MAC-MIGS

The Maxwell Institute Graduate School in Mathematical Modelling, Analysis and Computation

Quantum Annealing Dynamics

Anastasia Istratuca

Jacob Armstrong-Goodall

Razvan-Andrei Lascu

Supervisor: Professor Benedict Leimkuhler

Abstract

Since the world is in the middle of a race for achieving quantum "supremacy", every aspect related to quantum computing represents a popular research topic. By relying on applications from statistical physics and strong mathematical machinery, this report describes our exploration of the capabilities offered by a specialized quantum computer called *quantum annealer*. Our goals are to understand the differences between quantum and simulated annealing, to design numerical methods for implementing both simulated quantum annealing and simulated thermal annealing using a classical model, and finally to benchmark this model against the quantum annealer provided by D-Wave Systems. We believe that our work extends the literature by proposing a more complex numerical scheme than the one devised by Ma and Dudarev [MD11]. Additionally, we have extended the classical benchmarking model proposed by Subires et al [Sub+21], but further investigation is required in order to compare and contrast results.

Acknowledgements

We would like to endorse the considerable effort and support of our supervisor, Prof. Ben Leimkuhler, and thank him for the guidance, encouragement and advice he has provided throughout the duration of the project. We are deeply grateful that he was, more often than not, as engaged in the work for the project as we were, and that he always opened up (many!) new avenues to further explore the topic. We would also like to recognise the assistance of Dr. Pui-Wai (Leo) Ma, who very kindly provided access to the `C++` implementation for the paper which formed the basis of this project.

Author Contributions

We would like to emphasise that equal credit should be given to each of us. We really have worked together on every element of the project. We all have, in one way or another, contributed to the reading and research required to understand (very) challenging concepts, to the mathematical and computational implementation of the algorithms, to the numerical exploration and analysis of our results, and to the final output summarising our work.

Contents

List of Figures

1 Introduction

Quantum computing has started to permeate every aspect of computational and applied mathematics where classical computers display limitations. Although the hardware implementation of quantum computers is still work in progress, open-source simulators such as IBM's Qiskit allow the user to explore the software functionalities of a quantum computer. Quantum computers are theorised to have the potential to outperform classical computers on a number of important problems. There is already a collection of well-developed quantum algorithms, such as the Deutch-Jozsa algorithm, Grover's search algorithm, Simon's algorithm, etc., which are fully understood, but lack any practical application. A remarkable treatment on theoretical quantum computing can be found in [NC11].

There is a wide class of optimisation problems which are intractable using a classical computer because they fall into the NP-hard set of problems. We will briefly enumerate some of them in the following subsection. Thus, there has been a strong interest in specialised devices which perform quantum computing and which could address these challenging problems. Such devices are called *quantum annealers*, and the one we will focus on in our work was built in 1999 by the Canadian company D-Wave Systems.

Quantum annealing operates in the spirit of simulated annealing, which we will discuss in Chapter 1. In order to efficiently solve these difficult optimisation problems, quantum annealers also leverage the physical properties of quantum systems.

The foundations of quantum annealing will be discussed in Chapter 2, where it is shown that simulated quantum annealing (SQA) allows for shorter annealing times than simulated thermal annealing (SA).

The realisation of quantum annealers such as the D-Wave machine has opened up new potential directions for research. One such area consists of using simulations of the quantum annealing process as a benchmark for the performance of the annealer, and as a means to understanding the behaviour of the annealing process. To this end, a model is introduced in Chapter 3, where Langevin's equations of motion for multi-spin systems are integrated using a Hamiltonian and norm-preserving splitting method.

In order to effectively simulate the annealing process, we will delve into the implementation of the D-Wave device in Chapter 4. Then, in Chapter 5, we will present our numerical results, where we will make preliminary comparisons to results obtained using the D-Wave device.

1.1 Combinatorial Optimisation

The focus of combinatorial optimisation is on finding the "optimal" object (i.e. an object that maximises or minimises a particular function) from a finite set of mathematical objects. Problems of this type arise frequently in real world settings and throughout pure and applied mathematics, operations research and theoretical computer science. Typically, it is impractical to apply an exhaustive search, as the number of possible solutions grows rapidly with the "size" of the input to the problem. The aim of combinatorial optimisation is to find cleverer methods (i.e. algorithms) for exploring the solution space.

Mathematical optimisation seeks to minimise or maximise a cost function. Without loss of generality, we can consider only minimisation. Prominent examples of optimisation problems are the travelling salesman problem (TSP), the graph max-cut problem, Hamiltonian graph partitioning, or in physics, searching for the ground state of a spin glass. When solving optimisation problems one aims to build algorithms which run in polynomial time. In the case of the TSP, this could be impossible, so heuristics such as the Lin-Kernighan algorithm [LK73] have been developed to approximate optimal solutions of the TSP . Combinatorial optimisation problems can be ranked in terms of complexity as follows, from [Joh90]:

- P: The class of decision problems solvable on a deterministic Turing Machine (DTM)

obeying a polynomial bound on running time in terms of the input size.

- NP: Problems whose solutions can be verified in polynomial time on a non-deterministic Turing Machine (NDTM) (NP stands for non-deterministic polynomial time).
- NP-Complete: The problem X is NP-complete if there is a polynomial time algorithm reducing all NP problems to X, and X is itself NP. Since NP complete problems are reducible to one another in polynomial time, a polynomial algorithm for one can solve all of the others. Many combinatorial optimisation problems are NP-complete.
- NP-Hard: A problem X, not necessarily NP, is NP-hard if there exists an NP-complete problem Y for which there is a polynomial-time algorithmic reduction from Y to X. The full class of NP-hard problems are only solvable in polynomial time if P = NP.

Heuristics such as divide-and-conquer and iterative improvement are developed for approximately solving the problem, but we want to probe the average performance relative to computing requirements. For very large problems, this is linked to statistical mechanics which studies the statistical behaviour of systems in thermal equilibrium.

1.2 Statistical Mechanics

One of the most famous models in statistical mechanics is the Ising model whose original purpose was to study ferromagnetism. An extensive introduction to this model and its properties can be found in Cipra's paper [Cip87]. To understand the model, one needs to imagine a lattice such that at each vertex there is an atom of magnetic material. Every such atom has an intrinsic magnetic moment, called *spin*, and which could point either "up" or "down". Mathematically, we invoke the convention that a value of $+1$ is associated to an upward pointing spin, and a value of -1 is assigned to a downward pointing spin. What we have so far is a system of spins in a solid, and we are interested in studying its energy and dynamics. This could be done by considering the Hamiltonian of the system. For each configuration of spins $(s_1, ..., s_N)$, it can be written as:

$$\mathcal{H}(s_1, ..., s_N) = -\frac{1}{2}\left(\sum_{(i,j)\in\mathcal{I}} J_{ij}s_is_j\right) - \sum_{i=1}^{N} hs_i, \qquad s_i \in \{-1, +1\}. \tag{1.1}$$

The Ising model assumes that only spins whose corresponding vertices, usually the integer lattice, are connected by an edge can interact. Thus, long-range interactions are neglected, which is a strong assumption, but renders the system easier to analyse. For a ferromagnet, $J_{ij} > 0$, which means that spins s_i, s_j are forced to align and point in the same direction, i.e. $s_is_j = 1$. For an anti-ferromagnet, $J_{ij} < 0$, so spins point in opposite directions, i.e. $s_is_j = -1$. The extra parameter h represents an external magnetic field which guides spins to point in the direction dictated by the field. The Ising model is the discrete version of a more general model of ferromagnetism, called the Heisenberg model. The latter lies at the heart of our subsequent discussion.

One of the main difficulties in statistical mechanics is to calculate the partition function of a system:

$$Z(T) = \sum_{\omega} e^{-\mathcal{U}(\{r_1, r_2, ..., r_N\})/k_BT}, \tag{1.2}$$

where $\mathcal{U}(\{r_1, r_2, \ldots, r_N\})$ is the potential energy of the configuration of the system, k_B is the Boltzmann constant, and T is the temperature. Each configuration ω is defined by the set of atomic positions $\{r_i\}$, so summing over ω boils down to summing over all possible configurations of the system.

Barahona [Bar82] studies the problem of calculating the partition function and finding the ground state of the energy for Ising spin glass models. In particular, he shows that, in the case of two-dimensional finite lattices, these problems are polynomially solvable with respect to the size of the lattice. However, they become NP-Hard for the three-dimensional case, which is exactly the Heisenberg model. Its Hamiltonian is given by:

$$\mathcal{H}(S_1, ..., S_N) = -\frac{1}{2}\left(\sum_{(i,j)\in\mathcal{I}} J_{ij} S_i \cdot S_j\right) - \sum_{i=1}^{N} h \cdot S_i, \quad S_i \in \mathbb{R}^3, \quad \|S_i\| = 1, \tag{1.3}$$

where h is an external magnetic field. A coupling J_{ij} chosen randomly from $\{-1, +1\}$ is assigned to each pair (i, j) connected by an edge. The Heisenberg model is the continuous 3-dimensional version of the Ising model.

What we have been discussing so far are classical models which could be seen from a quantum-mechanical perspective as well. The Stern-Gerlach experiment [GS22] suggests that, besides the classical idea of a spin which we have already described, the electron has an extra degree of freedom which allows it to "self-rotate". The mathematical apparatus for this discovery was developed by Pauli (1927). Using the formalism of quantum mechanics, the spin state of an electron can be represented using the two-dimensional basis

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{1.4}$$

Formally, we aim to define an operator $\hat{S}_z$ which acts on the $\{|0\rangle, |1\rangle\}$ basis, as follows:

$$\hat{S}_z |0\rangle = \frac{\hbar}{2} |0\rangle, \quad \hat{S}_z |1\rangle = -\frac{\hbar}{2} |1\rangle. \tag{1.5}$$

These eigenvalue equations can be translated into the language of the classical model by observing that states $|0\rangle$ and $|1\rangle$ correspond to upward and downward pointing spins in the z-direction, respectively.

The general form of a state can be described as:

$$|\Psi\rangle = \alpha |0\rangle + \beta |1\rangle, \tag{1.6}$$

with $\alpha, \beta \in \mathbb{C}$. Hence, showing the action of the operator $\hat{S}_z$ on the state $|\Psi\rangle$ can be done by representing $\hat{S}_z$ as the 2×2 matrix:

$$\hat{S}_z = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{1.7}$$

One can easily check that this form indeed acts on the $\{|0\rangle, |1\rangle\}$ basis, as required. The expressions for the $\hat{S}_x$ and $\hat{S}_y$ operators are:

$$\hat{S}_x = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \hat{S}_y = \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}. \tag{1.8}$$

Then, one can finally define the *Pauli matrices* σ^α by:

$$\hat{S}_\alpha = \frac{\hbar}{2}\sigma^\alpha, \quad \alpha \in \{x, y, z\}. \tag{1.9}$$

Returning to our previous discussion, we can state the Hamiltonian of the Quantum Ising model

as:

$$\hat{\mathcal{H}} = -\frac{1}{2}\left(\sum_{(i,j)\in\mathcal{I}} J_{ij}\sigma_i^z\sigma_j^z\right) - \sum_{i=1}^{N} h\sigma_i^z. \tag{1.10}$$

1.3 A type of optimisation problem in statistical physics

In ideal systems, all particles are alike and the ground state is a unique configuration depending on a single parameter. An optimisation problem typically has distinct changeable elements and perhaps multiple control parameters. Systems where not all atoms are alike may simultaneously contain the presence of interactions favouring different and incompatible kinds of ordering. For example, spin glasses display competition between ferromagnetic and anti-ferromagnetic ordering. Consider an Ising-type Hamiltonian where the couplings J_{ij} are sampled from the Gaussian distribution. This corresponds to the Sherrington-Kirkpatrick (SK) model [SK75] of spin glasses with long-range interactions, i.e. all spins on the lattice interact:

$$\mathcal{H}_{\mathcal{J}} = -\sum_{\langle i,j\rangle} J_{ij}s_is_j - h\sum_i s_i, \tag{1.11}$$

where s_i is the spin at site i, h is the external field, J_{ij} are i.i.d. Gaussian random variables. We emphasise that the subscript $\mathcal{J}$ denotes a particular realisation of the couplings. Note that J_{ij} can be greater, less than or equal to zero and, once chosen, are fixed for all time. In this case, ferromagnetic and anti-ferromagnetic interactions compete with each other, which leads to a phenomenon called *frustration*: no spin configuration can simultaneously satisfy all couplings. As the name suggests, spin glasses have the same relationship to magnets as glass does to crystals, and they exhibit many degenerate ground states. Another famous spin glass model is Edwards-Anderson model [EA75] with nearest-neighbour interactions and no external field. An interesting connection between the travelling salesman problem and the SK model was found in [KT85]. Finding the ground state of the SK model is also known to be an NP-hard problem.

1.4 The QUBO and Ising formulations

Having discussed combinatorial optimisation and statistical mechanics, we now unify these two areas and show how, in the paradigm of quantum computing, they are used interchangeably. Quadratic Unconstrained Binary Optimisation (QUBO) problems are a class of NP-hard problems which can be treated using the infrastructure offered by D-Wave. Examples of combinatorial optimisation problems which can be translated into the QUBO formulation are the Maximum-cut and Graph partition problem. For a survey on QUBOs, one can consult [Koc+14].

A QUBO problem is, in general, formulated in terms of a symmetric matrix Q with real-valued entries Q_{ij}, and, more concretely, it is defined by the binary minimisation problem:

$$\min_{x_i\in\{0,1\}} \sum_{i\leq j} x_iQ_{ij}x_j. \tag{1.12}$$

The mapping $x_i = \frac{1}{2}(s_i + 1)$ sends $x_i = 0$ to $s_i = -1$ and $x_i = 1$ to $s_i = 1$. Hence, making this substitution in the minimisation problem (1.12) we obtain an equivalent minimisation problem:

$$\min_{s_i\in\{-1,1\}} -\frac{1}{2}\sum_{i\neq j} J_{ij}s_is_j - \sum_i h_is_i + C, \tag{1.13}$$

where $J_{ij} = -\frac{1}{2}Q_{ij}$, for $i \neq j$, $J_{ii} = 0$, $h_i = -\frac{1}{2}\sum_j Q_{ij}$, and the constant $C = \frac{1}{4}\sum_{i\leq j} Q_{ij}$. This latter formulation is exactly the problem of finding the ground state of the Ising Hamiltonian up to a constant. Once the ground state is calculated, it is just shifted by the value of C, and

thus the QUBO problem is also solved. More discussion about the QUBO-Ising duality can be found in [VV21], [Wil+21].

1.5 The need for a continuous model

Suppose we are interested in the properties of materials in their lowest energy state. As temperature goes down to zero, the distribution collapses into this ground state, but, in practice, low temperature is not sufficient for finding the ground state of a material. This must be done by annealing, for example by melting the substance, then lowering the temperature slowly, spending a long time near the freezing point. If the cooling is not done carefully, then it will get out of equilibrium and, either defects will develop, or the substance will form a glass which has no crystalline order and only meta-stable, locally optimal structures. To study the properties of this phase transition and the formation of topological deformities, it is necessary to describe the annealing dynamics in continuous time. This is the motivation for the Spin Vector Langevin (SVL) model, which is used in [Sub+21] to predict the formation of topological defects or "kinks" during the phase transition in a quantum spin system.

Our work is mainly motivated by the use of the SVL model as a source for benchmarking the dynamics of a quantum annealer. This model relies on the following simplification: each spin operator is assigned to an angle in plane such that $\sigma_i^z \to \sin\theta_i$ and $\sigma_i^x \to \cos\theta_i$. Therefore, the configuration of the classical system is not defined by the spins $(S_1, S_2, ..., S_N)$, rather by the set of their corresponding continuous angles $(\theta_1, \theta_2, ..., \theta_N)$. Then, under this assumption, the SVL model evolves under Langevin dynamics.

We aim to extend this two-dimensional model and benchmark the dynamics of the quantum annealer using the Heisenberg model which is three-dimensional, and which also exhibits rotational symmetry by restricting the spin-length to the unit sphere. Also, in order to perform simulations of the dynamics of this system, we use the the Langevin-type equation introduced by Ma and Dudarev [MD11]. This particular equation is the Landau-Lifschitz (LL) equation augmented by an additional stochastic term which adds the Langevin dynamics behaviour to the system. The exact expressions of the Heisenberg Hamiltonian and the equation we are treating are in Sections 3.1 and 3.2, respectively. It is worth mentioning that the LL equation is not analytically integrable, so we build an integrator based on splitting methods, following the methodology described in [LM16]. We also believe that our integrator extends the one proposed by Ma and Dudarev in [MD11]. The details of this construction are described in Section 3.3.

One important difficulty which became apparent due to time constraints is the investigation of the statistics of topological defects. This limitation of our work will be discussed in Chapter 6.

2 Foundations of Quantum and Simulated Annealing

2.1 The principles of Quantum Annealing

We see thermodynamics in action every day, water heats up, turns to steam, then cools and forms puddles. In principal, quantum annealing is a thermodynamics process and all that separates it from the everyday occurrences we observe is that it takes place at the scale of quantum particles, where classical laws break down and the specific trajectory of a particle is unknowable due to the Heisenberg uncertainty principle. The laws of thermodynamics state that closed systems conserve their energy, entropy increases and the minimum possible energy of a system occurs at absolute zero. By taking advantage of the links made between statistical mechanics and optimisation problems in 1.1, and considering the quantum analog of the laws of thermodynamics, quantum annealing seeks to use the peculiar properties of quantum systems to efficiently solve optimisation and sampling problems.

At the heart of quantum annealing one finds quantum bits, known as *qubits*. These are the quantum version of classical bits and exhibit properties from quantum physics. Inside the quantum computer, qubits are devised as circulating currents with a corresponding magnetic field which allows the information to flow. The currents have a predefined orientation, which is clockwise for state $|0\rangle$ and anti-clockwise for state $|1\rangle$. The distinguishing feature of qubits is that, unlike classical bits, their current state can be a linear combination of the states $|0\rangle$ and $|1\rangle$, a principle known as *superposition* [The+17]:

$$|\phi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle\,. \tag{2.1}$$

Here $\alpha, \beta \in \mathbb{C}$ describe how much the classical states $|0\rangle$ and $|1\rangle$ contribute to the superposed state of the qubit $|\phi\rangle$.

A key observation regarding Eq. (2.1) is that a qubit's superposed nature is not expressed through $|\phi\rangle$ being intermediate between $|0\rangle$ and $|1\rangle$. In other words, we do not know at any point the amplitudes α and β. Rather, as stated by Dirac [Dir47], the superposed state is expressed through the *probability* of $|\phi\rangle$ being intermediate between the corresponding *probabilities* of $|0\rangle$ and $|1\rangle$. That is, when we measure a qubit, we retrieve the probabilities $|\alpha|^2$ and $|\beta|^2$ of the states $|0\rangle$ and $|1\rangle$, respectively.

The process of quantum annealing for one qubit is shown in Figure 2.1. It starts with the qubit in the lowest-energy superposed state where there is only one minimum. As the system anneals, the potential barrier is lifted and the qubit has probability $\frac{1}{2}$ of ending in either state $|0\rangle$ or state $|1\rangle$. However, by applying an external magnetic field, one could interfere with the potential barrier, thus creating a lower valley where the qubit ends up with higher probability. For example, in Figure 2.1, the lowest-energy state achieved by the qubit is $|1\rangle$. The external magnetic field which makes the qubit minimise its energy is a control parameter chosen by the user of the quantum machine.

Of course, systems in general contain at least two qubits which can interact with each other. This interaction is called coupling and it is managed by a device built into the quantum machine

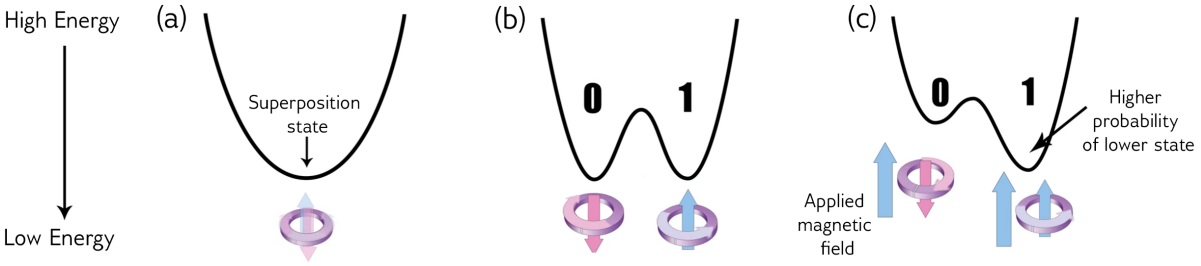


Figure 2.1: Energy diagram of the quantum annealing process [D-Wg]

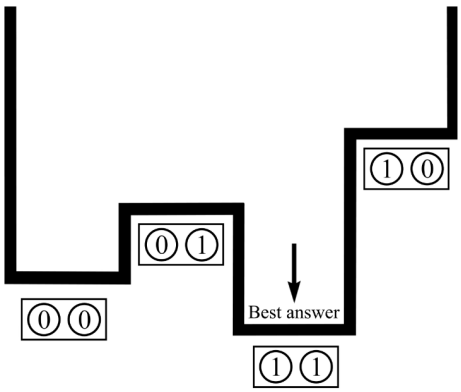


Figure 2.2: Energy diagram for two entangled qubits [D-Wg]

called a coupler [D-Wg]. Essentially, the coupler measures the coupling strength between qubits and its value is inserted into the quantum machine by the user. If the coupling is positive, the qubits will tend to end up in the same state (ferromagnetic), whereas if the coupling is negative, they will tend to opposite states (anti-ferromagnetic). Coupling is based on another phenomenon from quantum physics called entanglement [Dir47]. Roughly speaking, two entangled qubits can be considered as an individual object with four possible states: $|00\rangle, |01\rangle, |10\rangle, |11\rangle$. Thus, n entangled qubits result in 2^n possible states, which indicates that systems could become rapidly complex for a reasonably small number of qubits. For example, Figure 2.2 shows the energy diagram for a system of two entangled qubits with $|11\rangle$ being the lowest-energy state found by the annealer.

2.2 The probabilistic framework for Simulated Annealing

Simulated annealing was introduced by Kirkpatrick, Gelatt and Vecchi [KGV83] in 1983 as a method to solve optimisation problems for which the cost function to be minimised exhibits several local minima. The probabilistic framework for simulated annealing was proposed in a highly cited article of Bertsimas and Tsitsiklis [BT93] where they restrict to discrete-time Markov chains for which the cost function covers only a finite number of states.

Before we describe the method, it is essential to fix some notation. Let us define the cost function which we aim to minimise by $c : S \to \mathbb{R}$, where S is the set of states the system goes through. The proper subset $S_{\min} \subset S$ is the set of all minima of c. The decreasing function $T : \mathbb{N} \to (0, \infty)$ is called the *annealing schedule* and $T(t)$ is the temperature at time $t \in \mathbb{N}$. For every state $i \in S$, we can assign a transition probability $q_{ij} \in [0, 1]$, where $j \in N(i) \coloneqq \{j \in S : j \neq i \text{ is a neighbour of } i\}$. These probabilities guarantee that we could visit every state of the Markov chain. Furthermore, by the axioms of probability, we require that $\sum_{j \in N(i)} q_{ij} = 1$. It is also reasonable to assume that $j \in N(i)$ if and only if $i \in N(j)$. Now, the Markov chain is defined as $\{x(t) : t \in \mathbb{N}\}$, and it starts from an initial state $x(0) \in S$. In general, moving from state $x(t)$ to the next state $x(t+1)$ is described by Algorithm 1. If we are at state $x(t) = i$, then we can arbitrarily choose a neighbour state $j \in N(i)$ according to the transition probability q_{ij}.

The Metropolis algorithm introduced by Metropolis et al [Met+53] aims to sample elements from a distribution, say, ρ by simulating a Markov chain which is known to have ρ as stationary distribution. One could understand the reasoning behind the simulated annealing algorithm by making an analogy to the statistical physics context. A problem that has been of great importance in statistical physics is to sample points from the set S according to the well-known Gibbs distribution π_T given by

$$\pi_T(i) = \frac{1}{Z_T} \exp\left(\frac{-c(i)}{T}\right), \quad i \in S. \tag{2.2}$$

Algorithm 1 Going from state $x(t)$ to state $x(t+1)$

Require: $x(t) = i$
  if $c(j) \leq c(i)$ **then**
    $x(t+1) = j$
  end if
  if $c(j) > c(i)$ **then**
    $x(t+1) = j$ with probability $\exp\left(-\frac{c(j)-c(i)}{T(t)}\right)$
  else
    $x(t+1) = i$.
  end if

This is done by simulating the irreducible and aperiodic Markov chain $\{x_T(t) : t \in \mathbb{N}\}$ until equilibrium is achieved. In this case, we consider the temperature $T(t)$ to be constant and equal to T. Letting $T \to 0$, one can observe that the samples drawn from the distribution π_T are concentrated in $S_{\min}$. Returning to the optimisation picture, one can maintain the temperature T very low and sample points from the distribution π_T in order to make sure that the points sampled lie in $S_{\min}$ with large probability. However, there is a drawback to this strategy. If T happens to be too low, it will take an exaggerated amount of time for x_T to reach equilibrium. Hence, at the expense of having an additional increase of energy, we start the process from a sufficiently large temperature $T(t)$ and gradually decrease it according to the annealing schedule. Thus, one can avoid being trapped in local minima.

2.3 Convergence results

In theory, the discussion above settles the answer to the convergence of $x(t)$ to an element in the set $S_{\min}$. The next crucial question one might pose is related to how fast the algorithm converges. One of the first results in this sense was proved by Geman and Geman [GG84] (1984) who essentially established a bound on the family of annealing schedules. It is worth pointing out that if the rate of decrease for $T(t)$ is unreasonably slow, then the algorithm is impractical. Hence, the optimal schedule proposed by Geman and Geman is $T(t) = k/\log t$ or anything lower than logarithmic decay. The parameter k depends on the characteristics of the system and the cost function. Nevertheless, faster schedules, such as linear decrease, might lead to adequate outcomes in practice. A more general and sharper result regarding convergence was stated and proven by Hajek [Haj88].

Theorem 2.1 (Hajek (1988), Bertsimas and Tsitsiklis (1993))**.** *Suppose that there exists a collection of neighbours in S which can be ordered in a path starting from $i \in S$ and ending at some element in $S_{\min}$ such that the largest value of c along the path is equal to $c(i)+h$. Then we say that i is connected to $S_{\min}$ at height h. Let k^* be the smallest height for which i is connected to $S_{\min}$. Then the simulated annealing algorithm converges if and only if*

$$\lim_{t\to\infty} T(t) = 0 \tag{2.3}$$

and

$$\sum_{t=1}^{\infty} \exp\left(-k^*/T(t)\right) = \infty. \tag{2.4}$$

Here we emphasise that the necessary and sufficient condition for convergence does not depend on the explicit form of the schedule, but rather on the parameter k^*. The quantity k^* captures the difficulty of the system to overcome being trapped in local minima and reach the set $S_{\min}$. The condition that $T(t)$ has to convergence to 0 for large t is sensible given the nature of the algorithm just described.

We now offer a heuristic argument why the last condition of the previous theorem guarantees convergence. Consider a local minimum whose "valley" measures height k^*. Then, consider the infinite sequence of events which say that SA escapes from this local minimum. These events are clearly independent from each other, and the probability that each event succeeds is (roughly) $\exp\left(-k^*/T(t)\right)$. Then, by the Borel-Cantelli lemma, one could deduce that the escapes happen infinitely often with almost sure probability.

There are various annealing schedules that one might experiment with in practical applications, such as linear and quadratic schedules $T(t) = T_0/(1+\alpha t), T(t) = T_0/(1+\alpha t^2)$, or even an exponential one $T(t) = T_0 \exp(-\alpha t)$, where T_0 is the initial temperature. The empirical parameter $\alpha > 0$ is set at the beginning of the process and controls the cooling rate of the system. One is right to expect the exponential decay to be unacceptably fast, even with a carefully chosen α. However, a faster schedule is desired whenever one cannot access sufficiently powerful computational resources. There has been some recent progress on trying to mitigate this difficulties by adapting the cooling rate during the optimisation process. In particular, Karabin and Stuart [KS20] consider the exponential schedule above, but modify the cooling rate such that it varies as a function of the instantaneous temperature, $\alpha(T(t))$. The purpose of this change is to anneal slowly around regions containing important temperatures and anneal faster when the system is far from those regions. Therefore, the annealing schedule becomes

$$T(t) = T_0 \exp\left(-\int \alpha(T(t))\mathrm{d}t\right), \tag{2.5}$$

where the actual cooling is done the via finite-difference increments

$$T(t+\Delta t) = T(t) - \alpha(T)\Delta t. \tag{2.6}$$

Even though one can devise schemes for optimising the simulated annealing algorithm, the advantage of quantum computers hints to a potentially quicker, yet more sophisticated type of annealing.

2.4 Quantum annealing vs simulated annealing

Recently, there has been an interest in how the principles of quantum physics could improve the performance of simulated annealing [KN98], [MN08], [Bro+99], [Suz09]. Thus, the concept of quantum annealing was born. As already mentioned, the main difficulty of simulated annealing lies in the height of barriers which separate the minima of the cost function. A potential remedy to this issue is quantum tunneling for which the height of the barrier is irrelevant. One can potentially accelerate the convergence of simulated annealing by tunnelling directly through the barriers, without incurring the additional cost of energy to "climb" them. Instead of causing thermal fluctuations in the system, quantum annealing cleverly introduces a transverse field Γ into the system and slowly decreases it to induce quantum fluctuations. In the case when temperature is a variable in the system, it is kept fixed during the annealing process. In order to illustrate the distinction between the two types of annealing, we rely on the following example from statistical physics. Suppose one aims to find the ground state (global minimum) of an Ising model with interactions between all vertices i, j connected by an edge. The Hamiltonian whose ground state we aim to obtain is given by

$$\mathcal{H}_{\text{Ising}}(t) = -\frac{1}{2}\sum_{i,j} J_{ij} s_i s_j, \qquad s_i \in \{-1, +1\}. \tag{2.7}$$

In particular, J_{ij} are couplings such that $J_{ij} > 0$ denotes ferromagnetic coupling and $J_{ij} < 0$ denotes anti-ferromagnetic coupling. To introduce the transverse field term Γ, we need to

consider the quantum version of the Hamiltonian (2.7):

$$\hat{\mathcal{H}}(t) = \underbrace{-\frac{1}{2}\sum_{i,j} J_{ij}\sigma_i^z\sigma_j^z}_{\text{Quantum Ising Hamiltonian}} \underbrace{-\Gamma(t)\sum_{i=1}^{N}\sigma_i^x}_{\text{Tunnelling Hamiltonian}} \tag{2.8}$$

$$= \hat{\mathcal{H}}_{\text{Ising}} - \Gamma(t)\sum_{i=1}^{N}\sigma_i^x, \tag{2.9}$$

where σ^z and σ^z are the Pauli matrices defined in (1.9). Assuming that $\Gamma = 0$, we get that $\left[\sigma^z, \hat{\mathcal{H}}\right] = 0$. Heisenberg's equation of motion [Che+19] for σ^z reads

$$\frac{\mathrm{d}\sigma^z}{\mathrm{d}t} = -i\left[\sigma^z, \hat{\mathcal{H}}\right], \tag{2.10}$$

where $\hbar = 1$, and therefore σ^z is conserved with time, and hence, so is $\hat{\mathcal{H}}$. However, introducing the tunnelling Hamiltonian, we force the commutator $\left[\sigma^z, \hat{\mathcal{H}}\right]$ to stay constant and non-zero, thus inducing quantum fluctuations in the system's levels of energy.

The transverse field was shown to play an analogous role to temperature in the Hopfield model by Nishimori and Nonomura in [NN96], so (2.8) mimics the rationale behind the SA algorithm. We start the quantum annealing with $\Gamma(t)$ having large magnitude and gradually decrease it until it vanishes. The initial state of the system is exactly the ground state of the tunnelling Hamiltonian. Then we lower the transverse field with time. If the change of the transverse field is conducted slowly enough, the evolution of states is driven adiabatically, and the system reaches the ground state of the Ising Hamiltonian when the transverse field is asymptotically zero.

There is, however, a significant downside to the QA algorithm. The dynamics of the process is governed by the time-dependent Schrödinger's equation:

$$i\frac{\mathrm{d}}{\mathrm{dt}}\left|\Psi(t)\right\rangle = \hat{\mathcal{H}}(t)\left|\Psi(t)\right\rangle, \tag{2.11}$$

where $\Psi(x,t)$ is the (eigen)state which needs to be found at the lowest-energy level. If t_a is a sufficiently long annealing time, then one has that $\Gamma(t_a) \to 0$ as $t_a \to \infty$, and so

$$\hat{\mathcal{H}}_{\text{Ising}}\left|\Psi(t_a)\right\rangle = E_{\min}\left|\Psi(t_a)\right\rangle, \tag{2.12}$$

where $E_{\min}$ is the eigenvalue of the Quantum Ising Hamiltonian. This implies that $E_{\min}$ is the ground state of our "problem" Hamiltonian $\mathcal{H}_{\text{Ising}}$. For small-size systems, one can solve the time-dependent Schrödinger's equation using traditional numerical schemes. For instance, one such scheme is the Fourth Order Runge-Kutta method, which is employed in [Boy20] to solve Eq. (2.11) for the system described above having 16 spins. However, solving the time-dependent Schrödinger's equation (2.11) for large systems can be done efficiently only using quantum computers.

2.5 Annealing schedules for Quantum Annealing

From now on, we shall remove the $\hat{}$ symbol from the Hamiltonian $\hat{\mathcal{H}}$ to distinguish between the classical and the quantum versions. Instead, whenever the Pauli matrices are involved it should be assumed we are referring to the quantum Hamiltonian. If we denote $\mathcal{H}_{\text{TF}}(t) = -\Gamma(t)\sum_{i=1}^{N}\sigma_i^x$,

then we can write the total Hamiltonian as:

$$\mathcal{H} = \mathcal{H}_{\text{Ising}} + \mathcal{H}_{\text{TF}}(t). \tag{2.13}$$

One issue that remains to be discussed is to establish conditions under which the QA algorithm converges. A sufficient condition in this sense is given in the next section. More precisely, a specifically chosen schedule for decreasing Γ ensures that the annealing runs adiabatically. Additionally, one could generalise the Hamiltonian above by writing it as a linear homotopy between a potential term and a kinetic term, as follows:

$$\mathcal{H}(s) = f(s)\mathcal{H}_{\text{pot}} + (1 - f(s))\mathcal{H}_{\text{kin}}, \tag{2.14}$$

where $s = t/t_a$ and the function f is the annealing schedule such that $f(0) = 0$ and $f(1) = 1$. Hence, $\mathcal{H}(0) = \mathcal{H}_{\text{kin}}$ and $\mathcal{H}(1) = \mathcal{H}_{\text{pot}}$, which correspond to the transverse field Hamiltonian $\mathcal{H}_{\text{TF}}(0)$, and the Ising Hamiltonian $\mathcal{H}_{\text{Ising}}$, respectively.

2.6 Convergence and Computation Complexity

Quantum annealing (QA) was introduced in the sense discussed by Kadowaki and Nishimori in [KN98]. Related is the idea of quantum adiabatic evolution (QAE), as applied to the 3-SAT problem by Farhi et al. in [Far+00]. QA and QAE both rely on the quantum adiabatic theorem, the difference is their treatment in the literature where in QAE the focus is usually on the computational complexity of a particular problem, whereas QA is studied as a general algorithm - the quantum analog of thermal annealing. In this section, we discuss results on the convergence time for this algorithm and see how complexity bounds can be deduced for the most general class of problems. Our main reference is [NN96], based on the PhD thesis of Morita, where it is pointed out that polynomial dependence on system size should not be expected in such a general framework, since the model accounts for the very hardest cases of combinatorial optimisation problems. It is interesting to note that, in the ideal noise-free setting, QAE has been proven to be equivalent to the circuit model of quantum computing, see [MLM21].

2.6.1 Quantum Adiabatic theorem

The adiabatic theorem for quantum mechanics [Kat50] states that, although in general the Schrödinger equation (2.11) has no stationary solution, when a perturbation acts on the system slowly enough, and when initialised from a stationary state H_0, the system will pass through all of the stationary states corresponding to H_t for all t. Let the ground state of (2.11) be denoted by $|0(t)\rangle$ and the j^{th} eigenstate which corresponds to the eigenvalue $\epsilon_j(t)$ by $|j(t)\rangle$, so that

$$\mathcal{H}(t)|j(t)\rangle = \epsilon_j(t)|j(t)\rangle. \tag{2.15}$$

Then the adiabaticity condition for the real time Schrödinger equation is:

$$\frac{1}{\Delta_j(t)^2}\left|\langle j(t)|\frac{\mathrm{d}\mathcal{H}(t)}{\mathrm{d}t}|0(t)\rangle\right| = \delta \ll 1, \tag{2.16}$$

where $\Delta_j(t) = \epsilon_j(t) - \epsilon_0(t)$.

2.6.2 The General form of the Ising Model

Now consider the Ising model in its most general form:

$$\mathcal{H}_{\text{Ising}} = -\sum_{i=1}^{N} J_i\sigma_i^z - \sum_{i,j=1}^{N} J_{ij}\sigma_i^z\sigma_j^z - \sum_{i,j,k=1}^{N} J_{ijk}\sigma_i^z\sigma_j^z\sigma_k^z - \cdots, \tag{2.17}$$

where σ^x, σ^y and σ^z are the Pauli matrices. We can introduce a transverse field $\mathcal{H}_T$ with time dependent factor $\Gamma(t)$ to get

$$\mathcal{H} = \mathcal{H}_{\text{Ising}} - \Gamma(t)\mathcal{H}_T. \tag{2.18}$$

Here, Γ controls the strength of the transverse field which, in turn, determines the rate of transition between states. The quantum adiabatic theorem implies that if $\Gamma(t)$ starts off large and decreases very slowly, then in the infinite time limit the system will reach the ground state of (2.17). The condition for adiabaticity of the quantum system depends on the form of the transverse field. Here we introduce the three forms covered with proofs in [MN08].

2.6.3 Standard Form

First, for the simplest and most commonly studied form of the transverse field:

$$\mathcal{H}_T(t) = -\sum_{i=1}^{N} \sigma_i^z, \tag{2.19}$$

to ensure convergence to the ground state of (2.17), the time dependence of $\Gamma(t)$ must be

$$\Gamma(t) = a(\delta t + c)^{\frac{-1}{2N-1}}, \tag{2.20}$$

where a and c are constants of order $\mathcal{O}(N^0)$ and $t > t_0 > 0$. This time dependence guarantees convergence as $t \longrightarrow \infty$. It is important to note that as $t_0 \longrightarrow 0$, the coefficient a diverges, so t_0 must indeed be strictly greater than zero.

2.6.4 The Transverse Field with Ferromagnetic Coupling

If a ferromagnetic coupling term is included in the transverse field, that is, if the transverse field is defined as

$$\mathcal{H}_{TI}(t) = -\Gamma_{TI}(t)\left(\sum_{i=1}^{N}\sigma_i^x + \sum_{i,j}\sigma_i^x\sigma_j^x\right), \tag{2.21}$$

then the above convergence condition can be improved, in the sense that it theoretically decreases the annealing time required to a time dependency of $\Gamma(t) \propto t^{\frac{-1}{(N-1)}}$ for the quantum system $\mathcal{H}_{\text{Ising}} + \mathcal{H}_{TI}$.

2.6.5 Infinite Range Ferromagnetic Ising Model

If the transverse field is taken with the form:

$$\mathcal{H}_{MTI} = \Gamma_{MTI}(t)\prod_{i=1}^{N}(1-\sigma_i^x), \tag{2.22}$$

then the annealing time can be further reduced to the schedule

$$\Gamma_{MTI} \propto \frac{2^{N-2}}{\delta t}. \tag{2.23}$$

2.6.6 Computational Complexity

These convergence results can be used to directly derive very general results about the computational complexity of combinatorial optimisation problems that can be written in the form of (2.17). The above schedules do not grow polynomially with the size N of the system. Indeed,

it can be shown that the time taken for $\Gamma(t)_{TF}$ to obtain a lower bound ϵ sufficiently close to the ground state of $\mathcal{H}_{\text{Ising}}$ is

$$t_{TF} \approx \frac{1}{\delta}\left(\frac{1}{\epsilon}\right)^{2N-1}. \tag{2.24}$$

Likewise, for the infinite range ferromagnetic field, the time taken to achieve a close enough approximation to the ground state of (2.17) is

$$t_{MTI} \approx \frac{s^{N-1}}{\delta\epsilon}, \tag{2.25}$$

which is exponentially dependent on the size of the problem, N. Although the annealing times grow exponentially with N, they are much faster than the corresponding times required in simulated thermal annealing, which is what we discuss next in the quantum setting.

2.6.7 Thermal Annealing in Quantum Systems

The convergence condition for thermal SA was derived by [GG84] using the theory of inhomogeneous Markov chains, as described in the Quantum Monte Carlo context. The relation of annealing time to system size was derived as:

$$T(t) \propto \frac{N}{\log(\alpha t + 1)}. \tag{2.26}$$

Any classical, finite-dimensional spin model on a lattice can be associated to a quantum model, defined on the same lattice. This can be done by mapping every classical state s_i to the Pauli z-operator σ_i^z. This maps the classical Ising model $\mathcal{H}_{\text{Classical}} = \sum_{i,j=1}^{N} J_{ij} s_i s_j$ to the quantum Ising model $\mathcal{H}_{\text{Quantum}} = \sum_{i,j=1}^{N} J_{ij}\sigma_i^z\sigma_j^z$.

In this way, we can obtain an expression of classical thermal expectation value in terms of a quantum ground state expectation value. If we consider the classical Hamiltonian corresponding to (2.17) via the above mapping, then the thermal expectation value of the classical physical quantity $Q(\sigma_i^z)$ is:

$$\langle Q\rangle_T = \frac{1}{Z(T)}\sum_{\mathcal{S}} e^{-\beta\mathcal{H}_{\text{Ising}}} Q(\{s_i\}). \tag{2.27}$$

Above, the sum is taken over $\mathcal{S}$, i.e. all the configurations of the spin system, and $\{s_i\} = \{s_i : i = 1, ..., N\}$.

The thermal expectation value above is equal to the expectation value of Q by the quantum wave function

$$|\Psi(T)\rangle = e^{-\frac{\beta\mathcal{H}}{2}} \sum_{\mathcal{S}} |\{s_i\}\rangle, \tag{2.28}$$

where $|\{s_i\}\rangle$ is the basis state diagonalising each σ_i^z as s_i, and the sum runs over all possible such assignments. Now, define H_j by the sum of terms in the Hamiltonian (2.17), that is:

$$H_j = -J_j\sigma_j^z - \sum_k J_{jk}\sigma_j^z\sigma_k^z - \sum_{kl} J_{jkl}\sigma_j^z\sigma_k^z\sigma_l^z - \cdots. \tag{2.29}$$

Then, if we take $p = \max_i\{H_i\}$ and $T > 0$, the wave function (2.28) is the ground-state of the following Hamiltonian:

$$H_q(T) = -e^{-\beta p}\sum_j H_q^j(T) = -e^{-\beta p}\sum_j(\sigma_x^j - e^{\beta H_j}). \tag{2.30}$$

The adiabaticity condition applied to the quantum system (2.30) leads to the following condition

for the convergence of simulated thermal annealing:

$$T(t) = \frac{pN}{\log(\alpha t + 1)}, \tag{2.31}$$

where α is exponentially decreasing in N. This reproduces the Geman-Geman result for the convergence of simulated annealing.

2.6.8 Comparison to Quantum Adiabatic Evolution

In Quantum Adiabatic Evolution, a system of bits is encoded with a set of clauses $\mathcal{C} = \{C_i : i = 1, ..., M\}$. Each clause is true or false depending on the state of some subset of the bits. Then the problem Hamiltonian is written as:

$$\mathcal{H}_p(t) = \sum_i H_{C_i}, \tag{2.32}$$

where each H_{C_i} depends only on the clause C_i and acts only on the bits determining the truth of that clause.

We can cast this problem as an energy minimisation problem as they do in [Far+00], and it is possible to show that this maps to the general Ising Hamiltonian (2.17). This approach fits into the framework we are discussing, but, since here we mention the most pessimistic bounds (due to the generality of the results), it is instructive to study specific computationally hard problems to find possibilities for a quantum advantage.

Mathematical analysis of QA focuses generally on convergence conditions, as we do here, while studies of QAE focus on computational complexity of a specific problem in finite time frames. These two topics fit into the same framework, utilising the adiabatic theorem as the fundamental tool, but approach the problem from different angles. It was shown in the discussion above that the adiabaticity condition leads directly to the convergence condition in the infinite time limit, both for quantum annealing with the transverse field and thermal annealing in the quantum setting, thus linking SA, QA and QAE.

2.7 Numerical Methods For Simulated Quantum Annealing

2.7.1 Quantum Monte Carlo Evolution

Simulated Quantum annealing with Schrödinger dynamics is not practical as it involves numerically solving the quantum many-body problem. Normally, Monte-Carlo methods are used as a stochastic approximation to the dynamics of quantum systems. In [Shi+14], Seung Woo Shin et al. use the Metropolis algorithm to perform simulated annealing on two-dimensional vectors, and it is shown that this model for simulated quantum annealing correlates better with the D-Wave output than the standard Metropolis algorithm described in section 2.2. Later, continuous models were proposed, for example in [Sub+21] for quantum annealing benchmarking, and in [MD11] for Langevin dynamics of spin systems. Such a model will be the topic of Chapters 3 and 5, but, to end this chapter, we briefly discuss the methods introduced in [MN08], where it is shown that the convergence condition for the Monte Carlo Evolution is nearly the same as the one in the case of the transverse field term when considering the Schrödinger dynamics.

2.7.2 Path-Integral Monte Carlo

The path integral Monte Carlo uses the path integral, see (2.33), to obtain a classical system to which the algorithm is applied. As discussed in 2.4, the addition of a transverse field introduces quantum fluctuations which play the role of temperature in SA and the ground state search is carried out by decreasing the energy of that field. The path integral method [Son+97] maps

a $d-$dimensional quantum Ising system to a $(d+1)-$dimensional classical system which can be simulated on a classical computer. The Suzuki-Trotter formula is used to decompose the Hamiltonian and obtain the partition function of the classical system, hence the additional dimension is called the Trotter dimension. In a general sense this method can be used to obtain the expression:

$$Z(t) = \sum_{x \in \mathcal{S}} \exp\left(-\frac{F_0(x)}{T_0} - \frac{F_1(x)}{T_1(t)}\right). \tag{2.33}$$

In this partition function, F_0 is the cost function of a combinatorial optimisation problem we wish to solve and for which $\mathcal{S}$ is the set of all possible configurations. The term F_1 is linked to the kinetic energy and is in the form of one of the transverse field terms in Section 2.6. The quantum fluctuations are tuned according to the annealing schedule, such that T_1 increases with time, while T_0 is chosen to be sufficiently small.

The partition function is the normalisation parameter in the Boltzmann distribution, which is the equilibrium distribution for a fixed T_1. With that in mind, we can construct a Monte Carlo scheme with acceptance probability A as follows:

$$A(y, x; t) = g\left(\frac{q(y;t)}{q(x;t)}\right), \tag{2.34}$$

$$q(x;t) = \frac{1}{Z(t)} \exp\left(-\frac{F_0(x)}{T_0} - \frac{F_1(x)}{T_1(t)}\right). \tag{2.35}$$

It can be shown that path integral Monte-Carlo QA is strongly ergodic. Define the set of local maximum states of F_1 as:

$$\mathcal{S}_m = \{x | x \in \mathcal{S}, \forall y \in \mathcal{S}_x, F_1(y) \leq F_1(x)\}. \tag{2.36}$$

Define a metric $d(x, y)$ on $\mathcal{S}$ as the minimum number of steps needed to transition from x to y. Then:

$$R = \min_{(x,y)} \left\{ \max_x \{d(y, x) \mid y \in \mathcal{S}\} \;\middle|\; x \in \mathcal{S} \setminus \mathcal{S}_m \right\} \tag{2.37}$$

is the minimum of the maximal distances between two states in $\mathcal{S} \backslash \mathcal{S}_m$. That is, for a state x, find the farthest other state in $\mathcal{S} \backslash \mathcal{S}_m$ under the metric d, then take the minimum over all such pairs. Finally, in a given step, let L_0 and L_1 denote the largest possible change in F_0 and F_1, respectively. That is, let:

$$L_0 = \max\left\{|F_0(x) - F_0(y)| \middle| P(y, x) > 0,\ x, y \in \mathcal{S}\right\} \tag{2.38}$$

$$L_1 = \max\left\{|F_1(x) - F_1(y)| \middle| P(y, x) > 0,\ x, y \in \mathcal{S}\right\}. \tag{2.39}$$

Then the ergodicity can be stated as in the following theorem:

Theorem 2.2 (Morita and Nishimori(2008) [MN08])**.** *The inhomogeneous Markov chain generated by* (2.34) *is strongly ergodic and converges to the equilibrium state corresponding to the first term of the right-hand side of* (2.35)*, if*

$$T_1(t) \geq \frac{RL_1}{\log(t+2)}. \tag{2.40}$$

It is possible to generalise the acceptance probability (2.34), but this is not proven to be strongly ergodic. Another Monte Carlo scheme for the Schrödinger equation in imaginary time is introduced in [MN08] using Green's Function to calculate the partition function, but it is exceptionally complicated. Indeed, while Monte Carlo schemes have proven effective in simulating the behaviour of Hamiltonian systems, it is difficult to relate the time-step to real-time evolution. As discussed in Section 1.5, this is particularly important in studying phase transi-

tions. It is with this in mind that the stochastic Landau-Lifschitz (LL) equation is introduced as a candidate model for the time evolution of annealing dynamics. The LL equation is a differential equation describing the procession of spins in a magnetic material. It was introduced by Landau and Lifschitz in [LL92]. The addition of the stochastic term produces Langevin's equation for the dynamics of a spin system. In [MD11], Ma and Dudarev integrate this model with a stochastic term to numerically analyse the relaxation dynamics of a spin ensemble as it evolves towards the Gibbs distribution. In the next section, we introduce a similar scheme with a different splitting, a modified version of the solution to the stochastic component and an annealing parameter to simulate the annealing dynamics of a quantum computer.

3 The implementation of SA for the Heisenberg system

3.1 The Heisenberg model

For a system of correlated spins in the presence of an external field and where we consider the interactions between individual components of a single spin, we have the following Hamiltonian:

$$\mathcal{H}(S_1, ..., S_N) = -\frac{1}{2}\sum_{i=1}^{N} S_i \cdot DS_i - \frac{1}{2}\left(\sum_{(i,j)\in\mathcal{I}} J_{ij} S_i \cdot S_j\right) - \sum_{i=1}^{N} h_i \cdot S_i. \tag{3.1}$$

Here, the diagonal matrix $D = \mathrm{diag}(d_1, d_2, d_3) \in \mathbb{R}^{3\times 3}$ accounts for anisotropy, a phenomenon explained in the context of the Heisenberg model in e.g. [Sel+09]. Additionally, we let $|d_1| \leq |d_2| \leq |d_3|$, as in [FHL97]. Moreover, the second term in the definition above accounts for correlations between pairs (i, j) which are connected by an edge. Here, the constants J_{ij} represent the exchange coupling, so that $J_{ij} > 0$ gives ferromagnetic and $J_{ij} < 0$ gives anti-ferromagnetic behaviour of the system of spins. Finally, $h_i = \left(h_i^x, \quad h_i^y, \quad h_i^z\right)^T \in \mathbb{R}^3$ denotes an external magnetic field applied to the spin S_i.

The Hamiltonian (3.1) is usually referred to in the literature as the anisotropic (anti)-/ferromagnetic Heisenberg Hamiltonian, see e.g. [AL04]. Our choice for this particular type of Hamiltonian is guided by a number of papers where it was treated in connection to numerically solving the Landau-Lifschitz (LL) equation, see [AL04], [FHL97]. We keep the customary terminology from the papers we have studied and, as such, for fixed $i = 1, 2, ..., N$, we take a spin $S_i = \left(S_i^x, \quad S_i^y, \quad S_i^z\right)^T \in \mathbb{R}^3$ such that $\|S_i\| = 1$, i.e. it lies on the surface of the unit sphere.

For $N = 1$, we disregard the coupling term, and the Hamiltonian (3.1) reduces to the following single-spin Hamiltonian:

$$\mathcal{H}(S) = -\frac{1}{2} S \cdot DS - h \cdot S. \tag{3.2}$$

For $N > 1$, we neglect the anisotropy term and consider spins on the vertices of a weighted graph $G = (S, \mathcal{I})$, where $S = (S_1, S_2, \ldots, S_N)$, $\mathcal{I} = \{(i, j) : i \sim j\}$ and J_{ij} is the weight associated with edge (i, j), so that the Hamiltonian (3.1) becomes:

$$\mathcal{H}(S_1, ..., S_N) = -\frac{1}{2}\left(\sum_{(i,j)\in\mathcal{I}} J_{ij} S_i \cdot S_j\right) - \sum_{i=1}^{N} h_i \cdot S_i. \tag{3.3}$$

3.2 Langevin Dynamics of Spin System

The Langevin equation for the motion of a single spin S_i reads:

$$\mathrm{d}S_i = \underbrace{S_i \times \nabla_{S_i}\mathcal{H}\mathrm{d}t}_{\text{Conservative dynamics}} + \underbrace{\gamma S_i \times (S_i \times \nabla_{S_i}\mathcal{H})\,\mathrm{d}t}_{\text{Damping Term}} + \underbrace{\sqrt{\eta} S_i \times \mathrm{d}W_i}_{\text{Wiener noise}}. \tag{3.4}$$

This is a stochastic differential equation (SDE) which consists of three individual parts that we discuss separately.

Firstly, the conservative motion of the spin is given by $\mathrm{d}S_i = S_i \times \nabla_{S_i}\mathcal{H}\mathrm{d}t$, which simply describes the precession of the spin around an applied magnetic field, where the orientation of the spin is induced by the magnetic field.

Secondly, the damping term forces the spin into alignment with the magnetic field. A different damping parameter was derived phenomenologically by Gilbert in [Gil04] to make the equation more consistent with other theories and take large damping into account. The Gilbert damping term has identical γ, but introduces an additional parameter in front of the conser-

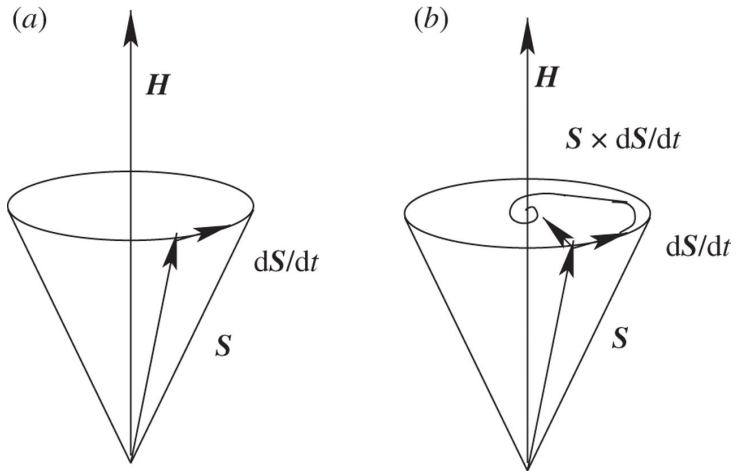


Figure 3.1: (a) Conservative spin motion (b) Spin motion affected by damping [Lak11].

vative term $\gamma^* = \gamma(1 + \alpha^2)$, which proportionately increases the rate of precession with the experimentally derived parameter α. In [Gil04], the value $\alpha = 0.5$ is mentioned for a ferromagnetic material. All our simulations are performed with the original Landau-Lifschitz damping. This behaviour is illustrated in Figure 3.1.

Hence, only considering these first two terms, we can write the equation of motion for one individual spin as:

$$\dot{S}_i = S_i \times \nabla_{S_i}\mathcal{H} + \gamma S_i \times (S_i \times \nabla_{S_i}\mathcal{H}), \tag{3.5}$$

which is the LL equation that was introduced as a model for the magnetisation of spins in a continuous setting [Lak11].

Besides dissipation, the dynamics of magnetisation also has to consider thermal fluctuation which influences the direction of the particle introduced in a thermal bath. A consequence of the theory of Brownian motion shows that a particle exchanges energy with its thermal bath, which causes the particle to deviate from its thermal equilibrium [Ell+15]. This idea was proposed by Brown [Bro63] and consists of randomness added to the LL equation through the term $\sqrt{\eta}S_i \times \mathrm{d}W_i$. Here $\mathrm{d}W_i$ is a Wiener process induced by the three-dimensional Gaussian white noise vector $W_i = \left(W_i^x, \quad W_i^y, \quad W_i^z\right)^T \in \mathbb{R}^3$ defined by:

$$\langle W_i \rangle = 0, \tag{3.6}$$

and

$$\langle W_i^\alpha(t), W_j^\beta(t') \rangle = 2\eta\delta_{ij}\delta_{\alpha\beta}\delta(t - t'), \tag{3.7}$$

where α, β indicate the Cartesian coordinates of the vector and i, j label two arbitrarily chosen spins. If we let T denote the temperature of the thermal bath, we can write the fluctuation-dissipation relation [MD11] which relates γ, T and η. That is,

$$\eta = \gamma\hbar k_B T, \tag{3.8}$$

where $\hbar$ and k_B are Planck's and Boltzmann's constants, respectively. The fluctuation-dissipation relation captures the intuitive notion that friction induces thermal fluctuations caused by the production of heat.

A system of N spins leads to a system of N coupled equations of the form (3.4) that must be solved simultaneously. The solutions of these equations for the Hamiltonian (3.3) result in the trajectories for a systems of correlated spins in the presence of an external field. By solving the system, we obtain a continuous time classical model for the dynamics of a spin system. We can introduce an annealing schedule into the model to investigate the dynamics of a quantum

annealer. We expect that if the field is perturbed slowly enough, as expressed in the adiabatic theorem in 2.6.1, then the stationary distribution of the Langevin equation will correspond to the ground state of the Hamiltonian, see Figure 3.2. Thus, we proceed to describing our approach to solving Eq. (3.4) for the Hamiltonians (3.2), and (3.3) - namely, for the single-spin and the system-of-spins cases.

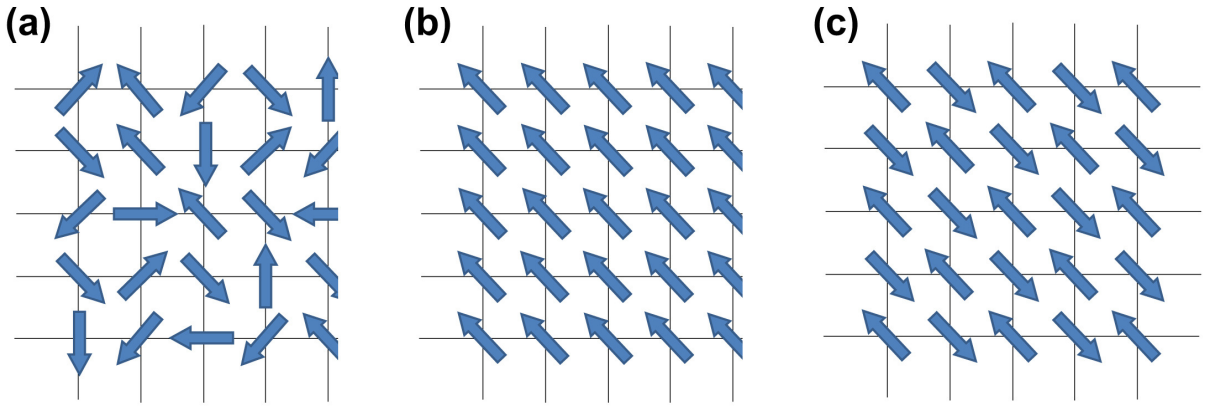


Figure 3.2: Typical spin configurations in the ordered state of $\mathcal{H}$: (a) The paramagnetic state (mixture of ferromagnetic and antiferromagnetic coupling terms), (b) The ferromagnetic state ($J > 0$), and (c) The antiferromagnetic state ($J < 0$), [KT15].

3.3 Splitting methods

In this section, we motivate the use of splitting methods over classical numerical methods, then briefly explain how they can be used both generally and in the current context, and finally illustrate how we apply them to obtain an approximate solution to Eq. (3.4).

3.3.1 Motivation

In physical applications, preserving one or more of the "geometric" properties of the governing equation of motion is of crucial importance. Examples of such properties include preservation of energy, momentum, phase space volume, time-reversal symmetry, symplectic structure and dissipation. To achieve this, an individual class of numerical methods was developed, namely geometric integrators, many of them tailored to problems analogous in nature to the one considered in this report (see, for example, [FHL97; AL04; TKL04]).

To begin with, we illustrate the advantage of the above integrators over non-geometric numerical methods. Similar to work carried out by Leimkuhler and Matthews in [LM16], we consider as an example the one-dimensional harmonic oscillator, governed by the following system of ordinary differential equations (ODEs):

$$\begin{cases} \dot{p} &= -q \\ \dot{q} &= p, \end{cases} \tag{3.9}$$

where $p, q : \mathbb{R} \to \mathbb{R}$. Further, the energy of the harmonic oscillator is described by the following autonomous Hamiltonian:

$$\mathcal{H} = \frac{p^2 + q^2}{2}, \tag{3.10}$$

which is conserved along exact trajectories. Taking a set of initial conditions which form a circle of radius 0.2 and centered at $(p, q) = (1, 0)$, we can apply both non-structure-preserving numerical methods, such as explicit and implicit Euler [But16], as well as geometric integrators, such as symplectic and centered symplectic Euler [FQ87], to simulate a long-time evolution.

In this framework, Figure 3.3a shows the evolution of the disc delimited by the initial conditions computed using explicit and implicit Euler. Clearly, the area of the disc is not preserved over time. In particular, the increase created by explicit Euler can be associated with a numerical production of energy, whereas the reduction caused by implicit Euler can be interpreted as a numerical dissipation of energy. Therefore, these methods do not preserve an underlying property of the system, as they modify the area delimited by the initial conditions.

On the other hand, Figure 3.3b illustrates the solution of the system (3.9) computed using the symplectic and centered symplectic Euler schemes, starting from the same initial conditions as in the previous simulation. Thus, we notice that even if the disc is deformed by the former method, the area delimited by the circle is still invariant. Moreover, using the latter approach, the shape is unmodified and the area is preserved.

Hence, the harmonic oscillator is an example which motivates the use of geometric integrators for numerically solving differential equations which exhibit intrinsic geometric properties.

3.3.2 Some geometric properties of the LLG equation

Having motivated the use of geometric integrators, we now explore some of the underlying geometric properties of Eq. (3.4).

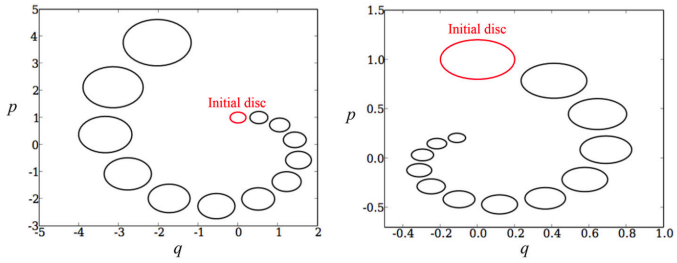
To begin with, one of the essential features that an integrator should preserve is the norm of the spin, namely $\|S_i\| = 1$, $i = 1, \dots, N$, where N is the number of spins in the system.

Secondly, important, but not essential, is the Poisson structure of the deterministic part of Eq. (3.4), specifically of Eq. (3.5). In fact, this property is not always preserved in our integrators.

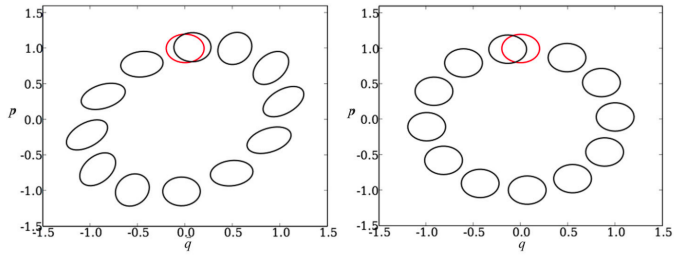
Finally, let us consider the stochastic part of the Eq. (3.4) for a single spin, in other words:

$$\mathrm{d}S = \sqrt{\eta} S \times \mathrm{d}W. \tag{3.11}$$

To rewrite the RHS of this equation, note that for two vectors $a = \begin{pmatrix} a_1, & a_2, & a_3 \end{pmatrix}^T \in \mathbb{R}^3$ and



(a) Evolution of disc determined by explicit Euler (left) and implicit Euler (right)



(b) Evolution of disc given by symplectic Euler (left) and centered symplectic Euler (right)

Figure 3.3: Comparison of geometric and non-geometric integrators [RHC18].

$b = \begin{pmatrix} b_1, & b_2, & b_3 \end{pmatrix}^T \in \mathbb{R}^3$, their cross product can be written as:

$$a \times b = \begin{pmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = -\text{skew}(a)b, \tag{3.12}$$

where we have defined:

$$\text{skew}(a) \coloneqq \begin{pmatrix} 0 & a_3 & -a_2 \\ -a_3 & 0 & a_1 \\ a_2 & -a_1 & 0 \end{pmatrix}. \tag{3.13}$$

Since the cross product is anti-commutative, i.e. $a \times b = -b \times a$, Eq. (3.11) can be rewritten as:

$$\mathrm{d}S = \sqrt{\eta} S \times \mathrm{d}W = -\sqrt{\eta} \mathrm{d}W \times S = \sqrt{\eta}\ \text{skew}(\mathrm{d}W) S, \tag{3.14}$$

where the last equality stems from the linearity of the skew-symmetric operator.

Heuristically, we can think of Eq. (3.14) as a SDE at the $\mathfrak{so}(n)$ Lie algebra level because of the skew$(\mathrm{d}W(t))$ term. To motivate why this is the case, let us now consider the special orthogonal group $\mathrm{SO}(n)$ defined by:

$$\mathrm{SO}(n) = \left\{ A \in \mathbb{R}^{n \times n} \middle| AA^T = A^T A = I_n,\ \det(A) = 1 \right\}, \tag{3.15}$$

whose elements can be interpreted geometrically as rotations. It can be proven that $\mathrm{SO}(n)$ is a Lie group, see e.g. [Hal03]. Then, we can compute the Lie algebra $\mathfrak{so}(n)$ of the special orthogonal group $\mathrm{SO}(n)$, which illustrates that the presence of skew-symmetric matrices in the definition of the cross product is expected. To achieve this, we give the following definitions:

Definition 3.1 ([Hal03])**.** *The Lie algebra of a matrix Lie group G is the set of all matrices X such that e^{tX} is in G for all $t \in \mathbb{R}$.*

Proposition 3.2 ([Hal03])**.** *The Lie algebra of $SO(n)$ consists of all real matrices A such that $A^T = -A$, i.e. of all real skew-symmetric matrices.*

Proof. A real matrix A belongs to $\mathrm{SO}(n)$ if and only if $A^T = A^{-1}$ and $\det(A) = 1$. Thus, e^{tA} is in $\mathrm{SO}(n)$ if and only if $\left(e^{tA}\right)^T = \left(e^{tA}\right)^{-1}$ and $\det\left(e^{tA}\right) = 1$. Using the properties of the matrix exponential, the first equality can be written further as $e^{tA^T} = e^{-tA}$ for all $t \in \mathbb{R}$. So this holds if and only if $A^T = -A$. Using Theorem 2.12 from [Hal03], the second equality from earlier becomes $\det\left(e^{tA}\right) = e^{\text{trace}(tA)} = e^{\text{trace}(A)t} = 1$ for all $t \in \mathbb{R}$. Hence, $\text{trace}(A) = 0$. But this second condition does not add anything new to the Lie algebra because we can go from $A^T = -A$ by taking trace on both sides and find that A has to have null trace. Definition (3.1) concludes the proof. □

We have given here a brief description of some of the geometric properties of Eq. (3.4). In the following section, we study how methods that preserve these properties can be developed.

3.3.3 Building Blocks

A subclass of geometric integrators is the collection of symplectic integrators. These can be built, for example, by considering Runge-Kutta methods and checking whether the coefficients automatically preserve the symplectic property. Another more straightforward approach hinges on splitting techniques, also known as Lie-Trotter methods [LM16].

The underlying idea of splitting schemes is that the vector field associated with the differential equation under consideration can be decomposed into several integrable parts. Subsequently, one determines the flow maps (or, sometimes, approximate flow maps) for each of the individual pieces and composes them to obtain an approximation of the solution to the original equation.

Consider, for example, the following simple ODE:

$$\dot{y} = f = f^1 + f^2, \tag{3.16}$$

so that the original vector field f is split into analytically integrable parts f^1 and f^2. Suppose that the flow maps for the equations:

$$\dot{y} = f^1, \quad \dot{y} = f^2, \tag{3.17}$$

are given by ϕ_h^1 and ϕ_h^2, respectively. Then the following two maps:

$$\hat{\phi}_h^{1,2} = \phi_h^1 \circ \phi_h^2 \tag{3.18}$$

$$\tilde{\phi}_h^{2,1} = \phi_h^2 \circ \phi_h^1 \tag{3.19}$$

are approximations of the flow map ϕ_h corresponding to the original ODE. Moreover, Leimkuhler and Matthews [LM16] state that any part in the splitting may be replaced by an approximation of the flow map, rather than the flow map itself. For example,

$$\bar{\phi}_h^{1,2} = \phi_{h/2}^1 \circ \phi_{h/2}^2 \tag{3.20}$$

is another symplectic numerical method.

In terms of convergence, it is shown in [LM16] that if f_1 is C^2, then the approximation (3.18) is second-order accurate, or equivalently:

$$\|\phi_h - \hat{\phi}_h^{1,2}\| \leq Ch^2. \tag{3.21}$$

In addition, they also illustrate that *symmetric* compositions have even convergence order, a fact which we make use of in our own integrators.

Splitting methods applied to ODEs can be extended to problems which involve Hamiltonians describing the energy for a system of spins. In fact, such an integrator is developed in [FHL97] for Eq. (3.5) with the following Hamiltonian:

$$\mathcal{H} = -\sum_i S_i \cdot S_{i+1}. \tag{3.22}$$

The first algorithms described in [FHL97] involve splitting $\mathcal{H} = \mathcal{H}_1 + \mathcal{H}_2$, where:

$$\mathcal{H}_1 = -\sum_i S_i \cdot S_{i+1}, \quad \mathcal{H}_2 = -\sum_i S_i \cdot S_{i-1}. \tag{3.23}$$

Another approach proposed in [FHL97] involves splitting the vector field on the RHS of Eq. (3.5) into odd- and even-spins interactions, known as the *staggered Red-Black method*.

Splitting methods can be broadened further to consider stochastic differential equations (SDEs). Such an example is illustrated in [LM16], where multiple geometric integrators for the following Langevin equation are developed:

$$\mathrm{d}q = M^{-1} p \mathrm{d}t \tag{3.24}$$

$$\mathrm{d}p = -\nabla U(q)\mathrm{d}t - \gamma p \mathrm{d}t + \sqrt{2\gamma k_B T} M^{1/2} \mathrm{d}W. \tag{3.25}$$

This is achieved by dividing the system (3.24) in three parts:

$$\mathrm{d}\begin{pmatrix} q \\ p \end{pmatrix} = \begin{pmatrix} M^{-1} \\ 0 \end{pmatrix} \mathrm{d}t + \begin{pmatrix} 0 \\ -\nabla U(q) \end{pmatrix} \mathrm{d}t + \begin{pmatrix} 0 \\ -\gamma p \mathrm{d}t + \sqrt{2\gamma k_B T} M^{1/2} \mathrm{d}W \end{pmatrix}. \tag{3.26}$$

The first two terms in the splitting can be integrated analytically, while the third is the Ornstein–Uhlenbeck process, with an established solution [Gar+85]. As explained in [LM16], different composition of the three maps give approximations with different orders of convergence.

In addition, an integration algorithm for Eq. (3.4) based on splittings was proposed in

[MD11]. In particular, the scheme uses the Suzuki-Trotter decomposition [HS05] whose symplectic structure is essential in this context. Their approach is to consider Eq. (3.4) for each spin S_i individually, and solve it for a generic Hamiltonian $\mathcal{H}$ independent of time. To achieve this, Ma and Dudarev split Eq. (3.4) into its deterministic and stochastic parts, as follows:

$$\mathrm{d}S_i = S_i \times \nabla_{S_i}\mathcal{H}\mathrm{d}t + \gamma S_i \times (S_i \times \nabla_{S_i}\mathcal{H})\,\mathrm{d}t \tag{3.27}$$

$$\mathrm{d}S_i = \sqrt{\eta} S_i \times \mathrm{d}W_i, \tag{3.28}$$

and analytically solve them. The solution to Eq. (3.27) involves switching to polar coordinates, analytically integrating the resulting equations and then switching back to Cartesian coordinates. We were not successful in applying the same procedure to our time-dependent Hamiltonians (3.2) and (3.3), and hence, resorted to a different splitting-composition method. Moreover, [MD11] does not provide a technique for solving Eq. (3.28). Since we also keep the stochastic term isolated in our schemes, we derive a solution to Eq. (3.28) ourselves, which is, to some extent, different from the one proposed in [MD11].

We do not dwell on the underlying details of the preceding methods. Rather, we stress that they establish a theoretical foundation for extending symplectic integrators for ODEs to splitting methods for systems of spins and SDEs, which is precisely the subject of this chapter.

3.3.4 A proposed method for the Single-Spin Case

We now introduce our approach to solving the Langevin equation (3.4) using splitting methods. For both of our choices of Hamiltonian (3.2) and (3.3), Eq. (3.4) is, to the best of our knowledge, not analytically integrable. We propose in this report two different splittings for each of the two Hamiltonians, as we chronologically established and implemented them.

We first focus on the single-spin Hamiltonian (3.2), whose gradient with respect to S is:

$$\nabla_S \mathcal{H} = -DS - h. \tag{3.29}$$

One technique of splitting equation (3.4) rests on the scheme developed for the rigid body Hamiltonian and the Euler equation in [LM16]. In particular, we first use the form of the Hamiltonian (3.2), or equivalently the form of its gradient (3.29), to separate the two terms $1/2\, S \cdot DS$ and $h \cdot S$. Upon substitution, the equation for the quadratic term becomes:

$$\mathrm{d}S = -S \times DS\mathrm{d}t - \gamma S \times (S \times DS)\mathrm{d}t + \sqrt{\eta} S \times \mathrm{d}W. \tag{3.30}$$

The equation for the linear term is analogous, with DS replaced by h above.

These can be split further by considering each column corresponding to the diagonal entries in the matrix $D = \mathrm{diag}(d_1, d_2, d_3)$ and each entry in the vector $\begin{pmatrix} h^x, & h^y, & h^z \end{pmatrix}^T$ individually, so that the gradient (3.29) becomes:

$$\nabla_S \mathcal{H} = -D_1 S - D_2 S - D_3 S - h_1 - h_2 - h_3, \tag{3.31}$$

where, for example, $D_1 = \begin{pmatrix} d_1, & 0, & 0 \end{pmatrix}^T$ and $h_1 = \begin{pmatrix} h^x, & 0, & 0 \end{pmatrix}^T$. Note that this second splitting step only affects the deterministic terms of Eq. (3.30).

To simplify the problem even more, we split the resulting equation again into its three constituting parts: conservative, dissipative and stochastic. This third splitting step leads to the following three equations, which we can now solve analytically:

$$\mathrm{d}S = -S \times D_1 S\mathrm{d}t \tag{3.32}$$

$$\mathrm{d}S = -\gamma S \times (S \times D_1 S)\,\mathrm{d}t \tag{3.33}$$

$$\mathrm{d}S = \sqrt{\eta} S \times \mathrm{d}W. \tag{3.34}$$

It is worth mentioning that one need not perform the last splitting, since Eq. (3.32) and Eq. (3.33) can be analytically integrated together, as we show later in this chapter. In addition, we note that it suffices to solve Eq. (3.32) and Eq. (3.33) for D_1, since the solutions for the other terms in the splitting, i.e. corresponding to D_2, D_3, and h_1, h_2, h_3, are analogous.

Hence, we now turn our attention to Eq. (3.32). After computing the cross product $S \times D_1 S$, we obtain the following system of equations:

$$\begin{pmatrix} \dot{S}^x \\ \dot{S}^y \\ \dot{S}^z \end{pmatrix} = \begin{pmatrix} 0 \\ -d_1 S^z S^x \\ d_1 S^y S^x \end{pmatrix}. \tag{3.35}$$

This is a system that we can now easily solve.

Clearly, the first equation gives $S^x(t) = S^x(0)$, while the other two equations constitute, upon substituting S^x, a coupled system of first order ODEs, which has the form of the harmonic oscillator system (3.9). One way to solve this system is by differentiating one of the equations again, and then substituting the other, leading to a second order linear ODE, as follows:

$$\ddot{S}^y = -S^x(0) d_1 \dot{S}^z = -\left(S^x(0) d_1\right)^2 S^y. \tag{3.36}$$

Using the characteristic equation and imposing the initial conditions on $S^y(0)$ and $\dot{S}^y(0)$, we obtain the following solution:

$$S^y(t) = S^z(0) \sin\left(d_1 S^x(0) t\right) + S^y(0) \cos\left(d_1 S^x(0) t\right). \tag{3.37}$$

The full derivation of this solution is given in Appendix A.1. Moreover, the solution S^z can be found analogously to be:

$$S^z(t) = S^y(0) \sin\left(d_1 S^x(0) t\right) - S^z(0) \cos\left(d_1 S^x(0) t\right). \tag{3.38}$$

This gives a complete solution S to Eq. (3.32).

We now focus on Eq. (3.33). Here, the double cross product $\gamma S \times (S \times D_1 S)$ yields:

$$\begin{pmatrix} \dot{S}^x \\ \dot{S}^y \\ \dot{S}^z \end{pmatrix} = \begin{pmatrix} \gamma d_1 S^x (S^y)^2 + d_1 S^x (S^z)^2 \\ -\gamma d_1 S^y (S^x)^2 \\ -\gamma d_1 S^z (S^x)^2 \end{pmatrix}. \tag{3.39}$$

Multiplying each equation by $2S^x$, $2S^y$ and $2S^z$, respectively, and letting $u^x \coloneqq (S^x)^2$, $u^y \coloneqq (S^y)^2$, and $u^z \coloneqq (S^z)^2$, we obtain the following equivalent system of equations:

$$\begin{pmatrix} \dot{u}^x \\ \dot{u}^y \\ \dot{u}^z \end{pmatrix} = \begin{pmatrix} 2\gamma d_1 u^x (u^y + u^z) \\ -2\gamma d_1 u^y u^x \\ -2\gamma d_1 u^z u^x \end{pmatrix}. \tag{3.40}$$

Note that, using $\|S\| = 1 \iff u^x + u^y + u^z = 1$, the first equation becomes:

$$\dot{u}^x = 2\gamma d_1 u^x \left(1 - u^x\right). \tag{3.41}$$

This is now a logistic differential equation, which can be solved, for example, by using partial fractions, as shown in Appendix A.2. The solution is:

$$u^x(t) = \frac{R \exp(2\gamma d_1 t)}{1 + R \exp(2\gamma d_1 t)}, \tag{3.42}$$

where $R = u^x(0)/(1 - u^x(0))$. This then gives $S^x(t) = \pm\sqrt{u^x(t)}$.

Next, to obtain a solution for u^y and u^z, we return to Eq. (3.40) and note that we need only integrate u^x, since, for example, the equation for u^y gives:

$$\int_{u^y(0)}^{u^y} \frac{1}{v}\mathrm{d}v = -2\gamma d_1 \int_0^t u^x \mathrm{d}t' = -2\gamma d_1 \int_0^t \frac{R\exp(2\gamma d_1 t)}{1+R\exp(2\gamma d_1 t)}\mathrm{d}t', \tag{3.43}$$

The integral on the RHS can be computed by substituting $w = 1 + R\exp(2\gamma d_1 t)$, as shown in Appendix A.2. The solution $u^y(t)$ becomes:

$$u^y(t) = u^y(0)\frac{1+R}{1+R\exp(2\gamma d_1 t)}. \tag{3.44}$$

This yields the solution $S^y(t) = \pm\sqrt{u^y(t)}$. Analogously, we find:

$$u^z(t) = u^z(0)\frac{1+R}{1+R\exp(2\gamma d_1 t)}, \tag{3.45}$$

so that $S^z = \pm\sqrt{u^z(t)}$. This completes the solution to Eq. (3.33).

Finally, we derive an analytical solution for Eq. (3.34). Since Eq. (3.14) is a SDE at the $\mathfrak{so}(n)$ Lie algebra level, as motivated in Section 3.3.2, the solution $S(t)$ is to be found at the group level. More specifically, suppose $X(t) \in \mathrm{SO}(n)$, for all $t \in \mathbb{R}$. If we write $X(t) = X_0 e^{\Theta(t)}$, where $X_0 \coloneqq X(0)$, then since $X(t)^T X(t) = I_n$, it follows that $e^{\Theta(t)^T} X_0^T X_0 e^{\Theta(t)} = I_n$, and so $e^{\Theta(t)^T} e^{\Theta(t)} = I_n$, or in other words $e^{\Theta(t)^T} = e^{-\Theta(t)}$, for all $t \in \mathbb{R}$, which gives $\Theta^T = -\Theta$.

This suggests the ansatz $S(t) = S(0)R(t)$, where $R(t)$ is an element of SO(3). Differentiating this with respect to t and substituting the result in Eq. (3.14) gives:

$$R'(t)S(0)\mathrm{d}t = \sqrt{\eta}\ \mathrm{skew}(\mathrm{d}W(t))R(t)S(0). \tag{3.46}$$

This holds for all $S(0)$ assumed to be non-zero. Multiplying both sides by $R^{-1}(t)$, integrating and exponentiating yields, after some manipulation:

$$S(t) = S(0)e^{\sqrt{\eta}\sqrt{t}\ \mathrm{skew}(W(1))}. \tag{3.47}$$

Now, suppose we have $\omega = \begin{pmatrix}\omega_1, & \omega_2, & \omega_3\end{pmatrix}^T$, where each component is a standard Gaussian random variable such that $\|\omega\| = 1$. Then:

$$\Omega \coloneqq \mathrm{skew}(\omega) = \begin{pmatrix} 0 & \omega_z & -\omega_y \\ -\omega_z & 0 & \omega_x \\ \omega_y & -\omega_x & 0 \end{pmatrix}. \tag{3.48}$$

Performing the usual matrix multiplication, we observe that $\Omega^3 = -\Omega$. Hence, by the general formula for the matrix exponential [Hal03], upon simplification, we have that:

$$e^{\sqrt{\eta t}\Omega} = I_3 + \sin(\sqrt{\eta t})\Omega + \left(1 - \cos(\sqrt{\eta t})\right)\Omega^2. \tag{3.49}$$

Therefore, the solution to Eq. (3.34) is:

$$S(t) = S(0)\left(I_3 + \sin(\sqrt{\eta t})\Omega + \left(1 - \cos(\sqrt{\eta t})\right)\Omega^2\right). \tag{3.50}$$

The full derivation of the procedure described above is given in Appendix A.3.

We can now obtain an approximation to the solution of Eq. (3.4) for the Hamiltonian (3.2) by composing the solutions we obtained for each of the terms in the splitting. We sacrifice computational effort in favour of accuracy and only using symmetric compositions in our

implementations. Hence, one such possible approximation can be expressed as:

$$\begin{aligned}\hat{S}^C_{h/2} &= S^{C,D_1}_{h/2} \circ S^{C,D_2}_{h/2} \circ S^{C,D_3}_{h/2} \circ S^{C,c_1}_{h/2} \circ S^{C,c_2}_{h/2} \circ S^{C,c_3}_{h/2} \\ \hat{S}^D_{h/2} &= S^{D,D_1}_{h/2} \circ \cdots \circ S^{D,c_3}_{h/2} \\ \hat{S}_h &= \hat{S}^C_{h/2} \circ \hat{S}^D_{h/2} \circ S^S_h \circ \left(\hat{S}^D_{h/2}\right)^{-1} \circ \left(\hat{S}^C_{h/2}\right)^{-1},\end{aligned} \tag{3.51}$$

where, S^C, S^D and S^S represent the solutions to the conservative, dissipation and stochastic equations (3.32), (3.33) and (3.34), respectively. Further, S^{D_1} denotes, for example, the solution to the equation corresponding to the D_1 part of the Hamiltonian, while $S_{h/2}$ indicates that the solution is evaluated at $t = h/2$. Here, $(\cdot)^{-1}$ stands for the reverse composition of maps.

This completes the description of our approach to constructing an approximate solution to Eq. (3.4) for the single-spin Hamiltonian (3.2) by exploiting splitting methods. In the subsequent section, we propose a different integrator for the system-of-spins Hamiltonian (3.3).

3.3.5 A proposed method for the System-of-Spins Case

Recall from the end of Section 3.3.3 that the method proposed by Ma and Dudarev [MD11] treats Eq. (3.4) for each spin S_i, $i = 1, \dots N$, independently. In contrast, we divide pairwise interactions of coupled spins S_i-S_j, and solve the resulting equation as outlined below.

First, note that the gradient of the Hamiltonian (3.3) with respect to one spin S_i is:

$$\nabla_{S_i}\mathcal{H} = -\frac{1}{2}\sum_{j:i\sim j} J_{ij}S_j - h_i, \tag{3.52}$$

where $i \sim j$ indicates that we only sum over the spins which are coupled with S_i.

As with the one-spin case, we simplify the problem by separating the interaction and external-field terms in the expression of the gradient (3.52). Upon substitution, for the h_i term and for one spin S_i, Eq. (3.4) simplifies greatly to:

$$\mathrm{d}S_i = -S_i \times h_i \mathrm{d}t - \gamma S_i \times (S_i \times h_i)\mathrm{d}t + \sqrt{\eta} S_i \times \mathrm{d}W_i, \tag{3.53}$$

which is identical to equation corresponding to the h term in the one-spin case.

For the interaction term, we split the Hamiltonian further by considering only pairwise couplings. In particular, we break the sum in Eq. (3.52) and treat each term individually. Consequently, for an arbitrary interaction S_i-S_j, Eq. (3.4) becomes:

$$\mathrm{d}S_i = -\frac{1}{2}J_{ij}S_i \times S_j \mathrm{d}t - \frac{1}{2}J_{ij}\gamma S_i \times (S_i \times S_j)\mathrm{d}t + \sqrt{\eta} S_i \times \mathrm{d}W_i. \tag{3.54}$$

Further, observe that the stochastic term in Eq. (3.53) and (3.54) is independent of h_i and S_j, respectively. Therefore, we can isolate it, and simply consider the equation:

$$\mathrm{d}S_i = \sqrt{\eta} S_i \times \mathrm{d}W_i, \tag{3.55}$$

for each spin S_i, $i = 1, \dots, N$. Notice that this equation is exactly the stochastic equation (3.34) for the one-spin case. Thus, we can use the solution defined in Eq. (3.50) for each spin S_i as an element of our splitting-composition method.

Now we need only solve the deterministic unit in the aforementioned equations. To this end, we can rewrite Eq. (3.54) by removing the stochastic term, both for an arbitrary interaction

S_i-S_j, and for its corresponding S_j-S_i reverse coupling. This yields the following coupled system:

$$\begin{cases} \dot{S}_i &= -\frac{1}{2}J_{ij}S_i \times S_j - \frac{1}{2}J_{ij}\gamma S_i \times (S_i \times S_j) \\ \dot{S}_j &= -\frac{1}{2}J_{ji}S_j \times S_i - \frac{1}{2}J_{ji}\gamma S_j \times (S_j \times S_i), \end{cases} \tag{3.56}$$

where $J_{ij} = J_{ji}$. We highlight here that we believe that this system of equations is, in fact, analytically integrable, and we give the partial solution in Appendix A.4. However, we could not derive the full solution, hence we resort to a different approach.

To simplify the problem, we decouple the system (3.56) by successively setting $\dot{S}_i = 0$ and $\dot{S}_j = 0$, so that $S_i(t) = S_i(0)$ and $S_j(t) = S_j(0)$, respectively. In addition, we solely turn our attention to the first case, since the second is, in fact, analogous. By doing so, note that Eq. (3.53) and (3.54) become equivalent, since they both now involve a vector which is independent of time, namely h_i and $S_j(0)$, respectively. Thus, it suffices to solve either of them.

Accordingly, let us consider the deterministic term in Eq. (3.54). Unlike in Section 3.3.4, we do not split this further by treating the conservative and damping terms independently. Rather, we solve the equation by directly separating each entry in the vector $S_j(0)$, yielding:

$$\dot{S}_i = -\frac{1}{2}J_{ij}S_i \times S_j^X(0) - \frac{1}{2}J_{ij}\gamma S_i \times (S_i \times S_j^X(0)), \tag{3.57}$$

where $S_j^X(0) = \left(S_j^x(0), 0, 0\right)^T$. In addition, we obtain two more equations for the $S_j^y(0)$ and $S_j^z(0)$ components. It is now enough to solve Eq. (3.57), as the remaining ones are analogous. We achieve this by first working through the cross product, which gives:

$$\begin{pmatrix} \dot{S}_i^x \\ \dot{S}_i^y \\ \dot{S}_i^z \end{pmatrix} = \begin{pmatrix} \frac{1}{2}J_{ij}\gamma S_j^x(0)(S_i^y)^2 + \frac{1}{2}J_{ij}\gamma S_j^x(0)(S_i^z)^2 \\ -\frac{1}{2}J_{ij}S_i^z S_j^x(0) - \frac{1}{2}J_{ij}\gamma S_i^x S_i^y S_j^x(0) \\ \frac{1}{2}J_{ij}S_i^y S_j^x(0) - \frac{1}{2}J_{ij}\gamma S_i^x S_i^z S_j^x(0) \end{pmatrix}. \tag{3.58}$$

Using the fact that $\|S_i\| = 1$, the first equation becomes:

$$\dot{S}_i^x = \frac{1}{2}J_{ij}\gamma S_j^x(0)\left(1 - (S_i^x)^2\right). \tag{3.59}$$

This can be solved by using partial fractions, as shown in Appendix A.5. The solution is:

$$S_i^x(t) = \frac{M\exp\left(J_{ij}\gamma S_j^x(0)t\right) - 1}{M\exp\left(J_{ij}\gamma S_j^x(0)t\right) + 1}, \tag{3.60}$$

where $M = \left(1 + S_i^x(0)\right)/\left(1 - S_i^x(0)\right)$. Now we need to solve the following coupled system of equations:

$$\begin{cases} \dot{S}_i^y &= -\frac{1}{2}J_{ij}S_i^z S_j^x(0) - \frac{1}{2}J_{ij}\gamma S_i^x S_i^y S_j^x(0) \\ \dot{S}_i^z &= \frac{1}{2}J_{ij}S_i^y S_j^x(0) - \frac{1}{2}J_{ij}\gamma S_i^x S_i^z S_j^x(0). \end{cases} \tag{3.61}$$

Multiplying the first equation by S_i^z, the second equation by $-S_i^y$ and adding them yields:

$$\dot{S}_i^y S_i^z - S_i^y \dot{S}_i^z = -\frac{1}{2}J_{ij}S_j^x(0)\left((S_i^y)^2 + (S_i^z)^2\right). \tag{3.62}$$

Dividing this equation by $(S_i^z)^2$, one could recognise the derivative of the quotient on the left-hand side, and arrive at:

$$\left(\frac{\dot{S}_i^y}{S_i^z}\right) = -\frac{1}{2}J_{ij}S_j^x(0)\left(1 + \left(\frac{S_i^y}{S_i^z}\right)^2\right). \tag{3.63}$$

Setting $v(t) \coloneqq S_i^y(t)/S_i^z(t)$, we obtain an equation which can be solved again using the separation technique, as shown in Appendix A.5. The solution is:

$$\frac{S_i^y}{S_i^z} = \frac{v(0)\cos\left(\frac{1}{2}J_{ij}S_j^x(0)t\right) - \sin\left(\frac{1}{2}J_{ij}S_j^x(0)t\right)}{\cos\left(\frac{1}{2}J_{ij}S_j^x(0)t\right) + v(0)\sin\left(\frac{1}{2}J_{ij}S_j^x(0)t\right)}. \tag{3.64}$$

If we denote the right-hand side of equation (3.64) by $g(t)$ and substitute $S_i^y = g(t)S_i^z$ into the equation for $\dot{S}_i^z$ in (3.61), we find that:

$$\dot{S}_i^z = \frac{1}{2}J_{ij}g(t)S_i^zS_j^x(0) - \frac{1}{2}J_{ij}\gamma S_i^xS_i^zS_j^x(0), \tag{3.65}$$

which becomes and ODE in the S^y variable. We can integrate the right-hand side as described in Appendix A.5, so that the solution $S_i^z(t)$ is given by:

$$S_i^z(t) = S_i^z(0)\frac{1+M}{\exp(-\frac{1}{2}J_{ij}\gamma S_j^x(0)) + M\exp(\frac{1}{2}J_{ij}\gamma S_j^x(0))}\left(\cos\left(\frac{1}{2}J_{ij}S_j^x(0)t\right) + v(0)\sin\left(\frac{1}{2}J_{ij}S_j^x(0)t\right)\right), \tag{3.66}$$

and the solution $S_i^y(t)$ is obtained immediately from equation (3.64).

Finally, to obtain an approximation to the solution of Eq. (3.4) for the Hamiltonian (3.3), we need only compose the maps derived above. We propose a similar procedure to Section 3.3.4, and use symmetric compositions only. Accordingly, one such possible approximation is:

$$\begin{aligned}\hat{S}_{h/2}^i &= \left(\bigcirc_{j:i\sim j}\left(S_{h/2}^{j,x}\circ S_{h/2}^{j,y}\circ S_{h/2}^{j,z}\right)\right)\circ S_{h/2}^{h_i,x}\circ S_{h/2}^{h_i,y}\circ S_{h/2}^{h_i,z}\circ S_h^{S_i}\\ \hat{S}_h &= \bigcirc_{i=1}^{N}\hat{S}_{h/2}^i\circ\bigcirc_{i=N}^{1}\left(\hat{S}_{h/2}^i\right)^{-1},\end{aligned} \tag{3.67}$$

where S^j denotes the solution associated with the specific interaction S_i-S_j. Further, S^α, $\alpha \in \{x, y, z\}$ refers to the solution to the equation corresponding to each component α of $S_j(0)$, while S^{S_i} is the solution to the stochastic equation for spin S_i. Here, $S_{h/2}$ indicates that the solution is evaluated at $t = h/2$, and $(\cdot)^{-1}$ stands for the reverse composition of maps.

Thus, in this section we described how the LL equation (3.4), together with the Hamiltonians (3.2) and (3.3), can be used to model the motion of one spin, as well as of a system of interacting spins, respectively. Additionally, we have described how splitting methods can be applied to the problem at hand, with an emphasis on the two new schemes proposed in this report. This can now be used to simulate quantum annealing, which is the topic of Chapter 5. However, we first give, in the subsequent section, a summary of how annealing is implemented on the D-Wave quantum machine.

4 D-Wave implementation of Quantum Annealing

Having motivated the use of quantum annealers for solving optimisation problems in Section 1.1, we shall now examine how the quantum annealing process is implemented on the publicly available D-Wave's quantum computer, based on their documentation [D-Wf].

As explained in Section 2.1, inside a quantum computer, qubits are devised as circulating currents with a corresponding magnetic field, which is also the case for the D-Wave's quantum machine. Moreover, each qubit has attached to it an external magnetic field, as well as a fixed number of couplers which depends on the specific architecture.

Thus, together, the strength of the external magnetic field and of the coupling, also known in D-wave [D-Wg] as biases and weights, respectively, define the problem which the quantum annealer subsequently solves. In particular, they define an energy landscape, so that at the end of the quantum annealing process the minimum energy of that landscape is returned.

Furthermore, the D-Wave's quantum processing unit (QPU) is a physical realisation of the quantum Ising spin system in a transverse field [D-Wg], so that the time-dependent Hamiltonian, which defines the energy for a problem, is the Ising Hamiltonian [Cip87], given by:

$$\mathcal{H}_{ising} = \underbrace{-\frac{A}{2}\left(\sum_i \sigma_i^x\right)}_{\text{Initial Hamiltonian}} + \underbrace{\frac{B}{2}\left(\sum_i h_i\sigma_i^z + \sum_{i>j} J_{i,j}\sigma_i^z\sigma_j^z\right)}_{\text{Final Hamiltonian}}. \tag{4.1}$$

Here, h_i and $J_{i,j}$ give the sign and magnitude of the the aforementioned external field and coupling strengths, and they are limited to those available in the QPU graph (namely the Pegasus or Chimera graphs [D-Wb]). Additionally, σ_i^x and σ_i^z are Pauli matrices operating on qubit q_i, and the physical implementation of the qubit q_i is given by the one-dimensional Ising spin s_i, taking values ± 1. Moreover, $A(t)$ and $B(t)$ are called the annealing functions, as they define the schedule according to which the system is annealed, and they are expressed as energies in units of Joules [D-Wg].

The two terms in (4.1) are the initial and final Hamiltonian. The former is also known as the *tunneling* Hamiltonian, since this is what causes the quantum tunneling phenomenon to take place, while the latter is also called the *problem* Hamiltonian, since its lowest-energy state provides the answer to the minimisation problem under consideration.

To sample from the Ising model (4.1), one option is to call the `sample_ising` method available in the `DWaveSampler` class, whose mandatory parameters are the qubit biases and coupling strengths in the form of Python dictionaries. The problem is then submitted to the quantum annealer through the Solver API (SAPI), and a digital-to-analog converter (DAC) attached to each qubit and coupler converts the submitted constants to electric current [D-Wa].

Subsequently, the annealing process starts at $t_0 = 0\mu$s and ends at $t_f\,\mu$s, parameter which can be specified by the user in the aforementioned function. When $t = 0$, $A(t) \gg B(t)$ and the system is in the lowest-energy of the initial Hamiltonian, when all the qubits are in their superposed state. Progressively, A decreases and B increases until $t = t_f$, so that the influence of the problem Hamiltonian is enhanced, while the impact of the tunneling Hamiltonian is diminished. If the annealing is performed slowly enough, when $t = t_f$, $B(t) \gg A(t)$ and the system is in the lowest-energy state of the final Hamiltonian, when all the qubits are in classical states.

As mentioned in Section 2.5, choosing an annealing schedule is a non-trivial task. Here, the values of the annealing functions A and B for each of D-Wave's QPUs (`Advantage_system3.2`, `Advantage_system4.1`, `Advantage_system5.1`, `DW_2000Q_6`) are available in the form of an `xlsx` file, along with their QPU-specific characteristics [D-Wc]. For instance, Figure 4.1 illustrates typical annealing functions A and B used in the D-Wave's Advantage systems [D-Wa]. Here, $s = t/t_f$ is the normalised annealing fraction and represents the annealing time, so that $0 \leq s \leq$

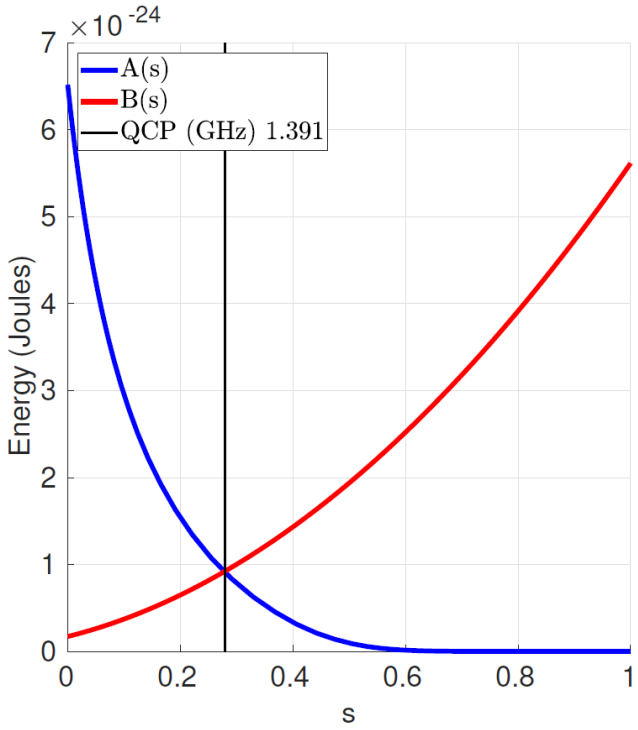


Figure 4.1: Annealing functions $A(s)$ and $B(s)$. The quantum critical point (QCP) is the point where the amplitudes of $A(s)$ and $B(s)$ are equal. [D-Wa]

1. Figure 4.1 also shows the quantum critical point (QCP), i.e. the point where $A(s) = B(s)$, which is where a phase transition takes place and topological defects may occur [Hod89].

The predefined values of $A(s)$ and $B(s)$ are described in terms of the hardware implementation of the quantum annealer. In particular, A is the energy difference between the two eigenstates of the rf-SQUID (Superconducting QUantum Interference Device) qubit used in the D-Wave's quantum machines with no external applied flux. On the other hand, $B = 2M_{AFM}I_p(s)^2$, where M_{AFM} is the maximum available mutual inductance achievable between pairs of flux qubit bodies and I_p is the magnitude of the current flowing through the body of the rf-SQUID loop [D-Wa].

While the A and B functions cannot be modified, the user can control the annealing schedule s or, in other words, the trajectory of the annealing functions. The default schedule is linear, so that $s = t/t_f$. This can be changed through the `anneal_schedule` parameter in the `sample_ising` method, as long as the submitted schedule adheres to certain restrictions [D-Wd].

In addition, there seems to be some flexibility in how the user can control the annealing process which occurs inside the D-Wave's quantum machines. Apart from varying the annealing schedule s as explained above, one can also *pause*, which makes the system wait for a given amount of time at a particular $A(s)$ and $B(s)$, or *quench* the schedule, which abruptly stops the annealing process near the specified point, or even initialise the qubits into a specific classical state and then anneal in *reverse*. Moreover, the values of A and B can be vertically shifted, known as *anneal offset*, allowing the user to control the anneal path of each qubit [D-Wa].

Nonetheless, one limitation of using the sampling methods available in the `DWaveSampler` class (or any predefined class in the `dwave-system` API) is that the returned values only provide information about the result of the annealing process, and not the intermediate mechanism. This restricts the comparisons we can perform with our simulations, which is the topic of the following section.

5 Numerical Experiments

Throughout this section, we explore various numerical experiments which can be mainly divided into the following two groups: validation and simulations. The former is concerned with investigating whether the integrators we have developed in Sections 3.3.4 and 3.3.5 behave correctly, while the latter aims to contrast the results obtained from our quantum annealing simulations with the outcomes retrieved from the D-Wave's quantum annealer (QA).

5.1 Validation

In the absence of an analytical solution to Eq. (3.4), we do not carry out experiments related to the error analysis of our integrators. It is worth mentioning that, since this is a stochastic equation, it is possible to approximate the order of the error associated with the two methods. In particular, this can be achieved by first computing a reference solution $\hat{S}_{h_{\text{ref}}}$ obtained by employing, for example, the integrator (3.51) with a very small stepsize h_{ref}. Similarly, solutions $\hat{S}_h$ are then approximated with larger stepsizes h, so that one can subsequently compute:

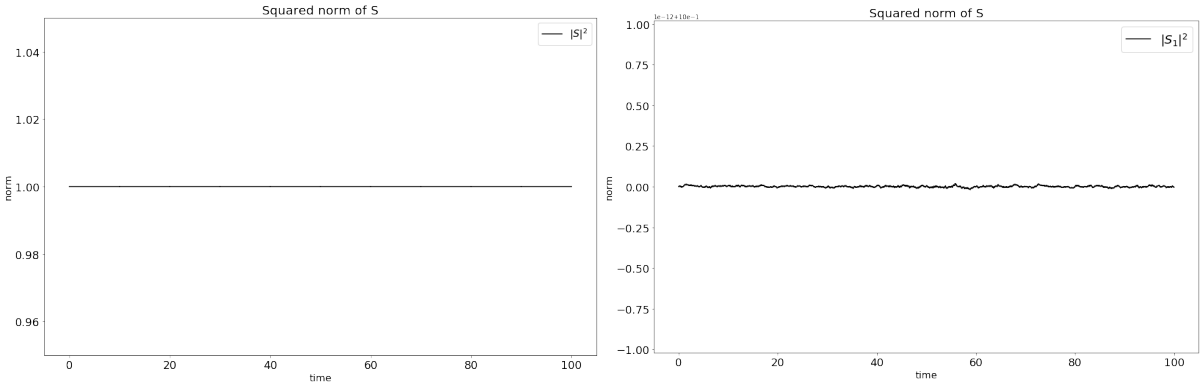
$$e = \left| \hat{S}_{h_{\text{ref}}} - \hat{S}_h \right|. \tag{5.1}$$

This can be repeated for the integrator (3.67). Due to the computational resources required to carry out this analysis, we exploit alternative techniques to validate our numerical methods.

First, one of the essential geometric properties described in Section 3.3.2 is the norm. This should satisfy $\|S_i(t)\| = 1, \forall t$, or equivalently each spin S_i should always stay on the unit sphere, with $i = 1, \dots N$, where N is the number of spins in the system, so that $N = 1$ gives the single-spin case. Accordingly, Figure 5.1 shows that this is indeed the case, up to machine precision, for both our integrators (3.51) and (3.67).

Next, we turn our attention to the $N = 1$ case. Since our integrator (3.51) is constructed by separating the conservative, damping and stochastic terms in Eq. (3.4), we verify whether each corresponding element in the composition performs as expected. To this end, we use the observations in Section 3.2 related to each of the aforementioned components.

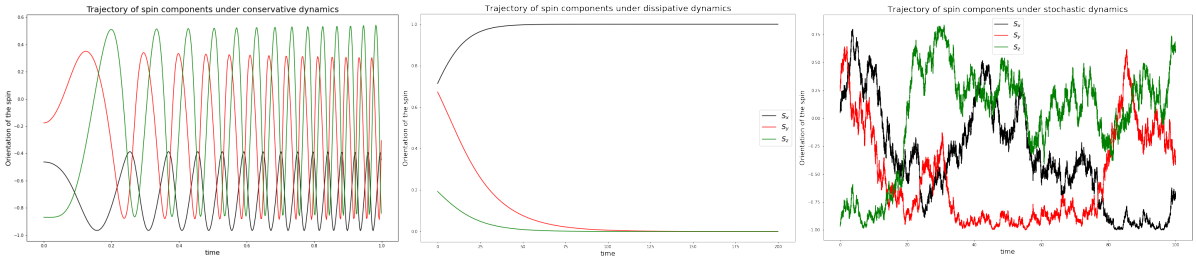
Thus, Figure 5.2a depicts the conservative dynamics generated by composing the component-wise solutions to Eq. (3.32). In particular, we see the spin orbiting around an applied magnetic field, namely $h = \begin{pmatrix} 1, & 1/4, & 1/6 \end{pmatrix}^T$. Further, Figure 5.2b illustrates the dissipative dynamics obtained similarly from Eq. (3.33). The damping term should cause the motion of the spin to simply disperse to a specific point, determined again by the external field, which is, indeed, what we observe in Figure 5.2b. Finally, Figure 5.2c reproduces the motion of a spin S under



(a) Single-spin case

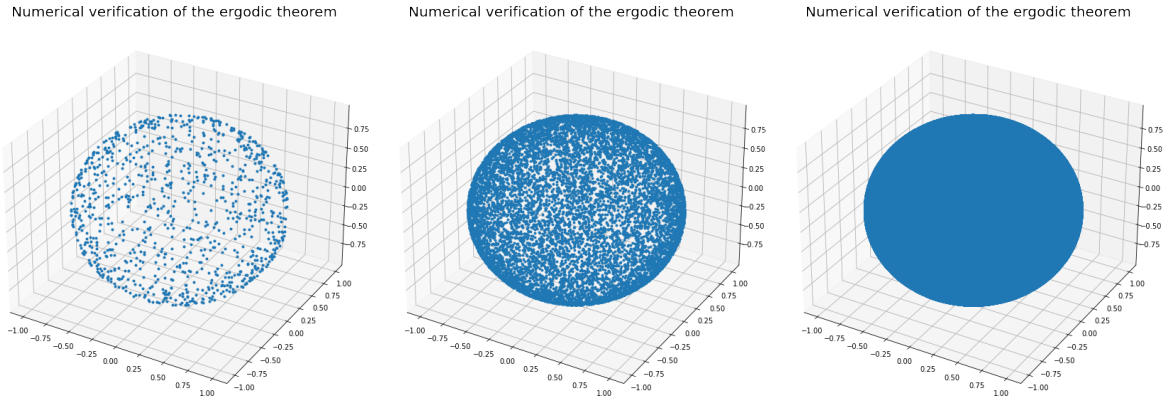
(b) System-of-spins case

Figure 5.1: Plot of the norm $\|S_i\|$ of one spin for integrators (3.51) and (3.67).



(a) Conservative dynamics (b) Dissipative dynamics (c) Stochastic dynamics

Figure 5.2: Plot of the individual dynamics in Eq. (3.4) for a single spin S.



(a) $T = 0.01$ (b) $T = 0.1$ (c) $T = 1$

Figure 5.3: Plot of the stochastic dynamics in Eq. (3.34) for different temperatures T.

the dynamics imposed by the governing stochastic equation (3.34). This resembles Brownian motion dynamics on the unit sphere, as expected.

In addition, since both integrators (3.51) and (3.67) share the solution to the stochastic equation (3.11), a relevant test is whether the solution (3.50) satisfies the ergodic theorem [Bir31], given in Appendix A.6. Roughly speaking, the theorem states that, given enough time, almost all trajectories of a dynamical system eventually stretch to every point in the phase space.

In the current context, we can verify the ergodic theorem either by fixing temperature T and letting the simulation run long enough, or by varying T, since this essentially governs the amount of stochasticity in the system. The latter procedure is illustrated in Figure 5.3, where we increase temperature from $T = 0.01$ in Figure 5.3a, to $T = 0.1$ in Figure 5.3b and $T = 1$ in Figure 5.3c. Indeed, as we gradually raise the temperature, we notice that the motion of the spin covers more and more points on the sphere, as anticipated.

Let us now focus on a system with $N > 1$ spins. Here, perhaps the most rigorous assessment of our integrator (3.67) consists of computing an observable for which we can establish a theoretical relationship. For example, we show in Appendix A.7 that for the rigid body Hamiltonian $\mathcal{H} = \frac{1}{2} S \cdot DS$, the following holds:

$$\langle \mathcal{H} \rangle_{\rho_\beta} = \frac{3}{2} T, \tag{5.2}$$

where $\langle \cdot \rangle_{\rho_\beta}$ denotes the average with respect to the stationary distribution ρ_β. By the ergodic theorem, $\langle \mathcal{H} \rangle_{\rho_\beta}$ is equal to the long-time average of $\mathcal{H}$, giving a tractable procedure for numerically verifying Eq. (5.2).

While we have not extended the derivation in Appendix A.7 to the system-of-spins Hamiltonian (3.3), intuitively we still expect the average energy to scale linearly with temperature.

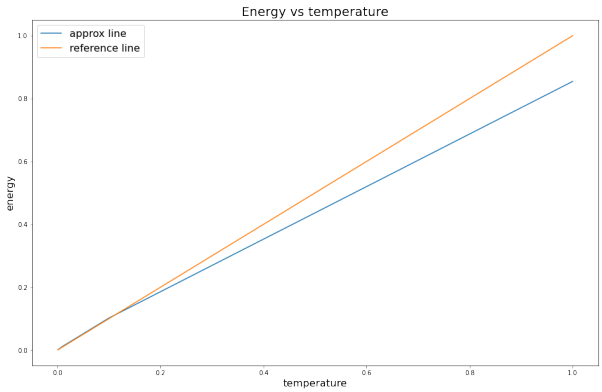
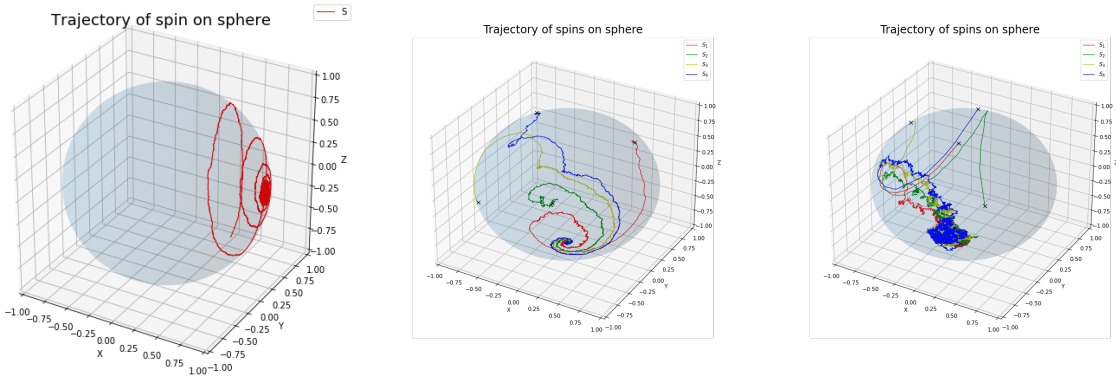


Figure 5.4: Plot of the mean energy $\langle \mathcal{H} \rangle$ for the Hamiltonian (3.3) against temperature T, with $N = 4$ spins, $n = 10^6$ steps, and stepsize $h = 10^{-4}$.



(a) $N = 1$ (b) $N = 4$ with SA (c) $N = 4$ with SQA

Figure 5.5: Plot of the trajectory of N spins on sphere.

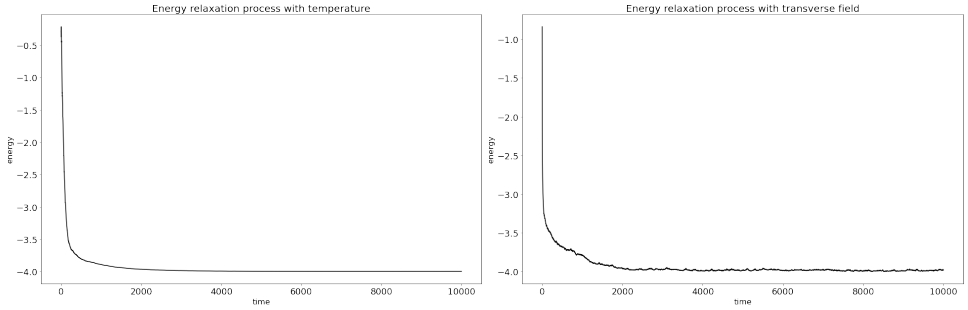
In fact, this hypothesis is also supported by our numerical experiments. Specifically, Figure 5.4 depicts the mean energy of $\mathcal{H}$ computed with $N = 4$ spins, $n = 10^6$ steps, and stepsize $h = 10^{-4}$. On the x-axis, we plot the temperature T, where $T \in \{10^{-3}, 10^{-2}, 10^{-1}, 1\}$, while on the y-axis, we plot the mean energy $\langle \mathcal{H} \rangle$ for each of the aforementioned T. Certainly, the resulting approximating line does not fully match the reference line. However, this can be justified, by the fact that, on the one hand we did not run long enough simulations, and, on in the other hand, we included too many points in the estimation of the average, that is, we did not discard sufficiently many points to allow the system to fully relax.

Finally, to visualise the complete dynamics governed by Eq. (3.4), we demonstrate the motion of $N \geq 1$ spins generated by our integrators (3.51) and (3.67) in Figure 5.5.

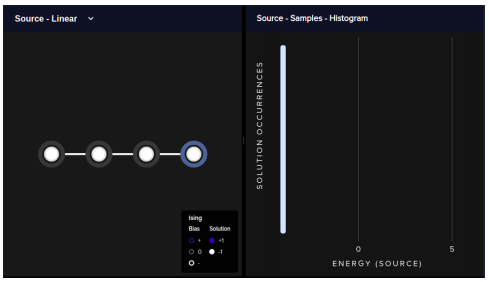
First, Figure 5.5a depicts the motion of one spin for the $N = 1$ case, where we use $D = \text{diag}(1, 1/2, 1/3)$ and apply an external field $h = \begin{pmatrix} 1, & 1/4, & 1/6 \end{pmatrix}^T$. Indeed, we observe a damped stochastic orbital motion on the unit sphere, until the spin S aligns with the external field h and subsequently oscillates around it on a small scale.

Further, Figures 5.5b and 5.5c illustrate the $N = 4$ case where we only use nearest-neighbour interactions, forming a chain of spins with open boundary conditions, i.e the first and last spin do not interact with each other. Additionally, we use $J_{i,j} = 1$ to force ferromagnetic interactions, so that, at the end, all the spins align with each other, and apply an external field $h_1 = \begin{pmatrix} 0, & 0, & -1 \end{pmatrix}^T$ to the first spin only. In addition, Figure 5.5b is generated using simulated annealing (SA) as described in Chapter 2, with temperature schedule $T_k = T_0/(1 + \alpha k)$, where $T_0 = 1$ is the initial temperature, k is the iteration number, and we take $\alpha = 1$. On the other

Figure 5.6: Chain of four spins with open boundary conditions and ferromagnetic couplings.



(a) SA (b) SQA



(c) D-Wave's QA

Figure 5.7: Simulated annealing, simulated quantum annealing and quantum annealing for a ferromagnetic chain of four spins.

hand, in Figure 5.5c we employ simulated quantum annealing (SQA) as it is also detailed in Chapter 2, with the transverse field $\Gamma(t) = 1/t$. In both cases, we notice the same behaviour: the spins align with each other and then with the applied external field. The difference in the type of motion essentially stems from the temperature T. In particular, in SA temperature is decreased, and consequently, so is the amount of stochasticity in the system, while in SQA temperature is low, yet constant, and therefore stochasticity is always present.

Having validated our integrators (3.51) and (3.67), we now direct our attention to performing simulations with specific system of spins and comparing the outcomes to the D-Wave's QA results.

5.2 Simulations

In this section, we only consider two systems of spins, namely a chain of four spins with open boundary conditions, and a frustrated system of eight spins. The former is a simple case which aims to provide a better understanding of the underlying procedure, while the latter is an intricate problem developed by Kadowaki and Nishimori [KN98] which exhibits meta-stability.

5.2.1 Chain of four spins

The configuration considered in this section is identical to the one described at the end of Section 5.1 for Figures 5.5b and 5.5c. Specifically, the system is highlighted in Figure 5.6, where full lines indicate ferromagnetic interactions. The global minimum for the system is achieved when

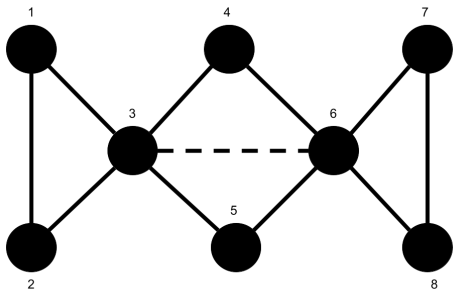


Figure 5.8: Frustrated system of eight spins as proposed by Kadowaki and Nishimori [KN98].

$S_i = \big(0, \quad 0, \quad -1\big)$, or, for the one-dimensional Ising spins, $s_i = -1$, $i = 1, \ldots, 4$, which yields:

$$\mathcal{H} = -(S_1 \cdot S_2 + S_2 \cdot S_3 + S_3 \cdot S_4) - (S_1 \cdot h_1) = -4. \tag{5.3}$$

Correspondingly, the energy relaxation for the system, as well as the D-Wave's QA result are given in Figure 5.7. In particular, Figure 5.7a illustrates the evolution of the chain's energy in time using SA with the temperature schedule $T_k = T_0/(1 + \alpha k)$, where $T_0 = 1$ is the initial temperature, k is the iteration number, and we take $\alpha = 1$. Note that $\mathcal{H} \approx -4$ within the first 10^3 iterations. Moreover, in Figure 5.7b we apply our integrator (3.67) to simulate the quantum annealing process with the transverse field $\Gamma(t) = 1/t$, where $t = kh$, h denoting the stepsize. Thus, we observe again that the energy dips to the ground state $\mathcal{H} \approx -4$, albeit more slowly. In both cases, we use $n = 10^5$ steps, with stepsize $h = 10^{-2}$.

Additionally, Figure 5.7c is retrieved from D-Wave's integrated development environment (IDE) Leap [D-We] which provides direct access to their Advantage quantum system. Following the procedure described in Section 4, we input the same values as above for the coupling and external magnetic field strength to the `sample_ising` function, and use the default annealing time 20μs and default schedule s. Hence, the left-hand side of Figure 5.7c shows the system configuration found from the aforementioned parameters, while the right-hand side is a histogram illustrating the number of times the annealer reaches the ground state $\mathcal{H} = -4$, where the Hamiltonian $\mathcal{H}$ is given by Eq. (4.1).

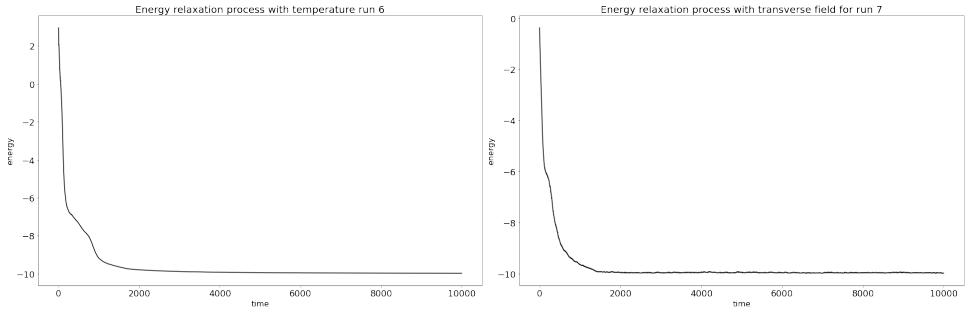
Consequently, we acknowledge that for systems as simple as chains of spins, both simulated and quantum annealing readily identify the corresponding global minimum. In order to reproduce the trapping phenomenon, that is, the system being temporarily frozen in a local minimum, in the subsequent section, we make use of a more complex configuration.

5.2.2 Frustrated system of eight spins

Let us now consider the system of eight spins proposed in [KN98], whose structure is reproduced in Figure 5.8. Here, full lines indicate ferromagnetic couplings, while dashed lines stand for antiferromagnetic couplings, which we impose by letting $J_{ij} = 1$ and $J_{ij} = -1$, respectively, making the system frustrated. In addition, similar to the previous experiment, we apply an external field $h_1 = \big(0, \quad 0, \quad -1\big)^T$ to the first spin only.

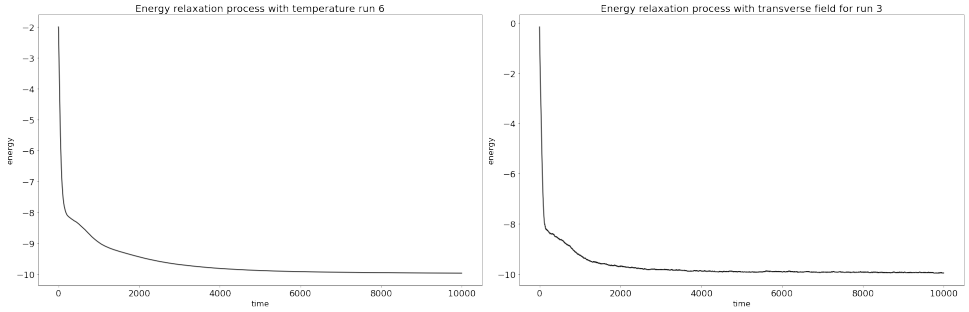
Hence, this system has a global minimum at $\mathcal{H} = -10$, which is achieved again when $S_i = \big(0, \quad 0, \quad -1\big)$, or, for the one-dimensional Ising spins, $s_i = -1$, $i = 1, \ldots, 8$. However, this specific organisation of spins also exhibits local minima at $\mathcal{H} \in \{-8, -6\}$. This is precisely what makes the previously mentioned configuration frustrated: the system cannot simultaneously achieve its ground state and comply with the requirements enforced by the couplings.

Accordingly, we anticipate that the associated energy $\mathcal{H}$ briefly stabilises at the $\mathcal{H} = -6$ and $\mathcal{H} = -8$ levels, before declining further to the ground state $\mathcal{H} = -10$. In order to test



(a) $\mathcal{H}$ trapped at -6 with SA (b) $\mathcal{H}$ trapped at -6 with SQA

Figure 5.9: Plot of the evolution of the energy $\mathcal{H}$ in time, with the system temporarily trapped in local minimum $\mathcal{H} = -6$.



(a) $\mathcal{H}$ trapped at -8 with SA (b) $\mathcal{H}$ trapped at -6 with SQA

Figure 5.10: Plot of the evolution of the energy $\mathcal{H}$ in time, with the system temporarily trapped in local minimum $\mathcal{H} = -8$.

our hypothesis, we employ our integrator (3.67) first with SA, exploiting the same temperature schedule as before, namely $T_k = T_0/(1 + \alpha k)$ with $T_0 = 1$ and $\alpha = 0.5$, and afterwards with SQA, adopting an exponential transverse field, given by $\Gamma(t) = e^{-t}$.

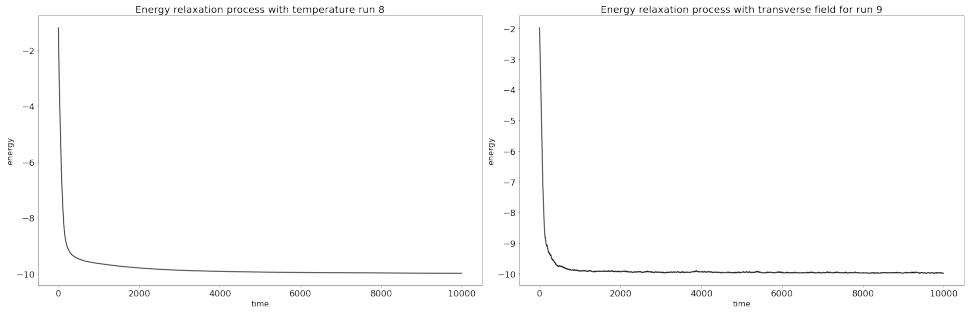
In all our simulations below we use $n = 10^5$ steps and stepsize $h = 10^{-2}$. Additionally, in an effort to emulate the format of the outcomes retrieved from the D-Wave's QA, we perform each of our experiments R times, where, due to the high computational cost, we favor small values of R, such as $R = 10$ or $R = 20$.

Hence, Figure 5.9 depicts the evolution of the energy $\mathcal{H}$ with time for the frustrated system of spins under consideration, with a temporary stagnation at energy level $\mathcal{H} = -6$. In particular, Figures 5.9a and 5.9b highlight this phenomenon when SA and SQA are applied to the problem, respectively. Moreover, it is worth mentioning that this arises within the first 2000 iterations. Subsequently, the energy progressively decreases to its lowest-energy level $\mathcal{H} = -10$.

If we now turn our attention to Figure 5.10, we recognise that the time-progression of the Hamiltonian $\mathcal{H}$ is comparable to the one in Figure 5.9. The predominant contrast between the two rests on where the energy level at which the system briefly stagnates occurs. Here, the local minimum emerges at $\mathcal{H} = -8$, rather than $\mathcal{H} = -6$, both when SA and SQA are engaged. This is illustrated in Figures 5.10a and 5.10b, respectively.

Critically, the most prevalent outcome we encounter in our simulations features the system settling directly in its ground state. This is highlighted in Figures 5.11a and 5.11b, where we apply both the SA and SQA approaches, respectively. Indeed, the two figures show the Hamiltonian $\mathcal{H}$ starting at a random point, associated with random initial conditions for the spins S_i, $i = 1, \ldots, 8$. Afterwards, the energy gradually decreases to the lowest-energy level $\mathcal{H} = -10$ within the first 10^3 iterations, without getting trapped in a local minimum.

Finally, we input the coupling values and the strength of the external magnetic field corre-



(a) Ground state $\mathcal{H} = -10$ with SA (b) Ground state $\mathcal{H} = -10$ with SQA

Figure 5.11: Plot of the evolution of the energy $\mathcal{H}$ in time, with the system declining directly to the global minimum $\mathcal{H} = -10$.

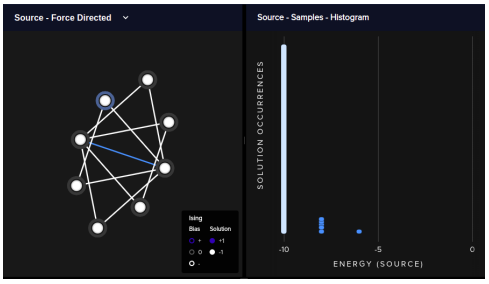


Figure 5.12: D-Wave's QA results for frustrated system of eight spins.

sponding to the configuration introduced at the beginning of this section in the `sample_ising` function on Leap and display the complete results in Figure 5.12. Perhaps not surprisingly, an important remark highlighted in this figure is that the D-Wave's QA struggles to find the global minimum of this problem, even though the current system involves a relatively small number of qubits, namely $N = 8$. In particular, the left-hand side of Figure 5.12 depicts the qubits on a graph, where the white and blue edges indicate ferromagnetic and anti-ferromagnetic couplings, respectively. The right-hand side of Figure 5.12 displays a histogram of the solutions identified by the D-Wave's QA, from which we can infer the corresponding probabilities. Thus, when using $t_f = 50\mu$s annealing time and 1000 runs, the global minimum $\mathcal{H} = -10$ is found with a probability of approximately 95%, the local minima $\mathcal{H} = -8$ is encountered with a probability of just above 5%, and the energy level $\mathcal{H} = -6$ arises less than 1% of the time.

To summarise, in this section, we have first validated our integrators (3.51) and (3.67) through various techniques. Subsequently, this allowed us to carry out diverse numerical experiments with different ensembles of spins and contrast the outcomes with the solutions generated with a physical quantum annealer, such as the one provided by D-Wave. As a cautious resolution, we can affirm that the two seem to produce comparable results, albeit more careful investigations and analyses with larger systems are necessary.

6 Conclusion and Future Work

6.1 Conclusion

6.1.1 Summary of Work

The study of quantum computers is motivated by the potential of such devices to revolutionise the field of scientific computing. It is expected that, provided they can be implemented, quantum computers will prove advantageous over classical computers, especially in the case of combinatorial optimisation.

In this report, we began with a summary of a class of such problems: Quadratic Unconstrained Binary Optimisation (QUBO). They can be linked to statistical mechanics by being written in the form of the Ising model and compared to a minimisation problem. This link was exploited previously in the case of simulated annealing as a heuristic algorithm for the approximate solution of NP-complete problems. We then continued, by analogy, to the concept of quantum annealing as a potentially faster route to such solutions. We discussed the mathematical foundations, noting that many results have been produced in great generality, but there are still unknowns regarding specific cases of optimisation problems. In the worst case scenarios, complexity will still grow exponentially with the size of the problem, as detailed in Section 2.6.

Section 3.3 introduced the classical model of the quantum system we aimed to simulate, i.e. the Heisenberg Hamiltonian under Langevin Dynamics. The rest of this section presented the main contribution of this project, namely two norm-preserving integrators based on splitting methods. Then, Section 4 explained how quantum annealing is implemented on D-Wave's machine, how we interact with the it via the IDE, and which control parameters are accessible to the user.

Finally, in Chapter 5, we used the previously mentioned integrators to obtain trajectories of spins and simulate the annealing process. This allowed us to compare our results with computations performed on D-Wave's quantum annealer by means of the IDE. In addition, we validated our method both mathematically and experimentally, and carried out simulations of a frustrated system, indicating signs of meta-stability in the relaxation of energy.

6.2 Future Work

6.2.1 Implementation in C or C++

Our prototypical code was written in Python, however, for large scale simulations, the current implementation is prohibitively slow. In order to proceed, we will need to translate the algorithm into a more efficient language such as `C` or `C++`, and potentially make use of HPC and parallelisation. Towards the end of this project, we came close to an analytic solution to a system of two coupled spins (see appendix A.4). Were we able to solve this, it is possible that the implementation will improve the accuracy of the integrator for large systems by decreasing the number of splittings needed. Parallel computing seems natural in this context since, by splitting the equations as was discussed in 3.3.5, we arrive at N independent sets of equations to solve. Solving these in parallel before composing the maps may be more efficient.

6.2.2 Compare Simulations to D-Wave Data

We wish to explore the similarity of our simulations to real data produced using LEAP, the D-Wave IDE. Ideally, we would run a large number of simulations and compare the resulting statistics to the D-Wave data, however, D-Wave provides only one free minute every month, making careful comparisons difficult. Greater access to the system would be a considerable advantage. It is still an unknown whether details of the relaxation process can be returned via the D-Wave IDE. This would also be of great interest. These comparisons are especially fascinating in the case of frustrated systems. Relevant examples for such systems are spin

glasses which have many meta-stable phases [Zam10] and degenerate ground states [KGV83]. As shown in Section 5.2.2, frustrated systems can pose challenges for the D-Wave machine as it freezes in a meta-stable state on some runs. It would be interesting to compare the annealing schedules discussed in Sections 2.5 and 2.6 to the results obtained using the D-Wave machine in order to have a better understanding about the effects of the annealing time on the results.

6.2.3 Explore the Formation of Topological Defects

Related to the existence of phase transitions and meta-stability is the concept of *topological defects*. They represent a collection of neighbouring particles for which the angle θ, specifying the direction of the spin, changes by 2π in one circuit of any closed contour enclosing the core of the vortex located at the origin. In the setting of a Heisenberg magnet this is known as a *hedgehog defect*, as shown in Figure 6.1. In short, when the transverse field is decreased, it is

Figure 6.1: Hedgehog defect in 3D, the vectors are normal to the sphere, all either inward or outward [NS18].

expected that a critical point is attained where the relaxation time diverges [Suz09]. For the Heisenberg model, this divergence generates quantum phase transition from the paramagnetic to the (anti)-/ferromagnetic state. Hence, the formation of topological defects or "kinks" occurs in the final state. In [Sub+21], the statistics of defects is studied using the kink-number operator [Dzi05]:

$$\mathcal{N} = \frac{1}{2}\sum_{i=1}^{N-1}(1-\sigma_i^z\sigma_{i+1}^z). \tag{6.1}$$

Unfortunately, this operator is specifically devised for quantum Ising models, so more investigation is needed to find a suitable operator for the quantum Heisenberg model. The Kibble-Zurek Mechanism (KZM) [Kib76], [Kib80], [Zur85], [Zur96] is the theory used to describe the formation of topological defects as a system goes through a phase transition. It essentially states that the average number of kinks scales as a function of the annealing time t_a :

$$\mathbb{E}[\mathcal{N}] \sim t_a^{-\alpha_{\mathrm{KZM}}}, \tag{6.2}$$

where $\alpha_{\mathrm{KZM}} = d\nu/(1+z\nu)$ is the KZM critical exponent for d-dimensional and point-like defects. The values the critical exponents ν and z can potentially take are discussed in [Sub+21]. As suggested in [Sub+21], Eq. (6.2) is a suitable metric to assess the performance of a quantum annealer.

References

[AL04] Teijo Arponen and Benedict J. Leimkuhler. "An Efficient Geometric Integrator for Thermostatted Anti-/Ferromagnetic Models". In: *BIT Numerical Mathematics* 44 (2004), pp. 403–424.

[Bar82] F Barahona. "On the computational complexity of Ising spin glass models". In: *Journal of Physics A: Mathematical and General* 15.10 (Oct. 1982), pp. 3241–3253. ISSN: 0305-4470, 1361-6447. DOI: 10.1088/0305-4470/15/10/028. (Visited on 12/14/2021).

[Bir31] George D Birkhoff. "Proof of the ergodic theorem". In: *Proceedings of the National Academy of Sciences* 17.12 (1931), pp. 656–660.

[Boy20] Òscar Boyle. *Quantum annealing via real-time Schrödinger dynamics*. 2020. URL: http://diposit.ub.edu/dspace/bitstream/2445/170922/1/BOYLE%5C%20GARCIA%5C%20O%5C%CC%5C%80SCAR_1338794_assignsubmission_file_TFG-Boyle_Garcia_Oscar.pdf.

[Bro+99] J. Brooke et al. "Quantum Annealing of a Disordered Magnet". In: *Science* 284.5415 (Apr. 1999), pp. 779–781. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.284.5415.779. (Visited on 12/13/2021).

[Bro63] William Fuller Brown. "Thermal Fluctuations of a Single-Domain Particle". In: *Phys. Rev.* 130 (5 June 1963), pp. 1677–1686. DOI: 10.1103/PhysRev.130.1677.

[BT93] Dimitris Bertsimas and John Tsitsiklis. "Simulated Annealing". In: *Statistical Science* 8.1 (1993), pp. 10–15. DOI: 10.1214/ss/1177011077.

[But16] John Charles Butcher. *Numerical methods for ordinary differential equations*. John Wiley & Sons, 2016.

[Che+19] Weng Cho Chew et al. "Hamilton Equations, Commutator, and Energy Conservation". In: *Quantum Reports* 1 (Dec. 2019), pp. 295–303. DOI: 10.3390/quantum1020027.

[Cip87] Barry A. Cipra. "An Introduction to the Ising Model". In: *The American Mathematical Monthly* 94.10 (1987), pp. 937–959. DOI: 10.1080/00029890.1987.12000742.

[D-Wa] D-Wave. *Annealing Implementation and Controls — D-Wave System Documentation documentation*. URL: https://docs.dwavesys.com/docs/latest/c_qpu_annealing.html (visited on 12/07/2021).

[D-Wb] D-Wave. *D-Wave QPU Architecture: Topologies*. URL: https://docs.dwavesys.com/docs/latest/c_gs_4.html#getting-started-topologies (visited on 12/13/2021).

[D-Wc] D-Wave. *QPU-Specific Characteristics*. URL: https://docs.dwavesys.com/docs/latest/doc_physical_properties.html#doc-qpu-characteristics (visited on 12/13/2021).

[D-Wd] D-Wave. *Solver Parameters*. URL: https://docs.dwavesys.com/docs/latest/c_solver_parameters.html#param-anneal-sched (visited on 12/13/2021).

[D-We] D-Wave. *The Leap Quantum Cloud Service*. URL: https://www.dwavesys.com/solutions-and-products/cloud-platform/ (visited on 12/15/2021).

[D-Wf] D-Wave. *Welcome to D-Wave — D-Wave System Documentation documentation*. URL: https://docs.dwavesys.com/docs/latest/c_gs_1.html (visited on 11/22/2021).

[D-Wg] D-Wave. *What is Quantum Annealing? — D-Wave System Documentation documentation*. URL: https://docs.dwavesys.com/docs/latest/c_gs_2.html (visited on 11/22/2021).

[Dir47] P.A.M. Dirac. *The Principles of Quantum Mechanics*. 3rd. London: Clarendon Press, 1947, p. 13.

[Dzi05] Jacek Dziarmaga. "Dynamics of a Quantum Phase Transition: Exact Solution of the Quantum Ising Model". In: *Phys. Rev. Lett.* 95 (24 Dec. 2005), p. 245701. DOI: 10.1103/PhysRevLett.95.245701.

[EA75] S F Edwards and P W Anderson. "Theory of spin glasses". In: *Journal of Physics F: Metal Physics* 5.5 (May 1975), pp. 965–974. DOI: 10.1088/0305-4608/5/5/017.

[Ell+15] Matthew O. A. Ellis et al. "The Landau–Lifshitz equation in atomistic models". In: *Low Temperature Physics* 41 (2015), pp. 705–712.

[Far+00] Edward Farhi et al. *Quantum Computation by Adiabatic Evolution*. Jan. 2000. arXiv: quant-ph/0001106 [quant-ph].

[FHL97] Jason Frank, Weizhang Huang, and Benedict J. Leimkuhler. "Geometric Integrators for Classical Spin Systems". In: *Journal of Computational Physics* 133 (1997), pp. 160–172.

[FQ87] Kang Feng and Meng-zhao Qin. "The symplectic methods for the computation of hamiltonian equations". In: *Numerical Methods for Partial Differential Equations*. Ed. by You-Ian Zhu and Ben-yu Guo. Berlin, Heidelberg: Springer Berlin Heidelberg, 1987, pp. 1–37. ISBN: 978-3-540-48126-3.

[Gar+85] Crispin W Gardiner et al. *Handbook of stochastic methods*. Vol. 3. springer Berlin, 1985, p. 106.

[GG84] Stuart Geman and Donald Geman. "Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* PAMI-6.6 (1984), pp. 721–741. DOI: 10.1109/TPAMI.1984.4767596.

[Gil04] Thomas L. Gilbert. "A phenomenological theory of damping in ferromagnetic materials". In: *IEEE Transactions on Magnetics* 40 (2004), pp. 3443–3449.

[GS22] Wālther Gerlāch and Otto Stern. "Der experimentelle Nachweis der Richtungsquantelung im Magnetfeld". In: *Zeitschrift für Physik* 9 (1922), pp. 349–352.

[Haj88] Bruce Hajek. "Cooling Schedules for Optimal Annealing." In: *Mathematics of Operations Research* 13.2 (1988), pp. 311–329. ISSN: 0364-765X. DOI: 10.1287/moor.13.2.311.

[Hal03] B.C. Hall. *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction*. Graduate Texts in Mathematics. Springer, 2003. ISBN: 9780387401225.

[Hod89] Hardy M Hodges. "Formation of topological defects in phase transitions". In: *Physical Review D* 39.12 (1989), p. 3557.

[HS05] Naomichi Hatano and Masuo Suzuki. "Finding Exponential Product Formulas of Higher Orders". In: *Quantum Annealing and Other Optimization Methods*. Ed. by Arnab Das and Bikas K. Chakrabarti. Berlin, Heidelberg: Springer Berlin Heidelberg, 2005, pp. 37–68. ISBN: 978-3-540-31515-5. DOI: 10.1007/11526216_2.

[Joh90] David S. Johnson. "A Catalog of Complexity Classes". In: *Algorithms and Complexity*. Elsevier, 1990, pp. 67–161. DOI: 10.1016/b978-0-444-88071-0.50007-2.

[Kat50] Tosio Kato. "On the Adiabatic Theorem of Quantum Mechanics". In: *Journal of the Physical Society of Japan* 5.6 (Nov. 1950), pp. 435–439. DOI: 10.1143/jpsj.5.435.

[KGV83] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. "Optimization by simulated annealing". In: *Science* 220.4598 (1983), pp. 671–680.

[Kib76] T W B Kibble. "Topology of cosmic domains and strings". In: *Journal of Physics A: Mathematical and General* 9.8 (Aug. 1976), pp. 1387–1398. DOI: 10.1088/0305-4470/9/8/029.

[Kib80] T.W.B. Kibble. "Some implications of a cosmological phase transition". In: *Physics Reports* 67.1 (1980), pp. 183–199. ISSN: 0370-1573. DOI: https://doi.org/10.1016/0370-1573(80)90091-5.

[KN98] Tadashi Kadowaki and Hidetoshi Nishimori. "Quantum annealing in the transverse Ising model". In: *Physical Review E* 58.5 (Nov. 1998), pp. 5355–5363. DOI: 10.1103/physreve.58.5355.

[Koc+14] Gary Kochenberger et al. "The unconstrained binary quadratic programming problem: a survey". In: *Journal of Combinatorial Optimization* 28.1 (July 2014), pp. 58–81. ISSN: 1382-6905, 1573-2886. DOI: 10.1007/s10878-014-9734-0. (Visited on 12/14/2021).

[KS20] Mariia Karabin and Steven J. Stuart. *Simulated Annealing with Adaptive Cooling Rates*. 2020. arXiv: 2002.06124 [physics.chem-ph].

[KT15] H. Kawamura and T. Taniguchi. *Chapter 1 - Spin Glasses*. Ed. by K.H.J. Buschow. Vol. 24. Handbook of Magnetic Materials. Elsevier, 2015, pp. 1–137. DOI: https://doi.org/10.1016/bs.hmm.2015.08.001.

[KT85] S. Kirkpatrick and G. Toulouse. "Configuration space analysis of travelling salesman problems". In: *Journal de Physique* 46.8 (1985), pp. 1277–1292. ISSN: 0302-0738. DOI: 10.1051/jphys:019850046080127700. (Visited on 12/14/2021).

[Lak11] M. Lakshmanan. "The fascinating world of the Landau-Lifshitz-Gilbert equation: an overview". In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* (2011). DOI: 10.1098/rsta.2010.0319.

[LK73] S. Lin and B. W. Kernighan. "An Effective Heuristic Algorithm for the Traveling-Salesman Problem". In: *Operations Research* 21.2 (Apr. 1973), pp. 498–516. ISSN: 0030-364X, 1526-5463. DOI: 10.1287/opre.21.2.498. (Visited on 12/13/2021).

[LL92] L. Landau and E. Lifschitz. "3 - On the theory of the dispersion of magnetic permeability in ferromagnetic bodies. Reprinted from Physikalische Zeitschrift der Sowjetunion 8, Part 2, 153, 1935." In: *Perspectives in Theoretical Physics*. Ed. by L.P. Pitaevski. Amsterdam: Pergamon, 1992, pp. 51–65. ISBN: 978-0-08-036364-6. DOI: https://doi.org/10.1016/B978-0-08-036364-6.50008-9.

[LM16] Ben Leimkuhler and Charles Matthews. *Molecular Dynamics*. Springer, 2016.

[MD11] Pui-Wai Ma and S. L. Dudarev. "Langevin spin dynamics". In: *Phys. Rev. B* 83 (13 Apr. 2011), p. 134418. DOI: 10.1103/PhysRevB.83.134418.

[Met+53] Nicholas Metropolis et al. "Equation of State Calculations by Fast Computing Machines". In: *The Journal of Chemical Physics* 21.6 (June 1953), pp. 1087–1092. DOI: 10.1063/1.1699114.

[MLM21] Ari Mizel, D. A. Lidar, and M. W. Mitchell. "Erratum: Simple Proof of Equivalence between Adiabatic Quantum Computation and the Circuit Model [Phys. Rev. Lett. 99, 070502 (2007)]". In: *Phys. Rev. Lett.* 127 (13 Sept. 2021), p. 139901. DOI: 10.1103/PhysRevLett.127.139901.

[MN08] Satoshi Morita and Hidetoshi Nishimori. "Mathematical foundation of quantum annealing". In: *Journal of Mathematical Physics* 49.12 (Dec. 2008), p. 125210. DOI: 10.1063/1.2995837.

[NC11] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. 10th. USA: Cambridge University Press, 2011. ISBN: 1107002176.

[NN96] Hidetoshi Nishimori and Yoshihiko Nonomura. "Quantum Effects in Neural Networks". In: *Journal of the Physical Society of Japan* 65.12 (Dec. 1996), pp. 3780–3796. DOI: 10.1143/jpsj.65.3780.

[NS18] Lynn Nadel and Daniel I. Stein. *1990 Lectures in Complex Systems*. Ed. by Lynn Nadel and Daniel L. Stein. CRC Press, Oct. 2018. DOI: 10.1201/9780429503573.

[RHC18] Dina Razafindralandy, Aziz Hamdouni, and Marx Chhay. "A review of some geometric integrators". In: *Advanced Modeling and Simulation in Engineering Sciences* 5.1 (June 2018), p. 16. ISSN: 2213-7467. DOI: 10.1186/s40323-018-0110-y.

[Sel+09] W. Selke et al. "Classical and quantum anisotropic Heisenberg antiferromagnets". In: *Condensed Matter Physics* 12 (2009), pp. 547–558.

[Shi+14] Seung Woo Shin et al. *How "Quantum" is the D-Wave Machine?* Jan. 2014. arXiv: 1401.7087 [quant-ph].

[SK75] David Sherrington and Scott Kirkpatrick. "Solvable Model of a Spin-Glass". In: *Phys. Rev. Lett.* 35 (26 Dec. 1975), pp. 1792–1796. DOI: 10.1103/PhysRevLett.35.1792.

[Son+97] S. L. Sondhi et al. "Continuous quantum phase transitions". In: *Reviews of Modern Physics* 69.1 (Jan. 1997), pp. 315–333. DOI: 10.1103/revmodphys.69.315.

[Sub+21] David Subires et al. *Benchmarking Quantum Annealing Dynamics: the Spin-Vector Langevin Model*. 2021. arXiv: 2109.09750v2 [quant-ph].

[Suz09] Sei Suzuki. "A comparison of classical and quantum annealing dynamics". In: *Journal of Physics: Conference Series* 143 (Jan. 2009), p. 012002. DOI: 10.1088/1742-6596/143/1/012002.

[The+17] T. Theurer et al. "Resource Theory of Superposition". In: *Phys. Rev. Lett.* 119 (23 Dec. 2017), p. 230401. DOI: 10.1103/PhysRevLett.119.230401.

[TKL04] Shan-Ho Tsai, M. Krech, and D. Landau. "Symplectic Integration Methods in Molecular and Spin Dynamics Simulations". In: *Brazilian Journal of Physics* 34 (June 2004), pp. 384–391. DOI: 10.1590/S0103-97332004000300009.

[VV21] Máté Tibor Veszeli and Gábor Vattay. *Mean Field Approximation for solving QUBO problems*. 2021. arXiv: 2106.03238 [quant-ph].

[Wil+21] Dennis Willsch et al. *Benchmarking Advantage and D-Wave 2000Q quantum annealers with exact cover problems*. 2021. arXiv: 2105.02208 [quant-ph].

[Zam10] Francesco Zamponi. *Mean field theory of spin glasses*. Aug. 2010. arXiv: 1008.4844 [cond-mat.stat-mech].

[Zur85] Wojciech Zurek. "Cosmological experiments in superfluid helium?" In: *Nature* 317 (1985), pp. 505–508.

[Zur96] W.H. Zurek. "Cosmological experiments in condensed matter systems". In: *Physics Reports* 276.4 (1996), pp. 177–221. ISSN: 0370-1573. DOI: https://doi.org/10.1016/S0370-1573(96)00009-9.

A Appendices

A.1 Appendix 1

Equation (3.36), namely:

$$\ddot{S}^y = -S^x(0)d_1\dot{S}^z = -\left(S^x(0)d_1\right)^2 S^y, \tag{A.1}$$

can be solved by first computing the roots of the associated characteristic equation:

$$\lambda^2 + \left(S^x(0)d_1\right)^2 = 0. \tag{A.2}$$

These are given by $\lambda_\pm = \pm iS^x(0)d_1$, so that the solution to the ODE (A.1) is:

$$S^y(t) = c_1 \sin\left(d_1 S^x(0)t\right) + c_2 \cos\left(d_1 S^x(0)t\right). \tag{A.3}$$

Imposing the initial condition on S^y yields:

$$S^y(0) = c_2. \tag{A.4}$$

From Eq. (3.35), we also have the initial condition $\dot{S}^y(0) = -d_1 S^x(0)S^z(0)$, which gives:

$$-c_1 d_1 S^x(0) = -d_1 S^x(0)S^z(0) \implies c_1 = S^z(0). \tag{A.5}$$

This gives the complete solution:

$$S^y(t) = S^z(0)\sin\left(d_1 S^x(0)t\right) + S^y(0)\cos\left(d_1 S^x(0)t\right). \tag{A.6}$$

A.2 Appendix 2

The solution to Eq. (3.41), namely:

$$\dot{u}^x = 2\gamma d_1 u^x\left(1 - u^x\right), \tag{A.7}$$

can be found by using partial fractions, as follows:

$$\int_{u^x(0)}^{u^x}\left(\frac{1}{v} + \frac{1}{1-v}\right)\mathrm{d}v = 2\gamma d_1 \int_0^t \mathrm{d}t', \tag{A.8}$$

which, upon rewriting, gives:

$$\ln\left|\frac{u^x}{1-u^x}\right| - \ln\left|\frac{u^x(0)}{1-u^x(0)}\right| = 2\gamma d_1 t. \tag{A.9}$$

Since $u^x + u^y + u^z = 1$, we have that $u^\alpha < 1$, $\forall t$, for $\alpha \in \{x, y, z\}$, so that the quantities inside the absolute values are positive. Thus:

$$\frac{u^x}{1-u^x} = \frac{u^x(0)}{1-u^x(0)}\exp(2\gamma d_1 t), \tag{A.10}$$

and letting $R \coloneqq u^x(0)/(1-u^x(0))$, we obtain:

$$u^x(t) = \frac{R\exp(2\gamma d_1 t)}{1 + R\exp(2\gamma d_1 t)}. \tag{A.11}$$

Further, to solve Eq. (3.43), namely:

$$\int_{u^y(0)}^{u^y} \frac{1}{v}\mathrm{d}v = -2\gamma d_1 \int_0^t u^x \mathrm{d}t' = -2\gamma d_1 \int_0^t \frac{R\exp(2\gamma d_1 t)}{1+R\exp(2\gamma d_1 t)}\mathrm{d}t', \tag{A.12}$$

we need only compute the integral on the RHS. This can be achieved by substituting $w = 1 + R\exp(2\gamma d_1 t)$, so that $\mathrm{d}w = 2\gamma d_1 R\exp(2\gamma d_1 t)$. Hence, the integral becomes:

$$\int_{1+R}^{1+R\exp(2\gamma d_1 t)} \frac{1}{w}\mathrm{d}w, \tag{A.13}$$

and substituting into Eq. (A.12) gives:

$$\ln\left|\frac{u^y}{u^y(0)}\right| = -\ln\left|\frac{1+R\exp(2\gamma d_1 t)}{1+R}\right|. \tag{A.14}$$

Using the fact that $w > 0$ and $u^y > 0$, $\forall t$, we obtain the following solution for u^y:

$$u^y(t) = u^y(0)\frac{1+R}{1+R\exp(2\gamma d_1 t)}. \tag{A.15}$$

A.3 Appendix 3

The solution to Eq. (3.34), namely:

$$\mathrm{d}S = \sqrt{\eta}S \times \mathrm{d}W, \tag{A.16}$$

can be found by further manipulation of Eq. (3.46) given by:

$$R'(t)S(0)\mathrm{d}t = \sqrt{\eta}\ \mathrm{skew}(\mathrm{d}W(t))R(t)S(0). \tag{A.17}$$

In particular, multiplying both sides by $R^{-1}(t)$ yields:

$$R'(t)R^{-1}(t)\mathrm{d}t = \sqrt{\eta}\ \mathrm{skew}(\mathrm{d}W(t)). \tag{A.18}$$

Furthermore,

$$\mathrm{d}\log R(t) = \sqrt{\eta}\ \mathrm{skew}(\mathrm{d}W(t)). \tag{A.19}$$

Integrating both sides and exponentiating, we obtain the following:

$$\begin{aligned} R(t) &= e^{\sqrt{\eta}\int_0^t \mathrm{skew}(dW(s))} \\ &= e^{\sqrt{\eta}\ \mathrm{skew}\left(\int_0^t dW(s)\right)} \\ &= e^{\sqrt{\eta}\ \mathrm{skew}(W(t))} \\ &= e^{\sqrt{\eta}\sqrt{t}\ \mathrm{skew}(W(1))}. \end{aligned} \tag{A.20}$$

Here, in the second equality we used that the skew-symmetric operator is linear, while in the last equality the Itô integral $\int_0^t \mathrm{d}W(s)$ is just the 3-dimensional Brownian motion $\sqrt{t}W(1)$ where each component $W^\alpha(1) \sim N(0,1)$ for $\alpha \in \{x, y, z\}$. Therefore,

$$S(t) = S(0)e^{\sqrt{\eta}\sqrt{t}\ \mathrm{skew}(W(1))}. \tag{A.21}$$

Further, we take $\omega = \begin{pmatrix}\omega_1, & \omega_2, & \omega_3\end{pmatrix}^T$ and Ω as in Section 3.3.4. Using $\Omega^3 = -\Omega$, by the

general formula for the matrix exponential [Hal03], we have that:

$$\begin{aligned}
e^{\sqrt{\eta t}\Omega} &= I_3 + \sqrt{\eta t}\Omega + \frac{1}{2!}\eta t\Omega^2 + \frac{1}{3!}(\eta t)^{\frac{3}{2}}\Omega^3 + \frac{1}{4!}(\eta t)^2\Omega^4 + ... \\
&= I_3 + \sqrt{\eta t}\Omega + \frac{1}{2!}\eta t\Omega^2 - \frac{1}{3!}(\eta t)^{\frac{3}{2}}\Omega - \frac{1}{4!}(\eta t)^2\Omega^2 - ... \\
&= I_3 + \left(\sum_{k=0}^{\infty}\frac{(-1)^k(\sqrt{\eta t})^{2k+1}}{(2k+1)!}\right)\Omega + \left(-\sum_{k=1}^{\infty}\frac{(-1)^k(\sqrt{\eta t})^{2k}}{(2k)!}\right)\Omega^2 \\
&= I_3 + \left(\sum_{k=0}^{\infty}\frac{(-1)^k(\sqrt{\eta t})^{2k+1}}{(2k+1)!}\right)\Omega + \left(1 - \sum_{k=0}^{\infty}\frac{(-1)^k(\sqrt{\eta t})^{2k}}{(2k)!}\right)\Omega^2 \\
&= I_3 + \sin(\sqrt{\eta t})\Omega + \left(1 - \cos(\sqrt{\eta t})\right)\Omega^2. \qquad \text{(A.22)}
\end{aligned}$$

Finally, the solution to Eq. (A.16) is therefore:

$$S(t) = S(0)\left(I_3 + \sin(\sqrt{\eta t})\Omega + \left(1 - \cos(\sqrt{\eta t})\right)\Omega^2\right). \qquad \text{(A.23)}$$

A.4 Appendix 4

The solution to Eq. (3.56), namely:

$$\begin{cases} \dot{S}_i &= -\frac{1}{2}J_{ij}S_i \times S_j - \frac{1}{2}J_{ij}\gamma S_i \times (S_i \times S_j) \\ \dot{S}_j &= -\frac{1}{2}J_{ji}S_j \times S_i - \frac{1}{2}J_{ji}\gamma S_j \times (S_j \times S_i), \end{cases} \qquad \text{(A.24)}$$

can be derived as follows. Using the triple product formula for the cross product, we have:

$$S_i \times (S_i \times S_j) = S_i(S_i \cdot S_j) - S_j(S_i \cdot S_i) = S_i(S_i \cdot S_j) - S_j \qquad \text{(A.25)}$$

and

$$S_j \times (S_j \times S_i) = S_j(S_j \cdot S_i) - S_i(S_j \cdot S_j) = S_j(S_j \cdot S_i) - S_i \qquad \text{(A.26)}$$

since $S_i \cdot S_i = S_j \cdot S_j = 1$. If we denote $w = S_i \cdot S_j$, and apply the Leibniz rule for the dot product, we get $\dot{w} = \dot{S}_i \cdot S_j + S_i \cdot \dot{S}_j$. Noting that $S_j \cdot (S_i \times S_j) = S_i \cdot (S_j \times S_i) = 0$, and combining both equations we arrive at:

$$\dot{w} = -J_{ij}\gamma(w^2 - 1) \qquad \text{(A.27)}$$

since $J_{ij} = J_{ji}$. This leads to:

$$\int_{w(0)}^{w}\frac{\mathrm{d}u}{u^2 - 1} = -J_{ij}\gamma\int_0^t \mathrm{d}t'. \qquad \text{(A.28)}$$

If we write:

$$\frac{1}{u^2 - 1} = \frac{1}{2}\left(\frac{1}{u-1} - \frac{1}{u+1}\right), \qquad \text{(A.29)}$$

then we observe the anti-derivative of the logarithm in the integral on the left-hand side, and therefore we finally arrive at:

$$w(t) = \frac{1 - K\exp(-2J_{ij}\gamma t)}{1 + K\exp(-2J_{ij}\gamma t)}, \qquad \text{(A.30)}$$

where $K \coloneqq (1 - w(0))/(1 + w(0))$. Using the relations above, we can further write (3.56) as:

$$\begin{cases} \dot{S}_i &= -\frac{1}{2}J_{ij}S_i \times S_j - \frac{1}{2}J_{ij}\gamma(S_i(S_i \cdot S_j) - S_j) \\ \dot{S}_j &= -\frac{1}{2}J_{ji}S_j \times S_i - \frac{1}{2}J_{ji}\gamma(S_j(S_j \cdot S_i) - S_i). \end{cases} \qquad \text{(A.31)}$$

Adding these two equations and using the fact that $S_i \times S_j = -S_j \times S_i$ gives:

$$\dot{u} = -\frac{1}{2} J_{ij} \gamma u(w-1), \tag{A.32}$$

where $u(t) \coloneqq S_i(t) + S_j(t)$. This leads to:

$$\int_{u(0)}^{u} \frac{\mathrm{d}r}{r} = -\frac{1}{2} J_{ij}\gamma \int_0^t (w(t')-1)\mathrm{d}t'. \tag{A.33}$$

Hence,

$$u(t) = u(0)\left(\frac{1+K}{1+K\exp\left(-2J_{ij}\gamma t\right)}\right)^{1/2}. \tag{A.34}$$

One could now, for example, use the first equation in (A.31) and it reduces to:

$$\dot{S}_i = -\frac{1}{2} J_{ij} S_i \times u - \frac{1}{2} J_{ij}\gamma((1+w)S_i - u). \tag{A.35}$$

We believe that this equation is analytically integrable, but we have not yet found a way to integrate it.

A.5 Appendix 5

The solution to equation (3.59), namely:

$$\dot{S_i^x} = \frac{1}{2} J_{ij}\gamma S_j^x(0)\left(1-(S_i^x)^2\right), \tag{A.36}$$

can be solved by taking $1-(S_i^x)^2$ to the left-hand side and separating into partial fractions:

$$\frac{1}{2}\int_{S_i^x(0)}^{S_i^x}\left(\frac{1}{1-v}+\frac{1}{1+v}\right)\mathrm{d}v = \frac{1}{2} J_{ij}\gamma S_j^x(0)\int_0^t \mathrm{d}t', \tag{A.37}$$

which, upon rewriting, gives:

$$\ln\left|\frac{1+S_i^x}{1-S_i^x}\right| - \ln\left|\frac{1+S_i^x(0)}{1-S_i^x(0)}\right| = J_{ij}\gamma S_j^x(0)t. \tag{A.38}$$

Since $\|S_i\| = 1$, we have that $S_i^x < 1$, $\forall t$, and $\forall i$, so that the quantities inside the absolute values are positive. Thus:

$$\frac{1+S_i^x}{1-S_i^x} = \frac{1+S_i^x(0)}{1-S_i^x(0)}\exp(J_{ij}\gamma S_j^x(0)t), \tag{A.39}$$

and letting $M \coloneqq (1+S_i^x(0))/(1-S_i^x(0))$, we obtain:

$$S_i^x(t) = \frac{M\exp(J_{ij}\gamma S_j^x(0)t)-1}{M\exp(J_{ij}\gamma S_j^x(0)t)+1}. \tag{A.40}$$

Further, to solve (3.63), namely:

$$\left(\frac{\dot{S_i^y}}{S_i^z}\right) = -\frac{1}{2} J_{ij} S_j^x(0)\left(1+\left(\frac{S_i^y}{S_i^z}\right)^2\right), \tag{A.41}$$

we set $v(t) \coloneqq S_i^y(t)/S_i^z(t)$, and obtain:

$$\int_{v(0)}^{v} \frac{\mathrm{d}w}{1+w^2} = -\frac{1}{2} J_{ij} S_j^x(0) \int_0^t \mathrm{d}t'. \tag{A.42}$$

On the left-hand side we recognize the anti-derivative of $\arctan(\cdot)$, and so:

$$\arctan\left(\frac{v - v(0)}{1 + v v(0)}\right) = -\frac{1}{2} J_{ij} S_j^x(0) t. \tag{A.43}$$

Applying $\tan(\cdot)$ to both sides and solving for $v(t)$, we finally arrive at:

$$v(t) = \frac{v(0)\cos\left(\frac{1}{2}J_{ij}S_j^x(0)t\right) - \sin\left(\frac{1}{2}J_{ij}S_j^x(0)t\right)}{\cos\left(\frac{1}{2}J_{ij}S_j^x(0)t\right) + v(0)\sin\left(\frac{1}{2}J_{ij}S_j^x(0)t\right)}, \tag{A.44}$$

where we have used the fact that $\tan(\cdot)$ is an odd function.

Finally, we solve (3.65), namely:

$$\dot{S}_i^z = \frac{1}{2} J_{ij} g(t) S_i^z S_j^x(0) - \frac{1}{2} J_{ij} \gamma S_i^x S_i^z S_j^x(0), \tag{A.45}$$

by taking S_i^z on the left-hand side and integrating:

$$\int_{S_i^z(0)}^{S_i^z} \frac{\mathrm{d}w}{w} = \frac{1}{2} J_{ij} S_j^x(0) \int_0^t g(r)\mathrm{d}r - \frac{1}{2} J_{ij}\gamma S_j^x(0) \int_0^t S_i^x(r)\mathrm{d}r. \tag{A.46}$$

The numerator of g is exactly the derivative of the denominator with respect to r up the constant in front of the corresponding integral, and so:

$$\frac{1}{2} J_{ij} S_j^x(0) \int_0^t g(r)\mathrm{d}r = \ln\left(\cos\left(\frac{1}{2}J_{ij}S_j^x(0)t\right) + v(0)\sin\left(\frac{1}{2}J_{ij}S_j^x(0)t\right)\right). \tag{A.47}$$

Upon manipulation, the second integral becomes:

$$-\frac{1}{2} J_{ij}\gamma S_j^x(0) \int_0^t S_i^x(r)\mathrm{d}r = -\frac{1}{2} J_{ij}\gamma S_j^x(0) \int_0^t \left(1 - \frac{2}{M\exp(J_{ij}\gamma S_j^x(0)r) + 1}\right)\mathrm{d}r \tag{A.48}$$

$$= -\frac{1}{2} J_{ij}\gamma S_j^x(0) t - \int_0^t \frac{-J_{ij}\gamma S_j^x(0)\exp(-J_{ij}\gamma S_j^x(0)r)}{M + \exp(-J_{ij}\gamma S_j^x(0)r)}\mathrm{d}r \tag{A.49}$$

$$= -\frac{1}{2} J_{ij}\gamma S_j^x(0) t - \ln\left(\frac{M + \exp(-J_{ij}\gamma S_j^x(0)r)}{M+1}\right). \tag{A.50}$$

Finally,

$$\int_{S_i^z(0)}^{S_i^z} \frac{\mathrm{d}w}{w} = \ln S_i^z - \ln S_i^z(0), \tag{A.51}$$

and combining everything together yields:

$$S_i^z(t) = S_i^z(0)\exp(-\frac{1}{2}J_{ij}\gamma S_j^x(0)t)\frac{1+M}{M+\exp(-J_{ij}\gamma S_j^x(0)t)}\cos\left(\frac{1}{2}J_{ij}S_j^x(0)t\right) + v(0)\sin\left(\frac{1}{2}J_{ij}S_j^x(0)t\right) \tag{A.52}$$

$$= S_i^z(0)\frac{1+M}{\exp(-\frac{1}{2}J_{ij}\gamma S_j^x(0)) + M\exp(\frac{1}{2}J_{ij}\gamma S_j^x(0))}\left(\cos\left(\frac{1}{2}J_{ij}S_j^x(0)t\right) + v(0)\sin\left(\frac{1}{2}J_{ij}S_j^x(0)t\right)\right). \tag{A.53}$$

A.6 Appendix 6

Theorem A.1. *[Bir31] For a measurable flow $(T_t)_{t\geq 0}$ in a σ-finite measure space (X,Σ,μ), and for any function $f \in L^1(X,\Sigma,\mu)$, if $\mu(X) < \infty$, then:*

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1} f(T^k x) = \frac{1}{\mu(X)}\int f \mathrm{d}\mu. \tag{A.54}$$

A.7 Appendix 7

If we consider the rigid-body Hamiltonian $\mathcal{H} = \frac{1}{2}S\cdot DS$, then:

$$\langle\mathcal{H}\rangle_{\rho_\beta} = \frac{3}{2}T, \tag{A.55}$$

where $\langle\cdot\rangle_{\rho_\beta}$ denotes the average with respect to the stationary distribution ρ_β. The general form of the Fokker-Planck equation [LM16] is:

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} = \mathcal{L}^+\rho. \tag{A.56}$$

For the Langevin equation (3.4), setting $\eta \coloneqq 2\gamma\beta^{-1}$, where $\beta = 1/(k_B T)$ we obtain that:

$$\mathcal{L}^+\rho = -\nabla\cdot\{S\times\nabla_S\mathcal{H}\rho + \gamma S\times(S\times\nabla_S\mathcal{H})\,\rho\} - \gamma\beta^{-1}\sum_{i,j\in\{x,y,z\}}\frac{\partial^2}{\partial S_i\partial S_j}(b_{ij}\rho), \tag{A.57}$$

where

$$B \coloneqq (b_{ij}) = \hat{S}^2 \quad \text{and} \quad \hat{S} = \mathrm{skew}(S) = \begin{pmatrix} 0 & S_z & -S_y \\ -S_z & 0 & S_x \\ S_y & -S_x & 0 \end{pmatrix}. \tag{A.58}$$

We observe that $\mathcal{L}^+ e^{-\beta\mathcal{H}} = 0$, and so $\rho_\beta = e^{-\beta\mathcal{H}}$ is the stationary distribution restricted to the unit sphere. Hence, by Stokes' Theorem, since $\partial\mathcal{S}^2 = 0$, it follows that:

$$\int_{\mathcal{S}^2} S\cdot\nabla\mathcal{H}e^{-\beta\mathcal{H}}\mathrm{d}\sigma_{\mathcal{S}^2} = \frac{1}{\beta}\int_{\mathcal{S}^2}(\nabla\cdot S)e^{-\beta\mathcal{H}}\mathrm{d}\sigma_{\mathcal{S}^2} = \frac{3}{\beta}\int_{\mathcal{S}^2} e^{-\beta\mathcal{H}}\mathrm{d}\sigma_{\mathcal{S}^2}, \tag{A.59}$$

because $\nabla\cdot S = 3$, and hence:

$$\langle S\cdot\nabla\mathcal{H}\rangle_{\rho_\beta} = \frac{3}{\beta}, \tag{A.60}$$

or, equivalently:

$$\langle\mathcal{H}\rangle_{\rho_\beta} = \frac{3}{2}T, \tag{A.61}$$

where we assume that Boltzmann's constant $k_B = 1$.