. For illustration purposes, we simulate a small chain, $\ell_0 = 9$, in a 10×10 square lattice with nearest-neighbor interactions. Such two-dimensional lattice models are commonly adopted in biophysical contexts, as they yield surface-to-volume ratios close to that of realistic small globular proteins. Moreover, we choose $M = 120$, allowing us to use Eq. (4.36) to fix the chemical potential μ as a function of ℓ_0, and we set $\nu_h = \nu_p = 0.5$. In the following section, we discuss separately the details concerning the Monte Carlo algorithm that is employed to simulate the HP-model considered in this section and the associated results.

4.6.1 Details on the Monte Carlo algorithm

Equation (4.56) enables us to simultaneously sample the heteropolymer's conformational and chemical space, using a Metropolis MC scheme. To this end, we first initialize the binary field Γ to fulfill all conformational constraints, i.e., to represent an arbitrary self-avoiding polymer configuration. Likewise, the initial binary fields η_a should obey the hard constraint set by the term $\prod_{\mathrm{i}} \delta[\Phi_{\mathrm{i}}]$ in Eq. (4.56). To propose conformational trial moves, we act at a random location with the deformation operator, Eq. (4.4), or with the terminal displacement operator, Eq. (4.5). New sequences are proposed by random point-wise mutations of the chemical residue at arbitrary positions in the lattice. These trial moves are accepted or rejected according to a Metropolis criterion defined by the probability density $\propto e^{\beta(\mathcal{L}-\mathcal{H})}$.

To initialize our Monte Carlo simulations, we generate the initial configuration of the binary field Γ_{ij} to (i) connect two arbitrarily chosen chain endpoints, (ii) satisfy the chain topology condition — Gauss' law in Eq. (4.2) —, and (iii) fulfill the self-avoidance condition.

To study the NP model employed in the illustrative application, we randomly initialize the field of binary variables η_{i}^{a} associated with the chemical residue species $a = h$ and $a = p$, in such a way that at each lattice point i either $\eta_{\mathrm{i}}^{h} = 1$ and $\eta_{\mathrm{i}}^{p} = 0$ or $\eta_{\mathrm{i}}^{h} = 0$ and $\eta_{\mathrm{i}}^{p} = 1$.

To generate a new element of the Markov chain, we propose "backbone deformation" moves, "end-point diffusion" moves, and "point-wise monomer mutation" moves, with equal probability.

To perform backbone deformation moves, we randomly pick a plaquette on the lattice and apply the corresponding chain deformation operator $\hat{C}_{\Box}$, defined in Eq. (4.4). The result of acting with this operator on a plaquette is to switch the value of each bond variable, as shown in Fig. 4.4. To shift the endpoints of the chain, we randomly select one of the two extremes of the chain, i, and one of the $2d$ oriented directions, $\mathbf{e}_k$, and we apply the operator $\hat{D}_{\mathrm{i}\,\mathrm{i}+\mathbf{e}_k}$ defined in Eq. (4.5), which acts on the backbone configuration as shown in Fig. 4.5.

To perform point-wise monomer mutations, we randomly select a lattice site i and a chemical flavor a, defining the operator $\hat{M}_{\mathrm{i}}(a)$, which turns the chemical element in site i into flavor a.

The Metropolis acceptance/rejection criterion for this set of moves is simply

$$\mathcal{A}\left(\Gamma, \eta \to \Gamma', \eta'\right) = \min\left(1, \frac{\mathcal{P}(\Gamma', \eta')}{\mathcal{P}(\Gamma, \eta)}\right), \tag{4.59}$$

with $\mathcal{P}(\Gamma, \eta)$ defined as

$$\mathcal{P}(\Gamma, \eta) \propto e^{\mathcal{L}[\Gamma,\eta]-\beta H_{\mathrm{chem}}[\Gamma,\eta]}, \tag{4.60}$$

and

$$\begin{aligned} \mathcal{L}[\Gamma, \eta] &= M \operatorname{Tr}\log T[\Gamma] + \mu \sum_{\mathrm{ij}} \Gamma_{\mathrm{ij}} + \sum_{a,\mathrm{i}} \mu_a\, \eta_{\mathrm{i}}^{a} \quad \text{and} \\ \mathcal{H}[\Gamma, \eta] &= H_{SA}[\Gamma] + \sum_{\mathrm{i}>\mathrm{j}} \sum_{a,b=1}^{D} \varepsilon_{ab}\, \Delta_{\mathrm{ij}}\; \chi_{\mathrm{i}}\eta_{\mathrm{i}}^{a}\, (1-\Gamma_{\mathrm{ij}})\, \chi_{\mathrm{j}}\eta_{\mathrm{j}}^{b}\,. \end{aligned} \tag{4.61}$$

Symbol	**Definition**	**Value**
M	Number of fermionic degenerate modes	120
m_{f}	Mass of fermions	1
β	Inverse temperature	$\{0.1, 1, 2\}$
e_1	Energy scale entering the repulsive term H_{SA} in Eq. (4.6)	$1,000$
–	Lattice format	10×10
D	Number of chemical elements	2
$\nu_h = \nu_p$	Abundance of h and p chemical species	0.5
$\mu_h = \mu_p$	Chemical potential associated with h and p chemical species	0
ℓ_0	Expected average chain's length (when no chemical interaction takes place)	9
N_{c}	Number of polymeric chains generated	1
N_{steps}	Number of Monte Carlo steps	10^6
N_{mc}	Number of Monte Carlo trajectories for each parameter choice	10
$\tilde{\Delta}_{0.1}$	Decorrelation time when $\beta = 0.1$	$1,800$
$\tilde{\Delta}_1$	Decorrelation time when $\beta = 1$	$1,980$
$\tilde{\Delta}_2$	Decorrelation time when $\beta = 2$	$5,040$

Table 4.1: Parameters adopted in the Monte Carlo simulations reproducing results in Figs. 4.9 to 4.11.

All the parameters appearing in these equations that we used in our numerics, along with the number and length of the Markov chains, are reported in Table 4.1. We found these to present good performance while reproducing the relevant physics.

Autocorrelation-time analysis: The conformational decorrelation time associated with a single Monte Carlo trajectory can be estimated from the autocorrelation function $A(\Delta)$, defined as

$$\begin{aligned} A(\Delta) = \frac{1}{N}\sum_{n=1}^{N}&\left(\frac{1}{2N_\mathrm{e}}\sum_{\langle \mathrm{i},\,\mathrm{j}\rangle}\Gamma_{\mathrm{ij}}^{(n)}\Gamma_{\mathrm{ij}}^{(n+\Delta)}\right)+ \\ &- \left(\frac{1}{N}\sum_{n=1}^{N}\frac{1}{2N_\mathrm{e}}\sum_{\langle \mathrm{i},\,\mathrm{j}\rangle}\Gamma_{\mathrm{ij}}^{(n)}\right)\left(\frac{1}{N}\sum_{n=1}^{N}\frac{1}{2N_\mathrm{e}}\sum_{\langle \mathrm{i},\mathrm{j}\rangle}\Gamma_{\mathrm{ij}}^{(n+\Delta)}\right)\Bigg]\,, \end{aligned} \tag{4.62}$$

where N_e is the number of edges in the lattice, $\Gamma_{\mathrm{ij}}^{(n)}$ and $\Gamma_{\mathrm{ij}}^{(n+\Delta)}$ are binary variables sampled at the n–th and $(n+\Delta)$–th element of the Markov chain, and the sum $\sum_n$ runs over the Monte Carlo steps. An analogous quantity is employed to estimate the correlation of the monomer field. To increase the statistics, we average the autocorrelation function over N_mc independent Monte Carlo trajectories. The autocorrelation analysis is reported in Fig. 4.9. For practical purposes, we consider configurational states to have reached maximum decorrelation after $\tilde{\Delta}_\beta$ Monte Carlo steps (see the vertical dotted lines in Fig. 4.9 and numerical values in Table 4.1). All results reported are derived by averaging over decorrelated configurations, sampled from N_mc independent Markov chains.

4.6.2 Results

This section presents the results generated by our Monte Carlo algorithm when applied to the illustrative example of an HP model.

The length distributions for several values of the inverse temperature β and parameters detailed in Tab. 4.1 are reported in Fig. 4.10. For these parameter values, the expected average chain length when neglecting chemical interactions is $l_0 = 9$. At low temperatures, where entropic contributions to the statistical properties of the ensemble are small, we expect the sampling of low-energy configurations to be enhanced. As non-covalent interactions among monomers can lower the protein's energy, and since the possible number of such interactions increases with the chain length, we expect to sample longer polymeric chains at low temperatures. Indeed, this is what we observe in our simulations. As the temperature increases, the average chain length tends towards the value of $l_0 = 9$ favoured by the chemical potential. In all considered cases, thanks to the large value

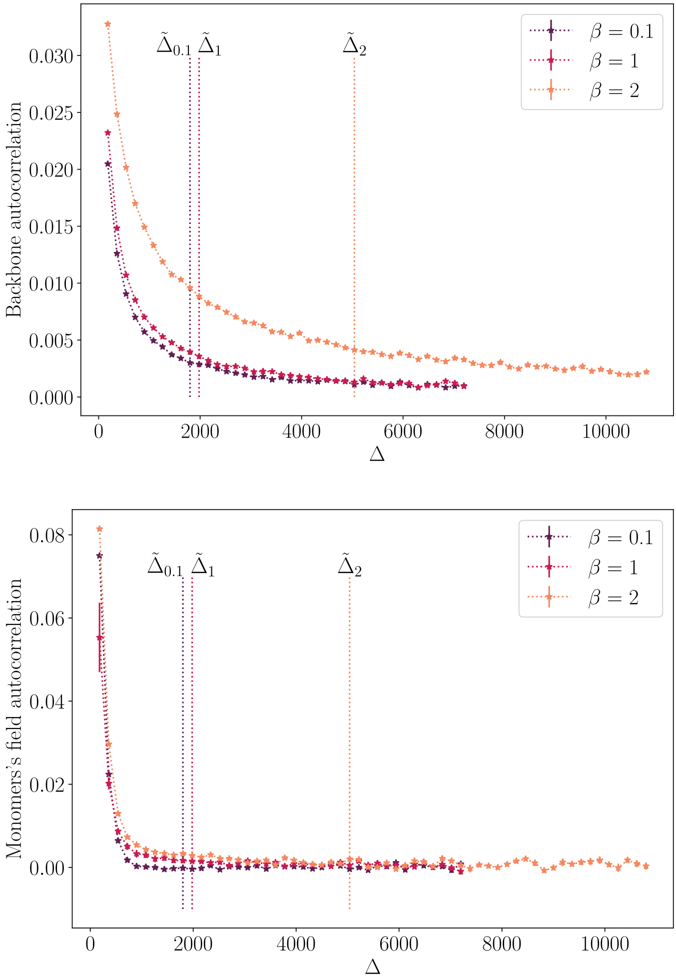


Figure 4.9: Autocorrelation analysis for simulations run at different inverse temperatures β. **(a)** Backbone auto-correlation and **(b)** monomer field autocorrelation as a function of Markov chain elements Δ. The vertical dotted lines show the choice of $\tilde{\Delta}_\beta$ for $\beta \in \{0.1, 1, 2\}$, reported in Tab. 4.1 and is used to estimate the distributions of $|\mathbf{i}_\mathrm{i} - \mathbf{i}_\mathrm{f}|$, of the chain's length ℓ, and of the abundance of H and P residues.

of $M = 120$ considered, the chain-length distribution is rather strongly peaked around the average. In the upper-left panel of Fig. 4.11, we report the end-point distance distributions (in lattice spacing units), which shows that the chain undergoes the hydrophobic collapse in the low-temperature regime. Next, we examine how soft interactions promote specific primary sequences. In the upper-right panel of Fig. 4.11, we report the measured relative occurrence of h residues at different temperatures. The latter were arbitrarily set to explore different regimes. As can be expected, even when $\mu_h = \mu_p$, the h-type residues are more frequent at low tempera-

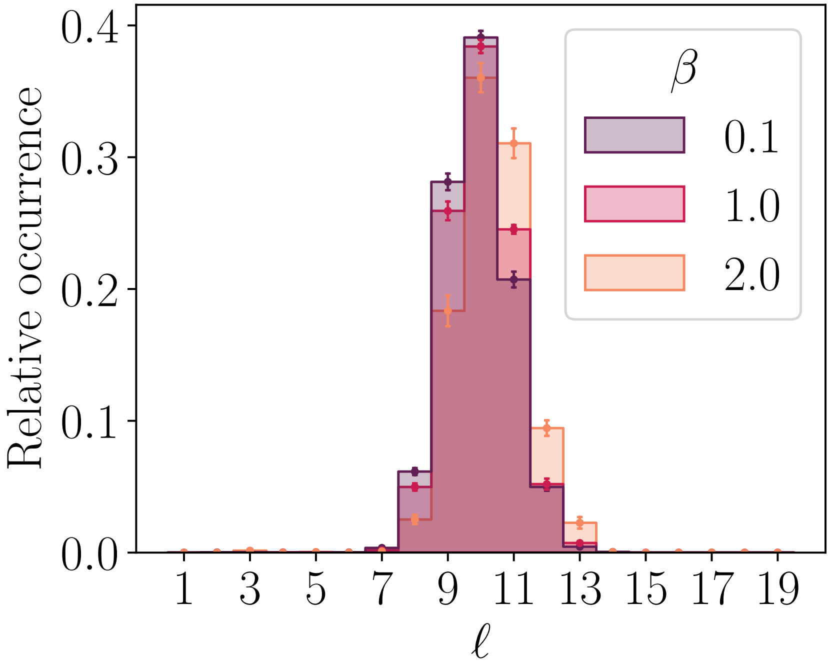


Figure 4.10: Chain-length distribution for various values of the inverse temperature ($\beta \in \{0.1, 1, 2\}$). Samples are drawn from ensembles where the chain's end-points are delocalized on the lattice. At low temperatures, where entropic effects are small and chemical non-covalent interaction plays a dominant role, the system prefers longer chains. As temperature increases, the chain-length distribution becomes peaked more closely to the value favoured by the chemical potential of $l_0 = 9$.

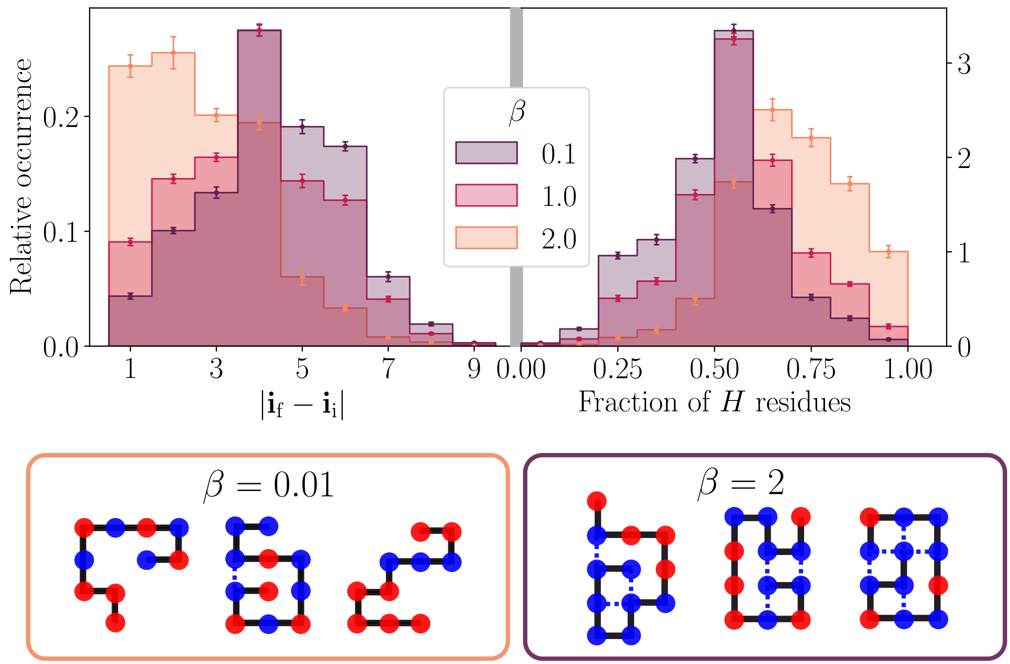


Figure 4.11: Predictions of the HP model obtained by MC simulation of the LGT, at different inverse temperatures. Upper left: distribution of end-point distance in units of lattice spacing. Upper right: relative occurrence of h-type residues. Bottom: samples of representative sequences and structures for two different β. Dotted lines highlight the network of hydrophobic interactions.

tures. Finally, to highlight sequence-structure correlations, in the bottom panel we report a sample of typical chains generated by our MC algorithm. In the low-temperature regime, compact structures are stabilized by the formation of a hydrophobic core, an effect well-known in proteins. To quantify this effect, we compare the average number of $h-h$ and $p-p$ non-covalent interactions at $\beta = 2$, finding respectively $N_{\mathrm{hh}} = 3.20 \pm 0.03$ and $N_{\mathrm{pp}} = 0.15 \pm 0.01$.

4.6.2.1 Computational efficiency

In this section, we showcase an example that illustrates why the LGT approach is particularly suited to sample maximally compact polymer configurations. Let us consider the problem of generating new compact self-avoiding structures by acting with chain deformation operators $\hat{C}_p$ on the specific Hamiltonian path state shown at the center of Fig. 4.12. This action yields a new viable compact structure when it acts on 4 out of the 9 plaquettes. The rejected moves are those for which $\hat{C}_p$ acts on a plaquette that contains a path turn, i.e., those highlighted in red in Fig. 4.12. Hence, when such a spiral-shaped Hamiltonian path is embedded in a larger two-dimensional lattice of size N, the chain deformation moves have a larger success rate, $\frac{(N-2)^2}{(N-1)^2}$, thus even approaching 100% efficiency in the thermodynamic limit.

The efficiency of the LGT moves is not surprising, since the action of the chain-deformation operator generates chain rearrangements that resemble those performed in graph-theory-based algorithms for sampling Hamiltonian paths [171,172] and state-of-the-art MC schemes for sampling maximally compact structures [173, 174]. However, our MC algorithm based on LGT is defined to tolerate small fluctuations in the chain length. Hence, it generates compact structures that are not always Hamiltonian paths. To showcase the computational efficiency of our LGT approach we use our scheme to generate nearly maximally compact configurations ($>98\%$ site occupancy) of a homopolymer chain in a poor solvent. This regime is reached by initializing the monomer's field to $\eta_{\mathrm{i}}^{h} = 1$ and $\eta_{\mathrm{i}}^{p} = 0$ everywhere. Then, the MC trial moves are chosen, with equal probability, to be either chain deformations or end-point displacements. Further, to simulate homopolymers on a lattice comprising N vertices, we set $\ell_0 = N-1$. The results obtained by setting $\beta = 1$ and $M = 200$, see Fig. 4.13, show that the MC decorrelation time τ grows only linearly with the median chain length ℓ.

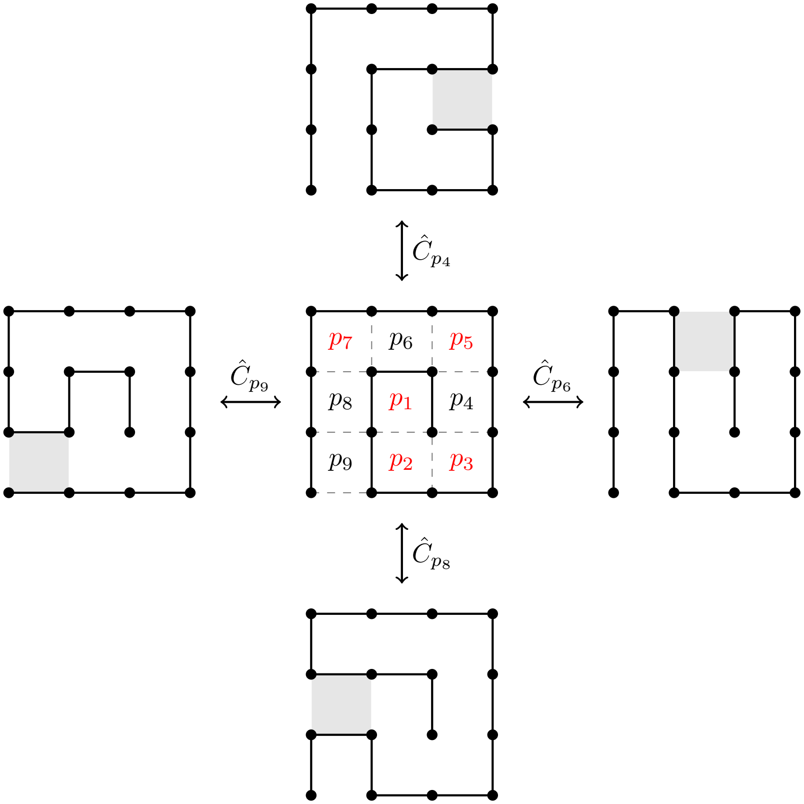


Figure 4.12: By acting with a single deformation operator (corresponding to plaquettes highlighted in gray), it is possible to derive 4 other Hamiltonian paths from the spiral-shaped configuration depicted in the center.

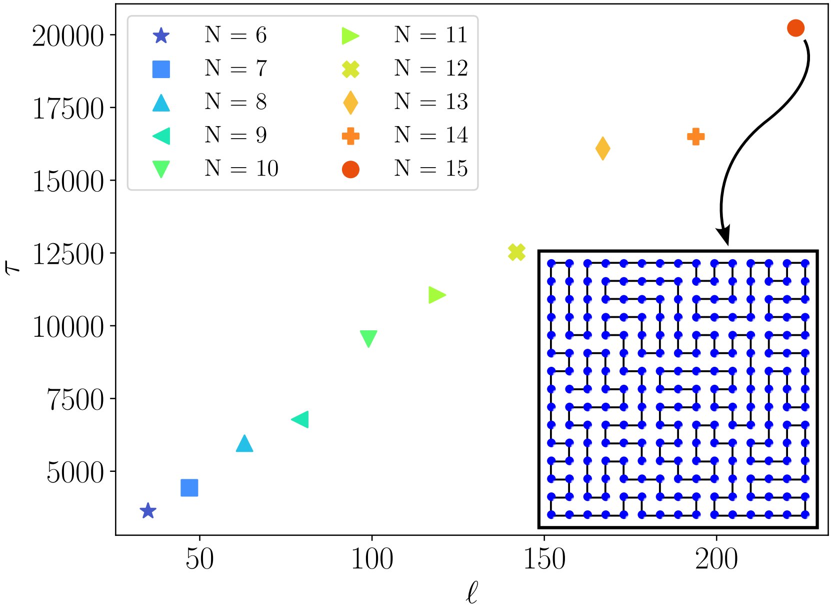


Figure 4.13: Number of MC steps needed to generate decorrelated compact structures ($> 98\%$ site occupancy) as a function of the median length ℓ (decorrelation time τ), for different two-dimensional lattice sides. τ was estimated as twice the exponent that leads the decay of the autocorrelation function defined in Eq. (4.62) of the SM. On the bottom-right corner we show a compact configuration obtained for the largest lattice.

4.7 Discussion and outlook

Starting from the pioneering work of Edwards [142] and DeGennes [143], field theory methods have been extensively employed in polymer science [183, 184], with applications ranging from material science [185] to biophysics [75, 186]. A key limitation of conventional polymer field theory is that it is very challenging to perform calculations beyond the mean-field approximation [183] because a sign problem prevents the application of stochastic methods. While successful heuristic applications of the complex Langevin equation have been reported [187, 188], a general understanding of the conditions under which these schemes converge is still lacking. In contrast, for even values of M our LGT does not suffer from a sign problem, and can thus be solved by MC. The $\mathbb{Z}_2$ gauge symmetry dictates conformational moves that automatically preserve the chain continuity, while self-avoidance and soft constraints can be implemented via the Metropolis scheme. Moreover, since the present LGT is natively formulated in terms of qubits, it sets the stage for developing quantum algorithms for polymer sampling that can capitalize on the ongoing fast development of quantum hardware.

The development of LGTs has been mainly driven by the quest for a non-perturbative solution of quantum chromo-dynamics (QCD). The present formulation of polymer physics may enable similar theoretical and computational field-theoretic methods to be applied to gain insight into the physical origin of the funneled structure of the protein energy landscape. To implement this program, we must recognize an additional symmetry acting on the residues' chemical types, which can be spontaneously broken. Remarkably, previous theoretical [189] and experimental [190] studies have shown that the key properties of proteins' thermodynamics are preserved by interchanging specific types of amino acids or grouping them into a few effective families. On the other hand, each protein is defined by a specific primary sequence, thus providing a different realization.

Important directions to pursue in the future include exploring the implications of this connection, along with assessing the computational efficiency of the LGT MC approach relative to conventional real-space MC, and designing quantum algorithms to simulate our LGT [191–193], capitalizing on recent rapid progress in the field.

PART III

ENTANGLEMENT WITNESSES

Chapter 5

Optimization of EWs in LGTs

Entanglement is a feature that only quantum systems exhibit [194] and constitutes the foundation of many real-world applications of quantum computing [58, 59, 62, 195, 196]. Quantifying the strength of quantum correlations within an entangled system typically requires prior knowledge of the full wave function or, at the very least, the portion of the wave function localized on a specific subsystem of interest. Unfortunately, this piece of knowledge demands in general exponentially many measurement operations [66, 197], even if the sought of more practical alternatives is an active field of research [67, 198].

Nonetheless, the mere presence of entanglement can be witnessed by specific observables, known as entanglement witnesses (EWs) [68, 70, 71, 199], which are prone to be directly incorporated into experimental protocols [200, 201]. The two core properties of such operators are: (i) they acquire non-negative expectation values when measured against *separable* states and (ii) their measurement potentially leads to negative outcomes. As a direct consequence, any state featuring a negative expectation value of an EW — and one such a state certainly exists — is reported to be entangled (i.e., non-separable).

As mentioned, entanglement characterizes all quantum systems, spanning from many-body systems [202] to single-particle systems [203], and its theoretical and practical implications have been widely studied. For instance, detecting and characterizing entanglement in quantum many-body systems is essential for their theoretical understanding [204–209] and the same holds for high-energy systems as well [210, 211].

In recent years, an increasing interest in embedding lattice gauge theoretical frameworks on quantum devices was prompted not only by the possibility of simulating high-energy physic and condensed-matter systems on quantum platforms [212–226], but also because LGTs offer a natural setting for the implementation of quantum error correction and mitigation techniques [227–233].

However, the study of LGTs from the perspective of quantum information theory offers additional challenges, especially regarding entanglement detection and its measures. This is because the presence of local gauge symmetries constrains physical states and observables, with direct implications to the underlying notion of separability [234–236], which inevitably impacts the definition of entanglement measures [237–244]. More generally, also when considering systems exhibiting global symmetries (e.g., fermionic systems), it is known that the implicit selection and superselection rules must be taken into consideration when studying entanglement [245–250].

In this chapter, we explore theoretical aspects concerning the definition of separable states in LGTs. In particular, in Section 5.1, we prove that at least for Abelian gauge symmetries, separability of physical states can be determined by testing the factorization of expectation values of operators such as $\hat{T}_{\mathrm{A}} \otimes \hat{T}_{\mathrm{B}}$, where $\hat{T}_R$ are gauge-invariant operators completely supported on region $R = \mathrm{A}, \mathrm{B}$. As a consequence, it is guaranteed that any entangled state arising in a LGT is detected by at least one gauge-invariant EW.

In Section 5.3 we generalize the optimization procedure of EWs [73] to the special case of gauge-invariant ones and in Section 5.4 we test the resulting protocol on a pure $U(1)$ LGT.

5.1 Theoretical formalism

This section introduces the notation and theoretical concerns related to separability and entanglement witnessing in the context of Abelian LGTs.

5.1.1 Terminology and notation

We start by considering a generic Abelian group $\mathcal{G}$ acting on degrees of freedom attached to the edges of a graph $G = (E, V)$, which consists of a vertex set, V, and an oriented edge set, E. Further, we associate with every edge, e, a local Hilbert space denoted by $\mathcal{H}_e$, so that the total Hilbert space is the tensor product $\mathcal{H}_{\text{tot}} = \bigotimes_{e \in E} \mathcal{H}_e$. The representation of the gauge transformation mediated by the generators $g \in \mathcal{G}$[1]acting on vertex x is

$$\hat{U}_x^g = \left(\bigotimes_{e=(x,w)\in E} \hat{g}_e \right) \otimes \left(\bigotimes_{e'=(w,x)\in E} \hat{g}_{e'}^{-1} \right), \tag{5.1}$$

where $\hat{g}_e$ is the representation of g on $\mathcal{H}_e$. Since our group is Abelian, it is possible to simultaneously diagonalize all generators $\hat{U}_x^g$, enabling the decomposition of the total Hilbert space as

$$\mathcal{H}_{\text{tot}} = \bigoplus_{\mathbf{Z}} \mathcal{H}(\mathbf{Z}) \quad \text{with} \quad |\psi\rangle \in \mathcal{H}(\mathbf{Z}) \Leftrightarrow \hat{U}_x^g \, |\psi\rangle = Z_x^g \, |\psi\rangle \ , \tag{5.2}$$

where $\mathcal{H}(\mathbf{Z})$ are eigenspaces identified with eigenvalues $\mathbf{Z}$ with respect to the set of generators. Indeed, to prove that generic states $|\psi_1\rangle \in \mathcal{H}(\mathbf{Z}_1)$ and $|\psi_2\rangle \in \mathcal{H}(\mathbf{Z}_2)$ are orthogonal when $\mathbf{Z}_1 \neq \mathbf{Z}_2$, it is sufficient to consider the quantity $\langle\psi_1|\hat{U}_x^{(g)}|\psi_2\rangle$, which can be equivalently expressed as $Z_{1,x}^{(g)} \langle\psi_1|\psi_2\rangle$ and as $Z_{2,x}^{(g)} \langle\psi_1|\psi_2\rangle$. Hence, since $\mathbf{Z}_1 \neq \mathbf{Z}_2$, there will be at least a group element g and a lattice point x where $Z_{1,x}^{(g)} \neq Z_{2,x}^{(g)}$, so that previous considerations directly imply $\langle\psi_1|\psi_2\rangle = 0$, confirming that $\mathcal{H}(\mathbf{Z}_1)$ and $\mathcal{H}(\mathbf{Z}_2)$ are orthogonal spaces. Due to local constraints characterizing the LGT, physical operators, $\hat{O}$, are not allowed to connect different sectors as they satisfy

$$\forall\, g \in \mathcal{G},\ \forall\, x \in V \quad , \quad \left[\hat{O}, \hat{U}_x^g\right] = \hat{\mathbb{0}} \,. \tag{5.3}$$

Let us specialize the discussion to the case where a bipartition of the initial graph G is introduced. By dividing the edges of G into two disjoint sets, E_1 and E_2, two subgraphs of G naturally arise; namely $\mathrm{A} = (E_1, V_1)$ and $\mathrm{B} = (E_2, V_2)$, where V_1 and V_2 denote the sets of vertices associated with the edges in E_1 and E_2, respectively. The local generators associated with lattice sites lying at the interface between the two subgraphs, $V_{\text{int}} =$

---

[1] In group theory, the generators of a group $\mathcal{G}$ are elements therein that can be combined to generate any transformation in $\mathcal{G}$.

$V_1 \cap V_2$, take the form

$$\hat{U}^g_x = \hat{U}^g_{\mathrm{A},x} \otimes \hat{U}^g_{\mathrm{B},x}\,, \tag{5.4}$$

with $\hat{U}^g_{R,x}$ being "broken" gauge generators entirely supported on region $R = \mathrm{A}, \mathrm{B}$. Similarly to Eq. (5.2), where full generators are used to define sectors, it is possible to use the broken gauge generators $\hat{U}^g_{R,x}$ to define *subsectors* $\mathcal{H}_\mathrm{R}(\mathbf{Z}_\mathrm{R})$ as

$$|\psi\rangle \in \mathcal{H}_\mathrm{R}(\mathbf{Z}_\mathrm{R}) \Leftrightarrow \hat{U}^g_{R,x}\,|\psi\rangle = Z^g_{R,x}\,|\psi\rangle\;, \tag{5.5}$$

so that the total Hilbert space takes the form

$$\mathcal{H}_\mathrm{tot} = \bigoplus_{\mathbf{Z}} \left(\bigoplus_{(\mathbf{Z}_\mathrm{A},\mathbf{Z}_\mathrm{B})}^{\prime} \mathcal{H}_\mathrm{A}(\mathbf{Z}_\mathrm{A}) \otimes \mathcal{H}_\mathrm{B}(\mathbf{Z}_\mathrm{B}) \right), \tag{5.6}$$

where the ($\prime$) in the summation indicates that only subsectors' labels, $\mathbf{Z}_\mathrm{A}$ and $\mathbf{Z}_\mathrm{B}$, that are compatible with the sector's ones, $\mathbf{Z}$, are to be considered. In fact, Eq. (5.4) establishes that such compatible labels are bound to satisfy

$$Z^g_x = Z^g_{A,x}\,Z^g_{B,x}, \quad \forall\, x \in V_\mathrm{int}\,,\; \forall g \in \mathcal{G}\,. \tag{5.7}$$

Gauge-invariant operators that are fully supported on region R are said to be *locally gauge-invariant* and commute with each broken gauge generator $\hat{U}^g_{R,x}$. Hence, they are not able to connect different subsectors.

5.1.2 Separability in Abelian LGTs

In a more general context than LGTs — particularly when the total Hilbert space admits a tensor product structure — the definition of separable pure states can be expressed in two equivalent ways. The first requires every pure separable state between regions A and B, $|\phi\rangle \in \mathcal{H}_\mathrm{A} \otimes \mathcal{H}_\mathrm{B}$, to admit a rewriting in the form $|\phi_\mathrm{A}\rangle \otimes |\phi_\mathrm{B}\rangle$. The second formulation requires that the expectation value of any observable taking the form of $\hat{T}_\mathrm{A} \otimes \hat{T}_\mathrm{B}$ factorize as $\langle\phi|\hat{T}_\mathrm{A} \otimes \mathbb{1}_\mathrm{B}|\phi\rangle\ \langle\phi|\mathbb{1}_\mathrm{A} \otimes \hat{T}_\mathrm{B}|\phi\rangle$.

Although these two definitions of separability are in principle equivalent, the first becomes ill-defined in LGTs, where pure states, defined within a sector, belong to a Hilbert space that is not supported by a tensor product structure, as Eq. (5.6) confirms. In this context, we are left with the second definition, which relies on the factorization of suitably chosen operators. Here, there are two viable options; the first is to consider generic operators supported on the two regions, $\hat{T}_{\mathrm{A}} \in \mathrm{B}(\mathcal{H}_{\mathrm{tot,A}})$ and $\hat{T}_{\mathrm{B}} \in \mathrm{B}(\mathcal{H}_{\mathrm{tot,B}})$ and test whether $\langle \hat{T}_{\mathrm{A}} \otimes \hat{T}_{\mathrm{B}} \rangle$ factorizes; the second possibility is to consider locally gauge-invariant operators only, i.e., test separability using operators such as $\hat{O}_{\mathrm{A}} \otimes \hat{O}_{\mathrm{B}}$, with $\hat{O}_{\mathrm{A}}$ and $\hat{O}_{\mathrm{B}}$ being, respectively, the embeddings of locally gauge-invariant operators of regions A and B in $\mathrm{B}(\mathcal{H}_{\mathrm{tot,A}})$ and $\mathrm{B}(\mathcal{H}_{\mathrm{tot,B}})$.

A legitimate concern is the relationship between the two possible definitions above. In what follows, we will distinguish between *generally-separable* and *gauge-separable* states, which refer to physical states that satisfy separability criteria involving generic operators and gauge-invariant operators, respectively.

Trivially, generally-separable states are also gauge-separable, for any gauge-invariant observable can be intended as a member of $\mathrm{B}(\mathcal{H}_{\mathrm{tot,A}}) \otimes \mathrm{B}(\mathcal{H}_{\mathrm{tot,B}})$.

It remains for us to explore the possibility that gauge-separable states might also be generically separable. To investigate this, we start by considering the specific case in which the gauge-separable state, $\hat{\rho}$, lies entirely in a subsector labeled by $(\mathbf{Z}_{\mathrm{A}}, \mathbf{Z}_{\mathrm{B}})$. In what follows, we prove that $\hat{\rho}$ satisfies separability criteria based on the factorization of any operator in the form $\hat{T}_{\mathrm{A}} \otimes \hat{T}_{\mathrm{B}}$, with $\hat{T}_{R} \in \mathrm{B}(\mathcal{H}_{\mathrm{tot,R}})$ (i.e., $\hat{\rho}$ satisfies the general-separability criteria).

By construction, see Eq. (5.6), the projector on $\mathcal{H}(\mathbf{Z}_{\mathrm{A}}) \otimes \mathcal{H}(\mathbf{Z}_{\mathrm{B}})$ has the tensor product structure $\hat{\pi}(\mathbf{Z}_{A}) \otimes \hat{\pi}(\mathbf{Z}_{B})$ and as $\hat{\rho}$ belongs to $\mathcal{H}_{\mathrm{A}}(\mathbf{Z}_{\mathrm{A}}) \otimes \mathcal{H}_{\mathrm{B}}(\mathbf{Z}_{\mathrm{B}})$, it is invariant under left and right multiplication by such projector. With these premises, $\langle \hat{T}_{\mathrm{A}} \otimes \hat{T}_{\mathrm{B}} \rangle_{\rho}$ writes

$$\begin{aligned} \langle \hat{T}_{\mathrm{A}} \otimes \hat{T}_{\mathrm{B}} \rangle_{\rho} &= \mathrm{Tr}\left[\hat{T}_{\mathrm{A}} \otimes \hat{T}_{\mathrm{B}} \left(\hat{\pi}(\mathbf{Z}_{A}) \otimes \hat{\pi}(\mathbf{Z}_{B})\, \hat{\rho}\, \hat{\pi}(\mathbf{Z}_{A}) \otimes \hat{\pi}(\mathbf{Z}_{B}))\right)\right] \\ &= \mathrm{Tr}\left[\hat{\pi}(\mathbf{Z}_{A})\, \hat{T}_{\mathrm{A}}\, \hat{\pi}(\mathbf{Z}_{A}) \otimes \hat{\pi}(\mathbf{Z}_{B})\, \hat{T}_{\mathrm{B}}\, \hat{\pi}(\mathbf{Z}_{B})\ \hat{\rho}\right]. \end{aligned} \tag{5.8}$$

Although $\hat{T}_{\mathrm{A}}$ and $\hat{T}_{\mathrm{B}}$ are generic operators, $\hat{\pi}(\mathbf{Z}_{A})\, \hat{T}_{\mathrm{A}}\, \hat{\pi}(\mathbf{Z}_{A})$ and $\hat{\pi}(\mathbf{Z}_{B})\, \hat{T}_{\mathrm{B}}\, \hat{\pi}(\mathbf{Z}_{B})$ are, respectively, gauge-invariant operators entirely localized on regions A and B. Recalling that $\hat{\rho}$ is assumed to be gauge-separable, the following holds

$$\begin{aligned}\langle \hat{T}_{\mathrm{A}} \otimes \hat{T}_{\mathrm{B}}\rangle_{\rho} = &\operatorname{Tr}\left(\hat{\pi}(\mathbf{Z}_{A})\, \hat{T}_{\mathrm{A}}\, \hat{\pi}(\mathbf{Z}_{A}) \otimes \hat{\mathbb{1}}_{\mathrm{B}}\ \hat{\rho}\right) \times \\ &\qquad\qquad \operatorname{Tr}\left(\hat{\mathbb{1}}_{\mathrm{A}} \otimes \hat{\pi}(\mathbf{Z}_{B})\, \hat{T}_{\mathrm{B}}\, \hat{\pi}(\mathbf{Z}_{B})\ \hat{\rho}\right) .\end{aligned} \tag{5.9}$$

Reabsorbing the projectors into $\hat{\rho}$, we are left with

$$\langle \hat{T}_{\mathrm{A}} \otimes \hat{T}_{\mathrm{B}}\rangle_{\rho} = \langle \hat{T}_{\mathrm{A}} \otimes \hat{\mathbb{1}}_{\mathrm{B}}\rangle_{\rho} \langle \hat{\mathbb{1}}_{\mathrm{A}} \otimes \hat{T}_{\mathrm{B}}\rangle_{\rho} \,, \tag{5.10}$$

proving that, at least in the special case of a state lying completely in a subsector, gauge-separability implies general-separability.

To generalize this statement to any gauge-separable *pure* state, it is sufficient to prove that any such state is bound to lie entirely on a single subsector, where the equivalence of gauge-separability and general-separablity is a consolidated result.

For instance, Eq. (5.2) forces any pure state arising in a LGT to lie within a single sector $\mathcal{H}(\mathbf{Z})$. Still, $\hat{\rho}$ is allowed to have components in multiple subsectors that are compatible with $\mathcal{H}(\mathbf{Z})$. Indeed, $\hat{\rho}$ allows in general to be decomposed as

$$\begin{aligned}&\hat{\rho} = \sum_{(\mathbf{Z}_{\mathrm{A}},\mathbf{Z}_{\mathrm{B}})}^{\prime} \ \sum_{(\mathbf{Z}^{*}_{\mathrm{A}},\mathbf{Z}^{*}_{\mathrm{B}})}^{\prime} \hat{\rho}(\mathbf{Z}_{\mathrm{A}}, \mathbf{Z}_{\mathrm{B}}; \mathbf{Z}^{*}_{\mathrm{A}}, \mathbf{Z}^{*}_{\mathrm{B}}) \quad \text{where} \\ &\hat{\rho}(\mathbf{Z}_{\mathrm{A}}, \mathbf{Z}_{\mathrm{B}}; \mathbf{Z}^{*}_{\mathrm{A}}, \mathbf{Z}^{*}_{\mathrm{B}}) = \hat{\pi}_{\mathrm{A}}(\mathbf{Z}_{A}) \otimes \hat{\pi}_{\mathrm{B}}(\mathbf{Z}_{B})\ \hat{\rho}\ \hat{\pi}_{\mathrm{A}}(\mathbf{Z}^{*}_{\mathrm{A}}) \otimes \hat{\pi}_{\mathrm{B}}(\mathbf{Z}^{*}_{\mathrm{B}}) \,.\end{aligned} \tag{5.11}$$

The ($\prime$)s in the summations stress that the subsectors' labels of regions A and B must be compatible with the ones associated with the sector $\mathcal{H}(\mathbf{Z})$ and thus they are required to satisfy Eq. (5.7). Since $\hat{\rho}$ is assumed to be gauge-separable, it surely factorizes the expectation value of any projector $\hat{\pi}(\mathbf{X}_{\mathrm{A}}) \otimes \hat{\pi}(\mathbf{X}_{\mathrm{B}})$, so that

$$\begin{aligned}&\operatorname{Tr}\left(\hat{\pi}(\mathbf{X}_{\mathrm{A}}) \otimes \hat{\pi}(\mathbf{X}_{\mathrm{B}})\, \hat{\rho}\right) = \\ &\quad = \operatorname{Tr}\left(\hat{\pi}(\mathbf{X}_{\mathrm{A}}) \otimes \hat{\mathbb{1}}_{\mathrm{B}}\, \hat{\rho}\right) \times \operatorname{Tr}\left(\hat{\mathbb{1}}_{\mathrm{A}} \otimes \hat{\pi}(\mathbf{X}_{\mathrm{B}})\, \hat{\rho}\right) = \\ &\quad = \sum_{\mathbf{Y}_{\mathrm{B}}}^{\prime} \operatorname{Tr}\left(\hat{\pi}(\mathbf{X}_{\mathrm{A}}) \otimes \hat{\pi}(\mathbf{Y}_{\mathrm{B}})\, \hat{\rho}\right) \times \sum_{\mathbf{Y}_{\mathrm{A}}}^{\prime} \operatorname{Tr}\left(\hat{\pi}(\mathbf{Y}_{\mathrm{A}}) \otimes \hat{\pi}(\mathbf{X}_{\mathrm{B}})\, \hat{\rho}\right) .\end{aligned} \tag{5.12}$$

By Eq. (5.7), once the sector, $\mathcal{H}(\mathbf{Z})$, and the subsector's labels of region A (or B) have been fixed, the labels associated with region B (or A) are univoquely determined. This also holds in the case where the regions A and B are not complementary to one another. To show this, let us start by denoting with C the complementary region to $\mathrm{A} \cup \mathrm{B}$. In doing so, we identify with V_{AB}, V_{BC}, and V_{AC} the lattice sites lying at the interface, respectively, of regions A and B, B and C, and A and C. This allows us to extend Eq. (5.7) as

$$\begin{aligned} Z^g_x &= Z^g_{\mathrm{A},x}\, Z^g_{\mathrm{B},x}, \quad \forall x \in V_{\mathrm{AB}}\,,\ \forall g \in \mathcal{G} \\ Z^g_x &= Z^g_{\mathrm{B},x}\, Z^g_{\mathrm{C},x}, \quad \forall x \in V_{\mathrm{BC}}\,,\ \forall g \in \mathcal{G} \\ Z^g_x &= Z^g_{\mathrm{A},x}\, Z^g_{\mathrm{C},x}, \quad \forall x \in V_{\mathrm{AC}}\,,\ \forall g \in \mathcal{G}\,. \end{aligned} \tag{5.13}$$

As no local operator supported on regions A or B can address degrees of freedom belonging to region C, the state whose separability is tested is indistinguishable from $\hat{\rho} = \mathrm{Tr}_{\mathrm{C}}\, \hat{\rho}_{\mathrm{tot}}$, where $\hat{\rho}_{\mathrm{tot}}$ is the density operator supported on the whole lattice. As, in this proof, we assumed $\hat{\rho}$ to be a *pure* state, then surely the labels associated with the subsector localized on region C are fully fixed. Because of this and of Eq. (5.13), when fixing the sector, $\mathcal{H}(\mathbf{Z})$, and a subsector $\mathcal{H}_{\mathrm{A}}(\mathbf{Z}_A)$ (or $\mathcal{H}_{\mathrm{B}}(\mathbf{Z}_B)$)[2], the labels associated with the subsector defined on B (or A) is univoquely determined.

With the above considerations, the terms $\hat{\pi}(\mathbf{Y}_{\mathrm{A}}) \otimes \hat{\pi}(\mathbf{X}_{\mathrm{B}})$ and $\hat{\pi}(\mathbf{X}_{\mathrm{A}}) \otimes \hat{\pi}(\mathbf{Y}_{\mathrm{B}})$ in Eq. (5.12) give a non-vanishing contribution only if $\mathbf{Y}_{\mathrm{A}} = \mathbf{X}_{\mathrm{A}}$ and $\mathbf{Y}_{\mathrm{B}} = \mathbf{X}_{\mathrm{B}}$, respectively. This leads us to

$$\mathrm{Tr}\left(\hat{\rho}(\mathbf{X}_{\mathrm{A}}, \mathbf{X}_{\mathrm{B}}; \mathbf{X}_{\mathrm{A}}, \mathbf{X}_{\mathrm{B}})\right) = \mathrm{Tr}\left(\hat{\rho}(\mathbf{X}_{\mathrm{A}}, \mathbf{X}_{\mathrm{B}}; \mathbf{X}_{\mathrm{A}}, \mathbf{X}_{\mathrm{B}})\right)^2, \tag{5.14}$$

which establishes that any pure gauge-separable state is allowed to have diagonal components in a single subsector. Still, this requirement is apparently met by any state of the form

$$\begin{aligned} \hat{\rho} = \hat{\rho}(\mathbf{Z}_{\mathrm{A}}, \mathbf{Z}_{\mathrm{B}}; \mathbf{Z}_{\mathrm{A}}, \mathbf{Z}_{\mathrm{B}}) + \\ + \sum'_{(\mathbf{X}_{\mathrm{A}}, \mathbf{X}_{\mathrm{B}}) \neq (\mathbf{Y}_{\mathrm{A}}, \mathbf{Y}_{\mathrm{B}})} \Big(\hat{\rho}(\mathbf{X}_{\mathrm{A}}, \mathbf{X}_{\mathrm{B}}; \mathbf{Y}_{\mathrm{A}}, \mathbf{Y}_{\mathrm{B}}) + \mathrm{h.c.}\Big)\,. \end{aligned} \tag{5.15}$$

Assuming such a state exists, it has to be semidefinite-positive, implying the existence of an operator, $\hat{B}$, featuring $\hat{\rho} = \hat{B}\hat{B}^\dagger$. The form of $\hat{\rho}$ in Eq. (5.15) forces $\hat{B}$ to take components in arbitrary subsectors other than the one labeled by $(\mathbf{Z}_{\mathrm{A}}, \mathbf{Z}_{\mathrm{B}})$, causing also $\hat{\rho}$ to have components in every such subsector. Since we already proved that $\hat{\rho}$ has diagonal components

---

[2] Which has to be in agreement with $\mathcal{H}_{\mathrm{C}}(\mathbf{Z}_C)$.

in only one subsector, we conclude that any off-diagonal term in Eq. (5.15) has to be discarded.

This proves that any pure gauge-separable state $\hat{\rho}$, is bound to lie in a single subsector, where gauge-separability and general-separability have been already proven to be equivalent.

Although the definitions of separability presented (i.e., general-separability and gauge-separability) are equivalent from a theoretical standpoint, from a more practical perspective, the definition relying only on gauge-invariant observables has the huge advantage of requiring to test much less operators with respect to the definition employing generic operators.

5.1.3 EW existence theorem

Outside the LGT framework, it is a consolidated result that every entangled state is detected at least by one EW [68, 199]. In what follows, we extend this result to the specific case of LGTs, where only gauge-invariant witnesses are to be considered.

Since any gauge-entangled state, $\hat{\rho}_{\text{ent}}$, is also generically-entangled, we are provided with at least one observable $\hat{W} \in \mathrm{B}(\mathcal{H}_{\text{tot}})$ reproducing a negative expectation value on $\hat{\rho}_{\text{ent}}$ and non-negative values on any separable state. Until now, we are guaranteed that any entangled state arising in the framework of LGTs is recognized by an EW, $\hat{W}$, that is not *a priori* gauge-invariant. In what follows, we illustrate how it is possible to distill out of $\hat{W}$ a gauge-invariant operator retaining the same detection capabilities as $\hat{W}$ when tested on physical states. For the purpose, let us consider a generic gauge-invariant state $\hat{\rho}$ and its sector-wise decomposition

$$\hat{\rho} = \sum_{\mathbf{Z}} \hat{\rho}(\mathbf{Z}) \quad \text{with} \quad \hat{\rho}(\mathbf{Z}) = \hat{\pi}(\mathbf{Z})\, \hat{\rho}\, \hat{\pi}(\mathbf{Z}) \in \mathrm{B}(\mathcal{H}(\mathbf{Z}))\,, \tag{5.16}$$

which enables the rewriting of $\langle \hat{W} \rangle_{\rho}$ in the form

$$\begin{aligned} \mathrm{Tr}\left(\hat{W}\,\hat{\rho}\right) = & \,\mathrm{Tr}\left[\hat{W}\,\left(\sum_{\mathbf{Z}} \hat{\pi}(\mathbf{Z})\,\hat{\rho}\,\hat{\pi}(\mathbf{Z})\right)\right] = \\ & \mathrm{Tr}\left[\left(\sum_{\mathbf{Z}} \hat{\pi}(\mathbf{Z})\,\hat{W}\,\hat{\pi}(\mathbf{Z})\right)\,\hat{\rho}\right] = \\ & \mathrm{Tr}\left(\hat{W}'\,\hat{\rho}\right). \end{aligned} \tag{5.17}$$

Since $\hat{W}$ was chosen to be an EW able to detect $\hat{\rho}_{\text{ent}}$, as a direct consequence of Eq. (5.17), also $\hat{W}'$ is an EW that qualifies $\hat{\rho}_{\text{ent}}$ as being entangled. Moreover, $\hat{W}'$ commutes with the generators of the gauge transformation and thus is gauge-invariant.

With this, we proved in particular that any gauge-invariant entangled state has a gauge-invariant EW capable of detecting it.

5.2 Construction of EWs in LGTs

To devise an EW that is compatible with the local symmetries that characterize the LGT, we start by considering two sets of local and gauge-invariant observables, $\{\hat{O}_{\text{A}}^{(i)}\}_i$ and $\{\hat{O}_{\text{B}}^{(i)}\}_i$, supported entirely on regions A and B, respectively. Then, the operator

$$\hat{C} = \sum_i \hat{O}_{\text{A}}^{(i)} \otimes \hat{O}_{\text{B}}^{(i)} \tag{5.18}$$

is by construction a gauge-invariant observable supported on $\text{A} \cup \text{B}$. As such, this operator will simultaneously preserve the boundary conditions of regions A and B, enabling its rewriting as

$$\begin{aligned} \hat{C} &= \sum_{\mathbf{Z}_{\text{A}},\mathbf{Z}_{\text{B}}} \hat{C}(\mathbf{Z}_{\text{A}},\mathbf{Z}_{\text{B}}) \qquad \text{with} \\ \hat{C}(\mathbf{Z}_{\text{A}},\mathbf{Z}_{\text{B}}) &= \hat{\pi}_{\text{A}}(\mathbf{Z}_A) \otimes \hat{\pi}_{\text{B}}(\mathbf{Z}_B)\, \hat{C}\, \hat{\pi}_{\text{A}}(\mathbf{Z}_A) \otimes \hat{\pi}_{\text{B}}(\mathbf{Z}_B)\,. \end{aligned} \tag{5.19}$$

The next step, consists in optimizing $\langle \hat{C}(\mathbf{Z}_{\text{A}},\mathbf{Z}_{\text{B}})\rangle$ on the set of separable states residing in each subsector $\mathcal{H}_{\text{A}}(\mathbf{Z}_A)\otimes\mathcal{H}_{\text{B}}(\mathbf{Z}_B)$, leading to real-valued extremals $\omega^{\max}_{\mathbf{Z}_{\text{A}},\mathbf{Z}_{\text{B}}}$ and $\omega^{\min}_{\mathbf{Z}_{\text{A}},\mathbf{Z}_{\text{B}}}$. At this point, the operators

$$\begin{aligned} \hat{W}^{+} &= \sum_{\mathbf{Z}_{\text{A}},\mathbf{Z}_{\text{B}}} \omega^{\max}_{\mathbf{Z}_{\text{A}},\mathbf{Z}_{\text{B}}}\, \hat{\pi}_{\text{A}}(\mathbf{Z}_A) \otimes \hat{\pi}_{\text{B}}(\mathbf{Z}_B) - \hat{C}(\mathbf{Z}_{\text{A}},\mathbf{Z}_{\text{B}}) \\ \hat{W}^{-} &= \sum_{\mathbf{Z}_{\text{A}},\mathbf{Z}_{\text{B}}} \hat{C}(\mathbf{Z}_{\text{A}},\mathbf{Z}_{\text{B}}) - \omega^{\min}_{\mathbf{Z}_{\text{A}},\mathbf{Z}_{\text{B}}} \hat{\pi}_{\text{A}}(\mathbf{Z}_A) \otimes \hat{\pi}_{\text{B}}(\mathbf{Z}_B) \end{aligned} \tag{5.20}$$

are by construction block semidefinite positive. Before validating these observables as EWs, it is necessary to make sure they can detect at least one entangled state. Formally, this corresponds to checking if the observables $\hat{W}^{\pm}$ have at least a negative eigenvalue. If they do, they meet all theoretical prescriptions to be qualified as EWs. In addition, they are inherently gauge-invariant, as they result from a linear combination of two gauge-invariant operators: the projectors $\hat{\pi}_{\mathrm{A}}(\mathbf{Z}_A) \otimes \hat{\pi}_{\mathrm{B}}(\mathbf{Z}_B)$ and $\hat{C}$ (see Eq. (5.20)).

Until now, we have outlined a viable protocol to construct gauge-invariant EWs; in the following section, we focus on the problem of optimizing EWs within LGTs.

5.3 EW optimization algorithm

As a starting point for this section, we review some basic definitions concerning the optimality of EWs. Afterward, we describe the specific concerns for LGTs and the implications that come from the presence of local constraints. As a concrete example, an explicit analytical result describing how to optimize a witness in a pure $U(1)$ LGT is provided.

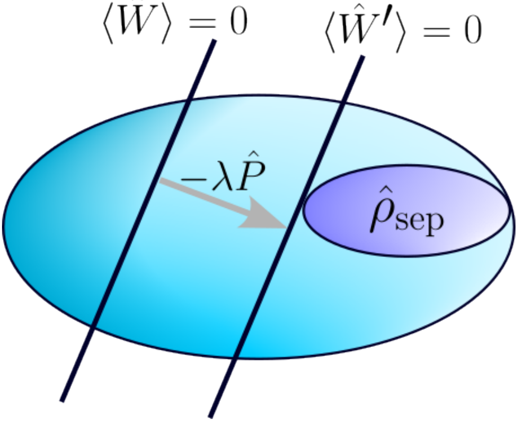


Figure 5.1: Both $\hat{W}$ and $\hat{W}'$ are valid EWs, as the set of separable states lies completely on one side, respectively, of the hyperplanes $\langle\hat{W}\rangle = 0$ and $\langle\hat{W}'\rangle = 0$. As $\hat{W}' = \hat{W} - \lambda\hat{P}$ detects all the entangled states that $\hat{W}$ does, along with additional ones, $\hat{W}'$ it is said to be *finer* than $\hat{W}$. Crucially, to qualify $\hat{W}'$ as an optimal EW, it is not sufficient to have a non-empty intersection between the hyperplane associated with $\hat{W}'$ and the set of separable states. On the contrary, $\hat{W}'$ is optimal if any possible positive operator, once subtracted from $\hat{W}'$ results in an operator that no longer meets the requirements for being an EW.

5.3.1 Optimality of EWs

Before specializing the discussion to the optimization of EWs within LGTs, it is conventient to first consider the more general case of EW optimization [73]. In this broader setting, there is no need to impose gauge invariance on the EW or its refinements, nor to account for the presence of superselection sectors.

Given an EW, $\hat{W}$, we say it is optimal if, for every positive operator $\hat{P}$, the operator $\hat{W}' = \hat{W} - \hat{P}$ is not an EW [73]. Concretely, this means that for every positive operator $\hat{P}$ there is a product state $|\psi_{\mathrm{sep}}\rangle$ satisfying

$$\langle\psi_{\mathrm{sep}}|\,\hat{W}\,|\psi_{\mathrm{sep}}\rangle < \langle\psi_{\mathrm{sep}}|\,\hat{P}\,|\psi_{\mathrm{sep}}\rangle \ , \tag{5.21}$$

which obstructs any potential refinement to the witness. Conversely, when this is not the case, we can search for a positive operator $\hat{P}$ without this obstruction and use it to construct a finer witness. After all, the positivity of $\hat{P}$ guarantees that $\hat{W}' = \hat{W} - \hat{P}$ detects all entangled states detected by $\hat{W}$, plus additional ones. Thus, the problem of optimizing an EW boils down to finding a suitable positive operator to subtract from it. A *a priori* strategy is to randomly select $\hat{P}$ and test if it is possible to remove a finite fraction of it from the EW. However, as the optimization protocol proceeds, this approach would be increasingly inefficient, as fewer $\hat{P}$ can be employed. To overcome this, as reported in Ref. [73], a convenient approach to find viable candidate directions for optimization, $\hat{P}$, consists in first identifying a set, $\mathcal{P}_W$, of separable states, $|\psi_{\mathrm{A}}\rangle \otimes |\psi_{\mathrm{B}}\rangle$, satisfying

$$\langle\psi_{\mathrm{A}}\psi_{\mathrm{B}}|\hat{W}|\psi_{\mathrm{A}}\psi_{\mathrm{B}}\rangle = 0\,. \tag{5.22}$$

Then, for any choice of $\hat{P}$ that does not annihilate *all* states in $\mathcal{P}_W$ — i.e., it exists a state $|\Psi_{\mathrm{sep}}\rangle \in \mathcal{P}_W$ such that $\hat{P}\,|\Psi_{\mathrm{sep}}\rangle \neq 0$ — we will have

$$\langle\hat{W} - \lambda\hat{P}\rangle_{\Psi_{\mathrm{sep}}} = -\lambda\langle\hat{P}\rangle_{\Psi_{\mathrm{sep}}} \ , \tag{5.23}$$

which, for any $\lambda > 0$, results in a negative expectation value, concluding that $\hat{P}$ can not be adopted to refine $\hat{W}$. Then, the viable choices for $\hat{P}$ are to be found among positive operators that annihilate any state belonging to $\mathcal{P}_W$, thus substantially simplifying the endeavor for potential optimization directions. Once a suitable direction, $\hat{P}$, has been singled out, we define

$$\lambda = \max\left\{\lambda \in \mathbb{R} \mid \hat{W} - \lambda\hat{P} \text{ is a witness}\right\}$$

to be the maximum fraction of $\hat{P}$ that can be removed from $\hat{W}$ while preserving $\hat{W}' = \hat{W} - \lambda\hat{P}$ as a valid EW. By construction, $\hat{W}' = \hat{W} - \lambda\,\hat{P}$ is the finest EW that can be obtained by subtracting the positive operator $\hat{P}$ to $\hat{W}$. This procedure is iterated until no viable positive operator $\hat{P}$ can be subtracted, namely when the EW is optimal.

5.3.2 Optimality of EWs in LGTs

In this section, we detail the case in which the optimization of the EW has to be consistent with local symmetries. In this case, the optimization process has to employ gauge-invariant positive operators, $\hat{P}$, having the form

$$\begin{aligned} &\hat{P} = \sum_{\mathbf{Z}_\mathrm{A},\mathbf{Z}_\mathrm{B}} \hat{P}(\mathbf{Z}_\mathrm{A},\mathbf{Z}_\mathrm{B}) \quad \text{with} \\ &\hat{P}(\mathbf{Z}_\mathrm{A},\mathbf{Z}_\mathrm{B}) = \hat{\pi}_\mathrm{A}(\mathbf{Z}_A)\otimes\hat{\pi}_\mathrm{B}(\mathbf{Z}_B)\;\hat{P}\;\hat{\pi}_\mathrm{A}(\mathbf{Z}_A)\otimes\hat{\pi}_\mathrm{B}(\mathbf{Z}_B)\,. \end{aligned} \tag{5.24}$$

An approach to refine $\hat{W}$ is to consider

$$\hat{W}' = \sum_{\mathbf{Z}_\mathrm{A},\mathbf{Z}_\mathrm{B}} \hat{W}(\mathbf{Z}_\mathrm{A},\mathbf{Z}_\mathrm{B}) - \lambda_{\mathbf{Z}_\mathrm{A},\mathbf{Z}_\mathrm{B}}\,\hat{P}(\mathbf{Z}_\mathrm{A},\mathbf{Z}_\mathrm{B})\,, \tag{5.25}$$

where subsector-specific multipliers $\lambda_{\mathbf{Z}_\mathrm{A},\mathbf{Z}_\mathrm{B}}$ are involved. It becomes apparent that the request of $\hat{W}'$ to be a valid EW can be equally expressed by asking each

$$\hat{W}'(\mathbf{Z}_\mathrm{A},\mathbf{Z}_\mathrm{B}) = \hat{W}(\mathbf{Z}_\mathrm{A},\mathbf{Z}_\mathrm{B}) - \lambda_{\mathbf{Z}_\mathrm{A},\mathbf{Z}_\mathrm{B}}\,\hat{P}(\mathbf{Z}_\mathrm{A},\mathbf{Z}_\mathrm{B}) \tag{5.26}$$

to be a valid EW. The finest refinement of $\hat{W}$ that can be achieved employing $\hat{P}$, makes use of the multipliers defined by

$$\lambda_{\mathbf{Z}_\mathrm{A},\mathbf{Z}_\mathrm{B}} = \max\left\{\lambda_{\mathbf{Z}_\mathrm{A},\mathbf{Z}_\mathrm{B}} \in \mathbb{R} \mid \hat{W}(\mathbf{Z}_\mathrm{A},\mathbf{Z}_\mathrm{B}) - \lambda_{\mathbf{Z}_\mathrm{A},\mathbf{Z}_\mathrm{B}}\,\hat{P}(\mathbf{Z}_\mathrm{A},\mathbf{Z}_\mathrm{B}) \text{ is a witness}\right\}\,. \tag{5.27}$$

The optimality criteria for $\hat{W}$ is satisfied if each of its projections is optimal. Stated differently, $\hat{W}$ is optimal if there is no subsector $\mathcal{H}_\mathrm{A}(\mathbf{Z}_A)\otimes\mathcal{H}_\mathrm{B}(\mathbf{Z}_B)$ where a positive gauge-invariant operator defined therein, $\hat{P}(\mathbf{Z}_\mathrm{A},\mathbf{Z}_\mathrm{B})$, can be further subtracted from $\hat{W}$. With this proviso, we approach the optimization process by focusing sequentially on individual sectors, where only the relevant projection of the witness is treated.

In the following section, we adopt this approach to optimize an EW within a pure $U(1)$ LGT.

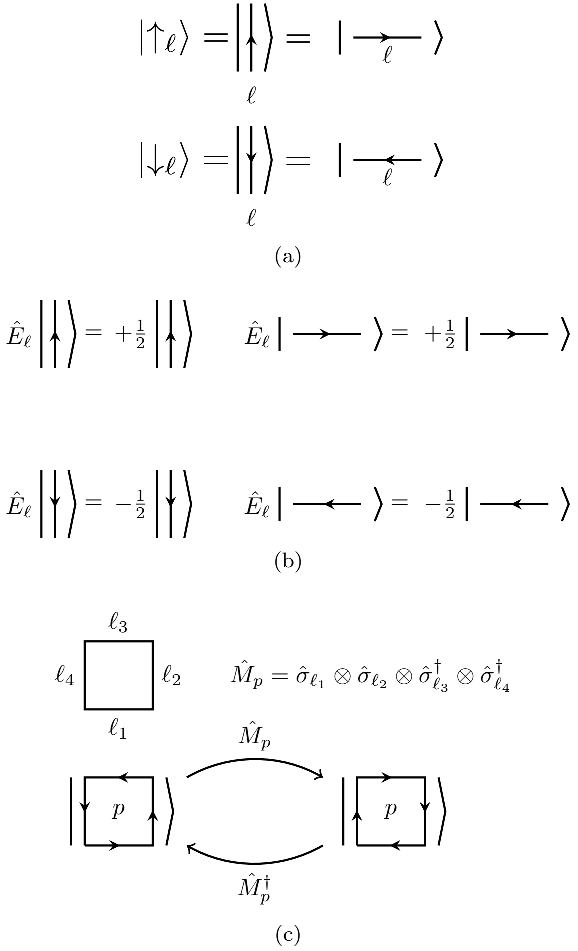


Figure 5.2: In **(a)** we graphically represent the basis vectors that span the local Hilbert spaces attached to edges. Horizontal (vertical) edges, ℓ, whose orientation points upward (rightward) are in the state $|\uparrow\rangle$. Otherwise, they are in the $|\downarrow\rangle$ state. In **(b)** we depict the action of electric field operators, while in **(c)** is represented the action of magnetic operators.

5.3.3 Analytical example concerning a pure $U(1)$ LGT

We consider a pure $U(1)$ LGT where electric charges are not involved and the only degrees of freedom are electric fields defined on the (oriented) lattice's edges, E. For illustration purposes, we model these degrees of freedom as two-dimensional Hilbert spaces spanned by the basis states $\{|\uparrow\rangle, |\downarrow\rangle\}$, whose graphical representation is depicted in Fig. 5.2a. In particular, we interpret vertical (horizontal) edges whose orientation points upward (right) to be in the $|\uparrow\rangle$ state, otherwise, they are in the $|\downarrow\rangle$ state. As the $U(1)$ group is entirely generated by a single group element, g, every rotation is expressed as $e^{i\theta g}$, and the generators acting at vertex x are

$$\hat{U}_x^{(e^{i\theta g})} = \prod_{(x,y)\in E} e^{i\theta\hat{g}_{(x,y)}} \otimes \prod_{(y,x)\in E} \left(e^{i\theta\hat{g}_{(y,x)}}\right)^{-1} . \tag{5.28}$$

From the elementary properties of the commutator, every operator commuting with

$$\hat{U}_x = \sum_{(x,y)\in E} \hat{g}_{(x,y)} - \sum_{(y,x)\in E} \hat{g}_{(y,x)} \tag{5.29}$$

also commutes with operators defined through Eq. (5.28). In the present framework, the representation of the group element g is given by $\hat{g} = \frac{1}{2}|\uparrow\rangle\langle\uparrow| - \frac{1}{2}|\downarrow\rangle\langle\downarrow|$. There are two kinds of gauge-invariant operators; these are (i) electric field operators, $\hat{E}_\ell = \hat{g}_\ell$ (whose action is depicted in Fig. 5.2b), and (ii) magnetic operators acting on plaquettes, p, $\hat{M}_p = \hat{\sigma}_{\ell_1} \otimes \hat{\sigma}_{\ell_2} \otimes \hat{\sigma}^\dagger_{\ell_3} \otimes \hat{\sigma}^\dagger_{\ell_4}$, where ℓ_i are the edges enclosing p and $\hat{\sigma}_\ell = |\downarrow\rangle\langle\uparrow|$ (see Fig. 5.2c). From these operators derive two kinds of non-diagonal gauge-invariant observables, namely $\hat{\mathcal{M}}^x_p = \hat{M}_p + \hat{M}^\dagger_p$ and $\hat{\mathcal{M}}^y_p = i\hat{M}_p - i\hat{M}^\dagger_p$, where ℓ_i are the edges enclosing the plaquette p (the action of such observables is depicted in Fig. 5.3). Gauge-invariant witnesses can be built by linearly combin-

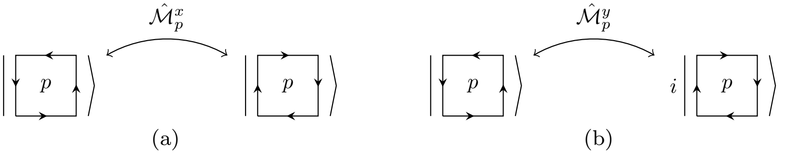


Figure 5.3: Action of $U(1)$ gauge-invariant observables having non diagonal entries. **(a)** If the plaquette p has a defined orientation, $\hat{\mathcal{M}}^x_p$ flips it; otherwise the state is annihilated. **(b)** $\hat{\mathcal{M}}^y_p$ applied to a positively (negatively) oriented plaquette gives a negatively (positively)-oriented plaquette with a $\pi/2$ $(-\pi/2)$ phase. If $\hat{\mathcal{M}}^y_p$ is applied to a plaquette having no orientation, the state is annihilated.

ing and/or taking tensor products of $\hat{E}_\ell$, $\hat{\mathcal{M}}^x_p$, and $\hat{\mathcal{M}}^y_p$. In the application below, we consider a rectangle $\mathcal{R}$ and two disjoint subregions of it, A and B (see Fig. 5.4a); the EW we optimize targets bipartite entanglement among these subregions. The superselection sector associated with region $\mathcal{R}$ is chosen as in Fig. 5.4b, so that two pairs of compatible superselection sectors associated with subregions A and B arise (see sketch in Fig. 5.4c).

$$\hat{W} = \hat{\mathbb{1}} + \left(\hat{\mathcal{M}}^x_{a_1}\otimes\hat{\mathcal{M}}^x_{b_2} + \hat{\mathcal{M}}^y_{a_1}\otimes\hat{\mathcal{M}}^y_{b_2}\right) + \left(\hat{\mathcal{M}}^x_{a_2}\otimes\hat{\mathcal{M}}^x_{b_1} + \hat{\mathcal{M}}^y_{a_2}\otimes\hat{\mathcal{M}}^y_{b_1}\right) . \tag{5.30}$$

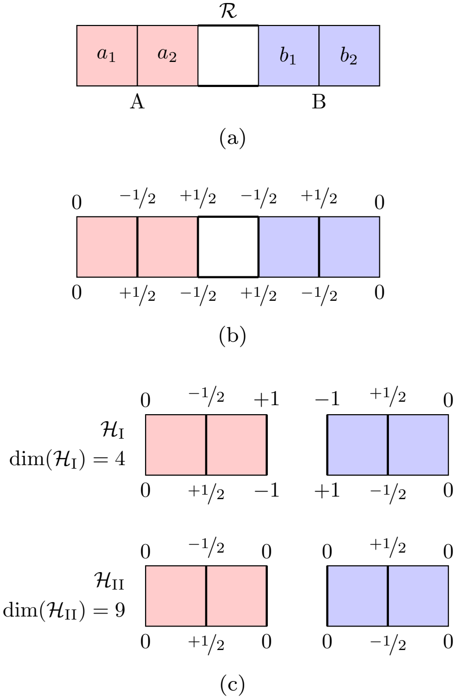


Figure 5.4: Details concerning geometric and boundary conditions choices for our proof-of-principle application. **(a)** bipartition of region $\mathcal{R}$ into two subregions A (red) and B (blue). **(b)** labels identifying the sector we choose for region $\mathcal{R}$. **(c)** labels associated with the two possible subsectors, $\mathcal{H}_\mathrm{I}$ and $\mathcal{H}_\mathrm{II}$, of regions A and B that compatible with $\mathcal{H}_\mathrm{R}$.

As a first step, we adopt the base elements represented in Figs. 5.5 and 5.6 and we use them to project the EW, respectively, onto $\mathcal{H}_{\mathrm{I}}$ and $\mathcal{H}_{\mathrm{II}}$, yielding

$$\begin{aligned}
\hat{W}_{\mathrm{I}} &= \hat{\mathbb{1}} + 2\left(\left|a^{(\mathrm{I})}b^{(\mathrm{I})}\right\rangle\!\left\langle a_1^{(\mathrm{I})}b_2^{(\mathrm{I})}\right| + \text{h.c.}\right)\\
\hat{W}_{\mathrm{II}} &= \hat{\mathbb{1}} + 2\left(\left|a^{(\mathrm{II})}\,b_2^{(\mathrm{II})}\right\rangle\!\left\langle a_1^{(\mathrm{II})}\,b^{(\mathrm{II})}\right| + \text{h.c.}\right) + \\
&\quad + 2\left(\left|a_2^{(\mathrm{II})}\,b^{(\mathrm{II})}\right\rangle\!\left\langle a^{(\mathrm{II})}\,b_1^{(\mathrm{II})}\right| + \text{h.c.}\right) .
\end{aligned} \tag{5.31}$$

It is possible to verify that both observables in Eq. (5.31) take non-negative expectation values when evaluated on separable states in $\mathcal{H}_{(\mathrm{I})}$ and in $\mathcal{H}_{(\mathrm{II})}$. To do so, it is sufficient to define the generic separable states

$$\begin{aligned}
\left|\Psi_{\mathrm{sep}}^{(\mathrm{I})}\right\rangle &= \left(\alpha\left|a^{(\mathrm{I})}\right\rangle + \alpha_1\left|a_1^{(\mathrm{I})}\right\rangle\right) \otimes \left(\beta\left|b^{(\mathrm{I})}\right\rangle + \beta_2\left|b_2^{(\mathrm{I})}\right\rangle\right)\\
\left|\Psi_{\mathrm{sep}}^{(\mathrm{II})}\right\rangle &= \left(\alpha\left|a^{(\mathrm{II})}\right\rangle + \alpha_1\left|a_1^{(\mathrm{II})}\right\rangle + \alpha_2\left|a_2^{(\mathrm{II})}\right\rangle\right) \otimes\\
&\qquad \left(\beta\left|b^{(\mathrm{II})}\right\rangle + \beta_1\left|b_1^{(\mathrm{II})}\right\rangle + \beta_2\left|b_2^{(\mathrm{II})}\right\rangle\right)
\end{aligned} \tag{5.32}$$

and optimize the complex-valued parameters appearing in Eq. (5.32) to minimize $\langle\Psi_{\mathrm{sep}}^{(\mathrm{I})}|\hat{W}_{\mathrm{I}}|\Psi_{\mathrm{sep}}^{(\mathrm{I})}\rangle$ and $\langle\Psi_{\mathrm{sep}}^{(\mathrm{II})}|\hat{W}_{\mathrm{II}}|\Psi_{\mathrm{sep}}^{(\mathrm{II})}\rangle$. As, in both cases, the minimal expectation value attained by $\hat{W}_{\mathrm{I}}$ and $\hat{W}_{(\mathrm{II})}$ is zero, then $\hat{W}$ satisfies the first condition to be a valid EW — namely it acquires non-negative expectation values when measured against *separable* states.

To confirm $\hat{W}$ as an effective EW, it is sufficient to consider the pure states

$$\begin{aligned}
\left|\Psi_{\mathrm{ent}}^{(\mathrm{I})}\right\rangle &= \frac{\left|a^{(\mathrm{I})}b^{(\mathrm{I})}\right\rangle - \left|a_1^{(\mathrm{I})}b_2^{(\mathrm{I})}\right\rangle}{\sqrt{2}}\\
\left|\Psi_{\mathrm{ent}}^{(\mathrm{II})}\right\rangle &= \frac{\left|a^{(\mathrm{II})}b_2^{(\mathrm{II})}\right\rangle - \left|a_1^{(\mathrm{II})}b^{(\mathrm{II})}\right\rangle}{\sqrt{2}} ,
\end{aligned} \tag{5.33}$$

and the related expectation values $\langle\Psi_{\mathrm{ent}}^{(\mathrm{I})}|\hat{W}_{\mathrm{I}}|\Psi_{\mathrm{ent}}^{(\mathrm{I})}\rangle = -1$ and $\langle\Psi_{\mathrm{ent}}^{(\mathrm{II})}|\hat{W}_{\mathrm{II}}|\Psi_{\mathrm{ent}}^{(\mathrm{II})}\rangle = -1$.

While $\hat{W}$ stands as a valid EW, it is not guaranteed to be optimal. In the following section, we undertake the task of optimizing $\hat{W}$ by following the protocol discussed in Section 5.3.2. Namely, we apply the refinement strategy in [73] by addressing individual subsectors.

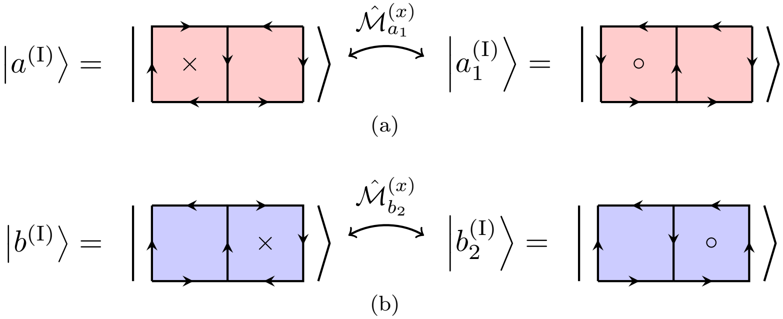
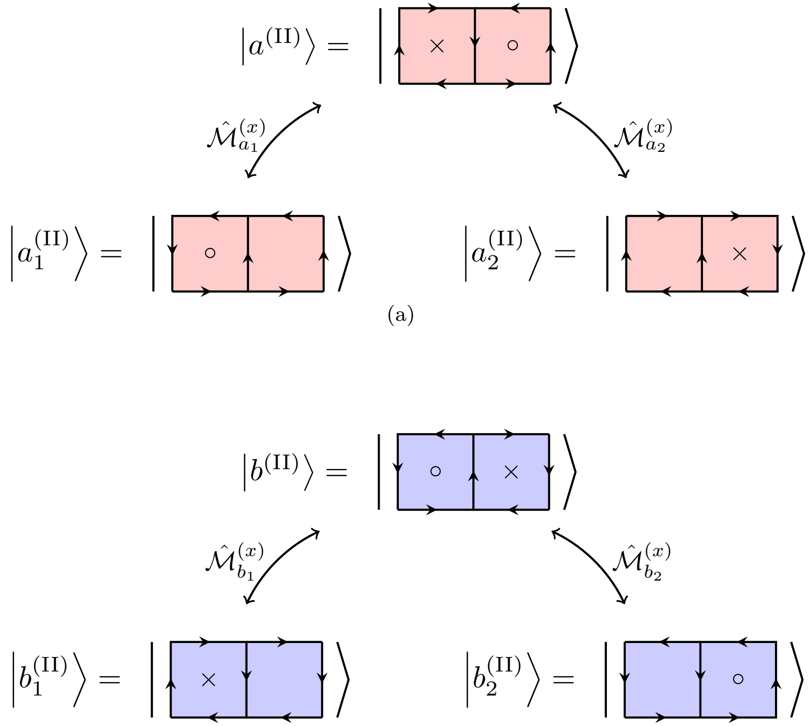


Figure 5.5: Base of subsector $\mathcal{H}_{\mathsf{I}} = \mathcal{H}_{\mathsf{I}}^{\mathrm{A}} \otimes \mathcal{H}_{\mathsf{I}}^{\mathrm{B}}$. In **(a)** we report the base of $\mathcal{H}_{\mathsf{I}}^{\mathrm{A}}$ and in **(b)** the base of $\mathcal{H}_{\mathsf{I}}^{\mathrm{B}}$.



(a)

(b)

Figure 5.6: Base of subsector $\mathcal{H}_{\mathsf{II}} = \mathcal{H}^{\mathrm{A}}_{\mathsf{II}} \otimes \mathcal{H}^{\mathrm{B}}_{\mathsf{II}}$. In **(a)** we report the base of $\mathcal{H}^{\mathrm{A}}_{\mathsf{II}}$ and in **(b)** the base of $\mathcal{H}^{\mathrm{B}}_{\mathsf{II}}$.

Concerning $\hat{W}_{\mathrm{I}}$, it is possible to verify that the separable states

$$\left|\psi_\alpha^{(\mathrm{I})}\right\rangle = \frac{1}{2}\Big(\left|a^{(\mathrm{I})}\right\rangle + \alpha\left|a_1^{(\mathrm{I})}\right\rangle\Big) \otimes \Big(\left|b^{(\mathrm{I})}\right\rangle - \alpha^*\left|b_2^{(\mathrm{I})}\right\rangle\Big), \tag{5.34}$$

with $\alpha \in \{+1,-1,i\}$, are such that $\langle\psi_\alpha^{(\mathrm{I})}|\hat{W}_{\mathrm{I}}|\psi_\alpha^{(\mathrm{I})}\rangle = 0$ i.e., according to the definitions in Ref. [73], they belong to the set $\mathcal{P}_W^{(\mathrm{I})}$. Based on Ref. [73], candidate positive operators to be subtracted from $\hat{W}_{(\mathrm{I})}$ must annihilate all elements in $\mathcal{P}_W^{(\mathrm{I})}$. In other words, as $\hat{P}$ operators are hermitian, their range has to be circumscribed to $\left(\mathcal{P}_W^{(\mathrm{I})}\right)^\perp$, which in this case is a 1-dimensional space spanned by the state $\left|\Omega^{(\mathrm{I})}\right\rangle = \frac{1}{\sqrt{2}}\big(|a^{(\mathrm{I})}b^{(\mathrm{I})}\rangle + |a_1^{(\mathrm{I})}b_2^{(\mathrm{I})}\rangle\big)$. This concludes that the only positive operator that can be adopted to refine $\hat{W}_{\mathrm{I}}$ is proportional to

$$\hat{P}^{(\mathrm{I})} = |\Omega^{(\mathrm{I})}\rangle\langle\Omega^{(\mathrm{I})}|\,. \tag{5.35}$$

Removing this positive operator from $\hat{W}_{\mathrm{I}}$, we register an increase of the fraction of detected entangled states from $6.30 \pm 0.08\,\%$ to values as high as $12.4 \pm 0.1\,\%$. These estimates are derived from evaluations performed on 100 batches, each containing 1,000 pure states randomly drawn from $\mathcal{H}_{\mathrm{I}}$.

Concerning the optimization of $\hat{W}_{\mathrm{II}}$, we notice that the states

$$\begin{aligned}
\left|\Psi_0^{(\mathrm{II})}\right\rangle &= \frac{1}{4}\Big(\sqrt{2}\left|a^{(\mathrm{II})}\right\rangle + \left|a_1^{(\mathrm{II})}\right\rangle + \left|a_2^{(\mathrm{II})}\right\rangle\Big)\otimes \\
&\qquad \otimes\Big(\sqrt{2}\left|b^{(\mathrm{II})}\right\rangle - \left|b_1^{(\mathrm{II})}\right\rangle - \left|b_2^{(\mathrm{II})}\right\rangle\Big) \\
\left|\Psi_\alpha^{(\mathrm{II})}\right\rangle &= \frac{1}{2}\Big(\left|a^{(\mathrm{II})}\right\rangle + \alpha\left|a_1^{(\mathrm{II})}\right\rangle\Big) \otimes \Big(\left|b^{(\mathrm{II})}\right\rangle - \alpha\left|b_2^{(\mathrm{II})}\right\rangle\Big), \\
\left|\Psi_\beta^{(\mathrm{II})}\right\rangle &= \frac{1}{2}\Big(\left|a^{(\mathrm{II})}\right\rangle + \beta\left|a_2^{(\mathrm{II})}\right\rangle\Big) \otimes \Big(\left|b^{(\mathrm{II})}\right\rangle - \beta\left|b_1^{(\mathrm{II})}\right\rangle\Big),
\end{aligned} \tag{5.36}$$

with $\alpha, \beta \in \{+1,-1,i\}$, belong to $\mathcal{P}_W^{\mathrm{II}}$ and span a 6-dimensional space. It is also straightforward to check that the states

$$\begin{aligned}
\left|\Omega_1^{(\mathrm{II})}\right\rangle &= \frac{1}{\sqrt{2}}\left(\left|a_1^{(\mathrm{II})}b_1^{(\mathrm{II})}\right\rangle - \left|a_2^{(\mathrm{II})}b_2^{(\mathrm{II})}\right\rangle\right), \\
\left|\Omega_2^{(\mathrm{II})}\right\rangle &= \frac{1}{\sqrt{2}}\left(\left|a^{(\mathrm{II})}b_2^{(\mathrm{II})}\right\rangle + \left|a_1^{(\mathrm{II})}b^{(\mathrm{II})}\right\rangle\right), \\
\left|\Omega_3^{(\mathrm{II})}\right\rangle &= \frac{1}{\sqrt{2}}\left(\left|a^{(\mathrm{II})}b_1^{(\mathrm{II})}\right\rangle + \left|a_2^{(\mathrm{II})}b^{(\mathrm{II})}\right\rangle\right),
\end{aligned} \tag{5.37}$$

span the orthogonal space to $\mathcal{P}_W^{\mathrm{II}}$ and thus can be employed to optimize $\hat{W}_{\mathrm{II}}$. Removing sequentially the positive operators $\hat{P}_n^{(\mathrm{II})} = |\Omega_n^{(\mathrm{II})}\rangle\langle\Omega_n^{(\mathrm{II})}|$, we estimate the fraction of detected entangled states at each iteration. The results, shown in Fig. 5.7, indicate an increase in the percentage of detected entangled states from $0.96 \pm 0.03\,\%$ up to $14.39 \pm 0.05\,\%$. This result is derived by testing at each refinement iteration 100 batches, each consisting of 1,000 random pure states, all belonging to $\mathcal{H}_{\mathrm{II}}$.

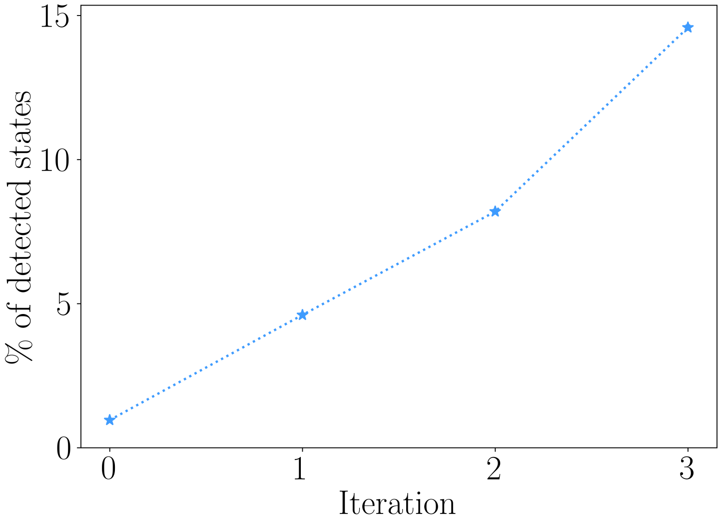


Figure 5.7: Fraction of detected entangled states throughout the optimization process of $\hat{W}_{\mathrm{II}}$. The positive operators employed are based on the directions in Eq. (5.37).

5.4 Numerical implementation and results

In this section, we discuss the details concerning the numerical implementation of the protocol outlined in Section 5.3.2. Further, we present additional results obtained by applying this algorithm to the case of the pure $U(1)$ LGT introduced in Section 5.3.3.

5.4.1 Algorithm implementation

In this section, we describe the details concerning the implementation of the algorithm we employed to derive the results in Sec. 5.4.2. The theoretical basis for our algorithm is detailed in Ref. [73]. Here, we discuss how the most crucial points are implemented.

The tasks that are needed to undertake a single refinement step are:

1. find the set of *separable* states $\mathcal{P}_W$ that produce a vanishing expectation value on $\hat{W}$,
2. consider positive operators, $\hat{P}$, that annihilate states in $\mathcal{P}_W$, and
3. evaluate the maximum amount of $\hat{P}$ that can be subtracted from $\hat{W}$ without invalidating the resulting EW.

Task 1. This problem traces back to finding the set $\mathcal{P}_W$ of states $|\psi_\mathrm{A}\rangle \otimes |\psi_\mathrm{B}\rangle$ such that $\langle \psi_\mathrm{A}\psi_\mathrm{B}|\hat{W}|\psi_A\psi_B\rangle = 0$, which is equivalent to optimizing the components of $|\psi_\mathrm{A}(\boldsymbol{\alpha})\rangle = \sum_n \alpha_n |n\rangle$, so that the partial operator $\hat{W}_B(\boldsymbol{\alpha}) = \langle \psi_\mathrm{A}(\boldsymbol{\alpha})|\hat{W}|\psi_\mathrm{A}(\boldsymbol{\alpha})\rangle$ has at least a vanishing eigenvalue. In practice, the algorithm optimizes $\boldsymbol{\alpha}$ based on the objective

$$f(\boldsymbol{\alpha}) = \hat{W}_\mathrm{B}(\boldsymbol{\alpha})\big|_{\text{min. eigenvalue}} \,. \tag{5.38}$$

Since $\hat{W}$ is an EW, $f(\boldsymbol{\alpha})$ is inherently a non-negative function that, when attaining its minima at zero for some parametrization $\bar{\boldsymbol{\alpha}}$, defines two states $|\psi_\mathrm{A}(\bar{\boldsymbol{\alpha}})\rangle$ and $|\psi_\mathrm{B}(\bar{\boldsymbol{\alpha}})\rangle \in \mathrm{Ker}(\hat{W}_\mathrm{B}(\bar{\boldsymbol{\alpha}}))$ whose tensor product state belongs to $\mathcal{P}_W$.

Due to limited machine precision, in our numerical implementation, we define the states belonging to $\mathcal{P}_W$ up to a threshold value $\epsilon = 10^{-14}$. Namely, we consider

$$\mathcal{P}_W^\epsilon = \{|\psi_\mathrm{A}\rangle \otimes |\psi_\mathrm{B}\rangle \mid -\epsilon \leq \langle \psi_\mathrm{A}\psi_\mathrm{B}|\hat{W}|\psi_\mathrm{A}\psi_\mathrm{B}\rangle \leq \epsilon\} \,. \tag{5.39}$$

Task 2. Once the set $\mathcal{P}_W^\epsilon$ is known, its orthogonal directions — if they exist — $\{|\phi_1\rangle, |\phi_2\rangle, \ldots, |\phi_m\rangle\}$ can be employed to build a positive operator, $\hat{P}$, that is to be subtracted from $\hat{W}$. The most general operator that can be considered writes

$$\hat{P} = \sum_{k=1}^{m} p_k \, |\Omega_k\rangle\langle\Omega_k| \, , \tag{5.40}$$

where p_k are non-negative values that sum up to 1 and $|\Omega_k\rangle$ are arbitrary linear combinations of $|\phi_i\rangle$ states. In our algorithm, $\hat{P}$ are chosen to be rank-1 operators achieved by setting all but one p_k to 0 and randomly choosing $|\Omega_k\rangle$ to be an arbitrary linear combination of $|\phi_i\rangle$ states.

Task 3. As a next step, it is necessary to evaluate the maximum λ such that $\hat{W} - \lambda\,\hat{P}$ is a valid EW (i.e., it is block semidefinite positive). In our application, we implement the scheme summarized in Fig. 5.8.

This protocol is a variation of the binary search algorithm that, despite converging at a slower pace, is more resilient to errors. To clarify the source of these errors, let us consider the problem of deciding whether $\hat{W}' = \hat{W} - \lambda\hat{P}$ is an EW. This is done by testing if $\hat{W}'$ acquires non-negative expectation values on separable states and testing their expectation values against $\hat{W}'$. This can *a priori* be fully settled by extensively enumerating all separable states. Clearly, this approach is impractical, so we must resort to a different strategy. Indeed, we implement an optimization protocol aiming at finding the parametrization $\boldsymbol{\alpha}$ that minimizes

$$g(\boldsymbol{\alpha}) = \hat{W}_{\mathrm{B}}(\boldsymbol{\alpha})\big|_{\text{min. eigenvalue}} \, , \tag{5.41}$$

with

$$\begin{cases} \hat{W}_{\mathrm{B}}(\boldsymbol{\alpha}) = \langle\psi_{\mathrm{A}}(\boldsymbol{\alpha})|\hat{W}|\psi_{\mathrm{A}}(\boldsymbol{\alpha})\rangle \\ |\psi_{\mathrm{A}}(\boldsymbol{\alpha})\rangle = \sum_n \alpha_n \, |n\rangle \end{cases} . \tag{5.42}$$

Despite this setting is formally identical to the one in Eq. (5.38), here $g(\boldsymbol{\alpha})$ is not guaranteed to be non-negative as, potentially, $\hat{W}'$ is not an EW. In principle, if the optimization technique were to be accurate, its outcome can be unambiguously employed to accept or reject $\hat{W}'$ as an EW. However, the numerical optimization is not guaranteed to converge to its global minimum, causing the interpretation of its results to be somewhat uncertain. To mitigate this we implement multiple times (10) the optimization procedure set by Eq. (5.41) and reject $\hat{W}'$ as an EW by checking if the attained minimum by $g(\boldsymbol{\alpha})$ is strictly lower than $-\gamma\varepsilon$, where $\gamma \in [0,1]$. To calibrate γ, we note that two distinct scenarios may arise. When γ is too small, the rejection criteria for candidate EWs $\hat{W}'$ becomes overly stringent. As a result, the accepted operators feature a $\mathcal{P}_{W'}$ set whose dimensionality does not improve compared to that of $\mathcal{P}_W$. Conversely, if γ is too large, the selection criteria may become too relaxed, leading to the incorrect classification of $\hat{W}'$ as an EW. In our numerical derivations, we adopted several values for this parameter, ranging from $\frac{1}{10}$ to $\frac{2}{3}$.

Until this point, we have discussed the mitigation techniques designed to assess whether $\hat{W}' = \hat{W} - \lambda\hat{P}$ satisfies the necessary conditions to qualify it as an EW. To further allow for some misclassification to take place, we modify the standard binary search as follows. Starting with a given search interval for λ, $[\lambda_{\text{min}}, \lambda_{\text{max}}]$, having width $\Delta = \lambda_{\text{max}} - \lambda_{\text{min}}$, we test block semidefinite positiveness of $\hat{W}' = \hat{W} - \lambda^*\hat{P}$, where $\lambda^* = \frac{\lambda_{\text{min}} + \lambda_{\text{max}}}{2}$. Based on the heuristic criteria described above, we accept or discard $\hat{W}'$ as a valid EW. As a consequence, when accepting $\hat{W}'$, the new search interval is updated to be

$$[\lambda'_{\text{min}}, \lambda'_{\text{max}}] \quad \text{with} \quad \begin{cases} \lambda'_{\text{min}} = \lambda + \frac{\Delta}{2} \\ \lambda'_{\text{max}} = \lambda_{\text{max}} \,. \end{cases} \tag{5.43}$$

On the countrary, when rejecting $\hat{W}'$, the new search interval is refined as

$$[\lambda'_{\text{min}}, \lambda'_{\text{max}}] \quad \text{with} \quad \begin{cases} \lambda'_{\text{min}} = \max\left\{0, \lambda - \frac{\Delta}{4}\right\} \\ \lambda'_{\text{max}} = \lambda_{\text{max}} - \frac{\Delta}{4} \,. \end{cases} \tag{5.44}$$

Importantly, the width of the refined interval equals $^{3\Delta}/_{4}$, guaranteeing convergence of the method.

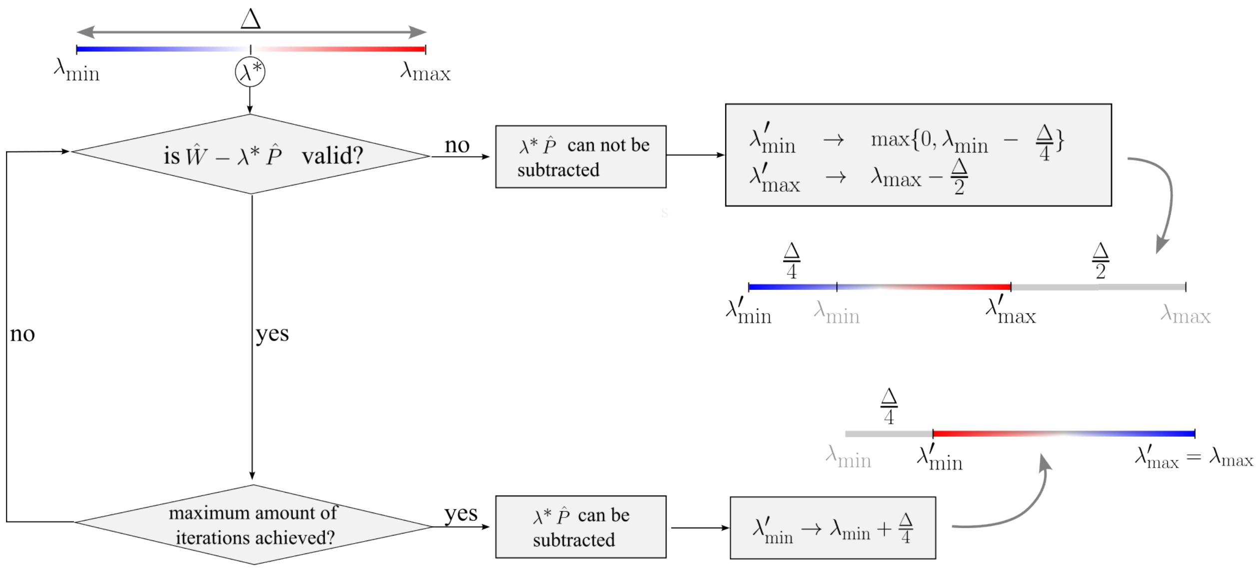


Figure 5.8: Variant of the binary search algorithm we implemented to find the maximum λ for which $\hat{W} - \lambda \hat{P}$ is a valid EW. For each λ^* tested, we minimize 10 times $\langle \psi_A \psi_B | \hat{W} - \lambda^* \hat{P} | \psi_A \psi_B \rangle$. Depending on the optimal values reached by optimizing Eq. (5.41), the interval's extremes are shifted. Irrespectively of the acceptance/rejection of λ^*, the width of the updated interval corresponds to $3/4$ of the original one, which guarantees the convergence of our method. Although this algorithm converges at a slower pace than standard binary-search protocols, it is less sensitive to misclassifications of $\hat{W}' = \hat{W} - \lambda^* \hat{P}$ as an EW.

5.4.2 Results

In section 5.3.3, the optimization of $\hat{W}_{\mathrm{I}}$ could be managed by subtracting a well-defined positive operator with no room for further freedom. The situation is completely different in the case of $\hat{W}_{\mathrm{II}}$, which allowed three directions for optimization, which, differently from what we did in the previous section, can be *a priori* linearly combined.

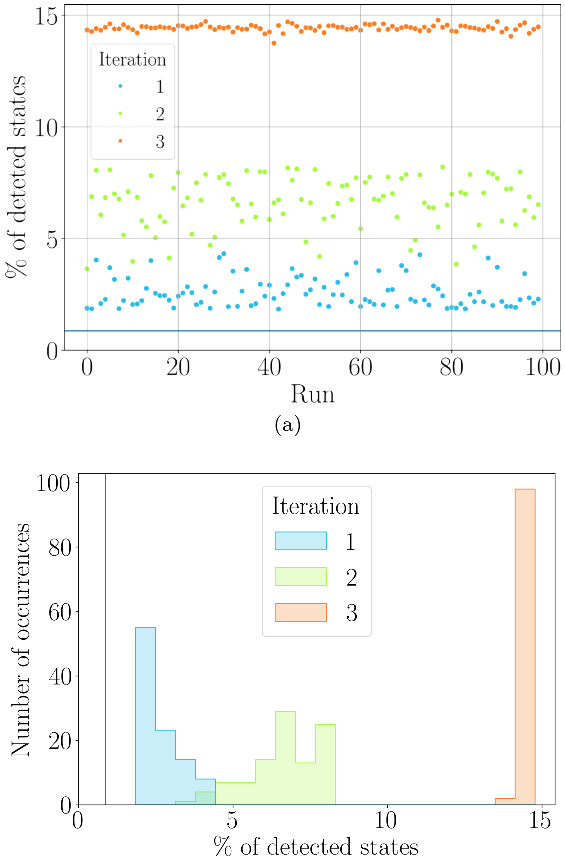
In this section, we explore the possibility of subtracting, at each optimization step, a positive operator $\hat{P}$ that arises from a random linear combination of the n available optimization directions $\{|\phi_1\rangle, \ldots, |\phi_n\rangle\}$. More specifically, we consider

$$\hat{P} = |\Omega\rangle\langle\Omega| \quad \text{with} \quad |\Omega\rangle = \sum_{i=1}^{n} \xi_i \, |\phi_i\rangle \tag{5.45}$$

and uniformly sampled complex components ξ. The full-optimization process, also referred to as *optimization trajectory*, is repeated 100 times and results concerning individual optimization trajectories are summarized in Fig. 5.9a, while Fig. 5.9b reports the statistics associated with the ensemble of trajectories. These results confirm that our algorithm is capable of selecting a viable optimization direction as the refinement process takes place. In addition to this, in all cases, the dimensionality of $\mathcal{P}_{W_3}$, where $\hat{W}_3$ is the entanglement wintess obtained after 3 optimization steps, satisfies the spanning property[3], meaning that it is guaranteed to be an optimal EW [73].

---

[3] If the $\mathcal{P}_W$ set associated with the EW $\hat{W} \in \mathrm{B}(\mathcal{H})$ is spanned by $\dim(\mathcal{H})$ linearly independent vectors, then $\hat{W}$ is said to satify the spanning property, which constitutes a sufficient condition for the optimality of $\hat{W}$.



(a)

(b)

Figure 5.9: The results shown concern the optimization of the EW defined in Eq. (5.30), within the superselection sector $\mathcal{H}_{\text{II}}$. Further details — including the lattice dimensionality, the selected region and its boundary conditions, the bipartition used to define bipartite entanglement, and the characterization of the sub-sector $\mathcal{H}_{\text{II}}$ — are provided in Fig. 5.4. **(a)** Estimated performance of individual optimization runs at each iteration step and **(b)** Histogram showing improvement of detection probability after multiple optimization rounds. The horizontal (vertical) line in the upper (lower) panel indicates the fraction of detected states that are accessible before undertaking the optimization procedure. The estimates reported in this graph are based on direct evaluation of the expectation value of EWs along the optimization trajectory on 10^5 randomly chosen pure states belonging to $\mathcal{H}_{\text{II}}$.

5.5 Outlook

Despite in the framework of LGTs the structure of the Hilbert space does not naturally support the common definiton of separable states based on the existence of a decomposition in terms of a tensor product, it is viable to characterize separable states resorting to an operational approach based on the factorization of the expectation value of operators in the form of $\hat{O}_{\mathrm{A}} \otimes \hat{O}_{\mathrm{B}}$. In this regard, the operators $\hat{O}_{\mathrm{A}}$ and $\hat{O}_{\mathrm{B}}$ can in principle be either generic bounded operators supported, respectively, on A and B or be gauge-invariant. In this work, we proved that at least when Abelian LGTs are involved, the derived definition of separability remains unaffected by the choice of operators that are employed to test the separability. Crucially, in experiments dealing with systems that display Abelian local symmetries, the separability of a state can be fully accessed by resorting to operators that can be realized within the LGT. An interesting direction for future research would be to investigate whether this result also holds for non-Abelian LGTs.

In addition to this, in the present work, we extended an existing protocol to optimize EWs [73] for the case of LGTs, where physically meaningful operators are required to be gauge-invariant.

To conclude, we tested the optimization procedure in the case of a pure $U(1)$ LGT, finding that the numerical algorithm we developed is capable of autonomously selecting an optimization direction and using it to improve the EW until the optimality criterium is met.

PART IV

CONCLUSIONS

Conclusions

The first part of this thesis is devoted to approaching relevant soft-matter applications either by directly employing quantum solvers, or by connecting the problem under study with a framework that is amenable to quantum simulation.

It should be apparent that, despite this thesis is concerned with "quantum computing", only a fraction of its content is directly connected with the use of quantum platforms. Paradoxical as it may seem, this is what we should be expecting when approaching real-world problems disposing of quantum computers that have of a limited capability — whether in terms of the number of qubits, their quality, or gate's fidelity. Indeed, it is not currently possible to feed a real-world problem in its *full complexity* to quantum computers and perform the simulation in its *full extent*. One way to overcome this challenge is by integrating the use of quantum technologies with state-of-the-art classical algorithms. In this case, classical algorithms can assist in reducing the number of variables describing the real-world problem while preserving relevant physical properties, minimizing the number of gates required in gate-based quantum simulations, or generating quantum annealing schedules that suppress diabatic transitions.

Given the remarkable progress in machine learning, deep learning, and other heuristic techniques in recent years, these algorithms arguably serve as the most natural complements to quantum computing at its current stage of technological maturity.

Another critical area of research in quantum computing focuses on understanding the origins of the quantum advantage that is observed in various applications. A key feature that has been identified in this context is entanglement. Although extensively studied, many aspects of entanglement remain open for further exploration. One such aspect is addressed in the final chapter of this thesis, where we analyze the interplay between quantum entanglement and lattice gauge theories, in which the degrees of freedom are subject to constraints.

In conclusion, it is my belief that research in quantum computing has reached a level of maturity sufficient to fuel innovative solutions to practical, real-world problems. The current technological limitations have stimulated a broad range of original approaches, which will continue to hold value even as higher-quality quantum platforms emerge.

PART V

APPENDICES

Appendix A

Ancillarization of polynomial binary optimization problems

Let us suppose that the integer optimization problem involves an array, $\mathbf{x} = \{x_1, \ldots, x_n\}$, of binary variables taking values 0 or 1 that are subjected to a set of constraints $L_k(\mathbf{x}) = 0$ ($k = 1, \ldots, m$). Further, assume that $\mathcal{C}(\mathbf{x})$ is the objective function whose minimization is addressed. The solution to this binary optimization problem where binary arrays are explicitly subject to constraints can also be found by minimizing the functional

$$\mathcal{F}(\mathbf{x}) = \mathcal{C}(\mathbf{x}) + \sum_k \lambda_k L_k^2(\mathbf{x}) \,, \tag{A.1}$$

where λ_i are sufficiently large multipliers. Importantly, the binary arrays entering Eq. (A.1) are no longer explicitly constrained, as configurations violating the constraints are associated with a significant increase of the new cost function $\mathcal{F}(\mathbf{x})$. If the expression in Eq. (A.1) contains only linear and quadratic terms, namely if it allows a rewriting as

$$\mathcal{F}(\mathbf{x}) = \left(\sum_{\mathrm{i}} C_{\mathrm{i}}\, x_{\mathrm{i}} + \sum_{\mathrm{i} \neq \mathrm{j}} C_{\mathrm{ij}}\, x_{\mathrm{i}}\, x_{\mathrm{j}} \right) + \sum_k \left(\sum_{\mathrm{i}} \ell_{\mathrm{i}}^{(k)}\, \hat{x}_{\mathrm{i}} + \sum_{\mathrm{i} \neq \mathrm{j}} \ell_{\mathrm{ij}}^{(k)}\, x_{\mathrm{i}}\, x_{\mathrm{j}} \right) , \tag{A.2}$$

where C_{i}, C_{ij}, $\ell_{\mathrm{i}}^{(k)}$, and $\ell_{\mathrm{ij}}^{(k)}$ are real parameters featuring both the cost function and the constraints, then it is a QUBO problem whose solutions can be encoded in the ground-space of the Ising-like Hamiltonian given by

$$\hat{H}_P = \left(\sum_{\mathrm{i}} C_{\mathrm{i}}\, \hat{q}_{\mathrm{i}} + \sum_{\mathrm{i} \neq \mathrm{j}} C_{\mathrm{ij}}\, \hat{q}_{\mathrm{i}}\, \hat{q}_{\mathrm{j}} \right) + \sum_{k} \left(\sum_{\mathrm{i}} \ell_{\mathrm{i}}^{(k)}\, \hat{q}_{\mathrm{i}} + \sum_{\mathrm{i} \neq \mathrm{j}} \ell_{\mathrm{ij}}^{(k)}\, \hat{q}_{\mathrm{i}}\, \hat{q}_{\mathrm{j}} \right) , \tag{A.3}$$

where $\hat{q}_{\mathrm{i}} = \frac{\hat{\mathbb{1}} - \hat{\sigma}_{\mathrm{i}}^{(z)}}{2}$, making Eq. (A.3) an Ising-like Hamiltonian.

Until now, we have considered the specific case where both the objective function $\mathcal{C}(\mathbf{x})$ and the penalties $L_i^2(\mathbf{x})$ contain terms up to second order. In cases where this condition is not met, the problem set in (A.1) would be a polynomial unconstrained binary optimization problem (PUBO). In this case, a viable approach involves identifying quadratic approximations of $\mathcal{F}(\mathbf{x})$ that share the same minima. As we will see later, this operation — known in the literature as *ancillarization* — does not require *a priori* knowledge of the minima's position. However, this comes at the cost of introducing additional variables, known as *ancillary"*variables.

We will proceed by degrees of increasing difficulty, and, without loss of generality, we will start by considering the case of $\mathcal{F}^{(3)}(\mathbf{x}) = x_1\, x_2\, x_3$. The ancillarization consists in introducing an ancillary variable, e.g., a_{12}, and imposing that low-energy states satisfy $a_{12} = x_1\, x_2$. Clearly, we can not reproduce this by involving the penalty term $(a_{12} - x_1\, x_2)^2$, as it would contain a cubic term. To circumvent this, let us consider a function involving x_1, x_2, and the ancilla a_{12} such that (i) has degenerate minima when $a_{12} = x_1\, x_2$ and (ii) contains only linear and quadratic terms. Observing that $f(\cdot)$ is symmetric with respect to the exchange of x_1 and x_2, enables reducing the number of parameters characterizing $f(\cdot)$.

It turns out that the most general expression with these features is

$$f(x_1, x_2, a_{12}) = \alpha x_1\, x_2 + \beta a_{12}\, (x_1 + x_2) + \gamma a_{12} \,, \tag{A.4}$$

where the parameters simultaneously satisfy

$$\begin{cases} \alpha > 0 \\ \beta < -\alpha \\ \gamma = -\alpha - 2\beta \,. \end{cases} \tag{A.5}$$

These conditions trivially imply $\beta < 0$, $\gamma > 0$, $\alpha + \gamma > 0$, and $\beta + \gamma > 0$. All combinations of binary values and the respective values assumed by

x_1	x_2	a_{12}	$f(x_1, x_2, a_{12})$
0	0	0	0
0	1	0	0
1	0	0	0
1	1	1	0
0	0	1	γ
0	1	1	$\beta+\gamma$
1	0	1	$\beta+\gamma$
1	1	0	$\alpha+\gamma$

Table A.1: Choosing the parameters defining $f(\cdot)$ as in Eq. (A.5), the quadratic function yields a vanishing result when its arguments satisfy $a_{12} = x_1\, x_2$ and a positive value otherwise.

$f(\cdot)$ are reported in Table A.1. The entries of the first four rows satisfy $a_{12} = x_1\, x_2$ and correspond to vanishing values of $f(\cdot)$, while the remaining violate the constraint and thus are correctly penalized in energy.

With this intermediate passage, the solutions of the integer optimization problems set by

$$\mathcal{F}^{(3)}(\mathbf{x}) = x_1\, x_2\, x_3 \quad \text{and} \quad \mathcal{F}^{(2)}(\mathbf{x}, a_{12}) = a_{12}\, x_3 + \lambda_{a_{12}}\, f(x_1, x_2, a_{12})\,, \tag{A.6}$$

provided the multiplier $\lambda_{a_{12}}$ is large enough, coincide.

To extend this strategy to arbitrary p-order cost functions, $\mathcal{F}^{(p)} = x_1\, \ldots\, x_p$, we introduce a first ancilla featuring $a_{p-1,p} = x_{p-1}\, x_p$ effectively down-grading the initial problem to integer problem to

$$\mathcal{F}^{(p-1)} = x_1\, \ldots\, x_{p-2}\, a_{p-1,p} + \lambda_{p-1,p}\, f(x_{p-1}, x_p, a_{p,p-1})\,, \tag{A.7}$$

which can be understood as a $(p-1)$–order function involving $(p+1)$ binary variables.

With this, we conclude that a generic p-order cost function can be reduced to a quadratic function that involves the p original variables, plus $p-2$ ancillary ones. As seen in Section 1.2.2, the final step involves embedding the Ising-like Hamiltonian in Eq. (1.14) onto the quantum platform, which includes mapping logical qubits to physical qubits and setting up the appropriate local fields and couplings.

Appendix B

D-Wave superconducting quantum solver

In this appendix, we discuss technological details concerning the specific quantum platform we leveraged in tackling the protein design problem in Chapter 3.

The qubits that constitute the quantum chip provided by D-Wave Inc. are realized with superconducting rings having Josephson junctions suitably disposed along them. Depending on the disposition of the Josephson Junctions and the regimes in which they operate, it is possible to distinguish different kinds of emergent qubit architectures: charge qubits, phase qubits, and flux qubits. In the D-Wave platform, a particular implementation of flux qubits is adopted. That is the Composite Josephson junction radio frequency superconducting quantum interference device (CCJ rf-SQUID) [251,252], where the main superconducting loop is interrupted by a secondary one that has two Josephson junctions on it. The main advantage of CCJ rf-SQUIDs is that, by regulating the magnetic flux applied on the loop containing JJs, it is possible to fine-tune the properties of individual qubits to compensate for fabrication defects. Apart from this feature, their working principle is not different from regular rf-SQUIDs, which we briefly illustrate.

To understand the physics that makes rf-SQUIDs effectively a two-level system, let us consider the behavior of a Josephson junction, a thin layer of insulating material separating two superconducting channels. When a potential $V(t)$ is applied to the two ends of a Josephson Junction, Cooper pairs tunnel through it, resulting in an effective current $I(t)$ flowing from one superconductor to the other. The equations regulating this process are given by

$$\begin{cases} I(t) & = I_c \sin\psi(t) \\ \frac{\partial\psi(t)}{\partial t} & = \frac{2e}{\hbar} V(t)\,, \end{cases} \tag{B.1}$$

where $\psi(t)$ is the phase difference between the wavefunctions residing on the two sides of the insulating layer and I_c is its critical current. The energy shift associated with this element is $E(\psi) = -E_J \cos\psi$, where E_J is a characteristic parameter of the Josephson junction. Then, by inserting a Josephson junction in a superconducting loop that is equipped with a linear inductor L, the total energy reads

$$E(\psi) = E_L\left(\psi - \psi_e\right)^2 - E_J \cos\psi\,, \tag{B.2}$$

where ψ_e corresponds to the phase shift induced by an external bias flux addressed to the superconducting wire and $E_L = \frac{1}{2}\left(\frac{\Phi_0}{2\pi L}\right)^2$ characterizes the energy scale associated with the inductance. When Eq. (B.2) describes a bistable potential, the two local minima become degenerate at $\psi_e = \pi$.

Moreover, due to the non-harmonicity of Eq. (B.2), the associated energy levels are not equally spaced, making it possible to decouple the two lowest-lying states from the others, resulting in an effective two-level system defining a qubit.

To fix the reference frame, let us consider the situation in which the energy potential has two degenerate minima and let us define $|0\rangle$ and $|1\rangle$ the two states localized on individual minima, corresponding to clockwise and anticlockwise persistent currents flowing in the superconducting loop. In this situation, the groundstate is the balanced superposition given by $|g\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle + |1\rangle\right)$, which corresponds to the groundstate of the single-qubit gate $-\hat{\sigma}^{(x)}$. By tuning away the external magnetic potential from $\frac{\Phi_0}{2}$, it is possible to raise the energy of one of the two minima, guiding the ground-state of the system towards $|0\rangle$ or $|1\rangle$. As a matter of fact, the final state is the ground-state of the single qubit operator of $\mu\hat{\sigma}^{(z)}$, where the sign of μ is fully determined by the instantaneous external magnetic field applied to the qubit.

Until now, no coupling between qubits has been introduced. From a technological point of view, one possible way of coupling rf-SQUIDS is making superconducting loops of different qubits overlap, enhancing their inductive interaction. Despite the effectiveness of this method, it does not provide a switchable and tunable coupling between qubits. To overcome this limitation, the interaction between adjacent rf-SQUIDS is mediated by an additional superconducting loop whose activity is regulated by a dedicated external magnetic field [253, 254].

Overall, the magnetic fields regulating individual qubits and their couplings define a transverse-field Ising Hamiltonian taking the form

$$\hat{H}_{\text{platform}} = -\sum_{\mathrm{i}} \alpha_{\mathrm{i}}\, \hat{\sigma}_{\mathrm{i}}^{(x)} + \sum_{\mathrm{i}} \beta_{\mathrm{i}}\, \hat{\sigma}_{\mathrm{i}}^{(z)} + \sum_{(\mathrm{i},\,\mathrm{j})\in\mathcal{G}} \gamma_{\mathrm{ij}}\, \hat{\sigma}_{\mathrm{i}}^{(z)}\, \hat{\sigma}_{\mathrm{j}}^{(z)}\,. \tag{B.3}$$

Importantly, this platform can not reproduce *arbitrary* Hamiltonians and thus can not be used for universal quantum computation. We also stress that the couplings γ_{ij} can be established only between topologically close qubits, as direct inductive phenomena mediate their interaction. To explicit this, we will refer to $\mathcal{G} = (Q, C)$ as the non-directed graph representing the device, where $Q = \{q_{\mathrm{i}}\}_{\mathrm{i}}$ is the set of physical qubits and $C = \{(q_{\mathrm{i}}, q_{\mathrm{j}})\}_{\mathrm{i,j}}$ are the available connections among them [255–257].

List of publications

Panizza V., Costa de Almeida R., and Hauke P. “Entanglement witnessing for lattice gauge theories.” J. High Energ. Phys. 2022, 196 (2022). https://doi.org/10.1007/JHEP09(2022)196

Panizza V., Hauke P., Micheletti C., Faccioli P. “Protein Design by Integrating Machine Learning and Quantum-Encoded Optimization” PRX Life 2, 043012 (2024). https://doi.org/10.1103/PRXLife.2.043012

Panizza V., Roggero A., Hauke P., Faccioli P. “Statistical Mechanics of Heteropolymers from Lattice Gauge Theory” PRL 134, 158101 (2025). https://doi.org/10.1103/PhysRevLett.134.158101

Panizza V. Costa de Almeida R., Bera A., Hauke P. “Optimization of entanglement witnesses subjected to selection rules” (in preparation)

Bibliography

[1] P. Benioff, "The computer as a physical system: A microscopic quantum mechanical hamiltonian model of computers as represented by turing machines," *Journal of Statistical Physics*, vol. 22, no. 5, pp. 563–591, 2005.

[2] R. P. Feynman, "Simulating physics with computers," *International Journal of Theoretical Physics*, vol. 21, no. 6-7, pp. 467–488, 1982.

[3] D. Deutsch, "Quantum theory, the church–turing principle and the universal quantum computer," *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, vol. 400, pp. 117–97, 1985.

[4] D. P. DiVincenzo, "The physical implementation of quantum computation," *Fortschritte der Physik*, vol. 48, no. 9-11, pp. 771–783, 2000.

[5] P. Shor, "Algorithms for quantum computation: discrete logarithms and factoring," in *Proceedings 35th Annual Symposium on Foundations of Computer Science*, pp. 124–134, 1994.

[6] L. K. Grover, "A fast quantum mechanical algorithm for database search," in *Proceedings of the Twenty-Eighth Annual ACM Symposium on Theory of Computing*, STOC '96, (New York, NY, USA), p. 212–219, Association for Computing Machinery, 1996.

[7] A. W. Harrow and A. Montanaro, "Quantum computational supremacy," *Nature*, vol. 549, no. 7671, pp. 203–209, 2017.

[8] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. S. L. Brandao, D. A. Buell, B. Burkett, Y. Chen, Z. Chen, B. Chiaro, R. Collins, W. Courtney, A. Dunsworth, E. Farhi, B. Foxen, A. Fowler, C. Gidney, M. Giustina, R. Graff, K. Guerin, S. Habegger, M. P. Harrigan, M. J. Hartmann, A. Ho, M. Hoffmann, T. Huang, T. S. Humble, S. V. Isakov, E. Jeffrey, Z. Jiang, D. Kafri, K. Kechedzhi, J. Kelly, P. V. Klimov, S. Knysh, A. Korotkov, F. Kostritsa, D. Landhuis, M. Lindmark, E. Lucero, D. Lyakh, S. Mandrà, J. R. McClean, M. McEwen, A. Megrant, X. Mi, K. Michielsen, M. Mohseni, J. Mutus, O. Naaman, M. Neeley, C. Neill, M. Y. Niu, E. Ostby, A. Petukhov, J. C. Platt, C. Quintana, E. G. Rieffel, P. Roushan, N. C. Rubin, D. Sank, K. J. Satzinger, V. Smelyanskiy, K. J. Sung, M. D. Trevithick, A. Vainsencher, B. Villalonga, T. White, Z. J. Yao, P. Yeh, A. Zalcman, H. Neven, and J. M. Martinis, "Quantum supremacy using a programmable superconducting processor," *Nature*, vol. 574, no. 7779, pp. 505–510, 2019.

[9] F. Pan and P. Zhang, "Simulation of quantum circuits using the big-batch tensor network method," *Phys. Rev. Lett.*, vol. 128, p. 030501, jan 2022.

[10] H.-S. Zhong, Y.-H. Deng, J. Qin, H. Wang, M.-C. Chen, L.-C. Peng, Y.-H. Luo, D. Wu, S.-Q. Gong, H. Su, Y. Hu, P. Hu, X.-Y. Yang, W.-J. Zhang, H. Li, Y. Li, X. Jiang, L. Gan, G. Yang, L. You, Z. Wang, L. Li, N.-L. Liu, J. J. Renema, C.-Y. Lu, and J.-W. Pan, "Phase-programmable gaussian boson sampling using stimulated squeezed light," *Phys. Rev. Lett.*, vol. 127, p. 180502, oct 2021.

[11] N. Quesada, J. M. Arrazola, and N. Killoran, "Gaussian boson sampling using threshold detectors," *Phys. Rev. A*, vol. 98, p. 062322, Dec 2018.

[12] Y. Kim, A. Eddins, S. Anand, K. X. Wei, E. van den Berg, S. Rosenblatt, H. Nayfeh, Y. Wu, M. Zaletel, K. Temme, and A. Kandala, "Evidence for the utility of quantum computing before fault tolerance," *Nature*, vol. 618, no. 7965, pp. 500–505, 2023.

[13] J. Tindall, M. Fishman, E. M. Stoudenmire, and D. Sels, "Efficient tensor network simulation of ibm's eagle kicked ising experiment," *PRX Quantum*, vol. 5, p. 010308, Jan 2024.

[14] A. D. King, A. Nocera, M. M. Rams, J. Dziarmaga, R. Wiersema, W. Bernoudy, J. Raymond, N. Kaushal, N. Heinsdorf, R. Harris, K. Boothby, F. Altomare, A. J. Berkley, M. Boschnak, K. Chern, H. Christiani, S. Cibere, J. Connor, M. H. Dehn, R. Deshpande, S. Ejtemaee, P. Farré, K. Hamer, E. Hoskinson, S. Huang, M. W. Johnson, S. Kortas, E. Ladizinsky, T. Lai, T. Lanting, R. Li, A. J. R. MacDonald, G. Marsden, C. C. McGeoch, R. Molavi, R. Neufeld, M. Norouzpour, T. Oh, J. Pasvolsky, P. Poitras, G. Poulin-Lamarre, T. Prescott, M. Reis, C. Rich, M. Samani, B. Sheldan, A. Smirnov, E. Sterpka, B. T. Clavera, N. Tsai, M. Volkmann, A. Whiticar, J. D. Whittaker, W. Wilkinson, J. Yao, T. J. Yi, A. W. Sandvik, G. Alvarez, R. G. Melko, J. Carrasquilla, M. Franz, and M. H. Amin, "Computational supremacy in quantum simulation," mar 2024. arXiv:2403.00910 [cond-mat, physics:quant-ph].

[15] A. Y. Kitaev, "Quantum computations: algorithms and error correction," *Russian Mathematical Surveys*, vol. 52, no. 6, pp. 1191–1249, 1997.

[16] K. R. Brown, J. Kim, and C. Monroe, "Co-designing a scalable quantum computer with trapped atomic ions," *npj Quantum Information*, vol. 2, no. 1, 2016.

[17] V. M. Schäfer, C. J. Ballance, K. Thirumalai, L. J. Stephenson, T. G. Ballance, A. M. Steane, and D. M. Lucas, "Fast quantum logic gates with trapped-ion qubits," *Nature*, vol. 555, no. 7694, pp. 75–78, 2018.

[18] C. D. Bruzewicz, J. Chiaverini, R. McConnell, and J. M. Sage, "Trapped-ion quantum computing: Progress and challenges," *Applied Physics Reviews*, vol. 6, no. 2, 2019.

[19] X.-F. Shi, "Quantum logic and entanglement by neutral rydberg atoms: methods and fidelity," *Quantum Science and Technology*, vol. 7, p. 023002, apr 2022.

[20] A. K. Fedorov, N. Gisin, S. M. Beloussov, and A. I. Lvovsky, "Quantum computing at the quantum advantage threshold: a down-to-business review," 2022.

[21] J. M. Pino, J. M. Dreiling, C. Figgatt, J. P. Gaebler, S. A. Moses, M. S. Allman, C. H. Baldwin, M. Foss-Feig, D. Hayes,

K. Mayer, C. Ryan-Anderson, and B. Neyenhuis, “Demonstration of the trapped-ion quantum ccd computer architecture,” *Nature*, vol. 592, no. 7853, pp. 209–213, 2021.

[22] G. K. Brennen, C. M. Caves, P. S. Jessen, and I. H. Deutsch, “Quantum logic gates in optical lattices,” *Phys. Rev. Lett.*, vol. 82, pp. 1060–1063, Feb 1999.

[23] D. Jaksch, J. I. Cirac, P. Zoller, S. L. Rolston, R. Côté, and M. D. Lukin, “Fast quantum gates for neutral atoms,” *Phys. Rev. Lett.*, vol. 85, pp. 2208–2211, Sep 2000.

[24] J. E. Mooij, T. P. Orlando, L. Levitov, L. Tian, C. H. van der Wal, and S. Lloyd, “Josephson persistent-current qubit,” *Science*, vol. 285, no. 5430, pp. 1036–1039, 1999.

[25] Y. Nakamura, Y. A. Pashkin, and J. S. Tsai, “Coherent control of macroscopic quantum states in a single-cooper-pair box,” *Nature*, vol. 398, no. 6730, pp. 786–788, 1999.

[26] J. M. Martinis, S. Nam, J. Aumentado, and C. Urbina, “Rabi oscillations in a large josephson-junction qubit,” *Phys. Rev. Lett.*, vol. 89, p. 117901, Aug 2002.

[27] J. Koch, T. M. Yu, J. Gambetta, A. A. Houck, D. I. Schuster, J. Majer, A. Blais, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, “Charge-insensitive qubit design derived from the cooper pair box,” *Phys. Rev. A*, vol. 76, p. 042319, Oct 2007.

[28] R. Barends, J. Kelly, A. Megrant, D. Sank, E. Jeffrey, Y. Chen, Y. Yin, B. Chiaro, J. Mutus, C. Neill, P. O’Malley, P. Roushan, J. Wenner, T. C. White, A. N. Cleland, and J. M. Martinis, “Coherent josephson qubit suitable for scalable quantum integrated circuits,” *Phys. Rev. Lett.*, vol. 111, p. 080502, Aug 2013.

[29] L. Casparis, T. W. Larsen, M. S. Olsen, F. Kuemmeth, P. Krogstrup, J. Nygård, K. D. Petersson, and C. M. Marcus, “Gatemon benchmarking and two-qubit operations,” *Phys. Rev. Lett.*, vol. 116, p. 150505, Apr 2016.

[30] D. Loss and D. P. DiVincenzo, “Quantum computation with quantum dots,” *Phys. Rev. A*, vol. 57, pp. 120–126, Jan 1998.

[31] D. Aharonov, W. van Dam, J. Kempe, Z. Landau, S. Lloyd, and O. Regev, "Adiabatic quantum computation is equivalent to standard quantum computation," in *45th Annual IEEE Symposium on Foundations of Computer Science*, pp. 42–51, 2004.

[32] M. Born and V. Fock, "Beweis des adiabatensatzes," *Zeitschrift für Physik*, vol. 51, no. 3-4, pp. 165–180, 1928.

[33] J. D. Biamonte and P. J. Love, "Realizable hamiltonians for universal adiabatic quantum computers," *Phys. Rev. A*, vol. 78, p. 012352, Jul 2008.

[34] A. Lucas, "Ising formulations of many np problems," *Frontiers in Physics*, vol. 2, 2014.

[35] T. Albash and D. A. Lidar, "Adiabatic quantum computation," *Rev. Mod. Phys.*, vol. 90, p. 015002, Jan 2018.

[36] Y. Ding, J. Gonzalez-Conde, L. Lamata, J. D. Martín-Guerrero, E. Lizaso, S. Mugel, X. Chen, R. Orús, E. Solano, and M. Sanz, "Toward prediction of financial crashes with a d-wave quantum annealer," *Entropy*, vol. 25, no. 2, p. 323, 2023.

[37] S. Ebadi, A. Keesling, M. Cain, T. T. Wang, H. Levine, D. Bluvstein, G. Semeghini, A. Omran, J.-G. Liu, R. Samajdar, X.-Z. Luo, B. Nash, X. Gao, B. Barak, E. Farhi, S. Sachdev, N. Gemelke, L. Zhou, S. Choi, H. Pichler, S.-T. Wang, M. Greiner, V. Vuletić, and M. D. Lukin, "Quantum optimization of maximum independent set using rydberg atom arrays," *Science*, vol. 376, no. 6598, pp. 1209–1215, 2022.

[38] S. Jeong, M. Kim, M. Hhan, J. Park, and J. Ahn, "Quantum programming of the satisfiability problem with rydberg atom graphs," *Phys. Rev. Res.*, vol. 5, p. 043037, Oct 2023.

[39] L. F. Pérez Armas, S. Creemers, and S. Deleplanque, "Solving the resource constrained project scheduling problem with quantum annealing," *Scientific Reports*, vol. 14, no. 1, 2024.

[40] W. Bożejko, A. Burduk, J. Pempera, M. Uchroński, and M. Wodecki, "Optimal solving of a binary knapsack problem on a d-wave quantum machine and its implementation in production systems," *Annals of Operations Research*, 2024.

[41] R. Harris, Y. Sato, A. J. Berkley, M. Reis, F. Altomare, M. H. Amin, K. Boothby, P. Bunyk, C. Deng, C. Enderud, S. Huang, E. Hoskinson, M. W. Johnson, E. Ladizinsky, N. Ladizinsky, T. Lanting, R. Li, T. Medina, R. Molavi, R. Neufeld, T. Oh, I. Pavlov, I. Perminov, G. Poulin-Lamarre, C. Rich, A. Smirnov, L. Swenson, N. Tsai, M. Volkmann, J. Whittaker, and J. Yao, "Phase transitions in a programmable quantum spin glass simulator," *Science*, vol. 361, no. 6398, pp. 162–165, 2018.

[42] P. Kairys, A. D. King, I. Ozfidan, K. Boothby, J. Raymond, A. Banerjee, and T. S. Humble, "Simulating the shastry-sutherland ising model using quantum annealing," *PRX Quantum*, vol. 1, p. 020320, Dec 2020.

[43] A. D. King, S. Suzuki, J. Raymond, A. Zucca, T. Lanting, F. Altomare, A. J. Berkley, S. Ejtemaee, E. Hoskinson, S. Huang, E. Ladizinsky, A. J. R. MacDonald, G. Marsden, T. Oh, G. Poulin-Lamarre, M. Reis, C. Rich, Y. Sato, J. D. Whittaker, J. Yao, R. Harris, D. A. Lidar, H. Nishimori, and M. H. Amin, "Coherent quantum annealing in a programmable 2,000 qubit ising chain," *Nature Physics*, vol. 18, no. 11, pp. 1324–1328, 2022.

[44] D. Ghamari, P. Hauke, R. Covino, and P. Faccioli, "Sampling rare conformational transitions with a quantum computer," *Scientific Reports*, vol. 12, no. 1, 2022.

[45] H. Bhatia, E. Tretschk, C. Theobalt, and V. Golyanik, "Generation of truly random numbers on a quantum annealer," *IEEE Access*, vol. 10, pp. 112832–112844, 2022.

[46] M. D'Arcangelo, L.-P. Henry, L. Henriet, D. Loco, N. Gouraud, S. Angebault, J. Sueiro, J. Forêt, P. Monmarché, and J.-P. Piquemal, "Leveraging analog quantum computing with neutral atoms for solvent configuration prediction in drug discovery," *Phys. Rev. Res.*, vol. 6, p. 043020, Oct 2024.

[47] "D-wave qpu architecture: Topologies." `https://docs.dwavesys.com/docs/latest/c_gs_4.html`. Accessed: 2025-02-03.

[48] W. Lechner, P. Hauke, and P. Zoller, "A quantum annealing architecture with all-to-all connectivity from local interactions," *Sci. Adv.*, vol. 1, p. e1500838, oct 2015.

[49] E. Lobe and A. Lutz, "Minor embedding in broken chimera and derived graphs is np-complete," *Theoretical Computer Science*, vol. 989, p. 114369, 2024.

[50] E. Pelofske, "4-clique network minor embedding for quantum annealers," *Phys. Rev. Appl.*, vol. 21, p. 034023, Mar 2024.

[51] Y. Sugie, Y. Yoshida, N. Mertig, T. Takemoto, H. Teramoto, A. Nakamura, I. Takigawa, S.-i. Minato, M. Yamaoka, and T. Komatsuzaki, "Minor-embedding heuristics for large-scale annealing processors with sparse hardware graphs of up to 102,400 nodes," *Soft Computing*, vol. 25, no. 3, pp. 1731–1749, 2021.

[52] A. Bonomi, T. De Min, E. Zardini, E. Blanzieri, V. Cavecchia, and D. Pastorello, "Quantum annealing learning search implementations," *Quantum Information and Computation*, vol. 22, no. 3,4, pp. 181–208, 2022.

[53] J. Marshall, A. Di Gioacchino, and E. Rieffel, "Perils of embedding for sampling problems," *Physical Review Research*, vol. 2, p. 023020, apr 2020. arXiv:1909.12184 [cond-mat, physics:quant-ph].

[54] "Programming the d-wave qpu: Setting the chain strength." `https://www.dwavesys.com/media/vsufwv1d/14-1041a-a_setting_the_chain_strength.pdf`, 2020. Accessed: 2025-01-26.

[55] "What is quantum annealing?." `https://docs.dwavesys.com/docs/latest/c_gs_2.html`. Accessed: 2025-02-03.

[56] A. Einstein, B. Podolsky, and N. Rosen, "Can quantum-mechanical description of physical reality be considered complete?," *Phys. Rev.*, vol. 47, pp. 777–780, May 1935.

[57] J. S. Bell and A. Aspect, "Speakable and unspeakable in quantum mechanics," *undefined*, 1987.

[58] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, "Teleporting an unknown quantum state via dual classical and einstein-podolsky-rosen channels," *Phys. Rev. Lett.*, vol. 70, pp. 1895–1899, Mar 1993.

[59] D. Bouwmeester, J.-W. Pan, K. Mattle, M. Eibl, H. Weinfurter, and A. Zeilinger, "Experimental quantum teleportation," *Nature*, vol. 390, no. 6660, pp. 575–579, 1997.

[60] D. Boschi, S. Branca, F. De Martini, L. Hardy, and S. Popescu, "Experimental realization of teleporting an unknown pure quantum state via dual classical and einstein-podolsky-rosen channels," *Phys. Rev. Lett.*, vol. 80, pp. 1121–1125, Feb 1998.

[61] A. K. Ekert, "Quantum cryptography based on bell's theorem," *Phys. Rev. Lett.*, vol. 67, pp. 661–663, Aug 1991.

[62] C. H. Bennett and S. J. Wiesner, "Communication via one- and two-particle operators on einstein-podolsky-rosen states," *Physical Review Letters*, vol. 69, no. 20, pp. 2881–2884, 1992.

[63] H. Alexander Semenovich, "Bounds for the quantity of information transmitted by a quantum communication channel," *Problems of Information Transmission*, vol. 9, pp. 177–183, 1973.

[64] M. B. Plenio and S. S. Virmani, "An Introduction to Entanglement Theory," *Quant. Inf. Comput.*, vol. 7, no. 1-2, pp. 001–051, 2007.

[65] C. Eltschka and J. Siewert, "Quantifying entanglement resources," *Journal of Physics A: Mathematical and Theoretical*, vol. 47, no. 42, p. 424005, 2014.

[66] K. Vogel and H. Risken, "Determination of quasiprobability distributions in terms of probability distributions for the rotated quadrature phase," *Phys. Rev. A*, vol. 40, pp. 2847–2849, Sep 1989.

[67] V. N. Ivanova-Rohling and N. Rohling, "Optimal choice of state tomography quorum formed by projection operators," *Phys. Rev. A*, vol. 100, p. 032332, sep 2019. arXiv:1810.09484 [cond-mat, physics:quant-ph].

[68] M. Horodecki, P. Horodecki, and R. Horodecki, "Separability of mixed states: necessary and sufficient conditions," *Physics Letters A*, vol. 223, pp. 1–8, nov 1996.

[69] B. M. Terhal, "Bell inequalities and the separability criterion," *Physics Letters A*, vol. 271, pp. 319–326, jul 2000.

[70] O. Gühne and G. Tóth, "Entanglement detection," *Physics Reports*, vol. 474, no. 1-6, pp. 1–75, 2009.

[71] D. Chruściński and G. Sarbicki, "Entanglement witnesses: construction, analysis and classification," *J. Phys. A: Math. Theor.*, vol. 47, p. 483001, dec 2014. arXiv:1402.2413 [math-ph, physics:quant-ph].

[72] J. Bae, D. Chruściński, and B. C. Hiesmayr, "Mirrored entanglement witnesses," *npj Quantum Information*, vol. 6, no. 1, 2020.

[73] M. Lewenstein, B. Kraus, J. I. Cirac, and P. Horodecki, "Optimization of entanglement witnesses," *Phys. Rev. A*, vol. 62, p. 052310, oct 2000. arXiv:quant-ph/0005014.

[74] R. A. Bertlmann, K. Durstberger, B. C. Hiesmayr, and P. Krammer, "Optimal entanglement witnesses for qubits and qutrits," *Phys. Rev. A*, vol. 72, p. 052331, Nov 2005.

[75] T. Garel, H. Orland, and E. Pitard, "Protein folding and heteropolymers," *Series on Directions in Condensed Matter Physics*, p. 387–443, 1997.

[76] V. Pande, A. Grosberg, and T. Tanaka, "Statistical mechanics of simple models of protein folding and design," *Biophysical Journal*, vol. 73, p. 3192–3210, dec 1997.

[77] S. Kmiecik, D. Gront, M. Kolinski, L. Wieteska, A. E. Dawid, and A. Kolinski, "Coarse-grained protein models and their applications," *Chemical Reviews*, vol. 116, no. 14, pp. 7898–7936, 2016.

[78] T. Zhang, P. H. Nguyen, J. Nasica-Labouze, Y. Mu, and P. Derreumaux, "Folding atomistic proteins in explicit solvent using simulated tempering," *The Journal of Physical Chemistry B*, vol. 119, no. 23, pp. 6941–6951, 2015.

[79] W. Zheng, G. L. Dignon, X. Xu, R. M. Regy, N. L. Fawzi, Y. C. Kim, R. B. Best, and J. Mittal, "Molecular details of protein condensates probed by microsecond-long atomistic simulations," *undefined*, 2020.

[80] S. Jalali, Y. Yang, F. Mahmoudinobar, S. M. Singh, B. L. Nilsson, and C. Dias, "Using all-atom simulations in explicit solvent to study aggregation of amphipathic peptides into amyloid-like fibrils," *Journal of Molecular Liquids*, vol. 347, p. 118283, 2022.

[81] I. Coluzza, “A coarse-grained approach to protein design: Learning from design to understand folding,” *PLoS ONE*, vol. 6, p. e20853, jul 2011.

[82] I. Coluzza, P. D. J. Van Oostrum, B. Capone, E. Reimhult, and C. Dellago, “Design and folding of colloidal patchy polymers,” *Soft Matter*, vol. 9, no. 3, pp. 938–944, 2013.

[83] C. Cardelli, V. Bianco, L. Rovigatti, F. Nerattini, L. Tubiana, C. Dellago, and I. Coluzza, “The role of directional interactions in the designability of generalized heteropolymers,” *Scientific Reports*, vol. 7, p. 4986, jul 2017.

[84] K. Yue, K. Fiebig, P. Thomas, H. Chan, E. Shakhnovich, and K. Dill, “A test of lattice protein folding algorithms.,” *Proceedings of the National Academy of Sciences*, vol. 92, no. 1, p. 325–329, 1995.

[85] C. Micheletti, F. Seno, A. Maritan, and J. Banavar, “Protein design in a lattice model of hydrophobic and polar amino acids,” *Physical review letters*, vol. 80, no. 10, p. 2237, 1998.

[86] J. M. Polson and M. J. Zuckermann, “Simulation of heteropolymer collapse with an explicit solvent in two dimensions,” *The Journal of Chemical Physics*, vol. 113, no. 3, pp. 1283–1293, 2000.

[87] A. Perdomo, C. Truncik, I. Tubert-Brohman, G. Rose, and A. Aspuru-Guzik, “Construction of model hamiltonians for adiabatic quantum computation and its application to finding low-energy conformations of lattice protein models,” *Physical Review A*, vol. 78, no. 1, 2008.

[88] A. Perdomo-Ortiz, N. Dickson, M. Drew-Brook, G. Rose, and A. Aspuru-Guzik, “Finding low-energy conformations of lattice protein models by quantum annealing,” *Scientific Reports*, vol. 2, no. 1, 2012.

[89] M. Ejtehadi, N. Hamedani, H. Seyed-Allaei, V. Shahrezaei, and M. Yahyanejad, “Stability of preferable structures for a hydrophobic-polar model of protein folding,” *Physical Review E*, vol. 57, no. 3, p. 3298–3301, 1998.

[90] S. Miyazawa and R. L. Jernigan, “Estimation of effective interresidue contact energies from protein crystal structures: quasi-chemical approximation,” *Macromolecules*, vol. 18, pp. 534–552, mar 1985.

[91] J. Skolnick, A. Godzik, L. Jaroszewski, and A. Kolinski, “Derivation and testing of pair potentials for protein folding. when is the quasichemical approximation correct?,” *Protein science*, vol. 6, no. 3, pp. 676–688, 1997.

[92] R. Dima, G. Settanni, C. Micheletti, J. Banavar, and A. Maritan, “Extraction of interaction potentials between amino acids from native protein structures,” *The Journal of Chemical Physics*, vol. 112, no. 20, p. 9151–9166, 2000.

[93] H. Li, C. Tang, and N. Wingreen, “Designability of protein structures: A lattice-model study using the miyazawa-jernigan matrix,” *Proteins: Structure, Function, and Genetics*, vol. 49, no. 3, p. 403–412, 2002.

[94] G. Tiana, M. Colombo, D. Provasi, and R. Broglia, “Deriving amino acid contact potentials from their frequencies of occurrence in proteins: a lattice model study,” *Journal of Physics: Condensed Matter*, vol. 16, no. 15, p. 2551, 2004.

[95] W. Chen and E. Shakhnovich, “Lessons from the design of a novel atomic potential for protein folding,” *Protein science*, vol. 14, no. 7, pp. 1741–1752, 2005.

[96] R. Goldstein, “Amino-acid interactions in psychrophiles, mesophiles, thermophiles, and hyperthermophiles: insights from the quasi-chemical approximation,” *Protein science*, vol. 16, no. 9, pp. 1887–1895, 2007.

[97] A. Robert, P. Barkoutsos, S. Woerner, and I. T., “Resource-efficient quantum algorithm for protein folding,” *NPJ Quantum Information*, vol. 7, no. 1, 2021.

[98] R. Wong and W.-L. Chang, “Fast quantum algorithm for protein structure prediction in hydrophobic-hydrophilic model,” *Journal of Parallel and Distributed Computing*, vol. 164, p. 178–190, 2022.

[99] A. Irbäck, L. Knuthson, S. Mohanty, and C. Peterson, “Folding lattice proteins with quantum annealing,” *Phys. Rev. Research*, vol. 4, p. 043013, oct 2022. arXiv:2205.06084 [cond-mat, physics:physics, physics:quant-ph, q-bio].

[100] C. Levinthal, “Are there pathways for protein folding?,” *Journal de chimie physique*, vol. 65, pp. 44–45, 1968.

[101] C. Levinthal, “How to fold graciously,” in *Mössbaun Spectroscopy in Biological Systems Proceedings*, (Urbana, Illinois), pp. 22–24, 1969.

[102] Sep 2024.

[103] C. Anfinsen, “Principles that govern the folding of protein chains,” *Science*, vol. 181, no. 4096, p. 223–230, 1973.

[104] C. Anfinsen and H. Scheraga, “Experimental and theoretical aspects of protein folding,” *Advances in Protein Chemistry*, vol. 29, pp. 205–299, 1975.

[105] A. Šali, E. Shakhnovich, and M. Karplus, “How does a protein fold?,” *Nature*, vol. 369, no. 6477, p. 248–251, 1994.

[106] J. D. Bryngelson, J. N. Onuchic, N. D. Socci, and P. G. Wolynes, “Funnels, pathways, and the energy landscape of protein folding: A synthesis,” *Proteins*, vol. 21, pp. 167–195, mar 1995.

[107] E. I. Shakhnovich and A. M. Gutin, “Engineering of Stable and Fast-Folding Sequences of Model Proteins,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 90, no. 15, pp. 7195–7199, 1993.

[108] M. Peto, A. Kloczkowski, V. Honavar, and R. L. Jernigan, “Use of machine learning algorithms to classify binary protein sequences as highly-designable or poorly-designable,” *BMC Bioinformatics*, vol. 9, p. 487, dec 2008.

[109] H. Li, R. Helling, C. Tang, and N. Wingreen, “Emergence of Preferred Structures in a Simple Model of Protein Folding,” *Science, New Series*, vol. 273, pp. 666–669, aug 1996.

[110] V. Shahrezaei and M. Ejtehadi, “Geometry selects highly designable structures,” *The Journal of Chemical Physics*, vol. 113, p. 6437–6442, oct 2000.

[111] F. Seno, M. Vendruscolo, A. Maritan, and J. Banavar, "Optimal Protein Design Procedure," *Physical Review Letters*, vol. 77, pp. 1901–1904, aug 1996.

[112] M. Peto, A. Kloczkowski, and R. L. Jernigan, "Shape-dependent designability studies of lattice proteins," *J. Phys.: Condens. Matter*, vol. 19, p. 285220, jul 2007.

[113] B. Kuhlman, G. Dantas, G. C. Ireton, G. Varani, B. L. Stoddard, and D. Baker, "Design of a novel globular protein fold with atomic-level accuracy," *Science*, vol. 302, no. 5649, pp. 1364–1368, 2003.

[114] K. Yue, K. M. Fiebig, P. D. Thomas, H. S. Chan, E. I. Shakhnovich, and K. A. Dill, "A Test of Lattice Protein Folding Algorithms," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 92, no. 1, pp. 325–329, 1995.

[115] J. Deutsch and T. Kurosky, "New algorithm for protein design," *Phys. Rev. Lett.*, vol. 76, pp. 323–326, jan 1996.

[116] A. Irbäck, C. Peterson, F. Potthast, and E. Sandelin, "Monte Carlo procedure for protein design," *Physical Review E*, vol. 58, pp. R5249–R5252, nov 1998.

[117] E. Shakhnovich, "Protein design: a perspective from simple tractable models," *Folding and Design*, vol. 3, pp. R45–R58, jun 1998.

[118] P. Huang, S. Boyken, and D. Baker, "The coming of age of de novo protein design," *Nature*, vol. 537, pp. 320–327, sep 2016.

[119] I. Korendovych and W. DeGrado, "*De novo* protein design, a retrospective," *Quarterly Reviews of Biophysics*, vol. 53, p. e3, 2020.

[120] A. Paladino, F. Marchetti, S. Rinaldi, and G. Colombo, "Protein design: From computer models to artificial intelligence," *WIREs Computational Molecular Science*, vol. 7, no. 5, 2017.

[121] C. Goverde, B. Wolf, H. Khakzad, S. Rosset, and B. Correia, "De novo protein design by inversion of the alphafold structure prediction network," *Protein Sci*, 2022.

[122] M. Jendrusch, J. O. Korbel, and S. K. Sadiq, "AlphaDesign: A *de novo* protein design framework based on AlphaFold," preprint, Bioinformatics, oct 2021.

[123] T. Takahashi, G. Chikenji, and K. Tokita, “Lattice protein design using bayesian learning,” *Physical Review E*, vol. 104, no. 1, 2021.

[124] I. Anishchenko, S. Pellock, T. Chidyausiku, T. Ramelot, S. Ovchinnikov, J. Hao, K. Bafna, C. Norn, A. Kang, A. Bera, F. DiMaio, L. Carter, C. Chow, G. Montelione, and D. Baker, “De novo protein design by deep network hallucination,” *Nature*, vol. 600, pp. 547–552, dec 2021.

[125] J. Watson, D. Juergens, N. Bennett, B. Trippe, J. Yim, H. Eisenach, W. Ahern, A. Borst, R. Ragotte, L. Milles, B. Wicky, N. Hanikel, S. Pellock, A. Courbet, W. Sheffler, J. Wang, P. Venkatesh, I. Sappington, S. Torres, A. Lauko, V. De Bortoli, E. Mathieu, S. Ovchinnikov, R. Barzilay, T. Jaakkola, F. DiMaio, M. Baek, and D. Baker, “De novo design of protein structure and function with rfdiffusion,” *Nature*, vol. 620, pp. 1089–1100, aug 2023.

[126] J. Ingraham, M. Baranov, Z. Costello, K. Barber, W. Wang, A. Ismail, V. Frappier, D. Lord, C. Ng-Thow-Hing, E. Van Vlack, S. Tie, V. Xue, S. Cowles, A. Leung, J. Rodrigues, C. Morales-Perez, A. Ayoub, R. Green, K. Puentes, F. Oplinger, N. Panwar, F. Obermeyer, A. Root, A. Beam, F. Poelwijk, and G. Grigoryan, “Illuminating protein space with a programmable generative model,” *Nature*, vol. 623, pp. 1070–1078, nov 2023.

[127] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, *et al.*, “Highly accurate protein structure prediction with alphafold,” *Nature*, vol. 596, p. 583–589, jul 2021.

[128] S.-J. Chen, M. Hassan, R. Jernigan, K. Jia, D. Kihara, A. Kloczkowski, S. Kotelnikov, D. Kozakov, J. Liang, A. Liwo, S. Matysiak, J. Meller, C. Micheletti, J. Mitchell, S. Mondal, R. Nussinov, K. Okazaki, D. Padhorny, J. Skolnick, T. Sosnick, G. Stan, I. Vakser, X. Zou, and G. Rose, “Protein folds vs. protein folding: Differing questions, different challenges,” *Proceedings of the National Academy of Sciences*, vol. 120, no. 1, p. e2214423119, 2023.

[129] M. H. Khatami, U. C. Mendes, N. Wiebe, and P. M. Kim, “Gate-based quantum computing for protein design,” *PLoS Comput Biol*, vol. 19, p. e1011033, apr 2023.

[130] C. Micheletti, F. Seno, and A. Maritan, “Recurrent oligomers in proteins: an optimal scheme reconciling accurate and concise backbone representations in automated folding and design studies,” *Proteins: Structure, Function, and Bioinformatics*, vol. 40, no. 4, pp. 662–674, 2000.

[131] R. Kolodny, P. Koehl, L. Guibas, and M. Levitt, “Small libraries of protein fragments model native protein structures accurately,” *Journal of molecular biology*, vol. 323, no. 2, pp. 297–307, 2002.

[132] A.-C. Camproux, R. Gautier, and P. Tuffery, “A hidden markov model derived structural alphabet for proteins,” *Journal of molecular biology*, vol. 339, no. 3, pp. 591–605, 2004.

[133] A. Pandini, A. Fornili, and J. Kleinjung, “Structural alphabets derived from attractors in conformational space,” *BMC bioinformatics*, vol. 11, no. 1, pp. 1–18, 2010.

[134] C. Mackenzie, J. Zhou, and G. Grigoryan, “Tertiary alphabet for the observable protein structural universe,” *Proceedings of the National Academy of Sciences*, vol. 113, no. 47, pp. E7438–E7447, 2016.

[135] P. Krupa, A. Hałabis, W. Zmudzinska, S. Ołdziej, H. Scheraga, and A. Liwo, “Maximum likelihood calibration of the unres force field for simulation of protein structure and dynamics,” *Journal of chemical information and modeling*, vol. 57, no. 9, pp. 2364–2377, 2017.

[136] R. Babbush, P. Love, and A. Aspuru-Guzik, “Adiabatic quantum simulation of quantum chemistry,” *Scientific Reports*, vol. 4, no. 1, 2014.

[137] P. Hauke, G. Mattiotti, and P. Faccioli, “Dominant reaction pathways by quantum computing,” *Physical Review Letters*, vol. 126, no. 2, 2021.

[138] C. Micheletti, P. Hauke, and P. Faccioli, “Polymer physics by quantum computing,” *Physical Review Letters*, vol. 127, p. 080501, aug 2021.

[139] F. Slongo, P. Hauke, P. Faccioli, and C. Micheletti, “Quantum-inspired encoding enhances stochastic sampling of soft matter systems,” *Science Advances*, vol. 9, no. 43, p. eadi0204, 2023.

[140] K. Lau and K. Dill, “A lattice statistical mechanics model of the conformational and sequence spaces of proteins,” *Macromolecules*, vol. 22, no. 10, pp. 3986–3997, 1989.

[141] A. Irbäck, L. Knuthson, and C. Mohanty, S.and Peterson, “Using quantum annealing to design lattice proteins,” *Physical Review Research*, vol. 6, p. 013162, feb 2024. arXiv:2402.09069 [cond-mat, physics:physics, physics:quant-ph, q-bio].

[142] S. Edwards, “The statistical mechanics of polymers with excluded volume,” *Proc. Phys. Soc.*, vol. 85, p. 613, 1965.

[143] P. DeGennes, “Scaling concepts in polymer physics, ithaca, ny: Cornell u,” 1979.

[144] N. Rom, D. Charutz, and D. Neuhauser, “Shifted-contour auxiliary-field monte carlo: circumventing the sign difficulty for electronic-structure calculations,” *Chemical Physics Letters*, vol. 270, no. 3-4, pp. 382–386, 1997.

[145] A. P. Lund, M. J. Bremner, and T. C. Ralph, “Quantum sampling problems, bosonsampling and quantum supremacy,” *npj Quantum Information*, vol. 3, no. 1, 2017.

[146] V. Maiorov and G. Crippen, “Contact potential that recognizes the correct folding of globular proteins,” *Journal of molecular biology*, vol. 227, no. 3, pp. 876–888, 1992.

[147] G. Crippen, “Easily searched protein folding potentials,” *Journal of molecular biology*, vol. 260, no. 3, pp. 467–475, 1996.

[148] F. Seno, C. Micheletti, A. Maritan, and J. R. Banavar, “Variational approach to protein design and extraction of interaction potentials,” *Physical Review Letters*, vol. 81, p. 2172–2175, sep 1998. arXiv:cond-mat/9804054.

[149] A. Rossi, C. Micheletti, F. Seno, and A. Maritan, “A self-consistent knowledge-based approach to protein design,” *Biophysical Journal*, vol. 80, no. 1, p. 480–490, 2001.

[150] P. Hauke, H. Katzgraber, W. Lechner, H. Nishimori, and W. Oliver, “Perspectives of quantum annealing: Methods and implementations,” *Reports on Progress in Physics*, vol. 83, no. 5, p. 054401, 2020.

[151] C. Chang, C.-C. Chen, C. Koerber, T. Humble, and J. Ostrowski, "Integer programming from quantum annealing and open quantum systems," sep 2020.

[152] J.-R. Jiang and C.-W. Chu, "Solving np-hard problems with quantum annealing," *2022 IEEE 4th Eurasia Conference on IOT, Communication and Engineering (ECICE)*, oct 2022.

[153] R. Au-Yeung, N. Chancellor, and P. Halffmann, "Np-hard but no longer hard to solve? using quantum computing to tackle optimization problems," *Frontiers in Quantum Science and Technology*, vol. 2, feb 2023.

[154] J. Raymond, R. Stevanovic, W. Bernoudy, K. Boothby, C. McGeoch, A. Berkley, P. Farré, J. Pasvolsky, and A. King, "Hybrid quantum annealing for larger-than-qpu lattice-structured problems," *ACM Transactions on Quantum Computing*, vol. 4, p. 1–30, apr 2023.

[155] M. Perry and R. Wagner, "Python module for simulated annealing," 2019.

[156] W. Krauth and M. Mezard, "Learning algorithms with optimal stability in neural networks," *Journal of Physics A: Mathematical and General*, vol. 20, no. 11, 1987.

[157] D. Kleinbaum, *Logistic Regression. Statistics in the Health sciences*, ch. Introduction to logistic regression, p. 1–38. Springer, New York, NY, 1994.

[158] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, pp. 273–297, sep 1995.

[159] R. Johnson and A. Mouhab, "A bayesian decision theory approach to classification problems," *Journal of Multivariate Analysis*, vol. 56, p. 232–244, feb 1996.

[160] J. Imhoff, "A polynomial training algorithm for calculating perceptrons of optimal stability," *Journal of Physics A: Mathematical and General*, vol. 28, p. 2173–2181, apr 1995.

[161] Y. Freund and R. Schapire, "Large margin classification using the perceptron algorithm," *Proceedings of the eleventh annual conference on Computational learning theory*, jul 1998.

[162] K.-L. Du, C.-S. Leung, W. Mow, and M. Swamy, “Perceptron: Learning, generalization, model selection, fault tolerance, and role in the deep learning era,” *Mathematics*, vol. 10, p. 4730, dec 2022.

[163] A. Liwo, J. Lee, D. Ripoll, and H. Scheraga, “Protein structure prediction by global optimization of a potential energy function,” *Proc. Natl. Acad. Sci. U.S.A.*, vol. 96, no. 10, pp. 5482–5485, 1999.

[164] Odziej *et al.*, “Physics-based protein-structure prediction using a hierarchical protocol based on the unres force field: Assessment in two blind tests.,” *Proc. Natl. Acad. Sci. U.S.A.*, vol. 102, p. 7547–7552, 2005.

[165] T. Pochart, P. Jacquot, and J. Mikael, “On the challenges of using d-wave computers to sample boltzmann random variables,” *2022 IEEE 19th International Conference on Software Architecture Companion (ICSA-C)*, 2022.

[166] K. Kremer and K. Binder, “Monte carlo simulation of lattice models for macromolecules,” *Computer Physics Reports*, vol. 7, p. 259, 1988.

[167] N. Madras and A. D. Sokal, “The pivot algorithm: A highly efficient monte carlo method for the self-avoiding walk,” *J. Stat. Phys.*, vol. 50, p. 109, 1988.

[168] N. Clisby, “Accurate estimate of the critical exponent for self-avoiding walks via a fast implementation of the pivot algorithm,” *Phys. Rev. Lett.*, vol. 104, p. 055702, 2010.

[169] H. Hsiao-Ping, V. Mehra, W. Nadler, and P. Grassberger, “Growth-based optimization algorithm for lattice heteropolymers,” *Physical Review*, vol. E 68, p. 021113, 2003.

[170] K. Yue, K. M. Fiebig, P. D. Thomas, H. S. Chan, E. I. Shakhnovich, and K. A. Dill, “Growth-based optimization algorithm for lattice heteropolymers,” *Proc. Natl. Acad. Sci. U.S.A.*, vol. 92, p. 146, 1995.

[171] R. Ramakrishan and J. Pekny, J.F.and Caruthers, “A combinatorial algorithm for effective generation of long maximally compact lattice chains,” *J. Chem. Physics*, vol. 103, p. 7592, 1995.

[172] R. Lua, A. L. Borovinskiy, and A. Y. Grosberg, "Fractal and statistical properties of large compact polymers: a computational study," *Polymer*, vol. 45, p. 717, 2004.

[173] M. K. Mansfield, "Unbiased sampling of lattice hamilton path ensembles," *J. Chem. Phys.*, vol. 125, p. 154103, 2006.

[174] M. Ubertini and A. Rosa, "Computer simulations of melts of ring polymers with nonconserved topology: A dynamic monte carlo lattice model," *Phys. Rev*, vol. E 104, p. 054503, 2021.

[175] A. Robert, P. Barkoutsos, S. Woerner, and I. Tavernelli, "Resource-efficient quantum algorithm for pro- tein folding," *NPJ Quantum Information*, vol. 7, p. 38, 2021.

[176] A. D. N. Perdomo-Ortiz, M. Drew-Brook, G. Rose, and A. Aspuru-Guzik, "Finding low-energy conformations of lattice protein models by quantum annealing," *Scientific Reports*, vol. 2, p. 571, 2012.

[177] F. Slongo, P. Hauke, P. Faccioli, and C. Micheletti, "Quantum-inspired encoding enhances stochastic sampling of soft matter systems," *Science Advances*, vol. 9, no. 43, 2023.

[178] J. B. Kogut, "An introduction to lattice gauge theory and spin systems," *Reviews of Modern Physics*, vol. 51, pp. 659–713, oct 1979.

[179] R. M. Gray, "Toeplitz and circulant matrices: A review," *Foundations and Trends in Communications and Information Theory*, vol. 2, no. 3, pp. 155–239, 2006.

[180] "Eigenvalues and eigenvectors of a circulant matrix," jun 2017.

[181] K. M. Abadir and J. R. Magnus, *Matrix Algebra.* Cambridge University Press, aug 2005.

[182] K. Yue and K. A. Dill, "Sequence-structure relationships in proteins and copolymers," *Physical Review E*, vol. 48, no. 3, pp. 2267–2278, 1993.

[183] G. H. Fredrickson and K. F. Middleton, "Self-consistent field theory and its applications in polymer science," *Annual Review of Physical Chemistry*, vol. 57, pp. 585–603, 2006.

[184] T. A. Vilgis, "Polymer theory: Path integrals and scaling," *Physics Reports*, vol. 336, pp. 167–254, 2000.

[185] G. Fredrickson, "The equilibrium theory of inhomogeneous polymers," *Oxford University Press*, 2005.

[186] H. Orland and A. Zee, "Rna folding and large n matrix theory," *Nuclear Physics B*, vol. 620, no. 3, pp. 456–476, 2002.

[187] X. Man, K. T. Delaney, M. C. Villet, H. Orland, and G. H. Fredrickson, "Coherent states formulation of polymer field theory," *Journal of Chemical Physics*, vol. 140, p. 024905, 2014.

[188] E. M. Lennon, G. O. Mohler, H. D. Ceniceros, C. J. García-Cervera, and G. H. Fredrickson, "Numerical solutions of the complex langevin equations in polymer field theory," *Multiscale Modeling & Simulation*, vol. 6, no. 4, pp. 1347–1370, 2008.

[189] T. Li, K. Fan, W. Jun, and W. Wei, "Reduction of protein sequence complexity by residue grouping," *Protein Engineering*, vol. 16, pp. 323–330, 2003.

[190] S. Kamtekar, J. Schiffer, H. Xiong, J. Babik, and M. Hecht, "Protein design by binary patterning of polar and nonpolar amino acids," *Science*, vol. 262, pp. 1680–5, 1993.

[191] M. Bañuls, R. Blatt, J. Catani, *et al.*, "Simulating lattice gauge theories within quantum technologies," *Eur. Phys. J. D*, vol. 74, p. 165, 2020.

[192] M. Dalmonte and S. Montagnero, "Lattice gauge theory simulations in the quantum information era," *Contemp. Phys*, vol. 57, p. 388, 2016.

[193] C. W. Bauer, Z. Davoudi, A. B. Balantekin, T. Bhattacharya, M. Carena, W. A. de Jong, P. Draper, A. El-Khadra, N. Gemelke, M. Hanada, D. Kharzeev, H. Lamm, Y.-Y. Li, J. Liu, M. Lukin, Y. Meurice, C. Monroe, B. Nachman, G. Pagano, J. Preskill, E. Rinaldi, A. Roggero, D. I. Santiago, M. J. Savage, I. Siddiqi, G. Siopsis, D. Van Zanten, N. Wiebe, Y. Yamauchi, K. Yeter-Aydeniz, and S. Zorzetti, "Quantum simulation for high-energy physics," *PRX Quantum*, vol. 4, p. 027001, May 2023.

[194] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, “Quantum entanglement,” *Rev. Mod. Phys.*, vol. 81, pp. 865–942, Jun 2009.

[195] A. Harrow, P. Hayden, and D. Leung, “Superdense coding of quantum states,” *Phys. Rev. Lett.*, vol. 92, p. 187901, May 2004.

[196] J. Yin, Y.-H. Li, S.-K. Liao, M. Yang, Y. Cao, L. Zhang, J.-G. Ren, W.-Q. Cai, W.-Y. Liu, S.-L. Li, R. Shu, Y.-M. Huang, L. Deng, L. Li, Q. Zhang, N.-L. Liu, Y.-A. Chen, C.-Y. Lu, X.-B. Wang, F. Xu, J.-Y. Wang, C.-Z. Peng, A. K. Ekert, and J.-W. Pan, “Entanglement-based secure quantum cryptography over 1,120 kilometres,” *Nature*, vol. 582, no. 7813, pp. 501–505, 2020.

[197] J. v. Matteo Paris, *Quantum State Estimation.* Springer Berlin, Heidelberg, 2004.

[198] S. Kuzmin, V. Mikhailova, I. Dyakonov, and S. Straupe, “Learning the tensor network model of a quantum state using a few single-qubit measurements,” *Phys. Rev. A*, vol. 109, p. 052616, May 2024.

[199] B. M. Terhal, “Bell inequalities and the separability criterion,” *Physics Letters A*, vol. 271, no. 5, pp. 319–326, 2000.

[200] M. Bourennane, M. Eibl, C. Kurtsiefer, S. Gaertner, H. Weinfurter, O. Grühne, P. Hyllus, D. Bruß, M. Lewenstein, and A. Sanpera, “Experimental detection of multipartite entanglement using witness operators,” *Physical Review Letters*, vol. 92, no. 8, 2004.

[201] N. Friis, G. Vitagliano, M. Malik, and M. Huber, “Entanglement certification from theory to experiment,” *Nature Reviews Physics*, vol. 1, no. 1, pp. 72–87, 2018.

[202] L. Amico, R. Fazio, A. Osterloh, and V. Vedral, “Entanglement in many-body systems,” *Reviews of Modern Physics*, vol. 80, no. 2, pp. 517–576, 2008.

[203] S. Azzini, S. Mazzucchi, V. Moretti, D. Pastorello, and L. Pavesi, “Single-particle entanglement,” *Advanced Quantum Technologies*, vol. 3, no. 10, p. 2000014, 2020.

[204] O. Baker and D. Kharzeev, "Thermal radiation and entanglement in proton-proton collisions at energies available at the cern large hadron collider," *Physical Review D*, vol. 98, no. 5, 2018.

[205] F. Alet and N. Laflorencie, "Many-body localization: An introduction and selected topics," *Comptes Rendus Physique*, vol. 19, no. 6, pp. 498–525, 2018.

[206] L. Pezzè, A. Smerzi, M. K. Oberthaler, R. Schmied, and P. Treutlein, "Quantum metrology with nonclassical states of atomic ensembles," *Rev. Mod. Phys.*, vol. 90, p. 035005, Sep 2018.

[207] D. A. Abanin, E. Altman, I. Bloch, and M. Serbyn, "Colloquium: Many-body localization, thermalization, and entanglement," *Reviews of Modern Physics*, vol. 91, no. 2, 2019.

[208] X.-G. Wen, "Choreographed entanglement dances: Topological states of quantum matter," *Science*, vol. 363, no. 6429, 2019.

[209] T. Hashizume, J. C. Halimeh, P. Hauke, and D. Banerjee, "Ground-state phase diagram of quantum link electrodynamics in $(2+1)$-d," *SciPost Phys.*, vol. 13, p. 017, 2022.

[210] R. Bellwied, "Quantum entanglement in the initial and final state of relativistic heavy ion collisions," *Journal of Physics: Conference Series*, vol. 1070, p. 012001, 2018.

[211] X. Feal, C. Pajares, and R. A. Vazquez, "Thermal behavior and entanglement in pb–pb and p–p collisions," *Physical Review C*, vol. 99, no. 1, 2019.

[212] U.-J. Wiese, "Ultracold quantum gases and lattice systems: quantum simulation of lattice gauge theories," *Annalen der Physik*, vol. 525, pp. 777–796, jul 2013.

[213] E. A. Martinez, C. A. Muschik, P. Schindler, D. Nigg, A. Erhard, M. Heyl, P. Hauke, M. Dalmonte, T. Monz, P. Zoller, *et al.*, "Real-time dynamics of lattice gauge theories with a few-qubit quantum computer," *Nature*, vol. 534, no. 7608, pp. 516–519, 2016.

[214] H. Bernien, S. Schwartz, A. Keesling, H. Levine, A. Omran, H. Pichler, S. Choi, A. S. Zibrov, M. Endres, M. Greiner, *et al.*, "Probing

many-body dynamics on a 51-atom quantum simulator," *Nature*, vol. 551, no. 7682, pp. 579–584, 2017.

[215] N. Klco, E. F. Dumitrescu, A. J. McCaskey, T. D. Morris, R. C. Pooser, M. Sanz, E. Solano, P. Lougovski, and M. J. Savage, "Quantum-classical computation of schwinger model dynamics using quantum computers," *Physical Review A*, vol. 98, no. 3, 2018.

[216] B. Yang, H. Sun, R. Ott, H.-Y. Wang, T. V. Zache, J. C. Halimeh, Z.-S. Yuan, P. Hauke, and J.-W. Pan, "Observation of gauge invariance in a 71-site bose–hubbard quantum simulator," *Nature*, vol. 587, no. 7834, pp. 392–396, 2020.

[217] A. Mil, T. V. Zache, A. Hegde, A. Xia, R. P. Bhatt, M. K. Oberthaler, P. Hauke, J. Berges, and F. Jendrzejewski, "A scalable realization of local u(1) gauge invariance in cold atomic mixtures," *Science*, vol. 367, no. 6482, pp. 1128–1130, 2020.

[218] M. C. Bañuls, R. Blatt, J. Catani, A. Celi, J. I. Cirac, M. Dalmonte, L. Fallani, K. Jansen, M. Lewenstein, S. Montangero, *et al.*, "Simulating lattice gauge theories within quantum technologies," *The European Physical Journal D*, vol. 74, no. 8, 2020.

[219] Y. Alexeev, D. Bacon, K. R. Brown, R. Calderbank, L. D. Carr, F. T. Chong, B. DeMarco, D. Englund, E. Farhi, B. Fefferman, *et al.*, "Quantum computer systems for scientific discovery," *PRX Quantum*, vol. 2, no. 1, 2021.

[220] M. Aidelsburger, L. Barbiero, A. Bermudez, T. Chanda, A. Dauphin, D. González-Cuadra, P. R. Grzybowski, S. Hands, F. Jendrzejewski, J. Jünemann, *et al.*, "Cold atoms meet lattice gauge theory," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 380, no. 2216, 2021.

[221] E. Zohar, "Quantum simulation of lattice gauge theories in more than one space dimension—requirements, challenges and methods," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 380, no. 2216, 2021.

[222] Z.-Y. Zhou, G.-X. Su, J. C. Halimeh, R. Ott, H. Sun, P. Hauke, B. Yang, Z.-S. Yuan, J. Berges, and J.-W. Pan, "Thermalization dynamics of a gauge theory on a quantum simulator," *arXiv.org*, 2022.

[223] J. Mildenberger, W. Mruczkiewicz, J. C. Halimeh, Z. Jiang, and P. Hauke, “Confinement in a

$$\mathbb{Z}_2$$

lattice gauge theory on a quantum computer,” *Nature Physics*, vol. 21, no. 2, pp. 312–317, 2025.

[224] X. Wang, E. Khatami, F. Fei, J. Wyrick, P. Namboodiri, R. Kashid, A. F. Rigosi, G. Bryant, and R. Silver, “Quantum simulation of an extended fermi-hubbard model using a 2d lattice of dopant-based quantum dots,” 2022.

[225] T. V. Zache, M. Van Damme, J. C. Halimeh, P. Hauke, and D. Banerjee, “Achieving the continuum limit of quantum link lattice gauge theories on quantum devices,” *arXiv.org*, 2022.

[226] N. H. Nguyen, M. C. Tran, Y. Zhu, A. M. Green, C. H. Alderete, Z. Davoudi, and N. M. Linke, “Digital quantum simulation of the schwinger model and symmetry protection with trapped ions,” May 2022.

[227] H. Osamu, Hirota A.S. and C. C.M., *Quantum Communication, Computing, and Measurement.* Springer Science & Business Media, 2012.

[228] H. Bombin, R. S. Andrist, M. Ohzeki, H. G. Katzgraber, and M. A. Martin-Delgado, “Strong resilience of topological codes to depolarization,” *Phys. Rev. X*, vol. 2, p. 021004, Apr 2012.

[229] A. G. Fowler, M. Mariantoni, J. M. Martinis, and A. N. Cleland, “Surface codes: Towards practical large-scale quantum computation,” *Phys. Rev. A*, vol. 86, p. 032324, Sep 2012.

[230] A. Dymarsky and A. Shapere, “Quantum stabilizer codes, lattices, and cfts,” *Journal of High Energy Physics*, vol. 2021, no. 3, 2021.

[231] C. Stahl, “Single-shot quantum error correction in intertwined toric codes,” *Phys. Rev. B*, vol. 110, p. 075143, Aug 2024.

[232] M. M. Wauters, E. Ballini, A. Biella, and P. Hauke, “Symmetry-protection zeno phase transition in monitored lattice gauge theories,” 2024.

[233] S. Kim, “Quantum error correction and $z(2)$ lattice gauge theories,” 2025.

[234] P. Buividovich and M. Polikarpov, “Entanglement entropy in gauge theories and the holographic principle for electric strings,” *Physics Letters B*, vol. 670, no. 2, pp. 141–145, 2008.

[235] R. M. Soni and S. P. Trivedi, “Aspects of entanglement entropy for gauge theories,” *Journal of High Energy Physics*, vol. 2016, jan 2016.

[236] W. Donnelly, “Decomposition of entanglement entropy in lattice gauge theory,” *Physical Review D*, vol. 85, no. 8, 2012.

[237] A. Gromov and R. A. Santos, “Entanglement entropy in 2d non-abelian pure gauge theory,” *Physics Letters B*, vol. 737, pp. 60–64, 2014.

[238] H. Casini, M. Huerta, and J. A. Rosabal, “Remarks on entanglement entropy for gauge fields,” dec 2013. arXiv:1312.1183 [cond-mat, physics:hep-th, physics:quant-ph].

[239] S. Aoki, T. Iritani, M. Nozaki, T. Numasawa, N. Shiba, and H. Tasaki, “On the definition of entanglement entropy in lattice gauge theories,” *Journal of High Energy Physics*, vol. 2015, no. 6, 2015.

[240] S. Ghosh, R. M. Soni, and S. P. Trivedi, “On the entanglement entropy for gauge theories,” *Journal of High Energy Physics*, vol. 2015, no. 9, 2015.

[241] D. Radičević, “Entanglement in weakly coupled lattice gauge theories,” *Journal of High Energy Physics*, vol. 2016, no. 4, pp. 1–31, 2016.

[242] K. Van Acoleyen, N. Bultinck, J. Haegeman, M. Marien, V. B. Scholz, and F. Verstraete, “Entanglement of distillation for lattice gauge theories,” *Physical Review Letters*, vol. 117, no. 13, 2016.

[243] H. Casini, M. Huerta, J. M. Magán, and D. Pontello, “Entanglement entropy and superselection sectors. part i. global symmetries,” *Journal of High Energy Physics*, vol. 2020, no. 2, 2020.

[244] D. Radičević, "Notes on entanglement in abelian gauge theories," 2022.

[245] P. Zanardi, "Quantum entanglement in fermionic lattices," *Physical Review A*, vol. 65, no. 4, 2002.

[246] M.-C. Bañuls, J. I. Cirac, and M. M. Wolf, "Entanglement in fermionic systems," *Physical Review A*, vol. 76, aug 2007.

[247] N. Friis, A. R. Lee, and D. E. Bruschi, "Fermionic-mode entanglement in quantum information," *Physical Review A*, vol. 87, feb 2013.

[248] A. P. Balachandran, T. R. Govindarajan, A. R. de Queiroz, and A. F. Reyes-Lega, "Entanglement and particle identity: A unifying approach," *Physical Review Letters*, vol. 110, feb 2013.

[249] P. Naaijkens, "Superselection sectors in quantum spin systems," 2014.

[250] N. Friis, "Reasonable fermionic quantum information theories require relativity," *New Journal of Physics*, vol. 18, p. 033014, mar 2016.

[251] R. Harris, F. Brito, A. J. Berkley, J. Johansson, M. W. Johnson, T. Lanting, P. Bunyk, E. Ladizinsky, B. Bumble, A. Fung, A. Kaul, A. Kleinsasser, and S. Han, "Synchronization of multiple coupled rf-squid flux qubits," *New Journal of Physics*, vol. 11, no. 12, p. 123022, 2009.

[252] R. Harris, T. Lanting, A. J. Berkley, J. Johansson, M. W. Johnson, P. Bunyk, E. Ladizinsky, N. Ladizinsky, T. Oh, and S. Han, "Compound josephson-junction coupler for flux qubits with minimal crosstalk," *Phys. Rev. B*, vol. 80, p. 052506, Aug 2009.

[253] A. M. v. d. Brink, A. J. Berkley, and M. Yalowsky, "Mediated tunable coupling of flux qubits," *New Journal of Physics*, vol. 7, pp. 230–230, 2005.

[254] R. Harris, A. J. Berkley, M. W. Johnson, P. Bunyk, S. Govorkov, M. C. Thom, S. Uchaikin, A. B. Wilson, J. Chung, E. Holtham, J. D. Biamonte, A. Y. Smirnov, M. H. S. Amin, and A. Maassen van den Brink, "Sign- and magnitude-tunable coupler for superconducting flux qubits," *Phys. Rev. Lett.*, vol. 98, p. 177001, Apr 2007.

[255] N. Dattani, S. Szalay, and N. Chancellor, “Pegasus: The second connectivity graph for large-scale quantum annealing hardware,” 2019.

[256] K. Boothby, P. Bunyk, J. Raymond, and A. Roy, “Next-generation topology of d-wave quantum processors,” 2020.

[257] “Zephyr topology of d-wave quantum processors.” `https://www.dwavesys.com/media/2uznec4s/14-1056a-a_zephyr_topology_of_d-wave_quantum_processors.pdf`, 2021. Accessed: 2025-01-25.