Designing and Probing Open Quantum Systems: Quantum Annealing, Excitonic Energy Transfer, and Nonlinear Fluorescence Spectroscopy

Dissertation advisor
Alán Aspuru-Guzik

Author
Alejandro Perdomo

Designing and Probing Open Quantum Systems: Quantum Annealing, Excitonic Energy Transfer, and Nonlinear Fluorescence Spectroscopy

Abstract

The 20th century saw the first revolution of quantum mechanics, setting the rules for our understanding of light, matter, and their interaction. The 21st century is focused on using these quantum mechanical laws to develop technologies which allows us to solve challenging practical problems. One of the directions is the use quantum devices which promise to surpass the best computers and best known classical algorithms for solving certain tasks. Crucial to the design of realistic devices and technologies is to account for the open nature of quantum systems and to cope with their interactions with the environment. In the first part of this dissertation, we show how to tackle classical optimization problems of interest in the physical sciences within one of these quantum computing paradigms, known as quantum annealing (QA). We present the largest implementation of QA on a biophysical problem (six different experiments with up to 81 superconducting quantum bits). Although the cases presented here can be solved on a classical computer, we present the first implementation of lattice protein folding on a quantum device under the Miyazawa-Jernigan model. This is the first step towards studying optimization problems in biophysics and statistical mechanics using quantum devices.

In the second part of this dissertation, we focus on the problem of excitonic energy transfer. We provide an intuitive platform for engineering exciton transfer dynamics and we show that careful consideration of the properties of the environment leads to opportunities to engineer the transfer of an exciton. Since excitons in nanostructures are proposed for use in quantum information processing and artificial photosynthetic designs, our approach paves the way for engineering a wide range of desired exciton dynamics. Finally, we develop the theory for a two-dimensional electronic spectroscopic technique based on fluorescence (2DFS) and challenge previous theoretical results claiming its equivalence to the two-dimensional photon echo (2DPE) technique which is based on polarization. Experimental realization of this technique confirms our theoretical predictions. The new technique is more sensitive than 2DPE as a tool for conformational determination of excitonically coupled chromophores and offers the possibility of applying two-dimensional electronic spectroscopy to single-molecules.

Contents

Abstract . . . iii
Table of contents . . . iv
Citations to previously published work . . . vi
Acknowledgments . . . vii

1 Introduction **1**

I Quantum annealing for lattice protein folding **6**

2 Construction of Hamiltonians for adiabatic quantum computation **7**
2.1 Introduction . . . 7
2.2 An adiabatic quantum algorithm for the HP model . . . 9
2.3 Construction of the lattice protein Hamiltonian for adiabatic quantum computation . . . 14
2.4 Case study: HPPH . . . 22
2.5 Converting an N-local Hamiltonian to a 2-local Hamiltonian . . . 25
2.6 Resources needed for a 2-local Hamiltonian expression in protein folding . . . 32
2.7 Conclusions . . . 33
2.8 Appendix: Extension of the pairwise interaction to three dimensions and N amino acids , $N = 2^M$ and $M \geq 3$ . . . 33

3 A study of heuristic guesses for adiabatic quantum computation **36**
3.1 Introduction . . . 36
3.2 Conventional adiabatic quantum computation (CAQC) . . . 40
3.3 Sombrero adiabatic quantum computation (SAQC) . . . 41
3.4 Hamiltonians for 3-SAT, numerical calculations and discussion . . . 44
3.5 Conclusions . . . 54

4 Lattice protein models by quantum-annealing **57**
4.1 Introduction . . . 57
4.2 Mapping of the lattice folding problem to the quantum hardware . . . 59
4.3 Results and discussion . . . 67
4.4 Conclusion . . . 68
4.5 Appendix 1. Transformation of the energy function of the lattice-folding model into the experimentally realizable spin-glass Hamiltonian . . . 68

4.6 Appendix 2. Embbedding of problem instances into the quantum hardware 73
4.7 Appendix 3. Experimental details 78
4.8 Appendix 4. Quantum simulations 83

II Excitonic energy transfer and nonlinear fluorescence spectroscopy 90

5 Engineering directed excitonic energy transfer 91
5.1 Introduction 91
5.2 Model Hamiltonian for the excitonic energy transfer dynamics 92
5.3 Conclusion 97

6 Phase-modulation 2D fluorescence spectroscopy 99
6.1 Introduction 99
6.2 Exciton-Coupled Dimer of Three-Level Molecules 102
6.3 Comparison between PM-2D FS and 2D Photon Echo Spectroscopy (2D PE) 105
6.4 Results and discussion 110
6.5 Conclusions 116
6.6 Appendix 1. Liposome sample preparation 116
6.7 Appendix 2. Phase-Modulation 2D Fluorescence Spectroscopy 117
6.8 Appendix 3. Computational modeling 118
6.9 Appendix 4. Error Analysis and Propagation of Uncertainties in PM-2D FS Signals 122

7 Future directions 128

References 131

Citations to previously published work

Chapters 2-6 have, apart from minor modifications, appeared as the following publications:

> "Construction of model Hamiltonians for adiabatic quantum computation and its application to finding low-energy conformations of lattice protein models," A. Perdomo, C. Truncik, I. Tubert-Brohman, G. Rose, and A. Aspuru-Guzik, *Phys. Rev. A.* **78**, 012320 (2008)
>
> "A study of heuristic guesses for adiabatic quantum computation," A. Perdomo-Ortiz, S. Venegas-Andraca, and A. Aspuru-Guzik, *Quantum. Inf. Process.*, 10, 33 (2011).
>
> "Finding low-energy conformations of lattice protein models by quantum-annealing," A. Perdomo-Ortiz, N. Dickson, M. Drew-Brook, S. Boixo, G. Rose, and A. Aspuru-Guzik, Under *peer review* in Scientific Reports, 2012.
>
> "Engineering directed excitonic energy transfer," A. Perdomo, L.Vogt, A. Najmaie, and A. Aspuru-Guzik, *Appl. Phys. Lett.*, **96**, 093114 (2010).
>
> "Conformation of self-assembled porphyrin dimers in liposome vesicles by phase-modulation 2D fluorescence spectroscopy," G. A. Lott*, A. Perdomo-Ortiz*, J. K. Utterback, J. R. Widom, A. Aspuru-Guzik, and A. H. Marcus (*Co-first authors), *Proc. Natl. Acad. Sci.*, **108**, 16521-16526 (2011).

The following articles were written to provide in-depth material with further examples and other details not included in the main chapters.

> "Simulating chemistry using quantum computers," I. Kassal*, J. D. Whitfield*, A. Perdomo-Ortiz, M-H. Yung, and A. Aspuru-Guzik (*Co-first authors). *Annu. Rev. Phys. Chem.*, **62**, 185-207 (2011).
>
> "Construction of energy functions for self-avoiding walks and the lattice heteropolymer model: resource efficient encoding for quantum optimization," A. Perdomo-Ortiz, B. O'Gorman, and A. Aspuru-Guzik. *In preparation*, 2012.
>
> "Conformation and electronic energy transport in membrane supported self-assembled porphyrin dimers by two-dimensional fluorescence spectroscopy," A. Perdomo-Ortiz, G. A. Lott, A. H. Marcus, A. Aspuru-Guzik, *In preparation*, 2012.

Acknowledgments

When deciding which PhD program to join, I remember receiving a phone call from Alán Aspuru-Guzik, a prospective faculty in the theoretical Chemistry division of the Department of Chemistry and Chemical Biology at Harvard University. With an offer to start at Harvard next Fall, Alán was recruiting students to start his research group. From that phone conversation, I had a great impression of Alán's personality: full of contagious energy and ready to push big research projects. It is an honor for me to be in the group of his first four PhD students (together with James D. Whitfield, Ivan Kassal and Leslie Vogt), and to see all the fast developments in the group since the beginning of his career as a faculty member. I could not thank Alán more for being not only an excellent and engaging advisor, but also for being a superb person who cares about his people.

The work developed in this dissertation would have not been possible without many useful discussions and help from many collaborators. Besides interactions within the Aspuru-Guzik group, I also had the good fortune to learn from and work with researchers from external research teams: the group of Andrew H. Marcus at the University of Oregon, Geordie Rose and his team from the company D-wave in Vancouver, Canada, the group of Jianfeng Du at the University of Science and Technology of China, and the groups of Karl Berggren and Moungi Bawendi at MIT.

For part I, Quantum Annealing for Lattice Protein Folding, I would like to thank my coauthors: Colin Truncik, Geordie Rose, Ivan Tubert-Brohman, Ivan Kassal, James D. Whitfield, Man-Hong Yung, Marshall Drew-Brook, Neil Dickson, and Salvador E. Venegas-Andraca. Other researchers with whom I had useful discussions include Bryan O'Gorman, Jacob Biamonte, Lisa Choe, Mohammad Amin, Nicolas Quesada, Rolando Somma, Ryan Babbush, Sergio Boixo, William MacReady and Profs. Edward Farhi, Jianfeng Du, and Sam Gutmann.

For part II, Excitonic Energy Transfer and Nonlinear Fluorescence Spectroscopy, I would like to thank my coauthors: Ali Najmaie, Geoffrey A. Lott, James K. Utterback, Julia R. Widom, Leslie Vogt, and Prof. Andrew H. Marcus. Other researchers with whom I had useful discussions include Alex Eisfeld, Dmitrij Rappoport, Francesco Marsili, Jacob Krich, Joel Yuen, Patrick Rebentrost, Semion Saikin, Vitor Manfrinato, and Profs. Karl Berggren, Keith Nelson, Robert Silbey, Shaul Mukamel, and Moungi Bawendi.

I want to thank my good friend Ali Najmaie, a postdoctoral fellow in the early years of my PhD, with whom I learned not only about the science we were doing on a daily basis, but also discussed and shared common views about the way science should be approached: letting the beauty of the scientific results and its development prevail over the unfortunate

overselling, desperate competition, and unethical aspects one might encounter and see in the world of academia.

A special thank you to Prof. Andrew H. Marcus who became my second advisor in the last track of my PhD, and the person Alán and I scientifically interacted the most during the past two years. Andy's passion for science and his "hands-on" approach made me feel like I was working with a closely colleague who is available most of the time and up for discussions, regardless of the distance. Doing science with Andy was a great experience, and I also want to thank him for being a great mentor and providing great advice and support when needed.

I also want to give a special thank you to Eric J. Heller and Eugene Shakhnovich for being part of my PhD committee and for their advice during my PhD track. The yearly progress meetings with them were very productive and a great opportunity to steer the projects in the right direction. Putting all my "crazy" ideas on the table and hearing their insights on the various projects was extremely valuable. In general, from my PhD committee, I learned how to prioritize projects and to go after the most relevant scientific questions.

Finally, I want to thank those who served as mentors and who inspired me, guided me, and encouraged me to pursue a PhD program abroad: my brother Oscar Mario, and my professors and good friends Maria Y. Ramirez, Julio C. Arce and Fabio Zuluaga from my *alma mater*, Universidad del Valle in Cali, Colombia. Also my previous advisors, Prof. Jason McNeill (Clemson Univ.), Prof. John R. Reynolds (then at Univ. of Florida, now at GeorgiaTech), and Jean-Luc Brédas (GeorgiaTech) contributed a lot to my formation and to the opportunity of me being at Harvard. Their guidance and the opportunity to do research in their groups were extremely valuable experiences, which helped me to make a smoother transition to the PhD program.

The dedication of this thesis goes to my family: to my siblings Carolina, Sandra and Oscar Mario, and especially to my parents, Oscar E. Perdomo and Lilia Ortiz, who are the essence of who I am as a person. Without their intense efforts to provide me a basic education, this whole story and sequence of dreams-coming-true would have been impossible. The experience as a PhD student at Harvard has certainly been the best time of my life, with an exponential growing in my personal and academic maturity, and I thank my mentors and all my friends, mentioned and not mentioned here, for being part of this incredible journey.

To my family

Chapter 1

Introduction

The 20th century saw the first revolution of quantum mechanics, setting the rules for our understanding of light, matter, and the interaction between them. The 21st century is focused on using these quantum mechanical laws to develop technologies which allows us to solve challenging practical computational problems and to make processes more efficient, e.g., light-harvesting in solar cells.

Crucial to the design of realistic devices and technologies is to account for the open nature of quantum systems and to cope with their interactions with the environment. In this dissertation, we will present mainly two applications of tunable quantum systems where the inclusion of the environment is crucial to theoretically predict the behavior of the systems. Next, we introduce each problem and describe their parts: the *quantum systems* and their *environments*, also referred to as their *baths* .

Part I. Quantum annealing for lattice protein folding

The first application studied in this dissertation is computational in nature. The idea of using quantum devices to outperform the best computers and best known classical algorithms for solving certain tasks is at the core of the now well-established field of quantum computation [142]. The main thesis we develop in the first part of the dissertation is the possibility of harnessing quantum-mechanical effects (a.k.a quantum computation) to study biophysical and statistical mechanics problems such as finding the low-energy conformations of proteins.

There are several models of, or ways of formulating, quantum computation. Example of the available paradigms include the circuit model [54, 19, 142], the adiabatic model [69], topological quantum computing [106, 139], one-way quantum computing [154, 155], and quantum walks [102]. Although equivalent in their computational capabilities, different problems are solved more naturally in different models. For example, of all the quantum-

computational models, adiabatic quantum computation (AQC) is perhaps the most naturally suited for studying and solving optimization problems [67, 93].

Finding efficient ways to solve hard optimization problems is one of the most important and sought-after research topics given its prevalence in several disciplines, ranging from the social sciences to the physical and natural sciences. For example, the protein folding problem [58, 134, 111, 147] consists of finding the lowest free-energy configuration or equivalently the native structure of a protein given its amino acid sequence. Understanding how proteins fold and their three-dimensional structure-function relationship is crucial to the understanding of enzymes and for the treatment of misfolded-protein diseases such as Alzheimer's, Huntington's, and Parkinson's disease. Due to the high computational cost of modeling proteins in atomistic detail [26, 175], coarse-grained descriptions of the protein folding problem, such as those found in lattice models [111, 147, 134, 146], have been very useful to gain insight about the folding mechanisms.

The first part of this thesis builds upon an instance of AQC known as quantum annealing (QA), which harnesses quantum-mechanical effects to speed up the solving of classical optimization problems. In this model, the state of the quantum device performing the computation remains in its ground state throughout the computation. The Hamiltonian $H(t)$ of the computer is changed slowly from a simple initial Hamiltonian H_i to a final Hamiltonian H_f whose ground state encodes the solution to the computational problem. The adiabatic theorem states that if the variation of the Hamiltonian is sufficiently slow, the easy-to-prepare ground state of H_i will be transformed continuously into the ground state of H_f. Escaping from the ground state reduces the performance of the computation and the maximum rate of change for the interpolation between H_i and H_f is mostly determined by the energy gap between the ground and first excited states during the evolution [131, 9, 194, 193].

In experimental implementations on quantum devices, the quantum levels used for the computation (the quantum system) are subject to noise and other uncontrollable degrees of freedom (the environment) which couple to the quantum system affecting its dynamics and the expected results from a purely coherent case, i.e., in the absence of an environment. For example, within the QA paradigm, maintaining the quantum system mostly in its ground state is ideal. Therefore, environmentally-induced excitations reducing the population of the ground state to other excited states would reduce the success probability of the computational process. The same environment might be responsible for returning these excitations back to the ground state, as part of a natural relaxation processes, therefore enhancing the robustness of the computation. Regardless of the nature of the interactions, excitations or relaxations, it is crucial to include them in a theoretical or numerical simulation of the quantum device dynamics, in order to properly characterize and predict its performance.

For example, the excellent agreement between theory and experiment presented in the last chapter of this first part (Chapter 4) would have not being possible without describing the quantum device within this framework of open quantum systems.

Proving the premise that we can solve optimization problems such as lattice protein folding in a quantum device running a quantum annealing protocol requires to overcome several theoretical and experimental challenges presented in detail throughout Chapers 2-4: from selecting a scalable experimental platform or quantum device capable of performing the QA algorithms described above, and understanding their limitations, to theoretically designing the algorithm to solve the problem instances.

The quantum hardware employed consists of an array of 128 coupled superconducting flux qubits [84, 97]. Physically, the quantum device represents an artificial Ising spin system with programmable spin-spin couplings and transverse magnetic fields, designed to solve instances of the following generally hard (technically called NP-hard [18]) classical optimization problem: given a set of local longitudinal fields $\{h_i\}$ and an interaction matrix $\{J_{ij}\}$, find the assignment $\mathbf{s}^* = s_1^* s_2^* \cdots s_N^*$, that minimizes the objective function $E(\mathbf{s})$, where

$$E(\mathbf{s}) = \sum_{1 \leq i \leq N} h_i s_i + \sum_{1 \leq i < j \leq N} J_{ij} s_i s_j, \tag{1.1}$$

$|h_i| \leq 1$, $|J_{ij}| \leq 1$, and $s_i \in \{+1, -1\}$.

The first theoretical challenge consisting of efficiently mapping the hard computational problem of interest (e.g., lattice folding) to a classical spin-glass energy of the form of Eq. 1.1 is solved in Chapter 2. The strategies presented there go beyond the construction of the energy function for the lattice folding problem and can be used as general building blocks for constructing other Hamiltonians to solve optimization problems of physical, chemical, and biophysical interest on a quantum device programmed to run a QA algorithm.

Chapter 3 is a contribution toward better algorithmic strategies one can use to improve the performance of the QA protocol. In QA, the initial Hamiltonians conventionally have a uniform superposition as ground state. We diverge from this practice by introducing a simple form of heuristics: the ability to start the quantum evolution with a state which is a guess to the solution of the problem. With this goal in mind, we explain the viability of this approach and the needed modifications to the conventional QA algorithm. Besides the possibility of introducing educated guesses as initial states, the new strategy allows for the possibility of restarting a failed adiabatic process from the measured excited state as opposed to restarting from the full superposition of states as in conventional QA. The outcome of the measurement can be used as a more refined guess state to restart the adiabatic evolution. This concatenated restart process is another heuristic that the conventional QA strategy cannot capture.

Finally, in chapter 4 we present the largest implementation of QA on a biophysical problem (six different experiments of up to 81 superconducting quantum bits). Although the cases presented here can be solved on a classical computer, we present the first implementation of lattice protein folding on a quantum device under the Miyazawa-Jernigan model. This is the first step towards studying optimization problems in biophysics and statistical mechanics using quantum devices.

Excitonic energy transfer and nonlinear fluorescence spectroscopy

Excitonic energy transfer (EET) has been studied in systems as varied as quantum dot (QD) nanostructures [171, 50], polymer chains [43], and photosynthetic complexes [121, 38]. Many applications of EET would benefit from controlling exciton dynamics. Recent work has shown that environment-induced decoherence can alter exciton dynamics [166, 158, 151], although controlling the transfer direction has only been achieved using external potentials [90].

Chapter 5 provides an intuitive platform for engineering exciton transfer dynamics. It builds upon the idea of engineering exciton transfer by designing appropriate system-bath interactions [71, 35]. We show that it is possible to design experimentally realizable systems where the environment can be used to direct the flow of energy. Since excitons in nanostructures are proposed for use in quantum information processing and artificial photosynthetic designs, our approach paves the way for engineering a wide range of desired exciton dynamics. We carefully describe the validity of the model and use experimentally relevant material parameters to show counter-intuitive examples of a directed exciton transfer in a linear chain of quantum dots.

The last chapter of this dissertation, Chaper 6, deviates from the design and tuning aspects of open quantum systems, and focuses on one of the experimental ways to probe its properties by two-dimensional electronic spectroscopy. The field of two-dimensional electronic spectroscopy provides valuable experimental techniques for unveiling the excited-state dynamics of complex systems. These techniques have recently been used to study excitonic energy transfer in photosynthetic systems [31, 44] and in general to study relaxation and dissipation processes in mesoscopic systems [138, 1]. Because of its high information content, 2D ES presents previously undescribed possibilities to extract quantum information from molecular systems and to determine model Hamiltonian parameters [212]. For example, experiments by Hayes and Engel extracted such information for the Fenna-Matthews-Olsen light harvesting complex [89]. Recently, it was demonstrated by Brinks et al. that single molecule coherences can be prepared using phased optical pulses and detected using fluorescence [30]. The latter experiments exploit the inherent sensitivity of fluorescence

and demonstrate the feasibility of controlling molecular quantum processes at the single molecule level.

We develop the theory for a two-dimensional electronic spectroscopic technique based on fluorescence (2DFS) and challenged previous theoretical results claiming its equivalence to the two-dimensional photon echo (2DPE) technique, which is based on polarization. Experimental realization of this technique confirms our theoretical predictions. The new technique is more sensitive than 2DPE as a tool for conformational determination of excitonically coupled chromophores and offers a unique possibility towards two-dimensional electronic spectroscopy on single-molecules. Fluorescence-based strategies to two-dimensional electronic spectroscopy (2D ES), such as those presented in the current work, could provide a route to extract high purity quantum information from single molecules. It may also be a means to study molecular systems in the ultraviolet regime where background noise due to solvent-induced scattering limits ultrafast experiments.

Part I

Quantum annealing for lattice protein folding

Chapter 2

Construction of Hamiltonians for adiabatic quantum computation

Summary

In this chapter we explore the use of a quantum optimization algorithm for obtaining low-energy conformations of protein models. We discuss mappings between protein models and optimization variables, which are in turn mapped to a system of coupled quantum bits. General strategies are given for constructing Hamiltonians to be used to solve optimization problems of physical, chemical, or biological interest via quantum computation by adiabatic evolution. As an example, we implement the Hamiltonian corresponding to the hydrophobic- polar model for protein folding. Furthermore, we present an approach to reduce the resulting Hamiltonian to two-body terms gearing toward an experimental realization.

2.1 Introduction

Finding the ensemble of low-energy conformations of a peptide given its primary sequence is a fundamental problem of computational biology, commonly known as the protein folding problem [32, 36, 152, 174, 134, 57, 77, 47]. The native fold conformation is usually assumed to correspond to the global minimum of the protein's free energy (according to the so-called thermodynamic hypothesis [61]), although some exceptions have been proposed [17, 120]. Thus, the protein folding problem can be described as a global optimization problem. Algorithms for quantum computers have been developed for many applications such as factoring [177] and the calculation of molecular energies [15]. In this chapter, we investigate the approach of using an adiabatic quantum computer for folding a highly simplified protein model.

The HP (H: hydrophobic, P: polar) lattice model [119] is one of the simplest protein

models implemented. Still its accuracy in predicting some of the folding behaviour of real proteins has made it a useful benchmark for testing optimization algorithms such as simulated annealing [183], genetic algorithms [46, 45, 51, 195, 181], and ant colony optimization [176]. Other heuristic methods such as hydrophobic core threading [16], chain growth [22, 95], contact interactions [192], and hydrophobic zippers [211] have also been considered. The HP model has also been useful for a qualitative investigation of the nature of the folding process and the interactions between proteins. The HP model depicted in Fig. 2.1 is defined by three assumptions: 1) There are only two kinds of amino acids or residues, hydrophobic (H) and polar (P); 2) residues are placed on a grid (typically a square grid for the 2D model and a cubic grid for the 3D model); 3) the only interaction among amino acids is the favorable contact between two H residues that are not adjacent in the sequence. The energy of this interaction is defined as -1 in arbitrary units, representing a hydrophobic effect which tends to fold the protein in a way that aggregates the H residues in a predominantly hydrophobic core, and leaves the P residues at the surface of the protein. The search for the native conformation of the protein is represented by a self-avoiding walk on the grid.

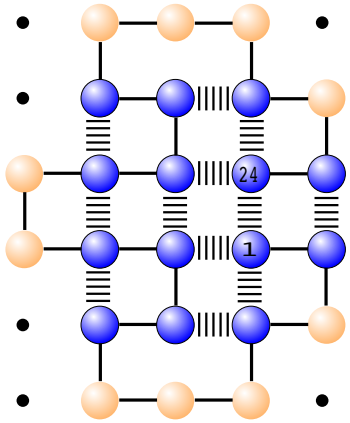

Figure 2.1: The lattice protein hydrophobic-polar (HP) model, showing the global energy minimum conformation for a sequence of 24 amino acids, HHPHPPPHHHHPPHHHHPP-PHPHH ($E = -12$). Blue (dark grey) beads represent hydrophobic residues (H) and orange (light grey) beads represent polar residues (P). The model consists of a self-avoiding chain with favorable ($E = -1$) energetic interactions among hydrophobic residues in contact. Contact between nearest neighbors in the primary sequence are unavoidable, and their contribution is not added to the calculated energy. Black dots represent lattice sites. Dotted lines represent favorable energetic interactions, solid lines represent the self-avoiding chain.

An important property of the model is that the number of possible conformations is roughly proportional to 2.7^N [119], where N is the length of the polypeptide chain. Proofs

of the NP-completeness of both the 2D and 3D HP models have been given [49, 20]. Due to this exponential growth, global optimality proofs become impractical when N reaches approximately 50 residues. For longer sequences, heuristics and stochastic algorithms have been employed for N up to 136 for the 3D HP model [192].

This chapter is structured as follows. Sec. 2.2 presents the general quantum algorithm and the terms of the Hamiltonian necessary to obtain the folded structure of the protein, and describes how to map the problem to arrays of coupled quantum bits [100, 80]. Sec. 2.3 explains the construction of the core component of the algorithm, the Hamiltonian that encodes the lowest energy conformation of the protein. In Sec. 2.4 we solve in detail the four amino acid sequence HPPH in a two-dimensional grid. In Sections 2.5 and 2.6 we discuss the resources necessary to carry out the reduction from a general k-body Hamiltonian to a two-body Hamiltonian, as a function of the size of the protein.

2.2 An adiabatic quantum algorithm for the HP model

We begin this section by describing the mapping of a sequence of N amino acids into binary variables, which will in turn be mapped to spin variables in the quantum mechanical version of the algorithm.

2.2.1 Mapping amino acids onto a lattice

The mapping of the coordinates of a sequence of N amino acids to a given grid of size $N \times N$ is developed as follows. We assume, without loss of generality, that the number of amino acids is a power of 2. A binary representation for the labels of the grid requires $\log_2 N$ binary variables to specify the position of an amino acid in each dimension, as shown in Fig. 2.2. The position of each of N amino acids in a D-dimensional lattice may thus be encoded by a bit string q composed of exactly $DN \log_2 N$ binary variables q_i. For example, for $N = 4$, $D = 2$, the length of the bit string q is 16 and therefore the number of configurations that can be explored is 2^{16}. Let q denote a particular configuration of the protein in the grid, written in the form

$$q = \underbrace{q_{16}q_{15}}_{y_4}\ \underbrace{q_{14}q_{13}}_{x_4}\ \underbrace{q_{12}q_{11}}_{y_3}\ \underbrace{q_{10}q_9}_{x_3}\ \underbrace{q_8q_7}_{y_2}\ \underbrace{q_6q_5}_{x_2}\ \underbrace{q_4q_3}_{y_1}\ \underbrace{q_2q_1}_{x_1}, \tag{2.1}$$

where x_i and y_i are the x and y coordinate of the i-th amino acid. Fig. 2.2 shows an example of the coordinate mapping given a specific sequence of residues or amino acids.

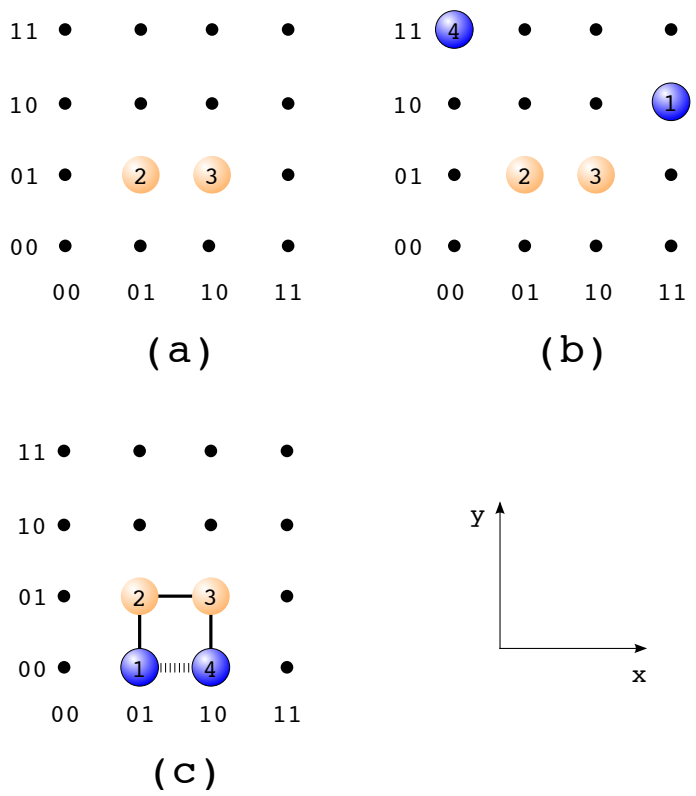

Figure 2.2: Grid-labeling conventions for a sequence of 4 amino acids, HPPH. **(a)** Amino acids 2 and 3 are fixed in the center of the grid to eliminate translational degeneracy. **(b)** One of the possible invalid configurations that might arise in the search and that would need to be discarded by the optimization algorithm. **(c)** Lowest-energy conformation for this example. The dotted line between amino acids 1 and 4 represents the hydrophobic interaction favored by the HP model. The configurations to optimize assume the form $q = q_{16}q_{15}q_{14}q_{13}$ 0110 0101 $q_4q_3q_2q_1$, where the set of variables $q_{16}q_{15}q_{14}q_{13}$ and $q_4q_3q_2q_1$ determine the position of amino acids 4 and 1, respectively. For the particular case in (b), $q = 1100\ 0110\ 0101\ 1011$.

In the quantum version of the problem, these configurations span a Hilbert space of dimension 2^{16}. The state vectors can be written as

$$|q\rangle \equiv |q_{16}\rangle\,|q_{15}\rangle \cdots |q_2\rangle\,|q_1\rangle\,. \tag{2.2}$$

We wish to implement a Hamiltonian which encodes the ground state of the protein on a spin-1/2 quantum computer [59], or, in particular onto an Ising-like Hamiltonian with a transverse magnetic field [29] (see Sec. 2.2.2). To do so, we realize the 16-qubit Hilbert space as a system of 16 spin-1/2 particles, with $|q_i = 0\rangle$ mapped to the spin state $|\sigma_i^z = +1\rangle$ and $|q_i = 1\rangle$ mapped to $|\sigma_i^z = -1\rangle$, with these spin states as the computational basis. In other words, the quantum version of the configuration states is related to spin variables

through the transformation

$$\hat{q}_i \equiv \frac{1}{2}(I - \hat{\sigma}_i^z), \tag{2.3}$$

with $I = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ and $\sigma^z = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$, the identity operator and the σ^z Pauli matrix represented in the computational basis, respectively.

In Sec. 2.3 we will derive an energy function in terms of the $ND\log_2 N$ binary variables used to describe all of the possible configurations for the N amino acids in a D-dimensional lattice. This energy function is constructed so that its minimum will yield the lowest-energy conformations of the protein. Eq. 2.3 provides the rule for the mapping of this energy function to a quantum Hamiltonian. Each q_i in the energy function will be replaced by an operator $\hat{q}_i$. The operator $\hat{q}_i$ is to be understood as a short hand notation for a quantum operator acting on the i-th qubit of the $ND\log_2 N$ multipartite Hilbert space, $\mathcal{H}_{NDlog_2N} \otimes \mathcal{H}_{NDlog_2N-1} \otimes \cdots \otimes \mathcal{H}_i \otimes \cdots \otimes \mathcal{H}_1$. The explicit form of $\hat{q}_i$ is given by $I \otimes I \otimes \cdots \otimes \hat{q}_i \otimes \cdots \otimes I$. Notice that the operator $\hat{q}_i$ as defined in Eq. 2.3 has been placed in the i-th position, and the identity operator acts on the rest of the Hilbert space. Products of the form $q_i q_j$ will be replaced by a quantum operator $\hat{q}_i\hat{q}_j$, which is a shorthand notation for the operators $\hat{q}_i$ and $\hat{q}_j$ acting on the i-th and the j-th qubits, respectively. As an illustrative example, consider an energy function dependent on four binary variables,

$$E(q_1, q_2, q_3, q_4) = 1 - q_1q_2 + q_1q_3 + q_2q_3q_4,$$

which will be mapped to a Hamiltonian acting on a four qubit Hilbert space, $\mathcal{H}_4 \otimes \mathcal{H}_3 \otimes \mathcal{H}_2 \otimes \mathcal{H}_1$. In the instance of this particular energy function the Hamiltonian will assume the form

$$\begin{aligned} \hat{H} &= I \otimes I \otimes I \otimes I - I \otimes I \otimes \hat{q} \otimes \hat{q} + I \otimes \hat{q} \otimes I \otimes \hat{q} + \hat{q} \otimes \hat{q} \otimes \hat{q} \otimes I \\ &\equiv I - \hat{q}_1\hat{q}_2 + \hat{q}_1\hat{q}_3 + \hat{q}_2\hat{q}_3\hat{q}_4. \end{aligned} \tag{2.4}$$

Following this mapping, transformation of any energy function to the quantum Hamiltonian is straightforward.

In order to eliminate redundancy due to translational symmetry, we fixed the two middle amino acids in a central position (see Fig. 2.2). This reduces the number of binary variables in the bit string from sixteen to eight. The variables corresponding to amino acids 1 and 4: $q_4q_3q_2q_1$ and $q_{16}q_{15}q_{14}q_{13}$, respectively, become the variables of interest, and the variables $q_8q_7q_6q_5$ and $q_{12}q_{11}q_{10}q_9$ corresponding to amino acids 2 and 3, become constant throughout the optimization process. In general, the $(N/2)^{th}$ amino acid is assigned to the $(N/2)^{th}$ grid point in all D dimensions. The $(N/2+1)^{th}$ amino acid is fixed to the $(N/2+1)^{th}$ grid point in the x direction and to the $(N/2)^{th}$ grid point in all other $D-1$ dimensions. As shown in Fig. 2.2, the final configuration we will try to optimize for the case of four amino acids takes the form $|q\rangle = |q_{16}q_{15}q_{14}q_{13}\rangle\,|0110\rangle\,|0101\rangle\,|q_4q_3q_2q_1\rangle$.

2.2.2 Adiabatic Quantum Computation

The goal of an adiabatic quantum algorithm is to transform an initial state into a final state which encodes the answer to the problem. A quantum state $|\psi(t)\rangle$ in the 2^n-dimensional Hilbert space for n qubits, evolves in time according to the Schrödinger equation

$$i\hbar\frac{d}{dt}|\psi(t)\rangle = \hat{H}(t)|\psi(t)\rangle, \tag{2.5}$$

where $\hat{H}(t)$ is the time-dependent Hamiltonian operator. The design of the algorithm takes advantage of the quantum adiabatic theorem [131], which is satisfied whenever $\hat{H}(t)$ varies slowly throughout the time of propagation $t \in [0,\tau]$. Let $|\psi_g(t)\rangle$ be the instantaneous ground state of $\hat{H}(t)$. If we construct $\hat{H}(t)$ such that the ground state of $\hat{H}(0)$, denoted as $|\psi_g(0)\rangle$, is easy to prepare, the adiabatic theorem states that the time propagation of the quantum state will remain very close to $|\psi_g(t)\rangle$ for all $t \in [0,\tau]$. One way to choose $\hat{H}(0)$ is to construct it in such a way that $|\psi_g(0)\rangle$ is a uniform superposition of all possible configurations of the system, i.e.

$$|\psi_g(0)\rangle = \frac{1}{\sqrt{2^n}}\sum_{q_i\in\{0,1\}}|q_n\rangle|q_{n-1}\rangle\cdots|q_2\rangle|q_1\rangle \tag{2.6}$$

summing over all 2^n vectors of the computational basis. Notice that an initial Hamiltonian of the form

$$\hat{H}(0) = \sum_{i=1}^{n}\hat{q}_x^i = \sum_{i=1}^{n}\frac{1}{2}(I-\hat{\sigma}_i^x) \tag{2.7}$$

would have as a non-degenerate ground state the vector $|\psi_g(0)\rangle$ defined in Eq. 2.6.

Similarly to the operator $\hat{q}$ from Eq. 2.3, we define

$$\hat{q}_x^i \equiv \frac{1}{2}(I-\hat{\sigma}_i^x), \tag{2.8}$$

with $I = \left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right)$ and $\sigma^x = \left(\begin{smallmatrix}0&1\\1&0\end{smallmatrix}\right)$, the identity operator and the σ^x-Pauli matrix represented in the computational basis, respectively.

For example, for the case of four qubits, $n=4$, $\hat{H}(0)$ is given by,

$$\hat{H}(0) = \sum_{i=1}^{4}\hat{q}_x^i = \hat{q}_x^1+\hat{q}_x^2+\hat{q}_x^3+\hat{q}_x^4 \tag{2.9}$$

$$= I\otimes I\otimes I\otimes\hat{q}_x + I\otimes I\otimes\hat{q}_x\otimes I + I\otimes\hat{q}_x\otimes I\otimes I + \hat{q}_x\otimes I\otimes I\otimes I. \tag{2.10}$$

To find the lowest energy conformation of the protein, one defines a Hamiltonian, $\hat{H}_{protein}$, whose ground state encodes the solution. Adiabatic evolution begins with $\hat{H}(0)$ and $|\psi_g(0)\rangle$, and ends in $\hat{H}_{protein} = \hat{H}(\tau)$. If the adiabatic evolution is slow enough, the state obtained

at time $t = \tau$ is $|\psi_g(\tau)\rangle$, the ground state of $\hat{H}(\tau) = \hat{H}_{protein}$. The details about the construction of $\hat{H}_{protein}$ will be provided in Sec. 2.3. A possible adiabatic evolution path can be constructed by the linear sweep of a parameter $t \in [0, \tau]$,

$$\hat{H}(t) = (1 - t/\tau)\hat{H}(0) + (t/\tau)\hat{H}_{protein}. \tag{2.11}$$

Even though Eq. 2.11 connects $\hat{H}(0)$ and $\hat{H}_{protein}$, determining the optimum value of τ is an important and non-trivial problem in itself. In principle, the adiabatic theorem states that over sufficient adiabatic time τ, the state $|\psi(\tau)\rangle$ will converge to the solution to the problem $|\psi_g(\tau)\rangle$. The magnitude of τ dictates the ultimate usefulness of the quantum algorithm proposed in this work. Farhi *et al.* [69, 67] showed promising numerical results for random instances of the Exact Cover computational problem.

Notice that the parameter τ determines the rate at which $\hat{H}(t)$ varies. Following the notation from Farhi *et al* [69], consider $\hat{H}(t) = \tilde{H}(t/\tau) = \tilde{H}(s)$, with instantaneous values of $\tilde{H}(s)$ defined by

$$\tilde{H}(s)\,|l; s\rangle = E_l(s)\,|l; s\rangle \tag{2.12}$$

with

$$E_0(s) \leq E_1(s) \leq \cdots \leq E_{N-1}(s) \tag{2.13}$$

where N is the dimension of the Hilbert space. According to the adiabatic theorem, if the gap between the two lowest levels, $E_1(s) - E_0(s)$, is greater than zero for all $0 \leq s \leq 1$, and taking

$$\tau \gg \frac{\varepsilon}{g_{min}^2} \tag{2.14}$$

with the minimum gap, g_{min}^2, defined by

$$g_{min} = \min_{0 \leq s \leq 1} (E_1(s) - E_0(s)), \tag{2.15}$$

and ε given by

$$\varepsilon = \max_{0 \leq s \leq 1} | \left\langle l = 1; s \left| \frac{d\tilde{H}}{ds} | l = 0; s \right. \right\rangle |, \tag{2.16}$$

then we can make

$$|\,\langle l = 0; s = 1\,|\psi(\tau)\rangle\,| \tag{2.17}$$

arbitrarily close to 1. In other words, the existence of a nonzero gap guarantees that $|\psi(t)\rangle$ remains very close to the ground state of $\hat{H}(t)$ for all $0 \leq t \leq \tau$, if τ is sufficiently large.

In the following sections, we derive the expression for an energy function which is mapped to $\hat{H}_{protein}$ using the procedure explained in Sec 2.2.1. The final expression for $\hat{H}_{protein}$ corresponds to an array of coupled qubits. We use H to denote both the Hamiltonians and the energy functions given that the mapping is straightforward as explained at the end of Sec. 2.2.1.

2.3 Construction of the lattice protein Hamiltonian for adiabatic quantum computation

Our goal in this section is to find an algebraic expression for an energy function in which the ground state represents the lowest energy conformation of a protein. Ideally, this energy function should contain the least possible number of terms. In order to optimize the computational resources, we desire terms with low locality, defined as the number of products of q_i's that appear in a certain term (e.g., a term of the form $h = q_1 q_3 q_4 q_6$ is 4-local).

2.3.1 Small computer science digression

Encoding positions of the amino acids in the grid in terms of Boolean variables makes it very convenient to use tools from computer science and basic Boolean algebra [165]. In this section, we will review these tools before using them to contruct arbitrary Hamiltonians that encode the spectrum of statistical mechanical models. We begin with some simple relations that are useful in the derivation of the Hamiltonian terms.

Consider two Boolean variables x and y. Expressions for the operations AND, OR, NOT can be written algebraically as:

$$\begin{aligned} f_{\text{AND}}(x,y) &= xy && \text{AND operation } (x \wedge y) \\ f_{\text{OR}}(x,y) &= x + y - xy && \text{OR operation } (x \vee y) \\ f_{\text{NOT}}(x) &= 1 - x && \text{NOT operation } (\neg x) \end{aligned}$$

An additional useful Boolean operator for the construction of Hamiltonian terms is XNOR. The output of the XNOR function is 0 unless all its arguments have the same value. The two-input version XNOR operation is also known as *logical equality*, here denoted as EQ,

$$f_{\text{EQ}}(x,y) = 1 - x - y + 2xy \qquad \text{XNOR operation}(x \text{ EQ } y)$$

The XNOR operator can be used to construct a very useful term for statistical mechanics Hamiltonians, an on-site repulsion penalty (described in Sec. 2.3.2 and illustrated in Fig. 2.3).

2.3.2 Hamiltonian terms for protein folding: the HP model

Most of the configurations represented by the bit strings q of Eq. 2.1 are invalid protein states. We seek a Hamiltonian that energetically favors valid configurations of the HP model by eliminating configurations in which more than one amino acid occupy the same

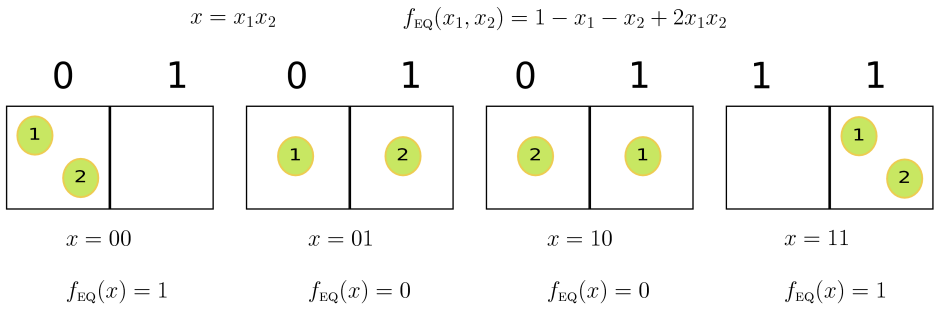

Figure 2.3: Illustrative example of one of the uses of the XNOR Boolean function in our scheme for the construction of Hamiltonians. Consider two particles 1 and 2 that are restricted to occupy either position 0 or 1 in the dimension shown, and let x_1 and x_2 encode the position particle 1 and particle 2 respectively. The Boolean function f_{EQ} can be interpreted as an onsite repulsion Hamiltonian which penalizes configurations where $x_1 = x_2$. The possible configurations are encoded in the bit string $x = x_1x_2$.

grid point, and discarding configurations that violate the primary sequence of amino acids. This Hamiltonian can be written as

$$H_{protein} = H_{onsite} + H_{psc} + H_{pairwise}, \tag{2.18}$$

where H_{onsite} is an onsite repulsion term for amino acids occupying the same grid point, H_{psc} is a primary sequence constraint term, and $H_{pairwise}$ is a pairwise interaction term that represents favorable hydrophobic interactions between adjacent hydrophobic amino acids.

Each protein configuration can be described by a string of $ND\log_2 N$ bits, where D is the number of dimensions and N is the number of amino acids. Without loss of generality, N is here contrained to be a power of two. Below, we describe each term in Eq. 2.18.

Onsite term, H_{onsite}

The first term in Eq. 2.18, H_{onsite}, prevents two or more amino acids from occupying the same grid point. For a given protein, at least one position variable must differ between each pair of amino acids for H_{onsite} to evaluate to zero. As an illustrative example, a simple one-dimensional two-site Hamiltonian is shown in Fig. 2.3 using the XNOR operation described in Sec. 2.3.1.

The general term for D dimensions and N amino acids is

$$H_{onsite}(N, D) = \lambda_0 \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} H^{ij}_{onsite}(N, D) \tag{2.19}$$

with

$$H^{ij}_{onsite}(N,D) = \prod_{k=1}^{D}\prod_{r=1}^{\log_2 N}\Big(1 - q_{f(i,k)+r} - q_{f(j,k)+r} + 2\, q_{f(i,k)+r}\, q_{f(j,k)+r}\Big) \tag{2.20}$$

and

$$f(i,k) = D(i-1)\log_2 N + (k-1)\log_2 N. \tag{2.21}$$

The terms enclosed by the parentheses in Eq. 2.20 are XNOR functions. The double product of these terms tests that all of these conditions are considered simultaneously by using AND relations. If all the binary variables describing the coordinates of the i-th and j-th amino acids are equal, then the series of products of XNOR functions is evaluated to $+1$. In this case, the energy penalty λ_0 with $\lambda_0 > 0$ is enforced. There will be no energy penalty, however, if even one of the binary variables for the i-th and j-th amino acids is different.

The function $f(i,k)$ is a pointer to the bit substring describing the coordinates of a particular amino acid. The index i points to the i-th amino acid and the index k points to the first bit variable of the k-th spatial coordinate. Here, $k = 1$ corresponds to the x coordinate, $k = 2$ to the y coordinate, and $k = 3$ to the z coordinate. For example, consider the case with $N = 4$ and $D = 2$. If we are interested in referring to the first binary variable describing the y coordinate ($k = 2$), for the third amino acid ($i = 3$), a direct substitution in Eq. 2.21 would yield $f(3,2) = 10$, which is indeed the variable we are interested in according to the convention established in Eq. 2.1.

Primary structure constraint, H_{psc}

The term H_{psc} in Eq. 2.18 evaluates to zero when two amino acids P and Q that are consecutive sequence-wise must be nearest neighbors on the lattice. Nearest-neighbors are defined as those points with a rectilinear (L_1) distance of $d_{PQ} = 1$ between them. We define a distance function that gives the base 10 distance squared between any two amino acids P and Q on the lattice,

$$d^2_{PQ}(N,D) = \sum_{k=1}^{D}\Big(\sum_{r=1}^{\log_2 N} 2^{r-1}(q_{f(P,k)+r} - q_{f(Q,k)+r})\Big)^2 \tag{2.22}$$

with $f(i,k)$ defined as in Eq. 2.21.

A simple way of defining H_{psc} is

$$H'_{psc}(N,D) = \lambda_1 \sum_{m=1}^{N-1}(1 - d^2_{m,m+1})^2 \tag{2.23}$$

Or, preferably,

$$H_{psc}(N,D) = \lambda_1\Big[-(N-1) + \sum_{m=1}^{N-1} d^2_{m,m+1}\Big]. \tag{2.24}$$

Unlike Eq. 2.23, the improved Hamiltonian in Eq. 2.24 is always 2-local regardless of the number of amino acids or the dimensionality of the problem, since $d^2_{PQ}(N,D)$ is always 2-local.

First, notice that for valid configurations, all $(N-1)$ terms in the sum will equal one, and $H_{psc}(N,D)$ evaluates to zero. If any of the $d^2_{m,m+1}$ terms is zero, meaning that two amino acids occupy the same location, then H_{onsite} will be drastically raised by the energy penalty λ_0. This can be achieved by setting $\lambda_0 > \lambda_1$, and $\lambda_1 = N$. After excluding configurations in which any $d^2_{m,m+1}$ are zero, only configurations with values of $d^2_{m,m+1} > 1$ are left. In these instances, $H_{psc}(N,D) > 0$ and λ_1 will play the role of an energy penalty since $\lambda_1 > 0$. Choosing $\lambda_1 = N$ and $\lambda_0 = N+1 > \lambda_1$ constrains unwanted or penalized configurations to eigenstates of $H_{protein}$ with energies greater than zero, while plausible configurations of the protein correspond to energies less than or equal to zero. Note that the minimum energy of the HP model, in the case of all hydrophobic sequences with the maximum number of favorable contacts, is always greater than $-N$. This is satisfied in general for N amino acids in either two or three dimensions.

Pairwise hydrophobic interaction term, $H_{pairwise}$

The HP model favors hydrophobic interactions by lowering the energy by 1 whenever non-nearest neighboring hydrophobic amino acids are a rectilinear distance of 1 away.

This kind of interaction is represented by the following general expression:

$$H_{pairwise}(N,D) = -\sum_{i=1}^{N}\sum_{j=1}^{N} G_{ij} H^{ij}_{pairwise} \tag{2.25}$$

Here G is an $N \times N$ symmetric matrix with entries G_{ij} equal to $+1$ when amino acids i and j are hydrophobic and non-nearest neighbors, and 0 otherwise. Note that G_{ij} is set to zero for amino acids that are neighbors in the protein sequence. Notice also that alternate definitions of G_{ij} could allow us to define lattice protein models that are more complex than the HP model. One of these models is the more realistic Miyazawa-Jernigan model [122] which includes interactions between 20 types of amino acids.

The form of $H^{ij}_{pairwise}$ depends on the spatial dimensionality of the problem. In two dimensions, we have

$$H^{ij}_{pairwise} = H^{ij,2D}_{pairwise}(N) = x^{ij,2D}_{+}(N) + x^{ij,2D}_{-}(N) + y^{ij,2D}_{+}(N) + y^{ij,2D}_{-}(N) \tag{2.26}$$

and in three dimensions,

$$\begin{aligned} H^{ij}_{pairwise} = H^{ij,3D}_{pairwise}(N) &= x^{ij,3D}_{+}(N) + x^{ij,3D}_{-}(N) + y^{ij,3D}_{+}(N) + y^{ij,3D}_{-}(N) \\ &\quad + z^{ij,3D}_{+}(N) + z^{ij,3D}_{-}(N) \end{aligned} \tag{2.27}$$

The terms on the right hand side of Eq. 2.27 are independent; each one serves to query whether the j-th amino acid is located, with respect with the i-th amino acid, to the right, left, above, below, in front, or behind as represented by $x^{ij,3D}_{+}$, $x^{ij,3D}_{-}$, $y^{ij,3D}_{+}$, $y^{ij,3D}_{-}$, $z^{ij,3D}_{+}$, and $z^{ij,3D}_{-}$ terms, respectively. If the j-th amino acid is located at a distance of exactly one in any direction, $H^{ij}_{pairwise}$ is set to +1; otherwise it is set to zero. There is a subtle but important condition embedded in these terms: they all vanish if the rightmost binary variable describing the i-th residue's coordinate of interest (say x for $x^{ij,3D}_{+}$ and $x^{ij,3D}_{-}$ or y for $y^{ij,3D}_{+}$ and $y^{ij,3D}_{-}$ or z for $z^{ij,3D}_{+}$ and $z^{ij,3D}_{-}$) does not end in 0, i.e., the coordinate has to correspond to an even number. This is why we intentionally double count each pair of amino acids in Eq. 2.25 by allowing both indexes i and j iterate from 1 to N. No special treatment is provided for the case where $i = j$, since the diagonal terms of G_{ij} are all zero due to the lack of amino acid self interaction. Finally, because we want the interaction to be attractive rather than repulsive, we use the minus sign in Eq. 2.25.

The case of N amino acids in a two dimensional grid for $N = 2^M$ and $M \geq 3$: The terms listed below correspond to the pairwise interaction Hamiltonian terms described above. The expressions below were constructed for $M \geq 3$. The four amino acid case ($M = 2$) is much simpler and will be discussed in Sec. 2.4. The expression for $x^{ij,2D}_{+}(N)$ is

$$\begin{aligned} x^{ij,2D}_{+}(N) &= (1 - q_{f(i,1)+1}) q_{f(j,1)+1} \prod_{s=2}^{\log_2 N} (1 - q_{f(j,1)+s} - q_{f(i,1)+s} + 2\, q_{f(j,1)+s}\, q_{f(i,1)+s}) \\ &\quad \prod_{r=1}^{\log_2 N} (1 - q_{f(i,2)+r} - q_{f(j,2)+r} + 2 q_{f(i,2)+r} q_{f(j,2)+r}) \end{aligned} \tag{2.28}$$

The first two factors of $x^{ij,2D}_{+}(N)$ (Eq. 2.28) treat the rightmost binary digit of the x position of the i-th and j-th amino acid. The first factor guarantees that the i-th residue is in an even position on the x-axis. For an interaction to be considered, the position of the j-th residue on the x-axis must be odd, as required by the second factor $q_{f(j,1)+1}$. The remaining factors of $x^{ij,2D}_{+}$ are XNOR functions that ensure that the rest of the binary digits that encode the x position are equal for the i-th and j-th amino acids. Finally, all the digits encoding the y position have to be equal, so that the i-th and j-th amino acids are nearest neighbors displaced only in the x-directionforcing the two residues to be in the same row. If all these conditions are satisfied, $x^{ij,2D}_{+}$ evaluates to +1; otherwise it evaluates to 0. These

conditions rely on the fact that adding 1 to an even number only changes the rightmost binary digit from 0 to 1.

The construction of $y_+^{ij,2D}$ follows the same procedure as that of $x_+^{ij,2D}$, namely,

$$\begin{aligned} y_+^{ij,2D}(N) \quad &= (1-q_{f(i,2)+1})q_{f(j,2)+1}\prod_{s=2}^{\log_2 N}(1-q_{f(j,2)+s}-q_{f(i,2)+s}+2\,q_{f(j,2)+s}\,q_{f(i,2)+s}) \\ &\prod_{r=1}^{\log_2 N}(1-q_{f(i,1)+r}-q_{f(j,1)+r}+2q_{f(i,1)+r}q_{f(j,1)+r}) \end{aligned} \tag{2.29}$$

The construction of $x_-^{ij,2D}$,

$$\begin{aligned} x_-^{ij,2D}(N) \;=\; & (1-q_{f(i,1)+1})q_{f(j,1)+1}\Big[1-\prod_{k=1}^{\log_2 N}(1-q_{f(i,1)+k})\Big](q_{f(j,1)+2}+q_{f(i,1)+2} \\ - \;& 2\,q_{f(j,1)+2}\,q_{f(i,1)+2})\prod_{r=3}^{\log_2 N}\Big[1-(q_{f(j,1)+r}+\prod_{u=2}^{r-1}q_{f(j,1)+u}-2\prod_{u=2}^{r}q_{f(j,1)+u}) \\ - \;& q_{f(i,1)+r}+2q_{f(i,1)+r}(q_{f(j,1)+r}+\prod_{u=2}^{r-1}q_{f(j,1)+u}-2\prod_{u=2}^{r}q_{f(j,1)+u})\Big] \\ & \prod_{s=1}^{\log_2 N}(1-q_{f(i,2)+s}-q_{f(j,2)+s}+2q_{f(i,2)+s}q_{f(j,2)+s}) \end{aligned} \tag{2.30}$$

involves several considerations. As in the expression for $x_+^{ij,2D}$, the first factor $(1-q_{f(i,1)+1})$ tests if the i-th amino acid is in an even position along the x-axis. Here, we are interested in querying whether the j-th amino acid is directly to the left of the i-th, and apply a different procedure than that of Eq. 2.28. We add $00\cdots01$ to the x coordinate of the j-th residue, thus moving "right" by one unit, and use the XNOR function to check if the result matches the x coordinate of the i-th amino acid. The problem is not as trivial as the case of $x_+^{ij,2D}$. Setting i at an even coordinate value along the axis of interest forces j to be in an odd coordinate. However, adding $00\cdots01$ to an odd binary number in general will change more digits than just the last digit due to carry bits. We used the circuit presented in Fig. 2.4 and the Boolean algebra introduced in Sec. 2.3.1 to obtain the general expression for the addition of $00\cdots01$ to an n-bit number. If we take $x = x_n x_{n-1}\cdots x_2 x_1$ and $y = 00\cdots01$, then the result $z = z_{n+1}z_n z_{n-1}\cdots z_2 z_1$ for the addition $z = x + y$ is the recursive algebraic

expression,

$$
\begin{aligned}
z_1 &= 0 \\
z_2 &= 1 - x_2 \\
z_k &= x_k + \prod_{u=2}^{k-1} x_u - 2 \prod_{u=2}^{k} x_u \quad \text{for} \quad 3 \leq k \leq n \\
z_{n+1} &= \prod_{u=2}^{n} x_u
\end{aligned}
$$

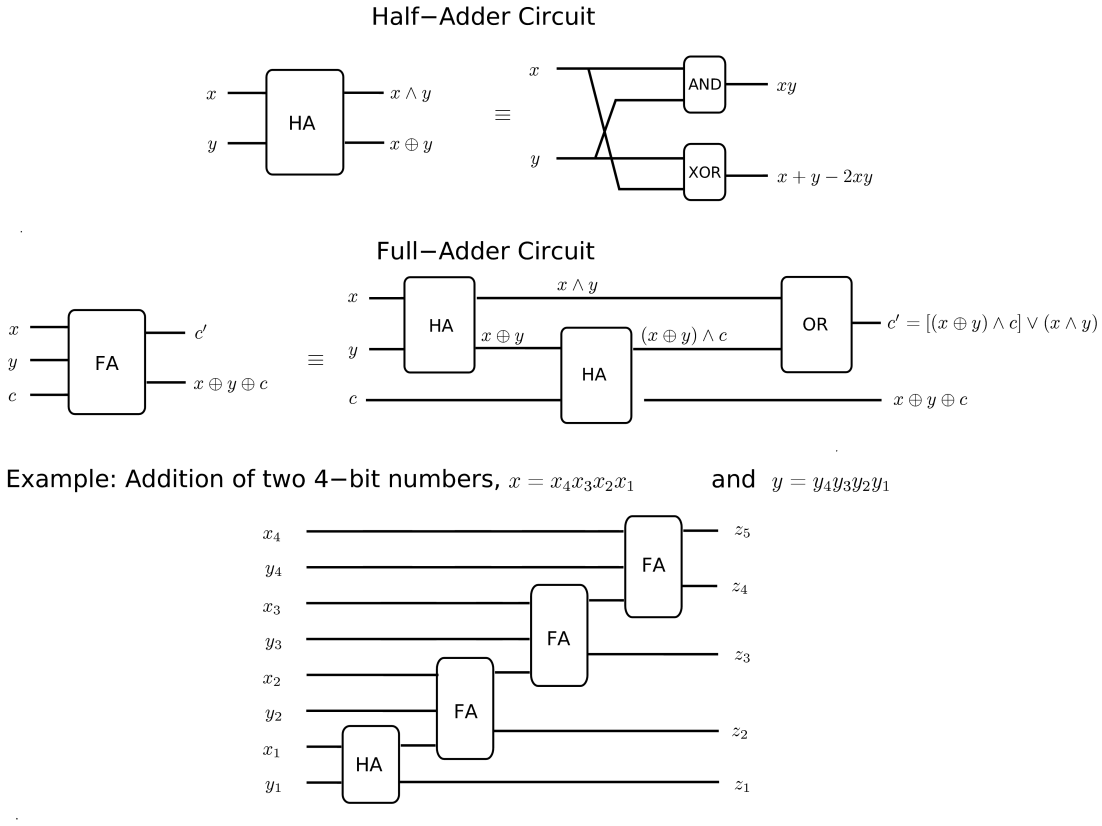

Figure 2.4: Half-adder and full-adder components for the addition circuit implemented in the pairwise interaction Hamiltonian. We show the implementation of these two components for the addition of two 4-bit numbers yielding $z = z_5 z_4 z_3 z_2 z_1$. The addition of n-bit numbers can be generalized trivially.

As in the case of $x_+^{ij,2D}$, we impose conditions that guarantee that the y coordinate is the same for both amino acids (that they are in the same row).

A special case arises when the j-th amino acid is at the rightmost position in the grid, with an x coordinate value of $11\cdots11$. When $00\cdots01$ is added to this coordinate, z_{n+1} evaluates to 1 and the n bits z_1 to z_n evaluate to 0. Since only the first n bits are used to compare coordinates, this z would be an undesirable match with an i-th amino acid positioned at $x = 00\cdots00$. Notice that a value of $x = 00\cdots00$ positions the i-th amino acid positioned at the minimal/leftmost position in the grid, for which $x_{-}^{ij,2D}$ should not even be considered. The factor $[1-\prod_{k=1}^{\log_2 N}(1-q_{f(i,1)+k})]$ in Eq. 2.30 sets the term $x_{-}^{ij,2D}$ to 0 if the x coordinate of the i-th amino acid is $00\cdots00$, taking care of both of these concerns.

The construction of $y_{-}^{ij,2D}$ follows the same procedure as that of $x_{-}^{ij,2D}$, namely,

$$\begin{aligned} y_{-}^{ij,2D}(N) \quad &= (1-q_{f(i,2)+1})q_{f(j,2)+1}\Big[1-\prod_{k=1}^{\log_2 N}(1-q_{f(i,2)+k})\Big](q_{f(j,2)+2}+q_{f(i,2)+2} \\ &- \; 2\,q_{f(j,2)+2}\,q_{f(i,2)+2})\prod_{r=3}^{\log_2 N}\Big[1-(q_{f(j,2)+r}+\prod_{u=2}^{r-1}q_{f(j,2)+u}-2\prod_{u=2}^{r}q_{f(j,2)+u}) \\ &- \; q_{f(i,2)+r}+2q_{f(i,2)+r}(q_{f(j,2)+r}+\prod_{u=2}^{r-1}q_{f(j,2)+u}-2\prod_{u=2}^{r}q_{f(j,2)+u})\Big] \\ &\prod_{s=1}^{\log_2 N}(1-q_{f(i,1)+s}-q_{f(j,1)+s}+2q_{f(i,1)+s}q_{f(j,1)+s}) \end{aligned} \tag{2.31}$$

The three-dimensional extension of these equations is presented in the Appendix.

2.3.3 Maximum locality and scaling of the number of terms in $H_{protein}$

In this section, we estimate the number of terms included in the total Hamiltonian $H_{protein}$ and present procedures required to reduce the locality of the terms to 2-local. These estimates assess the size of a quantum device necessary for eventual experimental realizations of the algorithm. The reduction of the locality of the terms involves ancillary qubits.

Each amino acid requires $D\log_2 N$ qubits to specify its position in the lattice. Since our algorithm fixes the position of two amino acids, the number of qubits needed to encode the coordinates of the $(N-2)$ remaining amino acids is $(N-2)D\log_2 N$. From the expressions given for H_{onsite}, H_{psc} and $H_{pairwise}$, one can deduce that the maximum locality is determined by $2D\log_2 N$ — the number of qubits corresponding to two amino acids. As described in Sec. 2.3.2, the H_{psc} term is always 2-local in nature regardless of the number of amino acids. For scaling arguments, it is crucial to point out that all possible 1-local and 2-local terms, that account for $(N-2)D\log_2 N$ and $\binom{(N-2)D\log_2 N}{2}$ total terms, repectively, appear in the expansion, but that not all possible 3-local or higher locality terms will be present. For example, the terms $q_iq_jq_k$, where the indexes i, j and k are associated with

Table 2.1: Scaling for number of quits as a function of problem size. The number of k-local terms obtained in the final expression for $H_{protein}$ as a function of the number of amino acids N, $N = 2^M$, and dimensions (D) of the lattice..

locality	Number of terms, T_k
$k = 0$	1
$k = 1$	$(N-2)D\log_2 N$
$2 \leq k \leq D\log_2 N$	$\binom{N-2}{2}\sum_{i=1}^{k-1}\binom{D\log_2 N}{i}\binom{D\log_2 N}{k-i} + (N-2)\binom{D\log_2 N}{k}$
$D\log_2 N < k \leq 2D\log_2 N$	$\binom{N-2}{2}\sum_{i=k-D\log_2 N}^{D\log_2 N}\binom{D\log_2 N}{i}\binom{D\log_2 N}{k-i}$
Total number of terms	$\sum_{k=0}^{2Dlog_2 N} T_k \sim N^{2D+2}$

three different amino acids, are not part of the expansion, since every term should only involve products of qubits describing two amino acids, regardless of its locality. Table 2.1 summarizes the number of k-local terms required to construct the protein Hamiltonian, $H_{protein} = H_{onsite} + H_{psc} + H_{pairwise}$. The alternative count from the combinatorial expressions of Table 2.1 scales as N^6 for $D = 2$ and as N^8 for $D = 3$. Table 2.1 provides the exact term count.

2.4 Case study: HPPH

With the goal of designing an experiment for adiabatic quantum computers with small numbers of qubits, we concentrate on the simplest possible instance of the HP-model – a four amino acid loop that contains a favorable interaction and therefore "folds".

In Sec. 2.4.1 we present the protein Hamiltonian, followed by the partitioning of the N-local Hamiltonian terms to 2-local. Finally, we present numerical simulations which confirm the local minimum through the use of the proposed algorithm.

2.4.1 Hamiltonian terms for the case of four amino acids in 2D

The onsite Hamiltonian for this example takes the form

Onsite term, H_{onsite}

$$H_{onsite}(N = 4, D = 2) = \lambda_0(H^{12}_{onsite} + H^{13}_{onsite} + H^{14}_{onsite} + H^{24}_{onsite} + H^{34}_{onsite}) \qquad (2.32)$$

with

$$H^{ij}_{onsite}(N=4,D=2)=\prod_{k=1}^{2}\prod_{r=1}^{2}\Big(1-q_{f(i,k)+r}-q_{f(j,k)+r}+2\,q_{f(i,k)+r}\,q_{f(j,k)+r}\Big) \tag{2.33}$$

and

$$f(i,k)=4(i-1)+2(k-1). \tag{2.34}$$

Note that H^{23}_{onsite} does not appear in Eq. 2.32 since, as described in Sec. 2.2.1, the two central amino acids are fixed in position and guaranteed not to occupy overlapping gridpoints that would contribute an energy penalty to the onsite term *a priori* . On the other hand, other terms involving amino acids 2 and 3 cannot be discarded, since these amino acids will affect their other neighbors through H_{psc} and they can participate in hydrophobic interactions through $H_{pairwise}$.

Primary structure constraint term, H_{psc}

The pairwise term

$$d^2_{PQ}(N=4,D=2)=\sum_{k=1}^{2}\Big(\sum_{r=1}^{2}2^{r-1}(q_{f(P,k)+r}-q_{f(Q,k)+r})\Big)^2 \tag{2.35}$$

with

$$H_{psc}(N=4,D=2)=\lambda_1\big(-3+d^2_{12}+d^2_{23}+d^2_{34}\big)=\lambda_1\big(-2+d^2_{12}+d^2_{34}\big) \tag{2.36}$$

takes advantage of the fact that $d^2_{23}=1$ by construction.

Pairwise term, $H_{pairwise}$

Finally, a pairwise interaction term is required to impose an energy stabilization for non-nearest neighbor hydrophobic amino acids that occupy adjacent sites in the lattice.

For the sequence HPPH,

$$G=\begin{pmatrix}0&0&0&1\\0&0&0&0\\0&0&0&0\\1&0&0&0\end{pmatrix} \tag{2.37}$$

and therefore,

$$H^{2D}_{pairwise}(N=4,D=2)=-(H^{14,2D}_{pairwise}+H^{41,2D}_{pairwise}). \tag{2.38}$$

For this particular case of interest

$$\begin{aligned}H^{ij,2D}_{pairwise}(N=4)=x^{ij,2D}_{+}(N=4)+x^{ij,2D}_{-}(N=4)+\\ y^{ij,2D}_{+}(N=4)+y^{ij,2D}_{-}(N=4).\end{aligned} \tag{2.39}$$

The explicit forms of these functions are:

$$x_+^{ij,2D}(N=4) = (1-q_{f(i,1)+1})q_{f(j,1)+1}(1-q_{f(j,1)+2}-q_{f(i,1)+2}+2\,q_{f(j,1)+2}q_{f(i,1)+2}) \prod_{s=1}^{2}(1-q_{f(i,2)+s}-q_{f(j,2)+s}+2q_{f(i,2)+s}q_{f(j,2)+s}), \quad (2.40)$$

$$y_+^{ij,2D}(N=4) = (1-q_{f(i,2)+1})q_{f(j,2)+1}(1-q_{f(j,2)+2}-q_{f(i,2)+2}+2\,q_{f(j,2)+2}\,q_{f(i,2)+s}) \prod_{s=1}^{2}(1-q_{f(i,1)+s}-q_{f(j,1)+s}+2q_{f(i,1)+s}q_{f(j,1)+s}), \quad (2.41)$$

$$x_-^{ij,2D}(N=4) = (1-q_{f(i,1)+1})q_{f(j,1)+1}q_{f(i,1)+2}(q_{f(j,1)+2}+q_{f(i,1)+2}-2q_{f(j,1)+2}q_{f(i,1)+2}) \prod_{s=1}^{2}(1-q_{f(i,2)+s}-q_{f(j,2)+s}+2\,q_{f(i,2)+s}q_{f(j,2)+s}), \quad (2.42)$$

$$y_-^{ij,2D}(N=4) = (1-q_{f(i,2)+1})q_{f(j,2)+1}q_{f(i,2)+2}(q_{f(j,2)+2}+q_{f(i,2)+2}-2\,q_{f(j,2)+2}q_{f(i,2)+2}) \prod_{s=1}^{2}(1-q_{f(i,1)+s}-q_{f(j,1)+s}+2\,q_{f(i,1)+s}q_{f(j,1)+s}). \quad (2.43)$$

After expanding all of the terms in H_{onsite}, H_{psc} and $H_{pairwise}$, we fix amino acids 2 and 3 as described in Sec. 2.2.1, substituting the variables $q_{12}q_{11}q_{10}q_9\ q_8q_7q_6q_5$ by the constant values 0110 0101 as shown in Fig. 2.2. The final expression for $H_{protein}$ now depends on the 8 binary variables encoding the coordinates of amino acids 1 and 4, $q_4q_3q_2q_1$ and $q_{16}q_{15}q_{14}q_{13}$, respectively. For convenience in notation, we relabel the coordinates of amino acid 4 from $q_{16}q_{15}q_{14}q_{13}$ to $q_8q_7q_6q_5$. After these substitutions, the final expression for the energy function $H_{protein}$ will be dependent on products involving the variables q_1 through q_8. Following the mapping explained at the end of Sec. 2.2.1, the quantum expression for $\hat{H}_{protein}$ is a $2^8 \times 2^8$ matrix. This Hamiltonian matrix defines the final Hamiltonian $\hat{H}(t=\tau)$ of the adiabatic evolution. The initial Hamiltonian representing the transverse field whose ground state is a linear superposition of all 2^8 states in the computational basis can be written as

$$\hat{H}_0 \equiv \hat{H}(t=0) = \sum_{i=1}^{8} \hat{q}_x^i = \sum_{i=1}^{8} \frac{1}{2}(I-\hat{\sigma}_i^x) \quad (2.44)$$

with

$$|\psi_g(t=0)\rangle = \frac{1}{\sqrt{2^8}} \sum_{q_i \in \{0,1\}} |q_8q_7q_6q_5q_4q_3q_2q_1\rangle \quad (2.45)$$

Finally, we can construct a time dependent Hamiltonian as shown in Eq. 2.11,

$$\hat{H}(t) = (1-t/\tau)\hat{H}_0 + (t/\tau)\hat{H}_{protein} \quad (2.46)$$

This time dependent Hamiltonian is also a $2^8 \times 2^8$ matrix as well. The instantaneous spectrum can be obtained by diagonalizing at every t/τ without need to specify τ. Since τ is the running time, we are interested in $0 \leq t/\tau \leq 1$. The spectrum of the corresponding $\hat{H}(t)$ for this four amino acid peptide HPPH is given in Fig. 2.5.

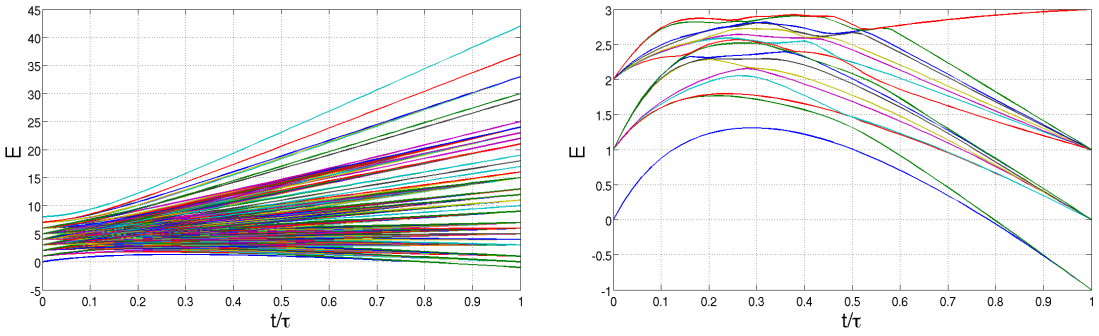

Figure 2.5: Spectrum of the instantaneous energy eigenvalues for the 8-local time dependent Hamiltonian used in the algorithm for the peptide HPPH (left). The plot to the right examines the lowest 15 states of the 256 states from the left.

Snapshots of the instantaneous ground state are shown in Fig. 2.6. Even though these snapshots do not correspond to explicit propagation of the Schrödinger equation, they indicate that the final $H_{protein}$ is correct and that it provides the correct answer if a sufficiently long time τ is allowed. Notice that at $t/\tau = 0$, the amplitude for all 256 states is equal, indicating a uniform superposition of all states; at $t/\tau = 1$, the readout corresponds to the two degenerate solutions of HPPH.

2.5 Converting an N-local Hamiltonian to a 2-local Hamiltonian

Motivated by the possibility of an experimental implementation, we explain how to reduce the locality of a Hamiltonian from k-local to 2-local while conserving its low-lying spectrum. We use Boolean reduction techniques [27, 23] for Hamiltonians contructed from energy functions with structure similar to $H_{protein}$, where all of terms are sums of tensor products of σ_z^i operators. By reducing the locality of the interactions, we introduce new ancilla qubits to represent higher order interactions with sums of at most 2-local terms. Here, we present an illutrative example with a relative simple energy function but the methodology can be immediately extended to higher locality energy functions such as the

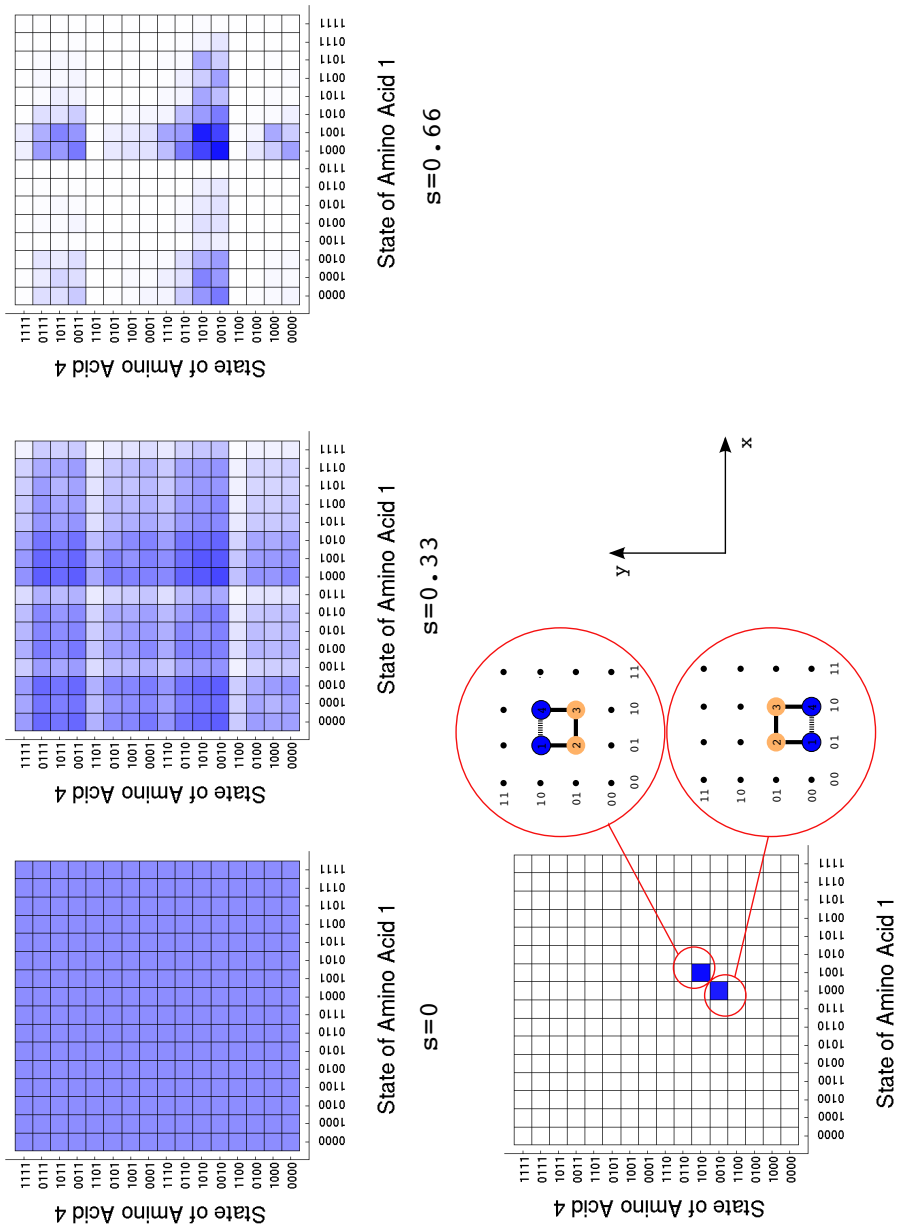

Figure 2.6: Snapshots of the instantaneous ground state for $H(t)$. The brightness of the box is proportional to $|c_n|^2$. Axis labels and state vectors for each particular box correspond to $|\psi\rangle = \sum_{n=0}^{255} c_n |n\rangle$ with $|n\rangle$ the n-th state vector out of the 256 possibilities given by $|q_{16}\rangle |q_{15}\rangle |q_{14}\rangle |q_{13}\rangle |q_4\rangle |q_3\rangle |q_2\rangle |q_1\rangle$. Notice that the x axis is given by $|q_4\rangle |q_3\rangle |q_2\rangle |q_1\rangle$ and the y axis given by $|q_{16}\rangle |q_{15}\rangle |q_{14}\rangle |q_{13}\rangle$. The final state corresponds to the two degenerate minima shown at the end

Table 2.2: Truth table for the energy function $H_{toy}(q) = 1+q_1-q_2+q_3+q_4-q_1q_2q_3+q_1q_2q_3q_4$.

q_4	q_3	q_2	q_1	$H(q_1, q_2, q_3, q_4)$
0	0	1	0	0
0	0	0	0	1
0	0	1	1	1
0	1	1	0	1
0	1	1	1	1
1	0	1	0	1
0	0	0	1	2
0	1	0	0	2
1	0	0	0	2
1	0	1	1	2
1	1	1	0	2
0	1	0	1	3
1	0	0	1	3
1	1	0	0	3
1	1	1	1	3
1	1	0	1	4

one resulting in $H_{protein}$.

Consider a 4-local energy function of the form

$$H_{toy}(q) = 1 + q_1 - q_2 + q_3 + q_4 - q_1q_2q_3 + q_1q_2q_3q_4. \tag{2.47}$$

As shown in Table 2.2, this energy function has a unique minimum energy given by $q = q_4q_3q_2q_1 = 0010$. The energy associated with this configuration is 0 in arbitrary units and all other possible values of the binary variables q_1, q_2, q_3 and q_4 have energies ranging from 0 to 4.

The goal is to obtain an energy function H' that preserves these energies along with their associated bit strings, but defines H' using only 1-local and 2-local terms. That is, the goal is to obtain a substitution for H_{toy} with the following form,

$$H'(\tilde{q}_1, \cdots, \tilde{q}_M) = c_0 + \sum_{i=1}^{M} c_i\tilde{q}_i + \sum_{i=1}^{M-1}\sum_{j=i+1}^{M} d_{ij}\tilde{q}_i\tilde{q}_j. \tag{2.48}$$

In Eq. 2.48 the new set of binary variables $\tilde{q}$ includes the original variables q_i as well as ancillary variables required to reduce locality. The extra ancillary bits raise the total number

Table 2.3: Truth table for the function $H_\wedge(q_i, q_j, \tilde{q}_n) = \delta(3\tilde{q}_n + q_i q_j - 2q_i\tilde{q}_n - 2q_j\tilde{q}_n)$ used for the locality reduction procedure described in Sec. 2.5.

$\tilde{q}_n$	q_i	q_j	$H_\wedge(q_i, q_j, \tilde{q}_n)$
0	0	0	0
0	0	1	0
0	1	0	0
1	1	1	0
1	0	0	3δ
1	0	1	δ
1	1	0	δ
0	1	1	δ

of variables to M.

Since the information contained within the problem and the solution we are seeking both rely on the original set of q variables (in the case of protein folding, for example, the string q encodes the positions of the amino acids in the lattice), we must be able to identify values corresponding to the original q, regardless of the substitutions made to convert a k-local function to a 2-local. The new energy function H' needs to have the energy values of the original function in its energy spectrum. In addition, the values of the bit string $\tilde{q}$ for these energies must match the same values of q in the original function. For the particular example of Eq. 2.47, consider the substitutions, $q_1 q_2 \to \tilde{q}_5$ and $q_3 q_4 \to \tilde{q}_6$. These two subtitutions introduce two new independent binary variables, $\tilde{q}_5$ and $\tilde{q}_6$ and regardless of the values of q_1, q_2, q_3 and q_4, they can take any value in $\{0, 1\}$. Since we want to preserve both the physical meaning of the original energy function, as well as its energy spectrum, we need to perform an action on the cases where the conditions $\tilde{q}_5 = q_1 \wedge q_2$ and $\tilde{q}_6 = q_3 \wedge q_4$ are not satisfied and lack any meaning in the context of the original energy function. One way to address this problem while keeping the original spectrum intact is to add a penalty function which enforces the conditions $\tilde{q}_5 = q_1 \wedge q_2$ and $\tilde{q}_6 = q_3 \wedge q_4$. For every substitution of the form $q_i q_j \to \tilde{q}_n$, consider a function of the form [23]

$$H_\wedge(q_i, q_j, \tilde{q}_n) = \delta(3\tilde{q}_n + q_i q_j - 2q_i\tilde{q}_n - 2q_j\tilde{q}_n). \tag{2.49}$$

As shown in Table 2.3, for $\delta > 0$, the function $H_\wedge(q_i, q_j, \tilde{q}_n)$ is greater than zero whenever $\tilde{q}_n \neq q_i \wedge q_j$ and it evaluates to zero whenever $\tilde{q}_n = q_i \wedge q_j$.

A two-local expression of the form presented in Eq. 2.48 can be obtained by adding one $H_\wedge(q_i, q_j, \tilde{q}_n)$ function for each substitution $q_1 q_2 \to \tilde{q}_5$ and $q_3 q_4 \to \tilde{q}_6$ and by making

the additional trivial substitutions $q_1 \to \tilde{q}_1$, $q_2 \to \tilde{q}_2$, $q_3 \to \tilde{q}_3$, and $q_4 \to \tilde{q}_4$, to conveniently change in notation to the set of binary variables $\tilde{q}$. For the case of the energy function of Eq. 2.47, the locality reduced version is

$$\begin{aligned} H_{toy,reduced}(\tilde{q}) &= 1 + \tilde{q}_1 - \tilde{q}_2 + \tilde{q}_3 + \tilde{q}_4 - \tilde{q}_5\tilde{q}_3 + \tilde{q}_5\tilde{q}_6 + H_\wedge(q_1, q_2, \tilde{q}_5) + H_\wedge(q_3, q_4, \tilde{q}_6) \\ &= 1 + \tilde{q}_1 - \tilde{q}_2 + \tilde{q}_3 + \tilde{q}_4 - \tilde{q}_5\tilde{q}_3 + \tilde{q}_5\tilde{q}_6 + \delta(3\tilde{q}_5 + \tilde{q}_1\tilde{q}_2 - 2\tilde{q}_1\tilde{q}_5 - 2\tilde{q}_2\tilde{q}_5) \\ &+ \delta(3\tilde{q}_6 + \tilde{q}_3\tilde{q}_4 - 2\tilde{q}_3\tilde{q}_6 - 2\tilde{q}_4\tilde{q}_6). \end{aligned} \tag{2.50}$$

Recall that the additional functions $H_\wedge(\tilde{q}_1, \tilde{q}_2, \tilde{q}_5)$ and $H_\wedge(\tilde{q}_3, \tilde{q}_4, \tilde{q}_6)$ increase the energy of $H_{toy,reduced}$ by at least δ whenever the conditions $\tilde{q}_5 = \tilde{q}_1 \wedge \tilde{q}_2$ and $\tilde{q}_6 = \tilde{q}_3 \wedge q_4$ are not satisfied. Table 2.4 shows the one-to-one mapping between the energies of non-penalized configurations of $H_{toy,reduced}(\tilde{q})$ and configurations presented in Table 2.2 associated with $H_{toy}(q)$. Even though there is a unique configuration $\{\tilde{q}_6 = \tilde{q}_3 \wedge \tilde{q}_4, \tilde{q}_5 = \tilde{q}_1 \wedge \tilde{q}_2, \tilde{q}_4, \tilde{q}_3, \tilde{q}_2, \tilde{q}_1\}$ associated with every $\{q_1, q_2, q_3, q_4\}$ with the same energy, it does not necessarily hold that the lowest 2^4 out of the 2^6 energies of $H_{toy,reduced}$ consist of the 2^4 energies of H_{toy}. For example, if we pick a small penalty δ in Table 2.4, say $0 \leq \delta \leq 4$, then some of the states penalized by either $H_\wedge(\tilde{q}_1, \tilde{q}_2, \tilde{q}_5)$ or $H_\wedge(\tilde{q}_3, \tilde{q}_4, \tilde{q}_6)$ can still have an energy within the energy values of H_{toy}. To avoid this situation, we can choose $\delta > \max(H_{toy})$ which will be sufficient to remove the energies of the penalized states from the region corresponding to energies of H_{toy}, therefore conserving the low-lying spectra of the original H_{toy}. Using the mapping explained at the end of Sec. 2.2.1, the quantum version of the 4-local energy function from Eq. 2.47 is:

$$\hat{H}_{toy} = I + \hat{\tilde{q}}_1 - \hat{\tilde{q}}_2 + \hat{\tilde{q}}_3 + \hat{\tilde{q}}_4 - \hat{\tilde{q}}_1\hat{\tilde{q}}_2\hat{\tilde{q}}_3 + \hat{\tilde{q}}_1\hat{\tilde{q}}_2\hat{\tilde{q}}_3\hat{\tilde{q}}_4. \tag{2.51}$$

The quantum version of the 2-local reduced form presented in Eq. 2.50 is,

$$\begin{aligned} \hat{H}_{toy,reduced} &= I + \hat{\tilde{q}}_1 - \hat{\tilde{q}}_2 + \hat{\tilde{q}}_3 + \hat{\tilde{q}}_4 - \hat{\tilde{q}}_5\hat{\tilde{q}}_3 + \hat{\tilde{q}}_5\hat{\tilde{q}}_6 + \delta(3\hat{\tilde{q}}_5 + \hat{\tilde{q}}_1\hat{\tilde{q}}_2 - 2\hat{\tilde{q}}_1\hat{\tilde{q}}_5 - 2\hat{\tilde{q}}_2\hat{\tilde{q}}_5) \\ &+ \delta(3\hat{\tilde{q}}_6 + \hat{\tilde{q}}_3\hat{\tilde{q}}_4 - 2\hat{\tilde{q}}_3\hat{\tilde{q}}_6 - 2\hat{\tilde{q}}_4\hat{\tilde{q}}_6) \end{aligned} \tag{2.52}$$

Notice that $\hat{H}_{toy}$ acts on a 2^4 dimensional Hilbert space, $\text{span}\{|\tilde{q}_4\rangle \otimes |\tilde{q}_3\rangle \otimes |\tilde{q}_2\rangle \otimes |\tilde{q}_1\rangle\}$, while $\hat{H}_{toy,reduced}$ acts on a 2^6 dimensional Hilbert space, $\text{span}\{|\tilde{q}_6\rangle \otimes |\tilde{q}_5\rangle \otimes |\tilde{q}_4\rangle \otimes |\tilde{q}_3\rangle \otimes |\tilde{q}_2\rangle \otimes |\tilde{q}_1\rangle\}$.

Due to the conservation of the spectrum and bit strings described above (as reflected in Tables 2.2 and 2.4), the solution obtained from an adiabatic quantum algorithm using either $\hat{H}_{toy}$ or $\hat{H}_{toy,reduced}$ as $\hat{H}_{final}$,

$$\hat{H}(t) = (1 - t/\tau)\hat{H}(0) + (t/\tau)\hat{H}_{final} \tag{2.53}$$

should be the same.

In the case of the 2-local Hamiltonian $\hat{H}_{toy,reduced}$, the solution to the optimization problem is obtained using an adiabatic algorithm after reading the qubits associated to

$\tilde{q}_4, \tilde{q}_3, \tilde{q}_2, \tilde{q}_1$ at $t = \tau$ from the space span$\{|\tilde{q}_6\rangle \otimes |\tilde{q}_5\rangle \otimes |\tilde{q}_4\rangle \otimes |\tilde{q}_3\rangle \otimes |\tilde{q}_2\rangle \otimes |\tilde{q}_1\rangle\}$ at $t = \tau$. Notice that the ancillary qubits in the six qubit version do not carry any physical information, as expected, since all of the valuable information was stored in the qubits coming from the original expression before the reduction. The cost of reducing the locality of a Hamiltonian to another which contains at most two-body interactions is the increase in the number of resources due to the additional ancillary bits.

Figure 2.7 shows the the eigenenergies of Eq. 2.53 vs. t/τ, where $\hat{H}_{final}$ is replaced by $\hat{H}_{toy}$ (see Figure 2.7(a)), and by $\hat{H}_{toy,reduced}$ with $\delta = 5$, (see Fig. 2.7(b)). As expected from Table 2.2 and 2.4, Fig. 2.7 illustrates the preservation of the subsystem corresponding to the variables $\tilde{q}_1, \tilde{q}_2, \tilde{q}_3$ and $\tilde{q}_4$ in the ground state of both the original and reduced-locality Hamiltonian. Degeneracy and overlap of lines in the spectra in Fig. 2.7 make it difficult to graphically convey that both spectra in Fig. 2.7 indeed have 16 states for $0 \leq$ eigenenergies ≤ 4. In Fig. 2.7(b) we plotted the first 19 eigenstates out of the 2^6 eigenstates corresponding to $\hat{H}_{toy,reduced}$. At $t/\tau = 1$, states with energy greater than 4 correspond to states which violate the AND condition introduced by the reduction process. Notice that there are two eigenstates with eigenvalue 5 in agreement with the table presented in Appendix **??** after substituting $\delta = 5$, and one state which corresponds to the one of the four-degenerate manifold with $E = 6$.

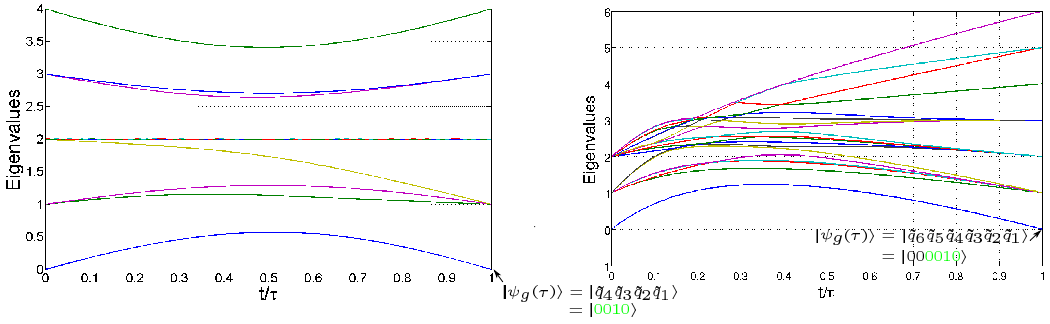

Figure 2.7: Spectrum comparison of the instantaneous energy eigenvalues for the 4-local toy Hamiltonian $\hat{H}_{toy}$ (left) and its corresponding 2-local version $\hat{H}_{toy,reduced}$(right). (left) Full spectrum of the 2^4 instantaneous eigenvalues for $\hat{H}_{toy}(\tilde{q},\tilde{q}_2, \tilde{q}_3, \tilde{q}_4)$. (right) First 19 instantaneous eigenvalues for the 2-local version of $\hat{H}_{toy}$, denoted as $\hat{H}_{toy,reduced}$ in text. The value used for δ is 5. The first 2^4 levels, $0 \leq$ eigenvalues ≤ 4, are associated to the original levels from $\hat{H}_{toy}$. The three remaining states with eigenvalues greater than 4 are penalized states which violate the conditions $\tilde{q}_n = \tilde{q}_i \wedge \tilde{q}_j$ (see Table 2.4 for details)

Table 2.4: Truth table for the energy function $H_{toy,reduced}(\tilde{q}) = 1+\tilde{q}_1-\tilde{q}_2+\tilde{q}_3+\tilde{q}_4-\tilde{q}_5\tilde{q}_3+\tilde{q}_5\tilde{q}_6+\delta(3\tilde{q}_5+\tilde{q}_1\tilde{q}_2-2\tilde{q}_1\tilde{q}_5-2\tilde{q}_2\tilde{q}_5)+\delta(3\tilde{q}_6+\tilde{q}_3\tilde{q}_4-2\tilde{q}_3\tilde{q}_6-2\tilde{q}_4\tilde{q}_6)$. The top of the table shows the 16 nonpenalized states that satisfy $\tilde{q}_5 = \tilde{q}_1 \wedge \tilde{q}_2$ and $\tilde{q}_6 = \tilde{q}_3 \wedge \tilde{q}_4$. These 16 states map one to one to the states in Table 2.2. A sample of the remaning 48 penalized states are shown in the second part of this table. The ellipses represent omitted rows.

$\tilde{q}_6$	$\tilde{q}_5$	$\tilde{q}_4$	$\tilde{q}_3$	$\tilde{q}_2$	$\tilde{q}_1$	$H'(\tilde{q}_1,\tilde{q}_2,\tilde{q}_3,\tilde{q}_4,\tilde{q}_5,\tilde{q}_6)$
0	0	0	0	1	0	0
0	0	0	0	0	0	1
0	1	0	0	1	1	1
0	0	0	1	1	0	1
0	0	0	1	1	1	1
0	0	1	0	1	0	1
0	0	0	0	0	1	2
0	0	0	1	0	0	2
0	0	1	0	0	0	2
0	1	1	0	1	1	2
1	0	1	1	1	0	2
0	0	0	1	0	1	3
0	0	1	0	0	1	3
1	0	1	1	0	0	3
1	1	1	1	1	1	3
1	0	1	1	0	1	4
0	1	0	0	1	0	δ
0	1	0	1	1	0	δ
0	0	0	0	1	1	$1+\delta$
0	1	1	0	1	0	$1+\delta$
0	1	1	0	1	0	$1+\delta$
1	0	0	1	1	0	$1+\delta$
1	0	1	0	1	0	$1+\delta$
0	0	0	1	1	1	$2+\delta$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
1	1	0	0	1	0	$1+4\ \delta$
1	1	0	1	0	0	$2+4\ \delta$
1	1	0	0	0	1	$3+4\ \delta$
1	1	1	0	0	0	$3+4\ \delta$
1	1	0	0	0	0	$2+6\ \delta$

2.6 Resources needed for a 2-local Hamiltonian expression in protein folding

For any k-local energy function, e.g., $h = q_1 q_2 \cdots q_k$, the reduction can be carried out iteratively, adding the penalty function $H_\wedge(q_i, q_j, \tilde{q}_n)$ for every substitution of the form $q_i q_j \to \tilde{q}_n$. For a k-local term, $(k-2)$ substitutions are required for the reduction to 2-local, and therefore require $(k-2)$ ancillary bits.

In the particular case of the protein Hamiltonian the reduction procedure needs to be repeated $(N-2)(N^D - D\log_2 N - 1)$ times, as described below. All the terms in the HP Hamiltonian include among at most interactions two amino acids, which results in a maximum locality of $2D\log_2 N$. In the following discussion, the cluster notation $[k][l]$ specifies the contributions of a particular $(k+l)-$local term into k variable coming from an amino acid with index i and l variables from an amino acid with index j. Since all the terms are of this form, to obtain a 2-local Hamiltonian, all products corresponding to each $[k]$ and $[l]$ of each cluster have to be converted to 1-local terms. We reduce terms for variables describing each amino acid in turn, for a total of $D\log_2 N$ variables. All possible combinations of two variables from the $D\log_2 N$ variables for an amino acid are substituted. The number of ancillary bits required for this substitution is $\binom{D\log_2 N}{2}$. These substitutions convert all terms of the form [3][0] and [2][1] to 2-local. To convert terms of the form [4][0] or [3][1] to 2-local we need to consider $\binom{D\log_2 N}{3}$ terms originally containing three variables from one amino acid. After employing an additional ancillary bit per term and applying the previous reduction step, all these terms collapse to 1-local with respect to the i-th amino acid, i.e., these terms will assume the form $[1][l]$. Iterating over the $D\log_2 N$ variables for a specific amino acid in order of increasing locality will give us the number of substitutions or ancilla bits needed per amino acid in order to reduce a particular cluster $[k]$ to [1] or 1-local. The total number of substitutions per amino acid corresponds to $\sum_{k=2}^{D\log_2 N}\binom{D\log_2 N}{k} = N^D - D\log_2 N - 1$. To carry out the procedure for all $(N-2)$ amino acids the number of ancilla qubits required is $(N-2)(N^D - D\log_2 N - 1)$. The number of qubits needed to represent a 2-local Hamiltonian version of the protein Hamiltonian is given by adding the number of ancillary qubits to the number of original $(N-2)D\log_2 N$ quantum bits,

$$\begin{aligned}\#\text{ of total qubits for a 2-local expression} &= (N-2)(N^D - D\log_2 N - 1) + (N-2)D\log_2 N \\ &= (N-2)(N^D - 1) \end{aligned} \tag{2.54}$$

Eq. 2.54 provides a closed formula for the number of qubits needed to find the lowest energy conformations for a protein with N amino acids in D dimensions in our encoding. In particular, for the case of a four amino acid peptide HPPH in two dimensions considered

in Sec. 2.4 requires 30 qubits.

2.7 Conclusions

We constructed the essential elements of an adiabatic quantum algorithm to find the lowest energy conformations of a protein in a lattice model. The number of binary variables needed to represent N amino acids on an $N \times N$ lattice is $(N-2)D\log_2 N$. The maximum locality of the final Hamiltonian, as determined by the interaction between pairs of amino acids using the mapping defined here, is $2D\log_2 N$.

General strategies to construct energy functions to map into other quantum mechanical Hamiltonians used for adiabatic quantum computing were presented. The strategies used in the construction of the Hamiltonian for the HP model can be used as general building blocks for Hamiltonians associated with physical systems where onsite energies and/or pairwise potentials are present.

We also demonstrated an application of the Boolean scheme for converting a k-local Hamiltonian into a 2-local Hamiltonian, aiming toward an experimental implementation in quantum devices. The resulting couplings, although 2-local, do not necessarily represent couplings among nearest neighbor quantum bits in a two-dimensional geometry. It is however known that the number of ancillary physical qubits required to embed an arbitrary N variable problem is upper-bounded by $N^2/(C-2)$, where C is the number of couplers allowed per physical qubit.

The most important question remaining to be explored in future work is the scaling of run time τ with respect to the number of amino acids N. Run time τ is dependent on the particular instance of the problem – in our case, to different protein sequences. It has been proposed that proteins have evolved towards a many-dimensional funnel-like potential energy surface [47]. The sequences that show a funnel-like structure might be easier to study using adiabatic quantum computation, because the funnel structure may facilitate annealing of the quantum wave function toward low energy conformations.

2.8 Appendix: Extension of the pairwise interaction to three dimensions and N amino acids , $N = 2^M$ and $M \geq 3$

This extension follows the principles presented in Sec. 2.3.2 and extends the terms of the Hamiltonian to the case of a three-dimensional lattice protein. The pairwise term for

the three-dimensional case is,

$$H^{3D}_{pairwise}(N) = -\sum_{i,j=1}^{N} G_{ij} H^{ij,3D}_{pairwise} \tag{2.55}$$

$$\begin{aligned} x_{+}^{ij,3D}(N) &= (1-q_{f(i,1)+1})q_{f(j,1)+1} \prod_{s=2}^{\log_2 N} (1 - q_{f(j,1)+s} - q_{f(i,1)+s} + 2\, q_{f(j,1)+s}\, q_{f(i,1)+s}) \\ &\quad \prod_{s=1}^{\log_2 N} (1 - q_{f(i,2)+s} - q_{f(j,2)+s} + 2q_{f(i,2)+s}q_{f(j,2)+s}) \prod_{r=1}^{\log_2 N} (1 - q_{f(i,3)+r} \\ &\quad - q_{f(j,3)+r} + 2q_{f(i,3)+r}q_{f(j,3)+r}), \end{aligned} \tag{2.56}$$

$$\begin{aligned} y_{+}^{ij,3D}(N) &= (1-q_{f(i,2)+1})q_{f(j,2)+1} \prod_{s=2}^{\log_2 N} (1 - q_{f(j,2)+s} - q_{f(i,2)+s} + 2\, q_{f(j,2)+s}\, q_{f(i,2)+s}) \\ &\quad \prod_{s=1}^{\log_2 N} (1 - q_{f(i,1)+s} - q_{f(j,1)+s} + 2q_{f(i,1)+s}q_{f(j,1)+s}) \prod_{r=1}^{\log_2 N} (1 - q_{f(i,3)+r} \\ &\quad - q_{f(j,3)+r} + 2q_{f(i,3)+r}q_{f(j,3)+r}), \end{aligned} \tag{2.57}$$

$$\begin{aligned} z_{+}^{ij,3D}(N) &= (1-q_{f(i,3)+1})q_{f(j,3)+1} \prod_{s=2}^{\log_2 N} (1 - q_{f(j,3)+s} - q_{f(i,3)+s} + 2\, q_{f(j,3)+s}\, q_{f(i,3)+s}) \\ &\quad \prod_{s=1}^{\log_2 N} (1 - q_{f(i,1)+s} - q_{f(j,1)+s} + 2q_{f(i,1)+s}q_{f(j,1)+s}) \prod_{r=1}^{\log_2 N} (1 - q_{f(i,2)+r} \\ &\quad - q_{f(j,2)+r} + 2q_{f(i,2)+r}q_{f(j,2)+r}), \end{aligned} \tag{2.58}$$

$$\begin{aligned} x_{-}^{ij,3D}(N) &= (1-q_{f(i,1)+1})q_{f(j,1)+1}\Big[1 - \prod_{k=1}^{\log_2 N} (1 - q_{f(i,1)+k})\Big](q_{f(j,1)+2} + q_{f(i,1)+2} \\ &\quad - 2\, q_{f(j,1)+2}\, q_{f(i,1)+2}) \prod_{r=3}^{\log_2 N} \Big[1 - (q_{f(j,1)+r} + \prod_{u=2}^{r-1} q_{f(j,1)+u} - 2\prod_{u=2}^{r} q_{f(j,1)+u}) \\ &\quad - q_{f(i,1)+r} + 2q_{f(i,1)+r}(q_{f(j,1)+r} + \prod_{u=2}^{r-1} q_{f(j,1)+u} - 2\prod_{u=2}^{r} q_{f(j,1)+u})\Big] \\ &\quad \prod_{s=1}^{\log_2 N} (1 - q_{f(i,2)+s} - q_{f(j,2)+s} + 2q_{f(i,2)+s}q_{f(j,2)+s}) \\ &\quad \prod_{r=1}^{\log_2 N} (1 - q_{f(i,3)+r} - q_{f(j,3)+r} + 2q_{f(i,3)+r}q_{f(j,3)+r}), \end{aligned} \tag{2.59}$$

$$
\begin{aligned}
y_{-}^{ij,3D}(N) &= (1-q_{f(i,2)+1})q_{f(j,2)+1}\Big[1-\prod_{k=1}^{\log_2 N}(1-q_{f(i,2)+k})\Big](q_{f(j,2)+2}+q_{f(i,2)+2} \\
&- 2\,q_{f(j,2)+2}\,q_{f(i,2)+2})\prod_{r=3}^{\log_2 N}\Big[1-(q_{f(j,2)+r}+\prod_{u=2}^{r-1}q_{f(j,2)+u}-2\prod_{u=2}^{r}q_{f(j,2)+u}) \\
&- q_{f(i,2)+r}+2q_{f(i,2)+r}(q_{f(j,2)+r}+\prod_{u=2}^{r-1}q_{f(j,2)+u}-2\prod_{u=2}^{r}q_{f(j,2)+u})\Big] \\
&\prod_{s=1}^{\log_2 N}(1-q_{f(i,1)+s}-q_{f(j,1)+s}+2q_{f(i,1)+s}q_{f(j,1)+s}) \\
&\prod_{r=1}^{\log_2 N}(1-q_{f(i,3)+r}-q_{f(j,3)+r}+2q_{f(i,3)+r}q_{f(j,3)+r}), \qquad (2.60)
\end{aligned}
$$

$$
\begin{aligned}
z_{-}^{ij,3D}(N) &= (1-q_{f(i,3)+1})q_{f(j,3)+1}\Big[1-\prod_{k=1}^{\log_2 N}(1-q_{f(i,3)+k})\Big](q_{f(j,3)+2}+q_{f(i,3)+2} \\
&- 2\,q_{f(j,3)+2}\,q_{f(i,3)+2})\prod_{r=3}^{\log_2 N}\Big[1-(q_{f(j,3)+r}+\prod_{u=2}^{r-1}q_{f(j,3)+u}-2\prod_{u=2}^{r}q_{f(j,3)+u}) \\
&- q_{f(i,3)+r}+2q_{f(i,3)+r}(q_{f(j,3)+r}+\prod_{u=2}^{r-1}q_{f(j,3)+u}-2\prod_{u=2}^{r}q_{f(j,3)+u})\Big] \\
&\prod_{s=1}^{\log_2 N}(1-q_{f(i,1)+s}-q_{f(j,1)+s}+2q_{f(i,1)+s}q_{f(j,1)+s}) \\
&\prod_{r=1}^{\log_2 N}(1-q_{f(i,2)+r}-q_{f(j,2)+r}+2q_{f(i,2)+r}q_{f(j,2)+r}). \qquad (2.61)
\end{aligned}
$$

Chapter 3

A study of heuristic guesses for adiabatic quantum computation

Summary

Adiabatic quantum computation (AQC) is a universal model for quantum computation which seeks to transform the initial ground state of a quantum system into a final ground state encoding the answer to a computational problem. AQC initial Hamiltonians conventionally have a uniform superposition as ground state. We diverge from this practice by introducing a simple form of heuristics: the ability to start the quantum evolution with a state which is a guess to the solution of the problem. With this goal in mind, we explain the viability of this approach and the needed modifications to the conventional AQC (CAQC) algorithm. By performing a numerical study on hard-to-satisfy 6 and 7 bit random instances of the satisfiability problem (3-SAT), we show how this heuristic approach is possible and we identify that the performance of the particular algorithm proposed is largely determined by the Hamming distance of the chosen initial guess state with respect to the solution. Besides the possibility of introducing educated guesses as initial states, the new strategy allows for the possibility of restarting a failed adiabatic process from the measured excited state as opposed to restarting from the full superposition of states as in CAQC. The outcome of the measurement can be used as a more refined guess state to restart the adiabatic evolution. This concatenated restart process is another heuristic that the CAQC strategy cannot capture.

3.1 Introduction

Adiabatic quantum computation (AQC) [69] is a promising paradigm of quantum computation because of its robustness [39, 123], and its intuitive mapping from NP-complete and NP-hard problems to potentially realizable Hamiltonians [69, 67, 93, 208]. Adiabatic

quantum computing is attractive because relevant optimization problems such as lattice models for protein folding can be readily formulated [148].

AQC algorithms involve the specification of a time-dependent Hamiltonian,

$$\hat{H}(t) = \hat{h}_i(t) + \hat{h}_{driving}(t) + \hat{h}_f(t), \tag{3.1}$$

This Hamiltonian has three important functions: (1) The initial Hamiltonian, $\hat{H}_i \equiv \hat{H}(0)$, encodes a ground state that is easy to prepare and that is used as the initial state for the quantum evolution. (2) The driving Hamiltonian, $\hat{h}_{driving}(t)$, is responsible for mediating the transformation of the initial ground state to any of other state. (3) The final Hamiltonian, $\hat{H}_f \equiv \hat{H}(\tau)$, is problem dependent and its ground state encodes the solution, $|\psi_{solution}\rangle$, to the computational problem. In the ideal case of a process being fully adiabatic, evolution under $\hat{H}(t)$ will keep the quantum state, $|\psi(t)\rangle$, in the ground state of $\hat{H}(t)$ throughout $0 < t < \tau$. If this condition is met, the final state at $t = \tau$ should coincide with the ground state of the final Hamiltonian, $\hat{H}_f$, i.e., $|\psi(\tau)\rangle = |\psi_{solution}\rangle$, if the process is adiabatic. The measurement at $t = \tau$ will provide the solution to the computational problem.

Since the original proposal by Farhi et al [69], a significant amount of progress has been made towards the design of final Hamiltonians for different computationally intractable problems such as NP-complete problems. For example, in this chapter we provide a detailed description of the construction of final Hamiltonian for the 3-SAT problem. In previous work, we described the respective construction of the final Hamiltonian for an NP-hard problem of interest in biology, the protein folding problem, which consists of finding the minimum energy configuration of a chain of interacting amino acids in a lattice [148]. Several choices for $\hat{h}_{driving}$ have been suggested. Farhi et al [69] proposed what we call in this work the conventional form used for $\hat{h}_{driving}(t)$: a linear ramp,

$$\hat{h}_{driving}(t) = (1 - t/\tau)\hat{H}_{driving}, \tag{3.2}$$

In Farhi *et al*'s scheme the term $\hat{h}_i(t)$ can be defined as $\hat{H}_i \equiv \hat{h}_{driving}(t = 0) = \hat{H}_{driving}$, namely, the term $\hat{H}_{driving}$ serves as the initial Hamiltonian $\hat{H}_i$ when its intensity is completely turned on at $t = 0$, and drives the quantum evolution until it is completely turned off at $t = \tau$. The final Hamiltonian is turned on with the functional form $\hat{h}_f(t) = (t/\tau)\hat{H}_f$. Scheduling the adiabatic evolution with this linear interpolation is not compulsory, thus different proposals have been studied such as the use of non-linear time-profile for auxiliary Hamiltonians [65] and optimal geodesic paths [161].

The question of whether AQC, or in general quantum computers, can be used for efficiently solving NP-complete problems is a difficult open question. Lessons that have

been learned [68] include that a poor choice of the initial Hamiltonian such as the one-dimensional projector as selected in Refs. [164] and [216], will lead inefficient AQC algorithms. Therefore, it is important to consider different strategies which might allow an escape from bottlenecks or trap states which might limit the use of AQC [8].

When attempting to tackle combinatorial and optimization problems with classical computers, a common approach to cope with intractability and NP-complete problems [72, 178], is to employ heuristic algorithms as an alternative to exhaustive search which scales exponentially with system size. In quantum computation, Hogg [92, 94] introduced heuristic techniques in quantum algorithms. He showed that using information about the structure of the problem as a heuristic guide can be used to enhance the performance of quantum search compared to the scheme proposed by Grover [76]. Hogg's proposal was suggested for the gate model for quantum computation, but to our knowledge, there are no studies of heuristics in AQC. The purpose of this chapter is to examine the following questions: How can we incorporate heuristics in AQC? Is there any advantage by doing so? What are the modifications to CAQC proposal from an algorithmic and experimental point of view?

Initial state selection. The simplest form of heuristics we could think of is to start the quantum evolution from a quantum state which is a guess to the solution, but this possibility is not available in any of the proposals for AQC. Physical intuition as well as constraints within the problem can be used to make an educated guess. To illustrate our idea, let's use a lattice model for protein folding [148] as an example. For this model, an educated guess for the initial state would be to choose a bit string which encodes an initial position for the amino acids in the spatial lattice such that no two amino acids are on top of each other, and that they are connected according to the sequence defining the protein to be folded. Conversely, CAQC would begin the computation with a quantum superposition of all possible states of the computational basis. This choice of initial state contains absurd configurations like all amino acids on top of each other, or assignments referring to configurations of amino acids which are fully disconnected or not properly linked according to the protein sequence. We conjecture here that the presence of these "non-sensical" states might act as trap states, making a smooth transition towards the desired final ground state more difficult and therefore increasing the time required for an adiabatic evolution. In other fields of computer science, physics and chemistry, one might also use classical methods or a mean field approach to find approximate solutions to be employed as educated guesses. For example, in the context of quantum simulation, a Hartree-Fock solution may be used as the initial state for an adiabatic preparation of an exact molecular wave function [15, 202, 201, 200].

In quantum mechanics, preparing a desired state is not a trivial task [110, 4] and it is not always possible to deterministically prepare a state from a superposition state by measurement. For NP-complete classical problems, like the 3-satisfiability problem (3-SAT) [72, 178] studied here, the simplest guess for the initial state is that of choosing one of the possible assignments from the solution space. We present a strategy which allows to design experimentally-realizable initial Hamiltonians whose ground state corresponds to the desired guess as the initial quantum state. This is essential for studying the importance of choosing a guess state instead of having full superposition of states as in CAQC [69]. In an algorithm like the one proposed here, both the initial guess and the final solution will be states of the computational basis. Therefore, the overlap between them is zero, unless one has guessed the right solution or the guess is included in the subspace of solutions, which is very unlikely. Regardless of this counterintuitive choice, $\langle \psi_{guess} | \psi_{solution} \rangle = 0$, we show that using this kind of heuristics can be of advantage. Even in the case of choosing the initial state by random guessing there is potential for outperforming CAQC.

Restarting the evolution. Notice that the method described here not only allows to begin the evolution from a guess state, but also allows for the possibility of restarting a failed AQC calculation from the measured state. This state can be used as a refined guess to restart the adiabatic evolution. An adiabatic processes is an idealized concept because real experiments have to be run in finite time and therefore there will always be probability of measuring a non-desired excited state which does not encode the solution to the computational task. The possibility of restarting the quantum evolution using the measured state as a guess is a feature which is not available in any of the AQC proposals to date, to our knowledge.

The incorporation of heuristics in AQC essentially involves two modifications to CAQC. We address both changes; the first modification involves the design of initial Hamiltonians for arbitrary guess states. The second modifications involves the change of the time profile of the driving Hamiltonian from the linear ramp to a non-linear time dependence with a "sombrero-like" time profile (see Fig. 3.1). Note that this second change is **not** the main point of the chapter since non-linear paths had been proposed before [65]. Also, it is not our purpose to explore what is the optimal selection for the driving term. It must be emphasized that the "sombrero-like" time profile is an essential feature needed if one is interested in the kind of heuristics we describe here, but this is not the case in the conventional way of doing AQC where it was used for auxiliary Hamiltonian terms [65]. Because of this distinctive feature and with the purpose of differentiating our heuristic strategy proposed with CAQC, we will refer to our method as Sombrero Adiabatic Quantum Computation (SAQC). The name should be associated with the algorithmic strategy (selection of initial guess, design

of initial Hamiltonian and sombrero-like profile for the driving Hamiltonian) which aims to incorporate heuristics in AQC, not only to the use of non-linear paths in AQC.

The chapter is divided as follows: in Section 3.2, we review the CAQC approach. Section 3.3 introduces the basic elements of the new implementation, SAQC. Finally, in Section 3.4, we present numerical calculations comparing the performance of both the CAQC and the SAQC algorithms based on the minimum gap, g_{min}, of their respective time-dependent Hamiltonians driving their corresponding time evolutions.

3.2 Conventional adiabatic quantum computation (CAQC)

The goal of AQC algorithms is that of transforming an initial ground state $|\psi(0)\rangle$ into a final ground state $|\psi(\tau)\rangle$, which encodes the answer to the problem. This is achieved by evolving the corresponding physical system according to the Schrödinger equation with a time-dependent Hamiltonian $\hat{H}(t)$. The AQC algorithm relies on the quantum adiabatic theorem [131, 127, 194, 203, 215, 9, 125, 7, 96, 52, 37], which states that if the quantum evolution is initialized with the ground state of the initial Hamiltonian, the time propagation of this quantum state will remain very close to the instantaneous ground state $|\psi_g(t)\rangle$ for all $t \in [0, \tau]$, whenever $\hat{H}(t)$ varies slowly throughout the propagation time $t \in [0, \tau]$. This holds under the assumption that the ground state manifold does not cross the energy levels which lead to excited states of the final Hamiltonian. Here, we denote by ground state manifold the first m curves associated with the lowest eigenvalue of the time-dependent Hamiltonian for $t \in [0, \tau]$, where m is the degeneracy of the final Hamiltonian ground state. An example of $m = 2$, is shown in Fig. 5 of Ref. [148].

Conventionally the adiabatic evolution path is the linear sweep of $s \in [0, 1]$, where $s = t/\tau$:

$$H(s) = (1 - s)H_{transverse} + sH_f. \tag{3.3}$$

$\hat{H}_{transverse}$ (see Eq. 3.9 below) is usually chosen such that its ground state is a uniform superposition of all possible 2^n computational basis vectors, for the case of an $n-$qubit system. Here, we choose the spin states $\{|q_i = 0\rangle, |q = 1\rangle\}$, which are the eigenvectors of $\hat{\sigma}_i^z$ with eigenvalues +1 and -1, respectively, as the basis vectors. Then the initial ground state is $|\psi_g(0)\rangle = \frac{1}{\sqrt{2^n}} \sum_{q_i \in \{0,1\}} |q_n\rangle |q_{n-1}\rangle \cdots |q_2\rangle |q_1\rangle$. Such an initial ground state is usually assumed to be easy to prepare, for example, by imposing a global transverse field. Since each state encodes a possible solution, this initial state assigns equal probability to all possible solutions to the computational problem.

3.3 Sombrero adiabatic quantum computation (SAQC)

For SAQC, the time-dependent Hamiltonian can be written as:

$$\hat{H}_{sombrero} = (1-s)\hat{H}_i + \mathrm{hat(s)}\hat{H}_{driving} + s\hat{H}_f. \tag{3.4}$$

We want the non-degenerate ground state of the initial Hamiltonian $\hat{H}_i$ to encode a guess to the solution, and the driving term, $\hat{H}_{driving}$, to couple the states in the computational basis. The function hat(s) is zero at the beginning and end of the adiabatic path; therefore $\hat{H}_{driving}$ acts only in the range s $\in (0,1)$ in a "sombrero-like" time dependence (see Fig. 3.1), which allows $\hat{H}_i$ ($\hat{H}_f$) to be fully turned on at the beginning (end) of the computation.

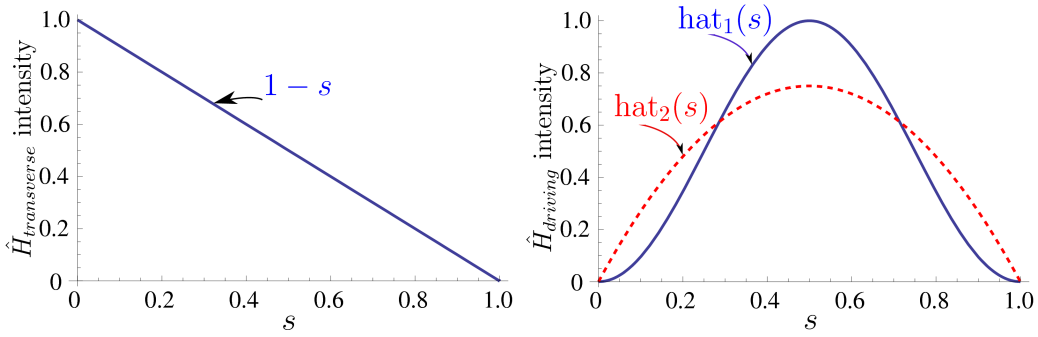

Figure 3.1: The main idea behind Sombrero Adiabatic Quantum Computation (SAQC) is to introduce heuristics in AQC, and having the possibility of restarting a failed AQC run from the measured excited state. In order to prepare an arbitrary state from any of the 2^N possible basis states from the computational basis of the N qubit system, we propose an initial Hamiltonian, $\hat{H}_i$ (see Eq. 3.5), in such a way that the desired initial guess state is the non-degenerate ground state of the designed initial Hamiltonian. The initial Hamiltonian is diagonal in the computational basis, and so is the final Hamiltonian for the case of classical problems such as the NP-complete problems, e.g., random 3-SAT. Since both, the initial and final Hamiltonians are diagonal, connecting them via a linear ramp as is usually done in CAQC (see left panel) will not lead the quantum evolution towards finding the ground state of the final Hamiltonian. To maintain the initial Hamiltonian uniquely and fully turned on at the beginning, $t = 0$, and the final Hamiltonian uniquely and fully turned on at the end of the computation, $t = \tau$, we introduce a driving Hamiltonian whose time profile intensity has a "sombrero-like" shape (see right panel) is such a way that it only acts during $0 < t < \tau$. Two examples of functions with this functional form are presented, where $\mathrm{hat}_1(s) = \sin^2(\pi s)$ and $\mathrm{hat}_2(s) = s(1-s)$. A desired feature of our algorithmic strategy is the possibility of introducing heuristics, and not that of introducing non-linear paths. The latter has been proposed previous publications [65, 13, 179], but here is employed as a consequence of the algorithmic strategy.

3.3.1 Design of the initial Hamiltonian for the guess state

As preparing an arbitrary initial non-degenerate ground state for adiabatic evolution is not a trivial task, we focus on easy to prepare initial guesses that consist of one of the states in the computational basis. The strategy proposed builds initial Hamiltonians such that the initial guess corresponds to the non-degenerate ground state of the initial Hamiltonian, as it is required by AQC. Additionally, this ground state would be non-degenerate.

Let us denote the states of the computational basis of an N qubit system as

$$|q_N\rangle\,|q_{N-1}\rangle\cdots|q_1\rangle \equiv |q_N\cdots q_1\rangle\,,$$

where $q_n \in \{0,1\}$. The proposed initial Hamiltonian, whose ground state corresponds to an arbitrary initial guess state of the form $|x_N\cdots x_1\rangle$, can be written as

$$\hat{H}_i = \sum_{n=1}^{N}\Big(x_n\hat{I} + \hat{q}_n(1-2x_n)\Big) = \sum_{n=1}^{N}\hat{h}_{x_n}, \tag{3.5}$$

where each x_n is a boolean variable, $x_n \in \{0,1\}$, while $\hat{q} \equiv \frac{1}{2}(\hat{I}-\hat{\sigma}^z)$ is a quantum operator acting on the n-th qubit of the multipartite Hilbert space $\mathcal{H}_N\otimes\mathcal{H}_{N-1}\otimes\cdots\otimes\mathcal{H}_n\otimes\cdots\otimes\mathcal{H}_1$. The operator $\hat{q}_n$ is given by

$$\hat{q}_n = \hat{I}_N\otimes\hat{I}_{N-1}\otimes\cdots\otimes(\hat{q})_n\otimes\cdots\otimes\hat{I}_1, \tag{3.6}$$

where $\hat{q}$ is placed in the nth position and the identity operators act on the rest of the Hilbert space.

The states constituting the computational basis, $|0\rangle$ and $|1\rangle$, are eigenvectors of $\hat{\sigma}_z$ with eigenvalues $+1$ and -1, and therefore they are also eigenstates of the operator $\hat{q}$ with eigenvalues 0 and 1 respectively. The logic behind the initial Hamiltonian in Eq. 3.5 then is clear: if $x_n = 0$, then $\hat{h}_{x_n=0} = \hat{q}_n$ but in the case of $x_n = 1$, then $\hat{h}_{x_n=1} = \hat{I} - \hat{q}_n$.

As an example, suppose one has a four qubit system, and one wishes to initialize the adiabatic computation with the state $|x_4 = 1, x_3 = 0, x_2 = 1, x_1 = 0\rangle \equiv |1010\rangle$ which one may choose either randomly or as an educated guess to the solution. According to Eq. 3.5, the initial Hamiltonian for the $|1010\rangle$ guess state should be constructed as

$$\hat{H}_i = \hat{h}_{x_4} + \hat{h}_{x_3} + \hat{h}_{x_2} + \hat{h}_{x_1} = (\hat{I}-\hat{q}_4) + \hat{q}_3 + (\hat{I}-\hat{q}_2) + \hat{q}_1, \tag{3.7}$$

and, clearly

$$\hat{H}_i\,|1010\rangle = \Big((\hat{I}-\hat{q}_4) + \hat{q}_3 + (\hat{I}-\hat{q}_2) + \hat{q}_1\Big)\,|1010\rangle = 0\,|1010\rangle\,. \tag{3.8}$$

In general, the 2^N states of the computational basis are all eigenstates of $\hat{H}_i$, and it can be easily verified that the spectrum of $\hat{H}_i$ are energies contained in $\{0,\cdots,N\}$. As

required, the ground state is also nondegenerate. The other states will have an eigenenergy which equals their Hamming distance to the ground state of the initial Hamiltonian.

3.3.2 Driving Hamiltonian

The encoding of an educated or a random guess into $\hat{H}_i$ (Eq. 3.5) makes both $\hat{H}_i$ and $\hat{H}_f$ (Eq. 3.4) diagonal in the computational basis. Therefore, connecting $\hat{H}_i$ and $\hat{H}_f$ with a linear ramp (as in Eq. 3.3), namely, omitting the operator $\hat{H}_{driving}$ in the quantum evolution, would yield zero probability of obtaining the state that encodes the unknown solution to the problem starting from the initial guess state. To avoid such a situation, $\hat{H}_{driving}$ must introduce non-diagonal terms in $\hat{H}_{sombrero}$ (see Eq. 3.4) that allows the initial state to transform from any arbitrary guess into the solution.

In order to make a fair comparison between CAQC and SAQC (see Eq. 3.3 and Eq. 3.4), we set

$$\hat{H}_{driving} = \hat{H}_{transverse} = \delta \sum_{n=1}^{N} \hat{q}_n^x, \tag{3.9}$$

in Eq. 3.4, where $\hat{q}_n^x$ stands for the quantum operator $\hat{q}^x$ acting on the nth qubit of the multipartite Hilbert space $\mathcal{H}_N \otimes \mathcal{H}_{N-1} \otimes \cdots \otimes \mathcal{H}_n \otimes \cdots \otimes \mathcal{H}_1$. The operator $\hat{q}_n^x$ is given by $\hat{I}_N \otimes \hat{I}_{N-1} \otimes \cdots \otimes (\hat{q}^x)_n \otimes \cdots \otimes \hat{I}_1$, where the operator $\hat{q}^x \equiv \frac{1}{2}(\hat{I} - \hat{\sigma}^x)$ has been placed in the nth position, and the $\hat{I}_i$'s are identity operators. From a physical point of view, the Hamiltonians $\hat{H}_{driving}$ and $\hat{H}_{transverse}$ can be related to a transverse magnetic field. The intensity of these Hamiltonians is tuned by varying the δ parameter. If we set δ to be the same for both adiabatic algorithms, all the dependence of the transverse field intensity lies on functions $(1 - s)$ in Eq. 3.3 and hat(s) in Eq. 3.4. A reasonable requirement for a fair comparison between CAQC and SAQC is that they both provide the same average intensity of the transverse magnetic field in $s \in [0, 1]$. A choice of hat(s) with the same average $\int_0^1 \text{hat}(s)ds = \int_0^1 (1 - s)ds = 1/2$, is hat$(s) = 3s(1 - s)$.

Even though nonlinear evolutions have been proposed in previous articles [65, 13, 179], our hat(s) function can be as simple or as complicated as desired, as long as hat(0) = hat$(1) = 0$ is fulfilled. There is plenty of room to optimize the performance by choosing a more convenient hat(s) for the adiabatic evolution, we emphasize that the additional advantage of SAQC is the possibility of choosing an initial guess. In the next section we present some results obtained based on one of the simple nonlinear function hat(s) = $3s(1-s)$ and discuss the performance of both CAQC and SAQC for random 3-SAT instances.

3.4 Hamiltonians for 3-SAT, numerical calculations and discussion

In order to provide a proof of concept for SAQC and to test the potential usefulness of both random and educated guesses in adiabatic evolution, we performed a numerical study on hard-to-satisfy 6- and 7-variable instances of the 3-SAT problem and compared our results with the CAQC approach. Let us now provide a succinct introduction to the 3-SAT problem as well as to briefly discuss its relevance in the fields of theoretical and applied computer science.

3.4.1 Construction of final Hamiltonians for satisfiability problems and design of numerical calculations

The K-SAT Problem. Let $A = \{e_1, e_2, \ldots, e_n, \bar{e}_1, \bar{e}_2, \ldots, \bar{e}_n\}$ be a set of Boolean variables $E = \{e_i\}$ and their negations $\bar{E} = \{\bar{e}_i\}$. Let us now construct a logical proposition P, defined as $P = \bigwedge_i [(\bigvee_{j=1}^{k} a_j)] = \bigwedge_i C_i$, where $a_j \in A$, i.e. P is a conjunction of clauses C_i over the set A, where each clause consists of the disjunction of k literals. Proposition P is a K-SAT instance and the solution of the K-SAT problem, for instance P, consists of finding a set of values for those binary variables upon which P has been built (i.e. a bitstring), so that replacement of such binary variables for their corresponding binary values makes $P = 1$, namely, proposition P is satisfied. 3-SAT is a particular case of K-SAT for K=3.

For example, let us examine the following instance of the 3-SAT problem. Let $E = \{x_1, x_2, x_3, x_4, x_5, x_6\}$ be a set of binary variables, and therefore the set of literals is $A = E \cup \bar{E} = \{x_1, x_2, \ldots, x_6, \bar{x}_1, \bar{x}_2, \ldots, \bar{x}_6\}$. Consider a 3-SAT instance specified by the proposition,

$$
\begin{aligned}
P \;=\; & (\bar{x_1} \vee \bar{x_4} \vee \bar{x_5}) \wedge (\bar{x_2} \vee \bar{x_3} \vee \bar{x_4}) \wedge (x_1 \vee x_2 \vee \bar{x_5}) \wedge (x_3 \vee x_4 \vee x_5) \wedge \\
& (x_4 \vee x_5 \vee \bar{x_6}) \wedge (\bar{x_1} \vee \bar{x_3} \vee \bar{x_5}) \wedge (x_1 \vee \bar{x_2} \vee \bar{x_5}) \wedge (x_2 \vee \bar{x_3} \vee \bar{x_6}) \wedge \\
& (\bar{x_1} \vee \bar{x_2} \vee \bar{x_6}) \wedge (x_3 \vee \bar{x_5} \vee \bar{x_6}) \wedge (\bar{x_1} \vee \bar{x_2} \vee \bar{x_4}) \wedge (x_2 \vee x_3 \vee \bar{x_4}) \wedge \\
& (x_2 \vee x_5 \vee \bar{x_6}) \wedge (x_2 \vee \bar{x_3} \vee \bar{x_5}) \wedge (\bar{x_2} \vee \bar{x_3} \vee \bar{x_4}) \wedge (x_2 \vee x_3 \vee x_6) \wedge \\
& (\bar{x_1} \vee \bar{x_2} \vee \bar{x_3}) \wedge (\bar{x_1} \vee \bar{x_4} \vee \bar{x_5}) \wedge (\bar{x_3} \vee \bar{x_4} \vee \bar{x_6}) \wedge (\bar{x_4} \vee \bar{x_5} \vee x_6) \wedge \\
& (\bar{x_2} \vee x_3 \vee \bar{x_6}) \wedge (x_2 \vee x_5 \vee x_6) \wedge (x_3 \vee x_5 \vee \bar{x_6}) \wedge (\bar{x_1} \vee x_3 \vee \bar{x_6}) \wedge \\
& (x_3 \vee \bar{x_5} \vee x_6) \wedge (x_4 \vee x_5 \vee x_6) \wedge (x_1 \vee x_2 \vee \bar{x_3})
\end{aligned}
$$

As this example suggests, finding solutions of even a modest 3-SAT instance can become difficult quite easily (in this case, P has only one solution: $x_1 = 1, x_2 = 1, x_3 = 0, x_4 = 1, x_5 = 0, x_6 = 0$.)

3-SAT is an NP-complete problem [72, 178], as opposed to 2-SAT which can be efficiently solved using a classical computer. Consequently, studying the properties of 3-SAT is

an important area of research, not only because a polynomial-time solution to 3-SAT would imply **P** = **NP**, but also because 3-SAT (due to its polynomial equivalence with K-SAT) may be used to model problems and procedures in theoretical computer science [2] as well as in several areas of applied computer science and engineering like artificial intelligence [73, 133].

For the purpose of simplifying our discussion, and without loss of generality, we randomly generated 3-SAT instances with a unique satisfying assignment (USA) and their number of clauses to number of variables ratio $\alpha \approx 4.26$. This value of α corresponds to the phase transition region where hard-to-satisfy instances are expected to be found [3, 132]. For completeness and to avoid any kind of bias in selecting this pool of instances, we selected 2^n different instances for every n variable case studied. More precisely, to exhaustively study the impact of different initial guesses with respect to unique solutions in the behavior of SAQC, we considered all 64 possible initial Hamiltonians $\hat{H}_i$ (using Eq. 3.5) for each one of the 64 randomly generated 6-variable USA instances. Similarly, we built 128 initial Hamiltonians for each 7-variable instance, one per possible initial guess (see Fig. 3.2). The instances were selected in such a way that their solutions had not only a USA, but also that there was no two instances with the same solution.

The generation of our USA 3-SAT instances took several steps, being the first one using the SAT instance generator developed by [137]. Unfortunately, as this generator does not warranty the production of 3-SAT instances with unique solutions, we took all generated 3-SAT instances and determined, by exhaustive bitstring substitution, whether such instances were USA or not. We iterated this process until we computed all 6-variable and 7-variable 3-SAT instances we needed.

The number of qubits used in our simulations is smaller than state-of-the-art calculations for AQC, such as those carried out by quantum Monte Carlo [208, 64]. Here we trade off carrying out few large calculations on many qubits for carrying out many calculations on fewer qubits. We wish to answer the question: What is the the impact of the initial guess in the spectral properties of the time-dependent Hamiltonian? To answer this, we explored the space of initial guesses in an exhaustive manner for the case of 6-variable and 7-variable SAT instances. We run a total of 81,920 and 327,680 for 6- and 7- variables respectively (see Fig. 3.2). As shown in the same figure, we also numerically explored the importance of the strength of the transverse external field for the performance of the algorithm.

Final Hamiltonians $\hat{H}_f$ are instance-dependent, i.e. the structure of each final Hamiltonian depends on the particular structure (conjunction of clauses) of each 3-SAT USA instance. Our final Hamiltonians $\hat{H}_f$ comply with the property that it must encode, in its ground state, the solution to the particular 3-SAT USA instance it was designed for [69, 199]. The design of the final Hamiltonian involves an intermediate step, where a classi-

cal cost or energy function is constructed for the particular instance of interest. Once this energy function is expressed in terms of binary variables, it can be easily transformed into a quantum Hamiltonian by performing the mapping indicated by Eq. 3.6, where each classical binary variable, q_n, is transformed into a quantum operator, $\hat{q}_n$. The energy function, H_f, associated with the final Hamiltonian, $\hat{H}_f$, can be constructed as a sum of other energy functions, h_{C_i} which involve only variables associated with one clause at a time,

$$H_f = \sum_i h_{C_i} \tag{3.10}$$

Each h_{C_i} is designed such that it is equal to 1 if clause C_i is unsatisfied and 0 if the clause is satisfied. Notice that the functions h_{C_i} contribute to the count of unsatisfied clauses which defines the spectrum of possible values for H_f, with $H_f = 0$ when all clauses are satisfied.

Formally, suppose $A = \{x_1, x_2, \ldots, x_n, \bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n\}$ is a set of n binary variables and their corresponding negations, P is a 3-SAT USA instance given by $P = \bigwedge_i C_i$, and each C_i is a disjuction of three elements of A, i.e. $C_i = a_\alpha \vee a_\beta \vee a_\gamma$ with $a_\alpha, a_\beta, a_\gamma \in A$ and indices α, β, γ are natural numbers, not necessarily consecutive. Finally, let $B = z_1 z_2 \ldots z_n$ be a set of n bits to be substituted in instance P. Then, h_{C_i} is given by

$$h_{C_i} = \begin{cases} 0, & \text{if substitution of } B = z_1 z_2 \ldots z_n \text{ in } a_\alpha \vee a_\beta \vee a_\gamma \text{ makes } C_i = 1 \\ 1, & \text{if substitution of } B = z_1 z_2 \ldots z_n \text{ in } a_\alpha \vee a_\beta \vee a_\gamma \text{ makes } C_i = 0. \end{cases}$$

To construct such a function for any arbitrary clause $C_i = a_\alpha \vee a_\beta \vee a_\gamma$, it is useful to note that the only assignment for which $C_i = 0$ is when $a_\alpha = 0$, $a_\beta = 0$, and $a_\gamma = 0$. Therefore, the function h_{C_i} by construction, should be 1 when $a_\alpha = 0$, $a_\beta = 0$, and $a_\gamma = 0$, and 0 otherwise. It can be easily checked that

$$h_{C_i} = (1 - a_\alpha)(1 - a_\beta)(1 - a_\gamma), \tag{3.11}$$

equals 1 when $a_\alpha = 0$, $a_\beta = 0$, and $a_\gamma = 0$ and 0 otherwise. Recall that each a_β represents a literal and therefore it could be representing the negation of a variable. One can always use the identity $\bar{x}_i = 1 - x_i$ to eliminate any $\bar{x}_i$, and obtain both h_{C_i} and H_f in terms of the x_i.

Consider for example the construction of the energy function h_C required for clause $C = \bar{x}_\alpha \vee \bar{x}_\beta \vee \bar{x}_\gamma$, i.e. C is a conjunction of three negated binary variables. In this case, C is satisfied by all possible 3-bit bitstrings except for 111 and, according to Eq. 3.11, the energy function assumes the form $h_{C_i} = (1 - \bar{x}_\alpha)(1 - \bar{x}_\beta)(1 - \bar{x}_\gamma) = x_\alpha x_\beta x_\gamma$.

As a last example consider a clause of the form $C = x_\alpha \vee x_\beta \vee x_\gamma$, where C is a conjunction of three non-negated binary variables taken from set A. It is clear that C will

be satisfied by all possible 3-bit bitstrings except for 000. According to Eq. 3.11, the energy function for this clause C is given by $h_C = (1 - x_\alpha)(1 - x_\beta)(1 - x_\gamma) = 1 - x_\alpha - x_\beta - x_\gamma + x_\alpha e_\beta + x_\beta x_\gamma + x_\alpha x_\gamma - x_\alpha x_\beta x_\gamma$. From the first equality of the previous equation, one can easily check that $h_C = 0$ for all possible 3-bit combinations except for $h_C(000) = 1$, as expected.

Once we have the final expression for the final classical energy function of Eq. 3.10, the final Hamiltonians $\hat{H}_f$ can be obtained using the mapping of Eq. 3.6, which relates the classical binary variables with quantum operators. Since we selected only USA instances for our study, each $\hat{H}_f$ has a non-degenerate ground state encoding the unique solution of one of our 3-SAT instances with corresponding ground eigenvalue equal to zero. The final Hamiltonians are the same for both strategies, CAQC and SAQC.

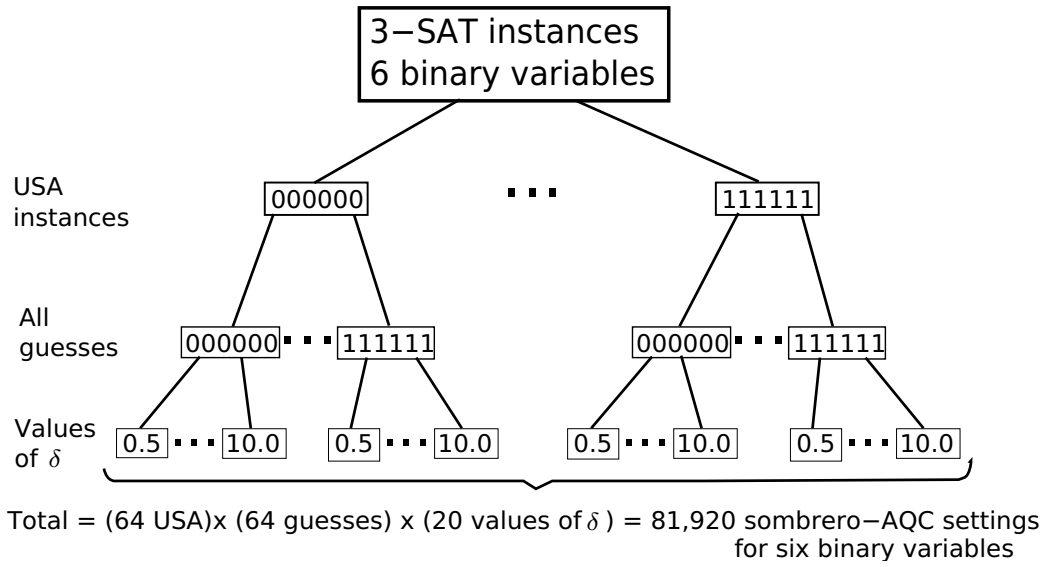

Figure 3.2: Scheme for 6 binary variables SAQC calculations. We generated 2^6 3-SAT unique satisfying assignment (USA) instances (first branching), each having as its only solution one of the 2^6 possible assignments. All 2^6 instances have a different state as solution, i.e. there is no chance for repeated instances. For each instance, we computed minimum-gap values associated with all possible settings of SAQC (Eq. 3.4) of all possible guesses (second branching), using 20 different values of $\delta \in \{0.5, 1.0, \ldots, 10.0\}$ (third branching). The same scheme was applied to 7 binary variable 3-SAT USA instances (not shown) for a total of (128USA) $\times$ (128guesses) $\times$ (20values of δ) $= 327,680$ SAQC settings.

The numerical results on the dependence of the minimum-gap value, g_{min}, as a function of δ are shown in Fig. 3.3. Curves were computed by taking the median of all bit strings that fulfilled the criteria specified in the legend boxes; namely either to produce i unsatisfied clauses (UC=i) when substituting the initial guess bit string in its corresponding instance,

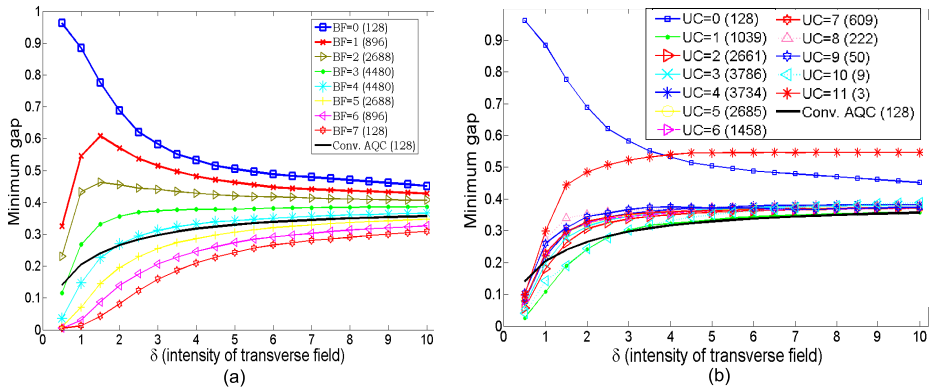

Figure 3.3: Summary of the 327,680 calculations for 7 variable 3-SAT instances of minimum-gap median values as a function of the transverse field intensity within groups sorted by (a) number of bit flips, BF, representing the Hamming distance between the initial guess and the solution and (b) number of unsatisfied clauses (UC). Plots include the CAQC (see Eq. 3.3) results for the same 128 different unique-satisfying assignment 3-SAT instances randomly generated for the case of 7 variables. Values in parentheses correspond to numbers of data points which contributed to the value of the median plotted in each curve. Results for 6 variable instances (not shown) are similar to the ones shown for 7 variable instances.

or to be j bit flips away from the solution (BF=j). BF represents the well known Hamming distance between the solution and the initial guess state. We focused on UC and BF because, in principle, the notion of closeness of an initial guess to the actual solution may be defined with either parameter. Data corresponding to a fixed value of δ is a statistical representation (median) of typical g_{min} values that would be expected for hard 3-SAT instances if the guess state belonged to a definite number of UCs or BFs under an experimental setup using SAQC. Such curves are compared with the minimum gap expected for CAQC.

3.4.2 Effects of the variation in the transverse field intensity on the minimum energy gap

The dependence of g_{min} values as a function of the transverse field intensity δ leaves open some important questions regarding the efficiency of adiabatic quantum algorithms, whether CAQC or SAQC. For example, what is the optimum value of δ which minimizes the running time of an adiabatic algorithm? How transferable is this optimum δ value among computational problems? Although we do not intend to do a thorough study of this question in this chapter, we would like to give some insight into this question and provide a qualitative discussion of what kind of results might be expected.

Following closely the notation from Farhi *et al* [69], consider $H(t) = \tilde{H}(t/\tau) = \tilde{H}(s)$,

with instantaneous values of $\tilde{H}(s)$ defined by

$$\tilde{H}(s)\,|E_l(s)\rangle = E_l(s)\,|E_l(s)\rangle \tag{3.12}$$

with

$$E_0(s) \leq E_1(s) \leq \cdots \leq E_{2^N-1}(s) \tag{3.13}$$

where 2^N is the dimension of the Hilbert space, and N the number of qubits or equivalently the number of binary variables in the SAT instance. According to the adiabatic theorem, if the gap between the two lowest levels, $E_1(s) - E_0(s)$, is greater than zero for all $0 \leq s \leq 1$, and taking,

$$\tau \gg \frac{\mathcal{E}}{g_{min}^2} \tag{3.14}$$

with the minimum gap, g_{min}, defined by,

$$g_{min} = \min_{0\leq s\leq 1}(E_1(s) - E_0(s)), \tag{3.15}$$

and $\mathcal{E}$ given by,

$$\mathcal{E} = \max_{0\leq s\leq 1}|\left\langle E_1(s)\right|\frac{d\tilde{H}}{ds}|E_0(s)\left.\right\rangle|, \tag{3.16}$$

then we can make the normed overlap

$$|\,\langle E_0(s=1)\,|\psi(\tau)\rangle\,| \tag{3.17}$$

arbitrarily close to 1. In other words, the existence of a nonzero gap guarantees that $|\psi(t)\rangle$ remains very close to the ground state of $H(t)$ for all $0 \leq t \leq \tau$, if τ is sufficiently large.

Even though we are aware of the new and more stringent conditions for adiabaticity [127, 194, 203, 215, 9, 125, 7, 96, 52, 37] and that Eq. 3.14 is just one of the inequalities to guarantee adiabatic evolution (though there is still lack of a sufficient and necessary condition according to Ref. [60]), we will base our discussion on Eq. 3.14 to illustrate that there is nothing anomalous in employing the additional Hamiltonian term in the full time-dependent Hamiltonian for SAQC. As well as in CAQC, the algorithmic complexity relies again in avoiding an exponentially narrowing of g_{min}. Along the way we find an important observation about the scaling of the running time as a function of the parameter δ.

Let us first determine an upper bound for $\mathcal{E}_{SAQC}$ in Eq. 3.14. Consider the Hamiltonian in Eq. 3.4 with hat$(s) = 3s(1-s)$ since this was the functional form used for our numerical calculations. We already discussed, at the end of Sec. 3.3.1, that the spectrum of H_i is contained in $\{0, 1, \cdots, N\}$ and, similarly, it can be easily shown that the spectrum of $\hat{H}_{driving}$ (see Eq. 3.9) is contained in $\{0, 1, \cdots, \delta N\}$. On the other hand, the spectrum of the final Hamiltonian, $\hat{H}_f$, is instance dependent, and its construction guarantees that

the maximum eigenvalue would be M which denotes the total number of clauses. This eigenvalue M would only appear in case we had an assignment which violates all of the clauses. Using these spectra upper bounds, we can establish an upper bound for $\mathcal{E}_{SAQC}$ in Eq. 3.14, i.e,

$$\begin{aligned}
\mathcal{E}_{SAQC} &= \max_{0\leq s\leq 1} |\left\langle E_1(s) \left| \frac{d\tilde{H}}{ds} | E_0(s) \right\rangle \right| \\
&= \max_{0\leq s\leq 1} |\left\langle E_1(s) \left| \hat{H}_f - \hat{H}_i + 3\delta \hat{H}_{driving} - 6\delta s \hat{H}_{driving} | E_0(s) \right\rangle \right| \\
&\leq \max \left(|\left\langle E_1(s) \left| \hat{H}_f | E_0(s) \right\rangle| + |\left\langle E_1(s) \left| \hat{H}_i | E_0(s) \right\rangle \right| \right. \\
&+ 3|\delta \left\langle E_1(s) \left| \hat{H}_{driving} | E_0(s) \right\rangle| + 6|\delta \left\langle E_1(s) \left| \hat{H}_{driving} | E_0(s) \right\rangle| \right) \\
&\leq M + N + 3|\delta|N + 6|\delta|N = N(\alpha + 1 + 9|\delta|)
\end{aligned} \tag{3.18}$$

Where we have used the triangle and Schwartz inequality and also the fact that $M = \alpha N$, with α close to 4.26 in this particular study. We can see that in the worst case scenario, $\mathcal{E}_{SAQC}$ scales linearly with the number of variables N, and linearly with the intensity of the magnetic field, $\mathcal{E}_{SAQC} = O(|\delta|N)$. A similar analysis gives also that $\mathcal{E}_{CAQC} = O(|\delta|N)$, and therefore, we showed that for SAQC, not surprisingly, the algorithmic complexity also relies on the scaling of g_{min}.

An interesting observation arise by analyzing the linear scaling of $\mathcal{E}$ with δ and using the numerical results for the dependence of the typical minimum gap values as a function of δ. There seem to be at least two distinguishable regimes for the dependence of g_{min} on δ for both CAQC and SAQC (Fig. 3.3). For relatively small values of $\delta \in [0.5, 1.5]$, g_{min} scales approximately linearly with δ and therefore $g_{min}^2 \sim \delta^2$. Since the running time is given by Eq. 3.14, and $\mathcal{E} \sim \delta$, the running time τ decays inversely proportional to δ within this linear regime. However, for large values of δ, in the 'stationary' regime where g_{min} is almost constant, increasing field intensity through δ would make both algorithms less efficient as running time τ would increase roughly linearly with δ.

Both CAQC and SAQC would benefit from an increase in the transverse field for small values of δ, but notice that g_{min} values for SAQC are more sensitive to δ, and soon become better on average than those for CAQC (Fig. 3.3). According to the previous discussion about running time as a function of δ, it would be ideal to choose δ near the end of the linear regime; in our calculations, δ somewhere in the interval (1,2). Further studies concerning the optimum value of δ as a function of the number of binary variables are needed, but we chose $\delta = 1.5$ for our analysis on the performance in SAQC and CAQC described in the following section.

3.4.3 Performance comparison between SAQC and CAQC

The data sorted with respect to BFs and UCs shows an increase of the minimum gap, g_{min}, as the Hamming distance from the initial guess to the solution decreases; the trend for UC is less apparent (see Fig. 3.3). Computing the number of UCs produced by a given initial guess can be done in polynomial time on a classical computer. Unfortunately there is no way to determine *a priori* how many bit flips the guess is from the solution, as that requires knowledge of the solution itself. Additionally, Fig. 3.4(a) shows that the SAQC implementation, using $\hat{H}_{driving}$ as defined in Eq. 3.9, does not necessarily favor states with low values of UC, but rather gives a homogenously distributed success probability between 25-45%, for $\delta = 1.5$. This is in accordance to the observation that solving 3-SAT hard instances is not necessarily guided by minimizing the number of UCs [101]. Given the above scenario, we analyzed the likelihood of better performance by choosing initial guesses at random.

In the following discussion, we use the term *significantly better* initial guess to mean an initial condition that leads a SAQC algorithm to be at least twice as fast as CAQC, i.e. running times for CAQC, τ_{CAQC}, and SAQC, τ_{SAQC}, are such that $\tau_{CAQC} \geq 2\,\tau_{SAQC}$ or, equivalently, $g_{min}^{SAQC} \geq \sqrt{2}\, g_{min}^{CAQC}$, assuming $\mathcal{E}_{CAQC} = \mathcal{E}_{SAQC}$.

For $\delta = 1.5$, choosing an initial state at random yields a probability greater than 50% of having $g_{min}^{SAQC} \geq g_{min}^{CAQC}$ (squares) as shown in Fig. 3.4(b). Moreover, the probability of *significantly better* performance, i.e. $\tau_{CAQC} \geq 2\,\tau_{SAQC}$ is $\approx 35\%$ (triangles). With the intention of predicting the performance of the SAQC protocol in the limit of large n, the third curve (circles) was produced using the following rationale: for USA instances, the number of bit configurations with a given value of $BF = m$ follows a binomial distribution $\binom{n}{m}$. In the limit of large n, the likelihood of choosing a state in the central region of the binomial distribution is the highest. This observation led us to concentrate on the performance of the most populated instance subsets, those that correspond to BF $= 3, 4$ for 7 variables. Here, the probability of significantly better performance is close to 40%.

Finally, we propose an algorithm based on SAQC. An initial guess is chosen either at random or by applying expert-domain knowledge and then encoded into the initial ground state of $\hat{H}_i$ (Eq. 3.5). An adiabatic passage based on SAQC is then performed either in serial or in parallel, depending on the availability of quantum hardware resources (see Fig. 3.5.). As an example of the potential usefulness of our algorithm, recall from Fig. 3.4 that the probability of *significantly better* performance using SAQC is 39% for $\delta = 1.5$.

One way to employ the probabilities we obtained from our numerical simulations in a more concrete scenario is: suppose one is assigned the task of using an adiabatic quantum computer and assume one uses Eq. 3.14 or any of the more stringent conditions for adiabaticity [127, 194, 203, 215, 9, 125, 7, 96, 52, 37] to estimate for how long one may need to

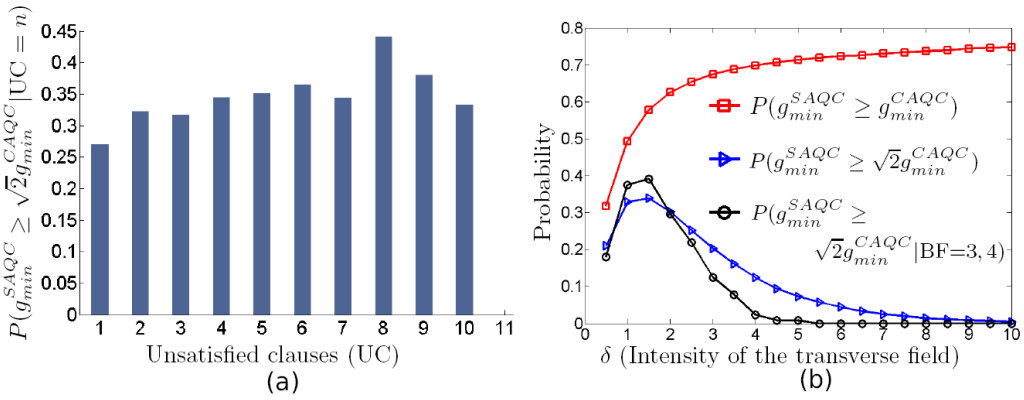

Figure 3.4: (a) $P(g_{min}^{SAQC} \geq \sqrt{2} g_{min}^{CAQC} | UC = n)$ is the conditional probability of choosing a state with UC= n and a SAQC minimum-gap large enough so that the performance of the SAQC is *significantly better* (at least twice faster) than the CAQC. The results were obtained for $\delta = 1.5$. Panel (a) shows that there is no correlation between the number of violated clauses and the g_{min} of the SAQC algorithm, for the hard-to-satisfy instances randomly chosen for this numerical study. Panel (b) shows the probability of choosing an initial state at random and satisfying the condition specified in the legend, for different values of the transverse field intensity, $0.5 \leq \delta \leq 10$. The conditional probability $P(g_{min}^{SAQC} \geq \sqrt{2} g_{min}^{CAQC} | BF = 3, 4)$ (triangles) aims to predict the performance of the SAQC algorithm in the case of large number n of qubits. In this limit, an initial state chosen at random will have with high probability a Hamming distance $BF \sim n/2$, given that they are binomially distributed, i.e., the number of n bit strings with BF $= m$ is equal to $\binom{n}{m}$

run an algorithm under the CAQC paradigm. Moreover, suppose that this estimated running time required to remain in the ground state with a high success probability is $\tau_{CAQC} =$ 2 days. Using the numerical results presented in this chapter, one can opt for performing the same task using SAQC as follows: suppose we have absolutely no information about the problem as it was in the case of this numerical study with random 3-SAT USA instances. Instead of running the CAQC algorithm for 2 days, pick a guess state at random from all the possible assignments and then use it as initial state for SAQC and run the algorithm with $\tau_{SAQC} = 1$ day. According to Fig. 3.4, the probability of having picked a state whose performance is as good as CAQC is 39%, for $\delta = 1.5$. If after measurement at the end of the first day the result is not a solution, we still can pick another state at random and let it run for one more day. By now the probability of having picked a state with the same performance as the CAQC in the two trials equals 63%. Note that in this simple probability calculations we are not taking into account the fact that even in the case where the state selected for the first run was not one of the 'ideal' ones (states corresponding to the 39 % of guesses for the results presented in Fig. 3.4), we still have a very good chance that the

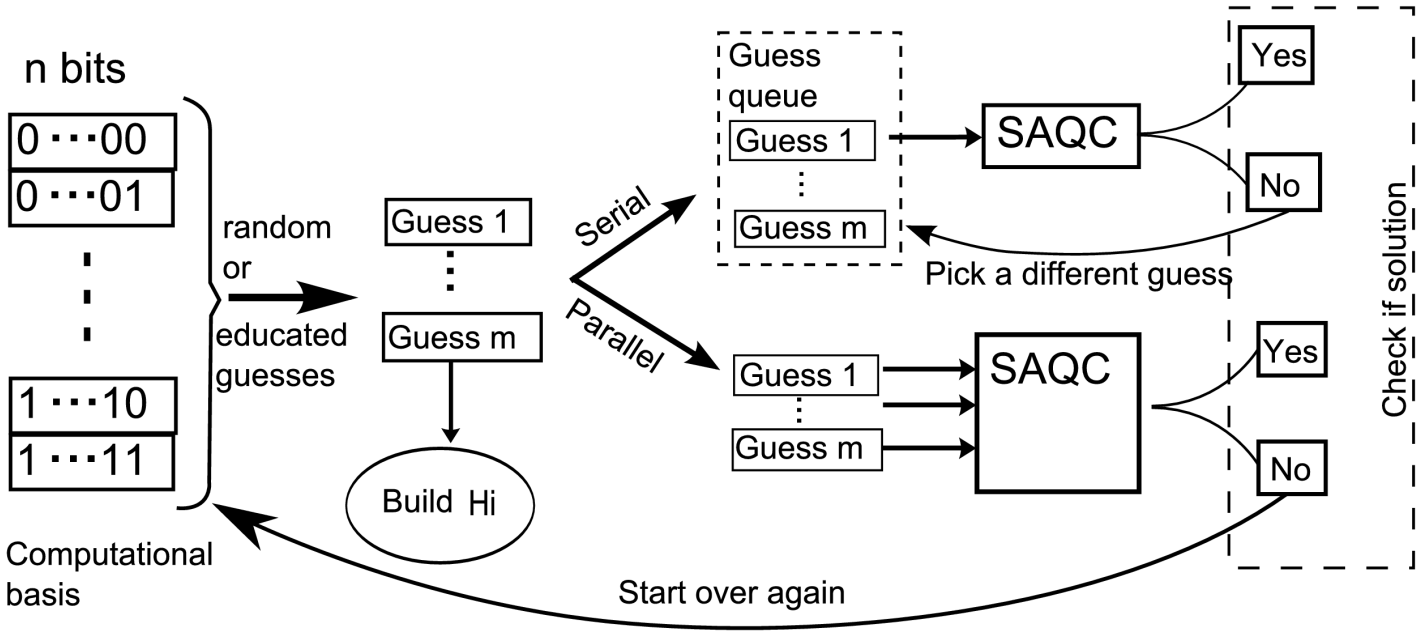

Figure 3.5: Implementation of an SAQC algorithm either in parallel or in serial. The algorithm begins by choosing a state from the computational basis. For each chosen initial state, an initial Hamiltonian is prepared according to Sec. 3.3.1. Next, choose an ideal time assuming a CAQC protocol will be run, and use that as a reference to run the SAQC protocol twice as fast. If only one AQC computer is available, one can still use the probabilistic speed up obtained in SAQC to run for example two adiabatic protocols instead of one, in serial mode. Once the first SAQC calculation is finished, one can efficiently check whether or not the result is a solution. In case that it is not a solution, one can submit an additional calculation, either randomly selecting another initial guess state or using the measured excited state. We call this latter option, "quantum heuristics" since the outcome of the near-adiabatic quantum evolution is used to refine the initial guess state for further experiments. In the case of having several adiabatic quantum computers at hand, one can do the same initial procedure of selecting guesses, but now submitting a different guess to a different node and running on each node twice as fast.

'non-ideal' state still delivers a right answer after the first measurement. This probabability will depend of course on how close is the chosen state from the set of 'ideal' ones.

Consequently, the execution of two SAQC algorithms in serial would take at most as much time as the execution of only one CAQC algorithm. By allowing us to choose two guesses to run in the same time as one case in CAQC, the probability of choosing a significantly better initial guess in these two SAQC executions increases from 39% to 63%. Furthermore, even when no significantly better initial guess is chosen and the process is not guaranteed to be fully adiabatic, there is still some probability that we measure the correct solution at the end of both executions.

3.5 Conclusions

In summary, we propose an algorithmic strategy which incorporates heuristics in adiabatic quantum computation. In particular, we study one of the most basic heuristic strategies consisting of initializing the computation with a desired initial state chosen by physical intuition, an educated guess, classical preprocessing of the problem, and/or by randomly choosing one of the possible assignments. This method allows to bridge powerful classical techniques such as heuristic optimization routines and/or mean-field calculations to obtain approximated solutions of the problem and use them as initial guess states for SAQC. The strategy presented allows for a parallel and/or a concatenated scheme. The parallel setup might be helpful if several adiabatic quantum platforms are available in which several guesses can be run simultaneously, one guess to run in each one of the adiabatic processing units. The idea of the concatenated scheme is to restart a failed adiabatic evolution with the measured excited state. Assuming a near-adiabatic trajectory, the measured excited state can be taken as a refined guess from which the quantum evolution is restarted. Notice this is not possible in the conventional approach which would restart the evolution from the full superposition, attributing equal probability to all the states, therefore "erasing" the information gained in the previous near-adiabatic evolution. Neither of the two features proposed above are possible in any of the different adiabatic quantum computation proposals to date. In addition, all the modifications proposed related to different adiabatic paths, auxiliary perturbations [66] can also be explored in the context of SAQC, which is also suitable to study quantum problems [103, 28, 144, 5, 136].

The numerical study performed in this chapter is a proof-of-principle to explore the importance and consequences of starting the adiabatic evolution with a guess state and to illustrate that incorporating this kind of heuristics in AQC is possible and might be advantageous. Our numerical simulations show that starting the adiabatic evolution with a guess state which has a zero-overlap with the solution, $\langle \phi_{guess} | \phi_{solution} \rangle = 0$, is not a

big concern. On the contrary, even when there is no hope to make an educated guess and selection of the initial state at random is the only available alternative, we obtain that approximately 40% of states might allow running the quantum algorithm at least twice as fast when compared to CAQC. This possibility of running the algorithm for shorter times but with several trials brings also additional advantages of getting the right answer in any intermediate measure. Moreover, these shorter runs would be less affected by decoherence effects.

Even though the procedure used for the performance comparison, CAQC *vs.* SAQC, is "reasonable" since it is based on the widely-used minimum gap criteria and its connection with the algorithm run-time (see Eq. 3.14), we are aware of the limitations of this analysis [127, 194, 203, 215, 9, 125, 7, 96, 52, 37]. We want to stress that the present numerical results are only encouraging indicators that heuristics in AQC might be a valuable algorithmic strategy for AQC, given the strong dependence of the value of the minimum gap as a function of the initial guess chosen. It is not our purpose to claim superiority of SAQC but to introduce the approach and the motivation behind it. For a more rigorous comparison of both schemes, CAQC and SAQC, we suggest numerical experiments which are not meant to be exhaustive but preferably involving larger size instances, and to explore different problems other than random 3-SAT. For example, we are interested in performing these studies for relevant instances of our recently developed AQC proposal for protein folding [148] and for adiabatic preparation of molecular ground states [15] where we expect the mean-field (Hartree-Fock) solution to be a better guess than a full superposition. Numerical propagation of the time-dependent Schrödinger equation, instead of inspection of the minimum gap after the Hamiltonian diagonalization, for these cases will provide a realistic simulation of the quantum computation.

Open questions to be explored further is the connections between SAQC, quantum phase transitions [172, 118], entanglement [145] and the effect of local minima [8]. The performance of the adiabatic algorithm in the limit of large n is still an open question [64, 209] which needs to be explored in the context of SAQC. It is not obvious that the same observations and conclusions/observations mentioned above for CAQC will hold for SAQC as well. For example, we think that SAQC might be a strategy to avoid the local minima traps described in Ref. [8]. From the study of the spectral properties of the random 3-SAT instances studied, we found that some initial states have a considerably larger gap and while others show a considerably smaller gap when compared with CAQC. Since the initial state in CAQC is a full superposition of all the possible states or solutions, including the ones with a large gap and others with small gap, we conjecture here that having a full superposition will not necessarily be the best choice, given that the presence of the states with small gaps could slow down the quantum evolution. An analysis beyond the gap criteria

would be needed to test this conjecture. Solving the time-dependent Schrödinger equation for the entire evolution is the most straightforward, yet numerically-challenging approach. In SAQC, the probability of obtaining these trap states can be avoided and even in the case of a failed evolution, the concatenated scheme may help to restart using a better guess. In contrast, in CAQC, the full superposition including states with considerably small gap might result in a bottleneck for the dynamics towards a successful computation. Further studies need to be done to verify this is indeed the case.

Chapter 4

Finding low-energy conformations of lattice protein models by quantum-annealing

Summary

Lattice protein folding models are a cornerstone of computational biophysics [196]. Although these models are a coarse grained representation, they provide useful insight into the energy landscape of natural proteins [146, 58, 134, 147, 111, 173]. Finding low-energy three-dimensional structures is an intractable problem [20, 48, 87] even in the simplest model, the Hydrophobic-Polar (HP) model. Exhaustive search of all possible global minima is limited to sequences in the tens of aminoacids [210]. Description of protein-like properties are more accurately described by generalized models, such as the one proposed by Miyazawa and Jernigan [135] (MJ), which explicitly take into account the unique interactions among all 20 amino acids. There is theoretical [6, 70, 99, 67, 167] and experimental [32] evidence of the advantage of solving classical optimization problems using quantum annealing [70, 99, 168, 53] over its classical analogue (simulated annealing [105]). In this chapter, we present a benchmark implementation of quantum annealing for a biophysical problem (six different experiments up to 81 superconducting quantum bits). Although the cases presented here can be solved in a classical computer, we present the first implementation of lattice protein folding on a quantum device under the Miyazawa-Jernigan model. This paves the way towards studying optimization problems in biophysics and statistical mechanics using quantum devices.

4.1 Introduction

The search for more efficient optimization algorithms is an important endeavor with prevalence on many disciplines ranging from the social sciences to the physical and natural sciences. Belonging to the latter, the protein folding problem [58, 134, 111, 147] consists

of finding the lowest free-energy configuration or, equivalently, the native structure of a protein given its amino acid sequence. Knowing how proteins fold elucidate their three-dimensional structure-function relationship which is crucial to the understanding of enzymes and for the treatment of misfolded-protein diseases such as Alzheimer's, Huntington's, and Parkinson's disease. Due to the high computational cost of modeling proteins in atomistic detail [26, 175], coarse-grained descriptions of the protein folding problem, such as those found in lattice models [111, 147, 134, 146], provide valuable insight about the folding mechanisms.

Harnessing quantum-mechanical effects to speed up the solving of classical optimization problems is at the heart of quantum annealing algorithms (QA) [70, 99, 168, 53]. In QA, quantum mechanical tunneling allows for more efficient exploration of difficult potential energy landscapes such as that of classical spin-glass problems. In our implementation of lattice folding, quantum fluctuations (tunneling) occurs between states representing different model protein conformations or folds.

The theoretical challenge is to efficiently map the hard computational problem of interest (e.g., lattice folding) to a classical spin-glass Hamiltonian: such mapping requiring a polynomial number of quantum bits (qubits) with the size of the problem (protein length) is described elsewhere [148]. Here we present a new mapping which, due to its exponential scaling with problem size, is not intended for large instances. The proposed mapping employs very few qubits for small problem instances, making it ideal for this first experimental demonstration and implementation on current quantum devices [97]. A combination of the existing polynomial mapping [148] and more advanced quantum devices would allow for the simulation of much larger instances of lattice folding and other related optimization problems.

Solving arbitrary problem instances requires a programmable quantum device to implement the corresponding classical Hamiltonian. We employ quantum annealing on the programmable device to obtain low-energy conformations of the protein model. We emphasize that nothing quantum mechanical is implied about the protein or its folding process; rather quantum fluctuations are a tool we use to solve the optimization problem.

The QA protocol performed here is also known as adiabatic quantum computation (AQC) [69, 67]. Of all the quantum-computational models, AQC is perhaps the most naturally suited for studying and solving optimization problems [67, 93]. For the experiments presented here, the small finite temperature of the superconducting device is enough to make the process less coherent than the original formulation of AQC, where the theoretical limit of zero temperature and quasi-adiabaticity are usually assumed [69, 67]. As we show in the discussion, numerical simulations including these unavoidable environmental effects accurately reproduce our experimental results.

Experimental implementations of QA or AQC are limited either by the number of qubits available in state-of-the-art quantum devices or by the programmability required to fulfill the problem specification. For example, the first realization of AQC was performed on a three-qubit NMR quantum device [182] and newer NMR implementations involve four qubit experiments [206]. Other experimental realizations of spin systems have been based on measuring bulk magnetization properties of the systems in which there is no control over the individual spins and the couplings among them [32, 204]. Quantum architectures using superconducting qubits [197, 207, 124, 91, 55, 141, 56, 100] offer promising device scalability while maintaining the ability to control individual qubits and the strength of their interaction couplings. During the preparation of this manuscript, an 84-qubit experimental determination of Ramsey numbers with quantum annealing was performed [24], underscoring the programmable capabilities of the device for problems with over 80 qubits. In this chapter, we present a quantum annealing experimental implementation of lattice protein models with general (Miyazawa-Jernigan [135]) interactions among the amino acids. Even though the cases presented here still can be solved on a classical computer by exact enumeration (the six-amino acid problem has only 40 possible configurations), it is remarkable that the device anneals to the ground state of a search space of 2^{81} possible computational outcomes. This study provides a proof-of-principle that optimization of biophysical problems such as protein folding can be studied using quantum mechanical devices.

4.2 Mapping of the lattice folding problem to the quantum hardware

The quantum hardware employed consists of 16 units of a recently characterized eight-qubit unit cell [84, 97]. Post-fabrication characterization determined that only 115 qubits out of the 128 qubit array can be reliably used for computation (see Fig. 4.1). The array of coupled superconducting flux qubits is, effectively, an artificial Ising spin system with programmable spin-spin couplings and transverse magnetic fields. It is designed to solve instances of the following (NP-hard [18]) classical optimization problem: Given a set of local longitudinal fields $\{h_i\}$ and an interaction matrix $\{J_{ij}\}$, find the assignment $\mathbf{s}^* = s_1^* s_2^* \cdots s_N^*$, that minimizes the objective function $E(\mathbf{s})$, where,

$$E(\mathbf{s}) = \sum_{1 \le i \le N} h_i s_i + \sum_{1 \le i < j \le N} J_{ij} s_i s_j, \tag{4.1}$$

$|h_i| \le 1$, $|J_{ij}| \le 1$, and $s_i \in \{+1, -1\}$.

Finding the optimal $\mathbf{s}^*$ is equivalent to finding the ground state of the corresponding

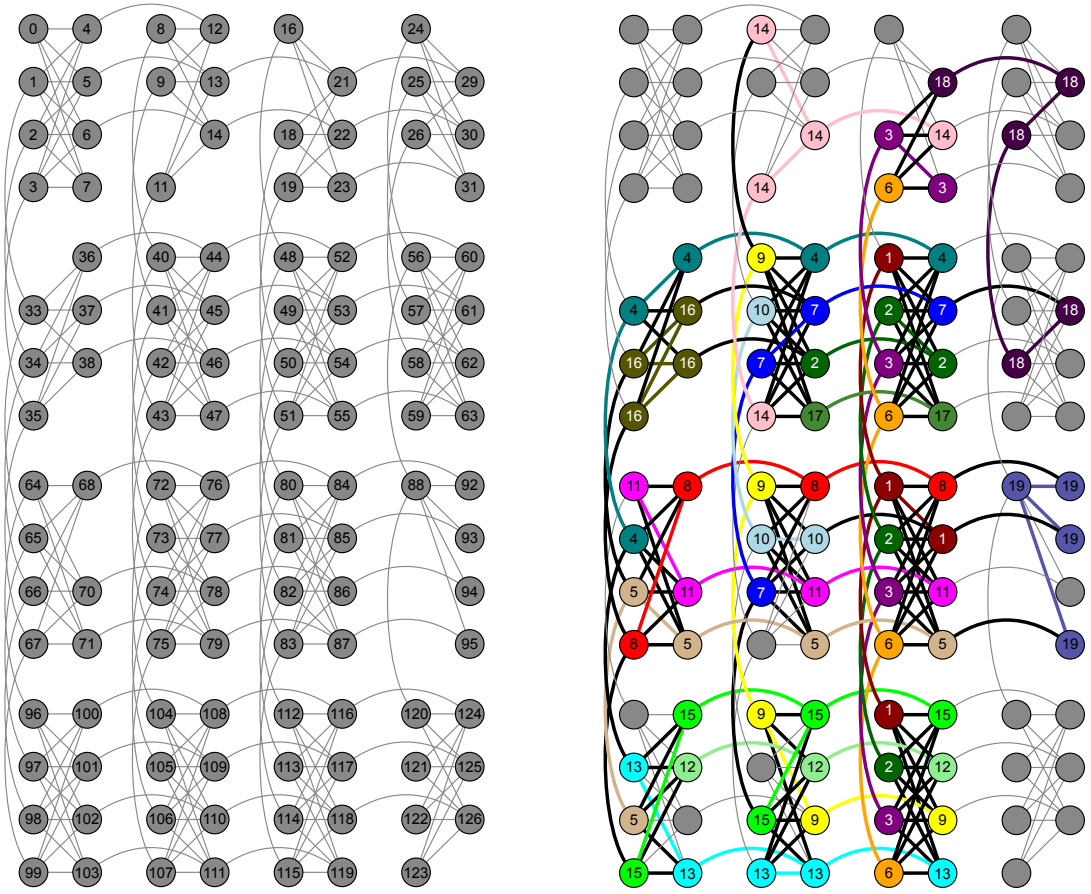

Figure 4.1: The array of superconducting quantum bits is arranged in 4×4 unit cells that consist of 8 quantum bits each. Within a unit cell, each of the 4 qubits in the left-hand partition (LHP) connects to all 4 qubits in the right-hand partition (RHP), and vice versa. A qubit in the LHP (RHP) also connects to the corresponding qubit in the LHP (RHP) of the units cells above and below (to the left and right of) it. (a) Qubits are labeled from 0 to 127 and edges between qubits represent couplers with programmable coupling strengths. Grey qubits indicate the 115 usable qubits, while vacancies indicate qubits under calibration which were not used. The larger experiments (Experiments 1,2, and 4) were performed on this chip, while the three remaining smaller experiments were run on other chips with the same architecture. (b) Embedding and qubit connectivity for Experiment 4, coloring the 81 qubits used in the experiment. Nodes with the same color represent the same logical qubit from the original 19-qubit Ising-like Hamiltonian resulting from the energy function associated with Experiment 4 (see Appendix 2 and Table 4.1). This embedding aims to fulfill the arbitrary connectivity of the Ising expression and allows for the coupling of qubits that are not directly coupled in hardware.

Ising classical Hamiltonian,

$$H_p = \sum_{1\leq i\leq N}^{N} h_i\sigma_i^z + \sum_{1\leq i<j\leq N}^{N} J_{ij}\sigma_i^z\sigma_j^z \tag{4.2}$$

where σ_i^z are Pauli matrices acting on the ith spin.

Experimentally, the time-dependent quantum Hamiltonian implemented in the superconducting-qubit array is given by,

$$H(\tau) = A(\tau)H_b + B(\tau)H_p, \qquad \tau = t/t_{run}, \tag{4.3}$$

with $H_b = -\sum_i \sigma_i^x$ responsible for quantum tunneling among the localized classical states, which correspond to the eigenstates of H_p (the computational basis). The time-dependent functions $A(\tau)$ and $B(\tau)$ are such that $A(0) \gg B(0)$ and $A(1) \ll B(1)$; in Fig. 4.2(b), we plot these functions as implemented in the experiment. t_{run} denotes the time elapsed between the preparation of the initial state and the measurement.

QA exploits the adiabatic theorem of quantum mechanics, which states that a quantum system initialized in the ground state of a time-dependent Hamiltonian remains in the instantaneous ground state, as long as it is driven sufficiently slowly. Since the ground state of H_p encodes the solution to the optimization problem, the idea behind QA is to adiabatically prepare this ground state by initializing the quantum system in the easy-to-prepare ground state of H_b, which corresponds to a superposition of all 2^N states of the computational basis. The system is driven slowly to the problem Hamiltonian, $H(\tau = 1) \approx H_p$. Deviations from the ground-state are expected due to deviations from adiabaticity, as well as thermal noise and imperfections in the implementation of the Hamiltonian.

The first challenge of the experimental implementation is to map the computational problem of interest into the binary quadratic expression (Eq. 4.2), which we outline next. In lattice folding, the sequence of amino acids defining the protein is viewed as a sequence of beads (amino acids) connected by strings (peptide bonds). This bead chain occupies points on a two- or three-dimensional lattice. A valid configuration is a self-avoiding walk on the lattice and its energy is calculated from the sum of interaction energies between nearest non-bonded neighbors on the lattice. By the thermodynamic hypothesis of protein folding [61], the global minimum of the free-energy function is conjectured to be the native functional conformation of the protein.

The hydrophobic-polar (HP) model is one of the simplest possible models for lattice folding [119]. In this model, the amino acids are classified into two groups, hydrophobic (H) and polar (P). To describe real protein energy landscapes a more elaborate description needs to be considered, such as the Mijazawa-Jernigan (MJ) model [135] which assigns the interaction energies for pairwise interactions among all twenty amino acids. The formulation

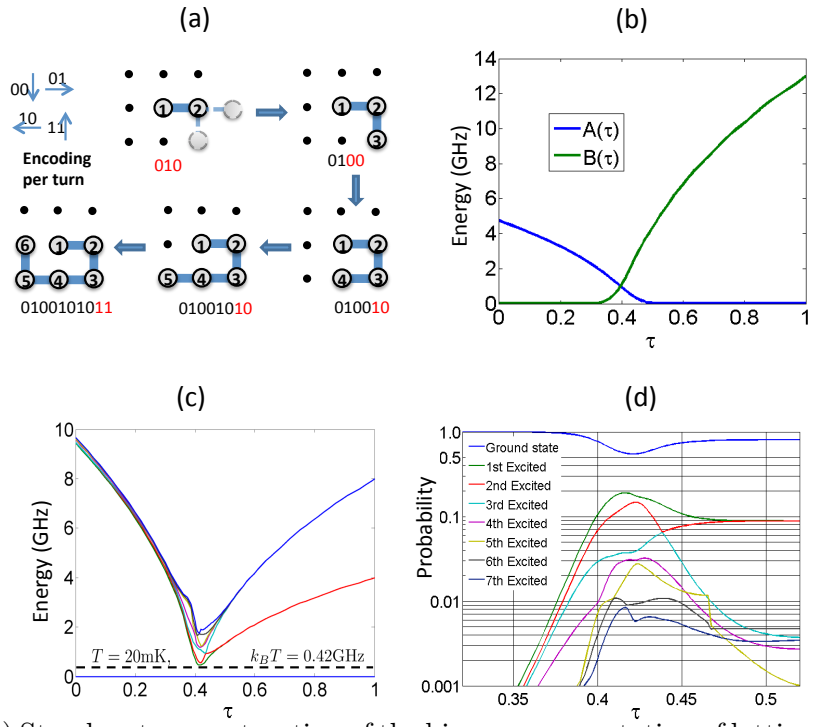

Figure 4.2: (a) Step-by-step construction of the binary representation of lattice protein. Two qubits per bond are needed and the bond directions are denoted as "00" (downwards), "01" (rightwards), "10" (leftwards), and "11" (upwards). The example shows one of the possible folds of an arbitrary six-amino-acid sequence. Any possible N-amino-acid fold can be represented by a string of variables $010q_1q_2q_3\cdots q_{\ell-1}q_\ell$ with $\ell = 2N-5$. (b)Time-dependence of the $A(\tau)$ and $B(\tau)$ functions, where $\tau = t/t_{run}$ with $t_{run} = 148\mu s$, (c) time-dependent spectrum obtained through numerical diagonalization, and (d) Bloch-Redfield simulations showing the time-dependent population in the first eight instantaneous eigenstates of the experimentally implemented 8-qubit Hamiltonian (Eq. 4.3) with H_p from Eq. 4.25. In panel (c), for each instantaneous eigenenergy curve we have subtracted the energy of the ground state, effectively plotting the gap of the seven-lowest-excited states with respect to the ground state (represented by the baseline at zero-energy). As a reference, we show the energy with the device temperature, which is comparable to the minimum gap between the ground and first excited state. In panel (d), populations are ordered in energy from top (ground state) to bottom. Although $\tau = t/t_{run}$ runs from 0 to 1, we show the region where most of the population changes occur. As expected, this is in the proximity of the minimum gap between the ground and first excited state around $\tau \sim 0.4$ [see panel(c)].

we used is general enough to take into account arbitrary interaction matrices for lattice models in two and three dimensions [149]. In particular, we solved a MJ model in 2D, the six amino-acid sequence of Proline-Serine-Valine-Lysine-Methionine-Alanine (PSVKMA in the one-letter amino-acid sequence notation). We solved the problem under two different experimental schemes (see Schemes 2 and 3 in Fig. 4.3), each requiring a different number of

resources. Solving the problem in one proposed experimental realization (Scheme 1) requires more resources than the number of qubits available (115 qubits) in the device. Scheme 2 and 3 are examples of the divide-and-conquer strategy, in which one partitions the problem in smaller instances and combines the independent set of results, thereby obtaining the same solution for the untractable problem. In the Appendix 1, we complement these four MJ related experiments with two small tetrapeptide instances (effectively HP model instances) for a total of six different problem Hamiltonians. We used the largest of these two instances (an 8 qubit experiment) for direct theoretical simulation of the annealing dynamics of the device. The results from our experiment and the theoretical model, which does not use any adjustable parameters (all are extracted experimentally from the device), are in excellent agreement (see panel (b), Fig. 4.5 in Sec. 4.5).

To represent each of the possible N-amino-acid configurations (folds) in the lattice, we encode the direction of each successive bond between amino acids; thus, for every N-bead sequence we need to specify $N-1$ turns corresponding to the number of bonds. For the case of a two dimensional lattice, a bond can take any of four possible directions; therefore, two bits per bond are required to uniquely determine a direction. More specifically, if a bond points upwards, we write "11". If it points downwards, leftwards or rightwards, we write "00", "10", or "01" respectively. Fixing the direction of the first bond reduces the description of any N-bead fold to $\ell = 2(N-2)$ binary variables, without loss of generality. As shown in Fig. 4.2(a), in the absence of external constraints other than those imposed by the primary amino acid sequence (see Appendix 1for an example with external constraints), we can fix the third binary variable to "0", forcing the third amino acid to go either straight or downward and reducing the number of needed variables to $\ell = 2N-5$. This constraint reduces the solution space by removing conformations which are degenerate due to rotational symmetry. Thus, a particular fold is uniquely defined by,

$$\boldsymbol{q} = \underbrace{01}_{turn1}\,\underbrace{0q_1}_{turn2}\,\underbrace{q_2q_3}_{turn3}\cdots\underbrace{q_{2N-6}q_{2N-5}}_{turn(N-1)} \tag{4.4}$$

An example of this encoding for a six-amino-acid sequence is represented in Fig. 4.2(a).

Using this mapping to translate between the amino acid chain in the lattice and the $2(N-1)$ string of bits, we constructed the energy function $E(\boldsymbol{q})$ in which $\boldsymbol{q}$ denotes the remaining $2N-5$ binary variables. Additionally, we penalized folds which exhibit two amino acids on top of each other, to favor self-avoiding walk configurations. The energy penalty chosen for each problem was sufficient to push the energy of invalid folds outside of the energy range of valid configurations (those with $E \leq 0$). Finally, we took into account the interaction energy among the different amino acids. A detailed construction of our energy function for the general case of N amino acids with arbitrary interactions is given elsewhere [149].

The experiment consists of the following steps: a) construction of the energy function to be minimized in terms of the turn encoding; b) reduction of the energy expression to a two-body Hamiltonian; and finally, c) embedding in the device. These last two steps need additional resources as explained below. We will focus on the simplest example (Experiment 3, Fig. 4.3) to show the procedure in detail. The embeddings for the other five experiments are provided in the Appendix 2. The energy function for Experiment 3, containing the contributions due to on-site penalties for overlapping amino acids, and pairwise interactions between amino acids is,

$$E(\boldsymbol{q}) \equiv E_{exp3}^{cubic} = -1 - 4q_3 + 9q_1q_3 + 9q_2q_3 - 16q_1q_2q_3 \tag{4.5}$$

where q_10 (q_2q_3) encodes the orientation of the fourth (fifth) bond (see Fig. 4.3). From Eq. 4.5 one can verify by substitution that the eight possible three-bit-variable assignments provide the desired energy landscape: the six conformations with $E \leq 0$ shown in blue in Fig. 4.3.

Eq. 4.5 describes the energy landscape of configurations but it is not quite ready for the device. Experimentally, we can specify up to two-body spin interactions, $\sigma_i^z\sigma_j^z$, and therefore, we need to convert this cubic energy function (Eq. 4.5) into a quadratic form resembling Eq. 4.1 (see Appendix 2). The resulting expression is

$$\begin{aligned} H_p^{\text{unembedded}} = (&7\sigma_1^z + 9\sigma_2^z + 8\sigma_3^z - 20\sigma_4^z + 9\sigma_1^z\sigma_3^z + 9\sigma_2^z\sigma_3^z \\ &- 16\sigma_1^z\sigma_4^z - 18\sigma_2^z\sigma_4^z - 18\sigma_3^z\sigma_4^z)/4 \end{aligned} \tag{4.6}$$

where the original binary variables and spin operators are related by $q_i \rightarrow (1 - \sigma_i^z)/2$. Experimental measurements of σ_i^z yield $s_i = +1$ ($s_i = -1$) corresponding to $q_i = 0$ ($q_i = 1$). Since $q_i = (1 - s_i)/2$, measurement of s_1, s_2, and s_3 allows us to reconstruct the bit string $q_10q_2q_3$ which encodes the desired fold.

One ancilla variable was added during the transformation of the three-variable cubic Hamiltonian into this quadratic four-variable expression. The meaning of the original variables s_1, s_2, and s_3 remains the same, allowing for the reconstruction of the folds. The energy of this four-variable expression will not change as long as the measurements of σ_1^z through σ_4^z result in values for $q_1q_2q_3q_4$ satisfying $q_4 = q_2 \wedge q_3$. This transformation ensures an energy penalty whenever this condition is violated.

The architecture of the chip lacks sufficient connectivity between the superconducting rings for a one-to-one assignment of variables to qubits (see Fig. 4.4). To satisfy the connectivity requirements of the four-variable energy function, the couplings of one of the most connected variables, q_4, were fulfilled by duplicating this variable inside the device such that $q_4 \rightarrow q_4$ and $q_{4'}$. In the form of Eq. 4.2 the final expression representing the energy

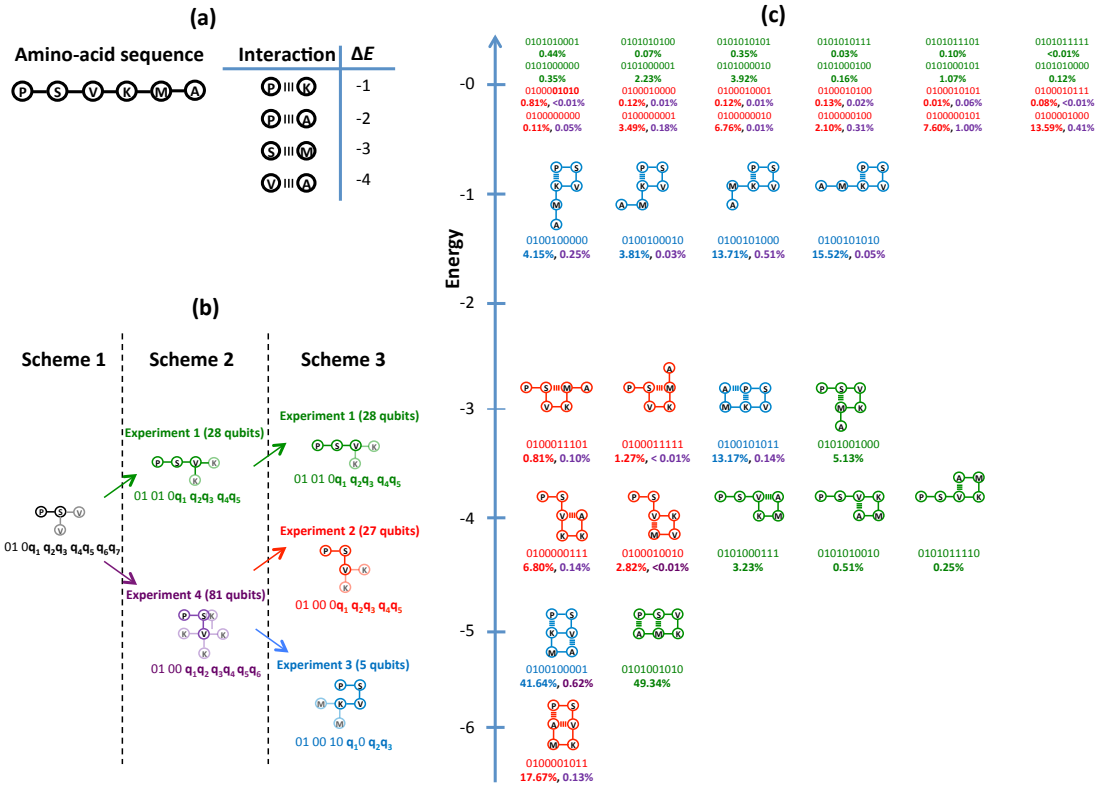

Figure 4.3: (a) Representation of the six-amino acid sequence, Proline-Serine-Valine-Lysine-Methionine-Alanine with its respective one-letter sequence notation, PSVKMA. We use the pairwise nearest-neighbor Miyazawa-Jernigan interaction energies reported in Table 3 of Ref. [135]. (b) Divide and conquer approach showing three different schemes which independently solve the six-amino acid sequence PSVKMA on a two-dimensional lattice. We solved the problem under Scheme2 and 3 (Experiments 1 through 4). (c) Energy landscape for the valid conformations of the PSVKMA sequence. Results of the experimentally-measured probability outcomes are given as color-coded percentages according to each of the experimental realizations described in panel (b). Percentages for states with energy greater than zero are 32.70%, 59.88%, 8.00%, and 95.97% for Experiments 1 through 4, respectively.

function of Experiment 3 is given by,

$$H_p = (7\sigma_2^z\sigma_1^z + 9\sigma_2^z + 8\sigma_3^z + 9\sigma_1^z\sigma_3^z + 9\sigma_2^z\sigma_{4'}^z - 2\sigma_{4'}^z - 16\sigma_1^z\sigma_{4'}^z - 18\sigma_2^z\sigma_{4'}^z - 18\sigma_{4''}^z - 18\sigma_3^z\sigma_{4''}^z - 36\sigma_{4''}^z\sigma_{4''}^z)/36 \tag{4.7}$$

This expression satisfies all requirements for the problem Hamiltonian (Eq. 4.3), the completion of which allows for the measurement of the energetic minimum conformation

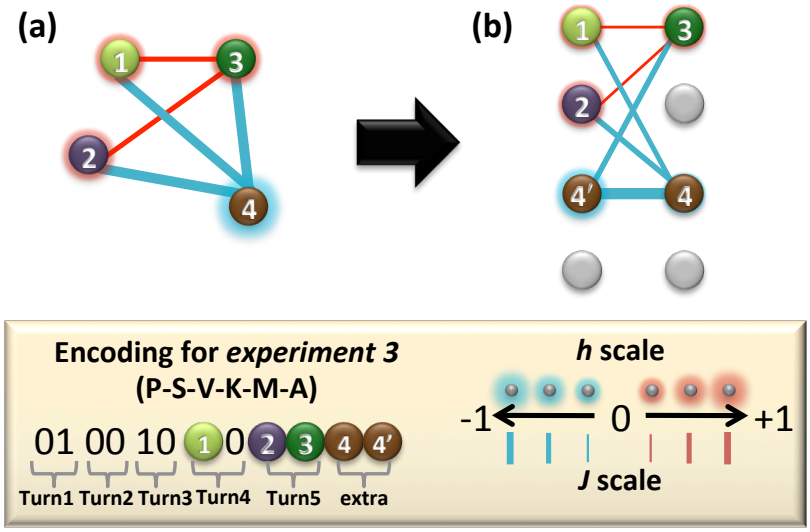

Figure 4.4: Graph representations of (a) the four-qubit unembedded energy function (Eq. 4.6) and (b) the five-qubit expression (Eq. 4.7) as was embedded into the quantum hardware. In graphs (a) and (b), each node denotes a qubit and the color and extent of its glow denotes the sign and strength of its corresponding longitudinal field, h_i. The edges represent the interaction couplings, J_{ij}, where color indicates sign and thickness indicates magnitude. Since we want the two qubits representing q_4 (q_4 and $q_{4'}$) to end up with the same value, we apply the maximum ferromagnetic coupling ($J = -1$) between them, which adds a penalty whenever this equality is violated. These maximum couplings are indicated in the figure by heavy lines. For the case of Experiment 3, the reconstruction of the binary bit stings representing the folds in Fig. 4.3, from the five-quibt experimental measurements can be recovered by $\boldsymbol{q}_{exp3} = 010010q_10q_2q_3|q_4q_{4'}$, with $q_i = 0$ ($q_i = 1$) whenever $s_i = 1$ ($s_i = -1$).

of this small peptide instance. The embedding of Eq. 4.7 into the hardware is shown in Fig. 4.4, where we label the five qubits used, q_1, q_2, q_3, q_4, and $q_{4'}$. Since we want the two qubits representing q_4 to end up with the same value, we apply the maximum ferromagnetic coupling ($J = -1$) between them, which adds a penalty whenever this equality is violated (last term in Eq. 4.7). These maximum couplings are indicated in Fig. 4.4 by heavy lines. The thinner lines show the remaining couplings used to realize the quadratic terms in Eq. 4.7, color coded according to the sign of the interaction and its thickness representing their strength. Note that every quadratic term in Eq. 4.7 has a corresponding coupler. Hereafter, we will denote the outcome of the five-qubit measurements as $\boldsymbol{q}_{expo} = 010010q_10q_2q_3|q_4q_{4'}$, with $q_i = 0$ ($q_i = 1$) whenever $s_i = 1$ ($s_i = -1$). Notice that only the bits preceding the divider character | contain physical information. These are the ones shown under each of the protein fold drawings associated with Experiment 3 (see Fig. 4.3).

4.3 Results and discussion

Similar embedding procedures to the one previously described were used for the larger experiments. For example, in Experiment 1, only 5 qubits define solutions of the computational problem. We needed 5 auxiliary qubits to transform the expression with 5-body interactions into an expression with only 2-body interactions. Embedding of this final expression required an additional of 18 qubits to satisfy the hardware connectivity requirements, for a total of 28 qubits. Table 4.1 in Appendix 2 summarizes the number of qubits required in each step through to the final experimental realizations.

Even though the quantum device follows a quantum annealing protocol, the odds of measuring the ground state are not necessarily high. For example, in the 81 qubit experiment, only 13 out of 10,000 measurements yielded the desired solution. We attribute these low-percentages to the analog nature of the device and to precision limitations in the real values of the local fields and couplings among the qubits in the experimental setup. When compared to other problem implementations, physical problems such as lattice folding lack the structure of the Ramsey number problem [24]. In the lattice folding problem implemented here, the parameters defining the problem instances are arbitrary and do not fall into certain integral distinct values as in the case of the Ramsey number experiment, making precision issues more pronounced in our implementation.

To gain insights into the dynamics and evolution of the quantum system, we numerically simulated the superconducting array with a Bloch-Redfield model of the 8-qubit experiment (see Appendix 4) which takes into account thermal fluctuations in the states due to the finite temperature (20mK) of the quantum device. For this 8-qubit experiment, the simulation predicted a ground state probability of 80.7 %, in excellent agreement with the experimentally observed value (80.3%). It is important to note that no adjustable parameters were used in our simulations to fit the data and all the parameters correspond to values measured directly from the quantum device. More details about the numerical simulations can be found in the Appendix 4.

As seen in Fig. 4.2(c), the temperature of the device is comparable with the minimum gap of the eight-qubit Hamiltonian. Therefore, we expect stronger excitation/relaxation near the gap closing, $\tau \approx 0.6$, due to exchange of energy with the environment, when compared to the other regimes of the annealing schedule where the gap is much larger than $k_B T$. In the absence of environment (a fully coherent process), our simulations indicate that that the success probability would be 100%, within numerical error. Fig. 4.2(d) shows that for the simulations at 20mK, the probability in the ground state goes down to $\sim$ 55%, but the same fluctuations make the system relax back to the ground state, yielding tan 80.27% success probability. This is due to the advantageous natural tendency of the

system to approach a thermal equilibrium which favors the ground state after crossing the minimum energy gap. As previously discussed in similar numerical simulations of quantum annealing algorithms [11], strong coupling to the bath and non-Markovianity would require going beyond the Bloch-Redfield model [10], but the agreement between experimental and simulated results support the validity of the quantum mechanical model used to describe the device. Previously reported temperature dependence predictions for the tunneling rate on the same qubits [97] and excellent agreement with the same level of theory used here reinforce the validity of our simulations for this 8-qubit instances.

4.4 Conclusion

We present the first quantum-mechanical implementation of lattice protein models using a programmable quantum device. We were able to encode and to solve the global minima solution for a small tetrapeptide and hexapeptide chain under several experimental schemes involving 5 and 8 qubits for the four-amino-acid sequence (Hydrophobic-Polar model) and 5, 27, 28, and 81 qubits experiments for the six amino-acid sequence under the Miyazawa-Jernigan model for general pairwise interactions. For the experiment with 8 qubits, we simulated the dynamics of the quantum device with a Redfield equation with no adjustable parameters, obtaining excellent agreement with experiment. Since the quantum annealing algorithm not only finds the ground state but also the low-lying excited states, it provides information about the relevant minimum energy compact structures of protein sequences [34] and it is useful to evaluate designability and stability such as that found in natural protein sequences, where the global minimum of free energy is well separated in energy from other misfolded states [14]. The approach employed here can be extended to treat other problems in biophysics and statistical mechanics such as molecular recognition, protein design, and sequence alignment [88].

4.5 Appendix 1. Transformation of the energy function of the lattice-folding model into the experimentally realizable spin-glass Hamiltonian

The energy function for the lattice model can be obtained as a sum of different contributions,

$$E_p(\boldsymbol{q}) = E_{onsite}(\boldsymbol{q}) + E_{pw}(\boldsymbol{q}) + E_{ext}(\boldsymbol{q}) \tag{4.8}$$

where $E_{onsite}(\boldsymbol{q})$ penalizes configurations with overlaps among any two amino acids, $E_{pw}(\boldsymbol{q})$ accounts for nearest-neighbor pairwise-interaction energies among non-bonded amino acids,

and $E_{ext}(\boldsymbol{q})$ refers to any external potentials other than the ones coming from interactions among the amino acids defining the protein. For amino acid sequences *in vacuo*, only E_{onsite} and E_{pw} are needed. The construction of these three-types of energy functions, in 2D and in 3D, for an arbitrary number of amino acids and interactions among them is explained in detail in Ref. [149]. Hereforth, we will only focus on the case of energy functions in 2D.

4.5.1 Case of the six-amino acid sequence PSVKMA (Experiments 1-4)

As explained in Sec. 4.2, the description of all possible 2D N-amino-acid fold in vacuo can be described by a bit string of length $2(N-1)$, with the first three bits held constant leaving $\ell = 2N-5$ binary variables as the computational variables of the problem,

$$\boldsymbol{q} = \underbrace{01}_{turn1}\,\underbrace{0q_1}_{turn2}\,\underbrace{q_2q_3}_{turn3}\cdots\underbrace{q_{2N-6}q_{2N-5}}_{turn(N-1)}. \tag{4.9}$$

For the case of $N=6$ (sequence PSVKMA), the problem is completely specified by the bit string

$$\boldsymbol{q}_{6AA} = \underbrace{01}_{turn1}\,\underbrace{0q_1}_{turn2}\,\underbrace{q_2q_3}_{turn3}\,\underbrace{q_4q_5}_{turn4}\,\underbrace{q_6q_7}_{turn5}. \tag{4.10}$$

By using the construction in Ref. [149], the 7-bit energy function describing the sequence PSVKMA (Scheme 1 in Fig. 4.3) is given by,

$$\begin{aligned}
E_{\text{PSVKMA}}(\boldsymbol{q}_{6AA}) = {} & -q_2 + 8q_1q_2 + 15q_2q_3 - 18q_1q_2q_3 - 3q_1q_4 + 12q_1q_2q_4 + 4q_3q_4 + 3q_1q_3q_4 \\
& - 6q_2q_3q_4 - 12q_1q_2q_3q_4 + 4q_2q_5 + 3q_1q_2q_5 - 15q_2q_3q_5 + 15q_4q_5 + 3q_1q_4q_5 \\
& - 6q_2q_4q_5 - 12q_1q_2q_4q_5 - 15q_3q_4q_5 + 28q_2q_3q_4q_5 - 2q_1q_2q_6 - 4q_3q_6 + 2q_2q_3q_6 \\
& + 13q_1q_2q_3q_6 - 2q_1q_4q_6 + 4q_1q_2q_4q_6 + 2q_3q_4q_6 + 13q_1q_3q_4q_6 + 4q_2q_3q_4q_6 \\
& - 37q_1q_2q_3q_4q_6 + 7q_5q_6 + 2q_2q_5q_6 + 13q_1q_2q_5q_6 + 4q_3q_5q_6 + 9q_2q_3q_5q_6 \\
& - 33q_1q_2q_3q_5q_6 - 20q_4q_5q_6 + 13q_1q_4q_5q_6 + 4q_2q_4q_5q_6 - 37q_1q_2q_4q_5q_6 + 9q_3q_4q_5q_6 \\
& - 33q_1q_3q_4q_5q_6 - 37q_2q_3q_4q_5q_6 + 99q_1q_2q_3q_4q_5q_6 - 4q_2q_7 + 4q_2q_3q_7 + 7q_4q_7 \\
& + 2q_2q_4q_7 + 13q_1q_2q_4q_7 + 4q_3q_4q_7 + 9q_2q_3q_4q_7 - 33q_1q_2q_3q_4q_7 + 4q_2q_5q_7 \\
& - 18q_4q_5q_7 + 9q_2q_4q_5q_7 - 33q_1q_2q_4q_5q_7 - 33q_2q_3q_4q_5q_7 + 62q_1q_2q_3q_4q_5q_7 + 7q_6q_7 \\
& + 2q_2q_6q_7 + 13q_1q_2q_6q_7 + 4q_3q_6q_7 + 9q_2q_3q_6q_7 - 33q_1q_2q_3q_6q_7 - 20q_4q_6q_7 \\
& + 13q_1q_4q_6q_7 + 4q_2q_4q_6q_7 - 37q_1q_2q_4q_6q_7 + 9q_3q_4q_6q_7 - 33q_1q_3q_4q_6q_7 - 37q_2q_3q_4q_6q_7 \\
& + 99q_1q_2q_3q_4q_6q_7 - 18q_5q_6q_7 + 9q_2q_5q_6q_7 - 33q_1q_2q_5q_6q_7 - 33q_2q_3q_5q_6q_7 \\
& + 62q_1q_2q_3q_5q_6q_7 + 53q_4q_5q_6q_7 - 33q_1q_4q_5q_7q_7 - 37q_2q_4q_6q_6q_7 + 99q_5q_2q_4q_5q_6q_7 \\
& - 33q_3q_1q_5q_6q_7 + 62q_1q_4q_4q_5q_6q_7 + 99q_2q_3q_4q_5q_6q_7 - 190q_1q_2q_3q_4q_5q_6q_7.
\end{aligned} \tag{4.11}$$

As shown in Fig. 4.3, expressions for each of the different experiments in Schemes 2 and 3 can be sequentially obtained by fixing the value of some of the variables in $E_{\mathrm{PSVKMA}}(\boldsymbol{q}_{6AA})$.

The energy function for Experiment 1 is obtained by evaluating $E_{\mathrm{PSVKMA}}(\boldsymbol{q}_{6AA})$ with $q_1 = 1$ (third amino-acid moves to the right) and $q_2 = 0$ (fourth amino-acid moves either down or right, exploiting upper/lower half-plane symmetry). After relabeling the five remaining variables so that their labels go from 1-5 instead of 3-7, i.e., $\boldsymbol{q}_{6AA} = 010q_1q_2q_3q_4q_5q_6q_7 \xrightarrow[\text{relabel}]{q_1=1,q_2=0} \boldsymbol{q}_{exp1} = 01010q_1q_2q_3q_4q_5$, the resulting expression describing the energy landscape for Experiment 1 is given by

$$\begin{aligned} E^{exp1}_{\mathrm{PSVKMA}}(\boldsymbol{q}_{exp1}) = &-3q_2 + 7q_1q_2 + 18q_2q_3 - 15q_1q_2q_3 - 4q_1q_4 - 2q_2q_4 + 15q_1q_2q_4 \\ &+ 7q_3q_4 + 4q_1q_3q_4 - 7q_2q_3q_4 - 24q_1q_2q_3q_4 + 7q_2q_5 + 4q_1q_2q_5 - 18q_2q_3q_5 \\ &+ 7q_4q_5 + 4q_1q_4q_5 - 7q_2q_4q_5 - 24q_1q_2q_4q_5 - 18q_3q_4q_5 + 20q_2q_3q_4q_5 \\ &+ 29q_1q_2q_3q_4q_5 \end{aligned} \tag{4.12}$$

The energy function for Experiment 4 is obtained by evaluating $E_{\mathrm{PSVKMA}}(\boldsymbol{q}_{6AA})$ with $q_1 = 0$ (third amino-acid moves down). After renaming the six remaining variables so that their labels span 1-6 instead of 2-7, i.e., $\boldsymbol{q}_{6AA} = 010q_1q_2q_3q_4q_5q_6q_7 \xrightarrow[\text{relabel}]{q_1=0} \boldsymbol{q}_{exp4} = 0100q_1q_2q_3q_4q_5q_6$, the resulting expression describing the energy landscape for Experiment 4 is given by

$$\begin{aligned} E^{exp4}_{\mathrm{PSVKMA}}(\boldsymbol{q}_{exp4}) = &-q_1 + 15q_1q_2 + 4q_2q_3 - 6q_1q_2q_3 + 4q_1q_4 - 15q_1q_2q_4 + 15q_3q_4 - 6q_1q_3q_4 \\ &- 15q_2q_3q_4 + 28q_1q_2q_3q_4 - 4q_2q_5 + 2q_1q_2q_5 + 2q_2q_3q_5 + 4q_1q_2q_3q_5 + 7q_4q_5 \\ &+ 7q_5q_6 + 2q_1q_4q_5 + 4q_2q_4q_5 + 9q_1q_2q_4q_5 - 20q_3q_4q_5 + 4q_1q_3q_4q_5 + 9q_2q_3q_4q_5 \\ &- 37q_1q_2q_3q_4q_5 - 4q_1q_6 + 4q_1q_2q_6 + 7q_3q_6 + 2q_1q_3q_6 + 4q_2q_3q_6 + 9q_1q_2q_3q_6 \\ &+ 4q_1q_4q_6 - 18q_3q_4q_6 + 9q_1q_3q_4q_6 - 33q_1q_2q_3q_4q_6 + 2q_1q_5q_6 + 4q_2q_5q_6 - 20q_3q_5q_6 \\ &+ 9q_1q_2q_5q_6 + 4q_1q_3q_5q_6 + 9q_2q_3q_5q_6 - 37q_1q_2q_3q_5q_6 - 18q_4q_5q_6 + 9q_1q_4q_5q_6 \\ &- 33q_1q_2q_4q_5q_6 + 53q_3q_4q_5q_6 - 37q_1q_3q_4q_5q_6 - 33q_2q_3q_4q_5q_6 + 99q_1q_2q_3q_4q_5q_6 \end{aligned} \tag{4.13}$$

The energy function for Experiment 2 is obtained by evaluating $E^{exp4}_{\mathrm{PSVKMA}}(\boldsymbol{q}_{exp4})$ with $q_1 = 0$ (fourth amino-acid moves either down or right). After renaming the five remaining variables so that their labels span 1-5 instead of 2-6, i.e., $\boldsymbol{q}_{exp4} = 0100q_1q_2q_3q_4q_5q_6 \xrightarrow[\text{relabel}]{q_1=0} \boldsymbol{q}_{exp2} = 01000q_1q_2q_3q_4q_5$, the resulting expression describing the energy landscape for Experiment 2 is given by

$$\begin{aligned} E^{exp2}_{\mathrm{PSVKMA}}(\boldsymbol{q}_{exp2}) = &\,4q_1q_2 + 15q_2q_3 - 15q_1q_2q_3 - 4q_1q_4 + 2q_1q_2q_4 + 7q_3q_4 + 4q_1q_3q_4 \\ &- 20q_2q_3q_4 + 9q_1q_2q_3q_4 + 7q_2q_5 + 4q_1q_2q_5 - 18q_2q_3q_5 + 7q_4q_5 + 4q_1q_4q_5 \\ &- 20q_2q_4q_5 + 9q_1q_2q_4q_5 - 18q_3q_4q_5 + 53q_2q_3q_4q_5 - 33q_1q_2q_3q_4q_5. \end{aligned} \tag{4.14}$$

Finally, the energy function for Experiment 3 is obtained by evaluating $E^{exp4}_{\mathrm{PSVKMA}}(\boldsymbol{q}_{exp4})$ with $q_1 = 1$, $q_2 = 0$(fourth amino-acid moves left) and $q_4 = 0$ (fifth amino-acid moves either down or left), exploiting the constrains imposed by the three fixed amino-acids (P,S, and V). After renaming the three remaining variables so that their labels are q_1, q_2 and q_3 instead of q_3, q_5, and q_7, i.e., $\boldsymbol{q}_{exp4} = 0100q_1q_2q_3q_4q_5q_6 \xrightarrow[\text{relabel}]{q_1=1,q_2=0,q_4=0} \boldsymbol{q}_{exp3} = 010010q_10q_2q_3$, the resulting expression describing the energy landscape for Experiment 3 is given by

$$E^{exp3}_{\mathrm{PSVKMA}}(\boldsymbol{q}_{exp3}) = -1 - 4q_3 + 9q_1q_3 + 9q_2q_3 - 16q_1q_2q_3 \tag{4.15}$$

4.5.2 Case of the four-amino acid sequence HPPH (Experiment 5)

Besides the six-amino acid sequence considered above, we also constructed the energy function for the simplest of all sequences within lattice protein models, the HPPH four-amino acid sequence within the HP model. For $N = 4$, we can specify any of its folds by the bit string $\boldsymbol{q}_{exp5} = 010q_1q_2q_3$. The three-bit energy function describing the energy landscape of Experiment 5 (see Fig. 4.5) is given by,

$$E_{HPPH}(\boldsymbol{q}_{exp5}) = -q_2 + 2q_1q_2 + 2q_2q_3 - 3q_1q_2q_3 \tag{4.16}$$

4.5.3 Case of the four-amino acid sequence HPPH under external constraints (Experiment 6)

A more realistic *in vivo* picture involves the presence of chaperone proteins assisting the folding dynamics towards the global minima. Chaperones, molecular docking, and molecular recognition are examples of problems which can be studied by adding external potentials, $E_{ext}(\boldsymbol{q})$, beyond the intrinsic interactions defined by the amino-acid chain, $E_{onsite}(\boldsymbol{q})$ and $E_{pw}(\boldsymbol{q})$ (see Eq. 4.8). The first consequence of adding an external potential $E_{ext}(\boldsymbol{q})$ (as the chaperone-like environment surrounding the small four-amino-acid sequence HPPH, illustrated in Fig. 4.5 by the pink-shaded area near the peptide) is that we can no longer exploit the symmetry of the solution space for upper and lower half plane conformations. Therefore, we cannot set the first variable of the turn associated with the third amino-acid to zero. Under external potentials, we specify arbitrary folds of the four-amino acid problem by $\boldsymbol{q}_{exp6} = 01q_1q_2q_3q_4$, where $q_1q_2(q_3q_4)$ encodes the orientation of the second (third) bond.

The external potential penalizes conformations in which either the third or fourth amino acid go into the chaperone region is:

$$\begin{aligned} E_{chap}(\boldsymbol{q}_{exp6}) = {} & \lambda^{d}_{ext}(1-q_1)(1-q_2) + \lambda^{r}_{ext}(1-q_1)q_2 + \lambda^{dr}_{ext}(1-q_1)(1-q_2)(1-q_3)q_4 \\ & + \lambda^{rd}_{ext}(1-q_1)q_2(1-q_3)(1-q_4) \end{aligned} \tag{4.17}$$

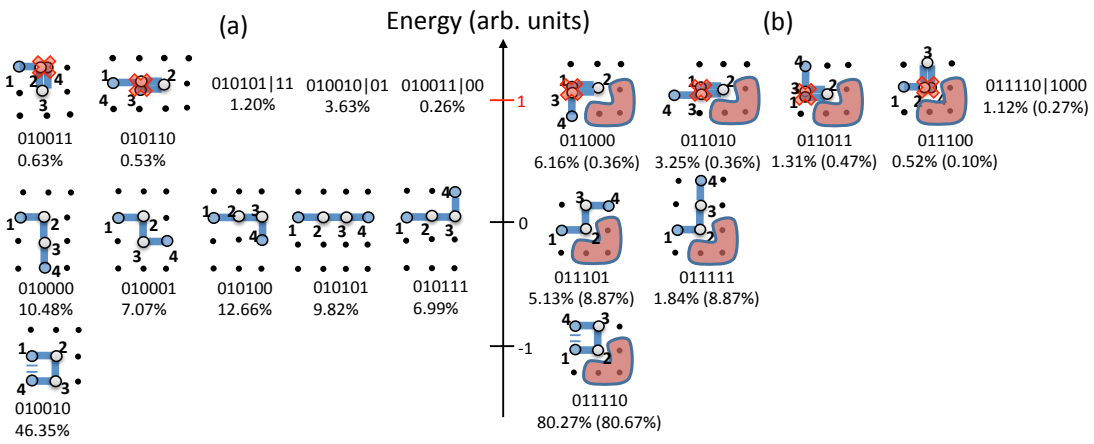

Figure 4.5: Energy landscape for the four amino-acid sequence HPPH, (a) *in vacuo* (Experiment 5), and (b) under the presence of a chaperone-like environment (Experiment 6) represented by the red-shaded region. In panel (a) [panel (b)], percentages indicate the experimentally measured probabilities of each state collected over 10,000 [28,672] runs of the quantum annealing algorithm described in Sec. 4.7.3. In the case of Experiment 6, numerical results from the Bloch-Redfield model discussed in Sec 4.8 are included in parenthesis. Although the variables involved in Experiment 5 (Experiment 6) are described by $\boldsymbol{q}_{exp5} = 010q_1q_2q_3|q_4q_{4'}$ ($\boldsymbol{q}_{exp6} = 01q_1q_2q_3q_4|q_5q_6q_{2'}q_{4'}$), under each fold we write only the physically-relevant variables which define the conformation. Since we show some experimental outcomes also for states with $E > 0$, then it is natural to find states which violate either the AND condition or the ferromagnetic condition; for these cases we explicitly write the auxilliary variables which went into the quantum hardware. For example, in Experiment 5 [panel (a)], the state $\boldsymbol{q}_{exp5} = 010101|11$ violates the AND condition since $q_4 \neq q_2q_3$. In the case of the state $\boldsymbol{q}_{exp5} = 010010|01$ the ferromagnetic condition for q_4 is violated since $q_4 \neq q_{4'}$. Each overlap of the amino acids with the chaperone raises energy by four units, whereas overlaps (red crossings) among amino acids in the chain raise energy by two units.

The penalty λ^d_{ext} raises energy only when the third amino-acid moves down ($q_1 = 0, q_2 = 0$), λ^r_{ext} raises energy only when the third amino-acid moves right ($q_1 = 0, q_2 = 1$), λ^{dr}_{ext} raises energy only when the third amino-acid moves down *and* the fourth-amino acid moves right ($q_1 = 0, q_2 = 0, q_3 = 0, q_4 = 1$), and the last penalty, λ^{rd}_{ext}, raises energy only when the third amino-acid moves down *and* the fourth-amino acid moves right ($q_1 = 0, q_2 = 1, q_3 = 0, q_4 = 0$). Each overlap of the amino acids with the chaperone increases energy by four units, i.e., $\lambda^d_{ext} = \lambda^r_{ext} = \lambda^{dr}_{ext} = \lambda^{rd}_{ext} = 4$.

When the third amino acid is also allowed to move upwards, the energy function for

the HPPH chain in vacuo is given by,

$$\begin{aligned} E_{HPPH}(\boldsymbol{q}_{exp6}) = q_1 - q_3 + q_1q_3 + 2q_2q_3 - 4q_1q_2q_3 + 2q_1q_4 \\ - 3q_1q_2q_4 + 2q_3q_4 - 4q_1q_3q_4 - 3q_2q_3q_4 + 7q_1q_2q_3q_4 \end{aligned} \tag{4.18}$$

After adding Eq. 4.17 and Eq. 4.18, the resulting energy function for the HPPH peptide in the presence of the "chaperone" environment illustrated in Fig. 4.5, is given by,

$$\begin{aligned} E_{HPPH,chap}(\boldsymbol{q}_{exp6}) = 4 - 3q_1 + 4q_2 - 4q_1q_2 - q_3 + q_1q_3 - 2q_2q_3 + 4q_4 - 2q_1q_4 \\ - 8q_2q_4 + 5q_1q_2q_4 - 2q_3q_4 + 5q_2q_3q_4 - q_1q_2q_3q_4, \end{aligned} \tag{4.19}$$

4.6 Appendix 2. Embbedding of problem instances into the quantum hardware

4.6.1 Reduction of high-order terms to a 2-body Ising-like Hamiltonian

Although the above energy expressions (Eqs. 4.11-4.19) describe the desired energy landscape, they are not suitable for experimental implementation. We need to reduce the degree of the high-order terms (cubic, cuartic, etc) to a quadratic expression (up to 2-body interactions). These high-order terms indicate many-body interactions which are not experimentally feasible within the current quantum device. To achieve this without altering the low-energy spectra ($E \leq 0$) where the target minima is supposed to be found, we use the technique described in Ref. [148, 149]. In Sec. 4.2, we presented the simplest case where only one reduction was required (expression for Experiment 3, Eq. 4.15). In the following we will focus on the next most complex case (Experiment 6, Eq. 4.19) which can be easily generalized to obtain any of the 2-body energy expressions for the larger experiments.

We introduce two ancilla binary variables, q_5 and q_6, and substitute Eq. 4.19 with products of the form q_1q_2 into q_5 and q_3q_4 into q_6. This substitution transforms the energy expression (Eq. 4.19) into a quadratic expression, e.g, the highest-order term which is quadratic, $q_1q_2q_3q_4$, is replaced by q_5q_6 which becomes quadratic, as desired. Under these substitutions, whenever we have six-variable assignments, $\boldsymbol{q}_{exp6} = 01q_1q_2q_3q_4|q_5q_6$, such that $q_6 = q_1 \wedge q_2 = q_1q_2$ and $q_6 = q_3 \wedge q_4 = q_3q_4$, we have the same energy spectrum as the one for the original quartic, four-variable expression. Since these two ancilla are new variables whose values are independent of the four original variables, we need to penalize six-variable assignments whenever $q_5 \neq q_1q_2$ and $q_6 \neq q_3q_4$. For every "collapse" of the form $q_iq_j \rightarrow r_k$, we add the penalty $E_\wedge(q_i, q_j, r_k; \delta_{ij}) = \delta_{ij}(3r_k + q_iq_j - 2q_ir_k - 2q_j\delta_{ij})$, where δ_{ij} is a positive number representing a penalty chosen (for more details see Ref. [149]) such that assignments violating this AND condition correspond to free-energies $E > 0$, outside the relevant search region ($E \leq 0$). The function $E_\wedge(q_i, q_j, r_k; \delta_{ij}) = 0$ only if $r_k = q_iq_j$ *and*

$E_\wedge(q_i, q_j, r_k; \delta_{ij}) > 0$ if $r_k \neq q_i q_j$. The six-variable expression resulting from the insertion of the new ancilla variables plus the penalty function becomes,

$$\begin{aligned} E^{2body}_{HPPH,chap}(\boldsymbol{q}) &= E_{HPPH,chap}(q_1, q_2, q_3, q_4; q_1q_2 \to q_5, q_3q_4 \to q_6) \\ &+ E_\wedge(q_1, q_2, q_5; \delta_{12}) + E_\wedge(q_3, q_4, q_6; \delta_{34}) \\ &= 4 - 3q_1 + 4q_2 + 6q_1q_2 - q_3 + q_1q_3 - 2q_2q_3 + 4q_4 - 2q_1q_4 - 8q_2q_4 + 4q_3q_4 \\ &+ 14q_5 - 12q_1q_5 - 12q_2q_5 + 5q_4q_5 + 10q_6 + 5q_2q_6 - 8q_3q_6 - 8q_4q_6 - q_5q_6 \end{aligned} \tag{4.20}$$

where according to the criteria in Ref. [149], we have chosen $\delta_{12} = 6$, and $\delta_{34} = 4$.

To rewrite this quadratic form in terms of the spin variables $\{s_i\}$, we apply the transformation $q_i \equiv \frac{1}{2}(1 - s_i)$ to each of the binary variables,

$$\begin{aligned} E^{Ising}_{HPPH,chap} &= 10 + \frac{13}{4}s_1 + \frac{3}{4}s_2 + \frac{7}{4}s_3 + \frac{1}{4}s_4 - 2s_5 - 2s_6 + \frac{3}{2}s_1s_2 + \frac{1}{4}s_1s_3 - \frac{1}{2}s_2s_3 \\ &- \frac{1}{2}s_1s_4 - 2s_2s_4 + s_3s_4 - 3s_1s_5 - 3s_2s_5 + \frac{5}{4}s_4s_5 + \frac{5}{4}s_2s_6 - 2s_3s_6 \\ &- 2s_4s_6 - \frac{1}{4}s_5s_6 \end{aligned} \tag{4.21}$$

After substracting the constant (independent term), we can fulfill the requirement that $|h_i| \leq 1$ and $|J_{ij}| \leq 1$ by scaling all coefficients of Eq. 4.21 down by the maximum absolute value of all coefficients. The renormalized quadratic expression is given by,

$$\begin{aligned} E^{unembedded}_{exp6}(\boldsymbol{s}) &= \frac{4}{13}(E^{ising}_{HPPH,chap} - 10) \\ &= (13s_1 + 3s_2 + 7s_3 + s_4 - 8s_5 - 8s_6 + 6s_1s_2 + s_1s_3 - 2s_2s_3 \\ &- 2s_1s_4 - 8s_2s_4 + 4s_3s_4 - 12s_1s_5 - 12s_2s_5 + 5s_4s_5 + 5s_2s_6 - 8s_3s_6 \\ &- 8s_4s_6 - s_5s_6)/13 \end{aligned} \tag{4.22}$$

The final Ising spin-glass Hamiltonian (before embedding into the quantum device) can be obtained by the substitution $s_i \to \sigma^z_i$.

$$\begin{aligned} H^{\text{unembedded}}_{exp6} &= (13\sigma^z_1 + 3\sigma^z_2 + 7\sigma^z_3 + \sigma^z_4 - 8\sigma^z_5 - 8\sigma^z_6 + 6\sigma^z_1\sigma^z_2 + \sigma^z_1\sigma^z_3 - 2\sigma^z_2\sigma^z_3 \\ &- 2\sigma^z_1\sigma^z_4 - 8\sigma^z_2\sigma^z_4 + 4\sigma^z_3\sigma^z_4 - 12\sigma^z_1\sigma^z_5 - 12\sigma^z_2\sigma^z_5 + 5\sigma^z_4\sigma^z_5 + 5\sigma^z_2\sigma^z_6 - 8\sigma^z_3\sigma^z_6 \\ &- 8\sigma^z_4\sigma^z_6 - \sigma^z_5\sigma^z_6)/13 \end{aligned} \tag{4.23}$$

4.6.2 Embedding into the quantum hardware

Eq. 4.23 does not fulfill the chip-connectivity requirements (see Fig. 4.6) for the primal graph representing Eq. 4.23. This limitation is fixed at the cost of adding two new qubits

serving as replicas of the two qubits which are linked by more than four connections. To enforce that the replicas of the i-th qubit ($\sigma_{i'}^z$) produce the same outcome as the original i-th qubit, we couple σ_i^z and $\sigma_{i'}^z$ with a strong ferromagnetic coupling, such that whenever the outcomes of the two variables are different they get penalized by a chosen penalty factor $\gamma_i > 0$. The function which performs this penalization for each replica i-th qubit is $H_{FM}(\{\sigma_i^z\};\gamma_i) = \gamma_i(1 - \sigma_i^z\sigma_{i'}^z)$. Notice that $H_{FM}(\{\sigma_i^z\};\gamma_i) = 0$, if $s_i = s_{i'} = \pm 1$, but $H_{FM}(\{\sigma_i^z\};\gamma_i) = 2\gamma_i$, if $s_i \neq s_{i'}$. For this study, a value of $\gamma_2 = \gamma_4 = 1$ suffices to leave assignments which violate this condition outside the region of interest with $E \leq 0$.

The redistribution of the connections among the original and primed qubits is given in the right panel of Fig. 4.6. The modified function taking into account the added ferromagnetic couplings is,

$$\begin{aligned}\tilde{H}_{exp6} &= H_{exp6}^{\text{unembedded}}(\sigma_2^z \to \{\sigma_2^z, \sigma_{2'}^z\}; \sigma_4^z \to \{\sigma_4^z, \sigma_{4'}^z\}) \\ &+ H_{FM}(\{\sigma_2^z\};\gamma_2 = 1) + H_{FM}(\{\sigma_4^z\};\gamma_4 = 1) \\ &= (13\sigma_1^z + 3\sigma_2^z + 7\sigma_3^z + \sigma_4^z - 8\sigma_5^z - 8\sigma_6^z + 6\sigma_1^z\sigma_{2'}^z + \sigma_1^z\sigma_3^z - 2\sigma_2^z\sigma_3^z \\ &- 2\sigma_1^z\sigma_{4'}^z - 8\sigma_{2'}^z\sigma_4^z + 4\sigma_3^z\sigma_4^z - 12\sigma_1^z\sigma_5^z - 12\sigma_2^z\sigma_5^z + 5\sigma_4^z\sigma_5^z + 5\sigma_{2'}^z\sigma_6^z - 8\sigma_3^z\sigma_6^z \\ &- 8\sigma_{4'}^z\sigma_6^z - \sigma_5^z\sigma_6^z)/13 + (1 - \sigma_2^z\sigma_{2'}^z) + (1 - \sigma_4^z\sigma_{4'}^z)\end{aligned} \tag{4.24}$$

Again, we subtract the independent constant terms from the insertion of the H_{FM} functions. The final expression, which is implementable in the quantum device is,

$$\begin{aligned}H_{exp6} = \tilde{H}_{exp6} - 2 &= (13\sigma_1^z + 3\sigma_2^z + 7\sigma_3^z + \sigma_4^z - 8\sigma_5^z - 8\sigma_6^z + 6\sigma_1^z\sigma_{2'}^z + \sigma_1^z\sigma_3^z - 2\sigma_2^z\sigma_3^z \\ &- 2\sigma_1^z\sigma_{4'}^z - 8\sigma_{2'}^z\sigma_4^z + 4\sigma_3^z\sigma_4^z - 12\sigma_1^z\sigma_5^z - 12\sigma_2^z\sigma_5^z + 5\sigma_4^z\sigma_5^z + 5\sigma_{2'}^z\sigma_6^z - 8\sigma_3^z\sigma_6^z \\ &- 8\sigma_{4'}^z\sigma_6^z - \sigma_5^z\sigma_6^z - 13\sigma_2^z\sigma_{2'}^z - 13\sigma_4^z\sigma_{4'}^z)/13\end{aligned} \tag{4.25}$$

The embeddings for Experiments 3 and 4 are shown in Fig. 4.4 and Fig. 4.1, respectively. The embeddings corresponding to Experiment 5 and 6 are represented in Fig. 4.6, while the embedding for the medium size problem instances (Experiments 1 and 2) are represented in Fig. 4.7.

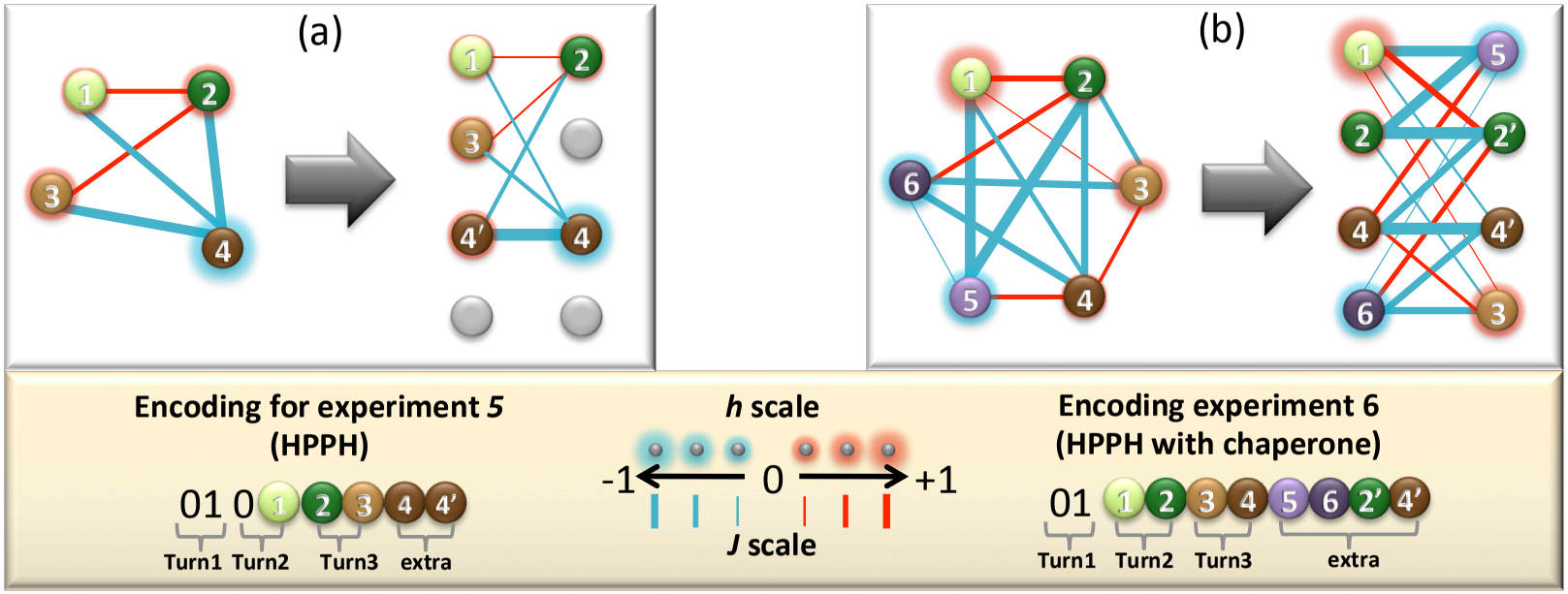

Figure 4.6: In the graphs presented in (a) and (b), each node denotes a qubit, and the color and extent of its glow denotes the sign and strength of its corresponding longitudinal field, h_i. The edges represent the interaction couplings, J_{ij}, where color indicates sign and thickness indicates magnitude. The maximum couplings are indicated in the figure by heavy lines. (a) Primal graph (left) and the embedded representation of the expression implemented in the quantum hardware for Experiment 5 (HPPH *in vacuo*). (b) Primal graph (left) and the embedded eight-qubit expression (Eq. 4.25) for Experiment 6 (HPPH in the chaperone-lke environment).

Table 4.1: Number of qubits needed for each one of the six experiments described in Fig. 4.3 and Fig. 4.5. The most compact version of the energy function corresponds to the second column. Each one of the steps, reduction of high-order terms in the energy function towards a 2-body Ising-like Hamiltonian and embedding of this Ising expression to fulfill the physical connectivity of the qubits in the device, requires more auxiliary qubits. The final column reports the number of qubits in the experimentally implemented expression of the energy function.

	Number of qubits needed		
Experiment #	energy function	Ising Hamiltonian	hardware-embedded expression
1	5	10	28
2	5	10	27
3	3	4	5
4	6	19	81
5	3	4	5
6	4	6	8

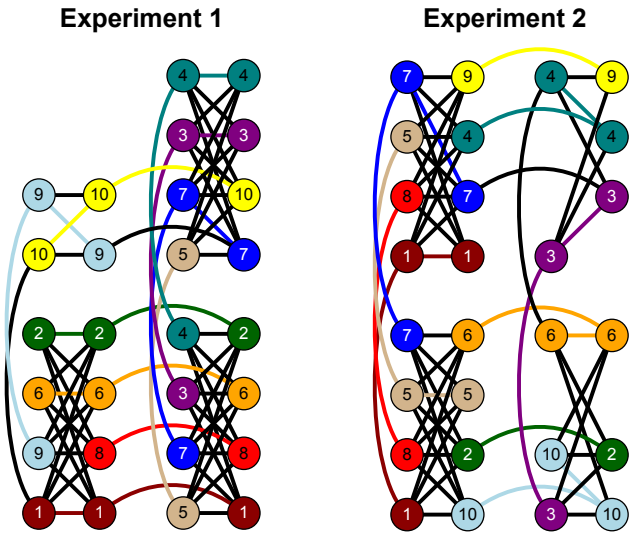

Figure 4.7: Embedding of Experiments 1 and 2 into the quantum hardware. The 28 qubits (27 qubits) from Experiment 1 (Experiment 2) have been relabeled to show the qubits which were strongly ferromagnetically coupled representing the same variable and biased to have the same experimental outcomes. Both problem instances resulted in ten-qubit spin-glass Hamiltonians after reducing their energy expressions to the Ising-like 2-body interaction expression. The additional qubits are part of the embedding procedure used to fulfill the arbitrary connectivity of the Ising expression, allowing for couplings of qubits that are not directly coupled in hardware.

4.7 Appendix 3. Experimental details

4.7.1 The processor chip

All experiments discussed herein were conducted on a sample fabricated in a four Nb layer superconducting integrated circuit process employing a standard Nb/AlOx/Nb trilayer, a TiPt resistor layer, and planarized SiO_2 dielectric layers deposited with a plasma-enhanced chemical vapour deposition process. Design rules included 0.25 μm lines and spaces for wiring layers and a minimum junction diameter of 0.6 μm. Experiments were conducted in an Oxford Instruments Triton 400 Cryofree DR at a temperature of 20 mK.

The sample processor chip contains a coupled array of 128 qubits of a design discussed in Ref. [82]. Each qubit is an rf-SQUID flux qubit with a double-well potential, as depicted in Fig. 4.8. They are magnetically coupled with sign and magnitude tunable couplers in a manner described in Ref. [98]. The array is built up of 16 eight-qubit unit cells. For example, Experiment 6 was conducted using a single unit cell (highlighted in Fig. 4.9**a**). The connectivity of qubits within the unit cell is shown schematically in Fig. 4.9**b**.

Three different chips available with this same architecture were used to run the different problem instances (Experiments 1-6, Fig. 4.3 and 4.5). Experiments 1, 2, and 4 were run in one chip, while Experiment 3 and 5 used a different chip. Experiment 6 used the same chip and unit cell used in Ref. [97]. Since all the chips have the same architecture and design but different calibration parameters, we will focus on the chip used to run Experiment 6, and report all the parameters used to run the numerical simulation reported in Sec. 4.8.

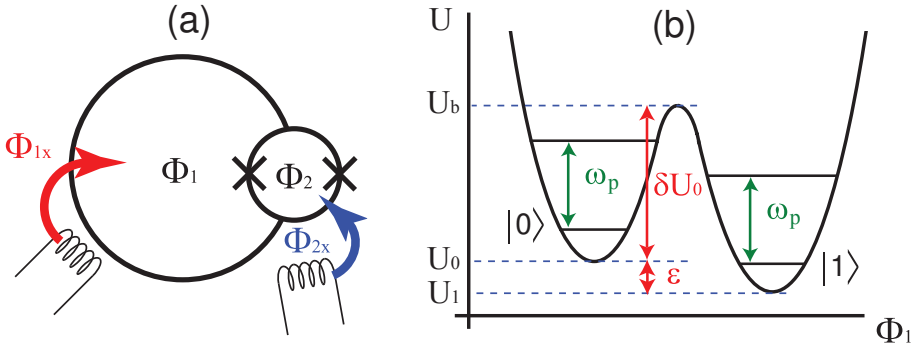

Figure 4.8: **(a)** Illustration of a single rf-SQUID flux qubit. Φ_{1x} is the flux bias applied to the major (qubit) loop, and Φ_{2x} is the flux bias applied to the minor (CJJ) loop. **(b)** Cross-section of the double-well potential of an rf-SQUID flux qubit, with 4 localized energy levels marked. Φ_{1x} primarily affects the qubit bias ϵ, whereas Φ_{2x} affects both the barrier height δU_0 and ϵ.

4.7.2 Magnetic Environment

The magnetic field in the sample space was controlled with three concentric high permeability shields and an innermost superconducting shield. Further active compensation of residual fields was achieved with compensation coils oriented along three axes, and used in conjunction with on-chip superconducting quantum interference device (SQUID) magnetometers located near each of the four corners of the processor block (Fig. 4.9**a**). Compensation coils were adjusted to minimise the magnetic field measured at the magnetometers while the chip was at 4.2 K. The chip was then thermally cycled just above and then back down through its superconducting transition temperature at this minimal field. We estimate that the chip was cooled through its superconducting transition with a field normal to the chip surface $|B_{\perp}| < 2.5$ nT, and that parallel to its surface $|B_{\parallel}| < 3.6$ nT over the area of active circuitry.

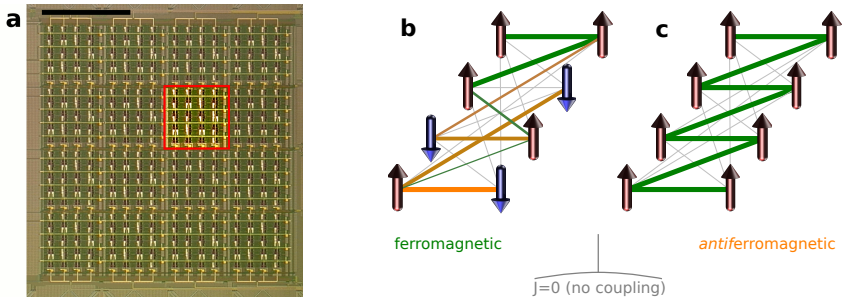

Figure 4.9: **(a)** Optical photograph of a portion of a partially fabricated 128 qubit chip. The block of eight qubits used in this experiment is outlined in red. **(b)** Artificial spins are connected in a complete bipartite graph $K_{4,4}$, and interact via couplers which are continuously tunable from ferromagnetic to antiferromagnetic interaction. A line between artificial spins indicates that a coupler is present. The colouring indicates one possible arrangement of coupler settings. **(c)** An example of how a linear ferromagnetic Ising spin chain could be implemented by selectively tuning some couplers to a ferromagnetic setting ($J < 0$) (green), and turning off the rest ($J = 0$).

4.7.3 Experimental method

The experiment discussed in the manuscript is outlined in Table 4.2. The steps in part **I** were performed once and would, in general, only be performed once for a new chip. The calibration step **I**-1 is performed by measuring the circulating current I_p in each qubit, and its dependence on the CJJ loop flux bias Φ_{2x}. From this information, one can extract the qubit critical current I_c and inductance L. Details of this procedure are discussed in detail in section IV.A of Ref. [82]. Given these qubit parameters, the effective inter-qubit coupling strength attained by the tunable couplers can be determined. This was done by measuring

the difference in magnetic flux coupled into a qubit B between states $|\uparrow\rangle$ and $|\downarrow\rangle$ of a qubit A. This coupled flux was measured as a function of the setting of the tunable coupler between qubits A and B, in a manner described in detail in Ref. [86].

Table 4.2: Outline of experiment

I. Initialisation

1. **Calibration:** measure intrinsic device parameters such as junction I_c, qubit inductance, transformer mutual inductances, etc.
2. **Homogenisation:** use on-chip programmable flux biases to ensure I_p of the different qubits match during annealing.

II. Annealing & read-out

1. **Set h, J**
2. **Anneal** (reduce $A(\tau)$ and increase $B(\tau)$)
3. **Read** state of spins

Once the device parameters for each qubit have been extracted, the effective junction I_c and inductance L of each qubit are tuned with on-chip tuning structures so as to make them as similar to each other as possible. The goal of this homogenisation procedure is to ensure that the circulating currents, I_p, of several qubits remain close to each other in magnitude while the qubits undergo annealing. This procedure is discussed in detail in Refs. [85] and [81]. On-chip tuning structures enabling this homogenisation are also described in Refs. [82] and [98]. Figure 4.10 shows the superimposed plots of the measured circulating current I_p (left) and tunnel splitting $A(\tau)$ (right) of each of the eight qubits used in this experiment after homogenisation. Qubit capacitance is extracted by measuring the spacing of macroscopic resonant tunnelling rate peaks[83]. At any point in Φ_{2x}, the standard deviation of the measured I_p across the 8 qubits is less than 25 nA. The uncertainty in each measurement of I_p is about 9 nA. The homogenised device parameters are summarised in Table 4.3.

The steps in **II** are performed repeatedly. Step **II**-1 is where the Hamiltonian parameters h_i and J_{ij} from Eq. 4.25 are programmed. For each such problem specification, steps **II**-2 and **II**-3 were repeated to allow collection of statistics about the relative probabilities of the possible states. For data presented in this chapter related to Experiment 6, **II**-1 was repeated 8 times, after each of which, **II**-2 and **II**-3 were repeated 4096 times, for a total of 32,768 repetitions of **II**-2 and **II**-3. However, step **II**-1 non-negligibly heated the chip, so in order to allow ample time for the chip to cool back to the base temperature, the first 512

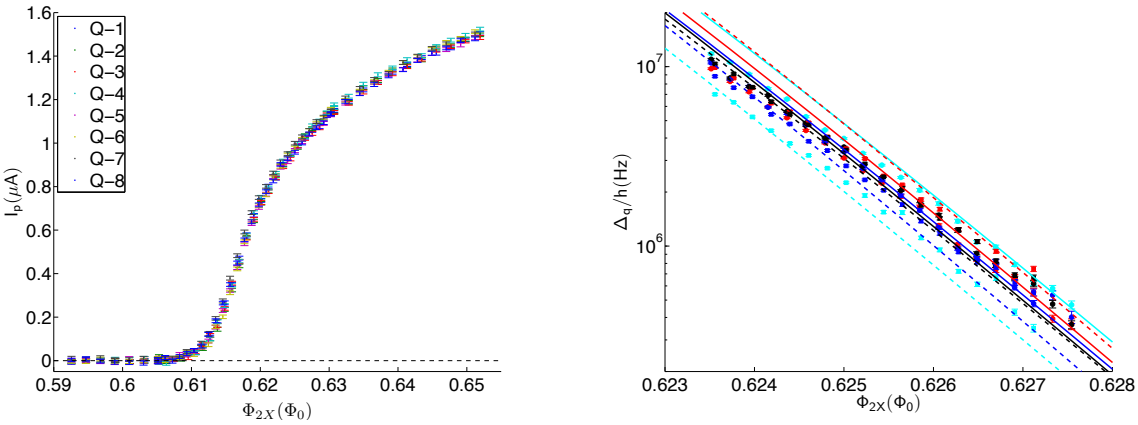

Figure 4.10: (Left) Measured circulating current I_p of each of the eight qubits used in this experiment after homogenisation (step **I**-2). (Right) Comparison of measured tunnel splitting $A(\tau)$ (labelled Δ_q) for the eight qubits used in the experiment, and the $A(\tau)$ fit to a physical model of the rf-SQUID. Fits of measured tunnel splitting $A(\tau)$ are used in conjunction with fits to I_p and MRT rate measurements to estimate parameters shown in Table 4.3.

repetitions after each execution of step **II**-1 were removed, leaving $8 \times 3,584$=$28,672$ total repetitions of **II**-2 and **II**-3. In the case of Experiments 1-5 the statistics were collected over 10,000 measurements in each experiment and enough thermalization time was allowed. Therefore, all data was included in the statistics without the need for removing any of the initial measurements. The experimental results of the probabilities measured are reported as percentages in Figs. 4.3 and 4.5.

Annealing was performed by raising the single qubit tunneling barrier. This is accomplished by changing Φ_{2x} linearly in time, from 0.592 Φ_0 to 0.652 Φ_0, over a period of 148 μs, as shown in Figure 4.11. Circulating current I_p shown in Figure 4.10 is plotted over exactly this range of Φ_{2x}. This also has the effect of changing parameters $A(\tau)$ and $B(\tau)$ from Eq. 4.3, as shown in Fig. 4.2(b), and as discussed in Ref. [79]. Control points α and β in Figure 4.11 correspond to the beginning and ending times of Fig. 4.2(b).

After the qubits have completed annealing, when Φ_{2x} has been set to Φ_0 as shown on the right in Figure 4.11, states of the spins are read with a hysteretic dc-SQUID readout, as described in Ref. [21].

4.7.4 Thermometry

In addition to a Ruthenium Oxide thermometer mounted on the dilution refrigerator mixing chamber, the effective qubit device temperature obtained during the measurements

Table 4.3: Total Josephson junction critical current, qubit inductance, inductance of loop 2, and junction capacitance extracted from circulating current and tunnel splitting measurements, and Macroscopic Resonant Tunneling (MRT) peak spacing.

Qubit	$I_c(\mu A)$	L_1 (pH)	L_2 (pH)	C(fF)
1	3.350	337.9	26	185
2	3.363	339.7	26	190
3	3.340	333.0	26	190
4	3.363	338.5	26	190
5	3.340	334.0	26	195
6	3.352	334.8	26	190
7	3.365	338.8	25	185
8	3.330	332.9	26	190

discussed in the manuscript was determined in two independent ways. The first is based on analysis of the single-qubit Macroscopic Resonant Tunnelling (MRT) rate, and its dependence on the qubit loop flux bias Φ_{1x}. Measurements and analysis of MRT rates for the devices used in this experiment are discussed in Ref [115]. The second is based on measurement of the equilibrium $P_\uparrow$ vs. Φ_{1x} attained at fixed barrier height (fixed value of Φ_{2x}). Both of these techniques are discussed in some detail in Ref. [83].

At a fixed barrier height achieved with a fixed value of Φ_{2x}, the equilibrium probability $P_\uparrow$ approaches the thermal distribution:

$$P_\uparrow(t \to \infty) = \frac{1}{2}\left[\frac{1}{2} + \tanh\left(\frac{I_p \Phi_{1x}}{k_B T_{th}}\right)\right] \tag{4.26}$$

where I_p is the value of circulating current obtained at that value of Φ_{2x} and T_{th} is the effective device temperature. Fitting a measurement of $P_\uparrow$ as a function of Φ_{1x} to Eq. 4.26, combined with a knowledge of I_p, allows us to extract T_{th}.

Measurement of T_{th} was performed on two of the devices at each temperature setting. An average of at least two independent measurements of the device temperature T_{th} of each of two qubits is compared against the mixing chamber thermometer temperature reading (T_{MXC}) in Figure 4.12. Uncertainty in T_{th} was dominated by the uncertainty in the fit transition width for each measurement, which was generally found to be larger than the standard deviation of the separate measurements.

The temperature extracted from MRT transition rate widths (T_{MRT}) is also plotted vs. T_{MXC} for temperatures below 40 mK, in Figure 4.12. From these plots it is clear that the two methods generally agree with each other as well as with the mixing chamber thermometer to within 3 mK over the temperature range used in the experiment.

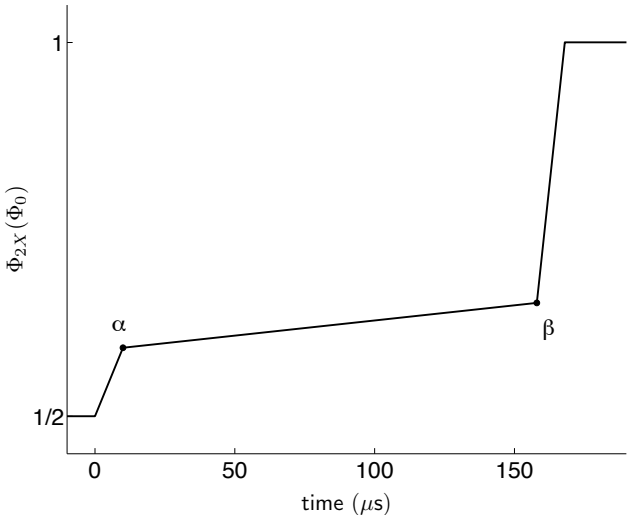

Figure 4.11: The annealing schedule is defined by the applied flux $\Phi_{2x}(t)$. The qubits make a transition between being monostable and bistable between control points $\alpha = (10\ \mu\text{s}, -0.592\ \Phi_0)$ and $\beta = (158\ \mu\text{s}, -0.652\ \Phi_0)$.

4.8 Appendix 4. Quantum simulations

To obtain better quantitative understanding of the behaviour of the system, a simulation was conducted to model this experiment. The agreement between the numerical simulations can be seen in panel (b) of Fig. 4.5, where both percentages [experiment (theory)] are reported next to each other for each one of the low-energy conformations.

Our simulation strategy is as follows: We first write a Hamiltonian for the superconducting circuit based on standard circuit models for capacitances, inductances, and Josephson junctions. This Hamiltonian is expected to correctly describe the behaviour of coupled rf-SQUIDs. We then numerically calculate the evolution of the system based on this Hamiltonian using quantum mechanical equations of motion which take into account coupling to an environment. Therefore, we predict the quantum evolutions for the same system Hamiltonian, the same coupling to environment, and the same type of noise spectral densities. This provides a fair comparison to the experimental data.

4.8.1 rf-SQUID Hamiltonian

A simplified version of the rf-SQUID qubit used in our processor is illustrated in Fig. 4.8**a**. (A more complete description of the actual qubits can be found in Ref. [82].) It has two main superconducting loops and therefore two flux degrees of freedom Φ_1 and

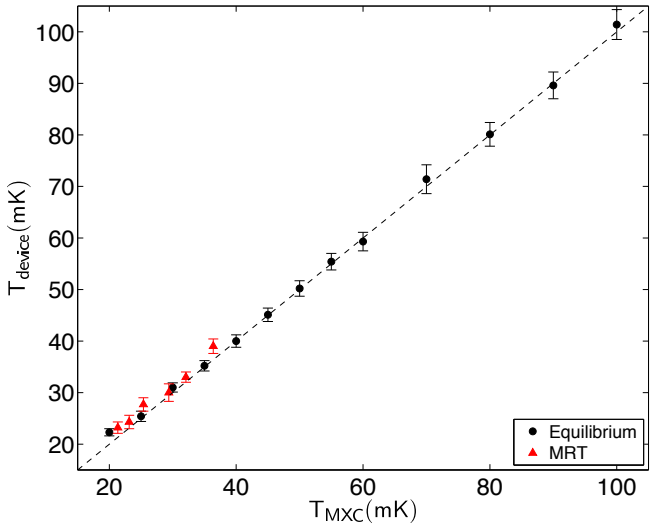

Figure 4.12: Plots of T_{th} (black circles) and T_{MRT} (red triangles) vs. the temperature measured with the Ruthenium Oxide thermometer mounted on the mixing chamber T_{MXC}.

Φ_2, subject to external flux biases Φ_{1x} and Φ_{2x}, respectively. The Hamiltonian of such an rf-SQUID is written as

$$H_{\text{SQUID}} = \frac{q_1^2}{2C_1} + \frac{q_2^2}{2C_2} + U(\Phi_1, \Phi_2) \tag{4.27}$$

where C_1 and C_2 are parallel and series combinations of the junction capacitances, q_1 and q_2 are the sum and difference of the charges stored in the two Josephson junctions respectively, and

$$\begin{aligned} U(\Phi_1, \Phi_2) = (\Phi_1 - \Phi_{1x})^2/2L_1 + (\Phi_2 - \Phi_{2x})^2/2L_2 \\ -2E_J \cos(\pi\Phi_2/\Phi_0)\cos(2\pi\Phi_1/\Phi_0), \end{aligned} \tag{4.28}$$

is a 2-dimensional potential with L_i being the inductances of the two loops and $\Phi_0 = h/2e$, the flux quantum. We have assumed symmetric Josephson junctions with Josephson energies $E_J = I_c\Phi_0/2\pi$, where I_c is the junctions' critical current. (A small asymmetry can be tuned away in situ in the physical implementation [82].)

At $\Phi_{1x} \approx \Phi_0/2$, the potential can become bistable and therefore form a two-dimensional double-well potential. If L_2 is small enough so that the deviation of Φ_2 from Φ_{2x} can be neglected, then the two-dimensional classical potential $U(\Phi_1, \Phi_2)$ can be approximated by a one-dimensional double-well potential, as shown in Fig. 4.8**b**. However, with our realistic qubit parameters, Φ_2 cannot be neglected and therefore is accounted for in all our numerical

calculations. When $\Phi_{1x} = \Phi_0/2$, the two wells are symmetric with no energy bias between them. One can tilt the potential by changing Φ_{1x} and establish an energy bias, as depicted in Fig. 4.8**b**. It is also possible to change the barrier height by changing Φ_{2x}.

An array of such qubits can be modelled by summing contributions of Eq. (4.27) from each device plus terms that describe magnetic coupling of the loops:

$$H_S = \sum_i H^{(i)}_{\text{SQUID}} + \sum_{i>j} H^{(ij)}_{\text{coupl}} \tag{4.29}$$

Coupling between qubits i and j can be modelled as a mutual inductance M_{ij} between loop 1 of each pair of coupled qubits:

$$H^{(ij)}_{\text{coupl}} = (\Phi^{(i)}_1 - \Phi^{(i)}_{1x})(\Phi^{(j)}_1 - \Phi^{(j)}_{1x})M_{ij}/L^{(i)}_1 L^{(j)}_1 \tag{4.30}$$

As discussed in Section 4.7.3 above, all parameters, i.e., inductances $L^{(i)}_\alpha$, capacitances $C^{(i)}_\alpha$, and Josephson critical currents $I^{(i)}_c$, are measured independently for each qubit.

To describe the system accurately we also need to introduce interaction with environment. Flux noise, which is the dominant noise in flux qubits, couples to the ith qubit as fluctuations $\delta\Phi^{(i)}_{\alpha x}$ of the external flux $\Phi^{(i)}_{\alpha x}$:

$$H_{\text{int}} = -\sum_{\alpha=1}^{2}\sum_i \frac{\Phi^{(i)}_\alpha - \Phi^{(i)}_{\alpha x}}{L^{(i)}_\alpha}\delta\Phi^{(i)}_{\alpha x} \tag{4.31}$$

The noise is much smaller for the smaller loop $\Phi^{(i)}_{2x}$ than for the larger loop $\Phi^{(i)}_{1x}$ due to the loop size. The flux noise $\delta\Phi^{(i)}_{\alpha x}$ is assumed to be uncorrelated between the qubits, which agrees with recent experimental observation[117].

Chip calibration and device parameter extraction

Device parameters were extracted for the simulations through a series of independent measurements of qubit circulating current, tunnel splitting Δ, and MRT peak spacing. A discussion of how these measurements are performed is given in Ref. [82]. Parameter values used in simulations are summarised in Table 4.3.

4.8.2 Quantum Simulation

To simulate the quantum mechanical dynamics of the system, we treat (4.29)-(4.31) as quantum mechanical Hamiltonians. In that case, the charge $q^{(i)}_\alpha$ is taken to be an operator, which is the momentum conjugate to the flux operator $\Phi^{(i)}_\alpha$ with commutation relation: $[\Phi^{(i)}_\alpha, q^{(i)}_\alpha] = i\hbar$. Unfortunately, it is impossible to calculate the dynamics of the system directly on the $2N$-dimensional continuous potential quantum mechanically. Instead, we

use energy discretization as a means to simplify the calculation. The simplest way to accomplish this is to treat an rf-SQUID as a 2-state system or qubit and replace (4.29) by a coupled qubit Hamiltonian. One may go further and keep more than two states per rf-SQUID in the calculation, as we shall discuss below.

We first numerically diagonalise the single rf-SQUID Hamiltonian (4.27) to obtain the lowest eigenvalues and eigenvectors. We treat the lowest few energy levels as the subspace relevant for computation. We then write the Hamiltonian in the basis of states that are localised within the wells. Such states are not true eigenfunctions of the Hamiltonian and, therefore, are metastable towards tunnelling to the opposite well. Hence, the resulting Hamiltonian in such a basis will have off-diagonal terms between states in the opposite wells but not between states within each well. The latter is because those states should be stationary within their own wells; any transition (relaxation) between them is only induced by the environment.

Let $|l\rangle$ denote localised states within the wells. We use even (odd) state numbers, i.e., $l = 2n$ $(2n{+}1)$, with $n = 0, 1, 2, ...$, to denote states that are localised in the left (right) well. For the lowest M energy levels (M is taken to be even), the effective $M \times M$ tunnelling Hamiltonian is written as

$$H_S = \sum_{l=0}^{M-1} E_l |l\rangle\langle l| + \sum_{n,m=0}^{M/2-1} K_{2n,2m+1}(|2n\rangle\langle 2m{+}1| + |2m{+}1\rangle\langle 2n|) \tag{4.32}$$

where E_l is the energy expectation value for state $|l\rangle$ and $K_{2n,2m+1}$ is the tunnelling amplitude between states $|2n\rangle$ and $|2m{+}1\rangle$, which exist in opposite wells. Notice that there is no matrix element between states on the same well: $\langle 2n|H_S|2m\rangle = \langle 2n{+}1|H_S|2m{+}1\rangle = 0$, which means that the states are metastable only towards tunnelling to the other side, or the states are quasi-eigenstates of the Hamiltonian within their own sides. All parameters of the tunnelling Hamiltonian, i.e., E_l and $K_{ll'}$ are extracted from the original rf-SQUID Hamiltonian (4.27). For the 2-state qubit model we keep only the lowest two energy levels of (4.32). The effective qubit Hamiltonian can be written as

$$H_{eff} = -\frac{1}{2}(\epsilon\sigma_z + \Delta\sigma_x) \tag{4.33}$$

where

$$\epsilon = E_0 - E_1, \qquad \Delta = -2K_{01}, \tag{4.34}$$

We also go beyond the 2-state model and keep 4 states per rf-SQUID. Those 4 states can be represented by two coupled qubits, one of which represents the direction of persistent current or flux, and the other one generating intrawell energy levels. We represent the first

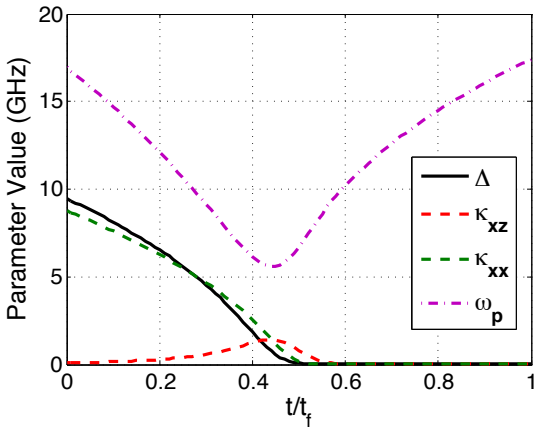

Figure 4.13: Parameters of the 4-level model for rf-SQUID qubit as a function of time during annealing.

(logical) qubit by Pauli matrices σ_α, and the extra (ancilla) qubit by Pauli matrices τ_α. The effective Hamiltonian for those two coupled qubits can be written as

$$H_{eff} = -\frac{1}{2}(\epsilon\sigma_z + \Delta\sigma_x) + \frac{1}{2}[\omega_p\tau_z + \kappa_{xz}\sigma_x(1+\tau_z) + \kappa_{xx}\sigma_x\tau_x]. \tag{4.35}$$

It is easy to show that (4.35) is equivalent, up to a constant energy, to Hamiltonian (4.32), with $M = 4$, if

$$\epsilon = E_0 - E_1 = E_2 - E_3, \qquad \omega_p = E_2 - E_0 = E_3 - E_1, \qquad \Delta = -2K_{01}, \tag{4.36}$$

$$\kappa_{xz} = K_{23} - K_{01} \approx K_{23}, \qquad \kappa_{xx} = 2K_{03} = 2K_{12}. \tag{4.37}$$

As can be seen, the coupling between logical and ancilla qubits are of XX+XZ type. Coupling qubits to each other is accomplished using σ_z operators which represent the direction of the induced flux. The ancilla qubits remain uncoupled from each other and from other qubits. It should be noted that the readout at the end of the evolution can only distinguish "left" well from "right" well in the double-well potential and cannot distinguish levels within each well. This is equivalent to reading out logical qubits and not ancilla qubits, but as we mentioned above, only logical qubits carry information.

To properly treat the environment, we need to write the interaction Hamiltonian (4.31) in the subspace of the lowest energy levels in terms of Pauli matrices. For quantum simulations we only consider noise coupling to the larger loop in Fig. 4.8**a**. Let us consider a single rf-SQUID and write the interaction Hamiltonian as

$$H_{\text{int}} = -\frac{\Phi_1 - \Phi_{1x}}{L_1}\delta\Phi_{1x} \tag{4.38}$$

We define the qubit persistent current by

$$I_p = \frac{1}{L_1} \left| \langle l | (\Phi_1 - \Phi_{1x}) | l \rangle \right| . \tag{4.39}$$

Here, we take I_p to be independent of $|l\rangle$ for the low lying states considered, although in reality there could be a small dependence. The interaction Hamiltonian can then be written as

$$H_{\text{int}} = -\frac{1}{2}(\sigma_z + \lambda \tau_x) Q, \tag{4.40}$$

where

$$Q = 2 I_p \delta \Phi_{1x}, \qquad \lambda = \frac{\langle 0 | H_{\text{int}} | 2 \rangle}{2 I_p L_1} = \frac{\langle 1 | H_{\text{int}} | 3 \rangle}{2 I_p L_1} \tag{4.41}$$

The matrix elements $\langle 0|H_{\text{int}}|2\rangle$ or $\langle 1|H_{\text{int}}|3\rangle$ are calculated directly via Eq. 4.38 using the eigenfunctions of the rf-SQUID Hamiltonian, Eq. 4.27). The values of I_p and λ can therefore be calculated numerically from the original rf-SQUID Hamiltonian. Only Q remains which should be characterised via its spectral density, which is the subject of Appendix A.

Quantum evolution of the system was calculated using a Markovian master equation for the density matrix described in Ref. [12]. Since the evolution is very slow (adiabatic) and temperature is low, only a small number of energy levels are expected to be occupied during the evolution. We write the density matrix in the instantaneous energy eigenstate basis and truncate it to the lowest 24 energy levels, which was found to sufficiently describe the type of evolution studied here. We use both 2-state and 4-state models for rf-SQUIDs, as described above, in our simulations. The result of the 4-level model simulation is shown in Fig. 4.5**b**.

4.8.3 Noise spectral density

The quantum noise operator $Q = 2 I_p \delta \Phi_{1x}$ is related to the flux noise as expected (for simplicity we only consider one rf-SQUID), and is characterised by its correlation function. Let us define the spectral density

$$S(\omega) = \int_{-\infty}^{\infty} dt \; e^{i\omega t} \langle Q(t) Q(0) \rangle = 4 I_p^2 S_\Phi(\omega) \tag{4.42}$$

where

$$S_\Phi(\omega) = \int_{-\infty}^{\infty} dt \; e^{i\omega t} \langle \delta \Phi_{1x}(t) \delta \Phi_{1x}(0) \rangle \tag{4.43}$$

is the spectral density of the flux noise. No direct measurement of $S_\Phi(\omega)$ at all frequencies is available. We assume the spectral density is a sum of low frequency and high frequency components: $S_\Phi(\omega) = S_{LF}(\omega) + S_{HF}(\omega)$. For the low frequency component we use

$$S_{LF}(\omega) = \frac{(A^2/k_B T) \hbar \omega |\omega|^{-\alpha}}{1 - e^{-\hbar\omega / k_B T}}, \tag{4.44}$$

with $\alpha \approx 1$, which at low ω behaves as 1/f noise: $\sim A^2|\omega|^{-\alpha}$. Parameter A is measured from low frequency noise measurement [116] and is found to be $A \approx 3$ nΦ_0.

The high frequency parts of the spectral density is assumed to be ohmic,

$$S_{HF}(\omega) = \left(\frac{\hbar^2}{4I_{p0}^2}\right) \frac{\eta\omega e^{-|\omega|/\omega_c}}{1 - e^{-\hbar\omega/k_B T}}, \tag{4.45}$$

where ω_c is the upper cutoff frequency, η is the dimensionless coupling coefficient, and I_{p0} is the value of persistent current at which η is measured. The coupling coefficient and the persistent current are found, using Macroscopic resonant tunnelling experiment (MRT), to be $\eta \approx 0.4$ and $I_{p0} \approx 1$ μA. The details of extraction of η via MRT are presented elsewhere [115].

This leaves no free parameters for the quantum simulations.

Part II

Excitonic energy transfer and nonlinear fluorescence spectroscopy

Chapter 5

Engineering directed excitonic energy transfer

Summary

We provide an intuitive platform for engineering exciton transfer dynamics. We show that careful consideration of the spectral density, which describes the system-bath interaction, leads to opportunities to engineer the transfer of an exciton. Since excitons in nanostructures are proposed for use in quantum information processing and artificial photosynthetic designs, our approach paves the way for engineering a wide range of desired exciton dynamics. We carefully describe the validity of the model and use experimentally relevant material parameters to show counter-intuitive examples of a directed exciton transfer in a linear chain of quantum dots.

5.1 Introduction

The widely-applied Förster theory for energy transfer links experimental results to estimates of system information, particularly in biological and nanoscale applications [170, 114]. The success of this theory is partly due to the simple expression of the kinetic rate constants as a product of electronic coupling and a spectral overlap factor which captures the complexity of the environment. Förster theory describes transport in the incoherent limit, but a more sophisticated approach, such as Redfield theory, is often required to describe energy transfer. However, the information essential to understanding the dynamics is buried within the structure of the equations. In this chapter, we employ a quantum kinetic rate approach to distill the information contained in equations into a simple, yet instructive, formula. We use this approach to design directed exciton transfer mediated by an environment.

Excitonic energy transfer (EET) has been studied in systems as varied as quantum dot (QD) nanostructures [171, 50], polymer chains [43], and photosynthetic complexes

[121, 38]. Many applications of EET would benefit from controlling exciton dynamics. Perfect state transfer, as studied in the quantum computing community, is achievable in certain engineered systems, but only at particular times during coherent evolution [42]. Recent works have shown that environment-induced decoherence can alter exciton dynamics [166, 158, 151], although controlling the transfer direction has only been achieved using external potentials [90]. This chapter builds upon the idea of engineering exciton transfer by designing appropriate system-bath interactions [71, 35]. We show that it is possible to design experimentally realizable systems where the environment can be used to direct the flow of energy.

5.2 Model Hamiltonian for the excitonic energy transfer dynamics

The Hamiltonian used in our simulation aims to capture dynamics in a single-exciton manifold [129] interacting with an environment,

$$\hat{H} = \hat{H}_s + \hat{H}_b + \hat{H}_{sb} \tag{5.1}$$

with

$$\hat{H}_s = \sum_{n=1}^{N} E_n \left|s_n\right\rangle \left\langle s_n\right| + \sum_{n \neq m} J_{mn} \left|s_m\right\rangle \left\langle s_n\right|. \tag{5.2}$$

This representation is in the *site basis* $\{|s_n\rangle\}$ of localized excitations on each of N sites, (e.g. QDs or chromophores), with excitation energy E_n for each site and inter-site coupling J_{mn}. The environment is described by a phonon bath,

$$\hat{H}_b = \sum_{\mathbf{q}} \hbar\omega_{\mathbf{q}} (b_{\mathbf{q}}^{\dagger} b_{\mathbf{q}} + 1/2), \tag{5.3}$$

where $b_{\mathbf{q}}^{\dagger}$ ($b_{\mathbf{q}}$) is the creation (destruction) operator for a phonon with wavevector $\mathbf{q}$. The system-bath interaction is assumed to be linear,

$$\hat{H}_{sb} = \sum_{n=1}^{N} \left|s_n\right\rangle \left\langle s_n\right| \sum_{\mathbf{q}} \hbar\omega_{\mathbf{q}} \Big(g_{\mathbf{q}}^{n} b_{\mathbf{q}}^{\dagger} + (g_{\mathbf{q}}^{n})^{*} b_{\mathbf{q}} \Big). \tag{5.4}$$

where $g_{\mathbf{q}}^{n}$ describes the site-specific coupling of electronic and vibrational degrees of freedom. As usual, we ignore the off-diagonal terms in the above equation, as they correspond to inter-site exciton transfer via absorption or emission of a phonon [129].

To describe the excitonic quantum dynamics we use Redfield theory [159, 153, 143], which is a reduced density matrix approach in the regime of weak system-bath coupling. The

formalism involves second-order perturbation theory in the system-bath interaction, $\hat{H}_{sb}$. This method assumes the Markov approximation, no initial correlations between system and bath degrees of freedom, and a thermalized bath. Since the secular approximation is not used, coherence to population transfers are taken into account, thereby providing a better physical description than the widely used Bloch equations [129, 143].

The equation of motion for the density operator $\hat{\rho}(t)$ in the excitonic energy basis representation, $\hat{H}_s \left|e_a\right\rangle = \epsilon_a \left|e_a\right\rangle$, is given by [159, 153, 143],

$$\frac{\mathrm{d}\rho_{ab}(t)}{\mathrm{d}t} = -i\omega_{ab}\rho_{ab}(t) + \sum_{cd} R_{ab,cd}\rho_{cd}(t), \tag{5.5}$$

with $\rho_{ab}(t) \equiv \left\langle e_a\right| \hat{\rho}(t) \left|e_b\right\rangle$ and $\omega_{ab} = (\epsilon_a - \epsilon_b)/\hbar$. The first term on the right hand side of Eq. 5.5 describes the fully coherent dynamics in the absence of $\hat{H}_{sb}$ and the second term describes the irreversible dynamics from the interaction with the phonon bath.

Correlations in bath-fluctuations on different sites [162] are taken into account by using the relation $g^n_{\mathbf{q}} g^{m*}_{\mathbf{q}} = g^m_{\mathbf{q}} g^{n*}_{\mathbf{q}} = g^2_{\mathbf{q}} e^{-R_{mn}/R_{\mathrm{corr}}}$, where $R_{mn} \equiv |\mathbf{R}_{mn}|$ is the distance between the sites and R_{corr} is the phonon correlation length [160].

Using this relation for the electron-phonon couplings, the cross-correlation $C_{mn}(\omega)$ can now be written as,

$$C_{mn}(\omega) = e^{-R_{mn}/R_{\mathrm{corr}}} C(\omega), \tag{5.6}$$

where the frequency correlation function $C(\omega) = 2\pi[n(\omega)+1](J(\omega)-J(-\omega))$, with $n(\omega)$ the Bose-Einstein distribution and the spectral density of the bath $J(\omega) = \sum_{\mathbf{q}} |g_{\mathbf{q}}|^2 \omega^2_{\mathbf{q}} \delta(\omega - \omega_{\mathbf{q}})$, where $\delta(\omega)$ is the Dirac delta function.

For the Hamiltonian specified in Eq. 5.1, the Redfield tensor elements are given by

$$\begin{aligned} R_{ab,cd} = {} & \Gamma_{db,ac}(\omega_{ca}) + \Gamma^*_{ca,bd}(\omega_{db}) \\ & - \delta_{bd} \sum_e \Gamma_{ae,ec}(\omega_{ce}) - \delta_{ac} \sum_e \Gamma^*_{be,ed}(\omega_{de}). \end{aligned} \tag{5.7}$$

where δ_{ij} is the Kronecker delta and

$$\Gamma_{ab,cd}(\omega_{dc}) = \frac{1}{2}\zeta_{ab,cd} C(\omega_{dc}) + \frac{i}{2\pi}\zeta_{ab,cd} \mathcal{P} \left(\int_{-\infty}^{\infty} \frac{C(\omega)}{\omega_{dc} - \omega} \, \mathrm{d}\omega \right), \tag{5.8}$$

$$\zeta_{ab,cd} = \sum_{n,m} (U^{-1})_{an} U_{nb} (U^{-1})_{cm} U_{md} e^{-R_{mn}/R_{\mathrm{corr}}}, \tag{5.9}$$

with $\mathcal{P}$ denoting the Cauchy principal value of the integral and $U_{na} = \left\langle s_n \right| \left|e_a\right\rangle$ the transformation matrix elements relating the site basis $\{|s_n\rangle\}$ and the excitonic basis $\{|e_a\rangle\}$. Since $|e_a\rangle = \sum_n U_{na} |s_n\rangle$, then $|U_{na}|^2$ can be interpreted as the contribution of the n-th site to the a-th eigenstate of $\hat{H}_s$.

Hereafter, we will focus on QDs as our prototypical experimental realization. We neglect possible inversion asymmetry of the crystal, and therefore the contribution of the piezoelectric coupling, and focus instead on the deformation potential coupling. As shown by Calarco *et al.* [33], the spectral density describing this coupling (in the absence of an external electric field) is given by

$$J(\omega) = \Theta(\omega)\eta\omega^3 e^{-\omega^2/\omega_c^2}, \tag{5.10}$$

where, $\Theta(\omega)$ is the Heaviside step function. We use typical values for GaAs QDs for our numerical simulations [126], giving $\eta = \frac{(D_e - D_h)^2}{4\pi^2\rho u^5 \hbar} = 0.035$ ps^2 and $\omega_c = \sqrt{2u^2/l^2} = 1.41$ ps^{-1}, where D_e (D_h) is the deformation coupling potential for electron (hole), u the speed of sound within the quantum dot, ρ its mass density and l the ground state localization length, assumed to be the same for electron and holes. We also assume a correlation length, R_{corr}, of 3 nm. We consider a typical GaAs QD with deformation potentials $D_e = -14.6$ eV, $D_h = -4.8$ eV, mass density of $\rho = 5.4$ g/cm^3, speed of sound $u = 5000$ cm/s, and radius of $l = 5$ nm.

Using Eqns. 5.7-5.9, the population transfer rates, k_{ab}, between the eigenstates $a \to b$ is given by,

$$k_{ab} = R_{bb,aa} = \zeta_{ab,ba} C(\omega_{ab}). \tag{5.11}$$

This equation is central to our insight into designing excitonic transfer. Though derived for a different regime, the form of Eq. 5.11 is similar to the widely used rate equation in Förster theory (incoherent limit). The latter has been successfully applied in fluorescence resonance energy transfer (FRET) to design chromophores for biosensing assays [114, 163, 130] and was recently verified experimentally for semiconductor QDs [104]. We use the simple structure of Eq. 5.11 to gain microscopic and experimentally relevant insight into engineering directed and optimized EET. While $C(\omega_{ab})$ depends on the overlap of system eigenenergies with the spectral properties of the phonon bath (e.g., lattice vibrations, solvent, protein environment, etc.), ζ depends on the transformation matrix U in Eq. 5.9, determined by the relative magnitude of electronic couplings with respect to site energies, site connectivity, and spatial correlation between sites. The aim is then to maximize (minimize) the product of these two factors in Eq. 5.11 to favor (suppress) the desired rates.

To illustrate the applicability of Eq. 5.11, we choose two three-site examples to highlight the importance of the phonon bath interaction to achieve directed EET (Fig. 5.1). Since multiplying a system Hamiltonian by a scalar does not change the maximum exciton transfer probability in the coherent limit, the distances and site energies for the two cases are chosen so that the Hamiltonians are related by a multiplicative factor ($3.5\hat{H}_s^{\mathbf{1}} = \hat{H}_s^{\mathbf{2}}$). As a consequence, the ζ values are roughly the same for cases **1** and **2** (Table 5.1); slight differences are introduced by the bath correlation term in Eq. 5.9. In the cases chosen, the

fully coherent evolution gives a maximum probability of finding an excitation of 5% (1%) for site 1(2). Therefore, any difference between the two examples is a result of interaction with the environment.

In contrast to the ζ factors, the $C(\omega_{ab})$ values for cases **1** and **2** differ by at least one order of magnitude due the position of the transition frequencies with respect to the spectral density (Table 5.1). Any changes to the system Hamiltonian can affect $C(\omega_{ab})$. In our examples, the scalar factor mentioned above changes the $C(\omega_{ab})$ such that the largest population transfer rate is switched from site 1 to site 2 since $\omega_{31}^{\mathbf{1}} \approx \omega_{32}^{\mathbf{2}}$ and $C(\omega_{31}^{\mathbf{1}}) \approx C(\omega_{32}^{\mathbf{2}})$ (Fig. 5.1).

From the values of ζ and $C(\omega_{ab})$, it is clear that case **1** is designed such that an excitation starting on site 3 will tend to transfer to site 1, but in case **2** the population will go to site 2, albeit at different rates. Simulations of the quantum dynamics according to Redfield theory confirms this result (Fig. 5.2). Over the typical exciton lifetime of 1 ns in QDs, we not only achieve directed transfer, but also population enhancement compared to both the maximum site population during fully coherent dynamics and the population expected at thermal equilibrium.

Moreover, while there are always experimental limitations in tuning parameters, the structure of Eq. 5.11 is valuable since it partitions the effects due to the system and bath. Using the calculated rates and/or a visual inspection of $C(\omega)$ (Fig. 1B), it is easy to determine the impact of varying a system parameter on exciton transfer. Future work will address exciton and electron transfer between sites with varying spectral density functions, as well as the role of aligning dipole moment orientations in engineering EET. We are also working to identify regimes in which preserved coherences enhance or reduce the efficiency of excitonic transfer. Of course, in situations where multiple excitons are present in the system due to incident light intensity, frequency range, and/or optical spectral density of

Table 5.1: Contributions of the system factor, $\zeta_{ab,ba}$, and of the overlap between transition frequency and phonon bath spectral properties, $C(\omega_{ab})$, to the calculation of the quantum kinetic rates k_{ab} from energy eigenstate $|e_a\rangle$ to $|e_b\rangle$. The two cases considered are described in Fig. 5.1.

	Case 1			**Case 1**		
$\|e_a\rangle \rightarrow \|e_b\rangle$	$\log\zeta$	$\log C$	$\log k$	$\log\zeta$	$\log C$	$\log k$
$3 \rightarrow 1$	-1.5	11.6	**10.0**	-1.5	7.9	**6.3**
$3 \rightarrow 2$	-2.4	10.7	**8.3**	-2.4	11.6	**9.1**
$2 \rightarrow 1$	-4.3	11.4	**7.1**	-4.3	10.2	**5.8**

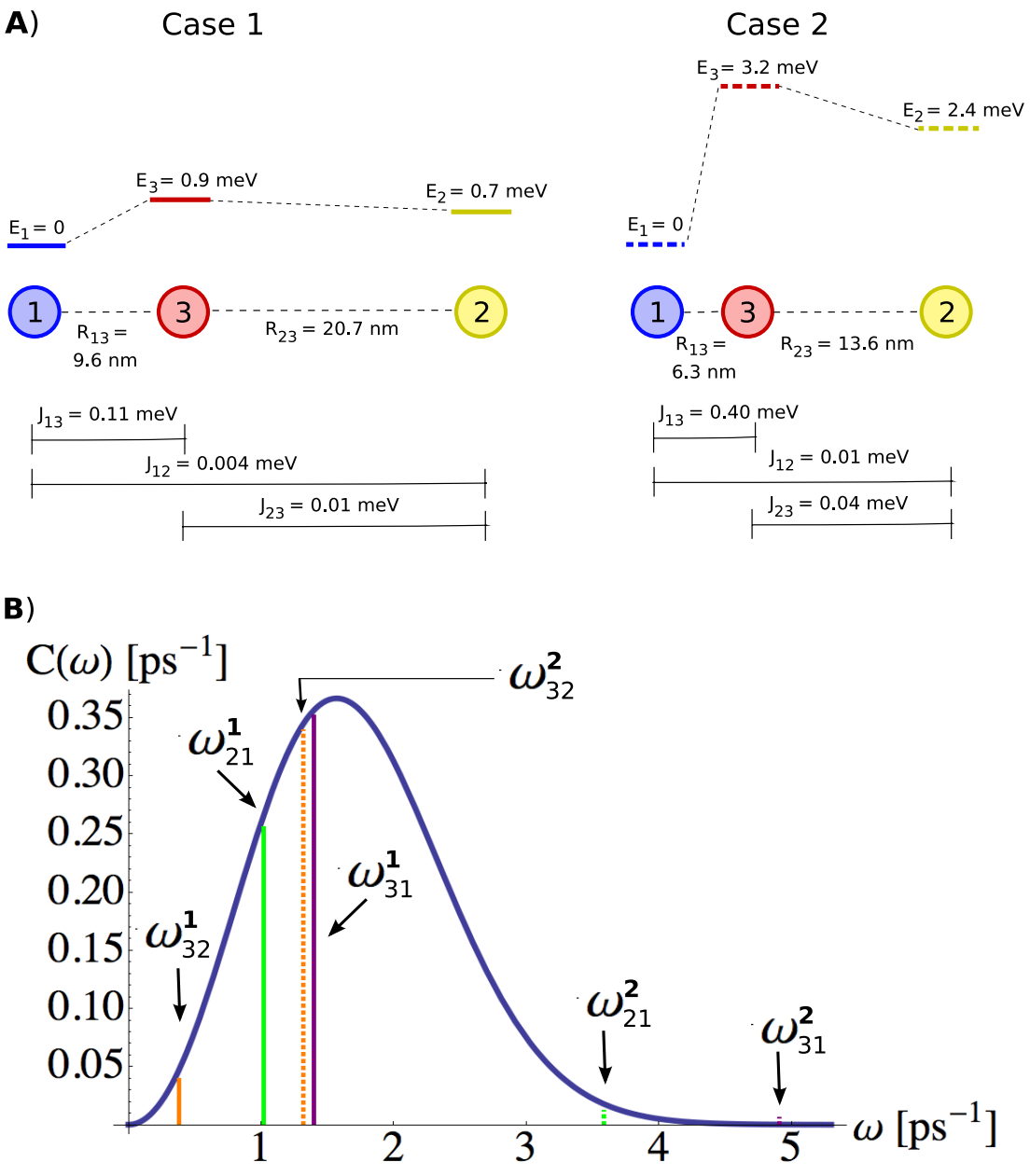

Figure 5.1: Two cases of a three-site system are considered. A) Scaled schematic of system spacing and energy levels with details of variable site energies (E_n) and intersite Förster coupling strengths (J_{mn}). For QDs with transition dipole moments aligned perpendicular to R_{mn}, $J_{mn} = 100\,\text{meV}/R_{mn}^3$, with R_{mn} in nanometers [140]. B) Frequency correlation function for a superohmic spectral density. Energy basis transition frequencies for case **1** (**2**) are indicated by solid (dashed) vertical lines.

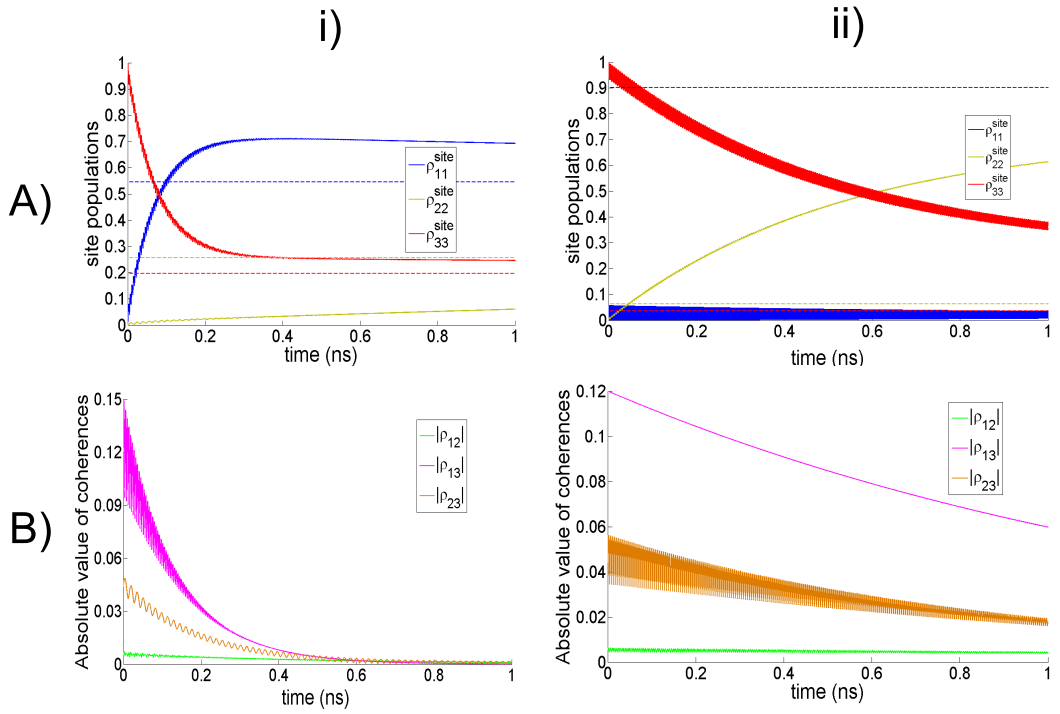

Figure 5.2: A) Site basis population probabilities for an excition starting on site 3 for case **1** (Ai) and case **2** (Aii) demonstrate the change in transfer dynamics obtained by scaling the Hamiltonian. Dashed lines indicate site populations at thermal equilibrium. B) Energy basis coherences for case **1** (Bi) and case **2** (Bii) display characteristic oscillations and damping over the course of 1 ns, a typical recombination time in QD systems.

the quantum dots, the Hamiltonian used to describe the system needs to be expanded accordingly [108]. To the extent that each exciton couples to the environment through a spectral density as described in this chapter, some of the intuition developed here should be transferable to these systems. However, there are a number of interesting complications, including many-body interactions among the excitons and their components; these effects are currently under investigation and are beyond the scope of the current communication.

5.3 Conclusion

In summary, we develop a framework for engineering environment-assisted and directed excitonic transfer in a network of coupled QDs based on a quantum kinetic rate approach. We emphasize the importance of how characteristic frequencies of a system fit within the spectral bath structure. Our examples utilize the factored and intuitive form of the population transfer rates equation, which separates the contributions from the system (electronic)

and bath (vibrational) degrees of freedom. This equation is similar in spirit to the rate equation for FRET, making it convenient to design interesting scenarios for environment-assisted transfer. Although we focus on QD examples, the principles presented here form the basis for engineering a wide range of desired EET in a variety of nanostructures or artificial molecular photosynthetic units.

Chapter 6

Conformation of self-assembled porphyrin dimers in liposome vesicles by phase-modulation 2D fluorescence spectroscopy

Summary

By applying a phase-modulation fluorescence approach to 2D electronic spectroscopy, we studied the conformation-dependent exciton coupling of a porphyrin dimer embedded in a phospholipid bilayer membrane. Our measurements specify the relative angle and separation between interacting electronic transition dipole moments and thus provide a detailed characterization of dimer conformation. Phase-modulation 2D fluorescence spectroscopy (PM-2D FS) produces 2D spectra with distinct optical features, similar to those obtained using 2D photon-echo spectroscopy. Specifically, we studied magnesium meso tetraphenylporphyrin dimers, which form in the amphiphilic regions of 1,2-distearoyl-sn-glycero-3-phosphocholine liposomes. Comparison between experimental and simulated spectra show that although a wide range of dimer conformations can be inferred by either the linear absorption spectrum or the 2D spectrum alone, consideration of both types of spectra constrain the possible structures to a T-shaped geometry. These experiments establish the PM-2D FS method as an effective approach to elucidate chromophore dimer conformation.

6.1 Introduction

The ability to determine three-dimensional structures of macromolecules and macromolecular complexes plays a central role in the fields of molecular biology and material science. Methods to extract structural information from experimental observations such as X-ray crystallography, NMR, and optical spectroscopy are routinely applied to a diverse array of problems, ranging from investigations of biological structure-function relationships

to the chemical basis of molecular recognition.

In recent years, two-dimensional optical methods have become well established to reveal incisive information about noncrystalline macromolecular systemsinformation that is not readily obtainable by conventional linear spectroscopic techniques. Two- dimensional optical spectroscopy probes the nanometer-scale couplings between vibrational or electronic transition dipole moments of neighboring chemical groups, similar to the way NMR detects the angstrom-scale couplings between adjacent nuclear spins in molecules [62]. For example, 2D IR spectroscopy probes the couplings between local molecular vibrational modes and has been used to study the structure and dynamics of mixtures of molecular liquids [113], aqueous solutions of proteins [63], and DNA [187]. Similarly, 2D electronic spectroscopy (2D ES) probes correlations of electronic transitions and has been used to study the mechanisms of energy transfer in multichromophore complexes. Such experiments have investigated the details of femtosecond energy transfer in photosynthetic proteinpigment arrays [31, 44, 1, 74], conjugated polymers [43], and semiconductors [214, 185].

Following the examples established by 2D NMR and 2D IR, 2D ES holds promise as a general approach for the structural analysis of noncrystalline macromolecular systems, albeit for the nanometer length scales over which electronic couplings occur. It is well known that disubstitution of an organic compound with strongly interacting chromophores can lead to coupling of the electronic states and splitting of the energy levels [112, 184, 128]. The arrangement of transition dipoles affects both the splitting and the transition intensities, which can be detected spectroscopically. Nevertheless, weak electronic couplings relative to the monomer linewidth often limits conformational analysis by linear spectroscopic methods alone. Two-dimensional ES has the advantage that spectroscopic signals are spread out along a second energy axis and can thus provide the information needed to distinguish between different model-dependent interpretations. Several theoretical studies have examined the 2D ES of molecular dimers [198, 186, 41, 107, 25], and the exciton-coupled spectra of multichromophore light harvesting complexes have been experimentally resolved and analyzed [157, 156, 169].

Because of its high information content, 2D ES presents previously undescribed possibilities to extract quantum information from molecular systems and to determine model Hamiltonian parameters [212]. For example, experiments by Hayes and Engel extracted such information for the FennaMatthewsOlsen light harvesting complex [89]. Recently, it was demonstrated by Brinks et al. that single molecule coherences can be prepared using phased optical pulses and detected using fluorescence [30]. The latter experiments exploit the inherent sensitivity of fluorescence and demonstrate the feasibility to control molecular quantum processes at the single molecule level. Fluorescence-based strategies to 2D ES, such as presented in the current work, could provide a route to extract high purity quan-

tum information from single molecules. It may also be a means to study molecular systems in the ultraviolet regime where background noise due to solvent-induced scattering limits ultrafast experiments.

Here we demonstrate a phase-modulation approach to 2D ES that sensitively detects fluorescence to resolve the exciton coupling in dimers of magnesium meso tetraphenylporphyrin (MgTPP), which are embedded in 1,2-distearoyl-sn-glycero-3- phosphocholine (DSPC) liposomal vesicles. MgTPP is a nonpolar molecule that preferentially enters the low dielectric amphiphilic regions of the phospholipid bilayer. At intermediate concentration, MgTPP forms dimers as evidenced by changes in the linear and 2D absorption spectra. Quantitative comparison between our measurements and simulated spectra for a broad distribution of selected conformations, screened by a global optimization procedure, shows that the information contained in linear spectra alone is not sufficient to determine a unique structure. In contrast, the additional information provided by 2D spectra constrains a narrow distribution of conformations, which are specified by the relative separation and orientations of the MgTPP macrocycles.

In our approach, called phase-modulation 2D fluorescence spectroscopy (PM-2D FS), a collinear sequence of four laser pulses is used to excite electronic population [190]. The ensuing nonlinear signal is detected by sweeping the relative phases of the excitation pulses at approximately kilohertz frequencies and by using lock-in amplification to monitor the spontaneous fluorescence. This technique enables phase-selective detection of fluorescence at sufficiently high frequencies to effectively reduce laboratory $1/f$ noise. Because the PM-2D FS observable depends on nonlinear populations that generate fluorescence, a different combination of nonlinear coherence terms must be considered than those of standard photon-echo 2D ES (referred to hereafter as 2D PE). In 2D PE experiments, the signala third-order polarization generated from three noncollinear laser pulsesis detected in transmission. The 2D PE signal depends on the superposition of well-known nonlinear absorption and emission processes, called ground-state bleach (GSB), stimulated emission (SE), and excited-state absorption (ESA) [138]. Analogous excitation pathways contribute to PM-2D FS. However, the relative signs and weights of contributing terms depend on the fluorescence quantum efficiencies of the excited-state populations. Equivalence between the two methods occurs only when all excited-state populations fluoresce with 100% efficiency [188]. Thus, self-quenching of doubly excited exciton population can give rise to differences between the spectra obtained from the two methodsdifferences that may depend, in themselves, on dimer conformation. For the conformations realized in the current study, we find that the PM-2D FS and 2D PE methods produce spectra with characteristic features distinctively different from one another.

6.2 Exciton-Coupled Dimer of Three-Level Molecules

Monomers of MgTPP have two equivalent perpendicular transition dipole moments contained within the plane of the macrocycle (see Fig. 6.1B, *Inset*). These define the directions of degenerate Q_x and Q_y transitions between the ground and lowest lying excited electronic states [75, 184, 112, 205]. Both transition moments contribute to the collective exciton interactions in a molecular complex, as illustrated in Fig. 6.1A. To specify dimer conformations, we adopt a molecular-frame coordinate system similar to that described in refs. [184] and [112]. For each monomer, a right-handed coordinate system is taken with the x and y axes lying parallel to the Q_x and Q_y transition directions, and the z axis perpendicular to the porphyrin plane. We adopt the convention that a conformation is specified by the monomer center-to-center vector $\mathbf{R}$, which is oriented relative to molecule 1 according to polar and azimuthal angles θ and ϕ. The relative orientation of molecule 2 is given by the Euler angles α and β. Because of the degeneracy of the Q_x and Q_y transitions, all of the results are independent of the third Euler angle, γ, which we set to zero from this point on [112].

For the Hamiltonian of a dimer of chemically identical three-level molecules in which system-bath effects are neglected, one defines the tensor product states $\{|ij\rangle\}$ where $i, j = g, x, y$ respectively label the states on monomer 1 and 2, and $\{|ij\rangle\}$ is the dimer Hilbert space basis. Notice x (y) is short-hand notation for the excited electronic state associated with the Q_x (Q_y) transition on each monomer.

Within this localized basis description, one can write the molecular Hamiltonian for the dimer

$$\hat{H} = \hat{H}^{(1)} + \hat{H}^{(2)} + \hat{V} = \hat{H}_0 + \hat{V}, \tag{6.1}$$

where $\hat{H}^{(1)}$ ($\hat{H}^{(2)}$ is the Hamiltonian associated with monomer 1 (monomer 2). Within the point-dipole approximation, the electronic coupling term can be expressed as

$$\hat{V} = \frac{1}{4\pi\varepsilon R^3}\hat{\boldsymbol{\mu}}_1 \cdot \left(1 - 3\frac{\mathbf{RR}}{R^2}\right) \cdot \hat{\boldsymbol{\mu}}_2, \tag{6.2}$$

with $\mathbf{R}$ the monomer center-to-center vector, $\hat{\boldsymbol{\mu}}_1$ ($\hat{\boldsymbol{\mu}}_2$) the dipole operator for monomer 1 (monomer 2), and ε the dielectric constant.

We simplify our notation by denoting the nine basis states $\{|l_i\rangle\}$, with $|l_1\rangle = |gg\rangle$, $|l_2\rangle = |xg\rangle$, $|l_3\rangle = |yg\rangle$, $|l_4\rangle = |gx\rangle$, $|l_5\rangle = |gy\rangle$, $|l_6\rangle = |xx\rangle$, $|l_7\rangle = |xy\rangle$, $|l_8\rangle = |yx\rangle$, $|l_9\rangle = |yy\rangle$. In this basis, the total Hamiltonian can be written as a nine-by-nine matrix of the form [112]:

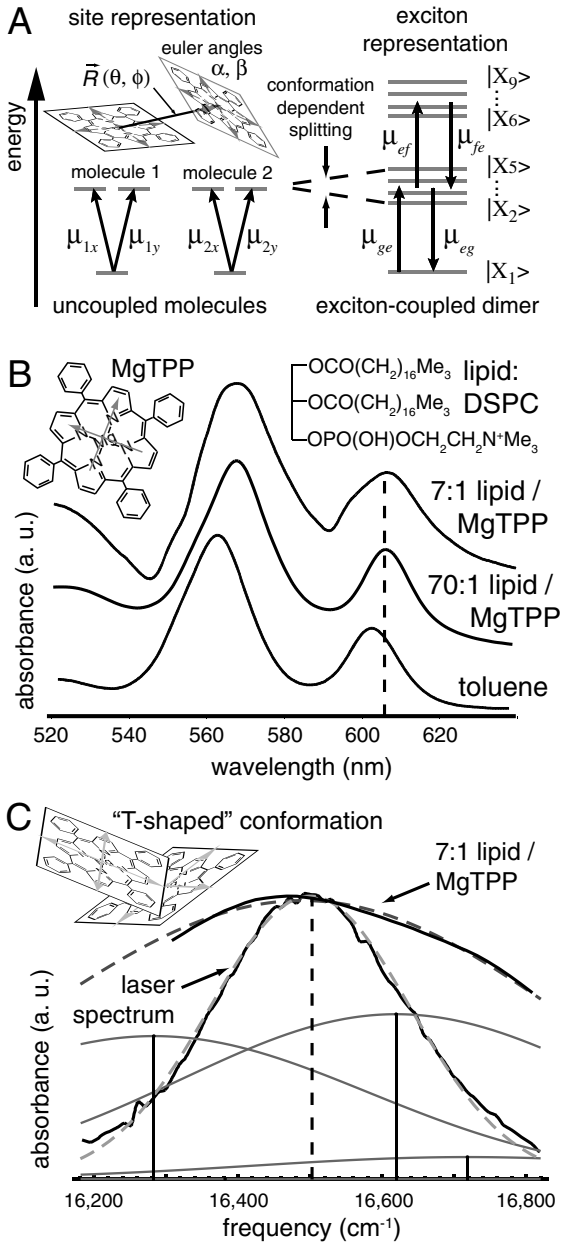

Figure 6.1: (A) Energy level diagram of two chemically identical three-level molecules, each with degenerate transition dipole moments directed along the x and y axes of the molecular frames. The inset shows a random configuration of two MgTPP monomers whose relative conformation is defined by the molecular center-to-center vector $\mathbf{R}$ and the angles θ, ϕ, α, and β. Electronic interactions results in an exciton-coupled nine-level system, with a single ground state, four non-degenerate singly-excited states, and four doubly-excited states. Multi-pulse excitation can excite transitions between ground, singly-excited, and doubly-excited state manifolds. (B) Absorption spectra of the MgTPP samples studied in this work. Spectra are vertically displaced for clarity. The samples correspond to MgTPP in toluene (bottom), aqueous liposome suspension with 70:1 DSPC:MgTPP (middle), and 7:1 DSPC:MgTPP (top). The dashed vertical line represents the lowest energy monomer transition energy used in our calculations. The insets show molecular formulas for MgTPP and lipid DSPC. (C) Overlay of the 7:1 DSPC:MgTPP absorbance and the laser pulse spectrum. The laser spectrum (solid black curve) has been fit to a Gaussian (dashed gray curve) with center frequency 15,501 cm^{-1} (606 nm), and FWHM = 327.0 cm^{-1} (12 nm). The linear absorbance (solid black curve) is compared to the simulated spectrum (dashed black curve), which is based on the T-shaped conformation shown in the inset. Also shown are the positions of the underlying exciton transitions (discussed in text).

$$\hat{H} \approx \begin{bmatrix} 0 & & & & & & & & & \\ & \epsilon_1 & V_{23} & V_{24} & V_{25} & & & & & \\ & V_{32} & \epsilon_1 & V_{34} & V_{35} & & & & & \\ & V_{42} & V_{43} & \epsilon_1 & V_{45} & & & & & \\ & V_{52} & V_{53} & V_{54} & \epsilon_1 & & & & & \\ & & & & & 2\epsilon_1 & & & & \\ & & & & & & 2\epsilon_1 & & & \\ & & & & & & & 2\epsilon_1 & & \\ & & & & & & & & 2\epsilon_1 & \end{bmatrix} \tag{6.3}$$

Here we have assumed all the diagonal contributions in the terms associated with $\hat{H}_0$, i.e., we have assumed that $\langle l_i|\hat{V}|l_i\rangle = 0$ for all l_i. To set the reference energy scale, we set $\epsilon_g^{(i)} = 0$ with $\hat{H}^{(i)}\,|g\rangle = \epsilon_g^{(i)}\,|g\rangle$, and therefore $\hat{H}_0\,|gg\rangle = (\epsilon_g^{(1)} + \epsilon_g^{(2)})\,|gg\rangle = 0\,|gg\rangle$. The value of ϵ_1 used in our simulation was 16,500.7 cm^{-1}, which corresponds to the monomer excitation energy associated with either of the degenerate Q_x or Q_y transitions for the 70:1 sample (see Fig. 6.1). Then $\hat{H}_0\,|l_k\rangle = \epsilon_k\,|l_k\rangle$ with $\epsilon_k = \epsilon_1$ for any of the states containing one excitation($k = 2-5$) and $\epsilon_k = 2\epsilon_1$ for the states containing two-excitations ($k = 6-9$). Diagonalization of the Hamiltonian is straightforward since it involves only the 4×4 block associated with the singly-excited state manifold. Note that the eigenenergies of the singly-excited state manifold correspond to the exciton transitions underlying the region of interest in the experimental and simulated linear spectra. The positions of these eigenenergies depend on the structural parameters of the dimer through the dependence on the couplings:

$$V_{ij} = \frac{1}{4\pi\varepsilon R^3}(\boldsymbol{\mu}_1)_{ij} \cdot \left(1 - 3\frac{\mathbf{RR}}{R^2}\right) \cdot (\boldsymbol{\mu}_2)_{ij} = \frac{\mu^2}{4\pi\varepsilon R^3}\kappa_{ij}^2. \tag{6.4}$$

Here the orientation factor κ_{ij}^2 is related to the directions of the transition dipole moments and this vector connecting their centers according to $\kappa_{ij}^2 = (\underline{\boldsymbol{\mu}}_1)_{ij} \cdot (\underline{\boldsymbol{\mu}}_2)_{ij} - 3[(\underline{\boldsymbol{\mu}}_1)_{ij} \cdot \underline{\mathbf{R}}\ \underline{\mathbf{R}}][\underline{\mathbf{R}} \cdot (\underline{\boldsymbol{\mu}}_2)_{ij}]$, where $\underline{\mathbf{R}} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$ is the monomer center-to-center unit vector, and $(\underline{\boldsymbol{\mu}}_n)_{ij} = \langle l_i|\hat{\boldsymbol{\mu}}_n|l_j\rangle/|\boldsymbol{\mu}|$ is the normalized transition dipole moment operator. The relationship between the square of the monomer transition dipole moment and its absorption coefficient α, is given by [78]:

$$|\boldsymbol{\mu}|^2 = \frac{3\varepsilon\hbar c}{\pi N_A}\int_{-\infty}^{\infty} d\bar{\nu}\frac{\alpha(\bar{\nu})}{\bar{\nu}}. \tag{6.5}$$

In Eq. 6.5, ε in the dielectric constant of the medium, $\hbar$ in Planck's constant divided by 2π, c is the speed of light, and N_A is Avogadro's number. The factor $\int_{-\infty}^{\infty} d\bar{\nu}\frac{\alpha(\bar{\nu})}{\bar{\nu}}$ is the optical linewidth of the $Q(0,0)$ transition, measured in wave numbers, and divided by its peak value. We estimated this number by numerical integration of the lineshape to be 44.3 $\mathrm{M}^{-1}\ \mathrm{cm}^{-1}$.

6.3 Comparison between PM-2D FS and 2D Photon Echo Spectroscopy (2D PE)

The PM-2D FS and 2D PE methods are conceptually similar, yet important distinguishing factors can result in their non-equivalence. The 2D PE signal can be interpreted as the third-order polarization of the sample, which is the source of the detected signal field. In contrast, PM-2D FS is a technique based on fluorescence-detection (2). The signal may be considered proportional to the fourth-order excited state population. We thus compare the signals of the two methods based on interpretation of 2D PE signals using third-order perturbation theory, and PM-2D FS signals using fourth-order perturbation theory.

We consider the semiclassical light-matter interaction Hamiltonian,

$$\hat{H}_{sc} = \hat{H}_0 + \hat{H}_{int}(t), \quad \hat{H}_{int}(t) = -\hat{\boldsymbol{\mu}} \cdot \mathbf{E}(t). \tag{6.6}$$

In PM-2D FS experiments, the electric field for P sequential collinear pulses polarized in the $\hat{\mathbf{x}}$ direction can be described by $\mathbf{E}(t) = \sum_j^P E_j(t)\hat{\mathbf{x}}$, where

$$E_j(t) = \lambda_j A_j(t - t_j)\cos(\omega_j(t - t_j) + \phi_j), \tag{6.7}$$

with λ_j the electric field maximum intensity, $A_j(t-t_j) = e^{-\frac{4\ln 2}{\tau_{\text{fwhm}}^2}(t-t_j)^2}$ the pulse envelope, and ω_j is the laser frequency for the j-th pulse. In 2D PE, the pulses are analogously described by $E_j(t) = \lambda_j A_j(t - t_j)\cos(\omega_j(t - t_j) - \mathbf{k}_j \cdot \mathbf{r})$.

Using density matrix formalism, the evolution of the system is described by the Liouville-von Neumann equation,

$$i\hbar\frac{\partial\hat{\rho}_I(t)}{\partial t} = [\hat{H}_{I,int}(t), \hat{\rho}_I(t)], \tag{6.8}$$

written in the interaction picture, where $\hat{\mathcal{O}}_I(t) \equiv e^{i\hat{H}_0(t-t_0)}\hat{\mathcal{O}}e^{-i\hat{H}_0(t-t_0)}$.

A formal and useful solution to Eq. 6.8 is,

$$\hat{\rho}_I(t) = \hat{\rho}_I(t_0) + \sum_{n=1}^{\infty}\hat{\rho}_I^{(n)}(t), \tag{6.9}$$

with,

$$\hat{\rho}_I^{(n)}(t) \equiv (-1)^n\left(\frac{i}{\hbar}\right)^n\int_{t_0}^{t}d\tau_n\int_{t_0}^{\tau_n}d\tau_{n-1}\cdots\int_{t_0}^{\tau_2}d\tau_1[\hat{H}_{I,int}(\tau_n),[\hat{H}_{I,int}(\tau_{n-1}),[\\ \cdots,[\hat{H}_{I,int}(\tau_1),\hat{\rho}_I(t_0)]\cdots]]] \tag{6.10}$$

The expectation value of any observable, $\langle\hat{\mathcal{O}}(t)\rangle \equiv \text{tr}\{\hat{\mathcal{O}}(t)\hat{\rho}(t)\} = \text{tr}\{\hat{\mathcal{O}}_I(t)\hat{\rho}_I(t)\}$ can be expressed as, $\langle\hat{\mathcal{O}}_I(t)\rangle = \sum_{n=0}^{\infty}\langle\hat{\mathcal{O}}_I^n(t)\rangle$, with, $\langle\hat{\mathcal{O}}_I^n(t)\rangle \equiv \text{tr}\{\hat{\mathcal{O}}_I(t)\hat{\rho}_I^{(n)}(t)\}$.

As mentioned above, the 2D PE signal is associated with the third-order polarization and therefore requires,

$$\mathbf{P}^{(3)}(t) \equiv \langle \hat{\boldsymbol{\mu}}^{(3)}(t) \rangle \equiv \mathrm{tr}\{\hat{\boldsymbol{\mu}}_I(t)\hat{\rho}_I^{(3)}(t)\} \tag{6.11}$$

while the PM-2D FS signal is associated with the fourth-order excited state population,

$$\hat{A}^{(4)}(t) = \langle \hat{A}(t) \rangle \equiv \mathrm{tr}\{\hat{A}(t)\hat{\rho}_I^{(4)}(t)\}, \tag{6.12}$$

with $A = \sum_\nu |\nu\rangle \langle\nu|$ the projector into all the states $\{|\nu\rangle\}$ of the excited state manifold.

We focus our discussion to the case of the nine-level model of the exciton-coupled dimer (see Fig. 6.1). 2D PE signals have been extensively derived and studied for this model [138, 40]. In Fig. 6.2, we show the double-sided Feynman diagrams (DSFD) contributing to the non-rephasing and rephasing signals, collected in the phase-matching direction $\mathbf{K}_I \equiv \mathbf{k}_1 - \mathbf{k}_2 + \mathbf{k}_3$ and $\mathbf{K}_{II} \equiv -\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3$, respectively. Neglecting dissipation for the moment, and assuming the rotating wave approximation in the impulsive limit [138], one obtains the following expressions for each one of the non-rephasing terms,

$$\mathbf{R}_{1a}^* \propto \sum_{e,e'} [\boldsymbol{\mu_{eg}\mu_{ge}\mu_{e'g}\mu_{ge'}}]_{\boldsymbol{e_1e_2e_3e_4}} e^{-i\omega_{eg}\tau} e^{-i\omega_{e'g}t} \tag{6.13}$$

$$\mathbf{R}_{2a} \propto \sum_{e,e'} [\boldsymbol{\mu_{eg}\mu_{ge'}\mu_{e'g}\mu_{ge}}]_{\boldsymbol{e_1e_2e_3e_4}} e^{-i\omega_{eg}\tau} e^{-i\omega_{ee'}T} e^{-i\omega_{eg}t} \tag{6.14}$$

$$\mathbf{R}_{3b}^* \propto \sum_{e,e',f} [\boldsymbol{\mu_{eg}\mu_{ge'}\mu_{fe}\mu_{e'f}}]_{\boldsymbol{e_1e_2e_3e_4}} e^{-i\omega_{eg}\tau} e^{-i\omega_{ee'}T} e^{-i\omega_{fe'}t} \tag{6.15}$$

Similarly, the rephasing terms are,

$$\mathbf{R}_{4a} \propto \sum_{e,e'} [\boldsymbol{\mu_{ge}\mu_{eg}\mu_{ge'}\mu_{ge'}}]_{\boldsymbol{e_1e_2e_3e_4}} e^{-i\omega_{ge}\tau} e^{-i\omega_{e'g}t} \tag{6.16}$$

$$\mathbf{R}_{3a} \propto \sum_{e,e'} [\boldsymbol{\mu_{ge}\mu_{e'g}\mu_{eg}\mu_{ge'}}]_{\boldsymbol{e_1e_2e_3e_4}} e^{-i\omega_{ge}\tau} e^{-i\omega_{e'e}T} e^{-i\omega_{e'g}t} \tag{6.17}$$

$$\mathbf{R}_{2b}^* \propto \sum_{e,e',f} [\boldsymbol{\mu_{ge}\mu_{e'g}\mu_{fe'}\mu_{ef}}]_{\boldsymbol{e_1e_2e_3e_4}} e^{-i\omega_{ge}\tau} e^{-i\omega_{e'e}T} e^{-i\omega_{fe}t} \tag{6.18}$$

Here, $e, e' \in \{X_2, X_3, X_4, X_5\}$ is the singly-excited manifold after diagonalization of the 4×4 block of the Hamiltonian in Eq. 6.3, and $f \in \{X_6, X_7, X_8, X_9\}$ is the doubly-excited state manifold, and $[\boldsymbol{\mu_{ab}\mu_{cd}\mu_{jk}\mu_{lm}}]_{\boldsymbol{e_1e_2e_3e_4}}$ denotes the three-dimensional orientational average product [40], $\langle (\boldsymbol{\mu_{ab}} \cdot \boldsymbol{e_1})(\boldsymbol{\mu_{cd}} \cdot \boldsymbol{e_2})(\boldsymbol{\mu_{jk}} \cdot \boldsymbol{e_3})(\boldsymbol{\mu_{lm}} \cdot \boldsymbol{e_4}) \rangle$, where $\boldsymbol{e_i}$ denotes the polarization of the ith pulse.

The detailed derivation of these expressions and their relation to the PM-2D FS terms will be published elsewhere [150]. In Fig. 6.2, we present the corresponding PM-2D FS non-rephasing and rephasing DSFDs obtained from the fourth-order perturbation expansion (Eq. 6.12). For our current purposes, we provide here the connection to the 2D PE expressions presented in Eqs. 6.13-6.18. For example, it can be shown that in the case of the non-rephasing contributions, the following relations between 2D PE and PM-2D FS hold: $\mathbf{R}^*_{1a} = \mathbf{Q}^*_{5a} \equiv \mathrm{GSB}_1$, $\mathbf{R}_{2a} = \mathbf{Q}_{2a} \equiv \mathrm{ESE}_1$, $\mathbf{R}^*_{3b} = \mathbf{Q}^*_{3b} \equiv \mathrm{ESA}_1$, and also $\mathbf{Q}^*_{3b} = \mathbf{Q}_{7b}$. For the rephasing signals we have: $\mathbf{R}_{4a} = \mathbf{Q}_{4a} \equiv \mathrm{GSB}_2$, $\mathbf{R}_{3a} = \mathbf{Q}_{3a} \equiv \mathrm{ESE}_2$, $\mathbf{R}^*_{2b} = \mathbf{Q}^*_{2b} \equiv \mathrm{ESA}_2$, and $\mathbf{Q}^*_{2b} = \mathbf{Q}^*_{8b}$.

Although most of the 2D PE and PM-2D FS contributions are equal, there are two key differences making their signals different:

1. Since PM-2D FS is a fluorescence detected technique, it is important to consider the nature of the resulting excited state of the the system after the interaction with the four ultrafast pulses. As a consequence, even though mathematically $\mathbf{Q}^*_{3b} = \mathbf{Q}_{7b}$ ($\mathbf{Q}^*_{2b} = \mathbf{Q}^*_{8b}$), they do not contribute equally, because the terms $\mathbf{Q}^*_{3b}$ ($\mathbf{Q}^*_{2b}$) end up in the singly-excited manifold $\{|e\rangle\}$ while $\mathbf{Q}_{7b}$ ($\mathbf{Q}^*_{8b}$) end up in the doubly-excited state $|f\rangle$. Since the quantum yield of singly- and doubly-excited states are different, we need to take into account this fact when simulating the signals. We introduced a multiplicative factor Γ in front of the diagrams ending in a doubly-excited population (see $\mathbf{Q}_{7b}$ and $\mathbf{Q}^*_{8b}$ in Fig. 6.2) to capture the relative quantum yield of this doubly-excited state compared to the singly-excited states. Due to the abundance of non-radiative relaxation pathways for highly excited states, one expects the relative quantum yield of the doubly-excited states to be significantly smaller than the singly-excited states. In a fully ideal coherent case, where two-photons are emitted via the pathway $|f\rangle \rightarrow |e\rangle \rightarrow |g\rangle$, then $\Gamma = 2$. In general, $0 \leq \Gamma \leq 2$. For the dimer studied in the current work, the value of $\Gamma = 0.31$ was obtained from the global optimization that compared simulated and experimental spectra. A visual illustration of these differences can be found in Fig. 6.5, where we compare for three different conformations PM-2D FS spectra ($\Gamma = 0.31$) to the corresponding 2D PE spectra ($\Gamma = 2$). Table 6.2 shows the sensitivity of the optimization target function to the parameter Γ around the optimal value of 0.31.

2. The GSB, SE and ESA terms add up differently for 2D PE and PM-2D FS. This is a consequence of the third-order versus fourth-order perturbation approach respectively. This is the main reason for the different appearances of PM-2D FS versus 2D PE spectra.

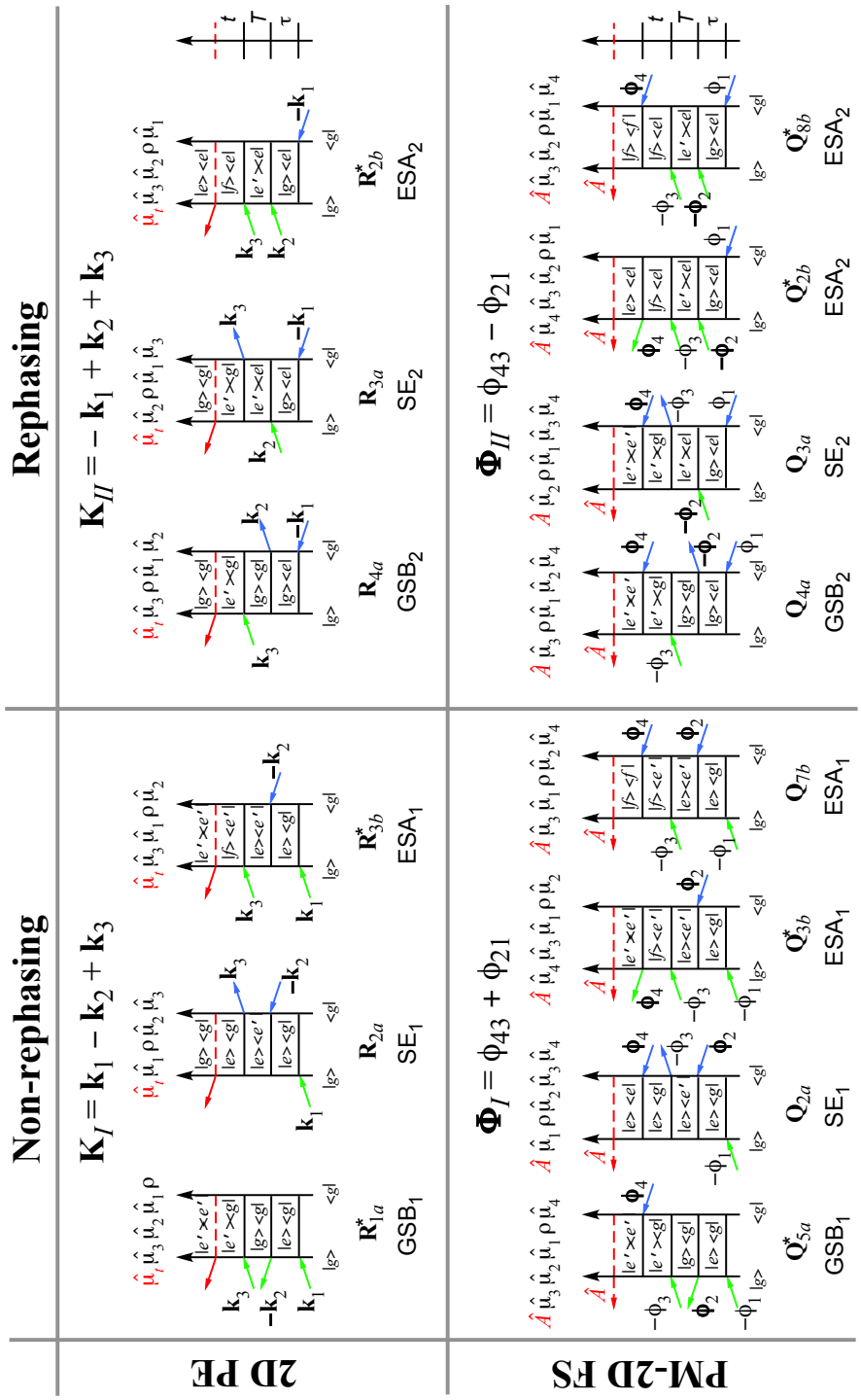

Figure 6.2: (Top) Four-level model used to describe the coupled dimers of MgTPP. The total dipole allows transitions from the ground state to the first-excited manifold and from the latter to the final doubly-excited state. (Bottom) Double-sided Feynman diagrams (DSFD) representing the different light-matter interactions contributing to the rephasing and non-rephasing signals measured experimentally. The sign associated with each diagram is determine by the number of arrows (dipole interactions) on the right vertical line of each ladder diagram ("bra" side). An even (odd) number of interaction pick up a plus (minus) sign for the term in consideration.

The non-rephasing and rephasing 2D PE signals are written:

$$\begin{aligned} S_{NRP}^{2DPE}(\tau,T,t) &\propto \mathbf{R}_{1a}^{*} + \mathbf{R}_{2a} - \mathbf{R}_{3b}^{*} \\ &\propto \mathrm{GSB}_1 + \mathrm{ESE}_1 - \mathrm{ESA}_1 \end{aligned} \tag{6.19}$$

$$\begin{aligned} S_{RP}^{2DPE}(\tau,T,t) &\propto \mathbf{R}_{4a} + \mathbf{R}_{3a} - \mathbf{R}_{2b}^{*} \\ &\propto \mathrm{GSB}_2 + \mathrm{ESE}_2 - \mathrm{ESA}_2 \end{aligned} \tag{6.20}$$

Taking account of the differences between the two methods mentioned above, and making use of Fig. 6.2, the non-rephasing and rephasing PM-2D FS signals are written:

$$\begin{aligned} S_{NRP}^{PM-2DFS}(\tau,T,t) &\propto -(\mathbf{Q}_{5a}^{*} + \mathbf{Q}_{2a} + \mathbf{Q}_{3b}^{*} - \Gamma\mathbf{Q}_{7b}) \\ &\propto -[\mathrm{GSB}_1 + \mathrm{ESE}_1 + (1-\Gamma)\mathrm{ESA}_1] \end{aligned} \tag{6.21}$$

$$\begin{aligned} S_{RP}^{PM-2DFS}(\tau,T,t) &\propto -(\mathbf{Q}_{4a} + \mathbf{Q}_{3a} + \mathbf{Q}_{2b}^{*} - \Gamma\mathbf{Q}_{8b}^{*}) \\ &\propto -[\mathrm{GSB}_2 + \mathrm{ESE}_2 + (1-\Gamma)\mathrm{ESA}_2] \end{aligned} \tag{6.22}$$

Although the signal expressions corresponding to the two techniques are closely related, the variable sign contribution of the ESA terms in the PM-2D FS expressions (Eqns. 6.21 and 6.22), in comparison to the well known negative sign ESA contribution in 2D PE spectroscopy (Eqns. 6.19 and 6.20), can lead to considerably different appearances of the 2D spectra. The differences in sign assignments of these terms arises from the commutator expansions of Eq. 6.10.

In the current work, we have considered the case where the population time $T = 0$ fs. To account for optical dephasing, inhomogeneous broadening and other dissipative processes, we multiplied each term given by Eqns. 6.19 - 6.22 by a phenomenological line broadening function which is assumed to be gaussian in both coherence times, τ and t. That is, the rephasing signals were multiplied by the factors $e^{-\tau^2/\sigma_{RP}^2}$ and e^{-t^2/κ_{RP}^2}. Similarly, we have used factors that contain the parameters σ_{NRP} and κ_{NRP} to describe the broadening of the non-rephasing signals. Fourier transformation of these equations to the ω_τ and ω_t domains provide the real, imaginary, and absolute value 2D spectra presented in Fig. 6.4, with very good agreement to experiment. We note that while the intensities and positions of 2D optical features are well accounted for by the molecular dimer Hamiltonian, the observed spectral lineshapes deviate markedly from this simple model. The asymmetric lineshapes could be due to a number of factors, including differences in the system-bath coupling and population times of the various excited states, as well as the effects of laser pulse overlap. Understanding the origins of the lineshape asymmetries is important to future studies.

6.4 Results and discussion

As discussed in Sec. 6.2, the effect of the interaction is to create an exciton-coupled nine-level system, with states labeled $|X_n\rangle$, comprised of a single ground state ($n = 1$), four singly-excited states ($n = 2 - 5$), and four doubly-excited states ($n = 6 - 9$). Transitions between states are mediated by the collective dipole moment, $\boldsymbol{\mu}_1 + \boldsymbol{\mu}_2$, which also depends on the structure of the complex.

In Fig. 6.1B are shown vertically displaced linear absorption spectra of MgTPP samples prepared in toluene, and 70:1 and 7:1 1,2-distearoyl-sn-glycero-3- phosphocholine (DSPC): MgTPP liposomes. For the 70:1 sample, the lineshape and position of the lowest energy $Q(0,0)$ feature, centered at 606 nm, underwent a slight red- shift relative to the toluene sample at 602 nm. For the elevated concentration 7:1 sample, the lineshape broadened, suggesting the presence of a dipole-dipole interaction and exciton splitting between closely associated monomer subunits.

In principle, it is possible to model the linear absorption spectrum in terms of the structural parameters $\mathbf{R}$, α, and β that determine the couplings V_{kl} and the collective dipole moments, and which ultimately determine the energies and intensities of the ground-state accessible transitions. To test the sensitivity of the linear absorption spectrum to different conformational models, we numerically generated approximately 1000 representative conformations and simulated their linear spectra (details provided in Appendix 6.8). By comparing experimental and simulated data, we established that a wide distribution of approximately 100 conformations can reasonably explain the linear absorption spectrum. Nevertheless, only a very small conformational sub-space could be found to agree with the experimental 2D spectra, and which is also consistent with the linear spectrum. In Fig. 6.1C is shown the simulated linear spectrum and the four underlying component transitions of the optimized "T-shaped" conformation. The linear spectrum corresponding to this conformation is composed of two intense spectral features at 16,283 cm^{-1} and 16,619 cm^{-1}, one weak feature at 16,718 cm^{-1}, and one effectively dark feature at 16,382 cm^{-1} (see Appendix 6.8). The relatively unrestrictive constraint imposed on dimer conformation by the linear spectrum is a consequence of the many possible arrangements and weights that can be assigned to the four overlapping Gaussian features with broad spectral width.

The PM-2D FS method uses four collinear laser pulses to resonantly excite electronic population, which depends on the overlap between the lowest energy electronic transition [the $Q(0,0)$ feature] and the laser pulse spectrum (as shown in Fig. 6.1C). We assigned the nonlinear coherence terms GSB, SE and ESA to time-ordered sequences of laser-induced transitions that produce population on the manifold of singly-excited states ($n = 2-5$) and the manifold of doubly excited states ($n = 6-9$). The theoretically derived expressions for

PM-2D FS were found to differ from those of 2D PE (details provided in Sec. 6.3). This is because ESA pathways that result in population on the doubly-excited states have a tendency to self-quench by, for example, exciton-exciton annihilation or other non-radiative relaxation pathways, so that these terms do not fully contribute to the PM-2D FS signal. In 2D PE experiments, signal contributions to ESA pathways interfere with opposite sign relative to the GSB and SE pathways, i.e. $S^{2DPE} = \mathrm{GSB} + \mathrm{SE} - \mathrm{ESA}$. In PM-2D FS experiments, quenching of doubly-excited state population leads to interference between GSB, SE and surviving ESA pathways with variable relative sign, i.e. $S^{PM-2DFS} = \mathrm{GSB} + \mathrm{SE} + (1 - \Gamma)\mathrm{ESA}$, where $0 \leq \Gamma \leq 2$ is the mean number of fluorescent photons emitted from doubly-excited states relative to the average number of photons emitted from singly excited states. In our analysis of PM-2D FS spectra (described below), we treated Γ as a fitting parameter to obtain the value that best describes our experimental data. As we show below, the difference between signal origins of the two methods can result in 2D spectra with markedly different appearances, depending on the specific dimer conformation.

In Fig. 6.3 are shown complex-valued experimental PM-2D FS data for the 7:1 lipid:MgTPP sample (top row), the 70:1 lipid:MgTPP (middle row), and the toluene sample (bottom row). Rephasing and non-rephasing data, shown respectively in panels A and B, were processed from independently detected signals according to their unique phase-matching conditions. The two types of spectra provide complementary structural information, since each depends on a different set of nonlinear coherence terms. Both rephasing and non-rephasing 2D spectra corresponding to the 7:1 liposome sample exhibit well resolved peaks and cross-peaks with apparent splitting 340 cm^{-1}. This is in contrast to the 2D spectra obtained from control measurements on the 70:1 liposome and toluene samples, which as expected exhibit only the isolated monomer feature due to the absence of electronic couplings in these samples. The 2D spectra of the 7:1 liposome sample are asymmetrically shaped, with the most prominent features a high energy diagonal peak and a coupling peak directly below it. We note that the general appearance of the 7:1 liposome PM-2D FS spectra is similar to previous model predictions for an exciton-coupled molecular dimer[198, 186, 41, 107, 109]. We next show that the information contained in these spectra can be used to identify a small sub-space of dimer conformations.

By extending the procedure to simulate linear spectra (described above), we numerically simulated 2D spectra for a broad distribution of conformations (see Appendix 6.8). We performed a least-square regression analysis that compared simulated and experimental spectra to obtain an optimized conformation consistent with both the 2D and the linear data sets. In our optimization procedure, we treated the fluorescence efficiency Γ of doubly-excited excitons as a parameter to find the value that best represents the experimental data. In Fig. 6.4, we directly compare our experimental and simulated PM-2D FS spectra for the

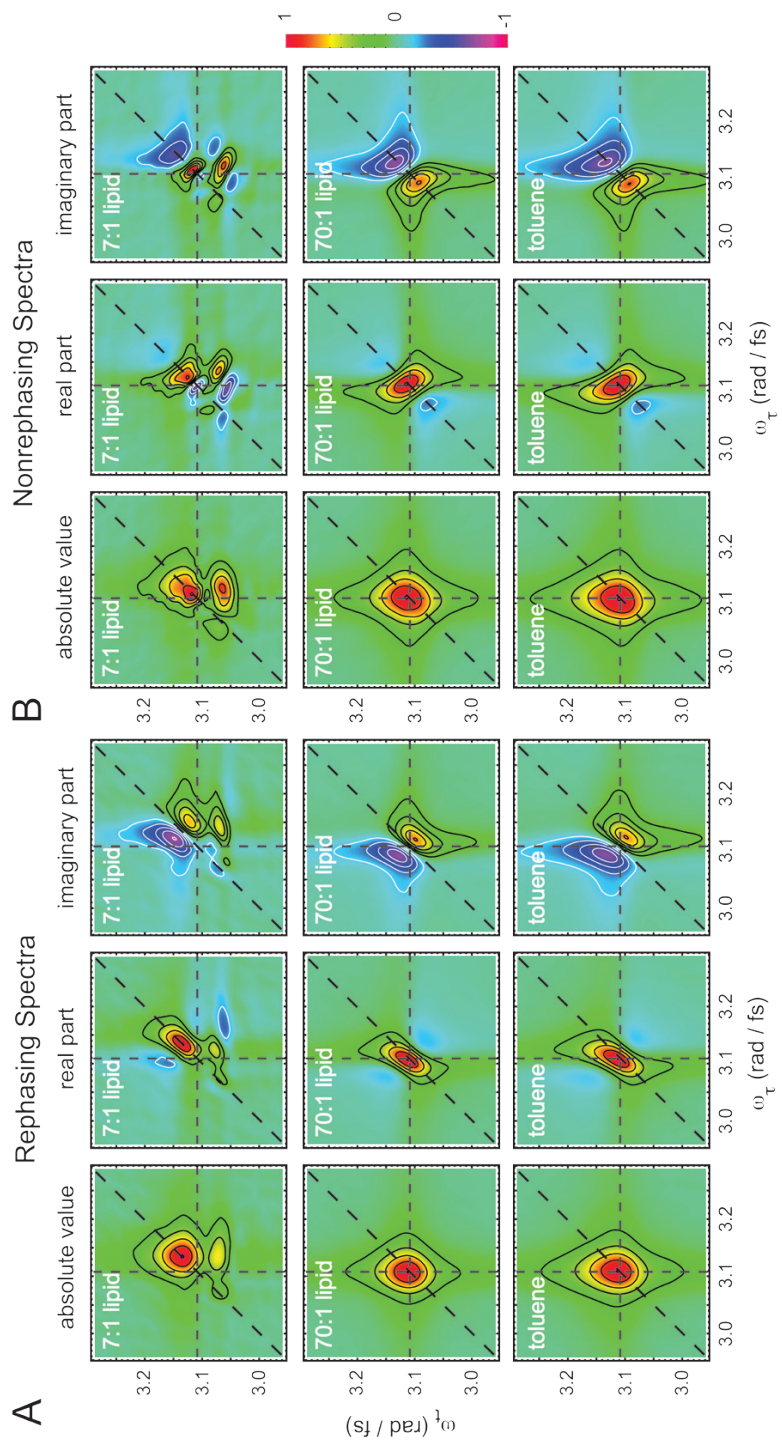

Figure 6.3: Comparison between rephasing (A) and non-rephasing (B) experimental 2D spectra corresponding to the MgTPP samples of Fig. 6.1B. Complex-valued spectra are represented as 2D contour plots, with absolute value (left column), real (middle column) and imaginary (right column) parts. The color scale of each plot is linear, and normalized to its maximum intensity feature. Positive and negative contours are shown in black and white, respectively, and are drawn at 0.8, 0.6, 0.4, 0.2, and 0.

optimized conformation. The values obtained for the parameters of this conformation are $\theta = 117.4°$, $\phi = 225.2°$, $\alpha = 135.2°$, $\beta = 137.2°$, $R = 4.2$ Å, and $\Gamma = 0.31$, with associated trust intervals: -16° $< \Delta\theta <$ 4°, $-11° < \Delta\phi < 11°$, $-11° < \Delta\alpha < 11°$, $-2° < \Delta\beta < 2°$, -0.05 Å$< \Delta R < 0.05$ Å,and $-0.1 < \Delta\Gamma < 0.1$ (details provided in Appendix 6.9). For both rephasing and non-rephasing spectra, the agreement between experiment and theory is very good, with an intense diagonal peak and a weaker coupling peak (below the diagonal) clearly reproduced in the simulation. A notable feature of the experimental 2D spectra is the asymmetric lineshape. A possible explanation for these asymmetries is the existence of distinct interactions between the various exciton states and the membrane environment. The discrepancy between experimental and simulated 2D lineshapes is an indication of a shortfall in the model Hamiltonian, which could be addressed in future experiments that focus on system-bath interactions.

In Fig. 6.5, we show the results of our calculations for three representative conformations. We compare simulated PM-2D FS spectra (with $\Gamma = 0.31$ optimized to the data, left column), 2D PE spectra (with $\Gamma = 2$, second column), and linear spectra (third column). It is evident that dimers with different conformations can produce very similar linear spectra. However, these same structures can be readily distinguished by the combined behaviors of both linear and 2D spectra. We note that for both PM-2D FS and 2D PE methods, the 2D spectrum depends on dimer conformation. However, we found that the qualitative appearance of simulated PM-2D FS spectra appear to vary over a greater range, and to exhibit a higher sensitivity to structural parameters in comparison to simulated 2D PE spectra.

Our confidence in the conformational assignment we have made is quantified by the numerical value of the regression analysis target parameter $\chi^2_{tot} = \chi^2_{linear} + \chi^2_{2D} = 7.39 + 9.87 = 17.26$, which includes contributions from both linear and 2D spectra. By starting with this conformation and incrementally scanning the structural parameters θ, ϕ, α, and β, we observed that χ^2_{tot} increased, indicating that the favored conformation is a local minimum when both linear and 2D spectra are included in the analysis (see Table 6.1). Similarly, we found that the value $\Gamma = 0.31$ corresponds to a local minimum (see Table 6.2). If only one of the two types of spectra is included, the restrictions placed on the dimer conformation are significantly relaxed. As shown in Fig. 6.5, conformations that depart from the optimized structure do not simultaneously produce 2D and linear spectra that agree well with experiment.

We found that the average conformation for the MgTPP dimer is a T-shaped structure with mean separation between Mg centers $R = 4.2$ Å. Close packing considerations alone would suggest the most stable structure should maximize $\pi - \pi$ stacking interactions. However, entropic contributions to the free energy due to fluctuations of the amphiphilic interior of the phospholipid bilayer must also be taken into account. It is possible that the average

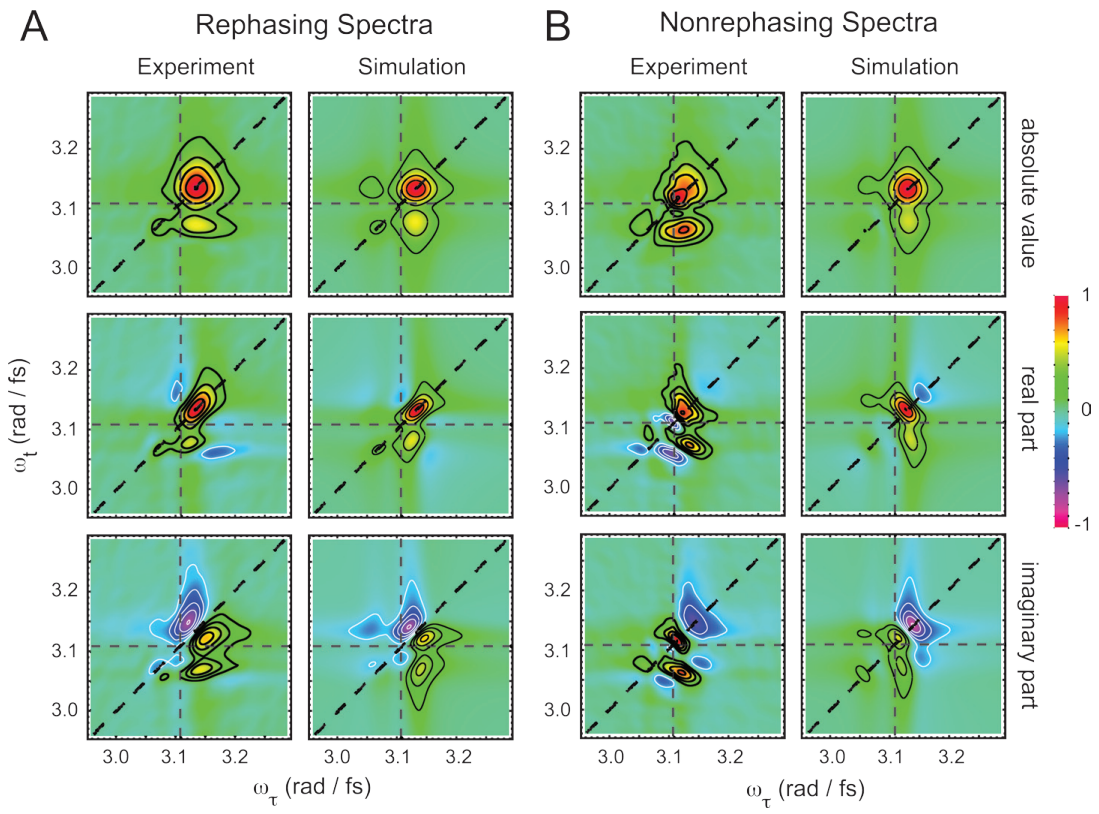

Figure 6.4: Comparison between rephasing (A) and non-rephasing (B) experimental (left columns) and simulated 2D spectra (right columns). Absolute value spectra (top), real part (middle) and imaginary part (bottom). The simulated spectra are based on the optimized T-shaped conformation depicted in Fig. 6.5 (top row, fourth column) and discussed in the text. Color scale and contours are the same as in Fig. 6.3.

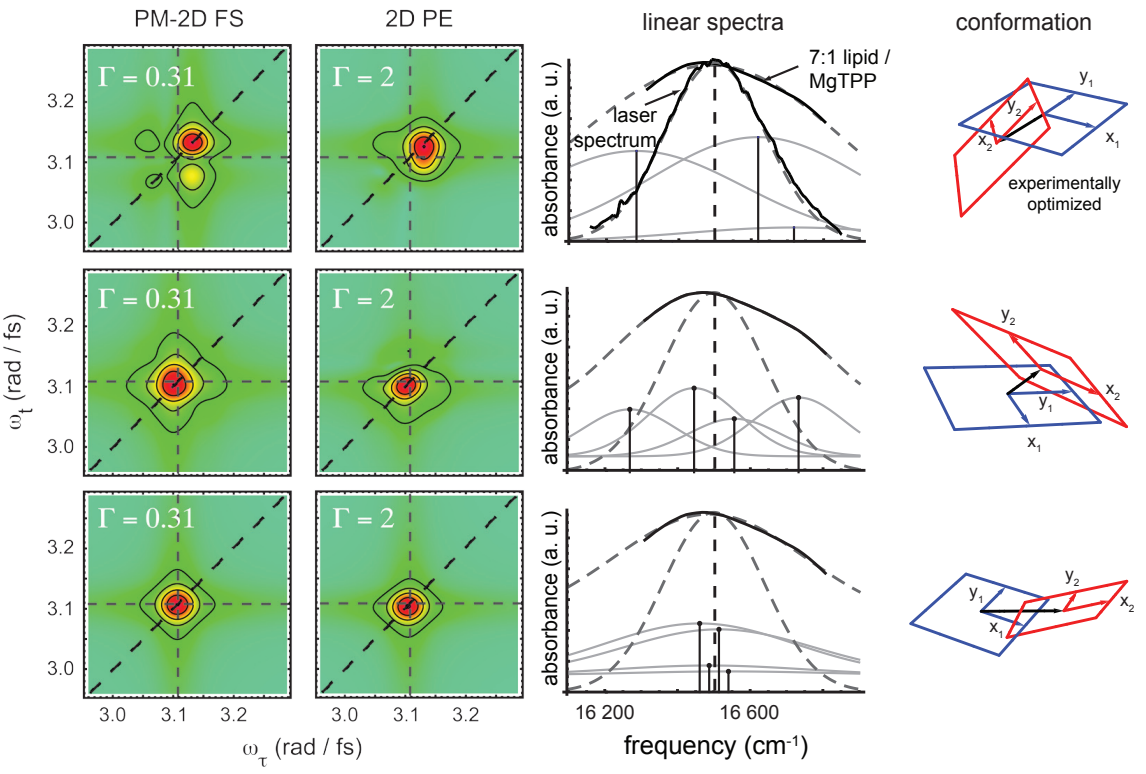

Figure 6.5: Comparison between simulated 2D and linear spectra for three selected dimer conformations. Each simulated linear spectrum (gray dashed curve) is compared to the experimental lineshape for the 7:1 DSPC:MgTPP sample. The laser spectrum is shown fit to a Gaussian (dashed gray curve) with center frequency 15,501 cm^{-1} (606 nm), and FWHM approximately 327.0 cm^{-1} (12 nm). Also shown are the positions of the underlying exciton transitions. Each of the three conformations produce a linear spectrum in agreement with experiment, while only the first (optimized) conformation produces simulated spectra that agree with PM-2D FS data (with $\Gamma = 0.31$). 2D PE spectra (with $\Gamma = 2$) are shown for comparison. Conformations are shown in the fourth column. The squares indicate the size of the MgTPP molecules, with monomer 1 in blue and monomer 2 in red with their respective Q_x and Q_y transition dipoles indicated. Top row: (optimized) conformation with $\theta = 117.4°$, $\phi = 225.2°$, $\alpha = 135.2°$, $\beta = 137.2°$, $R = 4.2$ Å. Middle row conformation with $\theta = 44.3°$, $\phi = 26.0°$, $\alpha = 29.2°$, $\beta = 138.6°$, $R = 3.7$ Å. Bottom row conformation with $\theta = 82.4°$, $\phi = 18.7°$, $\alpha = 47.9°$, $\beta = 124.0°$, $R = 7.6$ Å. Color scale and contours are the same as in Fig. 6.3.

conformation observed is the result of the system undergoing rapid exchange amongst a broad distribution of energetically equivalent structures. In such a dynamic situation, the significance of the observed conformation would be unclear. However, at room temperature the DSPC membrane is in its gel phase [213], and static disorder on molecular scales is expected to play a prominent role. It is possible that the observed dimer conformation - an anisotropic structure - is strongly influenced by the shapes and sizes of free volume pockets that form spontaneously inside the amphiphilic membrane domain. Future PM-2D FS experiments that probe the dependence of dimer conformation on temperature and membrane composition could address this issue directly.

6.5 Conclusions

We have shown that PM-2D FS can uniquely determine the conformation of a porphyrin dimer embedded in a non-crystalline membrane environment at room temperature. The appearance of the PM-2D FS spectra is generally very different from that produced by simulation of the 2D PE method. This effect is due to partial self- quenching of optical coherence terms that generate population on the manifold of doubly- excited states. In the current study on MgTPP chromophores in DSPC liposomes, we find that PM-2D FS spectra are quite sensitive to dimer conformation [157, 156, 169].

The PM-2D FS method might be widely applied to problems of biological and material significance. Spectroscopic studies of macromolecular conformation, based on exciton-coupled labels could be practically employed to extract detailed structural information. Experiments that combine PM-2D FS with circular dichroism should enable experiments that distinguish between enantiomers of chiral structures. PM-2D FS opens previously undescribed possibilities to study exciton-coupling under low light conditions, in part due to its high sensitivity. This feature may facilitate future 2D experiments on single molecules, or UV-absorbing chromophores.

6.6 Appendix 1. Liposome sample preparation

Samples were prepared according to the procedure described by MacMillan et al. **??**. MgTPP was purchased from Strem Chemicals (Boston), and used without further purification. 1.5 mg of MgTPP was dissolved in 20 mL of toluene, transferred to a 50 ml spherical flask, and the solvent was evaporated. In a separate flask, 12.8 mg of the phospholipid 1,2-distearoyl-sn-glycero-3-phosphocholine (DSPC, Sigma Aldrich) was dissolved in 20 mL of dichloromethane. The contents of the two flasks were combined to create a solution with 7:1 DSPC:MgTPP number ratio. The organic solvent was removed, and 30 ml of nanopure

water were added to the flask. The sample was alternately heated to 70°C and agitated by ultrasonication for a period of 15-30 minutes until an aqueous lipid / porphyrin emulsion was fully formed. The mixture was pre-filtered twice through glass wool, and then extruded through a 100-1000 nm pore nylon membrane (Avestin) to create a suspension of liposome vesicles. A second sample with 70:1 DSPC:MgTPP was prepared using the same procedure. It was confirmed using fluorescence microscopy that the MgTPP was localized to the membrane phase. An additional control sample was prepared by dissolving MgTPP in spectroscopic grade toluene.

6.7 Appendix 2. Phase-Modulation 2D Fluorescence Spectroscopy

The phase-modulation 2D fluorescence spectroscopy (PM-2D FS) method was described in detail elsewhere [190]. Samples were excited by a sequence of four collinear optical pulses with adjustable interpulse delays (see Fig. 6.6). The pulse sequence was produced using a high repetition regenerative amplifier (Coherent, RegA 9050, 250 kHz, pulse energy approximately 10 μJ), which was pumped by a Ti:sapphire seed oscillator (Coherent, Mira, 76 MHz, pulse energy approximately 9 nJ, pulse width approximately 35 fs) and a high power continuous wave ND:YVO4 laser (Coherent Verdi V-18, 532 nm). The amplified pulses were sent to two identical optical parametric amplifiers (Coherent, OPA 9400), with output pulse energies approximately 70 nJ. The relative phase of pulses 1 and 2, and pulses 3 and 4 were independently swept at distinct frequencies (5 kHz and 8 kHz, respectively) using acousto-optic Bragg cells. Electronic references were detected from the pulse pairs and sent to a waveform mixer to generate sum and difference sideband signals (13 kHz and 3 kHz, respectively). These reference waveforms were used to phase-synchronously detect the nonlinear fluorescence, which separately determined the nonrephasing and rephasing signals. The signal phase was calibrated to zero at the origin of the interferograms, i.e., when all interpulse delays were set to zero. The measured pulse spectrum at the sample was Gaussian with FWHM approximately 327 cm^{-1} (approximately 12 nm, shown in Fig. 1C). Separate dispersion compensation optics were used for each OPA, and the temporal pulse width determined by autocorrelation was approximately 60 fs for pulses 1 and 2, and approximately 80 fs for pulses 3 and 4. The sample cuvette was a flow cell (Starna Cells, 583.3/Q/3/Z15, path length 3 mm, 0.1-mL volume), which was fitted to a peristaltic pump (flow rate approximately 1 mL/min, approximately 6 mL reservoir volume). The excitation beam was focused into the sample using a 5-cm focal length lens. Fluorescence from the sample was collected using a 3-cm lens, spectrally filtered (620-nm long-pass, Omega Optical) and detected using an avalanche photo diode (Pacific Silicon Sensor). All mea-

surements were carried out at room temperature. The signals were measured as the delays between pulses 1 and 2 and between pulses 3 and 4 were independently scanned. Fourier transformation of the time-domain interferograms yielded the rephasing and nonrephasing 2D optical spectra.

6.8 Appendix 3. Computational modeling

The search for the porphyrin-dimer conformation consistent with both linear and 2D experimental data involved a constraint-nonlinear-global optimization with 13 variables. Optimizations performed separately on the linear and 2D spectra did not provide solutions consistent with both sets of experimental data. We therefore employed a joint target optimization function, which involved a least-square regression optimization using both sets of data – i.e., $\chi^2_{tot} = \chi^2_{linear} + \chi^2_{2D}$, which is described in the next section.

Construction of target function for linear spectra

The $Q(0,0)$ transition of the monomer in the lipid bilayer membrane has energy 16,500.7 cm^{-1} (see 70:1 lipid:MgTPP linear spectra shown in Fig. 6.1B). The $Q(0,0)$ feature contains contributions from both degenerate Q_x and Q_y transitions. Formation of the electronically coupled dimer results in four new transitions, which arise from the couplings between the states on each monomer. The energies of the resulting exciton transitions are given by the eigenvalues obtained from diagonalization of the 4×4 block of the Hamiltonian matrix (Eq. 6.3). The relative intensities of the exciton transitions are computed from the eigenvectors, which determine the transition dipole moments [112]. All of the transitions are broadened and modeled as Gaussians centered at their respective eigenvalues, with equal line widths σ_{lin}. The value of σ_{lin} was treated as an optimization parameter. The trial function used to reproduce the linear spectra can be written:

$$\text{trial}_{lin}(\theta, \phi, \alpha, \beta, R, a_0, \eta, \sigma_{lin}) = a_0 + \eta \left(\sum_{i=1}^{4} a_i(\theta, \phi, \alpha, \beta, R) e^{-[\bar{\nu} - \bar{\nu}_i(\theta,\phi,\alpha,\beta,R)]^2/\sigma^2_{lin}} \right) \quad (6.23)$$

In Eq. 6.23, a_0 accounts for background absorption, η is a multiplicative factor that uniformly adjusts the intensities a_i, and $\bar{\nu}_i$ are the eigen-energies of the transitions. All of the optimization parameters are determined by a least-square regression analysis when compared to experimental data. We isolated the experimental data inside the region-of-interest frequency window 16,300 cm^{-1}-16,810 cm^{-1}, which is centered around the uncoupled monomer transition energy ($\epsilon_1 = 16,500.7$ cm^{-1}). We denote the least-square sum as target_{lin}, and the contribution to the total optimization function is defined as $\chi^2_{lin} = 10^5$ target_{lin}. For example, the value of χ^2_{lin} corresponding to the best fit to both linear and

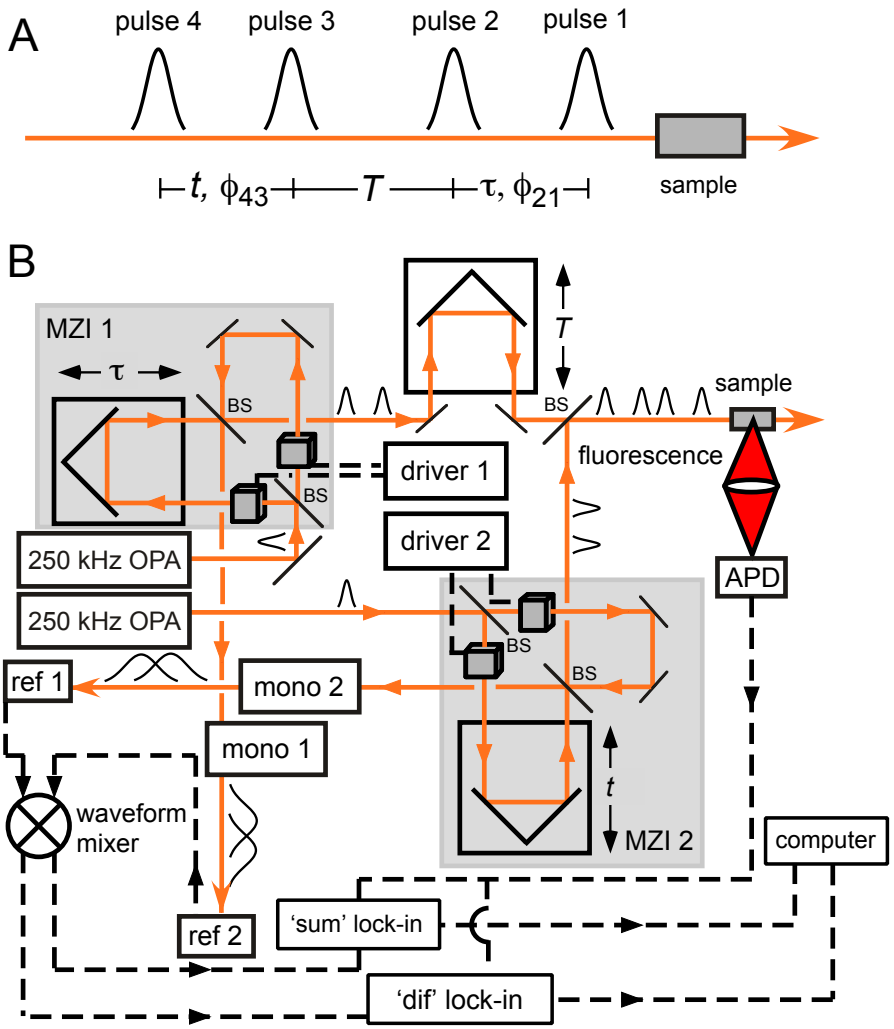

Figure 6.6: (A) Collinear sequence of optical pulses used in PM-2D FS experiments. The coherence, population, and measurement periods (τ, T, and t) are indicated, as well as the relative phase of pulses 1 and 2 (ϕ_{21}), and pulses 3 and 4 (ϕ_{43}). (B) Schematic of the PM-2D FS apparatus, described in the text and in [190]. The phases of the pulse electric fields are swept using acouto-optic Bragg cells, which are placed in the arms of two Mach-Zehnder interferometers (MZI 1 and MZI 2). The excitation pulses are made to be collinear before entering the sample. Reference waveforms are constructed from the pulse pairs from each interferometer. The reference signals oscillate at the difference frequencies of the acousto-optic Bragg cells (5 kHz and 8 kHz for ref 1 and ref 2, respectively). The reference signals are sent to a waveform mixer to construct sum and difference side band signals (3 kHz and 13 kHz). These reference side bands are used to phase-synchronously detect the fluorescence, which isolates the non-rephasing and rephasing population terms, respectively.

2D spectra is 7.39. The values of the eigen- energies for the optimized conformation are $\bar{\nu}_1 = 16,283$ cm^{-1}, $\bar{\nu}_2 = 16,382$ cm^{-1}, $\bar{\nu}_3 = 16,619$ cm^{-1}, and $\bar{\nu}_4 = 16,718$ cm^{-1}, with respective relative intensities $a_1 = 0.867$, $a_2 = 1.94 \times 10^{-13}$, $a_3 = 1.00$, and $a_4 = 0.133$.

Construction of the target function for the 2D spectra

The simulations of the 2D spectra involves the five geometrical parameters θ, ϕ, α, β, and R; the line-broadening parameters σ_{RP}, σ_{NRP}, κ_{RP}, and κ_{NRP} discussed above; and the doubly-excited state manifold fluorescence efficiency parameter Γ. For the least-square analysis of 2D spectra we used the experimental data ?in the frequency window $\omega_\tau \in [3.04$ rad fs^{-1}, 3.15 rad fs$^{-1}]$ and $\omega_t \in [3.04$ rad fs^{-1}, 3.15 rad fs$^{-1}]$, where the most intense diagonal peaks and cross-peaks were located. The least-square sum χ^2_{2D} includes the six sets of 2D experimental data, i.e., the real, imaginary and absolute value spectra for rephasing and non-rephasing signals. For example, the value of χ^2_{2D} for the best fit to both 2D linear and 2D spectra is 9.87.

Importance of the combined target function

Finding a single conformation that agrees well with the linear and 2D data proved to be a restrictive task, suggesting a definitive structural determination. For example, the optimization of either χ^2_{lin} or χ^2_{2D} by themselves did not result in solutions that were consistent with the other type of spectra. A single solution was only possible when the combined target function $\chi^2_{tot} = \chi^2_{lin} + \chi^2_{2D}$ was used. As shown in Fig. 6.5, it was possible to find examples for which χ^2_{lin} was smaller than the value obtained for the optimal conformation. Yet in these cases the 2D spectra departed significantly from the experimental data. Similarly, the optimization of only the target function χ^2_{2D} could lead to misleading results. In Table 6.1, we list values for the target function and its linear and 2D components for several values of the structural angles, which were scanned relative to the optimized conformation. We note that Table 6.1 contains some negative values for either χ^2_{lin} or χ^2_{2D}, indicating that a departure from the χ^2_{tot} minimum can yield improved agreement with one type of spectra at the expense of agreement with the other. The results presented in Table 6.1 suggests that the sensitivity of the search to structural parameters allows for a quantitative estimate of dimer conformation.

Table 6.1: Linear least-square target function $\chi^2_{tot} = \chi^2_{lin} + \chi^2_{2D}$ dependence on structural angles. Target function values are given relative to the reference values: $\chi^2_{lin} = 7.39$, $\chi^2_{2D} = 9.87$, and $\chi^2_{tot} = 17.26$, which correspond to the conformation with structural parameters $\theta = 117.4°$, $\phi = 225.2°$, $\alpha = 135.2°$, $\beta = 137.2°$, $R = 4.2\mathring{A}$, and $\Gamma = 0.31$, and line-broadening parameters $\sigma_{RP} = 108.1$ fs, $\sigma_{NRP} = 96.2$ fs, $\kappa_{RP} = 98.1$ fs, and $\kappa_{NRP} = 102.9$ fs.

	$\Delta\theta$			$\Delta\phi$			$\Delta\alpha$			$\Delta\beta$		
deg	$\Delta\chi^2_{lin}$	$\Delta\chi^2_{2D}$	$\Delta\chi^2_{tot}$	$\Delta\chi^2_{lin}$	$\Delta\chi^2_{2D}$	$\Delta\chi^2_{tot}$	$\Delta\chi^2_{lin}$	$\Delta\chi^2_{2D}$	$\Delta\chi^2_{tot}$	$\Delta\chi^2_{lin}$	$\Delta\chi^2_{2D}$	$\Delta\chi^2_{tot}$
-30	8.124	5.447	13.57	3.580	0.807	4.388	3.580	0.807	4.388	141.2	-0.353	140.9
-25	2.31	2.14	4.45	1.93	0.50	2.429	1.934	0.495	2.429	101.2	-0.340	100.9
-20	0.327	0.715	1.041	0.965	0.282	1.248	0.965	0.282	1.248	64.68	-0.306	64.38
-15	-0.018	0.150	0.132	1.077	0.144	1.221	1.077	0.144	1.221	34.84	-0.255	34.58
-10	0.080	-0.057	0.022	0.606	0.059	0.665	0.606	0.059	0.665	14.05	-0.190	13.86
-5	0.381	-0.100	0.285	0.557	0.014	0.571	0.557	0.014	0.571	3.031	-0.107	2.924
0	0	0	0	0	0	0	0	0	0	0	0	0
5	0.382	0.317	0.698	0.557	0.014	0.571	0.557	0.014	0.571	1.034	0.147	1.182
10	1.333	1.145	2.478	0.606	0.059	0.665	0.606	0.059	0.665	3.220	0.359	3.580
15	3.600	3.172	6.772	1.077	0.144	1.221	1.077	0.144	1.221	5.531	0.680	6.210
20	11.91	7.672	19.58	0.965	0.282	1.248	0.965	0.282	1.248	7.352	1.187	8.540
25	27.24	16.55	43.79	1.933	0.495	2.429	1.934	0.495	2.430	7.875	2.007	9.882
30	52.35	32.65	85.00	3.580	0.807	4.388	3.580	0.807	4.388	7.00	3.325	10.33

Table 6.2: Linear least-square target function $\Delta\chi^2_{2D}$ dependence on fluorescence efficiency Γ of the doubly-excited state manifold. Values are given relative to the optimized conformation with $\Delta\chi^2_{2D} = 9.87$ and $\Gamma = 0.31$.

Γ	0	0.2	0.31	0.4	0.6	0.8	1.0	1.2	1.4	1.6	1.8	2.0
$\Delta\chi^2_{2D}$	0.86	0.10	0	0.08	0.81	2.32	4.58	7.52	10.87	13.97	17.25	21.20

6.9 Appendix 4. Error Analysis and Propagation of Uncertainties in PM-2D FS Signals

In this section we calculate trust intervals for the structural parameter values we obtain for the MgTPP dimers embedded in DSPC liposomes. We discuss here the uncertainties in our results, which arise from two different sources: 1) the quality of the optimization search performed with the KNITRO package, and 2) the uncertainty in the reference experimental data used to construct the target function χ^2_{tot}, from which the values of the structural parameters are obtained.

To determine the quality of the KNITRO search, e.g., the absence of convergence to local minima, we performed a fine-resolution parameter scan (see Fig. 6.7) to verify the extent to which the values obtained by the program indeed correspond to global minimum, i.e., the best from the multi-start search. In Fig. 6.7 we plot the relative deviation $\Delta\chi^2_{tot}/\chi^2_{tot,ref} = (\chi^2 - \chi^2_{tot,ref})/\chi^2_{tot,ref}$ from the reference value of $\chi^2_{tot,ref}$, which can be interpreted as a relative error when moving away from the optimal conformation. Fig. 6.7 clearly shows that the structure found is the minimum, at least within ± 1 degree for the each of the angles, $\pm$ 0.05 Åfor the R distance, and $\pm$ 0.01 units in Γ. The few missing points in the scans for α and ϕ were removed because these converged to a higher local minima above the predominant-branch where the majority of points appear to lie. For all of the scans, one parameter was varied while the remaining parameters that entered the calculation of the 2D spectra were held constant. The lack of convergence we refer to here is due to the additional optimization required to relax the parameters needed for the linear spectra ($a_0, \{a_i\}, \eta$, and σ_{lin} in Eq. 6.25). Since the few data points that converged above the predominant-branch do not suggest an alternative minimum, it was not necessary to converge these points since enough were present to clearly show the behavior upon approaching the minimum.

The scans in Fig. 6.7 also serve to assess the degree of sensitivity. For example, it is clear that the scans are more sensitive to the parameters β, R, and θ, when compared to

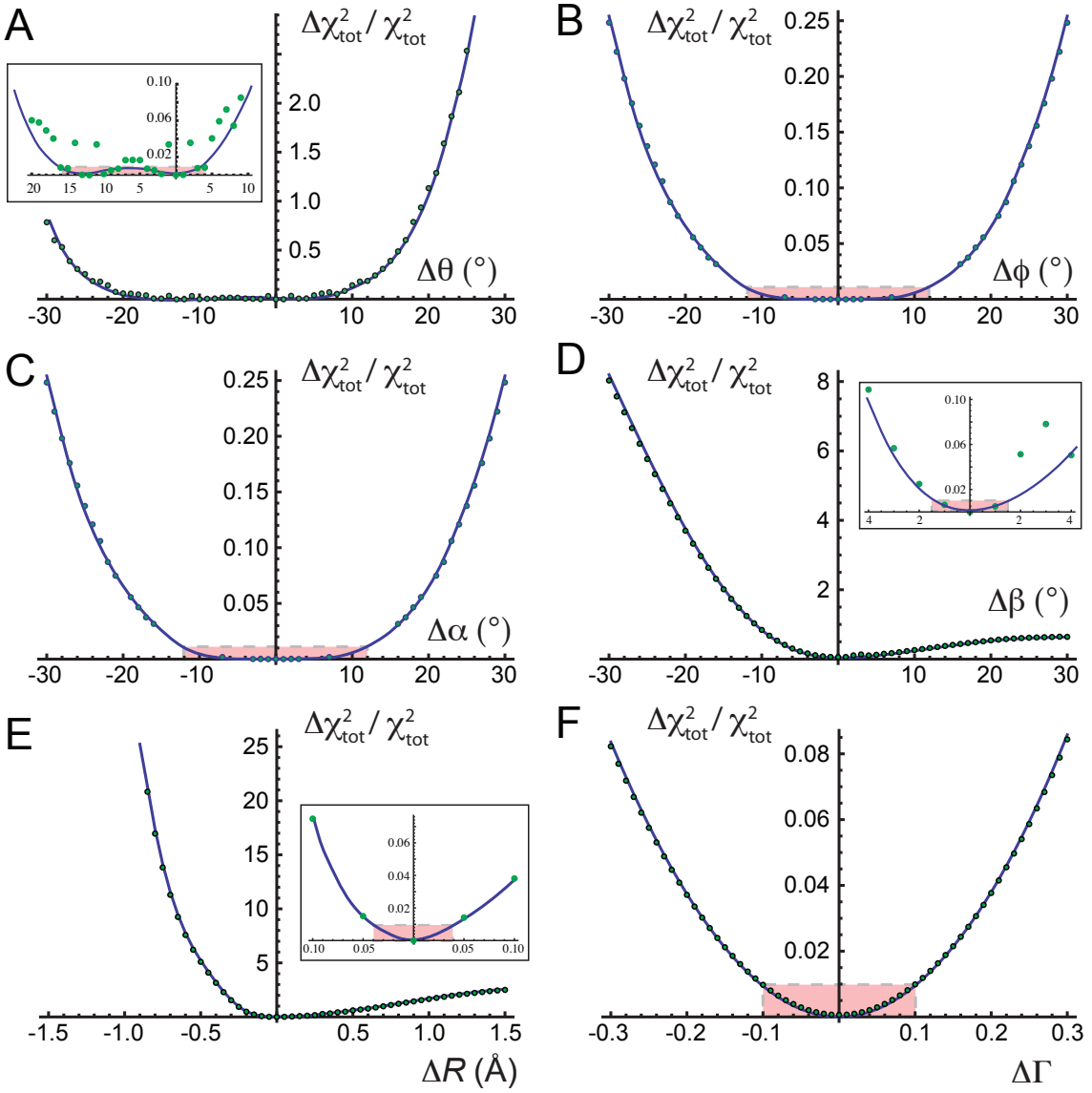

Figure 6.7: Relative deviation of the target function, $\Delta\chi^2_{tot}/\chi^2_{tot}$, from the optimized reference value, $\chi^2_{tot,ref}$, as a function of structural parameter uncertainties. Cross-sections of the target function are shown for the uncertainties (A) $\Delta\theta$, (B) $\Delta\phi$, (C) $\Delta\alpha$, (D) $\Delta\beta$, (E) ΔR, and (F) $\Delta\Gamma$, where $\Delta x = x - x_{ref}$, and x_{ref} is the value corresponding to the optimized conformation. The optimized conformation corresponds to a minimum of the multi-dimensional parameter surface. As indicated by the red shaded rectangles, trust intervals are directly read out from these plots, based on the $\sim 1\%$ relative error associated with the experimental data quality. The trust interval regions are expanded and shown as insets for the parameters $\Delta\theta$, $\Delta\beta$, and ΔR. The resulting intervals are $-16° < \Delta\theta < 4°$, $-11° < \Delta\phi < 11°$, $-11° < \Delta\alpha < 11°$, $-2° < \Delta\beta < 2°$, -0.05 Å$< \Delta R < 0.05$ Å, and $-0.1 < \Delta\Gamma < 0.1$.

other degrees of freedom such as α, ϕ, and Γ. As a consequence, under a certain fixed relative error, one expects that the uncertainty will be smaller for β and θ while slightly larger for α and ϕ.

Having established that our search routine is almost exact, we next address the error propagation due to uncertainties in the experimental measurements. In the following, we base our discussion on χ^2_{2D}, motivated by the assumption that $\Delta\chi^2_{tot}/\chi^2_{tot,ref} \approx \Delta\chi^2_{2D}/\chi^2_{2D,ref}$, i.e., that these relative errors are comparable. We thus use our estimate of $\Delta\chi^2_{2D}/\chi^2_{2D,ref}$, to read out the trust intervals directly from the scans shown in Fig. 6.7. This relative error was estimated to be approximately 1%, and it is indicated separately for each structural parameter by the red-shaded rectangles in Fig. 6.7.

We next explain the assumptions we have made to obtain the 1% estimate using standard error propagation analysis [189]. The 2D target function is defined according to

$$\begin{aligned}\chi^2_{2D} = \sum_{\omega^i_\tau,\omega^j_t} \quad & (\mathrm{AbsNRP_{sim}}(\omega^i_\tau,\omega^j_t) - \mathrm{AbsNRP_{exp}}(\omega^i_\tau,\omega^j_t))^2 + (\mathrm{ReNRP_{sim}}(\omega^i_\tau,\omega^j_t) - \\ & \mathrm{ReNRP_{exp}}(\omega^i_\tau,\omega^j_t))^2 + (\mathrm{ImNRP_{sim}}(\omega^i_\tau,\omega^j_t) - \mathrm{ImNRP_{exp}}(\omega^i_\tau,\omega^j_t))^2 + \\ & (\mathrm{AbsRP_{sim}}(\omega^i_\tau,\omega^j_t) - \mathrm{AbsRP_{exp}}(\omega^i_\tau,\omega^j_t))^2 + (\mathrm{ReRP_{sim}}(\omega^i_\tau,\omega^j_t) - \\ & \mathrm{ReRP_{exp}}(\omega^i_\tau,\omega^j_t))^2 + (\mathrm{ImRP_{sim}}(\omega^i_\tau,\omega^j_t) - \mathrm{ImRP_{exp}}(\omega^i_\tau,\omega^j_t))^2 \end{aligned} \tag{6.24}$$

In Eq. 6.24, the subscripts "sim" and "exp" indicate simulated and experimental spectra, respectively. The indices "i" and "j" indicate the 2D frequency coordinate. For the error propagation analysis, we include every data point from each of the six Fourier-transformed experimental signals ($\mathrm{AbsNRP_{exp}}, \mathrm{AbsRP_{exp}}, \mathrm{ReNRP_{exp}}, \mathrm{ReRP_{exp}}, \mathrm{ImNRP_{exp}}$, and $\mathrm{ImRP_{exp}}$) to define a variable with its own uncertainty. For simplicity, we define
$\mathrm{AbsNRP_{exp}}(\omega^i_\tau,\omega^j_t) \equiv f^{ij}_1$, $\mathrm{ReNRP_{exp}}(\omega^i_\tau,\omega^j_t) \equiv f^{ij}_2$, $\mathrm{ImNRP_{exp}}(\omega^i_\tau,\omega^j_t) \equiv f^{ij}_3$,
$\mathrm{AbsRP_{exp}}(\omega^i_\tau,\omega^j_t) \equiv f^{ij}_4$, $\mathrm{ReRP_{exp}}(\omega^i_\tau,\omega^j_t) \equiv f^{ij}_5$, and $\mathrm{ImRP_{exp}}(\omega^j_t,\omega^j_t) \equiv \omega_t f^{ij}_6$.

The sum in Eq. 6.24 is performed over the discrete frequency values inside the interval $\omega_\tau, \omega_t \in (3.04, 3.15)$ rad/fs. Since there are $N = 101$ points per frequency axis inside this interval, the number of terms in the summation contains $N^2 = 10,201$ variables of the form f^{ij}_k, for each k. Since we are dealing with $k = 1-6$, he number of independent variables in the error propagation analysis is 61,206. We define $z \equiv \chi^2_{2D}(\{f^{ij}_k\}) = \chi^2_{2D}(\{g_n\})$, where $g_n \equiv f^{ij}_k$, with n running from 1-61,206, denoting all the possible combinations of i, j, and k. Under the assumption of independent variables we can estimate the uncertainty of z by [189]

$$\Delta z \approx \sqrt{\sum_{n=1}^{61,206} \left(\frac{\partial z}{\partial g_n}\Delta g_n\right)^2}. \tag{6.25}$$

In terms of the g_n variables, Eq. 6.24 for χ^2_{2D} can be rewritten,

$$z = \sum_{n=1}^{61,206} (g_n^{\text{sim}} - g_n)^2 \tag{6.26}$$

The partial derivative can be calculated according to $\frac{\partial z}{\partial g_n} = -2(g_n^{\text{sim}} - g_n)$. Once the uncertainties Δg_n are calculated, the error in Eq. 6.25 can be easily calculated.

As previously stated, each of the g_n corresponds to a data point from any of the 2D spectra involved in the calculation of χ^2_{2D}. To estimate the uncertainty associated with each of 2D the 61,206 variables, we divide them into two groups; the first half ($n = 1-30,603$) associated with the absolute value, real and imaginary parts of the rephasing data, and the remaining half ($n = 30,604-61,206$) associated with that of the non-rephasing data. To simplify these calculations, we find a single uncertainty value representative for each of the two types of spectra. We denote these as Δg_{RP} and Δg_{NRP} for the rephasing and nonrephasing data respectively. Calculations of these uncertainties are illustrated in Fig. 6.8. The uncertainty is estimated from four different experimental runs performed on a ZnTPP monomer in dimethylformamide solution, which were processed using an identical procedure to the MgTPP samples studied here. The 2D absolute value rephasing and non-rephasing spectra of one data run are shown in Figs. 6.8A and 6.8B, respectively. In Figs. 6.8C and 6.8D are shown overlays of the absolute value rephasing and non-rephasing signals, $s_\omega^{\text{RP(NRP)}}$, for each of the four data runs along the diagonal profile, with $\omega_\tau = \omega_\tau = \omega$. Figs. 6.8E and 6.8F show the average signal $\bar{s}_\omega^{\text{RP(NRP)}} \equiv \langle s_\omega^{\text{RP(NRP)}} \rangle$, along a diagonal profile, where $\langle \cdots \rangle_{sets}$ ndicates the average performed over individual data sets. We similarly calculate the variance at each value of ω according to $\sigma^2_{RP(NRP)}(\omega) = \langle (s_\omega^{\text{RP(NRP)}} - \bar{s}_\omega^{\text{RP(NRP)}})^2 \rangle_{\text{sets}}$, which are shown in Figs. 6.8G and 6.8H.

The representative uncertainties, Δg_{RP} and Δg_{NRP}, are estimated as the frequency average of the standard deviations along the diagonal profiles, i.e., $\Delta g_{\text{RP(NRP)}} = \langle \sigma_{RP(NRP)}(\omega) \rangle_\omega$. The average over frequency was done to include most of the significant data, taking approximately twice the full-width at half-maximum from the main peak for both the rephasing and nonrephasing profile – i.e., over the interval $\omega \in (3.07, 3.20)$ rad fs^{-1}. By using the resulting values for $\Delta g_{\text{RP}} = 0.0086$ and $\Delta g_{\text{NRP}} = 0.016$ in Eq. 6.25, we find that $\Delta z / z_{ref} = \Delta\chi^2_{2D}/\chi^2_{2D} \approx \Delta\chi^2_{tot}/\chi^2_{tot} = 0.0096 \sim 1\%$. The value of $\chi^2_{2D} = 9,87$ used for this Having established that the expected error is $\sim 1\%$, we determine the trust intervals directly from the parameter scan plots shown in Fig. 6.7, as indicated by the red-shaded rectangles. These intervals correspond to $-16^o < \Delta\theta < 4^o$, $-11^o < \Delta\phi < 11^o$, $-11^o < \Delta\alpha < 11^o$, $-1^o < \Delta\beta < 1^o$, $-0.05\mathring{A} < \Delta R < 0.05\mathring{A}$, and $-0.1 < \Delta\Gamma < 0.1$, with $\Delta x \equiv x - x_{ref}$, and x_{ref} taken from the optimized outcomes.

We conclude this section by commenting on the uncertainty of the variable R. In

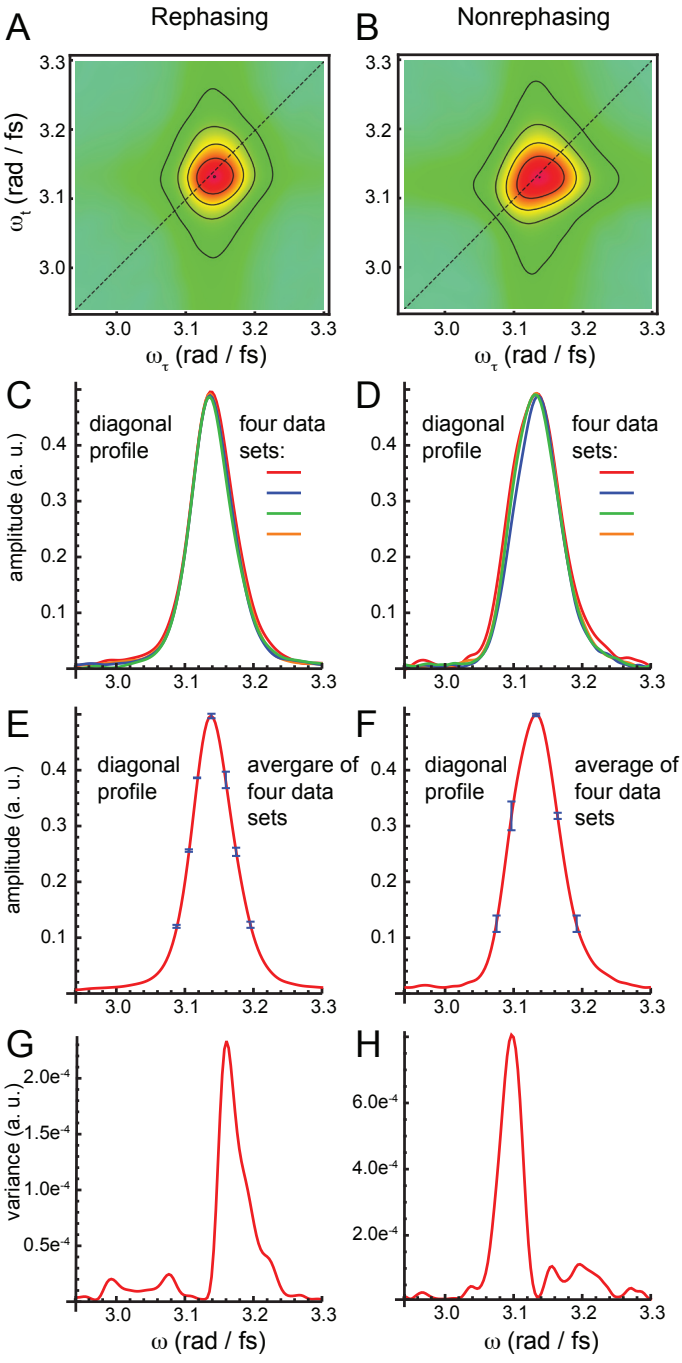

Figure 6.8: Experimental data runs performed on ZnTPP monomer in dimethylformamide solution, which were used for error propagation analysis. In panels (A) and (B) are shown, respectively, the 2D absolute value rephasing and non-rephasing spectra of a single representative data set. In panels (C) and (D) are shown overlays of the absolute value rephasing and non-rephasing signals for each of the four data runs along the diagonal profile. Panels (E) and (F) show the average of the four data sets along the diagonal profile. In panels (G) and (H) are shown the corresponding variances along the diagonal profile. By integrating the standard deviation of the data over the interval $\omega \in (3.07, 3.20)$ rad fs^{-1}, we obtain the relative uncertainties $\Delta g_{RP} = 0.0086$ and $\Delta g_{NRP} = 0.016$ (defined in text). These values are input to Eq. 6.25 to estimate the relative target function uncertainty $\Delta\chi^2_{tot}/\chi^2_{tot} = 0.0096 \approx 1\%$, which in turn establishes the trust intervals of the structural parameters relative to the optimized outcome.

addition to the uncertainties discussed above, an accurate estimate of ΔR must also account for its dependence on the calculated value of the monomer square transition dipole moment $|\boldsymbol{\mu}|^2$. Uncertainty in the estimation of $|\boldsymbol{\mu}|^2$ (Eq. 6.5) will appear in the electronic couplings (Eq. 6.4) as a rescaling of the end-to-end distance R. For example, too small an estimation of $|\boldsymbol{\mu}|^2$ will result in an apparent value of R that is also too small. Although we have attempted to make our estimate of $|\boldsymbol{\mu}|^2$ as accurate as possible, we cannot discount the possibility that a systematic error is present. We note that the values we have obtained for the angles θ, ϕ, α, and β constrain the conformation significantly. We therefore propose that further refinements in the conformation could be achieved through quantum chemical calculations. For example, semi-empirical calculations on the MgTPP dimer, in which only the distance R is varied, could be used to obtain its value where the energy minimum occurs. Given the degree of molecular detail provided by quantum chemical calculations, it should in principle be possible to capture the effects of steric interactions between bulky phenyl groups. Such an approach might be useful to further refine the values of the structural parameters within their trust intervals.

Chapter 7

Future directions

Part I. Quantum annealing for lattice protein folding

In the first part of this dissertation, we showed how to tackle classical optimization problems of interest to the physical sciences within the quantum computing paradigm known as quantum annealing (QA). We presented the largest implementation of QA on a biophysical problem (six different experiments with up to 81 superconducting quantum bits). Although the cases presented here can be solved on a classical computer, we presented the first implementation of lattice protein folding on a quantum device under the Miyazawa-Jernigan model. This is the first step towards studying optimization problems in biophysics and statistical mechanics using quantum devices.

There are several theoretical and experimental challenges which must be overcome before QA can compete with or outperform classical algorithms on substantially larger problem instances. The experimental challenge is clear, requiring new generations of scalable quantum annealing devices, where the quantum mechanical features responsible for the potential speed up over classical computers are preserved as the number of qubits increases. D-Wave, a company which has already released a commercial 128 qubit device, expects to make significant progress towards this goal by releasing a 512 qubit device before the end of the 2012.

Current work in progress tries to address some of the theoretical challenges towards scalability and better performance. For the implementation of any computational problem in quantum annealing, it is always desirable to use the least number of qubits to represent the search space which contains the global minima. An example of such improvement is presented in Chapter 4, where by using the "turn representation" for the different self-avoiding walks reduces the number of qubits needed with respect to the "spatial representation" where the coordinates of each amino acid is specified, as described earlier in Chapter 2. As discussed in Chapter 4, the construction of the energy function is a completely different

problem. In the case for the compact "turn representation", the construction of the Ising-like energy function to be optimized seems to scale exponentially with the size of the system, which is not the case of the construction presented in Chapter 2, where we proved a polynomial scaling with the number of amino acids for the construction of the energy function. In a paper in preparation, we present a solution to both problems: a polynomial scaling for the construction of the energy function within the most compact "turn representation".

Another exciting direction uses new algorithmic strategies which does not involve the construction of an Ising-like cost energy function. Such heuristic strategies combine ideas from machine learning and probabilistic computing [180] and would allow for the exploration of lattice protein folding instances with over a hundred amino acids, competitive with the best state-of-the-art classical algorithms.

Part II. Excitonic energy transfer and nonlinear fluorescence spectroscopy

In the second part of this dissertation, we focused on the problem of excitonic energy transfer. We provided an intuitive platform for engineering exciton transfer dynamics and we show that careful consideration of the properties of the environment allows for opportunities to engineer the transfer of an exciton. Since excitons in nanostructures are proposed for use in quantum information processing and artificial photosynthetic designs, our approach paves the way for engineering a wide range of desired exciton dynamics.

We also developed the theory for a two-dimensional electronic spectroscopic technique based on fluorescence (2DFS) and challenge previous theoretical results claiming its equivalence to the two-dimensional photon echo (2DPE) technique which is based on polarization. Experimental realization of this technique confirmed our theoretical predictions. The new technique is more sensitive than 2DPE as a tool for conformational determination of excitonically coupled chromophores and offers the possibility of applying two-dimensional electronic spectroscopy to single-molecules.

Only one year after publishing the theoretical proposal about engineering directed excitonic energy transfer (see Chapter 5), experimental realization of such quantum dot arrays were synthesized by the group of Prof. Sargent at the University of Toronto [191]. The femtosecond time resolution of the 2DFS technique would allow for study of the energy transfer dynamics in these coupled quantum dots, and for the first time provide insights into the dynamics of excitonic energy transfer in coupled quantum dots as potential candidates for artificial light-harvesting antennas.

Finally, we have characterized the excited state properties of 6-methylisoxanthoptherin (6-MI), a fluorescent analog of the natural nucleic acid base guanine. Since the 2DFS tech-

nique is based on excitonic coupling between excited states, it is important to characterize the electronic transitions moments of the isolated probe chromophore to model the effects of macromolecular structure on spectroscopic observables. Interactions between states of the 6-MI chromophore and those of the local environment perturb the electronic excitation and relaxation pathways, which is the origin of the probes sensitivity to DNA local macromolecular conformation. These local probes will provide insight into biological processes such as DNA replication.

References

[1] D. Abramavicius, B. Palmieri, D. V. Voronine, F. Šanda, and S. Mukamel. Coherent multidimensional optical spectroscopy of excitons in molecular aggregates; quasiparticle versus supermolecule perspectives. *Chem. Rev.*, **109**, 2350 (2009).

[2] S. Acharyya. Sat algorithms for colouring some special classes of graphs: Some theoretical and experimental results. *Journal on Satisfiability, Boolean Modeling and Computation*, **4**, 33 (2007).

[3] D. Achlioptas, A. Naor, and Y. Peres. Rigorous location of phase transitions in hard optimization problems. *Nature*, **435**, 759 (2005).

[4] D. Aharonov and A. Ta-Shma. Adiabatic quantum state generation and statistical zero knowledge. In *Proceedings of the thirty-fifth annual ACM symposium on Theory of computing*, pages 20–29, San Diego, CA, USA, 2003. ACM.

[5] D. Aharonov and A. Ta-Shma. Adiabatic quantum state generation. *SIAM J. Comput*, **37**, 47 (2007).

[6] P. Amara, D. Hsu, and J. E. Straub. Global energy minimum searches using an approximate solution of the imaginary time schroedinger equation. *J. Phys. Chem.*, **97**, 6715 (1993).

[7] A. Ambainis and O. Regev. An elementary proof of the quantum adiabatic theorem. arXiv:quant-ph/0411152 (2004).

[8] M. H. S. Amin. Effect of local minima on adiabatic quantum optimization. *Phys. Rev. Lett.*, **100**, 1305034 (2008).

[9] M. H. S. Amin. Consistency of the adiabatic theorem. *Phys. Rev. Lett.*, **102**, 220401 (2009).

[10] M. H. S. Amin, D. V. Averin, and J. A. Nesteroff. Decoherence in adiabatic quantum computation. *Phys. Rev. A*, **79**, 022107 (2009).

[11] M. H. S. Amin, C. J. S. Truncik, and D. V. Averin. Role of single-qubit decoherence time in adiabatic quantum computation. *Phys. Rev. A*, **80**, 022303 (2009).

[12] M. H. S. Amin, C. J. S. Truncik, and D. V. Averin. Role of single-qubit decoherence time in adiabatic quantum computation. *Phys. Rev. A*, **80**, 022303 (2009).

[13] M. Andrecut and M. K. Ali. Unstructured adiabatic quantum search. *Int. J. Theor. Phys.*, **43**, 925931 (2004).

[14] C. B. Anfinsen. Principles that govern the folding of protein chains. *Science*, **181**, 223 (1973).

[15] A. Aspuru-Guzik, A. D. Dutoi, P. J. Love, and M. Head-Gordon. Simulated quantum computation of molecular energies. *Science*, **309**, 1704 (2005).

[16] R. Backofen and S. Will. A constraint-based approach to structure prediction for simplified protein models that outperforms other existing methods. *Proc. XIX Intl. Conf. on Logic Programming*, pages 49–71 (2003).

[17] D. Baker and D. A. Agard. Kinetics versus thermodynamics in protein foldin. *Biochemistry*, **33**, 7505 (1994).

[18] F. Barahona. On the computational complexity of ising spin glass models. *J. Phys. A: Math. Gen.*, **15**, 3241 (1982).

[19] A. Barenco, C. H. Bennett, R. Cleve, D. P. DiVincenzo, N. Margolus, P. Shor, T. Sleator, J. A. Smolin, and H. Weinfurter. Elementary gates for quantum computation. *Phys. Rev. A*, **52**, 3457 (1995).

[20] B. Berger and T. Leighton. Protein folding in the hydrophobic-hydrophilic (HP) model is NP-complete. *J. Comput. Biol.*, **5**, 27 (SPR 1998).

[21] A. J. Berkley, M. W. Johnson, P. Bunyk, R. Harris, J. Johansson, T. Lanting, E. Ladizinsky, E. Tolkacheva, M. H. S. Amin, and G. Rose. A scalable readout system for a superconducting adiabatic quantum optimization system. *Supercond. Sci. Tech.*, **23**, 105014 (2010).

[22] T. C. Beutler and K. A. Dill. A fast conformational search strategy for finding low energy structures of model proteins. *Protein Sci.*, **5**, 2037 (1996).

[23] J. D. Biamonte. Nonperturbative k-body to two-body commuting conversion hamiltonians and embedding problem instances into ising spins. *Phys. Rev. A*, **77**, 052331 (2008).

[24] Z. Bian, F. Chudak, W. G. Macready, L. Clark, and F. Gaitan. Experimental determination of ramsey numbers with quantum annealing. arXiv:1201.1842v2 (2012).

[25] J. D. Biggs and J. A. Cina. Using wave-packet interferometry to monitor the external vibrational control of electronic excitation transfer. *J. Chem. Phys.*, **131**, 224101 (2009).

[26] J. Bohannon. Distributed computing: Grassroots supercomputing. *Science*, **308**, 810 (2005).

[27] E. Boros and P. Hammer. Pseudo-boolean optimization. *Discrete Appl. Math.*, **123**, 155 (2002).

[28] S. Bravyi. Efficient algorithm for a quantum analogue of 2-SAT. *quant-ph/0602108* (2006).

[29] S. Bravyi, D. P. Divincenzo, R. Oliveira, and B. M. Terhal. The complexity of stoquastic local hamiltonian problems. *Quantum Info. Comput.*, **8**, 361 (2008).

[30] D. Brinks, F. D. Stefani, F. Kulzer, R. Hildner, T. H. Taminiau, Y. Avlasevich, K. Mllen, and N. F. v. Hulst. Visualizing and controlling vibrational wave packets of single molecules. *Nature*, **465**, 905 (2010).

[31] T. Brixner, J. Stenger, H. M. Vaswani, M. Cho, R. E. Blankenship, and G. R. Fleming. Two-dimensional spectroscopy of electronic couplings in photosynthesis. *Nature*, **434**, 625 (2005).

[32] J. Brooke, D. Bitko, T. F. Rosenbaum, and G. Aeppli. Quantum annealing of a disordered magnet. *Science*, **284**, 779 (1999).

[33] T. Calarco, A. Datta, P. Fedichev, E. Pazy, and P. Zoller. Spin-based all-optical quantum computation with quantum dots: Understanding and suppressing decoherence. *Phys. Rev. A.*, **68**, 012310 (2003).

[34] C. J. Camacho and D. Thirumalai. Minimum energy compact structures of random sequences of heteropolymers. *Phys. Rev. Lett.*, **71**, 2505 (1993).

[35] J. Cao and R. J. Silbey. Optimization of exciton trapping in energy transfer processes. *J. Phys. Chem. A*, **113**, 13825 (2009).

[36] H. S. Chan and K. A. Dill. The protein folding problem. *Physics Today*, **46**, 24 (1993).

[37] J.-L. Chen, M. sheng Zhao, J. da Wu, and Y. de Zhang. Invariant perturbation theory of adiabatic process. arXiv:0706.0299 (2007).

[38] Y. Cheng and G. R. Fleming. Dynamics of light harvesting in photosynthesis. *Annu. Rev. Phys. Chem.*, **60**, 241 (2009).

[39] A. M. Childs, E. Farhi, and J. Preskill. Robustness of adiabatic quantum computation. *Phys. Rev. A.*, **65**, 012322 (2001).

[40] M. Cho. *Two-Dimensional Optical Spectroscopy.* CRC Press, Boca Raton, FL, 1st edition, 2009.

[41] M. Cho and G. R. Fleming. The integrated photon echo and solvation dynamics. II. peak shifts and two-dimensional photon echo of a coupled chromophore system. *J. Chem. Phys.*, **123**, 114506 (2005).

[42] M. Christandl, N. Datta, T. C. Dorlas, A. Ekert, A. Kay, and A. J. Landahl. Perfect transfer of arbitrary states in quantum spin networks. *Phys. Rev. A*, **71**, 032312 (2005).

[43] E. Collini and G. D. Scholes. Coherent intrachain energy migration in a conjugated polymer at room temperature. *Science*, **323**, 369 (2009).

[44] E. Collini, C. Y. Wong, K. E. Wilk, P. M. G. Curmi, P. Brumer, and G. D. Scholes. Coherently wired light-harvesting in photosynthetic marine algae at ambient temperature. *Nature*, **463**, 644 (2010).

[45] G. A. Cox and R. L. Johnston. Analyzing energy landscapes for folding model proteins. *J. Chem. Phys.*, **124**, 204714 (2006).

[46] G. A. Cox, T. V. Mortimer-Jones, R. P. Taylor, and R. L. Johnston. Development and optimisation of a novel genetic algorithm for studying model protein folding. *Theor. Chem. Acc.*, **112**, 163 (2004).

[47] T. E. Creighton, editor. *Protein folding*. New York: W. H. Freeman, 1992.

[48] P. Crescenzi, D. Goldman, C. Papadimitriou, A. Piccolboni, and M. Yannakakis. On the complexity of protein folding. *J. Comput. Biol.*, **5**, 597 (1998).

[49] P. Crescenzi, D. Goldman, C. Papadimitriou, A. Piccolboni, and M. Yannakakis. On the complexity of protein folding. *J. Comput. Biol.*, **5**, 423 (FAL 1998).

[50] S. A. Crooker, J. A. Hollingsworth, S. Tretiak, and V. I. Klimov. Spectrally resolved dynamics of energy transfer in Quantum-Dot assemblies: Towards engineered energy flows in artificial materials. *Phys. Rev. Lett.*, **89**, 186802 (2002).

[51] F. L. Custdio, H. J. C. Barbosa, and L. E. Dardenne. Investigation of the three-dimensional lattice hp protein folding model using a genetic algorithm. *Genet. Mol. Biol.*, **27**, 611 (2004).

[52] J. da Wu, M. sheng Zhao, J. lan Chen, and Y. de Zhang. Adiabatic approximation condition. arXiv:0706.0264v2 (2007).

[53] A. Das and B. K. Chakrabarti. Colloquium: Quantum annealing and analog quantum computation. *Rev. Mod. Phys.*, **80**, 106121 (2008).

[54] D. Deutsch. Quantum computational networks. *Proc. Roy. Soc. A*, **425**, 73 (1989).

[55] L. DiCarlo, J. M. Chow, J. M. Gambetta, L. S. Bishop, B. R. Johnson, D. I. Schuster, J. Majer, A. Blais, L. Frunzio, S. M. Girvin, and R. J. Schoelkopf. Demonstration of two-qubit algorithms with a superconducting quantum processor. *Nature*, **460**, 240 (2009).

[56] L. DiCarlo, M. D. Reed, L. Sun, B. R. Johnson, J. M. Chow, J. M. Gambetta, L. Frunzio, S. M. Girvin, M. H. Devoret, and R. J. Schoelkopf. Preparation and measurement of three-qubit entanglement in a superconducting circuit. *Nature*, **467**, 574 (2010).

[57] K. Dill, S. Bromberg, K. Yue, K. Fiebig, D. Yee, P. Thomas, and H. Chan. Principles of protein folding – a perspective from simple exact models. *Protein Sci.*, **4**, 561 (1995).

[58] K. A. Dill, S. B. Ozkan, M. S. Shell, and T. R. Weikl. The protein folding problem. *Ann. Rev. Biophys.*, **37**, 289 (2008).

[59] D. P. DiVincenzo. The physical implementation of quantum computation. *Fortschritte der Physik*, **48**, 771 (2000).

[60] J. Du, L. Hu, Y. Wang, J. Wu, M. Zhao, and D. Suter. Experimental study of the validity of quantitative conditions in the quantum adiabatic theorem. *Phys. Rev. Lett*, **101**, 060403 (2008).

[61] C. J. Epstein, R. F. Goldberger, and C. B. Anfinsen. Genetic control of tertiary protein structure - studies with model systems. *Cold. Spring. Harb. Sym.*, **28**, 439 (1963).

[62] R. R. Ernst, G. Bodenhausen, and A. Wokaun. *Principles of Nuclear Magnetic Resonance in One and Two Dimensions*. Oxford University Press, USA, Sept. 1990.

[63] C. Fang, J. D. Bauman, K. Das, A. Remorino, E. Arnold, and R. M. Hochstrasser. Two-dimensional infrared spectra reveal relaxation of the nonnucleoside inhibitor TMC278 complexed with HIV-1 reverse transcriptase. *Proc. Natl. Acad. Sci. USA.*, **105**, 1472 (2008).

[64] E. Farhi, J. Goldstone, D. Gosset, S. Gutmann, H. B. Meyer, and P. Shor. Quantum adiabatic algorithms, small gaps, and different paths. *0909.4766* (2009).

[65] E. Farhi, J. Goldstone, and S. Gutmann. Quantum adiabatic evolution algorithms with different paths. *quant-ph/0208135* (2002).

[66] E. Farhi, J. Goldstone, and S. Gutmann. Quantum adiabatic evolution algorithms with different paths. *quant-ph/0208135* (2002).

[67] E. Farhi, J. Goldstone, S. Gutmann, J. Lapan, A. Lundgren, and D. Preda. A quantum adiabatic evolution algorithm applied to random instances of an NP-Complete problem. *Science*, **292**, 472 (2001).

[68] E. Farhi, J. Goldstone, S. Gutmann, and D. Nagaj. How to make the quantum adiabatic algorithm fail. *International Journal of Quantum Information*, **06**, 503 (2008).

[69] E. Farhi, J. Goldstone, S. Gutmann, and M. Sipser. Quantum computation by adiabatic evolution. arXiv:quant-ph/0001106 (2000).

[70] A. B. Finnila, M. A. Gomez, C. Sebenik, C. Stenson, and J. D. Doll. Quantum annealing: A new method for minimizing multidimensional functions. *Chem. Phys. Lett.*, **219**, 343 (1994).

[71] K. M. Gaab and C. J. Bardeen. The effects of connectivity, coherence, and trapping on energy transfer in simple light-harvesting systems studied using the Haken-Strobl model with diagonal disorder. *J. Chem. Phys.*, **121**, 7813 (2004).

[72] M. Garey and D. Johnson. *Computers and Intractability. A Guide to the Theory of NP-Completeness.* W.H. Freeman and Co., NY, 1979.

[73] I. Gent and T. Walsh. The search for satisfaction. *Internal report, department of computer science, University of Strathclyde* (1999).

[74] N. S. Ginsberg, Y. Cheng, and G. R. Fleming. Two-Dimensional electronic spectroscopy of molecular aggregates. *Acc. Chem. Res.*, **42**, 1352 (2009).

[75] M. Gouterman. Optical spectra and electronic structure of porphyrins and related rings. In D. Dolphin, editor, *The Porphyrins, Physical Chemistry, Part A*, volume III, pages 1–156. Academic Press, New York, 1979.

[76] L. K. Grover. Quantum mechanics helps in searching for a needle in a haystack. *Phys. Rev. Lett.*, **79**, 325 (1997).

[77] M. Gruebele. The fast protein folding problem. *Annu. Rev. Phys. Chem.*, **50**, 485 (1999).

[78] J. L. Hardwick. Absorption and emission of electromagnetic radiation. In S. Wilson, editor, *Handbook of Molecular Physics and Quantum Chemistry.* Wiley, New York, 2003.

[79] R. Harris, A. Berkley, J. Johansson, M. Johnson, T. Lanting, P. Bunyk, E. Tolkacheva, E. Ladizinsky, B. Bumble, A. Fung, A. Kaul, A. Kleinsasser, and S. Han. Implementation of a quantum annealing algorithm using a superconducting circuit. arXiv:0903.3906v1 (2009).

[80] R. Harris, A. Berkley, M. Johnson, P. Bunyk, S. Govorkov, M. Thom, S. Uchaikin, A. Wilson, J. Chung, E. Holtham, J. Biamonte, A. Yu, M. Amin, and A. van den Brink. Sign- and magnitude-tunable coupler for superconducting flux qubits. *Phys. Rev. Lett.*, **98**, 177001 (2007).

[81] R. Harris, F. Brito, A. J. Berkley, J. Johansson, M. W. Johnson, T. Lanting, P. Bunyk, E. Ladizinsky, B. Bumble, A. Fung, A. Kaul, A. Kleinsasser, and S. Han. Synchronization of multiple coupled rf-SQUID flux qubits. *New J. Phys.*, **11**, 123022 (2009).

[82] R. Harris, J. Johansson, A. J. Berkley, M. W. Johnson, T. Lanting, S. Han, P. Bunyk, E. Ladizinsky, T. Oh, I. Perminov, E. Tolkacheva, S. Uchaikin, E. M. Chapple, C. Enderud, C. Rich, M. Thom, J. Wang, B. Wilson, and G. Rose. Experimental demonstration of a robust and scalable flux qubit. *Phys. Rev. B*, **81**, 134510 (2010).

[83] R. Harris, M. W. Johnson, S. Han, A. J. Berkley, J. Johansson, P. Bunyk, E. Ladizinsky, S. Govorkov, M. C. Thom, S. Uchaikin, B. Bumble, A. Fung, A. Kaul, A. Kleinsasser, M. H. S. Amin, and D. V. Averin. Probing noise in flux qubits via macroscopic resonant tunneling. *Phys. Rev. Lett.*, **101**, 117003 (2008).

[84] R. Harris, M. W. Johnson, T. Lanting, A. J. Berkley, J. Johansson, P. Bunyk, E. Tolkacheva, E. Ladizinsky, N. Ladizinsky, T. Oh, F. Cioata, I. Perminov, P. Spear, C. Enderud, C. Rich, S. Uchaikin, M. C. Thom, E. M. Chapple, J. Wang, B. Wilson, M. H. S. Amin, N. Dickson, K. Karimi, B. Macready, C. J. S. Truncik, and G. Rose. Experimental investigation of an eight-qubit unit cell in a superconducting optimization processor. *Phys. Rev. B.*, **82**, 024511 (2010).

[85] R. Harris, M. W. Johnson, T. Lanting, A. J. Berkley, J. Johansson, P. Bunyk, E. Tolkacheva, E. Ladizinsky, N. Ladizinsky, T. Oh, F. Cioata, I. Perminov, P. Spear, C. Enderud, C. Rich, S. Uchaikin, M. C. Thom, E. M. Chapple, J. Wang, B. Wilson, M. H. S. Amin, N. Dickson, K. Karimi, B. Macready, C. J. S. Truncik, and G. Rose. Experimental investigation of an eight-qubit unit cell in a superconducting optimization processor. *Phys. Rev. B*, **82**, 024511 (2010).

[86] R. Harris, T. Lanting, A. J. Berkley, J. Johansson, M. W. Johnson, P. Bunyk, E. Ladizinsky, N. Ladizinsky, T. Oh, and S. Han. Compound Josephson-junction coupler for flux qubits with minimal crosstalk. *Phys. Rev. B*, **80**, 052506 (2009).

[87] W. E. Hart and S. Istrail. Robust proofs of NP-Hardness for protein folding: General lattices and energy potentials. *J. Comput. Biol.*, **4**, 1 (1997).

[88] A. K. Hartmann and H. Rieger. *New Optimization Algorithms in Physics.* Wiley-VCH, 2004.

[89] D. Hayes and G. Engel. Extracting the excitonic hamiltonian of the Fenna-Matthews-Olson complex using Three-Dimensional Third-Order electronic spectroscopy. *Biophysical Journal*, **100**, 2043 (2011).

[90] A. A. High, E. E. Novitskaya, L. V. Butov, M. Hanson, and A. C. Gossard. Control of exciton fluxes in an excitonic integrated circuit. *Science*, **321**, 229 (2008).

[91] M. Hofheinz, H. Wang, M. Ansmann, R. C. Bialczak, E. Lucero, M. Neeley, A. D. O'Connell, D. Sank, J. Wenner, J. M. Martinis, and A. N. Cleland. Synthesizing arbitrary quantum states in a superconducting resonator. *Nature*, **459**, 546 (2009).

[92] T. Hogg. Highly structured searches with quantum computers. *Phys. Rev. Lett.*, **80**, 2473 (1998).

[93] T. Hogg. Adiabatic quantum computing for random satisfiability problems. *Phys. Rev. A.*, **67**, 022314 (2003).

[94] T. Hogg and . C. H. R. T. X. P. A. R. Center. Quantum search heuristics. *Phys. Rev. A.*, **61**, 052311 (2000).

[95] H. P. Hsu, V. Mehra, W. Nadler, and P. Grassberger. Growth-based optimization algorithm for lattice heteropolymers. *Physical Review E*, **68**, 021113 (2003).

[96] S. Jansen, M. Ruskai, and R. Seiler. Bounds for the adiabatic approximation with applications to quantum computation. *J. Math. Phys*, **48**, 102111 (2007).

[97] M. W. Johnson, M. H. S. Amin, S. Gildert, T. Lanting, F. Hamze, N. Dickson, R. Harris, A. J. Berkley, J. Johansson, P. Bunyk, E. M. Chapple, C. Enderud, J. P. Hilton, K. Karimi, E. Ladizinsky, N. Ladizinsky, T. Oh, I. Perminov, C. Rich, M. C. Thom, E. Tolkacheva, C. J. S. Truncik, S. Uchaikin, J. Wang, B. Wilson, and G. Rose. Quantum annealing with manufactured spins. *Nature*, **473**, 194 (2011).

[98] M. W. Johnson, P. Bunyk, F. Maibaum, E. Tolkacheva, A. J. Berkley, E. M. Chapple, R. Harris, J. Johansson, T. Lanting, I. Perminov, E. Ladizinsky, T. Oh, and G. Rose. A scalable control system for a superconducting adiabatic quantum optimization processor. *Supercond. Sci. Tech.*, **23**, 065004 (2010).

[99] T. Kadowaki and H. Nishimori. Quantum annealing in the transverse ising model. *Phys. Rev. E.*, **58**, 5355 (1998).

[100] W. M. Kaminsky, S. Lloyd, and T. P. Orlando. Scalable superconducting architecture for adiabatic quantum computation. arXiv:quant-ph/0403090 (2004).

[101] H. Kautz and B. Selman. The state of SAT. *Discrete Applied Mathematics*, **155**, 1514 (2007).

[102] J. Kempe. Quantum random walks - an introductory overview. *Contemp. Phys.*, **44**, 307 (2003).

[103] J. Kempe, A. Kitaev, and O. Regev. The complexity of the local hamiltonian problem. *SIAM Journal on Computing*, **35**, 10701097 (2006).

[104] D. Kim, S. Okahara, M. Nakayama, and Y. Shim. Experimental verification of förster energy transfer between semiconductor quantum dots. *Phys. Rev. B.*, **78**, 153301 (2008).

[105] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. Optimization by simulated annealing. *Science*, **220**, 671 (1983).

[106] A. Kitaev. Fault-tolerant quantum computation by anyons. *Ann. Phys.*, **303**, 2 (2003).

[107] P. Kjellberg, B. Brggemann, and T. Pullerits. Two-dimensional electronic spectroscopy of an excitonically coupled dimer. *Phys. Rev. B.*, **74**, 024303 (2006).

[108] V. I. Klimov. Spectral and dynamical properties of multiexcitons in semiconductor nanocrystals. *Annu. Rev. Phys. Chem.*, **58**, 635 (2007).

[109] J. Knoester. (2002) Optical Properties of Molecular Aggregates. Proceedings of the International School of Physics Enrico Fermi Course, eds M Agranovich and GC La Rocca Vol CXLIX (IOS Press, Amsterdam), pp 149186.

[110] D. Kohen and D. J. Tannor. Quantum adiabatic switching. *J. Chem. Phys.*, **98**, 3168 (1993).

[111] A. Kolinski and J. Skolnick. *Lattice Models of Protein Folding, Dynamics and Thermodynamics*. Chapman & Hall, 1996.

[112] M. Koolhaas, G. v. d. Zwan, F. v. Mourik, and R. v. Grondelle. Spectroscopy and structure of bacteriochlorophyll dimers. i. structural consequences of nonconservative circular dichroism spectra. *Biophysical Journal*, **72**, 1828 (1997).

[113] K. Kwak, S. Park, and M. D. Fayer. Dynamics around solutes and solutesolvent complexes in mixed solvents. *Proc. Natl. Acad. Sci. USA.*, **104**, 14221 (2007).

[114] J. R. Lakowicz. *Principles of Fluorescence Spectroscopy*. Springer, 3rd edition, Sept. 2006.

[115] T. Lanting, M. Amin, M. Johnson, F. Altomare, A. Berkley, S. Gildert, R. Harris, J. Johansson, P. Bunyk, E. Ladizinsky, E. Tolkacheva, and D. Averin. Probing high frequency noise with macroscopic resonant tunneling. arXiv:1103.1931v1 (2011).

[116] T. Lanting, A. J. Berkley, B. Bumble, P. Bunyk, A. Fung, J. Johansson, A. Kaul, A. Kleinsasser, E. Ladizinsky, F. Maibaum, R. Harris, M. W. Johnson, E. Tolkacheva, and M. H. S. Amin. Geometrical dependence of the low-frequency noise in superconducting flux qubits. *Phys. Rev. B*, **79**, 060509(R) (2009).

[117] T. Lanting, R. Harris, J. Johansson, M. H. S. Amin, A. J. Berkley, S. Gildert, M. W. Johnson, P. Bunyk, E. Tolkacheva, E. Ladizinsky, N. Ladizinsky, T. Oh, I. Perminov, E. M. Chapple, C. Enderud, C. Rich, B. Wilson, M. C. Thom, S. Uchaikin, and G. Rose. Cotunneling in pairs of coupled flux qubits. *Phys. Rev. B*, **82**, 060512 (2010).

[118] J. Latorre and R. Orus. Adiabatic quantum computation and quantum phase transitions. *Phys. Rev. A.*, **69**, 062302 (2004).

[119] K. F. Lau and K. A. Dill. A lattice statistical-mechanics model of the conformational and sequence-spaces of proteins. *Macromolecules.*, **22**, 3986 (1989).

[120] T. Lazaridisa and M. Karplus. Thermodynamics of protein folding: a microscopic view. *Biophysical Chemistry*, **100**, 367 (2003).

[121] H. Lee, Y. Cheng, and G. R. Fleming. Coherence dynamics in photosynthesis: Protein protection of excitonic coherence. *Science*, **316**, 1462 (2007).

[122] H. Li, C. Tang, and N. Wingreen. Designability of protein structures: A lattice-model study using the miyazawa-jernigan matrix. *Proteins*, **49**, 403 (2002).

[123] D. A. Lidar. Towards fault tolerant adiabatic quantum computation. *Phys. Rev. Lett.*, **100**, 160506 (2008).

[124] A. Lupascu, S. Saito, T. Picot, P. C. de Groot, C. J. P. M. Harmans, and J. E. Mooij. Quantum non-demolition measurement of a superconducting two-level system. *Nat. Phys.*, **3**, 119 (2007).

[125] R. MacKenzie, A. Morin-Duchesne, H. Paquette, and J. Pinel. Validity of the adiabatic approximation in quantum mechanics. *Phys. Rev. A.*, **76**, 044102 (2007).

[126] O. Madelung, editor. *Semiconductors: data handbook*. Springer, Berlin, 3rd edition, 2004.

[127] K. Marzlin and B. C. Sanders. Inconsistency in the application of the adiabatic theorem. *Phys. Rev. Lett.*, **93**, 160408 (2004).

[128] S. Matile, N. Berova, K. Nakanishi, J. Fleischhauer, and R. W. Woody. Structural studies by exciton coupled circular dichroism over a large distance:? porphyrin derivatives of steroids, dimeric steroids, and brevetoxin b? *J. Am. Chem. Soc.*, **118**, 5198 (1996).

[129] V. May and O. Kuhn. *Charge and Energy Transfer Dynamics in Molecular Systems*. Wiley-VCH, 2 edition, Mar. 2004.

[130] I. L. Medintz and H. Mattoussi. Quantum dot-based resonance energy transfer and its growing application in biology. *Phys. Chem. Chem. Phys.*, **11**, 17 (2009).

[131] A. Messiah. *Quantum Mechanics (Physics)*. Dover Publications, July 1999.

[132] M. Mezard, T. Mora, and R. Zecchina. Clustering of solutions in the random satisfiability problem. *Phys. Rev. Lett*, **94**, 197205 (2005).

[133] M. Mezard, G. Parisi, and R. Zecchina. Analytic and algorithmic solution of random satisfiability problems. *Science*, **297**, 812 (2002).

[134] L. Mirny and E. Shakhnovich. Protein folding theory: from lattice to all-atom models. *Annu. Rev. Biophys. Bio.*, **30**, 361 (2001).

[135] S. Miyazawa and R. L. Jernigan. Residue-residue potentials with a favorable contact pair term and an unfavorable high packing density term, for simulation and threading. *Journal of Molecular Biology*, **256**, 623 (1996).

[136] A. Mizel, D. A. Lidar, and M. Mitchell. Simple proof of equivalence between adiabatic quantum computation and the circuit model. *Phys. Rev. Lett.*, **99**, 070502 (2007).

[137] M. Motoki. Watanabe research group, Dept. of Math. and Computing Sciences, Tokyo Inst. of Technology, http://www.is.titech.ac.jp/ watanabe/gensat/.

[138] S. Mukamel. *Principles of Nonlinear Optical Spectroscopy*. Oxford University Press, USA, 1st edition, Jan. 1995.

[139] C. Nayak, S. H. Simon, A. Stern, M. Freedman, and S. D. Sarma. Non-abelian anyons and topological quantum computation. *Rev. Mod. Phys.*, **80**, 1083 (2008).

[140] A. Nazir, B. W. Lovett, S. D. Barrett, J. H. Reina, and G. A. D. Briggs. Anticrossings in frster coupled quantum dots. *Phys. Rev. B.*, **71**, 045334 (2005).

[141] M. Neeley, R. C. Bialczak, M. Lenander, E. Lucero, M. Mariantoni, A. D. O/'Connell, D. Sank, H. Wang, M. Weides, J. Wenner, Y. Yin, T. Yamamoto, A. N. Cleland, and J. M. Martinis. Generation of three-qubit entangled states using superconducting phase qubits. *Nature*, **467**, 570 (2010).

[142] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2000.

[143] A. Nitzan. *Chemical Dynamics in Condensed Phases: Relaxation, Transfer, and Reactions in Condensed Molecular Systems*. Oxford University Press, USA, June 2006.

[144] R. Oliveira and B. M. Terhal. The complexity of quantum system on a two-dimensional square lattice. *Quantum. Inf. Comput.*, **8**, 0900 (2008).

[145] R. Orus and J. I. Latorre. Universality of entanglement and quantum-computation complexity. *Phys. Rev. A.*, **69**, 052308 (2004).

[146] V. S. Pande. Simple theory of protein folding kinetics. *Phys. Rev. Lett.*, **105**, 198101 (2010).

[147] V. S. Pande, A. Y. Grosberg, and T. Tanaka. Heteropolymer freezing and design: Towards physical models of protein folding. *Rev. Mod. Phys.*, **72**, 259 (2000).

[148] A. Perdomo, C. Truncik, I. Tubert-Brohman, G. Rose, and A. Aspuru-Guzik. Construction of model hamiltonians for adiabatic quantum computation and its application to finding low-energy conformations of lattice protein models. *Phys. Rev. A*, **78**, 012320 (2008).

[149] A. Perdomo-Ortiz, B. O'Gorman, and A. Aspuru-Guzik. Construction of energy functions for self-avoiding walks and the lattice heteropolymer model: resource efficient encoding for quantum optimization. *In preparation* (2011).

[150] A. Perdomo-Ortiz, J. R. Widom, G. A. Lott, A. H. Marcus, and A. Aspuru-Guzik. Comparison of nonlinear fluorescent- detected and optical photon echo techniques: theoretical and experimental evidence of their differences. *In preparation* (2012).

[151] M. B. Plenio and S. F. Huelga. Dephasing-assisted transport: quantum networks and biomolecules. *New J. Phys.*, **10**, 113019 (14pp) (2008).

[152] S. S. Plotkin and J. N. Onuchic. Understanding protein folding with energy landscape theory part i: Basic concepts. *Quarterly Reviews of Biophysics*, 35 (2002).

[153] W. T. Pollard and R. A. Friesner. Solution of the redfield equation for the dissipative quantum dynamics of multilevel systems. *J. Chem. Phys.*, **100**, 5054 (1994).

[154] R. Raussendorf and H. J. Briegel. A one-way quantum computer. *Phys. Rev. Lett.*, **86**, 5188 (2001).

[155] R. Raussendorf, D. E. Browne, and H. J. Briegel. Measurement-based quantum computation on cluster states. *Phys. Rev. A*, **68**, 022312 (2003).

[156] E. L. Read, G. S. Engel, T. R. Calhoun, T. Mančal, T. K. Ahn, R. E. Blankenship, and G. R. Fleming. Cross-peak-specific two-dimensional electronic spectroscopy. *Proc. Natl. Acad. Sci. USA.*, **104**, 14203 (2007).

[157] E. L. Read, G. S. Schlau-Cohen, G. S. Engel, J. Wen, R. E. Blankenship, and G. R. Fleming. Visualization of excitonic structure in the Fenna-Matthews-Olson photosynthetic complex by Polarization-Dependent Two-Dimensional electronic spectroscopy. *Biophysical Journal*, **95**, 847 (2008).

[158] P. Rebentrost, M. Mohseni, I. Kassal, S. Lloyd, and A. Aspuru-Guzik. Environment-assisted quantum transport. *New J. Phys.*, **11**, 033003 (2009).

[159] A. G. Redfield. On the theory of relaxation processes. *IBM Journal of Research and Development*, **1**, 19 (1957).

[160] T. Renger and R. A. Marcus. On the relation of protein dynamics and exciton relaxation in pigmentprotein complexes: An estimation of the spectral density and a theory for the calculation of optical spectra. *J. Chem. Phys.*, **116**, 999710019 (2002).

[161] A. T. Rezakhani, W. Kuo, A. Hamma, D. A. Lidar, and P. Zanardi. Quantum adiabatic brachistochrone. *Phys. Rev. Lett.*, **103**, 080502 (2009).

[162] M. Richter, K. J. Ahn, A. Knorr, A. Schliwa, D. Bimberg, M. E. Madjet, and T. Renger. Theory of excitation transfer in coupled nanostructures - from quantum dots to light harvesting complexes. *physica status solidi (b)*, **243**, 2302 (2006).

[163] A. L. Rogach, T. A. Klar, J. M. Lupton, A. Meijerink, and J. Feldmann. Energy transfer with semiconductor nanocrystals. *J. Mater. Chem.*, **19**, 1208 (2009).

[164] J. Roland and N. J. Cerf. Quantum search by local adiabatic evolution. *Phys. Rev. A.*, **65**, 042308 (2002).

[165] K. Rosen. *Discrete Mathematics and its applications*. McGraw-Hill, 1999.

[166] E. Rozbicki and P. Machnikowski. Quantum kinetic theory of Phonon-Assisted excitation transfer in quantum dot molecules. *Phys. Rev. Lett.*, **100**, 027401 (2008).

[167] G. Santoro, R. Martonk, E. Tosatti, and R. Car. Theory of quantum annealing of an ising spin glass. *Science*, **295**, 24272430 (2002). 10.1126/science.1068774.

[168] G. E. Santoro and E. Tosatti. Optimization using quantum mechanics: quantum annealing through adiabatic evolution. *J. Phys. A.*, **39**, R393 (2006).

[169] G. S. Schlau-Cohen, T. R. Calhoun, N. S. Ginsberg, M. Ballottari, R. Bassi, and G. R. Fleming. Spectroscopic elucidation of uncoupled transition energies in the major photosynthetic light-harvesting complex, LHCII. *Proc. Natl. Acad. Sci. USA.*, **107**, 13276 (2010).

[170] G. D. Scholes. Long-range resonance energy transfer in molecular systems. *Annu. Rev. Phys. Chem.*, **54**, 57 (2003).

[171] G. D. Scholes and G. Rumbles. Excitons in nanoscale systems. *Nat Mater*, **5**, 683 (2006).

[172] R. Schutzhold and G. Schaller. Adiabatic quantum algorithms as quantum phase transitions: First versus second order. *Phys. Rev. A*, **74**, 060304 (2006).

[173] E. I. Shakhnovich. Proteins with selected sequences fold into unique native conformation. *Phys. Rev. Lett.*, **72**, 3907 (1994).

[174] E. Shaknovich. Protein folding thermodynamics and dynamics: Where physics, chemistry and biology meet. *Chem. Rev.*, 106 (2006).

[175] D. E. Shaw, P. Maragakis, K. Lindorff-Larsen, S. Piana, R. O. Dror, M. P. Eastwood, J. A. Bank, J. M. Jumper, J. K. Salmon, Y. Shan, and W. Wriggers. Atomic-Level characterization of the structural dynamics of proteins. *Science*, **330**, 341 (2010).

[176] A. Shmygelska and H. H. Hoos. An ant colony optimisation algorithm for the 2d and 3d hydrophobic polar protein folding problem. *BMC Bioinformatics*, 6 (2005).

[177] P. W. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM J. Comput.*, **26**, 1484 (1997).

[178] M. Sipser. *Introduction to the Theory of Computation*. PWS Publishing Co., 2005.

[179] M. S. Siu. Adiabatic rotation, quantum search, and preparation of superposition states. *Phys. Rev. A*, **75**, 0623375 (2007).

[180] V. N. Smelyanskiy, E. G. Rieffel, S. I. Knysh, C. P. Williams, M. W. Johnson, M. C. Thom, W. G. Macready, and K. L. Pudenz. A near-term quantum computing approach for hard computational problems in space exploration. arXiv:1204.2821v2 (2012).

[181] J. Song, J. Cheng, T. Zheng, and J. mao. A novel genetic algorithm for hp model protein folding. *Proceedings of the Sixth International Conference on Parallel and Distributed Computing, Applications and Technologies (PDCAT'05)*, pages 935–937 (2005).

[182] M. Steffen, W. van Dam, T. Hogg, G. Breyta, and I. Chuang. Experimental implementation of an adiabatic quantum optimization algorithm. *Phys. Rev. Lett.*, **90**, 067903 (2003).

[183] K. Steinhfel, A. Skaliotisa, and A. A. Albrecht. Relating time complexity of protein folding simulation to approximations of folding time. *Comp. Phys. Commun.*, **176**, 465 (2007).

[184] R. G. Stomphorst, R. B. M. Koehorst, G. van der Zwan, B. Benthem, and T. J. Schaafsma. Excitonic interactions in covalently linked porphyrin dimers with rotational freedom. *J. Porphyr. Phthalocya.*, **3**, 346 (1999).

[185] K. W. Stone, K. Gundogdu, D. B. Turner, X. Li, S. T. Cundiff, and K. A. Nelson. Two-Quantum 2D FT electronic spectroscopy of biexcitons in GaAs quantum wells. *Science*, **324**, 1169 (2009).

[186] V. Szocs, T. Palszegi, V. Luke, J. Sperling, F. Milota, W. Jakubetz, and H. F. Kauffmann. Two-dimensional electronic spectra of symmetric dimers: Intermolecular coupling and conformational states. *J. Chem. Phys.*, **124**, 124511 (2006).

[187] L. Szyc, M. Yang, E. T. J. Nibbering, and T. Elsaesser. Ultrafast vibrational dynamics and local interactions of hydrated DNA. *Angew. Chemie. Int. Edit.*, **49**, 3598 (2010).

[188] H. Tan. Theory and phase-cycling scheme selection principles of collinear phase coherent multi-dimensional optical spectroscopy. *J. Chem. Phys.*, **129**, 124501 (2008).

[189] J. R. Taylor. *An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements*. University Science Books, Sausalito, CA, 2nd edition, 1997.

[190] P. F. Tekavec, G. A. Lott, and A. H. Marcus. Fluorescence-detected two-dimensional electronic coherence spectroscopy by acousto-optic phase modulation. *J. Chem. Phys.*, **127**, 214307 (2007).

[191] G. Tikhomirov, S. Hoogland, P. E. Lee, A. Fischer, E. H. Sargent, and S. O. Kelley. DNA-based programming of quantum dot valency, self-assembly and luminescence. *Nat. Nanotechnol.*, **6**, 485 (2011).

[192] L. Toma and S. Toma. Contact interactions method: A new algorithm for protein folding simulations. *Protein Sci.*, **5**, 147 (1996).

[193] D. M. Tong. Quantitative condition is necessary in guaranteeing the validity of the adiabatic approximation. *Phys. Rev. Lett.*, **104**, 120401 (2010).

[194] D. M. Tong, K. Singh, L. C. Kwek, and C. H. Oh. Sufficiency criterion for the validity of the adiabatic approximation. *Phys. Rev. Lett.*, **98**, 150402 (2007).

[195] R. Unger and J. Moult. Genetic algorithms for protein folding simulations. *J. Mol. Biol.*, **231**, 75 (1993).

[196] A. Šali, E. Shakhnovich, and M. Karplus. How does a protein fold? *Nature*, **369**, 248 (1994).

[197] D. Vion, A. Aassime, A. Cottet, P. Joyez, H. Pothier, C. Urbina, D. Esteve, and M. H. Devoret. Manipulating the quantum state of an electrical circuit. *Science*, **296**, 886 (2002).

[198] D. V. Voronine, D. Abramavicius, and S. Mukamel. Coherent control of cross-peaks in chirality-induced two-dimensional optical signals of excitons. *J. Chem. Phys.*, **125**, 224504 (2006).

[199] M. Žnidaric. Scaling of the running time of the quantum adiabatic algorithm for propositional satisfiability. *Phys. Rev. A.*, **71**, 062305 (2005).

[200] H. Wang, S. Ashhab, and F. Nori. Efficient quantum algorithm for preparing molecular-system-like states on a quantum computer. *Phys. Rev. A*, **79**, 042335 (2009).

[201] H. Wang, S. Kais, A. Aspuru-Guzik, and M. R. Hoffmann. Quantum algorithm for obtaining the energy spectrum of molecular systems. *Phys. Chem. Chem. Phys.*, **10**, (2008).

[202] N. J. Ward, I. Kassal, and A. Aspuru-Guzik. Preparation of many-body states for quantum simulation. *J. Chem. Phys.*, **130**, 194105 (2009).

[203] Z. Wei and M. Ying. Quantum adiabatic computation and adiabatic conditions. *Phys. Rev. A.*, **76**, 024304 (2007).

[204] W. Wernsdorfer. Molecular nanomagnets: towards molecular spintronics. *Int. J. Nanotechnol.*, **7**, 497 (2010).

[205] Y. Won, R. A. Friesner, M. R. Johnson, and J. L. Sessler. Exciton interactions in synthetic porphyrin dimers. *Photosynth. Res.*, **22**, 201 (1989). 10.1007/BF00048299.

[206] N. Xu, J. Zhu, D. Lu, X. Zhou, X. Peng, and J. Du. Quantum factorization of 143 on a dipolar-coupling nmr system. arXiv:1111.3726v1 (2011).

[207] J. Q. You and F. Nori. Superconducting circuits and quantum information. *Phys. Today.*, **58**, 42 (2005).

[208] A. P. Young, S. Knysh, and V. N. Smelyanskiy. Size dependence of the minimum excitation gap in the quantum adiabatic algorithm. *Phys. Rev. Lett.*, **101**, 170503 (2008).

[209] A. P. Young, S. Knysh, and V. N. Smelyanskiy. First-order phase transition in the quantum adiabatic algorithm. *Phys. Rev. Lett.*, **104**, 020502 (2010).

[210] K. Yue and K. A. Dill. Forces of tertiary structural organization in globular proteins. *Proc. Natl. Acad. Sci. USA.*, **92**, 146 (1995).

[211] K. Yue, K. M. Fiebig, P. D. Thomas, H. S. Chan, E. I. Shaknovich, and K. A. Dill. A test of lattice protein folding algorithms. *Proc. Nat. Acad. Sci. USA*, **92**, 325 (1995).

[212] J. Yuen-Zhou and A. Aspuru-Guzik. Quantum process tomography of excitonic dimers from two-dimensional electronic spectroscopy. i. general theory and application to homodimers. *J. Chem. Phys.*, **134**, 134505 (2011).

[213] M. Zein and R. Winter. Effect of temperature, pressure and lipid acyl chain length on the structure and phase behaviour of phospholipidgramicidin bilayers. *Phys. Chem. Chem. Phys.*, **2**, 4545 (2000).

[214] T. Zhang, I. Kuznetsova, T. Meier, X. Li, R. P. Mirin, P. Thomas, and S. T. Cundiff. Polarization-dependent optical 2D fourier transform spectroscopy of semiconductors. *Proc. Natl. Acad. Sci. USA.*, **104**, 14227 (2007).

[215] Y. Zhao. Reexamination of the quantum adiabatic theorem. *Phys. Rev. A.*, **77**, 032109 (2008).

[216] M. Znidaric and M. Horvat. Exponential complexity of an adiabatic algorithm for an NP-complete problem. *Phys. Rev. A*, **73**, 0223295 (2006).