Coherent Control for Quantum Information Processing

Gaurav Bhole

St Hilda's College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Michaelmas Term, 2020

Abstract

Coherent Control for Quantum Information Processing

A thesis submitted for the degree of *Doctor of Philosophy*

Gaurav Bhole
St Hilda's College, Oxford
Michaelmas Term 2020

An exquisite control over the dynamics of a quantum system is essential for realizing any quantum technology. This thesis presents some practical and efficient strategies for coherent quantum control: a paradigm for controlling quantum dynamics using electromagnetic fields. Coherent control is a common framework for describing the physical implementation of quantum logic gates in several quantum information processing platforms including superconducting qubits and trapped ions. For the purpose of demonstration, however, liquid-state Nuclear Magnetic Resonance (NMR) is used in this thesis as it is an ideal test-bed for developing, testing and benchmarking coherent control tools and techniques.

Coherent control was revitalised in 2005 by the development of Gradient Ascent Pulse Engineering (GRAPE), a powerful optimal control technique that is widely used for designing shaped pulses for a variety of tasks. GRAPE, like any other closed-loop optimal control algorithm, requires multiple evaluations of the computationally expensive matrix exponential and the full time-evolution propagator. To this end, I present strategies to sidestep the explicit evaluation of matrix exponentials, thereby offering substantial speed-ups over conventional implementations of the GRAPE algorithm. The results are demonstrated experimentally in NMR, but also in simulations of ultracold molecules and superconducting qubits.

Beyond GRAPE, this thesis also considers methods for designing spin-echo sequences for implementing any arbitrary network of controlled-Z gates in a system of qubits with zz-couplings. First, a method based on linear-programming provides near time-optimal sequences in fully-coupled systems having up to a few hundred qubits. Further, an analytic method based on graph colouring finds near time-optimal spin echo sequences in engineered systems with any number of qubits having only nearest and next-nearest neighbour-couplings.

Acknowledgements

This DPhil has been a truly transformational experience for me and it would not have been possible without the support and guidance that I received from my supervisor, collaborators, friends and family. Firstly, I would like to take this opportunity to extend my gratitude and formally thank my supervisor, Jonathan Jones, who has not only been my research supervisor, but also a mentor, counsel and constant support in various aspects throughout my DPhil. His invaluable feedback along with meticulous guidance will be dearly missed. I would like to thank T. S. Mahesh for being instrumental in fostering my interest in the field of quantum control and quantum computing early on in my undergraduate years at IISER-Pune; it is due to the advice and varied learning opportunities provided by him that made the transition from undergraduate to DPhil effortless.

My thesis is mainly a result of several fruitful collaborations, and hence, I want to thank all the collaborators I have had the pleasure to work with: (1) Vlatko Vedral, Chiara Marletto and Maria Violaris for all the interesting yet perplexing theory proposals that kept on pushing me to devise new and better techniques for implementing them experimentally; (2) energetic and ever enthusiastic Andrew Baldwin for introducing me to protein NMR and being patient while I learned this topic outside my field of research, Charles Buchanan and Gogulan Karunanithy for helping with countless protein NMR experiments; (3) Peter Leek and Takahiro Tsunoda for all the simulating discussions which led to the unanticipated work in Hamiltonian engineering; (4) Dieter Jaksch, Jordi Mur-Petit and Michael Hughes for helping me broaden my horizons with the work on polar molecules.

I am thankful for my flatmates in Oxford, who have been like my family away from home and have been so friendly and kind. I am deeply indebted to my family members and friends who have always believed in me and encouraged me to pursue this endeavour. I want to say a heartfelt thank you to them for their unconditional love and support during these challenging three years and helping me unleash my full potential.

Last but not the least, I want to thank the Felix Trust for their generous financial support throughout my DPhil, which made this journey possible. While I have been working on my thesis, the world is being grappled with the pandemic affecting us in many ways. I would like to thank the University of Oxford for providing me with the relief fund in such unprecedented times.

Table of Contents

Chapter 1

Introduction

The second quantum revolution [1] we are currently undergoing promises quantum technologies [2] that have unprecedented capabilities in a wide range of areas like information processing [3–5], communication [6, 7], cryptography [8–12], simulation [13–15], imaging [16–18], sensing [19, 20] and metrology [21–23]. Much of the success of these quantum technologies is reliant on our ability to actively control and manipulate the dynamics of individual quantum systems. In other words, quantum control is a critical enabler of these emerging quantum technologies [1, 24–27]. My thesis is primarily concerned with developing quantum control methods for near-term quantum computers [28]. To begin, I provide an abstract description of a quantum computer and describe the role of quantum control that forms the core of this thesis.

1.1 Quantum computers and quantum control

A quantum computer is a physical device whose operational principles are based on the laws of quantum mechanics [3, 29, 30]. A quantum computer processes quantum information, *i.e.* it makes explicit use of quantum mechanical properties like interference, superposition and entanglement for encoding and manipulating information [4]. In a quantum computer, information is encoded in the quantum states of a two-level quantum system referred to as

a qubit [31]. Unlike a classical bit which is confined to only binary states, a qubit can exist in a superposition of the binary states, which provide an orthonormal basis. The state of a qubit can be represented by a state vector

$$|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle \quad \text{where,} \quad |\alpha|^2 + |\beta|^2 = 1 \quad \text{and} \quad \{\alpha, \beta\} \in \mathbb{C}. \tag{1.1}$$

Similarly, a system of n coupled qubits can exist in a superposition of 2^n states at the same time, and can be represented by a state vector with 2^n normalized complex amplitudes. By transforming these quantum states in a controlled and reversible manner, a quantum computer can manipulate the information encoded in the qubits [32]. Most commonly, such a transformation of quantum states is achieved through a sequence of discrete unitary operators $U_k U_{k-1} \ldots U_3 U_2 U_1$, each operating on a few selected qubits at a time. These unitary operators U_i are referred to as quantum logic gates or quantum gates [33].

The ultimate purpose of a quantum computer is to realize any arbitrary quantum computation, which, at the fundamental level is nothing but a manipulation of quantum information using quantum gates [33]. Quantum gates therefore serve as the building blocks of a quantum computation. Further, a finite number of quantum gates that form a universal set [34–36] suffices to efficiently realize *any* quantum computation [37, 38]. Thus, an ability to implement a universal set of quantum gates is undoubtedly one of the necessary requirements for building a quantum computer [39].

The framework of coherent quantum control [40–42] is a common paradigm for describing the physical implementation of quantum gates. The idea here is to use coherent light or other electromagnetic fields to control and steer the unitary evolution of the qubits towards the desired quantum gate [43, 44]. To elaborate, consider a system of qubits defined by an internal Hamiltonian $\mathcal{H}_0$ interacting with coherent control fields described by a time-dependent control Hamiltonian $\mathcal{H}_c(t)$ for $t \in [0, T]$. Assuming a closed quantum system, the time-evolution of these qubits takes place in accordance with the Schrödinger equation, and

can be formally represented by a unitary propagator

$$U(T) = \mathcal{T} \cdot \exp\left(-\frac{\mathrm{i}}{\hbar}\int_0^T \left[\mathcal{H}_0 + \mathcal{H}_c(t)\right]\,\mathrm{d}t\right), \tag{1.2}$$

where, $\mathcal{T}$ is the Dyson time-ordering operator. The problem of implementing a desired quantum gate U_D within the framework of coherent quantum control can be described as finding appropriate control fields $\mathcal{H}_c(t)$ that causes the propagator $U(T)$ to resemble U_D as accurately as possible—in the minimum possible time, using minimal resources, and in the presence of experimental imperfections. The main goal of my thesis is to develop practical and efficient quantum control methods for solving this problem of implementing quantum gates.

1.2 Liquid-state NMR: A testbed for quantum control

Several candidate platforms for the physical implementation of a quantum computer have been proposed and developed in the last 25 years [45, 46], the prominent ones being trapped ions [47–51], liquid-state NMR [52–56], superconducting circuits [57–64], linear optics [65–67] and spins in solid state systems [68–79]. Out of these, trapped ions (1995) and liquid-state NMR (1997) were amongst the earliest proposals. While the first ever experimental realization of a two-qubit quantum gate (CNOT) with cold trapped ions was accomplished in 1995 [80], it was not until 2003 that a complete implementation of the Deutsch's algorithm [81] was achieved [82]. On the other hand, by 2001 NMR already had the first demonstrations of several quantum algorithms (Deutsch [83, 84], Deutsch–Josza [85, 86], Grover [87–89], Shor [90]), quantum error-correction [91, 92] and a wide variety of quantum information processing (QIP) tasks [93–102] to its name. What could be the reason behind this rapid initial progress of NMR compared to the other proposals?

The answer lies in the ease with which arbitrary unitary transformations or quantum gates can be implemented in NMR. Established originally as a spectroscopic tool for analytic

chemistry, NMR has a long and illustrious history of controlling nuclear spins using radio-frequency (RF) pulses [103, 104]. Even before quantum information matured as an independent field, chemical and biochemical NMR experiments routinely used precisely sequenced and tailored RF pulses to manipulate hundreds of nuclear spins in organic molecules [105–108]. Some of the control techniques involved shaped pulses [109–112], composite pulses [113–115], refocussing sequences [116–118], average Hamiltonians [119] *etc.*, which when used in these 'conventional' NMR experiments were in essence strategies to realize arbitrary unitary transformations [120, 121]. For example, the NMR experiments involving selective population transfer [122–125] performed in the 1970s in fact correspond to implementing a specific instance of the CNOT gate [126], while other conventional NMR experiments like COSY [105] and INEPT [127] can be seen in retrospect as QIP tasks. Such prior knowledge and experience in control techniques was instrumental in providing a head start to NMR-QIP. The other contributing factor to this success was the ready availability of molecules with coupled spin-1/2 nuclei as 'natural' qubits, along with the sophisticated control hardware that was developed over the decades because of the highly demanding use of NMR in chemical and biochemical fields. As a result, it was possible to focus exclusively on algorithmic-level ideas of coherent control rather than worrying too much about the hardware, thus establishing NMR as a simple 'plug and play' device for implementing ideas from QIP.

NMR, however, is a non-scalable platform for building a quantum computer—a fact which researchers knew ever since it was first proposed in 1997—rendering it unviable for doing any useful QIP. So then, why did the researchers even bother doing QIP with NMR? The primary reason for those early proof-of-principle NMR-QIP demonstrations was to establish the viability of building quantum computers using other scalable platforms. Meaning, if it were not possible to implement such 'toy problems' on a plug and play device like NMR, then the chances with other platforms are definitely bleak. In this manner, NMR gave a head start on realizing and solving the coherent control problems that one could possibly face with other platforms. In fact, many ideas of coherent control that are used in several of the today's leading quantum computer platforms have their origins in NMR. See Ref. [128] for some examples of how control techniques developed in NMR have been extended to

trapped-ions and superconducting circuits.

These aforementioned characteristics of NMR make a compelling case for using it as testbed for developing and benchmarking new quantum control methods for QIP [129]. Consequently, the quantum control methods and their experimental demonstrations in the subsequent chapters of my thesis are presented in the context of NMR-QIP. The basic principles of NMR can be found in several standard textbooks [130–135] while its application to QIP has been extensively reviewed [136–148]. For a review focused on quantum control for NMR-QIP see Ref. [121].

1.3 Thesis outline

Methods for quantum control are described in Chapters 2 to 5, while their applications in various quantum architectures are presented in Chapters 6 and 7. Here, I provide a brief outline of each chapter.

- Chapter 2: Gradient Ascent Pulse Engineering (GRAPE) is an efficient gradient-based optimal control method that is widely used to derive near time-optimal control fields for implementing quantum gates. This chapter provides an overview of the GRAPE algorithm.

- Chapter 3: Numerical evaluation of matrix exponentials, usually performed using Padé approximants or eigendecomposition, is one of the computational bottle-necks of the GRAPE algorithm. This chapter proposes strategies to avoid the explicit evaluation of matrix exponentials, thereby providing a pragmatic approach for implementing the GRAPE algorithm by reducing the computational time required to derive the optimal control fields. Part of this chapter is published in Ref. [149].

- Chapter 4: Spin-qubits in liquid-state NMR and certain implementations of superconducting qubits have an 'always-on' $\sigma_z \otimes \sigma_z$ type coupling between the qubits. In

such systems, spin-echoes are a powerful control tool for realizing a controlled-Z (CZ) gate. In this chapter, a method based on linear programming is proposed to obtain spin-echo sequences for an efficient implementation of any network of CZ gates. This method provides time-optimal sequences in fully coupled systems with tens of qubits and near-time-optimal sequences in fully coupled systems with up to a few hundred qubits. This chapter is closely based on work published in Ref. [150].

- Chapter 5: Linear programming based methods from the previous chapter are computationally infeasible for fully coupled systems beyond a few hundred qubits. Moreover, engineered systems like superconducting qubits rarely have a fully coupled array of qubits—the couplings are often restricted to the nearest or next-nearest neighbours. For such systems, this chapter proposes an analytic method based on graph colouring to design spin-echo sequences that can efficiently implement any network of CZ gates. The proposed method provides near-time-optimal sequences in such locally connected systems with practically any number of qubits. This chapter is closely based on work published in Ref. [151].

- Chapter 6: This chapter presents two different quantum simulations performed on a liquid-state NMR quantum computer. 'Practical' GRAPE methods are used to derive the control fields for implementing quantum gates required for the simulation. To ensure a high-fidelity implementation, the quantum gates are made robust to the unwanted couplings of qubits with the environment as well as uncertainties due to miscalibration of control fields. One simulation which involves the study of entanglement swapping is published in Ref. [152], while the simulation of a quantum homogenizer is published as Ref. [153].

- Chapter 7: Qubits in engineered platforms like polar molecules and superconducting quantum circuits are often embedded in a multi-level system. To ensure a high fidelity implementation of the quantum gates in such systems, it becomes necessary to restrict the effect of control fields to only certain energy levels, while making them robust to several experimental errors at the same time. To this end, use of the GRAPE

algorithm is demonstrated for implementing high fidelity quantum gates. Part of the chapter that concerns polar molecules is published in Ref. [154].

Chapter 2

GRAPE

Gradient Ascent Pulse Engineering (GRAPE) [155] is a powerful numerical method for deriving solutions to the coherent quantum control problem described in the previous chapter. Originally devised for applications in NMR spectroscopy, GRAPE soon became a popular method for implementing robust quantum gates in NMR [156–179] as well as in other platforms like superconducting qubits [180–183], trapped ions [184–186] and nitrogen-vacancy centers in diamond [187–191]. In this chapter, the description of GRAPE will be confined to implementing quantum gates in NMR and so to provide a background and develop the necessary terminology, I first present the coherent control problem in the context of NMR.

2.1 NMR Hamiltonians

2.1.1 Internal Hamiltonian

The qubits in an NMR quantum computer are given by the energy eigenstates of spin-1/2 nuclei (usually ^{1}H, ^{13}C, ^{15}N, ^{19}F, ^{31}P) placed in a static magnetic field [140]. Consider an isolated atomic nucleus with spin quantum number $I = 1/2$ and an intrinsic angular momentum $\vec{I}$ of magnitude $\hbar\sqrt{I(I+1)}$. Associated with this spin is an intrinsic magnetic moment $\vec{\mu} = \gamma\vec{I}$, which depends on the gyromagnetic ratio γ, a characteristic constant for

each nuclear species. When placed in a static magnetic field of strength B (usually taken along the z-axis), the magnetic moment interacts with this magnetic field to produce a Zeeman splitting of energy $\Delta E = \hbar\gamma B$ between the two otherwise degenerate spin states (aligned or anti-aligned with the field). These two energy eigenstates define a qubit, which can be described by an internal Hamiltonian,

$$\mathcal{H}_0 = \hbar\gamma B I_z = \hbar\omega I_z, \tag{2.1}$$

where,

- $I_z = \sigma_z/2$ is the spin angular momentum operator along the z-axis (as per the NMR product operator formalism [192]),
- $\omega/2\pi$ is the Larmor frequency in Hertz.

Typically, B lies in the range of 2.3 to 21.1 Tesla, which corresponds to a Larmor frequency of about 100 MHz to 900 MHz for ^{1}H nuclei [137]. By applying oscillating magnetic fields at this frequency, which lies in the RF region of the spectrum, transitions (Rabi oscillations) between the Zeeman levels can be induced [193, 194]. Spin-1/2 nuclei are natural qubits, *i.e.* their Hilbert space exactly corresponds to that of a qubit, and hence form ideal qubits.

The spin-1/2 nuclei do not exist in isolation: they reside in a molecule along with other nuclei having arbitrary spin numbers. Thus a molecule having n distinct spin-1/2 nuclei can constitute an n-qubit quantum computer. Further, there are about 10^{17} to 10^{19} such identical molecules present in a typical NMR sample, as the emission from a single nucleus in the RF region has an extremely low energy (few μeV) and hence is impractical to detect directly. Thus, NMR is an ensemble quantum computer, in which all molecules are identical quantum computers working in unison.

The qubits in NMR can be identified in frequency space. Molecules in which each spin-1/2 nucleus belongs to a different species will have a distinct γ, and hence a distinct Larmor frequency. Such spin systems are referred to as heteronuclear. Molecules in which all spin-

1/2 nuclei belong to the same nuclear species, known as homonuclear spin systems, can also have distinct Larmor frequencies due to an effect known as the chemical shift [195]. In such molecules, the strength of the magnetic field experienced by every nucleus is different from the applied magnetic field due to a distinct chemical environment (in the absence of symmetry in the molecule), leading to distinct Larmor frequencies for every nucleus therein. From now on, I will specialize to the case of homonuclear spin systems. The internal Hamiltonian for a homonuclear spin system with n-uncoupled nuclei can be expressed as,

$$\mathcal{H}_0 = \sum_{i=1}^{n} \gamma(1-\delta_i)BI_{iz} = \sum_{i=1}^{n} \omega_i I_{iz}, \tag{2.2}$$

where,

- δ_i is the shielding experienced by the i^{th} spin,
- I_{iz} indicates $\sigma_z/2$ on the i^{th} spin,
- $\omega_i/2\pi$ is the effective Larmor frequency of the i^{th} spin,
- $\hbar = 1$ as we will be working in natural units from now on.

For ^{1}H, the shielding constant is typically of the order of 10 ppm (parts per million). At $B_0 = 14$ Tesla, this corresponds to an effective Larmor frequency in the range 600 MHz $\pm$ 6 kHz. In general, δ is a tensor as it depends on the spatial orientation of the molecule with respect to the applied magnetic field. However, in liquid state NMR the rapid molecular tumbling results in a single average (isotropic) value for δ, and it can thus be treated as a scalar.

Nuclear spins interact with each other by two mechanisms. The dominant source of interaction between spins in liquid-state NMR is known as the J-coupling [196, 197], which is mediated by electrons in chemical bonds between atoms [198]. In general, the J-coupling is a tensor and depends on the orientation of the molecule with respect to the applied magnetic field. However, due to rapid molecular tumbling in isotropic liquids, the J-coupling reduces

to a scalar and is hence also known as scalar coupling in liquid-state NMR. As the J-coupling is mediated through local interactions, it only occurs between nuclei within a molecule and not between nuclei in different molecules. The other type of interaction, known as the dipolar coupling, is due to the dipole–dipole coupling between nuclei. Like the J-coupling, the dipolar coupling also depends on the molecular orientation, but unlike J-coupling it is purely anisotropic, and so it is largely averaged away in liquids due to the rapid molecular tumbling, and its effects are usually only observable as a decoherence mechanism.

The scalar J-coupling has the form of a Heisenberg interaction [198], which allows us to write the internal Hamiltonian for a homonuclear spin system with n-coupled nuclei as,

$$\mathcal{H}_0 = \sum_{i=1}^{n} \omega_i I_{iz} + \sum_{i<j} 2\pi J_{ij} \left(\mathbf{I}_i \cdot \mathbf{I}_j\right) \tag{2.3}$$

$$= \sum_{i=1}^{n} \omega_i I_{iz} + \sum_{i<j} 2\pi J_{ij} \left(I_{ix}I_{jx} + I_{iy}I_{jy} + I_{iz}I_{jz}\right), \tag{2.4}$$

where J_{ij} is the strength of the J-coupling between spins i and j, and I_{ix}, I_{iy} are similar to I_{iz} defined earlier. Typically the strength of the J-coupling varies from a few Hz to a few hundred Hz. If the difference between the Larmor frequencies of the coupled nuclei is much greater than the coupling strength ($|\omega_i - \omega_j| \gg 2\pi|J_{ij}|$), then the off-diagonal terms $I_{ix}I_{jx}$ and $I_{iy}I_{jy}$ in the Hamiltonian can be ignored using first-order perturbation theory. This truncates the Heisenberg interaction to an Ising form,

$$\mathcal{H}_0 = \sum_{i=1}^{n} \omega_i I_{iz} + \sum_{i<j} 2\pi J_{ij} I_{iz} I_{jz}, \tag{2.5}$$

which makes the internal Hamiltonian purely diagonal. This approximation, known as the weak coupling approximation, is satisfied by all heteronuclear systems and a majority of the homonuclear systems chosen for QIP.

Unlike many platforms, the internal Hamiltonian of NMR is very well understood and the parameters (shielding values and J-couplings) can be accurately characterized. Moreover, the Larmor frequencies are virtually constant, thanks to the ability of modern NMR spec-

trometers to produce a highly stable and homogenous magnetic field. Also, the J-coupling is constant as the interaction between nuclei is determined by inherent molecular properties. A quantum system with such a robust and constant internal Hamiltonian is extremely useful for developing quantum controls, as one can accurately characterize the time-evolution and focus exclusively on the control part without worrying about any deviations in the internal Hamiltonian. Furthermore, NMR spin systems are an excellent manifestation of a closed quantum system as the spin-1/2 nuclei are embedded deep inside a molecule, effectively isolating them from the environment.

2.1.2 Control Hamiltonian

The state of a spin-1/2 nucleus can be manipulated by applying a linearly oscillating RF magnetic field in the transverse plane close to its Larmor frequency. In heteronuclear spin systems where the Larmor frequencies of various nuclei are separated by tens or hundreds of MHz, each nucleus can be addressed using a separate RF field that is resonant with its Larmor frequency. However, in homonuclear spin systems it is possible to use only a single RF field to address all the nuclei as they are separated by only a few kHz. This makes it difficult to selectively address individual nuclei in a homonuclear system, and hence, manipulating a homonuclear system is much harder than a heteronuclear system. The RF fields applied to a homonuclear spin system with n-coupled nuclei can be described by a control Hamiltonian,

$$\mathcal{H}_c(t) = 2A\cos\left(\omega_{\mathrm{rf}}t + \phi\right) F_x, \tag{2.6}$$

where,

- A, ϕ, and ω_{rf} are the amplitude, phase and frequency of the RF field respectively,
- $F_x = \sum_{i=1}^{n} I_{ix}$ is the total angular momentum operator along the x-axis, through which the RF fields couple with the nuclei, and similar for F_y and F_z (used later).

Note that A and ϕ, known as 'controls', are in general time dependent and can be tuned by the experimenter. However, the notation showing their explicit time dependence *i.e.* $A(t)$ and $\phi(t)$, is dropped for brevity. The amplitude range can typically go up to a few tens of kHz, while the phase usually varies between 0 and 2π. Under the influence of such RF fields, the dynamics of the spin system can be described by a net time-dependent Hamiltonian,

$$\mathcal{H}(t) = \mathcal{H}_0 + \mathcal{H}_c(t), \tag{2.7}$$

where $\mathcal{H}_0$ is the internal Hamiltonian as defined in Eqn. 2.5. Evaluating the dynamics in the lab frame along with such amplitude and phase modulated RF fields is very difficult. The dynamics can be much simplified by going into a frame which rotates at a frequency ω_{rf} about the z-axis. Such a rotating frame transformation when applied to $\mathcal{H}_0$ gives,

$$\widetilde{\mathcal{H}}_0 = \mathrm{e}^{\mathrm{i}\omega_{\mathrm{rf}}tF_z}\mathcal{H}_0\mathrm{e}^{-\mathrm{i}\omega_{\mathrm{rf}}tF_z} = \sum_{i=1}^{n}\Omega_i I_{iz} + \sum_{i<j} 2\pi J_{ij} I_{iz} I_{jz}, \tag{2.8}$$

where, $\Omega_i = \omega_i - \omega_{\mathrm{rf}}$ is defined as the resonance-offset of spin i. Note that the J-coupling strengths remain unchanged after the rotating frame transformation and this result holds true even for Heisenberg type couplings because of the axial symmetry of the Hamiltonian. Similarly, the rotating frame transformation when applied to $\mathcal{H}_c(t)$ gives,

$$\begin{aligned}\widetilde{\mathcal{H}}_c(t) &= \mathrm{e}^{\mathrm{i}\omega_{\mathrm{rf}}tF_z}\mathcal{H}_c(t)\mathrm{e}^{-\mathrm{i}\omega_{\mathrm{rf}}tF_z} \\ &= A\left[\cos\left(\phi\right)F_x + \sin\left(\phi\right)F_y\right] + A\left[\cos\left(2\omega_{\mathrm{rf}}t+\phi\right)F_x - \sin\left(2\omega_{\mathrm{rf}}t+\phi\right)F_y\right].\end{aligned} \tag{2.9}$$

The RF field in the rotating frame Hamiltonian $\widetilde{\mathcal{H}}_c(t)$ consist of two components: a static field applied at an angle ϕ from the x-axis in the xy-plane, and a fast oscillating field in the xy-plane at a frequency $2\omega_{\mathrm{rf}}$. The later component can be neglected as the average contribution of the rapidly oscillating terms is approximately zero. This is known as the rotating wave approximation. Although the fast oscillating terms produce a small shift in the Larmor frequencies of the spins, known as the Bloch–Siegert shift [199], as long as the first order correction $A^2/4\omega_i$ is small compared to ω_i, the rotating wave approximation holds

well. With this, the control Hamiltonian in the rotating frame can now be approximated as,

$$\widetilde{\mathcal{H}}_c(t) = A\left[\cos(\phi)\, F_x + \sin(\phi)\, F_y\right]. \tag{2.10}$$

Note that this control Hamiltonian can also be written in the Cartesian form where $\omega_x = A\cos(\phi)$ and $\omega_y = A\sin(\phi)$ describe the time-dependent amplitudes of the RF field in the x and y quadratures respectively. The dynamics of the system in the rotating frame can be described by a time-dependent Hamiltonian,

$$\widetilde{\mathcal{H}}(t) = \widetilde{\mathcal{H}}_0 + \widetilde{\mathcal{H}}_c(t). \tag{2.11}$$

Note that the explicit tilde signs on $\widetilde{\mathcal{H}}_0$ and $\widetilde{\mathcal{H}}_c(t)$ will be omitted from now on as we shall be exclusively working in the rotating frame of reference.

2.2 Framework of GRAPE

The time evolution of an NMR spin system under the influence of amplitude and phase modulated RF fields can be formally represented by a propagator

$$U(T) = \mathcal{T}\cdot\exp\left(-\mathrm{i}\int_0^T \left[\mathcal{H}_0 + A\left(\cos(\phi)\, F_x + \sin(\phi)\, F_y\right)\right]\,\mathrm{d}t\right). \tag{2.12}$$

GRAPE aims to find appropriate controls A and ϕ that drive the propagator $U(T)$ close to the desired quantum gate U_D in the minimum possible time while obeying certain optimality criteria and experimental constraints. In this section, I describe the framework used by GRAPE to find these 'optimal' controls.

2.2.1 Optimal control theory

GRAPE uses optimal control [200–204] — a popular strategy for deriving solutions to the coherent control problem. Optimal control theory (OCT) can be viewed as an extension of calculus of variations [205] to systems involving dynamical constraints [206] (Schrödinger's equation in our case). The aim of OCT is to find a control law for the dynamical system that achieves a certain optimality criterion. In the framework of OCT, this task is formulated as an optimization problem that requires finding the extremum of a given performance index. In our case, this performance index can be defined as a scalar, real-valued cost functional [207]

$$\mathcal{J} \coloneqq \mathcal{F}(U(T)) + \int_0^T \mathcal{L}(U(t), A, \phi, t)\ \mathrm{d}t, \tag{2.13}$$

where $\mathcal{F}$ is the terminal cost that determines how close the system evolution is to the target unitary and $\mathcal{L}$ is the running cost which defines the constraints on the system and controls. The terminal cost, as its name suggests, depends only on the final state of the system and is independent of the control amplitude and phases, while the running cost on the other hand depends on the actual path taken by the system in the time interval $[0, T]$. The optimization problem can be described as finding appropriate controls A and ϕ that maximize (or minimize) the cost functional $\mathcal{J}$ while obeying the dynamical constraints and pre-defined optimality criterion. As quantum systems are sensitive to noise and decoherence, it is often necessary to seek time-optimal solutions, *i.e.*, controls which optimize $\mathcal{J}$ for a minimum value of T.

The necessary conditions for finding the optimal controls are described by a set of equations given by Pontryagin's principle [201]. In low-dimensional systems and special cases, geometric or analytical methods [208–210] can be employed for solving these equations to obtain global optimal controls; see [211–218] for some examples. However, in most cases one needs to resort to numerical optimization algorithms [219] due to the complexity of the control landscape and high-dimensionality of the system. GRAPE is one such numerical method that performs a gradient-driven optimization in an elegant manner. For examples of other

numerical methods for optimal control see [44, 220–236]. In all these numerical methods, including GRAPE, it is usually possible to obtain only locally optimal solutions.

2.2.2 Controllability

The first relevant question in OCT is about the existence of solutions. This question has to do with the controllability of the system [42]: what set of unitary operations can be realized with the given internal Hamiltonian $\mathcal{H}_0$ and control Hamiltonian $\mathcal{H}_c(t)$. Formally, this question is answered by the Lie algebra generated by $\mathcal{H}_0$, $\mathcal{H}_c(t)$ and their commutators [208, 237–240]. For a system to be fully controllable, *i.e.*, to be able to realize any unitary operation, it is necessary that $\mathcal{H}_0$, $\mathcal{H}_c(t)$ and their commutators generate the entire Lie algebra [241]. Thus, the controllability of a system is linked to its ability to implement a universal set of quantum gates [242]. Note that controllability has nothing to do with the optimality of solutions or the total time and resources needed for realizing a particular unitary; it is merely concerned with the existence of solutions. Informally speaking, any system is fully controllable if every qubit is coupled with every other qubit (directly or indirectly) and there is a provision to realize single-qubit rotations for every qubit about two non-parallel axes (say σ_x and σ_y) [243]. The NMR systems discussed earlier satisfy these criteria and are fully controllable.

2.2.3 Piecewise constant controls

In the majority of optimal control methods, including GRAPE, the control law needs to be evaluated in an open-loop setting without any experimental feedback. This requires an explicit computation of the time-evolution of the quantum system, formally given by the propagator in Eqn. 2.12. As this propagator arises from a continuously varying control Hamiltonian which does not commute at different times, its computation is particularly hard. To simplify this computation, the control Hamiltonian is usually replaced by a piecewise constant Hamiltonian in sufficiently small time intervals. Considering a total of N intervals

or segments of duration Δt each, the propagator in Eqn. 2.12 can thus be approximated as,

$$U(T) = U_N U_{N-1} \dots U_j \dots U_3 U_2 U_1, \tag{2.14}$$

where for the j^{th} segment,

- $U_j = \exp\left[-\mathrm{i}\left(\mathcal{H}_0 + \mathcal{H}_j\right)\Delta t\right]$ is the sub-propagator,
- $\mathcal{H}_j = A_j\left[\cos\left(\phi_j\right)F_x + \sin\left(\phi_j\right)F_y\right]$ is the control Hamiltonian, and
- A_j and ϕ_j is the amplitude and phase respectively.

With this, the computation of the time-evolution of the quantum system reduces to evaluating a product of matrix exponentials. The piecewise constant approximation used here is not an issue as most physical implementations commonly use discretized controls rather than generalized continuous forms. Typically in NMR, the total number of segments N would vary from a few hundred to a few thousand depending on the exact type of quantum gate to be implemented and accordingly, Δt is of the order of a few μs.

2.2.4 Fidelity

The terminal cost $\mathcal{F}$ assesses the quality of controls by comparing the time-evolution $U(T)$ with the desired unitary U_D and summarizing the differences between them using a single real number. To this end, a fidelity [244, 245] measure between the two unitary operators $U(T)$ and U_D is chosen and usually defined as the Euclidean distance,

$$\mathcal{F} \coloneqq \|U_D - U(T)\|^2 = \|U_D\|^2 + \|U(T)\|^2 - 2\,\mathrm{Re}\,\langle U_D|U(T)\rangle\,, \tag{2.15}$$

where $\|A\| = \sqrt{\mathrm{Tr}\{A^\dagger A\}}$ is the Frobenius norm and $\langle A|B\rangle = \mathrm{Tr}\{A^\dagger B\}$ is the Hilbert–Schmidt inner product [5]. For closed quantum systems, $\|U(T)\|$ is always constant since the purity remains the same. Therefore, in closed systems it suffices to consider only the third

term in Eqn. 2.15 for defining the fidelity, which is the overlap between the two unitaries[†]:

$$\mathcal{F} = |\langle U_D | U(T) \rangle|^2 . \tag{2.16}$$

The square modulus in the above expression ensures that the fidelity metric is independent of any global phase difference between the two unitaries. The fidelity $\mathcal{F}$ is also normalized, usually by the dimension of the underlying Hilbert space. Such a normalized fidelity metric $\mathcal{F}$ can take real values between 0 and 1, where the null and unity correspond to $U(T)$ and U_D being orthogonal and identical respectively.

Of course, it is ideal to seek controls that implement a quantum gate with unity fidelity. However, finding such good controls is an extremely difficult task and even if they are found, errors in their implementation and other unaccounted uncertainties would often result in experimental fidelities that are lower than expected. Fortunately, implementing a quantum algorithm does not require an arbitrarily good control: it is sufficient that the error rate $1-\mathcal{F}$ of the quantum gates is below a so-called fault-tolerant threshold (around 10^{-3}) [247–251], as quantum error correction protocols [252–260] can then take over to suppress the errors to arbitrarily low values with only a polynomial resource overhead. However, this does not mean that no efforts need to be taken to improve the controls once the fault-tolerant threshold is reached. The overhead costs near the fault-tolerant threshold are extremely high for doing practically useful quantum computing applications and meeting them is well outside the reach of our present abilities. For example, a recent estimate shows that for running the Shor algorithm for factoring a 2048-bit number using gates with an error rate of 10^{-3} with currently known methods would require about 20 million physical qubits [261]. In order to bring down such large overhead costs to a more practicable level, it is therefore necessary to aim for controls that promise a significantly higher gate fidelity.

[†]For a non-unitary map, considering only the overlap between two operators as the fidelity definition will give an incorrect fidelity; it is necessary to use the full fidelity definition $\|\kappa_D - \kappa(T)\|^2$, where κ is the map and κ_D, $\kappa(T)$ is the desired and obtained operator respectively. Similarly, in open quantum systems with a non-unitary map, it is necessary to use density matrices (as there can be mixed states and purity can change), and in this case the correct fidelity definition is given by $\|\rho_D - \rho(T)\|^2$, where ρ_D is the density matrix for the desired state and $\rho(T)$ is the final state of the system [246].

2.2.5 Robustness to systematic errors

To ensure high experimental fidelities for quantum gates, the least that one can do is make the control fields robust to systematic errors. This involves accounting for known uncertainties and variations in the Hamiltonian parameters while designing the controls. Consider for example a very common error that occurs when the strength of the generated control field is systematically different to that requested. In such situations, it is desirable to seek a control field that is tolerant to a range of amplitudes around the nominal value. To this end, the control field can be optimized for an average fidelity function that encompasses fidelities from propagators simulated for a range of discrete values of amplitudes around the nominal value [155]. In a similar fashion, the control fields can be made robust to any other parameter or combination of parameters in the internal and control Hamiltonian.

In NMR, the two dominant sources of systematic error are due to inhomogeneities in the RF control field and the static magnetic field across the sample [262, 263]. Usually, a variation of about $\pm 20\%$ around the nominal value of RF amplitudes is observed across the NMR sample tube. Oftentimes, it is also possible to map out the exact distribution of these RF inhomogeneities for an NMR probe [264], in which case the control field can be optimized for an average fidelity function that is appropriately weighted according to the observed distribution instead of a simple average fidelity function as described earlier. It is important to ensure that the control field performs well not only for the specified range of amplitudes for which it is optimized but also for the amplitudes in between. By choosing non-periodic discrete values of amplitudes for optimization, the risk of behaviour where the control field performs poorly for intermediate amplitudes can be reduced [263].

The second source of error due to variations in the static magnetic field can give rise to a range of Larmor frequencies for the nuclei across the sample causing an uncertainty in the resonance offsets. Although this error can be handled in a similar way as above, there is no practical need to do so. In liquid-state NMR, a well shimmed magnet is able to limit the variation in the static field well within a Hz. The magnitude of this error is almost

comparable to the line-widths found in typical NMR samples chosen for QIP, and hence can be ignored.

2.2.6 Penalties

Quite often, the optimized control fields turn out to be experimentally unimplementable because they violate either some physical limits or hardware constraints. Such situations can be avoided by using an appropriate penalty function in the form of a running cost $\mathcal{L}$ during the optimization. The penalty function essentially attempts to restrict the search space, so that the optimization process can search for control parameters only within the pre-defined limits. Consider for example a common scenario where the control amplitudes need to be restricted on account of the maximum amount of power that the hardware can offer. In such a case, the amplitudes exceeding a threshold amplitude $A_{\max}$ can be penalised using a penalty function,

$$\mathcal{L} = \begin{cases} \sum_k |A_{\max} - A_k|^2, & |A_k| > A_{\max} \\ 0, & |A_k| \leq A_{\max} \end{cases} \tag{2.17}$$

It is also possible to modify this penalty function to include a Gaussian or any suitably shaped filter function that enforces control fields to have zero amplitude at the start and end of the sequence [265]. Note that during the optimization of the cost functional $\mathcal{J}$, it is important to balance the contribution of $\mathcal{F}$ and $\mathcal{L}$ by assigning appropriate weights to them. The penalty function needs to be strong enough that the unacceptable amplitudes are discouraged, but at the same time it should not be so severe that it drives away the high fidelity candidate solutions which are near the threshold. In addition to such physical constraints, penalty functions can also be used to produce hardware friendly control fields with a view to improve the experimental fidelity. For example, a smoothing regularisation penalty as described in [266] can be used to reduce sudden amplitude jumps between successive segments of the control sequence and thus mitigate any transient effects [263].

2.3 The GRAPE algorithm

Numerical optimization is an integral part of OCT. Most of the numerical techniques start with an initial randomized guess of controls which then need to be iteratively improved to maximize or minimize the cost function. The cost function needs to be evaluated at least once during every iteration and is a very time intensive computation, especially in the case of quantum systems. It is therefore important to choose a good numerical method that helps reach the maxima or minima using a small number of cost function evaluations. The choice of the numerical method is strongly dependent on various factors like the dimensionality of the problem, nature of the control landscape (convex, non-convex, smooth, noisy), type of controls (analytic, piecewise-constant), constraints, availability of gradients *etc.*

Gradient driven numerical optimization methods are known to have among the best convergence rates, particularly when applied to quantum control problems that involve high-dimensional parameter spaces [236]. However, in many scenarios it is very likely that the gradient information is either computationally expensive or unavailable, compelling one to resort to gradient-free approaches like Nelder–Mead simplex [267] or genetic algorithms [268–270]. This was also the case prior to GRAPE, when gradient based methods for gate design problems were largely ruled out due to the high cost of evaluating gradients [225]. For example, the conventional method of calculating approximate gradients using finite-differences method requires at least $\mathcal{O}(N)$ evaluations of the computationally expensive propagator for calculating every gradient [263]. In contrast, the GRAPE algorithm requires only two propagator evaluations for obtaining each gradient, thereby offering an efficient way to perform a gradient based optimization. This allows a larger number of control parameters to be efficiently optimized, thereby offering a high flexibility in designing the control fields.

The key realization in GRAPE is that the gradient of $\mathcal{F}$ with respect to any j^{th} control parameter changes only the j^{th} sub-propagator. For example, the calculation of the gradient

of $\mathcal{F}$ (as in Eqn. 2.16) with respect to the control amplitude A_j proceeds as,

$$\begin{aligned}
\frac{\partial \mathcal{F}}{\partial A_j} &= \frac{\partial}{\partial A_j} \left| \langle U_D | U(T) \rangle \right|^2 \\
&= \frac{\partial}{\partial A_j} \langle U_D | U(T) \rangle \langle U(T) | U_D \rangle \\
&= \frac{\partial}{\partial A_j} \langle U_D | U_N U_{N-1} \dots U_2 U_1 \rangle \langle U_N U_{N-1} \dots U_2 U_1 | U_D \rangle \\
&= \langle U_D | U_N U_{N-1} \dots U_{j+1} \frac{\partial U_j}{\partial A_j} U_{j-1} \dots U_2 U_1 \rangle \langle U(T) | U_D \rangle + \text{c.c} \\
&= 2 \, \text{Re} \left\{ \langle U_D | U_N U_{N-1} \dots U_{j+1} \frac{\partial U_j}{\partial A_j} U_{j-1} \dots U_2 U_1 \rangle \langle U(T) | U_D \rangle \right\}
\end{aligned} \tag{2.18}$$

where the derivative $\partial U_j / \partial A_j$ can be evaluated using standard techniques [155, 265] that are described later. The gradient $\partial \mathcal{F} / \partial \phi_j$ can also be calculated in a similar way. For efficient calculation of these gradients, GRAPE utilizes a well known technique of trading-off computational time with memory. During the calculation of the propagator $U(T)$ for evaluating the fidelity, each of the intermediate results U_1, $U_2 U_1$, $U_3 U_2 U_1$ *etc.* are stored and used for the gradient calculation. A backward evolution in time to store the combined propagators U_N, $U_N U_{N-1}$, $U_N U_{N-1} U_{N-2}$ *etc.* is also done. In this way, GRAPE requires only two full time-evolutions, one forwards and one backwards, for calculating the gradients.

Below, I summarize the basic GRAPE algorithm based on [155] and then discuss some of its extensions along with practicalities associated with the implementation of the algorithm.

1. Start with an initial randomized guess of control amplitudes $\{A_1, A_2, \dots, A_N\}$ and phases $\{\phi_1, \phi_2, \dots, \phi_N\}$.

2. For each segment $j = 1, 2, \dots N$, evaluate the sub-propagators using:

$$U_j = \exp \left[-\text{i} \left(\mathcal{H}_0 + A_j \left[\cos(\phi_j) F_x + \sin(\phi_j) F_y \right] \right) \Delta t \right]. \tag{2.19}$$

3. Calculate and store each forward propagator

$$F_j = U_j U_{j-1} \dots U_2 U_1, \quad \text{for} \quad j = 0, 1, \dots N, \quad \text{with} \quad F_0 = \mathbb{I}. \tag{2.20}$$

Similarly, calculate and store each backward propagator

$$B_j = U_N U_{N-1} \dots U_{j+1}, \quad \text{for} \quad j = 1, 2, \dots N, \quad \text{with} \quad B_N = \mathbb{I}. \tag{2.21}$$

4. Using $U(T) = F_N$, compute the fidelity function

$$\mathcal{F} = |\langle U_D | U(T) \rangle|^2 . \tag{2.22}$$

5. For $j = 1, 2, \dots N$, compute the gradients

$$\frac{\partial \mathcal{F}}{\partial A_j} = 2 \,\mathrm{Re} \left\{ \langle U_D | B_j \frac{\partial U_j}{\partial A_j} F_{j-1} \rangle \, \langle U(T) | U_D \rangle \right\}, \tag{2.23}$$

$$\frac{\partial \mathcal{F}}{\partial \phi_j} = 2 \,\mathrm{Re} \left\{ \langle U_D | B_j \frac{\partial U_j}{\partial \phi_j} F_{j-1} \rangle \, \langle U(T) | U_D \rangle \right\}. \tag{2.24}$$

where, the forms for $\partial U_j / \partial A_j$ and $\partial U_j / \partial \phi_j$ are described in the next section.

6. Update the controls using a method like steepest ascent, conjugate gradients or BFGS as described in the next section.

7. Iterate steps 2 through 6 until the fidelity reaches a desired threshold.

Note that the desired penalty functions $\mathcal{L}$ can be added as required in Step 4 and accordingly, their corresponding gradients in Step 5 in a fairly straightforward manner.

There are a number of implementations of the GRAPE algorithm available as software packages written in MATLAB (Dynamo [271], Spinach [272]), Python (QuTiP [273, 274]) and C [275]. However, a self-written code developed in MATLAB was used for this thesis. In the following sections, I describe some of the implementation details for various steps of the GRAPE algorithm.

2.3.1 Gradient-based update schemes

In this section, I describe some of the methods that can be used for updating the control parameters in GRAPE (Step 6). For convenience and brevity, I introduce two shorthand notations for the controls and their gradients as follows:

$$\mathbf{c} \coloneqq [A_1 \dots A_j \dots A_N \;\; \phi_1 \dots \phi_j \dots \phi_N]^\intercal \tag{2.25}$$

$$\mathbf{g} \coloneqq \left[\frac{\partial \mathcal{F}}{\partial A_1} \cdots \frac{\partial \mathcal{F}}{\partial A_j} \cdots \frac{\partial \mathcal{F}}{\partial A_N} \;\; \frac{\partial \mathcal{F}}{\partial \phi_1} \cdots \frac{\partial \mathcal{F}}{\partial \phi_j} \cdots \frac{\partial \mathcal{F}}{\partial \phi_N}\right]^\intercal . \tag{2.26}$$

Steepest ascent

The simplest way to update the controls is by moving them in small steps in the direction of the gradient, which is the direction of steepest ascent. The update rule can be written as

$$\mathbf{c}_{r+1} = \mathbf{c}_r + \epsilon \mathbf{g}_r \tag{2.27}$$

where r is the iteration index and ϵ is the step size or the search length. Provided the step size is sufficiently small, it can be a constant for every iteration. However, in many scenarios much better performance can be achieved if the step size is continually adjusted and optimized. The idea here is to move the controls to a point (along the gradient direction) where the fidelity improvement is maximum, before calculating the next gradient direction. Of course, evaluating the fidelity at every point along the gradient is simply infeasible due the complexity of fidelity calculation. Nevertheless, it suffices to determine a step size that can produce even a local maximum fidelity. One simple and effective approach here is to find the fidelity of two more points along the gradient direction, fit a quadratic polynomial through the three points, and locate the maximum of the quadratic to obtain the optimum step size [271, 276]. This approach, however, will allow a fast convergence only when near a local extrema where the fidelity curve, expanded as a Taylor series, can be approximated by a quadratic. In case the quadratic interpolation approach fails for a particular iteration, it is

possible to fall back to other simpler line search approaches like backtracking or bracketing and sectioning as described in [266].

Conjugate gradients

The method of steepest ascent produces a slow convergence in many practical cases, especially when the fidelity landscape has valleys and ridges. If the gradient direction is not pointing along the crest of a ridge, the steepest ascent method will lead to a zig-zag back and forth pattern across the crest of the ridge, until it slowly reaches the crest. This behaviour indicates that the direction of the steepest ascent is not always the best direction to proceed in, especially in the long run. To this end, the method of conjugate gradients (CG) [277, 278] can be used to rectify such behaviour and provide an accelerated convergence near a maximum. The CG method tries to estimate the curvature of the fidelity landscape by including the previous search directions along with the current gradient to come up with a new better search direction. Thus, by keeping a 'memory' of previous directions, the CG method can quickly determine that the best direction is simply along the ridge and not across the ridge.

The CG method is elaborately described in Ref. [279]. Hence, I present only the relevant implementation details here. Assuming the initial search direction $\boldsymbol{\lambda}_0$ is the same as in gradient ascent, the subsequent search directions $\boldsymbol{\lambda}_r$ in the CG method can be computed using:

$$\boldsymbol{\lambda}_r = \mathbf{g}_r + \beta_r \boldsymbol{\lambda}_{r-1}, \quad \beta_r = \frac{\mathbf{g}_r^\intercal(\mathbf{g}_r - \mathbf{g}_{r-1})}{\mathbf{g}_{r-1}^\intercal \mathbf{g}_{r-1}}. \tag{2.28}$$

where the above form of β corresponds to the Polak–Ribière formula [279]. Some other popular forms include that of Fletcher–Reeves [280], Hestenes–Stiefel [277] and Dai–Yuan [281]. With this, the update rule can be written as

$$\mathbf{c}_{r+1} = \mathbf{c}_r + \epsilon \boldsymbol{\lambda}_r. \tag{2.29}$$

Note that for $\beta = 0$, the CG method is essentially the same as the steepest ascent method.

For CG to perform correctly and to obtain the best convergence, it is essential to first move the controls to the highest point of fidelity along the gradient direction. This can be done using the line search methods described previously. An appropriate implementation of the CG method can significantly reduce the overall calculation time of GRAPE by reducing the number of iterations required to achieve a given fidelity by at least an order of magnitude.

BFGS

The first-order derivatives or gradients provide only local information (slope) about the fidelity landscape and have no knowledge about the non-local features like the curvature. This makes it is difficult to distinguish between stationary points during the search as the slope for any stationary point is zero. To avoid the search getting trapped in such stationary points, information from second-order derivatives (Hessian) can be used. The Hessian matrix has knowledge about the local curvature of the landscape and can thus easily distinguish between, say a maximum and a saddle point. It is worth noting that close to a maximum, the CG method also behaves like a second-order method but without requiring an explicit Hessian. However, far from maxima it has the same behaviour as that of a first-order method.

The standard second-order derivative update, also known as the Newton update, can be written as

$$\mathbf{c}_{r+1} = \mathbf{c}_r + \epsilon \mathbf{h}_r^{-1} \mathbf{g}_r, \tag{2.30}$$

where $\mathbf{h}^{-1}$ denotes the inverse Hessian. Evaluation of the exact Hessian is usually very expensive and hence a quasi-Newton approach namely the Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm [219, 282–285] is often used to obtain an approximate Hessian. The BFGS approach only needs first-order derivative information to approximate the inverse Hessian. Starting from $\mathbf{h}_0 = \mathbb{I}$ (*i.e.* standard gradient ascent), the Hessian matrix after the r^{th} iteration in the BFGS algorithm can be approximated by [271, 286]

$$\mathbf{h}_{r+1}^{-1} = \left[\mathbb{I} - \frac{\Delta\mathbf{c}_r \Delta\mathbf{g}_r^\intercal}{\Delta\mathbf{g}_r^\intercal \Delta\mathbf{c}_r}\right]^\intercal \mathbf{h}_r^{-1} \left[\mathbb{I} - \frac{\Delta\mathbf{c}_r \Delta\mathbf{g}_r^\intercal}{\Delta\mathbf{g}_r^\intercal \Delta\mathbf{c}_r}\right] + \frac{\Delta\mathbf{c}_r \Delta\mathbf{c}_r^\intercal}{\Delta\mathbf{g}_r^\intercal \Delta\mathbf{c}_r} \tag{2.31}$$

where, $\Delta \mathbf{c}_r = \mathbf{c}_{r+1} - \mathbf{c}_r$ and $\Delta \mathbf{g}_r = \mathbf{g}_{r+1} - \mathbf{g}_r$.

The standard MATLAB library function `fminunc` which implements the BFGS algorithm was used in this thesis for updating controls in GRAPE. Oftentimes, due to the large number of control parameters in optimal control, it becomes difficult to store the Hessian matrix. In that case, the limited memory variant of the BFGS algorithm known as *l*-BFGS [287, 288] may be used. *l*-BFGS is also found as a standard MATLAB routine named `fmincon`. The `fmincon` module permits explicit constraints on control variables to be included, but that option was not used for this thesis.

2.3.2 Evaluating gradients

In this section, I present the forms of the gradients $\partial U_j / \partial A_j$ and $\partial U_j / \partial \phi_j$ required in Step 5 of the GRAPE algorithm.

Approximate gradients

Evaluating the derivative of a matrix exponential is a non-trivial task, particularly in the case where the Hamiltonian contains non-commuting terms. Here, I describe how to obtain the approximate value of $\partial U_j / \partial A_j$ (similar calculations for $\partial U_j / \partial \phi_j$ hold). To begin, we use the standard formula for computing the derivative of a matrix exponential [289]:

$$\frac{\partial}{\partial x}\left\{\mathrm{e}^{f(x)}\right\} = \int_0^1 \mathrm{e}^{sf(x)} \frac{\partial f}{\partial x} \mathrm{e}^{(1-s)f(x)} \mathrm{d}s. \tag{2.32}$$

where $f(x)$ is a general matrix function. Accordingly, we get

$$\frac{\partial U_j}{\partial A_j} = -\mathrm{i}\left(\int_0^{\Delta t} U_j(\tau)\left[\cos\left(\phi_j\right) F_x + \sin\left(\phi_j\right) F_y\right] U_j(-\tau)\mathrm{d}\tau\right) U_j, \tag{2.33}$$

where, $U_j(\tau) = \exp[-\mathrm{i}(\mathcal{H}_0 + \mathcal{H}_j)\tau]$. In the limit where $\Delta t \ll ||\mathcal{H}_0 + \mathcal{H}_j||^{-1}$, the unitary $U(\tau)$ can be expanded to first order in τ, giving us,

$$\frac{\partial U_j}{\partial A_j} \approx -\mathrm{i}\left(\int_0^{\Delta t} (\mathbb{I} - \mathrm{i}(\mathcal{H}_0 + \mathcal{H}_j)\tau)\left[\cos(\phi_j) F_x + \sin(\phi_j) F_y\right](\mathbb{I} + \mathrm{i}(\mathcal{H}_0 + \mathcal{H}_j)\tau)\,\mathrm{d}\tau\right) U_j. \tag{2.34}$$

On integrating the above expression and dropping terms in Δt^2, we get,

$$\frac{\partial U_j}{\partial A_j} \approx -\mathrm{i}\Delta t\left[\cos(\phi_j) F_x + \sin(\phi_j) F_y\right] U_j. \tag{2.35}$$

This is a first-order approximation of the gradient. The gradient with respect to phase ϕ_j can be obtained in a similar way:

$$\frac{\partial U_j}{\partial \phi_j} \approx -\mathrm{i}\Delta t A_j\left[-\sin(\phi_j) F_x + \cos(\phi_j) F_y\right] U_j. \tag{2.36}$$

Note that this approximation puts a constraint on the segment duration Δt and the corresponding amplitude during that segment. The first order gradients are only effective when every segment in the GRAPE sequence produces a sufficiently small rotation angle. Another way to put this is that the approximation is valid only when the time evolution U_j approximately commutes with the control operators F_x or F_y.

Exact gradients

To ensure proper functioning and convergence of the update methods, especially the BFGS algorithm, it is necessary to use accurate gradients. Here, I present the method described in Refs. [271, 290] which can be used to obtain the exact value of gradients. Following the theorem in [290, 291], the exact derivative of the exponential of a sum of two non-commuting operators $\mathcal{A}$ and $x\mathcal{B}$ with respect to x at $x = 0$ is given by,

$$\left\langle \xi_l \left| \frac{\partial}{\partial x} \mathrm{e}^{\mathcal{A}+x\mathcal{B}} \right| \xi_m \right\rangle = \begin{cases} \langle \xi_l|\mathcal{B}|\xi_m\rangle\, \mathrm{e}^{\xi_l}, & \text{if } \xi_l = \xi_m \\ \langle \xi_l|\mathcal{B}|\xi_m\rangle\, \frac{\mathrm{e}^{\xi_l} - \mathrm{e}^{\xi_m}}{\xi_l - \xi_m}, & \text{otherwise,} \end{cases} \tag{2.37}$$

where $|\xi_l\rangle$ and ξ_l are the eigenvectors and eigenvalues, respectively, of the operator $\mathcal{A}$: $\mathcal{A}|\xi_l\rangle = \xi_l|\xi_l\rangle$. For a proof of this result refer the appendix of [271]. According to this result we need to calculate:

$$\frac{\partial U_j}{\partial A_j} = \frac{\partial}{\partial x}\exp\left\{-\mathrm{i}\left(\mathcal{H}_0 + (A_j + x)\left[\cos(\phi_j)F_x + \sin(\phi_j)F_y\right]\Delta t\right)\right\}\Big|_{x=0}, \tag{2.38}$$

$$\frac{\partial U_j}{\partial \phi_j} = \frac{\partial}{\partial x}\exp\left\{-\mathrm{i}\left(\mathcal{H}_0 + A_j\left[\cos(\phi_j + x)F_x + \sin(\phi_j + x)F_y\right]\Delta t\right)\right\}\Big|_{x=0}. \tag{2.39}$$

First, we need to mould the exponents in the above equations into the form $\mathcal{A} + x\mathcal{B}$. To this end, we can use $\mathcal{A} = -\mathrm{i}(\mathcal{H}_0 + \mathcal{H}_j)\Delta t$ and

$$\mathcal{B} = -\mathrm{i}\Delta t\left[\cos(\phi_j)F_x + \sin(\phi_j)F_y\right] \tag{2.40}$$

for calculating $\partial U_j/\partial A_j$, and

$$\mathcal{B} = -\mathrm{i}\Delta t A_j\left[-\sin(\phi_j)F_x + \cos(\phi_j)F_y\right] \tag{2.41}$$

for calculating $\partial U_j/\partial \phi_j$. Note that the result $\sin(x) \approx x$ and $\cos(x) \approx 1$ for small x has been used to arrive at the above expression for $\mathcal{B}$. By plugging in appropriate forms of $\mathcal{A}$ and $\mathcal{B}$ in Eqn. 2.37, the exact derivatives of $\partial U_j/\partial A_j$ and $\partial U_j/\partial \phi_j$ can be calculated.

The exact derivatives need the expensive routine of matrix diagonalization to be carried out for obtaining the eigenvalues and eigenvectors of $\mathcal{A}$. However, there is an advantage to this method of calculating exact gradients: the matrix exponential required in calculating U_j essentially comes for free as we already have the eigenvectors and eigenvalues of $\mathcal{A}$. Evaluation of U_j now just needs the diagonal matrix of eigenvalues to be exponentiated, which can be done efficiently. Although matrix diagonalization comes at a slightly higher computational overhead compared with a direct calculation of a matrix exponential, the additional cost is outweighed by the benefits of a faster convergence rate due to exact gradients. Therefore, it is advantageous to use exact gradients.

2.3.3 Propagators, complexity and scaling up

The calculation of the propagator $U(T)$ requires two fundamental operations: matrix exponentiation for calculating the sub-propagators U_j and matrix multiplication for concatenating the U_j to form $U(T)$. Matrix exponentiation is a non-trivial task. In general, it is not possible to evaluate the matrix exponential analytically, except for 2×2 matrices and some very special matrices. Finding efficient numerical strategies for matrix exponentiation has been a subject of extensive research in the past years [292, 293]. Some of the popular numerical methods include those based on Taylor-series approximation, Padé-series approximation, Chebychev polynomials and eigendecomposition. The standard library function `expm` in MATLAB uses Padé-series approximation combined with a scaling and squaring technique [294, 295].

Both operations, matrix exponentiation (using any method) and matrix multiplication for reasonable matrix sizes require about $\mathcal{O}(\mathcal{D}^3)$ basic arithmetic operations, where $\mathcal{D}$ is the dimension of the Hilbert space. The mandatory calculation of the propagator $U(T)$ in every iteration severely slows down the GRAPE algorithm, thus causing a computational bottleneck. Further, the dimension of the Hilbert space grows exponentially with the number of qubits, preventing the GRAPE algorithm from scaling up to a large number of qubits. The combination of these two factors is responsible for restricting the viability of GRAPE to only a handful of qubits on standard desktop computers. The use of cluster computers along with efficient parallel processing can push this limit up to a few tens of qubits [296–299], but, that does not address the fundamental issue of scalability.

One possible solution towards a scalable GRAPE algorithm is described in Ref. [129]. The idea is as follows. In a large-scale practical quantum computer, it is very unlikely that every qubit is coupled to every other qubit; the coupling is often restricted to nearest or next-nearest neighbours. As a result, it is possible to divide the entire system of qubits into smaller (possibly overlapping) sub-systems and optimize the controls for a net fidelity function that is a sum of fidelities restricted to the sub-systems. Accordingly, the desired unitary operation

can also be defined as a tensor product of individual target unitaries across each sub-system. In this manner, the size of the quantum systems to be simulated can be extensively reduced. To ensure that the dominant dynamics of the system are properly captured in this sub-system approximation, it is important to have stronger couplings internal to at least one sub-system. In spite of that, there will be errors due to the evolution of the couplings between subsystems which are not accounted for during the simulation. These errors can be efficiently estimated [129] at a modest computational cost and corrected using phase tracking [300], refocussing sequences [117, 118] (details in Chapter 4) and incoherent averaging techniques [301] (details in Chapter 6).

2.3.4 Initialization

Along with efficient methods and strategies for implementing the GRAPE algorithm, it is equally important to have the proper initial values set for the various input parameters. Here, I describe some useful guidelines on setting up the GRAPE algorithm to help achieve proper implementation and convergence.

Firstly, the total evolution time T must be chosen in accordance with the fundamental quantum speed limits [302, 303] and the minimum time required to implement the particular quantum gate [304]. Along with these, factors like the desired fidelity, threshold amplitude and the range of RF inhomogeneities to be tolerated also need to be accounted for. This issue is more important when designing non-trivial two-qubit gates. For example, the CZ gate in NMR requires the qubits to evolve for a minimum duration $1/2J$, where J is the coupling strength between the two qubits involved. Thus, the T needs to be set slightly higher (a few ms) than the minimum duration. If the gate needs to be made robust to RF inhomogeneities (say $\pm 10\%$ around the nominal RF), then roughly speaking, T should be at least 10% to 20% higher than the minimum time in order to obtaining reasonable fidelities.

The number of discrete RF inhomogeneity values for optimization should also be chosen carefully. Having too many values will simply slow down the computation and is often

unnecessary. The chosen values also need to be properly spaced so that the fidelity is optimizing for intermediate values. As described earlier, it is useful to have a few non-periodic values to avoid a behaviour where GRAPE does not optimize well enough for the intermediate values as compared to the chosen values.

Although there is no systematic approach on how to choose initial guess controls, it is useful to start from smoothly varying amplitudes. For a proper functioning of BFGS update, it is important that the control amplitudes and phases are of the same order of magnitude. To this end, the amplitudes can be normalized between 0 and 1 so that they are of similar magnitude as that of the phases.

Chapter 3

Practical GRAPE

In the previous chapter, we identified that the GRAPE algorithm has two computational bottle-necks: matrix exponentiation for evaluating the sub-propagators U_j and matrix multiplication for computing the full propagator $U(T)$. While there is no way to side-step the latter, in this chapter, I describe some strategies to evaluate the sub-propagators while avoiding explicit matrix exponentials. Along with this, methods for finding exact gradients at no additional computational overhead are also described. These strategies improve the computational speed and convergence of GRAPE at the same time and when combined with the sub-system approach, allow larger sized sub-systems to be considered, thus paving the way towards better scalable control.

3.1 Single-spin GRAPE

The dynamics of an isolated single spin can be evaluated analytically. Therefore, the GRAPE algorithm for a single spin deserves a separate treatment from the regular multi spin case. In this section, I present analytical forms of the sub-propagators and gradients in single spin GRAPE, followed by a discussion on how to efficiently implement and use single spin GRAPE.

3.1.1 Evaluating sub-propagators and gradients

Consider an isolated single spin with internal Hamiltonian $\mathcal{H}_0 = \omega I_z$ and control Hamiltonian $\mathcal{H}_j = A_j[\cos(\phi_j)I_x + \sin(\phi_j)I_y]$ for the j^{th} segment. The analytical form of the propagator for the j^{th} segment is

$$U_j = \exp\left(-\mathrm{i}\left[\mathcal{H}_0 + \mathcal{H}_j\right]\Delta t\right) = \begin{pmatrix} a-b & -\mathrm{i}c \\ -\mathrm{i}c^* & a+b \end{pmatrix}, \tag{3.1}$$

where

$$a = \cos(\kappa), \quad b = \mathrm{i}\omega\lambda\sin(\kappa), \quad c = \mathrm{e}^{-\mathrm{i}\phi_j}A_j\lambda\sin(\kappa),$$

$$\kappa = \frac{\Delta t}{2}\sqrt{A_j^2+\omega^2}, \quad \text{and} \quad \lambda = \frac{1}{\sqrt{A_j^2+\omega^2}}.$$

Note that only the off-diagonal elements in U_j depend on the phase ϕ_j. As a result, the analytic gradient of U_j with respect to ϕ_j takes a particularly simple form

$$\frac{\partial U_j}{\partial \phi_j} = \begin{pmatrix} 0 & -c \\ c^* & 0 \end{pmatrix}, \tag{3.2}$$

where the definition of c is the same as that in Eqn. 3.1. On the other hand, since A_j is present in all four elements of U_j, the analytic gradient of U_j with respect to A_j has a more complicated form

$$\frac{\partial U_j}{\partial A_j} = \begin{pmatrix} d+e & -\mathrm{i}f \\ -\mathrm{i}f^* & d-e \end{pmatrix}, \tag{3.3}$$

where

$$d = -\mathrm{i}A_j\omega\lambda^3\sin(\kappa), \quad e = -\mathrm{i}\frac{\Delta t}{2}A_j\lambda\left[\omega\lambda\cos(\kappa) + \mathrm{i}\sin(\kappa)\right],$$

$$f = \mathrm{e}^{-\mathrm{i}\phi_j}\lambda\left[\frac{\Delta t}{2}A_j^2\lambda\cos(\kappa) - A_j^2\lambda^2\sin(\kappa) + \sin(\kappa)\right],$$

and the definition of κ and λ is the same as that in Eqn. 3.1.

3.1.2 Discussion

Implementation

All matrices in single-spin GRAPE are 2×2 unitary matrices. When implementing single-spin GRAPE it is useful to treat the four elements of any matrix as four separate complex numbers and not use any matrix structures. Avoiding the matrix structure and treating its elements as individual entities when doing any mathematical operation can provide a speed-up of nearly three orders of magnitude over conventional implementations when MATLAB is used. Further speed-up is possible by fixing the amplitude to a constant value and modulating only the phase across the segments [305]. In such purely phase-modulated pulses, the diagonal elements for all sub-propagators are constant and hence, need to be evaluated only once. It is also straightforward to calculate the second derivatives of the sub-propagator with respect to the phases to obtain an analytic Hessian for accelerated convergence. One might wonder if a speed-up in single-spin GRAPE is really necessary, and if so, then what are its applications?

Good initial guess pulse

Consider the case of designing a GRAPE pulse for performing single-qubit gates on a system of n-coupled spins. In such a scenario, single-spin GRAPE can be used to design an approximate pulse by treating the coupled spins as n-isolated spins and optimizing for the target unitary for each of the individual spins. This approximate pulse can be further used to provide a very good initial guess for GRAPE to perform an optimization on the actual system with n-coupled spins. Such a two-stage optimization can greatly reduces the overall computational time for GRAPE.

Bandwidth selective pulses

In chemical and biochemical NMR, it is often necessary to implement pulses that affect spins only within a defined frequency bandwidth in a particular manner. In such cases, the pulse can be designed using single-spin GRAPE by envisaging the defined frequency range as a system of uncoupled spins placed at various discrete frequencies therein, and optimizing for the desired unitary operation for each spin. This approach has been used for protein NMR applications in collaboration with the Baldwin group at Oxford, but is not described in detail here. See Refs. [306, 307] for examples of some complementary approaches.

3.2 Multi-spin GRAPE

Based on the published work in Ref. [149], in this section, I describe an approach to obtain a high-accuracy approximation of the sub-propagator using only a single matrix multiplication. Although matrix multiplication has the same $\mathcal{O}(\mathcal{D}^3)$ scaling as that of matrix exponentiation, it is significantly cheaper to multiply two matrices than to calculate a matrix exponential using any of the numerical methods. As a result, this approach can provide a constant but substantial speed gain over conventional approaches, irrespective of the number of qubits in the system.

3.2.1 Evaluating sub-propagators and gradients

Consider a coupled n-spin system with internal Hamiltonian $\mathcal{H}_0$ and control Hamiltonian $\mathcal{H}_j = A_j[\cos(\phi_j)F_x + \sin(\phi_j)F_y]$ for the j^{th} segment. The propagator for the j^{th} segment can be written as

$$U_j = \exp\left(-\mathrm{i}\left[\mathcal{H}_0 + \mathcal{H}_j\right]\Delta t\right) = \exp\left(-\mathrm{i}\left[\mathcal{H}_0 + A_j\cos(\phi_j)F_x + A_j\sin(\phi_j)F_y\right]\Delta t\right). \tag{3.4}$$

Here, the control Hamiltonian $\mathcal{H}_j$ lies in the xy-plane making an angle ϕ_j with the x-axis. By performing an appropriate phase transformation, $\mathcal{H}_j$ can be made to align with the x-axis, thus allowing an equivalent representation of the sub-propagator:

$$U_j = Z_j X_j Z_j^\dagger, \tag{3.5}$$

where,

$$Z_j = \exp\left[-\mathrm{i}\phi_j F_z\right] \quad \text{and} \quad X_j = \exp\left[-\mathrm{i}\left(\mathcal{H}_0 + A_j F_x\right)\Delta t\right].$$

The matrix Z_j is diagonal in the computational basis and can be efficiently computed as it does not require a full matrix exponential. However, X_j still requires a full matrix exponential to be calculated as it contains a sum of two non-commuting terms. Nevertheless, it is possible to approximate X_j using the Trotter–Suzuki formula as

$$X_j \approx \exp[-\mathrm{i}\mathcal{H}_0\Delta t/2]\exp[-\mathrm{i}\Omega_j F_x \Delta t]\exp[-\mathrm{i}\mathcal{H}_0\Delta t/2], \tag{3.6}$$

given that Δt is sufficiently small. Note that the outer two terms in Eqn. 3.6 are constant and need to be evaluated only once for the entire GRAPE calculation. Furthermore, the central term is a simple scalar multiple of F_x and so will always be diagonal in the Hadamard basis. This basis transformation can be combined with the fixed outer terms, giving

$$U_j \approx Z_j \left(W_1 \exp[-\mathrm{i}A_j F_z \Delta t] W_2\right) Z_j^\dagger, \tag{3.7}$$

with all matrix exponentials now diagonal in the computational basis. The basis transformations

$$W_1 = \exp[-\mathrm{i}\mathcal{H}_0\Delta t/2]\mathrm{H}^{\otimes n} \quad \text{and} \quad W_2 = \mathrm{H}^{\otimes n}\exp[-\mathrm{i}\mathcal{H}_0\Delta t/2], \tag{3.8}$$

where $\mathrm{H}^{\otimes n}$ is the n-qubit Hadamard gate, are the same for every propagator, and so are evaluated only once for the entire GRAPE calculation. Further, the multiplication by diagonal matrices can be performed rapidly as there is no need for a full matrix multiplication. Thus, the evaluation of propagator U_j now requires only a single, full matrix multiplication.

Next, I describe how to evaluate exact gradients. In the previous chapter, we have seen that the BFGS algorithm requires exact gradients for proper functioning and to ensure good convergence. Although we have an approximate sub-propagator here, it is important that the gradients calculated with respect to the approximate propagator are exact [308]. Finding exact gradients with respect to A_j and phase ϕ_j in this case is particularly straightforward as they are both present in distinct terms in Eqn. 3.7. The gradient of U_j with respect to A_j can be written as

$$\frac{\partial U_j}{\partial A_j} = -\mathrm{i}\Delta t Z_j \left(W_1 F_z \exp[-\mathrm{i}A_j F_z \Delta t] W_2\right) Z_j^\dagger, \tag{3.9}$$

since only the middle term in Eqn. 3.7 depends on A_j. Similarly, only the Z_j terms in Eqn. 3.7 depend on the phase ϕ_j, allowing us to write the gradient of U_j with respect to ϕ_j as

$$\frac{\partial U_j}{\partial \phi_j} = \mathrm{i}U_j F_z - \mathrm{i}F_z U_j = \mathrm{i}[U_j, F_z]. \tag{3.10}$$

3.2.2 Discussion

Errors

Since the sub-propagator calculations shown above involve certain approximations, they are bound to have some error. The Trotter–Suzuki approximation made in Eqn. 3.6 is accurate to third order in δt. As the full propagator fidelity depends quadratically on the size of the error, this approximation will be accurate to sixth order in Δt when calculating fidelities and their associated gradients. In practice, these errors are essentially negligible for pulse engineering calculations.

The errors in this approach can be considered most simply for a one-spin system with $\mathcal{H}_0 = \omega I_z$ and $\mathcal{H}_j = A_j I_x$, which permits analytical calculations. The fidelity (as per Eqn. 2.16) between the exact analytic value of U_j, and its approximation, Eqn. 3.7, is

$$\mathcal{F} = 1 - \frac{\omega^2 A_j^2 (\omega^2 + 4A_j^2)}{2304} \Delta t^6 + O(\Delta t^8). \tag{3.11}$$

For realistic frequencies and time steps in NMR systems, the infidelity of the approximate approach will be very small. Offset frequencies rarely exceed 15 kHz, and for the low RF powers used during long control pulses the nutation rate is usually below 5 kHz. Even for these extreme values, the error for a 10 μs time step is below 5×10^{-5}, and for the lower frequencies and smaller time steps normally used the error will be far smaller. A detailed guideline which compares the size of time steps and the corresponding fidelity obtained using this method is given in a recent study in Ref. [309].

Evaluating the error for pulse engineering in a real system is more complicated, reflecting the larger matrices involved, the large number of relevant frequencies, and the need to evaluate the error in the entire combined propagator, and not just the sub-propagators. The worst case occurs when each sub-propagator is identical: in this case the errors grow linearly with the number of steps, and so the infidelity, which depends on the square of the error, grows quadratically with the number of steps. However such a case is quite unrealistic in practice, as variations in the amplitude and phase of control fields mean that errors will partly cancel. Simulations indicate that the infidelities remain small in realistic cases, typically around 10^{-4} for the full propagator.

Reducing errors

The infidelity in Eqn. 3.7 arises from the fact that $A_j F_x$ does not commute with $\mathcal{H}_0$, and can be minimised by making $A_j F_x$ as small as possible. This can be achieved by moving part of $A_j F_x$ into $\mathcal{H}_0$, writing

$$\mathcal{H}'_0 = \mathcal{H}_0 + A_{\text{off}} F_x, \quad \mathcal{H}'_j = A'_j F_x = (A_j - A_{\text{off}}) F_x, \tag{3.12}$$

where the offset frequency A_{off} is chosen to make the values A'_j as close to zero as possible, for example by setting it to half the maximum amplitude expected. While this will increase the infidelity of some sub-propagators, on average the infidelity will be decreased, and the fidelity of the full propagator will improve. The basis transformation operators, Eqn. 3.8,

have to be calculated for the new value of $\mathcal{H}_0'$, but as this only needs to be done once for the entire calculation this gain in fidelity comes at no cost in time.

A more accurate approach is to choose A_{off} as the mean amplitude for the particular propagator being calculated: in this case the basis transformation operators have to be recalculated each time, but this still only requires one full matrix exponential for each propagator. If even higher accuracy is required then it is possible to use two different offset values, one for small values of A_j and one for large values, choosing the appropriate basis transformation in each case. Our simulations suggest that using a single value of A_{off} can improve the infidelity of the propagator by a factor around 15, while using two values can give an improvement of about 200, leading to propagator infidelities around 10^{-6}.

This could be improved still further by using a larger number of offsets, effectively trading memory for time [310], but these later gains are smaller than the early ones. When the numbers of offsets is large then all values of A_j' are small in comparison with all other frequencies, and in this limit the infidelity falls quadratically with the number of offsets, as expected from Eqn. 3.11. If very precise pulse engineering is required then it is simpler to use approximate techniques in the early stages of optimization, and switch to exact calculations using full matrix exponentials once the pulse fidelity is high enough to justify this.

Speed gains

The speed gains achieved by this approximate approach arise from avoiding full matrix exponentials by performing all calculations in an appropriate eigenbasis where the operators are diagonal. Since both matrix multiplication and numerical matrix exponentiation have a computational complexity of $\mathcal{O}(\mathcal{D}^3)$, these gains are approximately constant, and independent of the dimension of the Hamiltonian.

As three of the matrices in Eqn. 3.6 are diagonal most steps can be carried out using efficient partial matrix multiplication, and only one full matrix product is required. The exact speed up achieved can be quite complex, due to implicit parallel computation on larger matri-

ces. While this reduces the wall clock time, the CPU time will be increased to handle the overheads of parallelisation. This will affect the observed speed-up in a way that depends upon matrix size [298] as matrix exponentiation is more easily parallelised than our efficient algorithm. The speed gain achievable will depend on both the spin system chosen and the precise code used, but MATLAB simulations for the 3-spin homonuclear system corresponding to the three ^{13}C nuclei in alanine [91] indicate that the calculation of sub-propagators can be sped up by a factor of around 18, while full propagators (which require an additional full matrix multiplication for each sub-propagator) are sped up by a factor of around 13. This speed gain means that a four-spin system can be simulated more rapidly with our approximate approach than an exact simulation of a three-spin system.

Efficient matrix multiplications

As matrix multiplication is now the computational bottleneck, it is vital that this is carried out as efficiently as possible. Full matrix multiplication, with computational complexity of $\mathcal{O}(\mathcal{D}^3)$[†], should only be used when actually necessary: when multiplying a full matrix by a diagonal matrix only $\mathcal{O}(\mathcal{D}^2)$ steps are required.

If the algorithm is coded in a low level language then it is easy to ensure this, but in higher level languages, such as MATLAB, it is necessary to code carefully. Diagonal matrices should be stored not as matrices but as vectors, to avoid unnecessary operations. In particular the exponentials of diagonal matrices must be calculated using direct exponentiation of the individual elements. To combine the individual matrices, Eqn. 3.7 can be written as

$$\{\phi\}W_1\{A\}W_2\{\phi^*\} \tag{3.13}$$

where braces indicate diagonal matrices stored as vectors. The multiplications by the two

[†]Note that it is possible to perform matrix multiplication with a lower computation complexity than $\mathcal{O}(\mathcal{D}^3)$. For example, Strassen's algorithm [311, 312] has complexity $\mathcal{O}(\mathcal{D}^{2.807})$, while the Coppersmith-Winograd algorithm [313] is $\mathcal{O}(\mathcal{D}^{2.376})$, and this has subsequently been refined to $\mathcal{O}(\mathcal{D}^{2.373})$ [314]. However these fast methods are only useful for large matrices, and so are not used here. Strassen's algorithm can be faster than conventional multiplication for $\mathcal{D} \sim 100$, and is used in some matrix libraries, but the more recent methods are only faster than Strassen's algorithm for extremely large matrices.

outer diagonal matrices can be combined, and the overall process written as

$$\Phi .* (W_1 * (\Omega .* W_2)), \quad \text{with} \quad \Phi = \phi .* \phi^{\dagger}, \tag{3.14}$$

where $*$ is the MATLAB operator for a full matrix multiplication, $.*$ is the operator for an element-by-element multiplication, and $\phi^{\dagger}$ indicates the adjoint of the vector ϕ. Note that only a single full matrix multiplication is required.

Combining the phase multiplications into a single matrix Φ is particularly useful when developing pulses which are robust to variations in the RF coupling strength. These are simulated by evaluating propagators with the control amplitude set to a range of values [315] (e.g., 95%, 100% and 105% of the nominal value). In such cases the matrix Φ is the same for all the different coupling strengths, and need only be calculated once.

When programming in MATLAB it is also important to think carefully about memory handling. In particular it is quicker to evaluate Eqn. 3.14 one multiplication at a time, storing intermediate results in explicit variables. If the multiplications are carried out in one line then temporary variables are created to hold intermediate results, and subsequently destroyed. If equivalent calculations are carried out many times, as happens when evaluating propagators, it is quicker to reuse previously allocated variables. These minor issues are almost irrelevant in conventional GRAPE calculations, where the time needed for matrix exponentiation dominates over everything else, but become important once all these slow stages have been removed.

It might appear possible to speed up calculations further, for example by using the structure in F_z to avoid repeatedly calculating the same exponential terms. This would certainly be sensible when programming in a low level language, but in practice such tricks actually slow MATLAB down. It can be difficult to predict precisely what will give the fastest MATLAB code, and experimentation is the best approach.

3.3 Phase-only GRAPE

A typical GRAPE pulse usually has both amplitudes and phases varying across its segments. However, in this section I consider pulses in which only the phase is permitted to vary while having a fixed constant amplitude across all segments. In such purely phase modulated pulses, it is possible to obtain the exact propagator without evaluating any matrix exponentials or even a single full matrix multiplication. Moreover, using a single predetermined amplitude for each pulse sidesteps the need for amplitude penalties and can also reduce the effects of transients [263].

3.3.1 Evaluating sub-propagators and gradients

Consider a coupled n-spin system with internal Hamiltonian $\mathcal{H}_0$ and control Hamiltonian $\mathcal{H}_j = A_0[\cos(\phi_j)F_x + \sin(\phi_j)F_y]$ for the j^{th} segment. Note that the amplitude A_0 is a constant for all the segments. Analogous to Eqn. 3.5, the propagator for the j^{th} segment can be written as,

$$U_j = Z_j X_0 Z_j^{\dagger}, \tag{3.15}$$

where,

$$Z_j = \exp\left[-\mathrm{i}\phi_j F_z\right] \quad \text{and} \quad X_0 = \exp\left[-\mathrm{i}\left(\mathcal{H}_0 + A_0 F_x\right)\Delta t\right].$$

Here, X_0 is constant for all the segments and can thus be evaluated only once and then stored for reuse. As discussed before, the diagonal matrix Z_j can be treated as a column vector to perform element-wise multiplication with X_0. As a result, the only computation required for evaluating U_j is the element-wise multiplication of a column vector with a constant matrix. Further, the exact gradient of U_j with respect to the phase can be calculated in a similar way as in Eqn. 3.10:

$$\frac{\partial U_j}{\partial \phi_j} = \mathrm{i}U_j F_z - \mathrm{i}F_z U_j = \mathrm{i}[U_j, F_z]. \tag{3.16}$$

Note that the exact second order derivatives with respect to the phase are easy to obtain in this case and can be used to construct an accurate Hessian matrix for accelerated convergence.

3.3.2 Discussion

Speed gains

Here, the only matrix operation required for evaluating the sub-propagator is the multiplication of a full matrix by a diagonal matrix, which has a complexity of $\mathcal{O}(\mathcal{D}^2)$. As a result, this approach provides a substantial speed-up over the method discussed in the previous section wherein the sub-propagator evaluation has a cost of $\mathcal{O}(\mathcal{D}^3)$ because of a full matrix multiplication. Moreover, an exact sub-propagator can be obtained with this approach. The most expensive computation which now remains is that of evaluating the full propagator, which is unavoidable.

Bang-bang control

Using a bang-bang type ansatz [316, 317] can be helpful in reducing the number of matrix multiplications required for evaluating the full propagator. Just like phase-only GRAPE, a bang-bang pulse sequence consist of segments having a fixed constant amplitude (with variable phase) known as bangs, but unlike phase-only GRAPE, each bang is surrounded by delays on either sides during which the spins evolve under the internal Hamiltonian [234]. A typical bang-bang sequence can thus be represented as

$$U(T) = D_N U_N D_{N-1} \dots D_j U_j D_{j-1} \dots D_2 U_2 D_1 \tag{3.17}$$

where

$$U_j = \exp\left(-\mathrm{i}\left[\mathcal{H}_0 + A_0\cos(\phi_j)F_x + A_0\sin(\phi_j)F_y\right]\Delta t\right), \tag{3.18}$$

$$D_j = \exp\left(-\mathrm{i}\mathcal{H}_0\tau_j\right). \tag{3.19}$$

Since the total number of bangs is small in a bang-bang sequence, the full propagator $U(T)$ requires very few matrix multiplications. The optimization of the phases ϕ_j and the delays τ_j is also fairly straightforward and inexpensive. The sub-propagator U_j and gradient $\partial U_j/\partial\phi_j$ can be evaluated efficiently in the exact same way as in phase-only GRAPE. The calculation of the delay propagator D_j is also simple in the case where $\mathcal{H}_0$ is diagonal, since no explicit matrix exponential is required. Even if $\mathcal{H}_0$ is not a diagonal matrix, it needs to be diagonalized only once, and the eigenvalues and eigenvectors can be stored and reused for an efficient calculation of the delay propagator for any τ_j as required. Irrespective of whether $\mathcal{H}_0$ is diagonal or not, the calculation of the delay gradient proceeds as follows:

$$\frac{\partial D_j}{\partial\tau_j} = \frac{\partial}{\partial\tau_j}\exp\left(-\mathrm{i}\mathcal{H}_0\tau_j\right) = \frac{\partial}{\partial\tau_j}\sum_{k=0}^{\infty}\frac{\left(-\mathrm{i}\tau_j\mathcal{H}_0\right)^k}{k!} = \left(-\mathrm{i}\mathcal{H}_0\right)\sum_{k=1}^{\infty}\frac{\left(-\mathrm{i}\tau_j\mathcal{H}_0\right)^{(k-1)}}{(k-1)!} = -\mathrm{i}\mathcal{H}_0 D_j \tag{3.20}$$

In order to restrict GRAPE from producing negative values of times, the time τ_j in the delay propagator can be squared so that it has a lower bound of zero. This rescaling of τ_j causes the delay propagator and delay gradients to take the form

$$D_j = \exp\left(-\mathrm{i}\mathcal{H}_0\tau_j^2\right), \quad \frac{\partial D_j}{\partial\tau_j} = -2\mathrm{i}\tau_j\mathcal{H}_0 D_j \tag{3.21}$$

respectively.

Chapter 4

Hamiltonian Engineering

Implementation of any non-trivial two-qubit gate requires the presence of an interaction or coupling between the qubits. In most platforms, including trapped ions and superconducting qubits, these couplings are generated as and when required through the application of external control fields [318, 319]. In other cases, exemplified by NMR, the couplings are inherently present and thus form a part of the internal Hamiltonian as we saw earlier [146]. Many other proposed implementations fit into one of these two paradigms or a mixture [320]: for example NV centers are NMR-like when controlling the spins around a single center [321] but generate long-range couplings 'on-demand' [322].

With 'always-on' couplings (non-switchable couplings in the internal Hamiltonian), free evolution provides a universal set of quantum gates when combined with single-qubit rotations. However, the two-qubit gate obtained from a mere free-evolution is not a convenient one as it corresponds to a complex pattern of evolutions with contributions from all interactions that are present in the internal Hamiltonian. As a result, it is useful to replace the internal Hamiltonian with a more convenient average Hamiltonian in which only the desired interactions are retained while the undesirable interactions are suppressed.

One brute-force approach here is to employ the GRAPE algorithm to engineer a target average Hamiltonian [323] or to directly design the required two-qubit gate itself as described in the previous chapters. An optimal control technique like GRAPE has the ability to im-

plicitly handle all the unwanted interactions when designing pulse sequences and is therefore a very popular tool in NMR-QIP. However, a more elegant and historic approach for generating average Hamiltonians in NMR is using a Hahn spin-echo [116] (described in the next section). Sequences based on spin-echoes have been studied in the past [117, 118, 133] for refocussing all interactions, or for isolating a single interaction while refocussing everything else. Although these sequences provide a time-optimal way to implement a single two-qubit gate, they do not scale well—with the total time as well as the number of echo pulses required—when implementing a network of two-qubit gates.

A more general problem is therefore to simultaneously *rescale* the strength of interactions in the internal Hamiltonian so as to implement multiple two-qubit gates in parallel. In this chapter, I describe a method for finding a spin-echo based rescaling sequence with the shortest possible total time and with a fairly small number of echo pulses. In its simplest form, this method is practical for systems of up to about 20 qubits, but for larger systems I describe a pragmatic method for finding short, although not perfectly optimal, rescaling sequences, which works with more than 100 qubits. Related ideas have been explored in systems with other similar Hamiltonians [324, 325].

A natural question arises here: why bother developing schemes for utilizing always-on interactions in a non-scalable platform like NMR? This is especially relevant when promising platforms like trapped ions and superconducting qubits generate couplings on-demand, which is also much more convenient. Firstly, the dynamic generation of couplings requires additional hardware to generate the control fields along with additional calibrations. Such overhead requirements become particularly significant as the number of qubits scale up. Secondly, it is also possible to tune superconducting qubits—a scalable platform—to generate always-on NMR-like zz-Ising couplings [326–328]. Although unpopular, Dr Peter Leek's group in Oxford plans to operate superconducting qubits in such a configuration. In fact, the rescaling problem described in this chapter was motivated by Takahiro Tsunoda from the Leek group which was then jointly addressed by Prof. Jonathan Jones and me.

This chapter is based on the publication in Ref. [150] and the following sections are largely

reproduced from the published work. In the initial sections below, I introduce the terminology of spin echoes and reprise previous results for refocussing sequences. This is followed by a description of our rescaling method with some sample applications in systems with a small number of qubits. Finally, the extension of our rescaling method to larger systems is presented.

4.1 Spin echoes

As described earlier in Chapter 2, the internal Hamiltonian of an NMR spin system consists of one spin interactions (resonance offsets) and two spin interactions (J-couplings). Here we consider a system of q spins where the i^{th} spin has a resonance offset Ω_i and the pair of spins i and j have a J-coupling frequency ω_{ij}. Note that the coupling frequency here is denoted by ω_{ij} unlike the notation J_{ij} used earlier in Eqn. 2.8. The Hamiltonian for this system in the weak coupling limit is,

$$\mathcal{H} = \sum_i \Omega_i I_{iz} + \sum_{i<j} \omega_{ij} I_{iz} I_{jz}, \tag{4.1}$$

where, following NMR notation [192], factors of $\hbar$ have been dropped, and I_{iz} is the Pauli spin-1/2 operator σ_z acting on the i^{th} spin. In practice some of these interactions could either be zero or set to zero, allowing them to be dropped. For example, working in a suitable rotating frame allows some of the resonance offsets to be set to zero, while many J-couplings can be negligible in partially coupled spin systems. However, for generality and completeness, we shall initially consider fully coupled systems with q non-zero resonance-offsets and $p = q(q-1)/2$ non-zero J-couplings between the p pairs of spins.

During a period τ of free evolution each spin evolves under all the $q+p$ interactions, given by the propagator $U = \exp\{(-\mathrm{i}\mathcal{H}\tau)\}$. Since this Hamiltonian is diagonal in the chosen z-basis, all terms in the Hamiltonian commute, and the one-spin and two-spin evolutions can be

summarised by the acquired phases

$$\Phi_i = \Omega_i \tau, \quad \phi_{ij} = \omega_{ij} \tau. \tag{4.2}$$

Note that the acquired phase Φ_i corresponds to a single-qubit z-rotation of spin i while ϕ_{ij} is a controlled-phase-shift between spins i and j. To sculpt the effective Hamiltonian into the desired form it is necessary to isolate the spin interactions which we want while suppressing the unwanted interactions. This essentially requires controlling the evolution of the spins such that the unwanted interactions finally acquire a phase of zero while letting the required interactions evolve to reach the desired values.

A spin echo is a period of free evolution interrupted by π rotations in the middle of the evolution period. The notation π^i is used to denote a pulse which causes a 180 degree rotation on spin i, about the x-axis unless otherwise stated. Note that a single-qubit π rotation is the same as a NOT (X) gate up to a global phase. The effect of a pair of π^i rotations is to negate the effective frequency Ω_i of the spin for the time period between the two pulses. Thus, the sequence $\tau\, \pi^i\, \tau\, \pi^i$, where time periods indicate free evolution under the internal Hamiltonian, will refocus the offset Ω_i as the phase Φ_i acquired during the first period of evolution gets nullified by the phase $-\Phi_i$ acquired in the second half.

Clearly, the one-spin interaction of any given spin i is only affected by π rotations applied to spin i, and so individual interactions can be controlled independently. However, for two-spin interactions, a π rotation applied to either spin i or j reverses the frequency ω_{ij} while a simultaneous π rotation applied on both spins i and j leaves ω_{ij} unchanged. A sequence $\tau\, \pi^i\, \tau\, \pi^i$ will thus refocus Ω_i and ω_{ij}, whereas a sequence $\tau\, \pi^{i,j}\, \tau\, \pi^{i,j}$ will refocus both Ω_i and Ω_j but not ω_{ij}. In this manner, the two-spin interactions can be controlled, but this control cannot be achieved independently from one-spin interactions.

A general spin echo sequence comprises a series of free evolution time periods τ_m, sometimes called delays, separated by π pulses applied to one or more spins. As long as the total number of π pulses applied to a given spin is even the overall evolution can still be summarised by

a set of phases, but now

$$\Phi_i = \Omega_i \sum_m S_m^i \tau_m, \quad \phi_{ij} = \omega_{ij} \sum_m S_m^i S_m^j \tau_m, \tag{4.3}$$

where S is a sign matrix, containing only the elements ± 1, with a sign change whenever a π pulse is applied to spin i. For convenience we will also refer to the two-qubit sign matrix $S_m^{ij} = S_m^i S_m^j$, although this is obviously not independent from the one-qubit matrix. The complete sign matrix can be obtained by combining the one- and two-spin matrices.

4.2 Refocusing

Methods for removing all the interactions (sometimes called decoupling), or for isolating one single interaction while refocusing everything else, have been widely studied. The most effective methods to achieve this rely on choosing sign matrices whose rows are taken from Walsh–Hadamard matrices [117, 118, 329], so that each row is a Walsh function [330]. These matrices differ from other Hadamard matrices in that they are only defined for dimensions equal to a power of 2, the rows are not normalised, and the ordering of the rows is different.

A Walsh function W_n is defined by a vector with length equal to a power of 2 and with all the entries equal to ± 1. For W_0 all the entries are $+1$, while for every other W_n half the entries are $+1$ and half are -1, with the entries arranged such that there are n regularly spaced sign changes along the vector. For example the 4 by 4 Walsh–Hadamard matrix contains the four rows

$$\begin{pmatrix} W_0 \\ W_1 \\ W_2 \\ W_3 \end{pmatrix} = \begin{pmatrix} +1 & +1 & +1 & +1 \\ +1 & +1 & -1 & -1 \\ +1 & -1 & -1 & +1 \\ +1 & -1 & +1 & -1 \end{pmatrix} \tag{4.4}$$

Strictly the name of the Walsh function must specify the number of columns as well as the number of sign changes, but this is left implicit here: the number of columns is equal to the

smallest power of 2 larger than the highest Walsh number considered.

In a system of three spins it is possible to remove all three one-spin and all three two-spin interactions by using four equal time periods τ and a sign matrix obtained by choosing $S^i = W_i$, avoiding W_0. This relies on two key properties of Walsh functions. Firstly all Walsh functions except W_0 contain an equal number of ± 1 values, and so all one-qubit interactions will be refocused when equal length time periods are used. Secondly the product of two Walsh functions is itself a Walsh function [330], defined by

$$W_p \circ W_q = W_{p \oplus q} \tag{4.5}$$

where $\circ$ indicates element wise multiplication, sometimes called the Schur product [331], and $\oplus$ indicates bitwise addition modulo two. Thus all two-qubit interactions will also be refocused.

A decoupling network is easily modified [117, 118] to retain a single interaction: to retain a one-spin interaction Ω_i use $S_i = W_0$ for this spin, while to retain a coupling ω_{ij} set $S_i = S_j$ so that $S_{ij} = W_0$. To take a concrete example the coupling ω_{12} can be isolated in a three spin system by choosing $S_1 = S_2 = W_1$ and $S_3 = W_2$. The π pulses required can be deduced by applying a pulse to a spin whenever the corresponding row of S changes sign, including a final π pulse if the S row ends in -1, giving the sequence

$$\tau\, \pi^3\, \tau\, \pi^{1,2}\, \tau\, \pi^3\, \tau\, \pi^{1,2}. \tag{4.6}$$

Note that ω_{12} evolves with sign $+1$ at every stage, and so is retained at full strength. The total evolution time required is given by $4\tau = \phi_{12}/\omega_{12}$. If this expression gives a negative time then this can be resolved by applying additional π pulses to one spin at the beginning and end of the sequence to negate the evolution frequency.

Now consider how this approach scales to a system of q spins. Retaining a single interaction can be done efficiently: the number of time periods required is given by the smallest power

of 2 larger than q, which is upper bounded by $2q$, while the number of individual π pulses required is clearly less than $2q^2$, which corresponds to applying a pulse to every spin after every time period. Next, consider a more careful analysis of the total number of pulses required. Retaining a single two-qubit interaction while refocusing all one-qubit interactions can be most efficiently achieved by assigning the two coupled qubits to the Walsh function W_1, with the remaining qubits assigned from W_2 up to W_{q-1} in sequence. As W_n contains n sign changes it requires n π pulses when n is even, and $n+1$ pulses when n is odd, with the additional pulse needed to restore the Hamiltonian to its original sign. The overall number of pulses required is then obtained by summing these pulse counts from $n = 1$ to $q-1$, remembering to include W_1 twice, giving a total of

$$\frac{q(q-1)}{2} + \left\lceil \frac{q-1}{2} \right\rceil + 2 \leq \frac{q^2}{2} + 2.$$

Thus the total number of pulses required to retain a single two-qubit interaction is approximately $q^2/2$, which is still $O(q^2)$ but with a smaller pre-factor. As the single interaction is retained at full strength this is also a minimum time solution.

4.3 Rescaling

So far, we have only considered retaining a single one-spin or two-spin interaction while refocusing the remainder. However, a more general problem is to *rescale* the size of interactions in the Hamiltonian. In other words, we desire to achieve a certain set of non-zero target phases for *all* the spin interactions. The obvious approach is just to place spin echo sequences which isolate the individual interactions back to back. As there are a total of $r = q + p = q(q+1)/2$ single-spin and two-spin interactions to be considered it is clear that the number of time periods is $O(q^3)$, and the number of pulses is $O(q^4)$. The total time

required is given by the sum of the times required to evolve under each individual interaction,

$$T = \sum_i \left| \frac{\Phi_i}{\Omega_i} \right| + \sum_{i<j} \left| \frac{\phi_{ij}}{\omega_{ij}} \right|. \tag{4.7}$$

This naïve approach is expensive, both in time and the number of π pulses. We thus want to find a more efficient sequence, by carrying out as many evolutions as possible in parallel. Although this might sound challenging, we propose here a straightforward way to achieve this using linear programming. This approach also greatly reduces the number of pulses and time periods required.

4.3.1 Setting up the problem

We consider a system of q coupled spins described by $r = q(q+1)/2$ one- and two-spin interactions. Our aim is to rescale all r interactions simultaneously such that they reach the desired target phases, which for generality we assume to be all different. We begin by constructing an overcomplete Walsh basis by building a one-spin sign matrix with the rows given by Walsh functions numbered 2^j, where $j = 0, 1, \ldots (q-1)$. Next, we use this matrix to construct the two-spin sign matrix by taking products of corresponding rows in the one-spin matrix.

Combining these by stacking the two matrices together gives the complete sign matrix S of r rows and $s = 2^q$ columns. The single-spin functions correspond to binary numbers with precisely one bit set, while the two-spin functions correspond to binary numbers with precisely two bits set. As these numbers are all distinct it is guaranteed that the complete sign matrix has enough flexibility to permit every interaction to be controlled separately. This is quite different from refocusing sequences, where many functions are repeated.

4.3.2 Linear Programming

This overcomplete Walsh basis guarantees that solutions to Eqn. 4.3 can be found for any target values of $\{\Phi_i\}$ and $\{\phi_{ij}\}$ by choosing 2^q appropriate values of $\{\tau_m\}$, but it is not obvious how these can be found. As the basis is overcomplete, multiple solutions will exist but these can be distinguished by requiring that all the times $\{\tau_m\}$ must be non-negative and by preferring the solutions with the shortest value of total time $T = \sum_m \tau_m$. These criteria for desirable solutions suggest a powerful method, namely linear programming [332].

The general linear programming problem varies some inputs (here the times $\{\tau_m\}$) seeking to minimize some linear function of these inputs (here the total time T) subject to a number of equality constraints (here Eqn. 4.3) and inequality constraints (here, that each $\tau_m \geq 0$). We adopted a simple approach, using the inbuilt Matlab function `linprog`.

It is important to consider the computational complexity of linear programming, as this determines how the time required to find a solution scales with the number of qubits q. The precise computational complexity of linear programming is known to be poorly defined, depending on both the algorithm used and the details of the problem, and with the worst case behaviour being very different from the typical case [332]. The Matlab function `linprog` has a choice of two algorithms: the original simplex algorithm developed by Dantzig [333], and a more modern interior point algorithm [334]. Both algorithms typically have computational complexity between $O(n^2)$ and $O(n^3)$, where n, the dimension of the problem, can be taken as the sum of the number of rows and columns in the constraint matrix, so that here $n = r + s \approx 2^q$. We investigated this question experimentally by simply timing the program. All results are for the simplex algorithm unless otherwise stated.

4.3.3 Extracting solutions

As the linear programming algorithm is fundamentally trying to minimise T, subject to the positivity constraint and the target phases, the algorithm prefers solutions where many of

the $\{\tau_m\}$ are zero. (This is not specific to this problem, but is a general feature of linear programming solutions [335].) The linear programming solution has at most only as many non-zero times as the number of constraints r in the problem. If the problem involves extensive refocusing rather than rescaling then solutions with an even smaller number of non-zero times can be found.

It is obviously not necessary to explicitly implement the time periods of length 0, and so the overcomplete sign matrix, S, can be replaced by a reduced matrix, R, by selecting only r or fewer columns from S which correspond to non-zero evolution times.

4.3.4 Optimizing the solutions

One subtlety is that the order of columns in the R matrix does not affect the phases produced, but different orderings of these columns can lead to pulse sequences with different numbers of pulses. As minimising the number of pulses is desirable it is useful to explore different permutations of the R matrix, seeking for the arrangement which gives the smallest number of sign changes.

If the matrix is not too large then exhaustive permutation can be practical, but in larger cases it is more sensible to select a number of random permutations and keep the best one found. Experience so far suggests that different permutations can require numbers of pulses that differ by a factor of around two. We also find that the pulse pattern corresponding to the original R matrix is typically relatively good, although rarely the absolute best.

4.3.5 Building the pulse sequence

From this optimal reduced matrix R, a pattern of pulses can be generated by applying a π pulse to every spin whose sign changes. It is important to remember that all interactions start at $+1$ and must end at $+1$, which can be modelled by adding initial and final columns to R containing entirely $+1$. These additional columns have evolution times set to zero, and

so are not actually implemented, but the resulting sign changes make it necessary to apply pulses to some spins at the start and end of the sequence. This also ensures that the number of π pulses applied to each spin is even, which is required to create true spin echoes.

As the reduced matrix has $r \approx q^2/2$ times the final pulse sequence will have $O(q^2)$ time periods and $O(q^3)$ individual π pulses, which is a very significant improvement on naïve methods.

4.4 Example calculations

We illustrate our method of rescaling with the help of two specific examples before returning to the general case, when we will consider the computational time required to find these optimal solutions.

4.4.1 Homonuclear 3-spin fully coupled system

First consider the example of iodotrifluoroethene [336], C_2F_3I, with the three ^{19}F nuclei forming our spin system, in a magnetic field such that the Larmor frequency of ^1H nuclei is 600 MHz, with the excitation frequency set in the middle of the spectral range. The molecular structure and the Hamiltonian of the system is shown in Fig. 4.1 (A).

This system has $r = 6$ interactions altogether. Our aim here is to rescale all three of the two-spin interactions such that they have equal effective strength, while refocusing the three one-spin interactions. This is equivalent to achieving the target phases $\{\Phi_1, \Phi_2, \Phi_3\} = \{0, 0, 0\}$ and $\{\phi_{12}, \phi_{23}, \phi_{13}\} = \{\pi, \pi, \pi\}$. The quantum circuit shown in Fig. 4.1 (A) uses the coupling gate [99, 300], which is equivalent to a CZ gate, with the notation indicating the complete symmetry of a controlled gate implemented through zz couplings, with no distinction between control and target spins. Actual CZ gates would require additional single-spin z-rotations, so that $\{\Phi_1, \Phi_2, \Phi_3\} = \{\pi, \pi, \pi\}$. These can be either be implemented

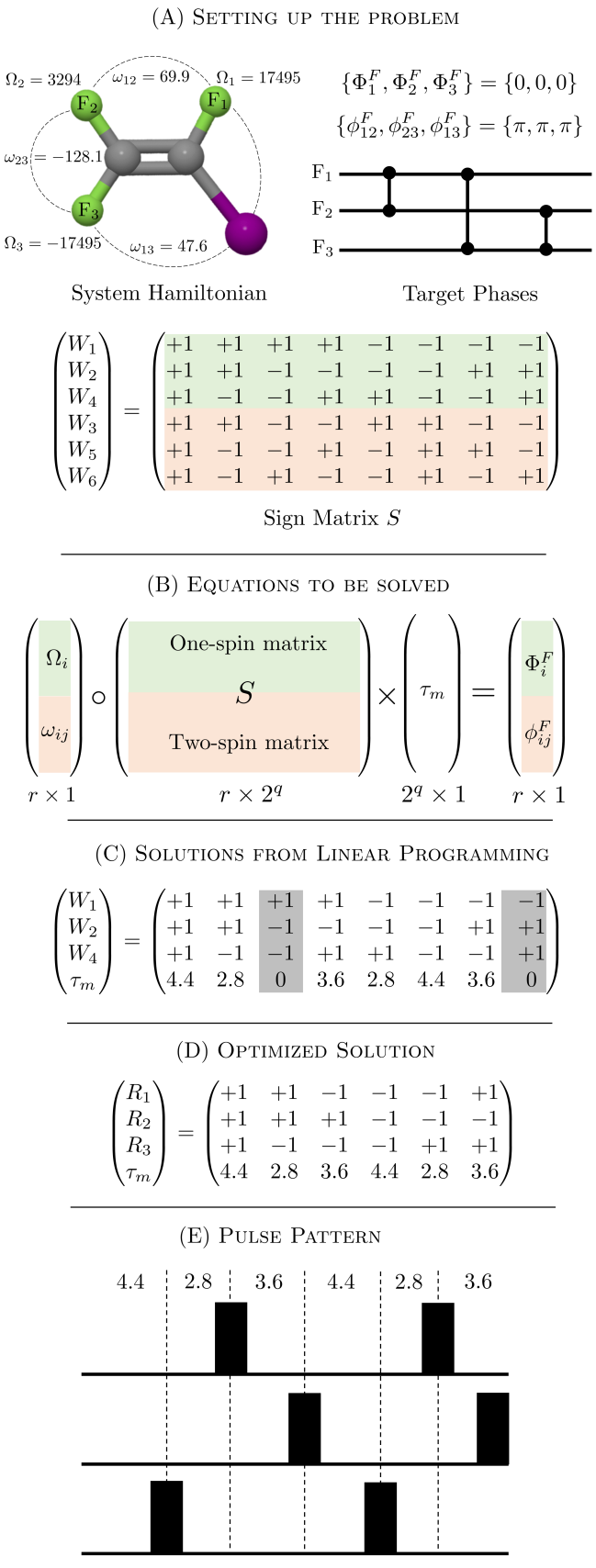


Figure 4.1: Implementation of our algorithm using the three ^{19}F nuclei in iodotrifluoroethene, forming a three-spin full coupled homonuclear NMR system. All interactions are written in Hz and times in ms. The stages of the implementation are described in the main text.

directly in the pulse sequence [317, 337, 338], or tracked and corrected at the end [99, 300]. Both approaches have been used in the literature, with tracking being particularly popular for fixed phases which are integer multiples of $\pi/2$, as seen here, as this is relatively simple in many experimental implementations.

Here we will consider both cases, beginning with tracking, so that the single-qubit target phases are all 0. Note that even though the target phases are all 0 or π, the corresponding evolution times will be different as the one- and two-spin interactions have different values. Thus, we expect to require up to six time periods, consistent with $r = 6$ interactions.

We begin by constructing an overcomplete basis from a one-spin sign matrix with three rows and eight columns, corresponding to the Walsh functions $\{W_1, W_2, W_4\}$, with the explicit form shown in Fig. 4.1 (A). The two-spin interaction matrix can be constructed by taking the three unique products of rows in the one-spin interaction matrix. The resulting set of equations is shown schematically in Fig. 4.1 (B). Here $\circ$ indicates an element wise (Schur) product as before, while $\times$ indicates a conventional matrix product.

As the sign matrix has more columns than rows there is no unique solution, but linear programming can be used to find a solution which minimises the total time while keeping all time periods positive. The program finds an optimal solution which uses only six columns as expected, by assigning zero time periods to two of the columns, as shown in Fig. 4.1 (C). The total time T of the evolution is 21.6 ms, which in this case is identical to the time required for a naïve implementation with sequential gates. This is a general rule for rescaling multiple couplings in three-spin systems, as it is impossible to control any coupling without affecting all the others.

However, even in this case, the resulting pulse sequence, requiring eight pulses, is much simpler than naïve sequential designs, which require 18 pulses. We can further improve this by permuting the columns of the reduced sign matrix R, as shown in Fig. 4.1 (D), to find a solution with a smaller number of sign changes. From the reduced sign matrix, we can now calculate the pattern of π pulses, Fig. 4.1 (E), with a π pulse applied wherever we encounter

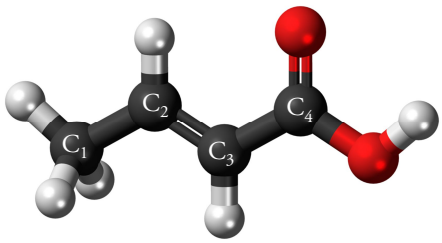
a sign change. Since there is one sign change in row R_2 it might seem that we need only one π pulse. However, since the row ends in -1 we must apply another π pulse at the end, leading to a total of six π pulses. Similarly, reduced sign matrix rows might in some cases begin with -1, and in this case it is necessary to apply a π pulse before the first time period.

Next we consider the case where tracking is not used, so that the single-qubit target phases of π need to be implemented directly. In this case a naïve sequential approach would require a small additional evolution time divided up by 12 additional pulses. By contrast our efficient method can achieve this with no additional evolution time or pulses by slightly unbalancing the identical pairs of times seen in the case above. It is always possible to do this as long as all the single-qubit offset frequencies are large in comparison with the two-qubit coupling strengths, which is usually the case and can always be achieved by changing the central excitation frequency.

We have assumed here that the π pulses are implemented perfectly, that is there is no evolution under the drift Hamiltonian while the π pulse is applied. This will be approximately true in a heteronuclear spin system, while in a homonuclear system equivalent pulses can be designed using techniques such as Gradient Ascent Pulse Engineering (GRAPE) [155] which are discussed in the previous chapters.

4.4.2 Homonuclear 4-spin linear chain

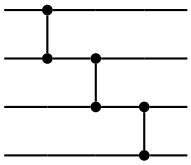
Next consider an NMR system provided by four ^{13}C nuclei in labelled crotonic acid [339]. The four spins form a rough linear chain, with large one-spin interactions, moderate nearest neighbour couplings and significantly smaller long range couplings as shown in Fig 4.2. Our typical target for this system is to refocus all the one-spin interactions and the three small couplings, but to retain the three nearest neighbour couplings. We also want to rescale the retained couplings such that they have the same target phases. We have a total of 10 interactions from which we need to refocus 7 while rescaling 3. This can also be represented,



C_1	C_2	C_3	C_4	(Hz)
-11962.2	41.6	1.5	7.1	C_1
	7306.0	69.6	1.2	C_2
		3972.1	72.3	C_3
			10626.1	C_4

Figure 4.2: Molecular structure of fully labelled crotonic acid and the corresponding Hamiltonian for four ^{13}C nuclei forming a four-spin chain like homonuclear NMR system.

up to single-spin phases, by the quantum circuit shown below.



(4.8)

In other words we want the target phases to be

$$\{\Phi_1, \Phi_2, \Phi_3, \Phi_4\} = \{0, 0, 0, 0\} \tag{4.9}$$

and

$$\{\phi_{12}, \phi_{13}, \phi_{14}, \phi_{23}, \phi_{24}, \phi_{34}\} = \{\pi, 0, 0, \pi, 0, \pi\}. \tag{4.10}$$

We begin by constructing the overcomplete basis by building a sign matrix S with 10 rows and 16 columns which correspond to the Walsh functions $\{W_1, W_2, W_4, W_8\}$ and their six distinct products, which are $\{W_3, W_5, W_6, W_9, W_{10}, W_{12}\}$. Running linear prediction and following the procedure as in the previous example, we now find a solution with 9 time periods

$$\begin{pmatrix} R_1 \\ R_2 \\ R_3 \\ R_4 \\ \tau_m \end{pmatrix} = \begin{pmatrix} +1 & +1 & -1 & -1 & -1 & -1 & +1 & +1 & +1 \\ +1 & +1 & +1 & -1 & -1 & -1 & -1 & +1 & +1 \\ +1 & +1 & +1 & +1 & -1 & -1 & -1 & -1 & -1 \\ +1 & -1 & -1 & +1 & +1 & -1 & -1 & -1 & +1 \\ 3.5 & 1.3 & 1.8 & 3.0 & 1.8 & 3.0 & 1.8 & 1.7 & 1.3 \end{pmatrix} \tag{4.11}$$

which can be implemented with 10 pulses. Note that since our target consists mostly of zeros the number of time periods required here is less than r. In this case the total time required is only 19.2 ms, significantly shorter than the naïve sequential time of 26.1 ms.

As a second example, we now consider the case where we only seek to implement the first and third CZ gates, so that the coupling phase are now

$$\{\phi_{12}, \phi_{13}, \phi_{14}, \phi_{23}, \phi_{24}, \phi_{34}\} = \{\pi, 0, 0, 0, 0, \pi\}. \tag{4.12}$$

In this case only 6 time periods are required, although once again 10 pulses are needed to implement the desired sign changes. The total time required is now 12.0 ms, which is identical to the time required to implement just the first gate by naïve methods. Thus in this case our algorithm has found a solution which implements two gates in parallel, at no additional cost in total time. Clearly it is not possible to implement the gates any quicker than this, as the time required to implement the slowest gate on its own provides a firm lower bound. An upper bound is provided by the time required to implement each of the r evolutions in sequence, Eq. 4.7, which is simply the sum of the times required for each gate.

4.4.3 General solutions

We have repeated calculations of this kind in a large variety of fictitious spin systems with increasing numbers of spins, up to $q = 18$, and solutions have always been found. For larger values of q these solutions are always more time-efficient than the sequential approach, and are usually far quicker. The greatest savings are found in cases where a moderate number of gates need to be implemented in parallel, and particularly when unused long-range couplings are significantly weaker than the couplings being controlled.

Until now we have described the problem as if there was a unique optimal solution which the linear programming locates. In fact there are multiple equivalent solutions, from among which the linear programming chooses one. These alternative solutions can be easily gener-

ated by permuting the columns of the S matrix before running the algorithm, but as they all have the same number of individual time periods and take the same total time there is no good reason to do this.

We note in passing that these optimal solutions are only optimal for implementations containing only delays and π pulses. If it is desired to implement an evolution corresponding to a weak long-range coupling then it may be quicker to use SWAP gates and related methods to implement long-range interactions through a chain of stronger short-range interactions [340, 341]. Even in such cases, however, the ideas described here can be used to assist in the design of such indirect gates.

4.4.4 Computation time complexity

The principal downside of this approach is that the computational time needed to run linear programming increases with the size of the basis set, and this grows as $s = 2^q$ for the current method. As discussed above a computation time scaling proportional to $(s+r)^2$ or $(s+r)^3$, where r is the number of equality constraints, is likely, which for large q is dominated by the exponential growth in s. This is confirmed by a plot of time required on a desktop computer (Intel Core i7-9700, 3.0–4.7 GHz, with 12 MB cache and 40 GB RAM), shown in Fig. 4.3. The linear behaviour at large q on this semi-log plot indicates exponential computational time complexity, and the gradient is consistent with the time required scaling as about 4^q.

Although attempts have been made to parallelize linear programming algorithms, progress so far has been limited [342]. Thus, this method seems to be practical only up to a small number of spins, perhaps $q = 20$. Indeed much above 20 qubits it becomes difficult even to hold the S matrix in memory on a desktop computer, although this could be sidestepped with a customised algorithm, in which the elements of S are calculated as they are used, rather than being calculated once and then stored in memory. While it is true that 20 is quite a large number of spins in the context of conventional NMR or even NMR QIP, we do not want to restrict ourselves to NMR spin systems but to extend to more general

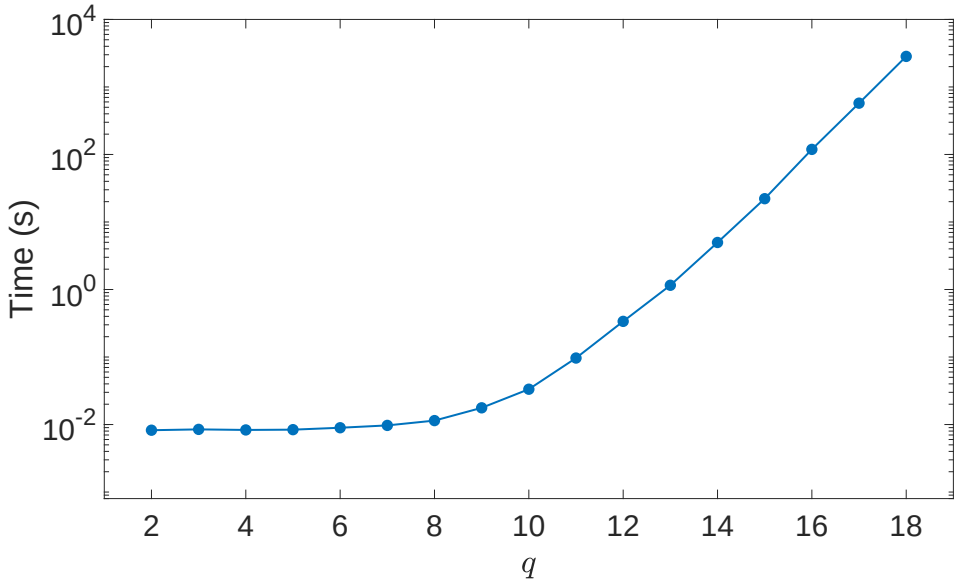


Figure 4.3: The computation (wall-clock) time required to find a linear programming solution on a desktop computer as a function of q. For small values of q this time is almost constant, but rises exponentially for larger q, rendering this method impractical beyond around 20 qubits. Timings are shown for the interior point algorithm, but results were very similar for the simplex algorithm. Error bars (estimated by repetition) are comparable to the size of the symbols, and the line simply joins the individual points. Note that the curve is almost flat for $q < 7$ because the timings there are dominated by fixed-time overheads linked to data-handling in Matlab, causing them to be nearly independent of q and hence approximately constant.

quantum systems which have the potential for a scalable quantum computing architecture. Fortunately, a more sophisticated approach is available which takes time only polynomial in q, albeit with a high power.

4.5 Stabilizing solutions

Having located an optimal solution it is important to consider whether it can be implemented reliably. This will obviously require the ability to implement large numbers of π pulses (single-qubit NOT gates) reliably, but we do not consider that problem here. Instead we consider the more fundamental question of the precision which is necessary for the free evolution times which lie between the pulses.

So far we have assumed that these times can be implemented with effectively perfect precision, but any real implementation will be built around a clock, so that any delay period must be rounded to the nearest multiple of some underlying cycle time. We show in Fig. 4.4 the effects of such rounding on the two three-qubit circuits considered in the previous section. This figure shows the infidelity $\mathcal{I} = 1 - \mathcal{F}$, where $\mathcal{F}$, the fidelity between the desired propagator, U, and the propagator actually implemented, V, is the propagator fidelity [343]

$$\mathcal{F} = \left| \frac{\mathrm{tr}(U^\dagger V)}{\mathrm{tr}(U^\dagger U)} \right|^2, \tag{4.13}$$

as seen in Chapter 2. Many other fidelity definitions are in use [244, 245, 344–347], some of which are more suitable for the experimental measurement of fidelities, but for comparing two unitary operations these are all closely related.

The lower line, plotted in blue, shows the infidelity for the first three-qubit circuit implemented on iodotrifluoroethene, in which the single-qubit target phases are all zero. Clearly this circuit works well, achieving an acceptable fidelity with a time resolution of $1\,\mu$s which is easily achievable in NMR experiments. The upper line, plotted in red, corresponds to the second three-qubit circuit, where the single-qubit target phases are all π. This circuit

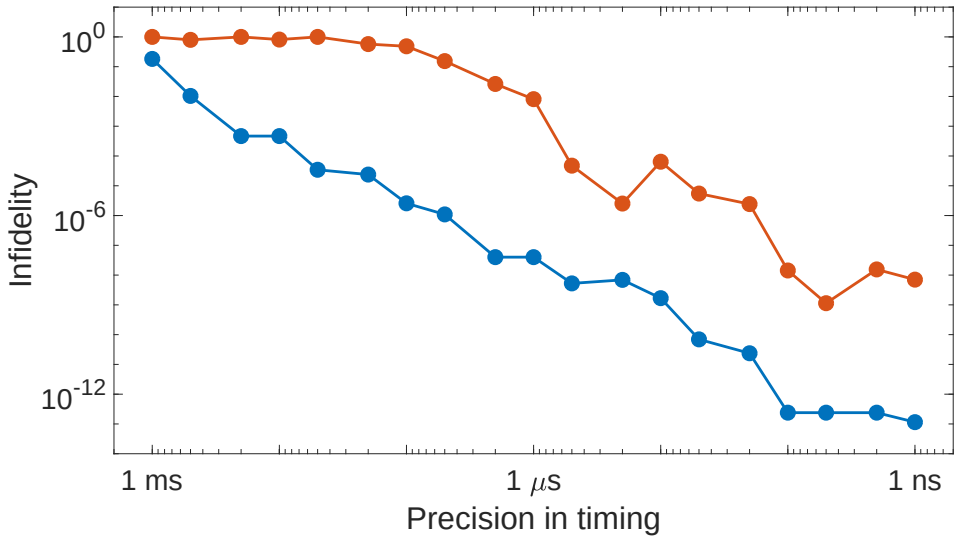


Figure 4.4: The infidelity of three-qubit circuits as a function of the precision of the underlying clock. Both the infidelity $\mathcal{I} = 1 - \mathcal{F}$ and the time resolution are plotted on logarithmic scales, with the time resolution running from 1 ms to 1 ns. The lower line, plotted in blue, shows the infidelity for the first three-qubit circuit, in which the single-qubit target phases are all zero. The upper line, plotted in red, corresponds to the second three-qubit circuit, where the single-qubit target phases are all π.

is unacceptably sensitive, requiring an unrealistic time resolution of about 1 ns to reach an acceptable fidelity.

The sensitivity of the second circuit arises because it uses evolution at the one-spin frequencies to achieve one-qubit phases. In NMR one-spin frequencies (typically several kHz) are much larger than the coupling frequencies that generate two-qubit phases (typically tens of Hz), and so very small errors in evolution time will give large phase errors. (This will be less of a problem in other technologies where the qubits are individually addressed, as the offset frequency can in principle be chosen at will, and so be set near zero.) The first circuit overcomes this by using pairs of *identical* time periods to cancel the single-spin evolution entirely, and when these identical time periods are rounded they remain identical, and so perfectly generate the desired single-qubit phase equal to 0. If a non-zero single-qubit phase is required it can be generated in other ways, most simply by altering the relative phase of two π pulses [317, 338].

However, this restriction is not sufficient to guarantee a stable solution. The four-qubit network described in Eq. 4.11 has all single-qubit phases equal to zero, and yet contains

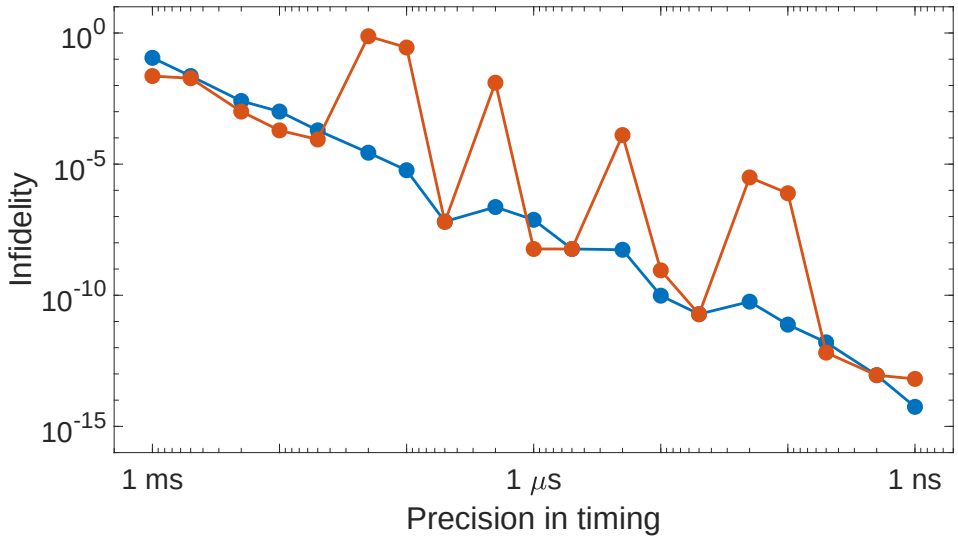


Figure 4.5: The infidelity of a four-qubit circuit (upper curve in red) and its stabilized equivalent (lower curve in blue) as a function of the precision of the underlying clock.

time periods without corresponding balancing pairs. The resulting infidelity is shown as the upper red curve in Fig. 4.5, and shows a very erratic dependence on time resolution. Fortunately there is an easy method to stabilize such networks. If a reduced sign matrix R is replaced by its negative $-R$ with the evolution times unchanged then this negated network will clearly generate identical two-spin phases to the original network (since the two-spin signs are negated twice, and so left unchanged) but the opposite one-spin phases. Thus if a matrix R is combined with $-R$ and all the evolution times are halved the combined network will generate the original two-spin phases, but will have all one-spin phases equal to zero, even in the presence of rounding errors in time periods. The infidelity of the stablized four-qubit network is shown as the lower blue curve in Fig. 4.5, and shows a greatly improved dependence on time resolution.

By its construction the stabilized network will require the same total evolution time as the original network, and the same computational time to find it, and so the cost of stabilization is confined to a doubling in the number of time periods and π pulses required for the implementation. This doubling can be reduced by subsequent optimization, locating equivalent evolutions in R and $-R$. Of the nine sign combinations found in Eq. 4.11, six can also be found in $-R$, and so the corresponding evolution times can be combined. Thus the symmetrized circuit can be reduced from 18 periods down to only 12 distinct evolution

periods.

As symmetrization is effectively essential to generate a stable network in the presence of large single-qubit interactions it is simpler just to generate symmetric networks by design. Since such networks are guaranteed to remove single-qubit phases, there are only $p = q(q-1)/2$ two-qubit phases that can be controlled. These phases can be controlled using only p underlying times, each of which must occur twice, once in R and once in $-R$. Thus it is unsurprising that a symmetrised four-qubit network contains twelve time periods. Such networks can be found by setting no constraints on the single-qubit phases, since these will be cancelled by the second half of the network. It is sufficient to use only columns from the first half of S in the linear programming search, as columns from the second half will be used in $-R$. Thus the search for symmetrized networks is actually faster than for direct networks.

4.6 Rescaling in larger spin systems

The fundamental problem which slows down the rescaling algorithm is the exponentially large size of the overcomplete basis. This basis contains $s = 2^q$ columns, from which the linear programming selects at most $r = q(q+1)/2$, equal to the number of interactions, and in many cases fewer. For moderate values of q, these numbers r and s are quite similar, but the difference grows rapidly with q, leading one to wonder whether there is some way to cut down the size of S before starting the linear programming step. Is it really necessary to include a very large number of columns, the great majority of which will eventually be discarded? In this regard we note that once one has identified the appropriate reduced sign matrix then the times required can be found by direct inversion of this r by r square matrix. However it is clear that most of the hard work is done in locating the appropriate columns used to construct the reduced sign matrix.

Starting from the other extreme, one could just select r random columns from the full sign

matrix and try to invert this. However, this process can fail in a number of ways. Firstly, the reduced matrix might not be of full rank (although this is unlikely when $s \gg r$) and so may not have an inverse. Even for a reduced matrix with full rank, the set of times obtained from the inversion process is very likely to include some negative times which are not physically implementable. Lastly, even if all the times are non-negative, the total time will not normally be the desired minimum, and so the sequence will not be time-optimal.

Between these extremes there is a middle way: using linear programming on a reduced, but still overcomplete, basis set. The current linear programming approach starts from the largest conceivable basis set, containing all of the $s = 2^q$ possible sign patterns, which guarantees finding an optimal solution but also makes the process slow. One might imagine choosing some subset of columns at random, and attempting linear programming on this subset. For large values of q the gap between the full size $s = 2^q$ and the minimum size r becomes very significant. It is thus worth exploring how many columns need to be picked so that linear programming generally finds a solution. There is no guarantee that such solutions will be time optimal, but as long as the random choice contains all the components of at least one optimal solution, then linear programming will find this. Given the very large number of equivalent solutions identified for moderate values of q it seems plausible that this could be achieved with quite a small subset.

4.6.1 The RROS method

In the random reduced overcomplete set (RROS) method, instead of using all s columns of the sign matrix S as in the exhaustive approach, we choose just kr columns from S, for some $k > 1$, and run linear programming. Of course, one does not have to explicitly construct the entire S matrix and then choose the kr columns as these kr columns correspond to the binary representations of kr distinct decimal numbers chosen randomly from 0 to $s - 1$. Our experience so far suggests that the probability of finding a possible solution, which achieves the desired phases using only positive evolution times, increases as k increases, with

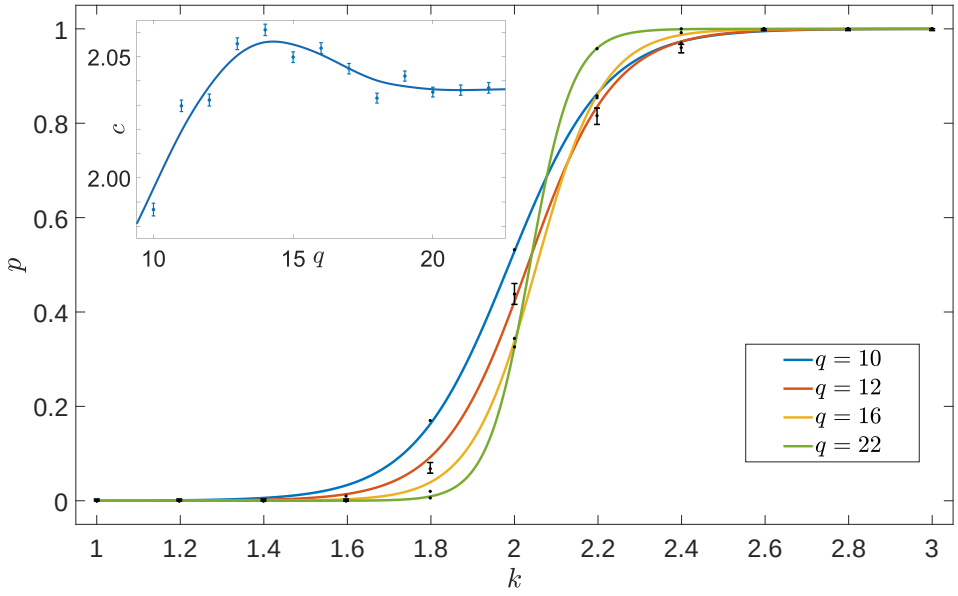


Figure 4.6: The effectiveness of the RROS method selecting kr columns at random, showing how the probability p of the algorithm finding a solution depends on q and k. Calculations were performed for q ranging from 10 to 22 but only four selected values are shown. Error bars on the $q = 12$ points were estimated using Bayesian credible intervals; error bars for other curves are very similar. Fitted sigmoidal curves were used to estimate the value of k at which this success probability reached 50%. The inset shows the location of this transition point as a function of q. Error bars are estimated using error propagation from the sigmoidal fit results; the smooth curve has no significance and is plotted simply to guide the eye.

a transition point around $k = 2$, at which the probability of a random set giving a solution reaches 50%.

This observation is substantiated by the empirical evidence in Fig. 4.6. RROS was run 500 times for values of q ranging from 10 to 22 with k varied between 1 and 3, and the fraction of occasions f when linear programming found a suitable solution was calculated. Error bars on these estimates were calculated using Bayesian credible intervals [348], corresponding to the region of the probability density function within ±34% of the median, equivalent to 1 standard deviation for a Normal distribution. It is clear that the probability of success rises sharply as k passes some critical value, with this transition becoming sharper as q is increased. To help locate this transition point a sigmoidal logistic function [348]

$$f(k) = \frac{1}{1 + \exp[-b(x - c)]} \tag{4.14}$$

was fitted for each value of q, with c being the transition point at which the success probability

passes 50%, and b indicating the sharpness of the transition. Although this function was chosen for convenience it clearly fits fairly well. A plot of the value of c as a function of q shown as an inset to Fig. 4.6 suggests that the transition lies just above $k = 2$. For large values of q this transition becomes sharp, so that for $k \geq 2.5$ it is almost certain that a solution will be found.

If a solution is found then this solution will be time-optimal for the subset of columns chosen, but there is no guarantee that this will be the overall optimum, taking the shortest possible time. Unsurprisingly the probability of finding a solution reaching the shortest possible time increases as k increases, but investigating this question in detail is challenging, as for large values of q the overall time-optimal solution cannot be located in a reasonable time. Nevertheless our preliminary studies suggest that for large q the quality of solutions plateaus around $k \approx 4$, and so there is little point going beyond this in practice. For small values of q it seems to be necessary to use a slightly larger value of k, but in these cases it is more sensible just to use direct solution of the full S matrix anyway.

4.6.2 Results for large numbers of qubits

The use of a smaller basis set permits RROS to be extended to much larger values of q. This was investigated by running the algorithm for q between 10 and 60, as shown in Fig. 4.7. For RROS the time required to perform linear programming depends not only on the randomly chosen Hamiltonian and target phases but also on the precise choice of columns, and so timings were repeated 10 times using different choices. As expected for $k = 4$ a solution was located in every case.

The simplex algorithm was found to be slightly faster than the interior point algorithm for $q \leq 31$, but the interior point algorithm was faster for $q \geq 32$, and became much faster at high q. The discontinuity in the interior point timings between $q = 31$ and $q = 32$ may indicate a change in the precise algorithm used by Matlab.

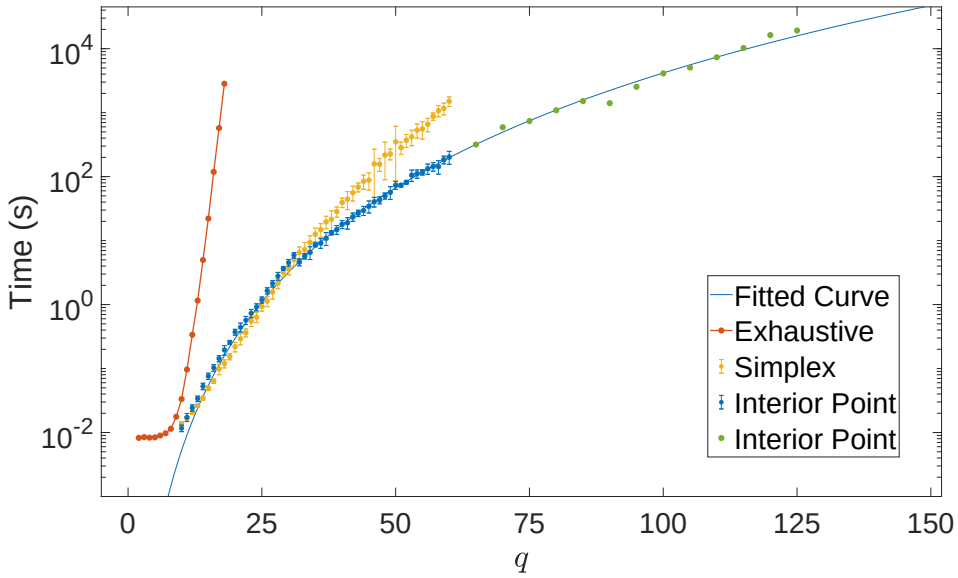


Figure 4.7: The computation time required to perform linear prediction using RROS with $k = 4$. The simplex method was found to be fastest for $q \leq 31$, with the interior point method faster for $q \geq 32$. Error bars for $q \leq 60$ indicated the mean and SD over 10 repetitions. The smooth curve with the form Ar^3 is fitted to interior point timings in the range $32 \leq q \leq 60$. Data points for $q > 60$ are single repetitions and were not included in the fit. The fact that these points lie very close to the extrapolated fit suggests that our fitted curve is a good model. The red points on the left show timings for exhaustive calculations, demonstrating the huge time gains possible with RROS.

For RROS the dimension of the problem is $n = r + kr$ and so we expect a computational complexity between $O(r^2)$ and $O(r^3)$. The smooth curve in the figure was fitted to the timings from the interior point algorithm for the range $32 \leq q \leq 60$ using a power function, and takes the form Ar^3, consistent with our expectations (including the power as an additional fitting parameter gave a value indistinguishable from 3). The timings for the simplex algorithm appear to lie between $O(r^4)$ and $O(r^5)$, which is much slower than the interior point.

With the simplex algorithm the number of time periods required was always approximately equal to r (in some cases a small number of time periods were negligibly short and so could be dropped). The number of pulses required was determined for the R matrix as originally found, without further optimisation, and was typically around $qr/2 \approx q^3/4$, confirming our previous expectations. With the interior point algorithm the solutions always use all kr time periods, with a large number of very short times rather than a clear division into zero and non-zero times. These initial solutions can then be simplified by choosing a smaller subset

of $k'r$ columns, with $k' \approx 1.2$, corresponding to the columns with the largest time values in the original solution. Surprisingly the computational time for these recalculations appears to be linear in k, while a quadratic or cubic dependence might have been expected.

The calculation was then extended to higher values of q, using only a single repetition, and these data points were found to lie very close to the extrapolated curve. As $r \approx q^2/2$ this gives a pragmatic computational time scaling of $O(q^6)$, which is polynomial in the number of qubits. Extrapolating this curve still further suggests that calculations with up to about 150 qubits could be practical, but the Matlab implementation `linprog` runs out of memory above $q = 125$. This limit should be solvable by a custom implementation of linear programming.

4.6.3 Stability of solutions

With large numbers of qubits it is not easy to directly evaluate the fidelity of a solution as even writing down the diagonal elements of the underlying propagator requires 2^q terms. It is, however, possible to estimate the infidelity directly using

$$\mathcal{I} \approx \sum_{i<j} \frac{(\delta\phi_{ij})^2}{16} \tag{4.15}$$

where $\delta\phi_{ij}$ is the error in the coupling angle for the pair of qubits i and j, and the approximation is only valid when the infidelity is small. This assumes that the solution has been symmetrised so that the errors in the single-spin angles are all zero. Networks with large numbers of qubits have higher infidelities than those shown in section 4.5, reflecting the larger number of phases which need to be controlled. This is not a specific limitation of our methods, but is inherent in any attempt to control a large number of qubits simultaneously.

4.7 Conclusions

We have presented a practical algorithm which is guaranteed to find the minimum time solution to rescaling z and zz terms in the internal Hamiltonian of a quantum computer with up to about 20 qubits. Above 20 qubits the direct approach becomes intractable, but random sampling will extend this (although the solution might not be quite time optimal) to more than 100 qubits. Beyond about 150 qubits the q^6 scaling of the computational time for a fully coupled system renders any known approach impractical. It is, however, very unlikely that any system of that size would still be fully coupled, as couplings are usually only substantial between nearby spins. In partially coupled systems r is linear in q rather than quadratic, and allowing for this should permit the process to be extended to hundreds or even thousands of qubits.

Although described in the language of NMR the techniques used are applicable to any equivalent system, where an always-on two-qubit interaction commutes with the single-qubit background terms. It may be of particular value in solid state platforms, such as superconducting circuits, in which 2D lattices of qubits are developed with a sparse coupling network (generally nearest-neighbour). In such very large systems it is likely that the underlying structure in the pattern of interactions will allow symmetries to be exploited, potentially simplifying the problem greatly. One approach for doing this is discussed in the following chapter.

Chapter 5

Practical Hamiltonian Engineering

In the previous chapter, linear-programming based methods for finding time-optimal rescaling sequences in fully-coupled systems of up to around 20 qubits, and near-optimal sequences for around 100 qubits were described. These methods, however, do not scale to larger systems, as the computational time required to find these sequences grows rapidly with the dimension of the linear programming problem. The minimum dimension of the problem is of the order of total number of interactions in the system, which in turn grows quadratically with the number of qubits in the case of a fully-coupled system. Thus, for about 100 qubits, the linear programming problem is at least 5000 dimensional, making it computationally infeasible to go much beyond this.

In practical quantum systems however, it is rarely the case that every qubit is coupled to every other qubit in the system: the couplings are generally only significant between nearby qubits. For example, superconducting qubits are often engineered in a square-lattice [349] and can have couplings limited to only nearest or next-nearest neighbour qubits if circuits are well microwave engineered [350]. To this end, here I describe an analytical method based on graph colouring for rapidly finding near-optimal rescaling sequences in these highly practical locally coupled systems. Remarkably in this partially-coupled scenario, which is very realistic for large-scale superconducting circuits, the refocusing and rescaling of always-on couplings can be *efficiently* programmed, requiring only linear time to design control sequences with

a linear number of control pulses.

In the following sections, the refocusing scheme will first be applied to the smallest possible case of a nearest-neighbour square lattice, and then generalised to arbitrary lattice sizes. The square lattice is initially assumed to be engineered with identical couplings between nearest neighbours and no long range couplings at all, but both of these restrictions will subsequently be partly relaxed. Throughout the chapter, again, it will be assumed that ideal single qubit NOT gates can be applied to any combination of qubits, despite the presence of the internal Hamiltonian. Experimental implementations will need to use either general approaches such as GRAPE pulses [155], or specific methods tailored to this precise situation [351–353].

This chapter is based on the publication in Ref. [151] and the following sections are largely taken from the published work. My main contribution here was convincing the other authors about the existence and complexity of analytical graph colouring solutions by performing linear programming based rescaling simulations for the partially coupled systems. These simulations indicate that the solutions described below are close to being time optimal.

5.1 Square lattices

We begin with the smallest possible square lattice, containing four qubits with identical nearest neighbour couplings $\omega_{j,j+1}$ and ω_{14} only. Non nearest-neighbour pairs have no coupling, so $\omega_{13} = \omega_{24} = 0$. We can refocus all the couplings in the system, as well as any offsets present, using the circuit of spin-echoes shown in Fig. 5.1(a). Here, the unitary U describes the evolution of the system under the background Hamiltonian $\mathcal{H}$ given by the propagator $U = \exp\{(-\mathrm{i}\mathcal{H}t/\hbar)\}$, while the NOT gates are represented by X. As there is no coupling to be refocused between qubits 1 and 3 it is possible to apply the same pattern of NOT gates to them both, and similarly for qubits 2 and 4, so only two distinct patterns of NOT gates are required.

This network can also be extended to retain a single coupling, as shown in Fig. 5.1(b). In this

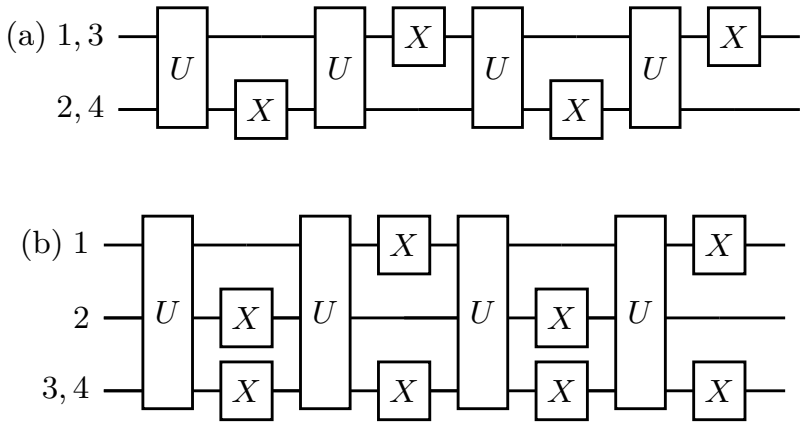


Figure 5.1: Networks to (a) refocus all offsets and couplings in a square of four qubits with only nearest-neighbour couplings, and (b) keep only the coupling between qubits 3 and 4. Here X indicates a NOT gate while U indicates evolution under the background Hamiltonian. Labels on the left indicate which qubits the NOT gates are applied to.

network the same pattern of NOT gates is applied to qubits 3 and 4, and the corresponding coupling is retained, while the remaining couplings are refocused [117]. These networks can be extended to retain *any* pattern of couplings in a square lattice system.

This extension works by colouring the square lattice, as described in [117]. The system can be described as a noncomplete graph, with vertices corresponding to qubits and edges to couplings, with only some of the possible edges present. The graph can be coloured by assigning a colour to each vertex, and is said to be properly coloured, corresponding to complete refocusing, if no two connected vertices are the same colour. Thus in the fully-decoupled square of four qubits we can colour qubits 1 and 3 black and qubits 2 and 4 white, while to retain a single coupling the pair of qubits involved must be assigned the same third colour, say red, as shown in Fig. 5.2(a). To implement a colouring, we use patterns of NOT gates corresponding to distinct Walsh functions [117, 150] for each colour.

This colouring pattern can be tessellated across a lattice, as shown in Fig. 5.2(b), by colouring surrounding qubits alternately black and white. Here we show a four-by-four patch containing sixteen qubits, which can be embedded in a larger lattice, retaining a single coupling while refocusing all the other interactions. The required pulse sequence can be obtained by assigning black qubits B to the first Walsh pattern, white qubits W to the second, and the

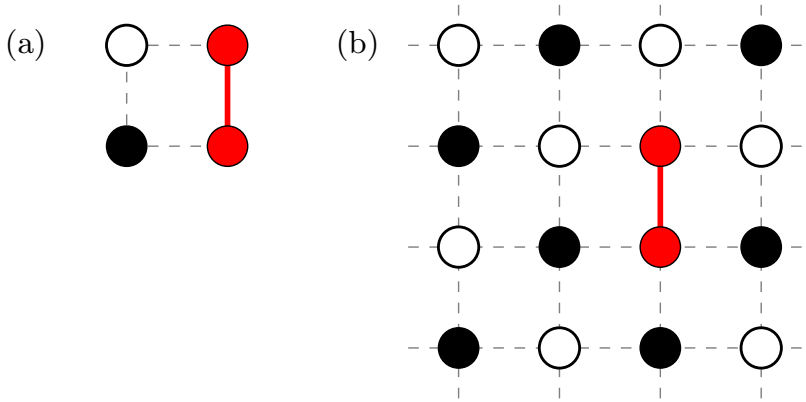


Figure 5.2: (a) Retaining a nearest neighbour coupling in a square requires three colours: decoupled qubits are coloured black and white, while the coupled qubits are coloured red. Grey dashed lines show couplings which have been refocused. (b) The same result can be achieved in a larger array by colouring surrounding qubits alternately black and white.

two red qubits R to the third, to obtain the sequence

$$U\,X_{\mathrm{WR}}\,U\,X_{\mathrm{BR}}\,U\,X_{\mathrm{WR}}\,U\,X_{\mathrm{BR}}, \tag{5.1}$$

where X_{WR} indicates that NOT gates are applied to the white and red qubits, and so on. Just like the sequence for the four-qubit system, this sequence requires only 4 time periods, but now requires $2q+4$ pulses for a system of q qubits. The total time required to implement the network for a π evolution, corresponding to a CNOT gate [150], is $T = 1/2J$, where the nearest-neighbour couplings are of size $\omega = 2\pi J$. This is the same time as is needed for an isolated coupling, as the retained coupling evolves at full strength.

The approach above will refocus all single-qubit interactions, but it is simple to modify the X gates in the network to implement single-qubit rotations directly. This relies on the identity [317]

$$\pi_{\phi_2}\pi_{\phi_1} = 2(\phi_2 - \phi_1)_z, \tag{5.2}$$

so that applying two π rotations around axes in the xy-plane separated by an angle δ is equivalent to performing a z-rotation through an angle 2δ. These rotations can be performed by modifying any pair of X gates in any refocusing network, and as each qubit is controlled separately different rotations can be applied to different qubits at no cost in time or pulse count.

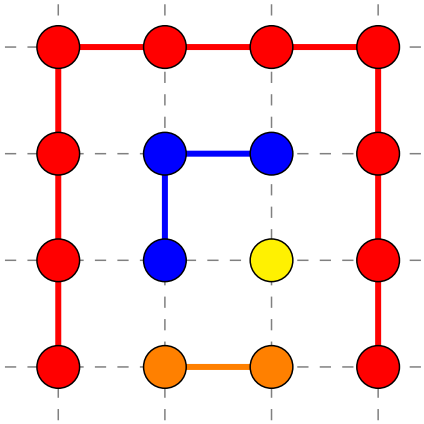


Figure 5.3: Retaining a more complex pattern of couplings in a square lattice: this pattern requires four colours.

5.2 Parallel gates

If it is desired to retain several different coupling interactions then this can be achieved most simply by applying such patterns back to back, changing the colouring at each stage, but it is more efficient to as far as possible perform evolutions in parallel. For simple cases this can be achieved as shown in Fig. 5.3. Qubits which are part of the same coupling island, that is qubits which are connected either directly or indirectly by retained couplings, have been coloured the same colour. (Note that the single yellow qubit forms an island on its own.)

This simple approach will only be successful when, as here, all the couplings which could appear in an island are retained. In such cases this simple colouring strategy will retain all the desired couplings while refocusing all the couplings between islands, reproducing the desired pattern in one go. As before, the total time required is just $T = 1/2J$, the same as for an isolated coupling.

It might appear that this approach would require the number of colours to equal the number of islands, but in fact it is never necessary to use more than four, since islands which are completely disconnected (that is, islands which are not connected by couplings in the underlying Hamiltonian) can be safely coloured the same. By the four-colour-map theorem this will never require more than four colours [354]. These four colours can be assigned to four

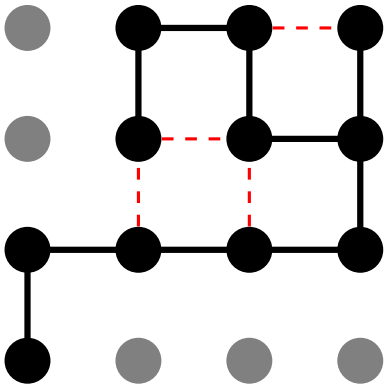


Figure 5.4: A target pattern of couplings which cannot be implemented with a single colouring pattern. Although the grey qubits are easily decoupled from each other and from the black island, it is impossible to retain all the desired black couplings without also retaining the undesired couplings shown as dashed red lines.

Walsh functions [150], requiring T to be divided up into eight equal time periods. Assigning the most common colours to W_1 and W_2, which both require two pulses, and the two rarer colours to W_3 and W_4, which require four pulses, means that the total number of pulses required in a system of q qubits lies between $2q$ and $3q$.

Although this method will work for some simple target patterns, it will not work in general, since many target patterns have one or more missing couplings. Consider, for example, the pattern shown in Fig. 5.4, where the black couplings must be retained, and the dashed red couplings must not be retained although they connect qubits within the main island. Any colouring which implements all the black couplings that must be retained will also implement the unwanted red couplings, and so this target pattern cannot be implemented with any single colouring pattern.

5.3 Multiple colourings

There is, however, a simple method for achieving *any* target pattern with two sequential colourings, each using four colours. As an example a pair of colourings which implements the target couplings in Fig. 5.4 is shown in Fig. 5.5. Pattern (a) assigns two colours to the odd-numbered rows and two more to the even numbered rows, thus ensuring horizontal

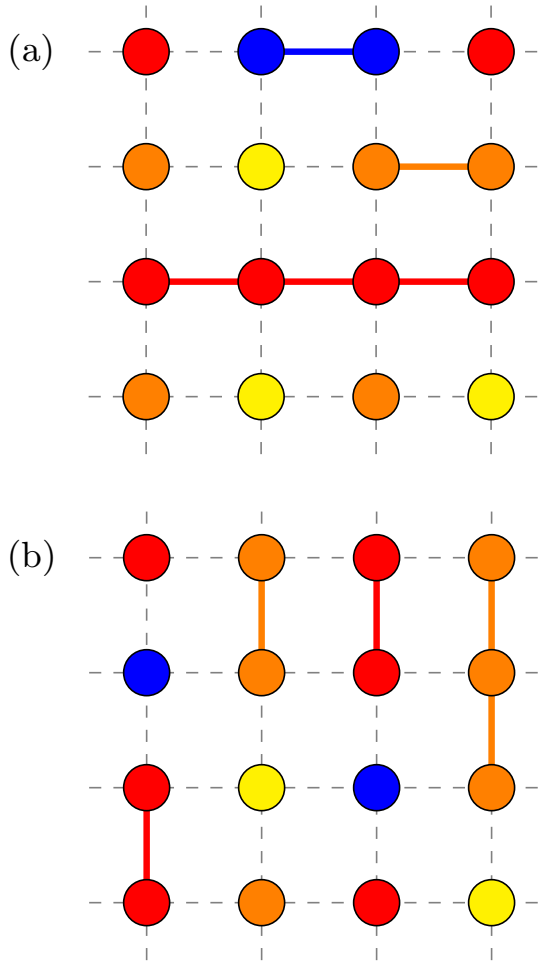


Figure 5.5: A solution implementing the target couplings in Fig. 5.4 using two sequential colouring patterns, (a) and (b).

couplings can be controlled while all vertical couplings will be refocused. Along a row, the colour of the first qubit is arbitrary, but the following qubit must be the same colour if the corresponding coupling is to be retained, and the other colour if it is to be refocused. Pattern (b) assigns colours to columns instead of rows to control vertical couplings in an analogous way.

As any subset of horizontal couplings can be selected by the first colouring, and any subset of vertical couplings can be selected by the second, any target pattern at all can be implemented in this way, in a total time $2T$. This total time will be divided into 16 equal time periods, separated by no more than $6q$ individual pulses.

It is useful to compare this result with timings found by our previous algorithm [150] based on

linear programming with an exhaustive basis set, which is guaranteed to find a true minimum-time solution. The computational complexity of this algorithm renders it impractical for large arrays, but for four-by-four arrays with nearest-neighbour couplings it is perfectly practical. We have analysed a very large number of randomly chosen targets in four-by-four arrays, and in every case the optimal solution required a total time of either T (in cases with no missing couplings, so a single colouring is possible) or $1.5\,T$ (in cases where this is not possible). Since there are more than 16 million possible targets with a four-by-four array, it is not feasible to study all of these by exhaustive linear programming, which takes around 10 seconds to find the optimal time. It is, however, possible to use the reduced-random-overcomplete-set (RROS) approach described in the previous chapter to find an upper bound on the time required for every target. We find that around 90% of targets require time $1.5\,T$, with the remaining 10% requiring a range of times between T and $1.5\,T$.

The implementation time required for our colouring based networks, $2T$, is slightly longer than the absolute minimum required, but these colourings are far easier to design, with a computation time scaling only *linearly* with the total number of qubits, and so can be applied to systems any number of qubits. The number of NOT gates required is also greatly reduced, from $O(q^2)$ for linear programming based solutions to $O(q)$.

5.4 Next-nearest neighbours

Until now we have assumed that only nearest-neighbour interactions are important, with all others being too small to matter. In practice, real physical systems are likely to also have non-zero couplings at longer range. It is reasonable to expect such couplings to drop with distance in well-engineered devices, making the next-nearest neighbour interactions, across the diagonals of the square array, the most significant. In the original two-colour refocusing scheme qubits connected by next-nearest neighbour interactions will be the same colour, both black or both white, and so these interactions will be retained rather than refocused, leading to a significant error in the final gate implementation.

This is not a problem with the general four-colour scheme, as the use of different colours for alternate rows and columns guarantees that diagonally connected qubits will be of different colours. Thus these networks automatically suppress any unwanted diagonal couplings, which can therefore simply be ignored. Suppressing even longer range couplings is more complicated, but can be achieved using larger numbers of colouring patterns with more colours used in each pattern, as described in the next section.

5.5 Longer range couplings

While longer range couplings are likely to be significantly weaker than nearest neighbour couplings, they will not in general be zero, and so could lead to significant errors. We begin by considering the case of error couplings between qubits separated by two edges horizontally or vertically (qubits separated by one horizontal and one vertical edge give rise to the diagonal error couplings considered previously). As before we will first colour the horizontal edges and then the vertical edges. When colouring horizontal edges we now have to divide the qubits into *three* groups of rows, rather than two, so that members of each group are at least three rows apart. Within each row a further problem arises, as we cannot directly retain couplings along two sequential horizontal edges without also retaining the error coupling between the two end qubits. Instead we have to use two separate colourings, retaining the two edges separately. This is most simply achieved by retaining odd and even numbered edges in separate steps. Within each row it is necessary to use three colours so that a colour is never repeated two qubits apart.

Putting this together we use two colouring patterns for the horizontal edges, and two for the vertical edges, making four patterns overall. Each pattern requires three separate colours for each of three sets of rows or columns, making nine colours in total. As before all diagonal error couplings will be automatically refocussed if the corresponding horizontal and vertical errors are refocussed.

This approach can be generalised to error couplings over any distance: to correctly handle all couplings between qubits n horizontal or vertical edges apart it is sufficient to use $2n$ colouring patterns, each of which has $(n+1)^2$ different colours. For the simplest case of nearest-neighbour couplings we have $n = 1$ and so two different four-colourings suffice, as described earlier.

5.6 Different evolution times

The parallel gates approach, however, hides a further important assumption: it is not sufficient to retain two different couplings if they are required to evolve for different times, either because different couplings must evolve through different angles, or because apparently equivalent couplings will have different strengths, and so will need different evolution times to achieve the same angle.

It might seem necessary to apply such gates using different echo sequences, but in fact they can be partly combined. Consider two couplings in the same group, where one requires evolution for a total time T_A and the second for a time T_B, with $T_B > T_A$. The naive approach is to use two different colouring patterns, one implementing the first coupling for time T_A, and another implementing the second coupling for T_B. In fact these periods can be carried out partly in parallel: during the first period, which lasts for time T_A, *both* couplings are retained, while for the second period, which lasts for time $T_B - T_A$, only the second coupling is retained. Thus both couplings can be carried out in a total time T_B, and the method generalises for any number of distinct couplings strengths.

It follows that *any* pattern of couplings can be achieved in an evolution time equal to the sum of the longest evolution times required for horizontal and vertical couplings, which itself is no more than $2T_{\max}$, where $T_{\max} = 1/2J_{\min}$ is the evolution time required for the slowest gate. However the resulting sequences will contain $O(q)$ evolution times and $O(q^2)$ pulses, while designing such sequences requires sorting the evolution times into ascending order,

with computational time complexity $O(q \log q)$. They will also be impractical to implement experimentally, as the differences between very similar times may be smaller than the clock resolution.

Rather than implementing a very large number of distinct evolution times precisely, it makes more sense to use a much smaller number of evolutions to approximate all the desired times. This is most easily achieved by dividing the range of evolution times by successive powers of two, in effect encoding each evolution time as a binary number. Using k different evolution times results in k-bit precision, with an exponential increase in precision with a linear increase in the number of evolution times used. By accepting a degree of approximation one can reduce the time complexity from $O(q \log q)$ to $O(kq)$, and reduce the pulse count from $O(q^2)$ to $O(kq)$, where the constant k depends on the accuracy required. For example, using 20 distinct delays will allow angles to be approximated to a precision better than 10^{-6}.

5.7 Conclusions

The colouring techniques described here allow the efficient control of interactions in qubit arrays of arbitrary size, provided the couplings are constrained to be local. While the resulting pulse sequences take slightly longer to implement than the absolute minimum time required, the computational time is vastly reduced, from $O(4^q)$ for exhaustive linear programming, or $O(q^6)$ for RROS [150], right down to $O(q)$, rendering them practical in systems with thousands or even millions of qubits. The number of control pulses is also greatly reduced, from $O(q^2)$ to $O(q)$, thus reducing implementation errors. The method can handle unwanted next-nearest neighbour couplings, and is easily extended in a scalable way to systems with variable coupling strengths or evolution angles.

Although it is not yet common to use always-on couplings as a computational resource in superconducting qubits, such implementations might fit well into the context of Near-term Intermediate Scale Quantum (NISQ) applications [28]. For example, the native interaction

can be used for hardware-efficient implementations of the variational quantum eigensolver (VQE) [355] and quantum approximate optimisation algorithms (QAOA) [356]. Since entangling gates are realized without external fields, there is less need to calibrate the control hardware, as the calibration problem can be focused only on local control. Furthermore, quantum error mitigation methods might allow slow native interactions to output useful results by extended time evolution [357–359]. Efficient refocusing and rescaling methods, such as those described here, will be essential to extend experiments to the next generation of devices.

Chapter 6

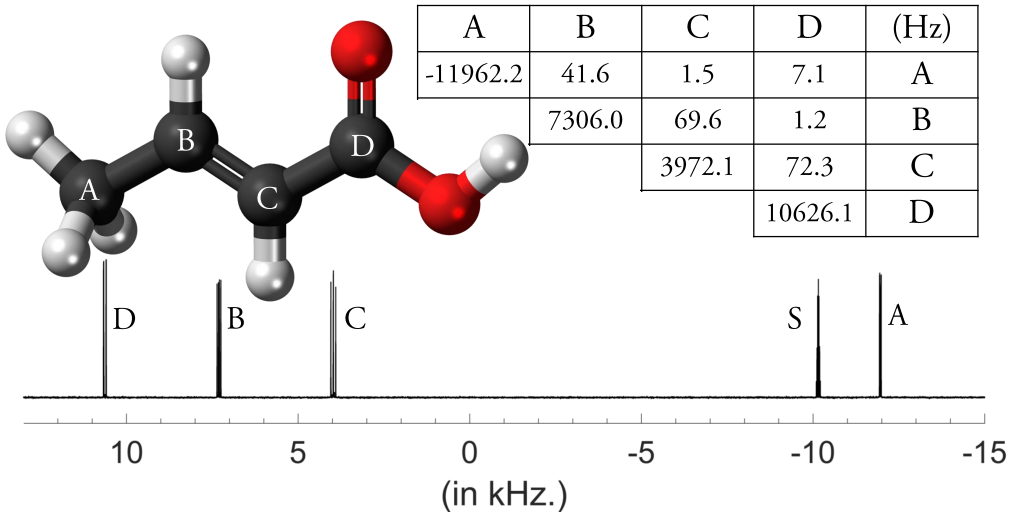
NMR Quantum Simulations

Liquid-state NMR has been extensively used to provide proof-of-principle demonstrations of various quantum algorithms, quantum simulations and other quantum information related phenomena [145]. Several of these demonstrations employ the GRAPE algorithm for controlling the unitary dynamics of the nuclear spins [156–179]. In this chapter, I illustrate the advantages and utility of the practical GRAPE methods developed in Chapter 3 through two different quantum simulations, published as references [152] and [153]. To begin, I provide some relevant details of the four-qubit NMR quantum processor that was used in both these quantum simulations.

6.1 A four-qubit NMR quantum computer

6.1.1 Qubits

The experimental qubits were provided by the four ^{13}C nuclei in a sample of fully labelled crotonic acid (Fig. 6.1) dissolved in deuterated acetone [339]. The Hamiltonian parameters (resonance offsets and J-couplings) as obtained on a 600 MHz (^{1}H frequency) Varian Unity Inova spectrometer at 300 K are given in Fig. 6.1. The spin system can be approximated by a linear chain, with strong nearest-neighbour couplings and much weaker next-nearest-



A	B	C	D	(Hz)
-11962.2	41.6	1.5	7.1	A
	7306.0	69.6	1.2	B
		3972.1	72.3	C
			10626.1	D

Figure 6.1: The molecular structure of ^{13}C labelled crotonic acid and Hamiltonian parameters: the diagonal elements denote the offset frequencies with respect to a base frequency of ≈ 150 MHz while the off-diagonal elements give the J-couplings. The multiplet labelled S in the spectrum comes from the solvent, deuterated acetone. Measured T_2 relaxation times were around 1.3 s for each spin, while T_1 varied between 10 s for spin A and 22 s for spin D.

neighbour couplings. The ^{13}C NMR spectrum of the thermal equilibrium state with ^{1}H decoupling (Fig. 6.1) shows that the four multiplets (one from each spin as indicated) are well separated and so can be individually addressed.

6.1.2 Quantum gates

In addition to four ^{13}C nuclei (the system qubits) crotonic acid also contains five ^{1}H nuclei (environment qubits) which have strong interactions with the system qubits. (The remaining three nuclei comprise two ^{16}O nuclei, which are spin-0 and so can be safely ignored, and a sixth ^{1}H nucleus in the hydroxyl group which undergoes rapid chemical exchange, averaging out its interactions with the main spins [131]). Thus, in order to design GRAPE pulses for the 4-spin system, it is necessary to effectively 'decouple' the environment qubits from the system qubits.

The traditional approach is to apply ^{1}H decoupling to the environment qubits throughout the experiment, usually with a composite pulse based broadband decoupling sequence such as WALTZ-16 [360]. In principle such decoupling can completely trace out the environ-

ment qubits, leaving a simple four-spin system. In practice, however, it is not possible to achieve completely effective decoupling without using high RF powers which are ruled out by hardware limits and the effects of sample heating. To support these observations, a high fidelity GRAPE pulse designed for an ideal 4-spin system was taken and its fidelity on a 9-spin system with WALTZ-16 decoupling was simulated for various decoupling strengths. Simulations suggested that for a near-perfect decoupling of protons, and therefore a high fidelity on the 9-spin system, it is necessary to perform WALTZ-16 decoupling at nearly five times the maximum advised decoupling power for our NMR equipment. As a result, this approach was ruled out.

The ideal approach is to design GRAPE pulses for the 9-spin system wherein the protons are overall left alone, *i.e.*, the target operator for the protons is an identity. Although this entirely avoids the need for actual ^{1}H decoupling, designing GRAPE pulse sequences for a $2^9 = 512$ dimension system is computationally infeasible, at least on a standard desktop computer. (Note that the three proton spins that constitute the methyl group are chemically equivalent, and can be treated as a single spin-1/2 entity under special circumstances [99]. However, for generality crotonic acid is treated as a 9-spin system.)

Here, I describe an approach that totally avoids actual proton decoupling yet has a computational complexity similar to that of a four-qubit system. The idea is to leave the ^{1}H nuclei completely untouched throughout the GRAPE pulse which implements the quantum gate, so that they can be treated as remaining in fixed energy eigenstates. In this case the ^{1}H nuclei are entirely passive, and can be thought of as providing a fixed frequency shift, which is different in each molecule, depending on the proton spin states in that particular molecule. As there are five proton spins there are $2^5 = 32$ possible spin states $\{00000, 00001 \ldots 11111\}$, although these give rise to only 16 genuinely distinct eigenstates as the three ^{1}H nuclei in the methyl group are completely equivalent. By considering each of these 16 eigenstates one at a time, 16 separate 'internal' Hamiltonians, each of dimension $2^4 = 16$ are obtained. The GRAPE pulses are then designed by optimising for all 16 internal Hamiltonians simultaneously, defining the overall fidelity as the average of the individual fidelities [155]. Therefore,

at modest computational overhead, the GRAPE pulses can be made robust to the environment qubits. Although this approach was independently worked out, it was later discovered that a very similar technique has already been used before [301].

6.1.3 Initialization

Thermal states

As mentioned earlier in Chapter 2, NMR is an ensemble quantum computing technique. NMR samples for QIP generally utilize a macroscopic ensemble of nearly 10^{17} to 10^{19} identical non-interacting molecules, wherein each molecule is an independent quantum computer comprising of one or more spin-1/2 nuclei as qubits. The state of qubits in NMR is determined by the collective state of all spins present in the ensemble and therefore best described using the density matrix formalism. The most natural and readily accessible state of the spin ensemble is where the spins are in thermal equilibrium with the environment. For a single qubit, this thermal equilibrium state can be expressed using the density matrix [132]:

$$\rho_{\text{eq}} = \frac{1}{\mathcal{Z}} \exp\left(-\frac{\mathcal{H}_0}{k_{\text{B}}T}\right) = \frac{1}{\mathcal{Z}} \begin{bmatrix} \mathrm{e}^{+\hbar\omega/2k_{\text{B}}T} & 0 \\ 0 & \mathrm{e}^{-\hbar\omega/2k_{\text{B}}T} \end{bmatrix}, \tag{6.1}$$

where $\mathcal{H}_0$ is the internal Hamiltonian, k_{B} is the Boltzmann's constant, T is the temperature of the environment and $\mathcal{Z} = \mathrm{e}^{+\hbar\omega/2k_{\text{B}}T} + \mathrm{e}^{-\hbar\omega/2k_{\text{B}}T}$ is the partition function that normalizes the state. The off-diagonal elements (coherences) of ρ_{eq} at thermal equilibrium are all zero [132], while the diagonal elements represent the Boltzmann distribution of populations in the ground ($E = -\hbar\omega/2$) and excited state ($E = +\hbar\omega/2$) respectively. At room temperature, where most liquid-state NMR experiments are performed, the thermal energy $k_{\text{B}}T \approx 25\,\text{meV}$ is much larger than the Zeeman splitting energy, which is of the order of only a few μeV for typical Larmor frequencies (below 1 GHz). Therefore, in the high temperature limit

($k_{\mathrm{B}}T \gg \hbar\omega$), the exponentials in ρ_{eq} can be approximated using a Taylor series to give

$$\rho_{\mathrm{eq}} = \frac{1}{2}\begin{bmatrix} 1+\hbar\omega/2k_{\mathrm{B}}T & 0 \\ 0 & 1-\hbar\omega/2k_{\mathrm{B}}T \end{bmatrix} = \frac{\mathbb{I}}{2} + \epsilon I_z, \tag{6.2}$$

where $\epsilon = \hbar\omega/2k_{\mathrm{B}}T \approx 10^{-5}$ is the spin polarization. Such a low spin polarization value indicates that both energy levels are almost equally populated, with only a slightly higher probability of occupancy of the ground state (about 1 part in 10^5). Thus, NMR states at room temperature thermal equilibrium are highly mixed. The thermal equilibrium state of a homonuclear n-spin system can be similarly written as

$$\rho_{\mathrm{eq}} = \frac{\mathbb{I}_n}{2^n} + \epsilon \sum_{i=1}^{n} I_{iz}, \tag{6.3}$$

where $\mathbb{I}_n$ is the n-spin identity matrix, $\epsilon = \hbar\omega/2^n k_{\mathrm{B}}T$ and the following additional approximations are made: (1) the J-couplings (few hundred Hz) are ignored as they are very small compared to the Larmor frequency (hundreds of MHz); (2) the Larmor frequency for all spins is considered to be identical as they all belong to the same nuclear species and also the chemical shift differences are small (few tens of kHz) compared to the Larmor frequency.

In order to implement a quantum algorithm or quantum simulation, it is usually necessary to initialize the qubits to a known pure state like the ground state $|000\ldots\rangle$. Initialization requires the quantum system to be placed in the desired pure state independent of its state prior to the initialization process. As a result, initialization is a non-unitary process, and cannot be achieved using only quantum logic gates. There are two obvious ways to prepare the ground state in any quantum system. One approach is to make a strong measurement on the qubits and use the result of that to transform the qubit's state into the ground state. However, as we will see in the next section, strong projective measurements cannot be performed in NMR, and only a weak ensemble measurement is possible in which the spin state is implicitly measured by the environment without actually changing it. The other approach is to cool the spins to a low enough temperature T such that the energy difference

ΔE between the ground and first excited state of the qubits satisfies $\Delta E \gg k_{\mathrm{B}}T$. To achieve this, either $T \ll 0.05\,\mathrm{K}$ is necessary, in which case the sample would no longer be in the liquid state, or if the sample is held at room temperature then a magnet with a field strength $B \approx 150000\,\mathrm{T}$ will be required, which is far beyond the largest NMR magnet currently available ($\sim$28 T). (Note that there is also a third way, which involves the preparation of non-equilibrium states wherein the spin temperature is very different from the bulk temperature. Techniques for lowering spin temperature by enhancing spin polarization have been used to make pure states in one-qubit and two-qubit systems [361–363] whereas in larger systems they have been used to get very large polarisation enhancements [157, 364], but not to get near pure states [365, 366].)

Pseudo-pure states

Even though the creation of a pure state sounds impractical in NMR, it is still possible to perform arbitrary quantum computations on a mixed state, provided the mixed state is made to be 'effectively pure' [54, 367] or 'pseudo pure' [52, 55]. Such a pseudo-pure state (PPS) can be represented by a density matrix

$$\rho_{\mathrm{pps}} = (1-\alpha)\frac{\mathbb{I}_n}{2^n} + \alpha\,|\psi\rangle\langle\psi|\,, \tag{6.4}$$

where $\mathbb{I}_n/2^n$ corresponds to the maximally mixed state (equal probability of system being found in any particular state) and the factor α is the excess probability of finding the system in the desired pure state, also referred to as the purity of the state [145]. Under unitary transformations, the pure component of the PPS evolves in the exact same way as a pure state would while the maximally mixed state remains unaffected and does not contribute towards the evolution. Thus, the PPS essentially mimics a pure state in terms of its dynamical behaviour.

In NMR descriptions, it is common to drop the maximally mixed state from the density matrix and describe the state using just the remaining component, referred to as the devia-

tion density matrix. This is because the maximally mixed state does not contribute to any observable signal in NMR as signals from different components of the identity matrix cancel out. In other words, all observables in NMR are traceless (described in the next section) and so the measurements are not sensitive to the maximally mixed state. Therefore, in a PPS, only the pure component will contribute to the observable signal. The size of this signal depends on the purity of the state. It was shown in Ref. [53] that the maximum possible signal that can be obtained from a PPS in conventional NMR falls exponentially ($\alpha \propto n/2^n$) with the number of qubits. The poor scaling of this signal intensity means that conventional NMR cannot be used to perform QIP with large numbers of qubits [53]. More fundamentally, the absence of entanglement in thermal pseudo-pure states has led some authors to question whether NMR devices are really quantum at all [368], implying that NMR experiments may be only simulations of simulations. It has, however, proved impossible to develop a fully classical model of NMR QIP [369], and NMR quantum computations run using entangled states produce identical results to those with pseudo-pure states with the exception of the increased signal size [362].

For a single spin, the thermal state I_z is already a PPS because of the excess population in the ground state compared to the excited state:

$$\rho_{\text{eq}} \sim \epsilon I_z = \begin{bmatrix} \frac{\epsilon}{2} & 0 \\ 0 & -\frac{\epsilon}{2} \end{bmatrix} = -\frac{\epsilon}{2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \epsilon \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \sim \epsilon \left|0\right\rangle \left\langle 0\right| . \tag{6.5}$$

However, this is not true in larger spin systems. For example, in a homonuclear two-spin system, the state corresponding to $|00\rangle\langle 00|$ PPS is $I_{1z} + I_{2z} + 2I_{1z}I_{2z}$ and not the thermal state $I_{1z} + I_{2z}$. The relative thermal state populations are $\{1, 0, 0, -1\}$, whereas the desired pattern for a PPS is $\{1, -1/3, -1/3, -1/3\}$. In order to generate a PPS from a thermal state in larger spin systems, it is therefore necessary to average the populations in all other energy levels while leaving the ground state untouched. Clearly, this task cannot be accomplished using only quantum gates as the eigenvalues of the two states (the diagonal elements, in this case) are different, and any unitary transformation preserves the eigenvalues. There are two

widely used non-unitary procedures for preparing a PPS, namely, temporal averaging [367] and spatial averaging [52, 55]. Here, I provide a brief overview of these methods and state the protocol used for initializing the 4-qubit quantum processor.

Temporal averaging

Although best known in NMR, temporal averaging can in principle be applied in any QIP technology. Temporal averaging is achieved by performing several different experiments and incoherently averaging their outcomes (to trace out the temporal information). If done correctly this can be interpreted as equivalent to averaging the states of the system before the measurement to give the desired density matrix. This approach works because of the linearity of quantum mechanics: the averaged output of a computation on individual input states $\sum_k U\rho_k U^\dagger$ is equivalent to the output of the computation on a single averaged state $U(\sum_k \rho_k)U^\dagger$.

Consider for instance the preparation of the ground state PPS $\{1, -1/3, -1/3, -1/3\}$ in a two-spin system. Here, the PPS can be achieved by averaging the thermal state $\rho_1 = \{1, 0, 0, -1\}$ and two other states $\rho_2 = \{1, 0, -1, 0\}$ and $\rho_3 = \{1, -1, 0, 0\}$ which can be generated by applying a network of CNOT gates to the thermal state. In general, a ground state PPS in an n-spin system can be obtained by averaging out the differences between the $2^n - 1$ smaller populations in the thermal state, while leaving the ground state alone. This would involve averaging the $2^n - 1$ cyclic permutations of these populations. In such exhaustive temporal averaging, clearly the number of initial states required and correspondingly the number of separate experiments that need to be performed increases exponentially with the number of qubits in the system. However, it is possible to substantially reduce the number of experiments in temporal averaging through a careful choice of selected initial states with appropriate weighting [86, 370]. For very large spin systems, it is even possible to choose a randomized permutation to prepare an approximate PPS [367].

Spatial averaging

Spatial averaging is similar to temporal averaging, except that the averaging takes places over spatial dimensions of spin ensembles instead of time. Spatial averaging methods in NMR commonly make use of magnetic field gradients. There are special coils in the NMR spectrometer to provide pulsed magnetic fields with strengths that vary approximately linearly in some spatial dimension of the NMR sample. The corresponding variation temporarily destroys the homogeneity of the static magnetic field, causing each spin in the ensemble to attain a position-dependent Larmor frequency and hence, a position-dependent phase. In other words, the effect of a gradient is to apply a z-rotation to each spin, where the angle of the z-rotation varies linearly over the spin ensemble. Subsequent measurements observe the average over the entire ensemble, and so an average over the phases. If the the gradient is strong enough then the effect is similar to dephasing each qubit.

There are, however, two differences between this process and dephasing. Firstly the dephasing is a collective process, and so systems in decoherence-free subspaces will not be affected (*e.g.*, zero-quantum coherences [145]). Secondly this dephasing is not genuinely non-unitary: the phase experienced by a member of the ensemble is dependent on its position in space, which is approximately constant, and so the phase evolution can be reversed by spin echoes. Non-unitary behaviour only arises when the spatial information is destroyed, either by the final observation process (which traces out position) or by diffusion or macroscopic motion within the sample (which is negligible when gradients are applied close together, but can become significant on longer time scales). In these cases, the dephasing due to gradients becomes irreversible and thus indistinguishable from true decoherence [91].

Magnetic field gradients in NMR-QIP are commonly used as 'crush' pulses to introduce non-unitary behaviour in the form of artificially induced decoherence. The effect of an ideal crush gradient can be described as crushing or destroying the coherences (off-diagonal elements, except zero-quantum coherences) from the density matrix while retaining only the populations (diagonal elements). The original spatial averaging sequence [52] which makes

use of crush gradients for creating a ground state PPS in a homonuclear two-spin system is as follows [145]:

$$\begin{aligned}
I_{1z}+I_{2z} &\xrightarrow{R_x^2(\pi/3)} I_{1z}+\frac{1}{2}I_{2z}-\frac{\sqrt{3}}{2}I_{2y} \\
&\xrightarrow{G_z} I_{1z}+\frac{1}{2}I_{2z} \\
&\xrightarrow{R_x^1(\pi/4)} \frac{1}{\sqrt{2}}I_{1z}-\frac{1}{\sqrt{2}}I_{1y}+\frac{1}{2}I_{2z} \\
&\xrightarrow{1/2J} \frac{1}{\sqrt{2}}I_{1z}+\frac{1}{\sqrt{2}}2I_{1x}I_{2z}+\frac{1}{2}I_{2z} \\
&\xrightarrow{R_{-y}^1(\pi/4)} \frac{1}{2}I_{1z}-\frac{1}{2}I_{1x}+\frac{1}{2}2I_{1x}I_{2z}+\frac{1}{2}I_{2z}+\frac{1}{2}2I_{1z}I_{2z} \\
&\xrightarrow{G_z} \frac{1}{2}I_{1z}+\frac{1}{2}I_{2z}+\frac{1}{2}2I_{1z}I_{2z}
\end{aligned} \tag{6.6}$$

where $R_x^i(\theta)$ is a rotation about x-axis by an angle θ for the i^{th} spin, G_z denotes a crush gradient and $1/2J$ indicates evolution under the zz-coupling Hamiltonian for a time $1/2J$. Note that the spatial averaging methods generally produce a PPS with reduced purity as compared to exhaustive temporal averaging approaches. For instance, the populations $\{3/4,-1/4,-1/4,-1/4\}$ for the PPS obtained from the above spatial averaging example correspond to a lesser purity than the exhaustive temporal averaging methods in which the obtained PPS populations $\{1,-1/3,-1/3,-1/3\}$ give the highest possible purity. This is because spatial averaging methods frequently sacrifice some of the available purity in order to simplify the PPS preparation sequence (the method of controlled transfer gates [371] is an exception). Nonetheless, spatial averaging requires only a single experiment as opposed to multiple experiments in temporal averaging.

The ground state PPS for the 4-qubit quantum processor was prepared by spatial averaging following the methods in Ref. [372]. For convenience, the pulse sequence is shown in Fig. 6.2. The relative populations of the 16 energy levels from an ideal implementation of this PPS pulse sequence are $\{15/16,-1/16,\cdots-1/16\}$. All quantum gates in the PPS sequence were implemented using robust GRAPE pulses as described in the previous section. It was observed that the use of regular GRAPE pulses (that require ^{1}H decoupling) in the PPS sequence resulted in nearly 50% signal loss as compared to the expected PPS signal

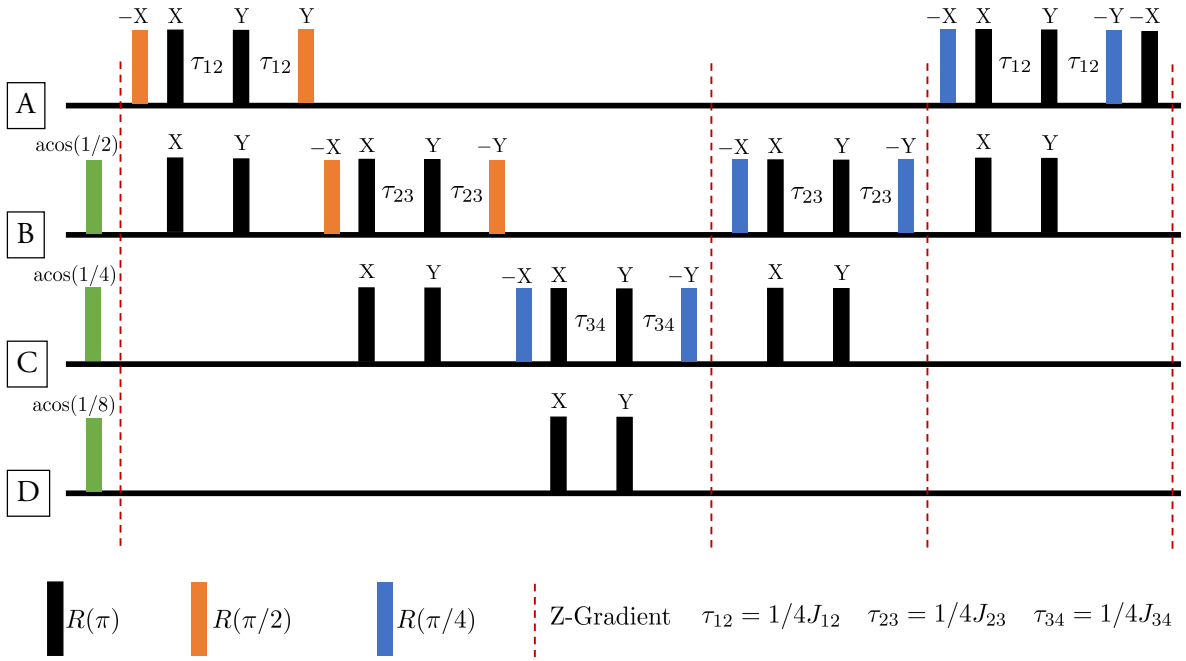


Figure 6.2: Pulse sequence for preparing a $|0000\rangle$ PPS.

intensity. In fact, mitigating this signal loss was the main motivation behind developing robust GRAPE pulses. To further improve the observed signal intensity of the PPS, the nuclear Overhauser effect (nOe) [373] from ^{1}H nuclei was used to generate a non-thermal initial state with enhanced polarization [137]. This was achieved through the application of a pair of composite π pulses on the protons separated by a delay of 6 seconds and a delay of 3 seconds between the second π pulse and the PPS sequence. Note that the PPS sequence assumes equal polarization on all spins to begin with, while the nOe enhancement for different ^{13}C spins within crotonic acid will in general be different (for example, the polarization enhancement for the methyl carbon is much greater than the carbonyl carbon). Suitable polarization on all spins can be achieved through an appropriate choice of initial flip angles (shown in green rectangles in Fig. 6.2).

6.1.4 Readout

A quantum computation is useful only if there is a provision to measure specific qubits and characterize the final state of the quantum system [39]. Ideally, it is desirable have the

ability to make measurements that project the state of the qubits onto the computational basis or any orthogonal basis. In such projective measurements, the measuring apparatus is usually made to interact strongly with the quantum system to cause the qubit's state to decohere instantaneously (compared to the duration of computational steps) into one of its eigenstates. Such 'strong' projective measurements, however, are not possible in conventional NMR.

As discussed previously, it is impractical to measure the magnetic moment of a single nuclear spin and hence, a large ensemble of nuclear spins is used in NMR. The process of measuring qubits in NMR is essentially the same as acquiring an NMR spectrum: the RF coil mounted next to the NMR sample tube records the oscillating magnetic field produced by the collective magnetic moment of the precessing spin ensemble to give a time-domain signal, which is then Fourier-transformed to give the familiar NMR spectrum. Usually, this measuring coil is the same as that used for producing RF pulses for controlling the spins. Since the measuring apparatus in NMR is permanently present and cannot be isolated from the spins at any point of time, it remains only weakly coupled to the spins and causes little or no decoherence. As a result, the spins decohere gradually (as opposed to instantaneously during a projective measurement) during which the RF coil merely monitors the state of the system without actually disturbing it, thereby performing a weak ensemble measurement.

The signal measured by the RF coil, also known as the free induction decay (FID), actually corresponds to the transverse component of the magnetic field produced by the precessing spins (the longitudinal component does not contribute to the signal as it does not precess). Formally speaking, the FID is a measure of the expectation values of the two traceless observables σ_x and σ_y. Due to the weak ensemble nature of measurements where information about individual spin states is unavailable but only their 'average behaviour' can be accessed, it is possible to simultaneously measure the expectation value of two non-commuting observables. The basic NMR measurement process (for a homonuclear n-spin system) can

be expressed as acquiring an FID $f(t)$ where

$$f(t) \propto \mathrm{Tr}\left[D^{\dagger}\rho(t)\right], \qquad D = \sum_{i=1}^{n}\left(I_{ix} - \mathrm{i}I_{iy}\right), \qquad \rho(t) = \mathrm{e}^{-\mathrm{i}\mathcal{H}_0 t}\rho(0)\mathrm{e}^{\mathrm{i}\mathcal{H}_0 t}. \tag{6.7}$$

The FID, which depends on $\rho(t)$, is a sum of components oscillating at frequencies corresponding to differences between frequencies in the internal Hamiltonian $\mathcal{H}_0$, which decay with a time constant T_2^* largely determined by T_2 relaxation and by the extent of the remaining inhomogeneities in the static magnetic field. These components can be separated by Fourier transformation, to give a spectrum comprising a series of lines with frequencies determined by $\mathcal{H}_0$ but with complex intensities that depend on $\rho(0)$ (the final state of the spins after the computation).

The form of the detection operator D means that only certain components of $\rho(0)$ are detectable. In NMR terminology these are called single quantum coherences, and can be thought of as states where one spin is in a superposition state and the rest are in eigenstates. Such states generate a signal line in the multiplet (group of lines) at a frequency corresponding to transitions of the spin in a superposition state, while the exact position of the line indicates the state of the other spins and the complex amplitude of the line depends on the amplitude and phase of that component in the overall state.

Consider the simplest case of a single spin for which the states $|0\rangle$ and $|1\rangle$ correspond to the NMR states I_z and $-I_z$ respectively. Applying a $R_y(\pi/2)$ pulse, also known as a read-out or detection pulse in NMR, converts these states to $\pm I_x$ which appear as absorption or emission lines in the NMR spectrum. For a two-spin system in a $|ab\rangle\langle ab| = (-1)^a I_{1z} + (-1)^b I_{2z} + (-1)^{a+b} 2 I_{1z} I_{2z}$ PPS, applying a read-out pulse on both the spins will convert the population terms to observable single quantum coherences $(-1)^a I_{1x} + (-1)^b I_{2x}$ while the remaining term becomes unobservable double and zero quantum coherence. The phases of the NMR signal (absorption or emission) can be used to extract the two-qubit state information. The exact details about the measurement and characterization procedure followed for each of the 4-qubit quantum simulations will be explained in their respective sections.

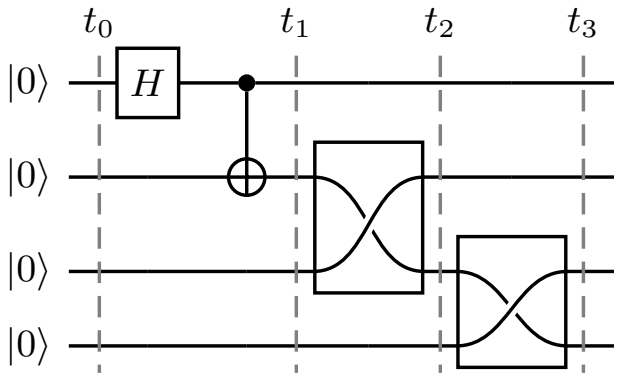


Figure 6.3: A quantum network for transferring the entanglement between a neighbouring pair of qubits to a distant pair of qubits.

6.2 Entanglement swapping

Based on the published work [152], in this section, I describe an NMR experiment for the quantum simulation of an entanglement swapping protocol. This protocol has connections with Marletto and Vedral's recently proposed scheme for witnessing quantum effects in gravity [374]. Since the theoretical connection between these protocols is well elaborated in the published work and also outside the scope of this thesis, here I focus only on the experimental aspects, essentially treating the quantum simulation as a task of quantum control. Consequently, my objective is to implement the quantum network shown in Fig. 6.3. The underlying theory for this work was developed by Prof. Vlatko Vedral and Dr Chiara Marletto, whereas I was involved in designing and performing the NMR experiments.

6.2.1 Description of the quantum simulation

We consider a system of four qubits in a linear chain labelled A, B, C and D along the chain. To begin with, qubits A and B are prepared in a maximally entangled state, for example, any Bell state. As shown in Fig. 6.3, the entangled state can be moved down the chain using SWAP gates acting locally on qubits B and C and qubits C and D. Qubits B and D are assumed to be 'quantum' whereas C can be either 'semi-classical' or 'quantum'. Depending on the nature of the mediator qubit C, the movement of entanglement is affected. The goal

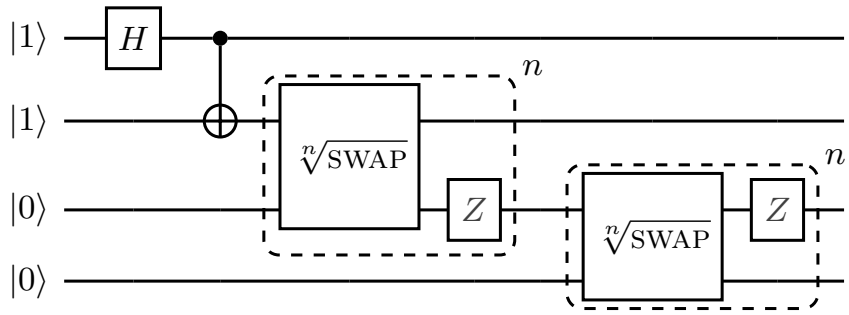


Figure 6.4: The quantum network for the simulation. The Z gates shown in grey are applied probabilistically to simulate decoherence, as described in the main text. The gates in each dashed box are repeated n times. Note that this circuit used for the simulation is slightly different from the original proposal, but easier to implement in NMR [152].

here is to show that if qubit C is quantum, then at the end of the network, qubits A and D are entangled; whereas if C is semi-classical then the transfer of entanglement does not happen, thus leaving qubits A and B fully entangled. In the context of Marletto and Vedral's proposal [374], qubit C manifests a gravitational field which can mediate local interactions. The proposal aims to indirectly test the non-classicality of a physical system that need not obey quantum theory (here, the gravitational field) by detecting whether or not the system is capable of entangling two quantum systems (in our case, the transfer of entanglement down the chain of qubits). Thus, a successful entanglement transfer through the mediator qubit C serves as an indirect witness of the non-classical nature of the gravitational field.

6.2.2 NMR experiments

Initialization

The circuit implemented in our experiment in shown in Fig. 6.4. To begin, a PPS was prepared using spatial averaging methods as described in Fig. 6.2. All gates used in the PPS preparation sequence as well as the quantum circuit were implemented using robust GRAPE pulses as discussed earlier. The initial state was chosen to be $|1100\rangle$ as this leads to a singlet entangled state, which is the most robust of the four Bell states to naturally occurring decoherence processes. The initial state $|1100\rangle$ can be obtained by applying selective NOT

gates to qubits A and B after implementing the PPS sequence.

SWAP gates and dephasing

The basic experiment looks at the effects of dephasing on the use of SWAP gates to move an entangled state down a chain. The idea is to break up the SWAP_{BC} and SWAP_{CD} gates into n steps using

$$U_{BC} = \sqrt[n]{\text{SWAP}_{BC}}, \qquad U_{CD} = \sqrt[n]{\text{SWAP}_{CD}}, \tag{6.8}$$

so that the SWAP process can be implemented by applying n U_{BC} gates and then n U_{CD} gates, and the effects of dephasing can be investigated by dephasing qubit C after every gate (in the case where qubit C is classical). For theoretical investigations it is useful to consider the limit $n \to \infty$, while for experiments n will inevitably be fairly small. The experiments were implemented using $n = 8$ which seems to be a sweet spot where the theoretical errors arising from n being too small and the experimental errors arising from n being too large are roughly balanced. Poor experimental results are obtained if the same GRAPE pulse is applied repeatedly, as the small errors which inevitably occur in any experimental implementation build up linearly on repeated application [375]. Instead multiple GRAPE pulses were designed for each gate, by using different random starting points in the GRAPE search. As each pulse has different implementation errors the total error will only grow with the square-root of the number of gates, leading to visibly better results.

Dephasing—a non-unitary process—can be realized in NMR using either spatial averaging or temporal averaging as discussed earlier. Here, we use the method of temporal averaging in which spectra are recorded both with and without the application of Z gates at any given point and the results combined. Temporal averaging is feasible if just a single dephasing step needs to be performed, but for the interrupted SWAP gate we need to dephase the qubit n times (or $2n$ for both SWAPs), and so we need to perform 2^n separate experiments, corresponding to the 2^n different combinations of Z gates or not. This exhaustive temporal averaging process swiftly becomes impractical as it requires 256 combinations for $n = 8$, or

65536 combinations for both SWAPs.

Instead we adopted the method of randomized temporal averaging [367], in which exhaustive averaging is replaced by averaging over a sample of possible sequences. In this randomized approach, dephasing can be thought of as applying or not applying a Z gate with probability 50%, so any given pattern of Z gates can be thought of as a random sample from the dephased possibilities. Ensemble-averaged dephasing is just the average over all these possibilities, and if the fidelity of the ensemble average to some theoretical model is high then the fidelity of almost all members of the ensemble must individually be fairly high. Thus averaging is not really necessary to see the desired overall behaviour, but is simply helpful to 'average out' the (relatively small) error terms. The limiting case, just observing a singly random chosen pattern, will work unless the randomly chosen pattern happens to be one of the small number which work really badly, such as the case where no Z gates are applied at all. However in the case that n is large enough the probability of choosing a really bad pattern is essentially negligible. With small n a bad pattern is more likely, but this can be avoided by making 'quasi random' choices, for example by insisting that a Z gate is applied exactly half the time.

To this end, we specifically designed two further gates, $V_{BC} = Z_C U_{BC}$ and $V_{CD} = Z_C U_{CD}$, and used a randomly chosen sequence of U and V gates, containing four of each. When the number of stages is small, as used here, the final result depends on the precise pattern of gates used, and so results were averaged over 16 different dephasing patterns.

Readout

Entangled states are not directly detectable in NMR: the observed signal depends on the expectation values of single-spin off-diagonal operators [376], and these are all zero for entangled states. Here, however, we use an indirect entanglement witness [377].

Consider the initial entangled state of qubits A and B, and the effect of performing a Hadamard operation on qubit A before observing the NMR spectrum. This will gener-

ate a pair of lines in the multiplet of transitions of spin A, one with positive intensity and another with negative intensity, called an antiphase doublet. However it also generates another antiphase doublet in the multiplet of spin B, even though spin B was not directly excited. There are no signals in the multiplets corresponding to spins C or D. In general, if we have an entangled singlet state between two qubits then applying a Hadamard gate to one of them generates antiphase doublets in both multiplets, so we can very easily follow the progress of the entangled state from AB to AD by applying a Hadamard to qubit A and observing the signal on spins B, C and D.

As the lines are antiphase doublets there is a potential worry that the two lines, which have equal and opposite intensity, will partly cancel, and this will certainly occur if they are too close to one another. The separation is given by the size of the relevant J-coupling constant. For crotonic acid the relevant coupling sizes are $J_{12} = 41.6\,\text{Hz}$, $J_{13} = 1.5\,\text{Hz}$, and $J_{14} = 7.1\,\text{Hz}$, while the linewidth is around $1\,\text{Hz}$. This means that cancellation will be a concern for the middle state (qubits A and C entangled), but the initial and final states which we want to observe should be clearly resolved.

Experimental results

The experimental results are shown in Fig. 6.5. Each spectrum is obtained after applying a Hadamard gate to spin A with a two step phase cycle to reduce errors [145]; the ^{1}H environment nuclei were decoupled throughout signal acquisition. Individual multiplets were then cut out of spectra like that shown in Fig. 6.1, permitting an expanded horizontal scale, and then rearranged into spin order. Antiphase signal on both spins r and s with a coupling J_{rs} indicates that spins r and s were entangled; smaller peaks are due to remaining experimental errors.

An AB singlet state was initially prepared, as shown by the antiphase doublets split by J_{AB} on both spins A and B in spectrum (a). The transfer of entanglement from spins A and B to spins A and D is clearly seen in (b), while dephasing suppresses this transfer as shown in

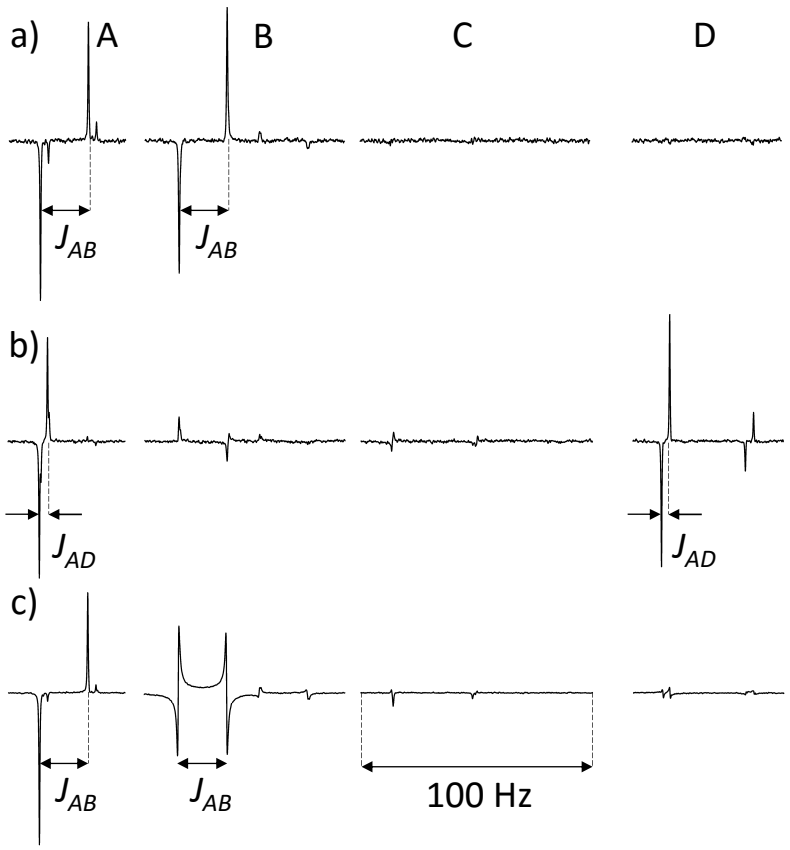


Figure 6.5: Experimental spectra of the ^{13}C spins (A, B, C, D): a) in the initial singlet state; b) after the SWAP gates; c) after the SWAP gates interrupted by dephasing on spin C. All spectra are plotted on the same vertical scale, and the horizontal frequency scale is as indicated. The higher signal-to-noise ratio in c) reflects the averaging over 16 different dephasing patterns.

(c). This suppression can be seen as a generalisation of the quantum Zeno effect in NMR [378] where rapid dephasing suppresses coherent evolution. The phase shift on spin B in spectrum (c) arises from the zz component of the SWAP Hamiltonian which commutes with the dephasing process and so is not suppressed by the Zeno effect.

The small additional peaks are due to experimental errors, and simulations suggest that these can be largely modelled by weak uniform depolarisation during the SWAP process. This apparent depolarisation arises from the accumulation of small errors in the U and V gates, which can be treated as random because of the way in which particular implementations of these gates are selected from a larger set.

Conclusions

We have implemented an entanglement swapping protocol on an NMR quantum processor comprising of four qubits arranged in an approximate linear chain configuration. The experiment demonstrates the local transfer of entanglement created at one end of the chain to the other through a mediator qubit which may be non-classical. The experimental results are consistent with the theoretical predictions: when the mediator behaves classically, the transfer of entanglement does not take place whereas in the case of a 'quantum' mediator, entanglement is established across the chain. This proof-of-principle demonstration of entanglement swapping serves as an indirect witness for illustrating the quantum nature of the gravitational field [374].

6.3 Quantum homogeniser

The quantum homogeniser can transform a qubit initialised in any state to any other state. It was originally proposed as a theoretical model for analysing many-body entanglement within unitary thermalisation, with an additional possible application as a quantum safe [379, 380]. More generally, it can be used to implement processes such as quantum informa-

tion scrambling and pure state preparation using only unitary interactions. In this section, I describe a 4-qubit NMR implementation of the quantum homogeniser along with a comparison of the limiting cases of transforming a qubit from a mixed to a pure state and the reverse process. This work is published as Ref. [153] and as in the previous simulation, my role here was designing and performing the NMR experiments.

6.3.1 Description of the quantum simulation

The quantum homogeniser is a machine consisting of N identical reservoir qubits. These each interact, one by one, with the system qubit (the qubit whose state is to be transformed) via a unitary partial swap:

$$U = \cos\eta\,\mathbb{I} + \mathrm{i}\sin\eta\,S. \tag{6.9}$$

The partial swap is a combination of the identity, $\mathbb{I}$ (which does nothing to the two input qubits), and the SWAP operation, S (which swaps the states of the two input qubits), weighted by the coupling strength parameter η. The quantum circuit used in our simulation is shown in Fig. 6.6. This circuit simulates the homogenisation between two system qubits, initially in the pure state $|0\rangle$, and two reservoir qubits initially in the maximally mixed state ρ_{MM}, and has been designed for implementation on a linear chain where two-qubit gates are only available between adjacent qubits. Partial swap gates are implemented between qubits B and C, while full SWAPs are used to rotate the system and reservoir registers, permitting indirect contact between any pair of qubits.

The four qubit state before the final readout stage depends on the coupling strength η, but in general is an entangled state. However the final readout stage involves an implicit partial trace leaving four separate single-qubit states, all of which lie along the z-axes of the respective Bloch spheres, permitting the state to be fully characterised by measurements in the computational basis. We can write for each qubit X

$$\rho_X = \frac{\mathbb{I}}{2} + f_X(\eta) \times \frac{\sigma_z}{2} \tag{6.10}$$

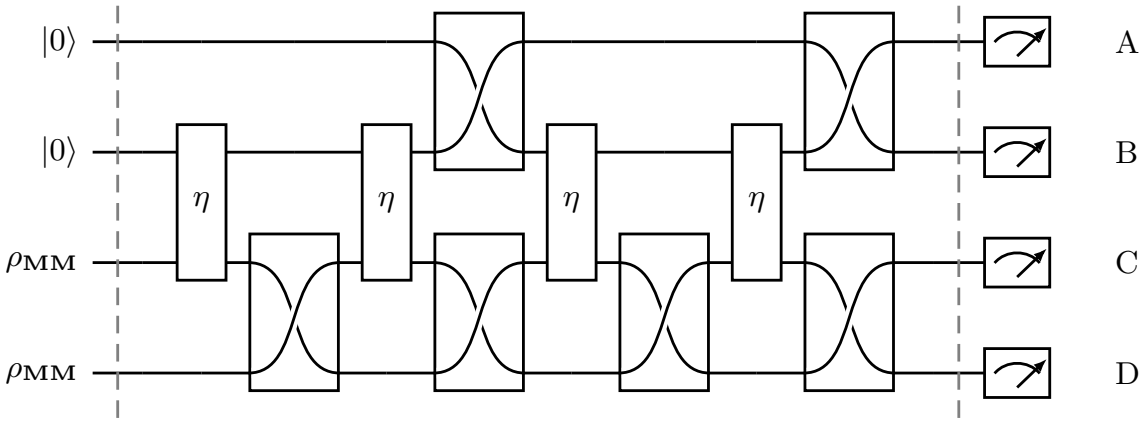


Figure 6.6: A quantum homogeniser with two system qubits and two reservoir qubits designed to operate on a linear chain system where two qubit gates are possible only between adjacent qubits. Homogenisation is achieved using the partial swap gates, labelled as η, which connect the two middle qubits, while full SWAP gates are used to rotate the system and reservoir registers to bring other qubit pairs into contact. The dashed lines divide the simulation into its initialisation, homogenisation and readout phases. Qubit labels on the right hand side correspond to spins in the NMR spin system in Fig. 6.1.

where f_X lies between 0 and 1 and corresponds to the difference between the probabilities of finding $|0\rangle$ or $|1\rangle$ when measuring qubit X in the computational basis. In particular we find the forms

$$f_B = \cos^4(\eta) \tag{6.11}$$

and $f_C = 1 - f_B$, so that f_B falls smoothly from 1 to 0, while f_C rises in the opposite way. Similarly

$$f_A = 4\cos^2(\eta) - 9\cos^4(\eta) + 8\cos^6(\eta) - 2\cos^8(\eta) \tag{6.12}$$

also falls from 1 to 0, but following a more complex pattern, with $f_D = 1 - f_A$ once again rising in the opposite way.

As the coupling strength η is increased the homogenisation becomes more effective, and for circuits with the same number of system and reservoir qubits the states are completely interchanged in the limiting case $\eta = \pi/2$. This symmetry permits the roles of the system and reservoir qubits to be interchanged, and so this homogenisation circuit can equally well be viewed as a process that randomises the pure qubits or polarises the mixed qubits.

6.3.2 NMR experiments

Initialisation

The thermal equilibrium state for the four-spin system is given by the deviation density matrix $\rho_{\text{eq}} \sim (A_z + B_z + C_z + D_z)$, where $A_z = (\sigma_z/2) \otimes \mathbb{I} \otimes \mathbb{I} \otimes \mathbb{I}$, and similarly for the other three terms. In this simulation, spins A and B need to be initialized in the state $|00\rangle$ and spins C and D in the maximally mixed state, which corresponds to the deviation density matrix $A_z + B_z + 2A_zB_z$. In principle, this two-spin ground state PPS can be prepared using any of the state preparation techniques discussed earlier. However, in this particular simulation, it suffices to use the simpler state $A_z + B_z$ as the initial state because there is no visible signal from the $2A_zB_z$ term. In the final readout stage here, we are only interested in measuring the single-qubit populations (integral of each of the four multiplets), unlike the previous simulation which involved the readout of multi-qubit entangled states (where the line structure within each multiplet is important). As a result, it was necessary to prepare a PPS in the previous simulation, whereas, here it suffices to use $A_z + B_z$ as the initial state which can be trivially prepared from the thermal state by applying 90° excitation pulses to spins C and D followed by a crush gradient [145].

Homogenisation

The homogeniser circuit was implemented using robust phase-only GRAPE pulses as discussed earlier. The GRAPE pulses were designed to implement each logic gate in Fig, 6.6, except that when two SWAP gates are written vertically above one another a single GRAPE pulse was used to implement both gates together. Separate pulses were prepared for ten different partial swap gates, varying η in 10° steps between 0 (an identity gate, for which the simplest implementation is just to omit the gate entirely) and 90°, corresponding to a full SWAP.

Readout

The circuit in Fig. 6.6 assumes conventional projective measurements in the computational basis, so that it is necessary to repeat the experiment many times to estimate the z-component of the Bloch vector describing each qubit. Here the ensemble nature of NMR comes into its own, as NMR spectra directly reveal the desired expectation value.

Direct observation of the NMR signal (FID) [131] reveals the expectation value of the x and y components of the Bloch vector. To observe the z-component we first apply a crush gradient, dephasing any pre-existing xy components, and then apply a 90° pulse to excite all four spins. The integrated signal intensity of each of the four multiplets seen in Fig. 6.1 is then proportional to the desired z-component. Note that integrating to find the total signal in each multiplet is equivalent to performing a partial trace [381].

Because we are integrating each multiplet it is not necessary to apply ^{1}H decoupling even during readout. However it is desirable to do so, as this reduces the width of each multiplet and so reduces the effects of noise in the integrated signal. To obtain accurate integrals it is important to process the data carefully, paying attention to phasing and baseline correction [382]. As NMR signal intensities are only *proportional* to the desired z-component it is essential to obtain a suitable reference intensity against which all other intensities can be normalised. In the results below we use two different choices of normalisation, which emphasise different features of the experimental results. The use of normalised intensities means that experimental errors can take measured values of f_X slightly outside the theoretical limits of ± 1, and such apparently unphysical values should not cause concern.

Results

The experimental results are shown in Fig. 6.7, with the two sub-figures corresponding to the two different normalisation choices. In each case the lines show the expected polarisations for each of the four qubits calculated as described in the previous section, while data-points

show the measured polarisations on the corresponding spins. Each experiment was repeated ten times, with the error bars showing the standard deviation around the mean.

The upper panel shows the results of normalising each intensity with respect to the average intensity of the A and B multiplets in the initial state. The experimental signal intensity is systematically lower than expected, with almost all data-points lying below the theoretical lines. The agreement for spins A and B is generally better than for spins C and D, although in all cases agreement is best close to $f_X = 0$.

This general pattern of signal loss is easily explained as arising from the inevitable errors in any experimental implementation. Decoherence will normally lead to signal loss, and so will coherent errors which become effectively incoherent when averaged over the experimental ensemble, for example over different RF powers in different parts of the sample. Finally the readout process, including the initial crush gradient and the implicit partial trace, itself removes all terms in the density matrix other than single spin z-magnetisation. Thus errors of any kind can *only* appear as a change in the measured values of f, usually reducing these towards zero.

The experimental asymmetry observed between the erasure of pure spins A and B and the polarisation of initially mixed spins C and D, in comparison with the symmetry of the theoretical predictions, is also easily understood. For spins A and B the signal loss acts in the same direction as the quantum homogeniser process, but for spins C and D it acts against the desired process, making the effects easier to see.

The lower panel shows the results when each spin was individually normalised against its intensity in the $\eta = 0$ spectrum (for spins A and B) or $\eta = 90°$ spectrum (for spins C and D). This removes the effects of signal loss during the SWAP gates, and for C and D also removes losses due to partial swap gates. The experimental data-points now lie much closer to the theoretical predictions and the expected symmetry is largely restored.

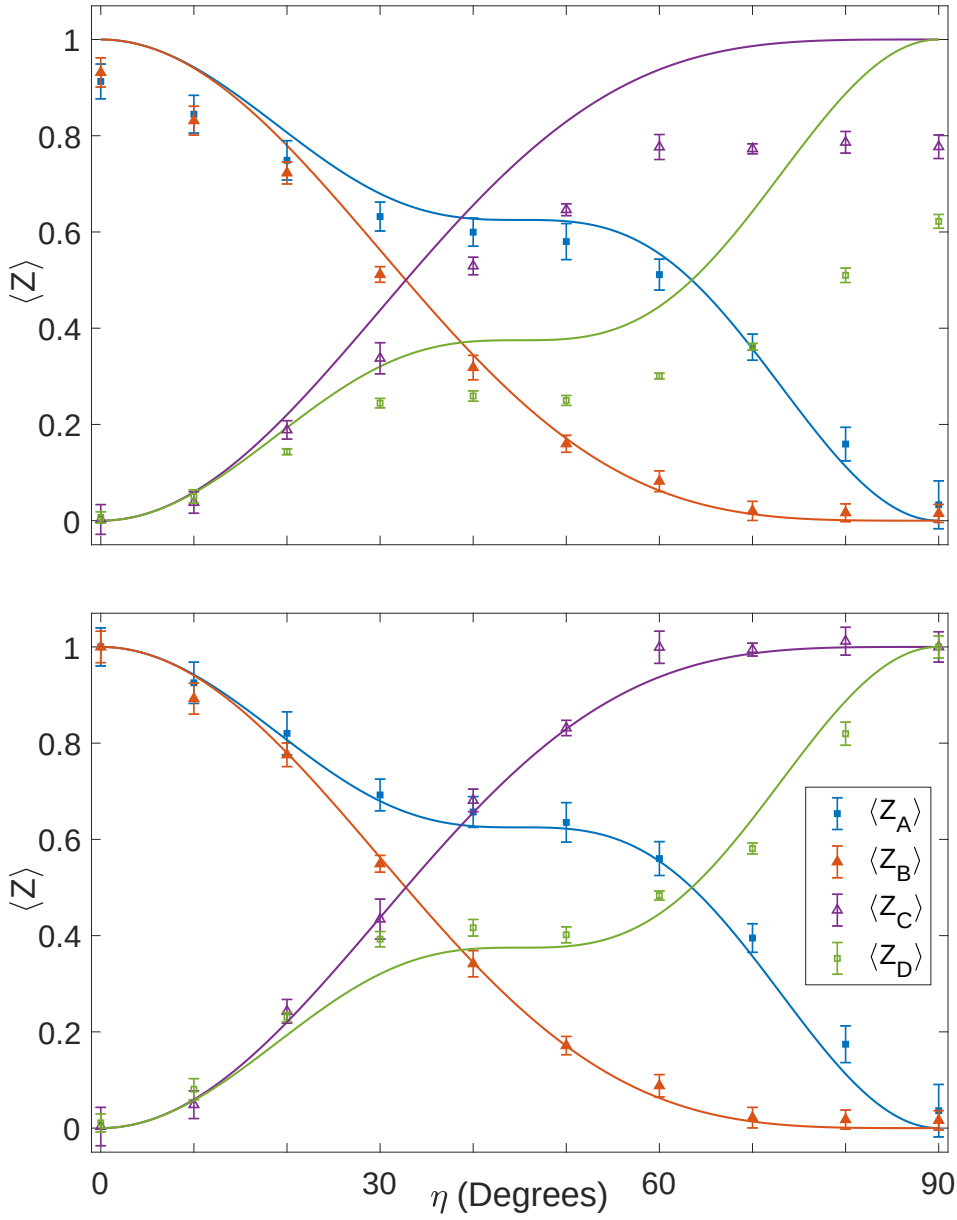


Figure 6.7: Experimental results and theoretical predictions for the quantum homogeniser. Data-points show the results from ten repetitions of the experiment. For the upper panel signal strengths were normalised against the initial state, while in the lower panel the signal from each spin is normalised against its own behaviour.

Interpretations

The experimental data show the convergence of the system qubit to the state of the reservoir qubits. This has been tested for the limiting cases of transforming a qubit from a mixed state to a pure state and from a pure to a mixed state, demonstrating that the homogenisation is effective regardless of the initial states of the system or reservoir qubits. After accounting for the bias towards mixed states caused by decoherence, the experimental results are consistent with the theoretical symmetry in the evolution of how the states evolve, where the states for the pure-to-mixed homogenisation vary inversely compared to the states for the mixed-to-pure homogenisation.

Pure state preparation—Figure 6.8 plots the von Neumann entropy

$$S = -\left(\frac{1+f}{2}\right)\log_2\left(\frac{1+f}{2}\right) - \left(\frac{1-f}{2}\right)\log_2\left(\frac{1-f}{2}\right) \tag{6.13}$$

of the theoretical and experimental qubit states against coupling strength. Theoretical curves were calculated using equation 6.11 for f_B and $f_C = 1 - f_B$. The qubit B is the system qubit for the pure-to-mixed homogenisation, while C is the system qubit for the mixed-to-pure homogenisation. As expected, the qubit being transformed from a mixed to a pure state decreases in entropy, with the effect being strongest for strong coupling, while the qubit being transformed from a pure to a mixed state increases in entropy. Since all the interactions are unitary, the total von Neumann entropy of the combined homogeniser and system must remain constant. Hence it can be deduced that the entropy decrease of the mixed-to-pure system qubit C must be accompanied by an increase in the entropy of the homogeniser (qubits A and B), which is the irreducible entropic cost associated with preparing a pure state.

Scrambling—The initial pure or mixed state of a system qubit becomes indistinguishable from the original state of the homogeniser qubits, which is a special case of information scrambling. As explained in [379], the information about the system's initial state becomes

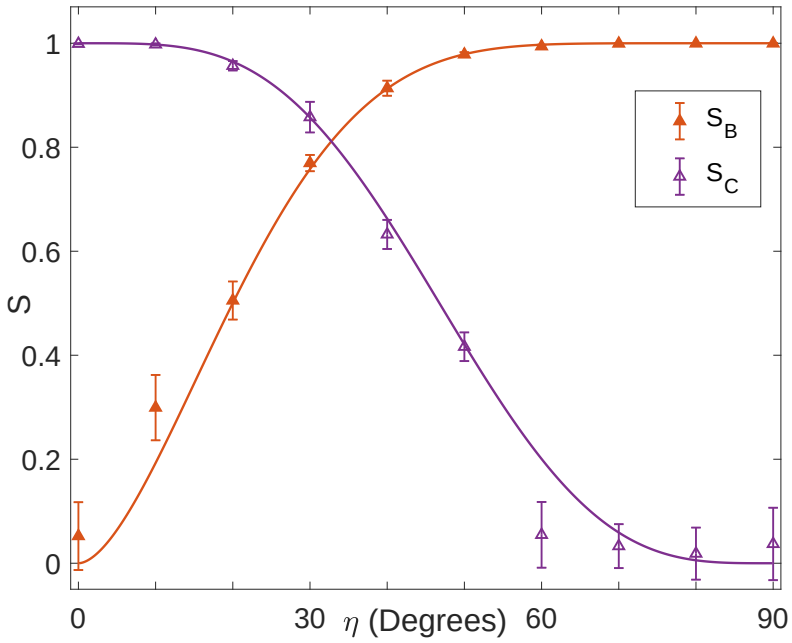


Figure 6.8: Von Neumann entropies of qubits B and C against coupling strength, calculated from the experimental results and plotted against the theoretical predictions.

hidden in mutual correlations between the homogeniser qubits. If there were no mutual correlations, one would expect the sum of the von Neumann entropies of the four qubits to equal two for all coupling strengths. The actual sum of the von Neumann entropies is in Figure 6.9. While this is two for the cases of an identity or a SWAP operation, for intermediate coupling strengths it is larger. This indicates that the negative contribution to von Neumann entropy from mutual correlations has been unaccounted for, which is due to considering only reduced density operators to describe the qubit states.

Reusability— The same homogeniser could be reused for a second time (or more) to successfully homogenise more system qubits. Qubit A can be interpreted as interacting with a homogeniser that has already been 'used' once, to homogenise B. Similarly, qubit D can be interpreted as being homogenised by a homogeniser that has already been used to homogenise C. These second system qubit states are closely aligned with the first system qubit counterparts for weak coupling, and diverge for strong coupling, in figure 6.7. This indicates that the homogeniser is minimally changed from its original state in the weak coupling regime, allowing it to perform just as effective a homogenisation on the second system qubit as it did on the first. Hence, the homogeniser can be reused to some extent to give the same

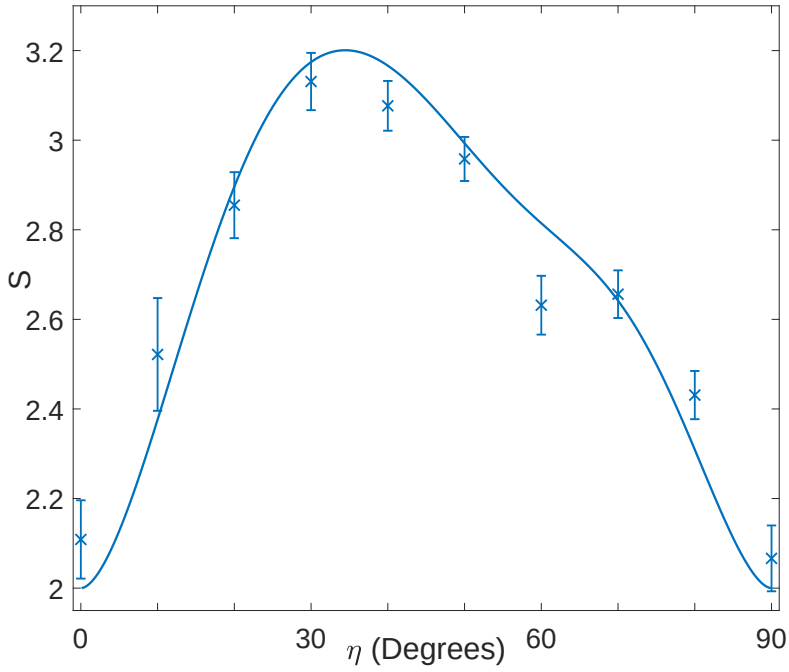


Figure 6.9: Total von Neumann entropies of the four individual qubits, calculated from the experimental results and plotted against the theoretical predictions.

incremental changes in system qubit state, in the weak coupling regime.

Conclusions

We have performed an NMR demonstration of the quantum homogeniser, using the partial swap on a system of four qubits. This demonstrates the principle behind a machine that can perform processes such as information erasure and the preparation of pure states using entirely unitary interactions. The experiments show the homogenisation of a pure state and of a mixed state. The asymmetry in the evolution of the states with coupling strength can be explained by the decoherence within the experiment. After accounting for this, the results are consistent with the theoretical symmetric evolution of the pure and mixed qubit states.

Chapter 7

Other Quantum Systems

In this chapter, I illustrate the utility of the GRAPE algorithm in implementing robust quantum gates in two different quantum technology platforms, namely, polar molecules [383–386] and superconducting qubits [326, 387, 388]. Unlike NMR spins which are natural qubits, a qubit in both these systems is defined as part of a larger multi-level system and hence, from a quantum control perspective, there is a common challenge associated with these two systems, which is to restrict the action of a designed quantum gate to the qubit space only, without affecting other energy levels. Based on the published work in Ref. [154], I first describe the implementation of an entangling gate using BFGS-GRAPE in a system of polar molecules. This collaborative project was about a new theoretical proposal for controlling polar molecules using magnetic and microwave fields, and my primary contribution was towards developing the code for GRAPE optimization. Next, I provide a brief overview of the ongoing work with Takahiro Tsunoda and Dr Peter Leek on GRAPE pulse design for implementing robust gates in superconducting qubits.

7.1 Polar molecules

Polar molecules are an emerging platform for quantum technologies based on their long-range electric dipole–dipole interactions, which open new possibilities for quantum information

processing [383, 384, 389, 390], simulation [391–395] and sensing [396, 397]. The discussion in this section is concerned with molecular species like CaF or RbCs which have been laser cooled to temperatures below 10 μK [398–401] and trapped in a two-dimensional array using an optical lattice [402, 403] or an array of optical tweezers [401, 404, 405]. The aim here is to design a high-fidelity fast entangling gate that is robust with respect to fluctuations in the trapping and control fields and to small thermal motional excitations. Without going into the details of the experimental proposal explained in Ref. [154], here I simply describe the problem from a quantum control perspective. I begin with a description of the Hamiltonian for a polar molecule influenced by magnetic and microwave (MW) fields.

Single-molecule Hamiltonian

Consider a polar molecule influenced by a homogenous magnetic field of strength B_0. This applied magnetic field separates the Zeeman components of the hyperfine levels within the rotational manifold of the molecule. The qubit space $\{|1\rangle, |0\rangle\}$ is defined by selecting a particular pair of Zeeman states within a rotational manifold (as shown in Fig. 7.1) that can be coupled using microwave (MW) fields with negligible off-resonant excitation to other states. The Hamiltonian for this system in the $\{|1\rangle, |0\rangle\}$ basis using natural units ($\hbar = 1$) can be written as:

$$\mathcal{H} = \frac{\overline{E}}{2}\mathbb{I}_2 + \frac{\omega}{2}\sigma_z + \Omega\cos(\omega_{\text{mw}}t)\,\sigma_x, \tag{7.1}$$

where,

- $\overline{E} = E_1(B) + E_0(B)$, with $E_j(B)$ the energy of state $|j\rangle$ as a function of B,
- $\mathbb{I}_n$ is the $n \times n$ identity matrix,
- $\omega = E_1(B) - E_0(B)$ is the splitting energy,
- $\omega_{\text{mw}} \approx \omega$ and Ω are the frequency and amplitude respectively of the MW field,
- $\sigma_z = |1\rangle\langle 1| - |0\rangle\langle 0|$ and $\sigma_x = |1\rangle\langle 0| + |0\rangle\langle 1|$ are the Pauli operators.

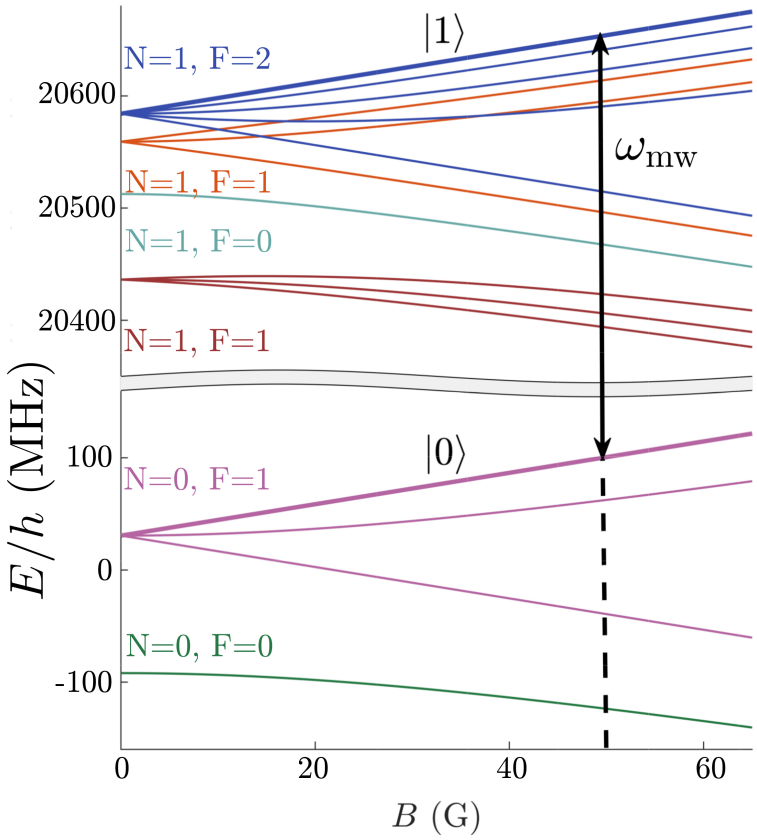


Figure 7.1: Energy levels of the $N = 0$ and $N = 1$ rotational manifolds of the vibrational ground state of $^{40}Ca^{19}F$ as a function of magnetic field B, where N is the rotational quantum number. The arrow indicates the transition addressed by the MW field to dress the states $|0\rangle = |N = 0, F = 1, M_F = 1\rangle$ and $|1\rangle = |N = 1, F = 2, M_F = 2\rangle$, where F is the total angular momentum quantum number of the molecule, and M_F its projection on the z axis, defined by the magnetic field. At $B_0 = 50\,\mathrm{G}$, the resonant transition frequency, ω, is 20.778 GHz.

In order to simplify the dynamics of the system, it is useful to express the above Hamiltonian in a rotating frame. Moving into a frame that rotates at a frequency ω_{mw} and invoking the rotating wave approximation (assuming $|\Delta|, \Omega \ll \omega_{\mathrm{mw}}$, where $\Delta = \omega - \omega_{\mathrm{mw}}$ is the detuning), we obtain the rotating frame Hamiltonian

$$\widetilde{\mathcal{H}} = \frac{\overline{E}}{2}\mathbb{I}_2 + \frac{\Delta}{2}\sigma_z + \frac{\Omega}{2}\sigma_x = \frac{1}{2}\begin{pmatrix} \overline{E}+\Delta & \Omega \\ \Omega & \overline{E}-\Delta \end{pmatrix}, \tag{7.2}$$

whose eigenenergies are

$$E_{\uparrow,\downarrow} = \frac{\overline{E}}{2} \pm \frac{1}{2}\sqrt{\Delta^2+\Omega^2}. \tag{7.3}$$

This Hamiltonian is analogous to single-qubit Hamiltonians encountered in other quantum-information platforms such as trapped ions or superconducting circuits. Single-qubit operations can be implemented by changing Ω which can be controlled in the MW regime. Note that although written here as a real number, Ω in general represents a complex parameter, which corresponds to amplitude control in two independent quadratures or equivalently amplitude and phase control of the MW field.

Two-molecule Hamiltonian

Now consider a system of two molecules subject to the same magnetic and MW fields described by the Hamiltonian

$$\mathcal{H} = \mathcal{H}^{\mathrm{A}} \otimes \mathbb{I}_2 + \mathbb{I}_2 \otimes \mathcal{H}^{\mathrm{B}} + \mathcal{H}_{\mathrm{ddi}}, \tag{7.4}$$

where $\mathcal{H}^{\mathrm{A}}$ and $\mathcal{H}^{\mathrm{B}}$ are the single-molecule Hamiltonians as in Eqn. 7.1 for molecules A and B respectively, and

$$\mathcal{H}_{\mathrm{ddi}} = V\sigma_x^{\mathrm{A}} \otimes \sigma_x^{\mathrm{B}} = V\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \tag{7.5}$$

is the dipole–dipole interaction (DDI) and V is the coupling strength. Writing this Hamiltonian in the rotating frame and making the rotating wave approximation as before (assuming $|\Delta|, |\Omega|$, and $|V| \ll \omega_{\mathrm{mw}}$), we get

$$\widetilde{\mathcal{H}} = \overline{E}\mathbb{I}_4 + \frac{\Delta}{2}(\sigma_z^{\mathrm{A}} \otimes \mathbb{I}_2 + \mathbb{I}_2 \otimes \sigma_z^{\mathrm{B}}) + \frac{\Omega}{2}(\sigma_x^{\mathrm{A}} \otimes \mathbb{I}_2 + \mathbb{I}_2 \otimes \sigma_x^{\mathrm{B}}) + V(\sigma_+^{\mathrm{A}} \otimes \sigma_-^{\mathrm{B}} + \sigma_-^{\mathrm{A}} \otimes \sigma_+^{\mathrm{B}}) \quad (7.6)$$

where $\sigma_+^j = |1\rangle\langle 0|$ and $\sigma_-^j = (\sigma_+^j)^\dagger$ are the raising and lowering operators respectively in the qubit space of molecule j. The terms involving Δ and Ω arise from the single-molecule coupling to the MW field, while the last term describes the dipole–dipole interaction in the rotating frame. This comprises exchange processes of the form $|01\rangle \leftrightarrow |10\rangle$. Double-flip processes (*i.e.*, transitions $|00\rangle \leftrightarrow |11\rangle$) involve the absorption or emission of two MW photons and are neglected in the rotating wave approximation. In the basis $\{|11\rangle, |10\rangle, |01\rangle, |00\rangle\}$, this Hamiltonian can be written in matrix form as

$$\widetilde{\mathcal{H}} = \overline{E}\mathbb{I}_4 + \begin{pmatrix} \Delta & \Omega/2 & \Omega/2 & 0 \\ \Omega/2 & 0 & V & \Omega/2 \\ \Omega/2 & V & 0 & \Omega/2 \\ 0 & \Omega/2 & \Omega/2 & -\Delta \end{pmatrix}. \quad (7.7)$$

It is useful to express $\widetilde{\mathcal{H}}$ in the triplet-singlet basis $\{|11\rangle, |\Psi^+\rangle, |00\rangle, |\Psi^-\rangle\}$, where $|\Psi^\pm\rangle = (|01\rangle \pm |10\rangle)/\sqrt{2}$, to see explicitly how the symmetric and antisymmetric subspaces decouple:

$$\begin{aligned} \widetilde{\mathcal{H}} = \overline{E}\mathbb{I}_4 \;&+\; \Delta(|11\rangle\langle 11| - |00\rangle\langle 00|) \\ &+\; \frac{\Omega}{\sqrt{2}}\left[(|00\rangle\langle\Psi^+| + |11\rangle\langle\Psi^+|) + \mathrm{H.c.}\right] \\ &+\; V|\Psi^+\rangle\langle\Psi^+| - V|\Psi^-\rangle\langle\Psi^-|. \end{aligned} \quad (7.8)$$

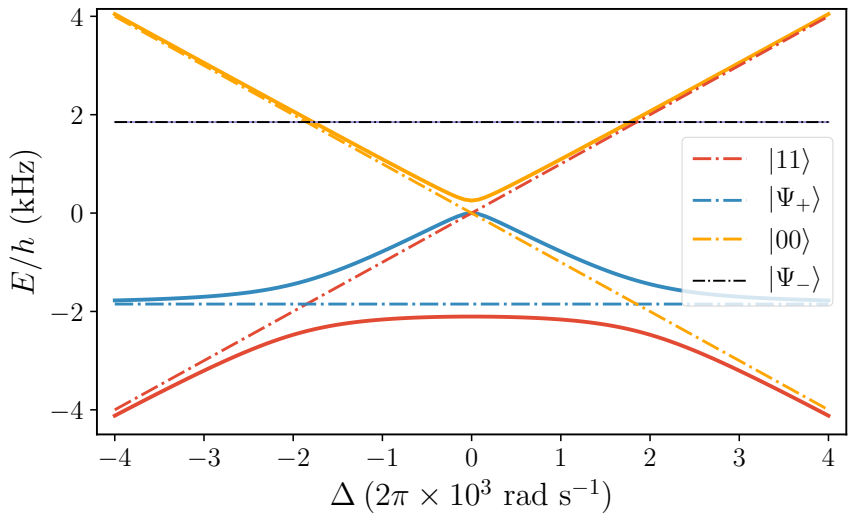


Figure 7.2: Energies of the two-molecule states. Dash-dotted lines represent the eigenenergies of $\widetilde{\mathcal{H}} - \overline{E}\mathbb{I}_4$ for $V = -1850\,\text{Hz}$ and $\Omega = 0$, while solid lines indicate the eigenenergies with the same V and $\Omega = 2\pi \times 731\,\text{rad/s}$.

The matrix form for the above Hamiltonian in the $\{|11\rangle, |\Psi^+\rangle, |00\rangle, |\Psi^-\rangle\}$ basis is

$$\widetilde{\mathcal{H}} = \overline{E}\mathbb{I}_4 + \begin{pmatrix} \Delta & \Omega/\sqrt{2} & 0 & 0 \\ \Omega/\sqrt{2} & V & \Omega/\sqrt{2} & 0 \\ 0 & \Omega/\sqrt{2} & -\Delta & 0 \\ 0 & 0 & 0 & -V \end{pmatrix}, \tag{7.9}$$

which makes it clear that the symmetric subspace spanned by $\{|11\rangle, |\Psi^+\rangle, |00\rangle\}$ (top 3×3 block matrix) does not mix with the antisymmetric subspace spanned by $|\Psi^-\rangle$. In the absence of DDI and MW coupling, the three symmetric states have equal energy when $\Delta = 0$. The presence of DDI shifts the energy of $|\Psi^+\rangle$ by V, which separates the three-level crossing into three distinct two level crossings (see Fig. 7.2). These become avoided crossings when $\Omega \neq 0$. The avoided crossing between $|00\rangle$ and $|11\rangle$ remains at $\Delta = 0$, while $|\Psi^+\rangle$ has avoided crossings at $\Delta = \pm V$ with $|00\rangle$ and $|11\rangle$ respectively.

A simple entangling gate

A non-zero DDI allows separate addressing of the transitions $|11\rangle \leftrightarrow |\Psi^+\rangle$ and $|00\rangle \leftrightarrow |\Psi^+\rangle$. The simplest way to implement an entangling gate is to consider a coherent transfer, *e.g.*, from $|11\rangle$ to $|\Psi^+\rangle$. In the triplet-singlet basis, this transfer corresponds to inverting the states $|11\rangle$ and $|\Psi^+\rangle$, which can be described by a unitary transformation such as

$$U_D = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & \mathrm{e}^{\mathrm{i}\beta} \end{pmatrix}, \tag{7.10}$$

where β is the phase acquired by the $|\Psi^-\rangle$ component, which is set by the value of V and the duration of the gate. A straightforward way to achieve a unitary transformation like this is by applying a Gaussian pulse at a constant detuning Δ. The total duration of this pulse τ_{gate} must be greater than $2\pi/|V|$ to be able to resolve the two transitions $|11\rangle \leftrightarrow |\Psi^+\rangle$ and $|00\rangle \leftrightarrow |\Psi^+\rangle$, and the length should be chosen so that the accumulated phase on the unexcited transition is a multiple of 2π. A Gaussian pulse under these conditions provides high fidelity in the absence of any errors in detuning Δ, but performs poorly when there are uncertainties in Δ; see Fig. 7.3(a). The fidelity defined as $\mathcal{F} = \mathrm{Tr}[U_D^\dagger U(\tau_{\text{gate}})]$, where $U(\tau_{\text{gate}})$ is the unitary time-evolution operator on the whole two-qubit space for the pulse, is observed to drop to ≈ 0.95 for detuning errors $\approx 100\,\mathrm{Hz}$. It is worth noting the similarity between this protocol of inverting a particular transition to implement the entangling gate and the selective line inversion experiments in NMR that resemble a CNOT like operation [85, 121–125]. Although a Gaussian pulse performs well on resonant transitions, it usually gives an overall lower gate fidelity because of the induced Bloch–Siegert shifts [199] on transitions that are off-resonance. To this end, we use a generalized driving scheme like GRAPE to design a high-fidelity gate that is robust to various experimental uncertainties as described below.

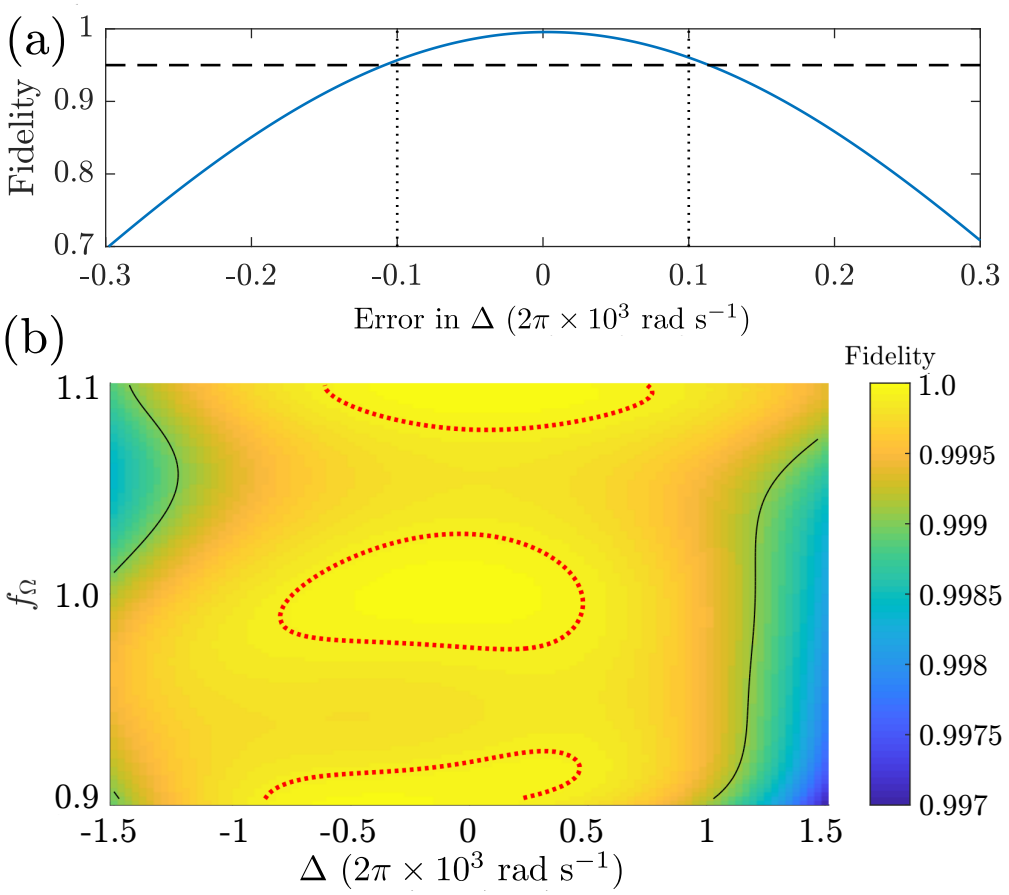


Figure 7.3: (a) Fidelity of the protocol with a Gaussian pulse (solid line) as a function of a constant error in detuning whose nominal value is taken as $\Delta = -2\pi \times 1970\,\text{rad/s}$. The horizontal dashed line marks 0.95 fidelity while vertical dotted lines mark errors of $\pm 100\,\text{Hz}$. The Gaussian pulse has maximum amplitude $2\pi \times 50 \times 10^3\,\text{rad/s}$, duration $\tau_{\text{gate}} = 0.5\,\text{ms}$ and root-mean-squared width $\tau_{\text{rms}} = 0.118\,\text{ms}$. (b) Fidelity of the protocol with the GRAPE pulse as a function of detuning and relative Rabi frequency f_Ω; see text for details. The red dotted lines indicate $\mathcal{F} = 0.9999$ while the black solid lines indicate $\mathcal{F} = 0.999$. The GRAPE pulse is of duration $\tau_{\text{gate}} = 0.5\,\text{ms}$ and has maximum amplitude $2\pi \times 50 \times 10^3\,\text{rad/s}$.

Robust entangling gate

In order to design the entangling gate in Eqn. 7.10 using GRAPE, it is not necessary to consider the full two-qubit Hilbert space for optimization. Since the symmetric and antisymmetric subspaces of the system do not mix with each other and the action of U_D is limited to the symmetric subspace only, it is sufficient to perform the optimization solely on the 3×3 symmetric subspace. If a GRAPE pulse is designed to have high fidelity in the 3×3 subspace, then its fidelity in the full two-qubit space is necessarily high, at least up to an irrelevant phase difference between the two subspaces. Thus, the target unitary operator for GRAPE optimization in the symmetric subspace is the top 3×3 block matrix of U_D in Eqn. 7.10, while the internal and control Hamiltonian respectively take the form:

$$\mathcal{H}_0 = \begin{pmatrix} \Delta & 0 & 0 \\ 0 & V & 0 \\ 0 & 0 & -\Delta \end{pmatrix}; \quad \mathcal{H}_c(t) = \Omega_x(t) I_x + \Omega_y(t) I_y, \tag{7.11}$$

where $\Omega_x(t)$ and $\Omega_y(t)$ are MW control fields along the X and Y quadratures respectively and I_x and I_y are the Pauli spin-1 operators that have the matrix form:

$$I_x = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}; \quad I_y = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & -\mathrm{i} & 0 \\ 1 & 0 & -\mathrm{i} \\ 0 & 1 & 0 \end{pmatrix}. \tag{7.12}$$

The results for BFGS-GRAPE optimization are shown in Fig. 7.3(b). The MW control fields were made robust to fluctuations within $\pm 10\%$ of the nominal value of control field $\Omega(t)$, where $\Omega(t) = \sqrt{\Omega_x^2(t) + \Omega_y^2(t)}$. This was achieved by maximizing the average fidelity for three values of Rabi frequency, $f_\Omega \times \Omega(t)$, with $f_\Omega = \{0.9, 1.0, 1.1\}$. Similarly, the pulse was also made robust to a range of detunings $\leq 1\,\mathrm{kHz}$. The fidelity reaches very high values, $\mathcal{F} > 0.9999$, for wide regions of the parameter space, and remains above the quantum-error-correction threshold, $\mathcal{F} > 0.999$, for errors $\lesssim 1\,\mathrm{kHz}$ in detuning and $\leq 10\%$ in the Rabi frequency (see Fig. 7.3). Note that the fidelities reported in the figure have been calculated

by evolving the two-molecule state within the whole 4×4 space, as described below, while the GRAPE pulse is optimized taking into account the 3×3 symmetric space.

In addition to robustness against errors in Δ and $\Omega(t)$, the pulse was also made robust to thermal fluctuations. Thermal fluctuations lead to the occupation of the excited motional states of the trap causing the internal state of the molecule to couple with the motional degrees of freedom as elaborated in Ref. [154]. The effect of motional excitation can be summarized as an energy shift in the internal Hamiltonian that depends on the particular motional state ($|gg\rangle, |ge\rangle, |eg\rangle, |ee\rangle$ where g and e denote ground and excited state respectively). Robustness in this case can be achieved by simply optimizing the pulse over an incoherent mixture of various internal Hamiltonians that represent the corresponding motional state. Recall that a similar type of averaging was performed in the previous chapter for designing GRAPE pulses that are robust to unwanted couplings with the environment qubits.

When both molecules are in the motional ground states of the trap (as assumed in the previous GRAPE optimization), it was possible to design the two-qubit gate by solely utilizing the 3×3 symmetric subspace. This was because of the symmetry of the internal Hamiltonian in Eqn. 7.7 which permitted the triplet-singlet separation. However, when the molecules are not in the ground state, the symmetry of the Hamiltonian is no longer maintained. Due to the coupling of qubit states with the motional degrees of freedom, energy shifts appear as diagonal terms in the internal Hamiltonian, thereby breaking its symmetry. As a result, the antisymmetric state can no longer completely uncouple from the symmetric subspace and therefore the system requires a treatment in the whole 4×4 Hilbert space when designing a pulse that is robust against thermal motional excitations.

The matrix forms of the four internal Hamiltonians for averaging (in $\{|11\rangle, |10\rangle, |01\rangle, |00\rangle\}$

basis) are as follows:

$$\mathcal{H}_{gg} = \begin{pmatrix} \Delta & 0 & 0 & 0 \\ 0 & 0 & V_{gg} & 0 \\ 0 & V_{gg} & 0 & 0 \\ 0 & 0 & 0 & -\Delta \end{pmatrix}, \tag{7.13}$$

$$\mathcal{H}_{ge} = \begin{pmatrix} \Delta + \delta f_{\text{trap}} & 0 & 0 & 0 \\ 0 & 0 & V_{ge} & 0 \\ 0 & V_{ge} & \delta f_{\text{trap}} & 0 \\ 0 & 0 & 0 & -\Delta \end{pmatrix}, \tag{7.14}$$

$$\mathcal{H}_{eg} = \begin{pmatrix} \Delta + \delta f_{\text{trap}} & 0 & 0 & 0 \\ 0 & \delta f_{\text{trap}} & V_{eg} & 0 \\ 0 & V_{eg} & 0 & 0 \\ 0 & 0 & 0 & -\Delta \end{pmatrix}, \tag{7.15}$$

$$\mathcal{H}_{ee} = \begin{pmatrix} \Delta + 2\delta f_{\text{trap}} & 0 & 0 & 0 \\ 0 & \delta f_{\text{trap}} & V_{ee} & 0 \\ 0 & V_{ee} & \delta f_{\text{trap}} & 0 \\ 0 & 0 & 0 & -\Delta \end{pmatrix}, \tag{7.16}$$

where $V_{gg} = -1862\,\text{Hz}$, $V_{ge} = -1877\,\text{Hz}$, $V_{eg} = -1877\,\text{Hz}$, $V_{ee} = -1892\,\text{Hz}$ were the values taken for the energy shifts in the DDI corresponding to the motional excitations while $\delta f_{\text{trap}} = 500\,\text{Hz}$ was chosen as the difference in trap frequency for $|0\rangle$ and $|1\rangle$ states for both the traps. Calculations for GRAPE pulse optimization were performed in the two-qubit computational basis whereas the fidelity at the end of every GRAPE iteration was evaluated by transforming the net pulse propagator into the triplet-singlet basis. The matrix for transforming a unitary

in $\{|11\rangle, |10\rangle, |01\rangle, |00\rangle\}$ basis to $\{|11\rangle, |\Psi^+\rangle, |00\rangle, |\Psi^-\rangle\}$ basis is

$$\mathrm{H} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{7.17}$$

Note that the ordering $\{|11\rangle, |\Psi^+\rangle, |\Psi^-\rangle, |00\rangle\}$ of basis states in the above transformation matrix is different from the order $\{|11\rangle, |\Psi^+\rangle, |00\rangle, |\Psi^-\rangle\}$ we have considered so far. Accordingly, the target unitary in this modified basis ordering can be obtained by reshuffling the rows of matrix U_D in Eqn. 7.10:

$$U_D^{'} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & \mathrm{e}^{\mathrm{i}\beta} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{7.18}$$

For the GRAPE optimization, however, the $(3,3)$ entry of $U_D^{'}$ was ignored and a modified target operator was defined as

$$O_D = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{7.19}$$

Optimizing for this modified operator corresponds to ignoring the antisymmetric subspace spanned by $|\Psi^-\rangle$. Essentially, the pulse was optimized only for the symmetric subspace because, as mentioned previously, if the pulse has a high-fidelity with respect to the symmetric subspace then we can be certain that the fidelity in the whole two-qubit space is necessarily high and there is no 'leakage' onto the antisymmetric subspace. In this manner, the GRAPE pulse was made robust against thermal motional excitations.

7.2 Superconducting qubits

In this section, I outline the ongoing work on the use of GRAPE pulses for implementing quantum gates in superconducting qubits. The quantum system discussed here can be described as an assembly of macroscopic superconducting electrical circuit elements that produce the energy spectrum of a quantized weakly-anharmonic oscillator [388]. Typically, the lowest two energy levels of this oscillator are used to define a qubit, which has a resonance frequency in the MW regime. One of the main challenges when implementing any gate in such a system is preventing leakage out of the qubit subspace. To this end, it is beneficial to design control pulses that are optimized to avoid leakage [406]. Usually, a third energy level is considered in the pulse optimization to account for the leakage, provided that its population is minimized during the entire duration of the pulse. Along with this, it is also necessary to implement the gate in the shortest possible duration as the coherence times of superconducting qubits are extremely short owing to their macroscopic character. Below, I illustrate how GRAPE can be used for designing time-optimal pulses that aim to prevent leakage. I begin by describing the system Hamiltonian.

Single-qubit Hamiltonian

Consider a slightly anharmonic three-level energy spectrum with only nearest level coupling. The internal Hamiltonian for such a system in the lab frame and in natural units ($\hbar = 1$) can be written as:

$$\mathcal{H}_0 = \sum_j \omega_j \left|j\right\rangle\!\left\langle j\right| = \omega_0 \left|0\right\rangle\!\left\langle 0\right| + \omega_1 \left|1\right\rangle\!\left\langle 1\right| + \omega_2 \left|2\right\rangle\!\left\langle 2\right| \tag{7.20}$$

where $|j\rangle\langle j|$ is the projection operator for the j^{th} energy level with ω_j being the corresponding energy and

$$|0\rangle = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}; \quad |1\rangle = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}; \quad |2\rangle = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}. \tag{7.21}$$

The ground state energy ω_0 is usually set to zero. The anharmonicity of the system is $\delta = \omega_2 - 2\omega_1$. The resonance frequency of the qubit is typically around 5 GHz while the anharmonicity is of the order of a few hundred MHz. Under the influence of a near resonant MW driving field, the Hamiltonian of the system is

$$\mathcal{H} = \sum_j \omega_j |j\rangle\langle j| + \Omega(t)(a + a^\dagger), \tag{7.22}$$

where $\Omega(t) = \Omega_x(t)\cos(\omega_d t) + \Omega_y(t)\sin(\omega_d t)$ with ω_d being the MW drive frequency and a, $a^\dagger$ are the annihilation and creation operators respectively. These ladder operators for the three-level system can be expressed as:

$$a = |0\rangle\langle 1| + \lambda |1\rangle\langle 2| = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & \lambda \\ 0 & 0 & 0 \end{pmatrix}; \quad a^\dagger = |1\rangle\langle 0| + \lambda |2\rangle\langle 1| = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & \lambda & 0 \end{pmatrix}, \tag{7.23}$$

where λ is the relative strength of $|0\rangle \leftrightarrow |1\rangle$ transition to the $|1\rangle \leftrightarrow |2\rangle$. For a transmon qubit [407–411] considered in this work, which is only a weakly-anharmonic system, $\lambda = \sqrt{2}$, as in a regular quantum harmonic oscillator. The operator $(a + a^\dagger)$ can be interpreted as a σ_x operator on the $\{|0\rangle, |1\rangle\}$ and $\{|1\rangle, |2\rangle\}$ subspaces as

$$a + a^\dagger = (|0\rangle\langle 1| + |1\rangle\langle 0|) + \sqrt{2}(|1\rangle\langle 2| + |2\rangle\langle 1|) = \sigma_x^{0,1} + \sqrt{2}\sigma_x^{1,2} := \sigma_x. \tag{7.24}$$

Transforming the lab frame Hamiltonian in Eqn. 7.22 into the rotating frame (at drive frequency ω_d) and making the rotating wave approximation, we get

$$\widetilde{\mathcal{H}} = \sum_j \Delta_j \left|j\right\rangle\!\left\langle j\right| + \frac{\Omega_x(t)}{2}\sigma_x + \frac{\Omega_y(t)}{2}\sigma_y, \tag{7.25}$$

where $\Delta_1 = \omega_1 - \omega_d$ and $\Delta_2 = \omega_2 - \omega_d = \delta + 2\Delta_1$ are the detunings and $\sigma_y \coloneqq \sigma_y^{0,1} + \sqrt{2}\sigma_y^{1,2} = \mathrm{i}(a^\dagger - a)$ analogous to Eqn. 7.24. When $\omega_d = \omega_1$, *i.e.*, the drive frequency is resonant with the qubit frequency, $\Delta_1 = 0$ and $\Delta_2 = \delta$ is the anharmonicity. The matrix form for the rotating frame Hamiltonian with a resonant drive is

$$\widetilde{\mathcal{H}} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \delta \end{pmatrix} + \frac{\Omega_x(t)}{2} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & \sqrt{2} \\ 0 & \sqrt{2} & 0 \end{pmatrix} + \frac{\Omega_y(t)}{2} \begin{pmatrix} 0 & -\mathrm{i} & 0 \\ \mathrm{i} & 0 & -\mathrm{i}\sqrt{2} \\ 0 & \mathrm{i}\sqrt{2} & 0 \end{pmatrix}. \tag{7.26}$$

Single-qubit gates can be implemented by controlling $\Omega_x(t)$ and $\Omega_y(t)$. An important consideration here when designing a GRAPE pulse is the definition of the fidelity function and the form of the target unitary operator used for optimization. The GRAPE fidelity function was defined as [181]

$$\mathcal{F} = \frac{1}{d^2}\left|\langle U_D|PU(T)P\rangle\right|^2 = \frac{1}{d^2}\left[\left|\left\langle 0|U_D^\dagger U(T)|0\right\rangle + \left\langle 1|U_D^\dagger U(T)|1\right\rangle\right|^2\right], \tag{7.27}$$

where U_D is the target unitary, $d = 2$ is the dimension of the qubit subspace, $U(T)$ is the GRAPE propagator and $P = |0\rangle\langle 0| + |1\rangle\langle 1|$ is a projector onto the qubit subspace. The target unitary for any single qubit rotation is given by $\exp\left[-\mathrm{i}\theta\sigma_x^{0,1}/2\right]$. The formulation of fidelity function in Eqn. 7.27 ignores the subspace spanned by the $|2\rangle$ energy level and by the same argument in the previous section on polar molecules, the fidelity in the qubit subspace spanned by $\{|0\rangle, |1\rangle\}$ can be maximized to achieve an overall high fidelity for the GRAPE pulse in the three level system. Although maximizing for such a fidelity function ensures that the entire population remains in the qubit subspace at the end of the GRAPE pulse, it is highly likely that the level $|2\rangle$ becomes populated during intermediate times as the GRAPE

pulse progresses. At any point in time, it is undesirable to have a significant population in level $|2\rangle$ because once level $|2\rangle$ is occupied then there is a chance of leakage onto higher energy levels ($|3\rangle, |4\rangle, \dots$) of the oscillator. As a result, the fidelity function in Eqn. 7.27 alone is insufficient to guarantee that there is no leakage outside the qubit subspace.

One way to avoid leakage is by including more energy levels in the pulse optimization and restricting the occupancy of higher energy levels. However, this approach is computationally infeasible, especially when multiple qubits are involved, as the dimension of Hilbert space increases rapidly. Alternatively, a penalty function can be defined which tracks the population in level $|2\rangle$ at every time and imposes a penalty when level $|2\rangle$ becomes populated. Thus, by modifying the above fidelity function to include such a penalty, the population in level $|2\rangle$ can be minimized not only at the end of the pulse but also during the entire duration of the pulse. In addition, we found that using a smoothing penalty (mentioned in Chapter 2) to obtain smoothly varying control fields is useful to prevent the populations in level $|2\rangle$ at any point in time; as smooth controls do not contribute to high frequency components. The resulting GRAPE pulse is very similar to DRAG (Derivative Removal by Adiabatic Gate) pulses [406] which are popular in superconducting qubit implementations for preventing leakage.

Two-qubit Hamiltonian

Superconducting qubits are like macroscopic artificial atoms that can be coupled through MW resonators using techniques of circuit quantum electrodynamics (cQED) [62]. Coupling between two superconducting qubits is usually generated dynamically, unlike the J-couplings in NMR spins which are inherently present. However, the particular device used by the Leek lab for this work was tuned to generate always-on NMR-like couplings, as mentioned earlier in Chapter 4. The internal Hamiltonian for such a two-qubit system in a doubly rotating frame with a resonant drive for each qubit is given by (note that each qubit can be addressed

using a separate drive, like in heteronuclear NMR spin systems):

$$\widetilde{\mathcal{H}}_0 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \delta_{\mathrm{A}} \end{pmatrix} \otimes \mathbb{I} + \mathbb{I} \otimes \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \delta_{\mathrm{B}} \end{pmatrix} + J \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \otimes \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \tag{7.28}$$

where $\delta_{\mathrm{A/B}}$ is the anharmonicity for qubit A/B and J is the coupling strength ($\sim$ few MHz) between the two (transmon) qubits. Usually the coupling between two transmons is transverse (*e.g.*, $\sigma_x \otimes \sigma_x + \sigma_y \otimes \sigma_y$) [412–414], but the system can also be engineered to generate a longitudinal coupling of the form $\sigma_z \otimes \sigma_z$ as in the above Hamiltonian [328, 415]. The qubits can be controlled through two independent MW drives, which can be expressed by the rotating frame control Hamiltonian

$$\widetilde{\mathcal{H}}_c(t) = \left[\frac{\Omega_x^{\mathrm{A}}(t)}{2} \sigma_x \otimes \mathbb{I} + \frac{\Omega_y^{\mathrm{A}}(t)}{2} \sigma_y \otimes \mathbb{I} \right] + \left[\frac{\Omega_x^{\mathrm{B}}(t)}{2} \mathbb{I} \otimes \sigma_x + \frac{\Omega_y^{\mathrm{B}}(t)}{2} \mathbb{I} \otimes \sigma_y \right]. \tag{7.29}$$

For GRAPE optimization, the fidelity calculation can be performed in the same way as described for the single-qubit case, except that the operator P for projecting onto the qubit subspace now becomes $P = |00\rangle\langle 00| + |01\rangle\langle 01| + |10\rangle\langle 10| + |11\rangle\langle 11|$.

The first preliminary goal of this project was to experimentally test GRAPE pulses for single-qubit gates on a two-qubit device. Although we managed to design high-fidelity single-qubit gates using GRAPE, the pulses did not, at the time this thesis was written, perform as expected in experiments, and this remains a matter of ongoing investigation.

Chapter 8

Conclusions

This thesis presented some practical strategies for coherent quantum control and corresponding example demonstrations for implementing quantum logic gates in experimental platforms such as NMR, polar molecules and superconducting qubits.

First, in Chapter 2, I described Gradient Ascent Pulse Engineering (GRAPE), a powerful optimal control algorithm for designing shaped EM pulses to achieve a wide variety of control related tasks. The popularity of GRAPE stems from the ease with which it can be adapted in many experimental platforms as well as the convenience and flexibility it offers in designing highly robust pulses by simultaneously handling several errors. The downside of GRAPE, however, is that it is computationally expensive and its complexity grows exponentially with the number of qubits in the system. Like any other open-loop optimal control algorithm, GRAPE requires several full-fledged quantum simulations of the system, for which it is necessary to perform the computationally intensive operations of matrix exponentiation and matrix multiplication. Chapter 3 described strategies to avoid the explicit evaluation of matrix exponentials, thereby speeding up the GRAPE algorithm by over an order of magnitude and giving constant gains, independent of the size of the system. These modest speed gains offer a more pragmatic approach to pulse engineering, even though they fail to address the fundamental issue of scalability of the GRAPE algorithm.

Beyond GRAPE, this thesis also presented methods for designing spin-echo sequences that

utilize inherent, non-switchable zz-couplings between qubits (as found in NMR and certain realizations of superconducting qubits) for implementing any arbitrary network of CZ gates. The method described in Chapter 4, based on linear programming, finds time-optimal pulse sequences for fully-coupled systems having up to few tens of qubits and near time-optimal sequences in systems with a few hundred qubits. Further, the analytic method based on graph colouring, described in Chapter 5, gave near time-optimal sequences for engineered systems with any number of qubits having only nearest and next-nearest neighbour couplings.

Some natural questions one might ask here concern how the above methods compare and when to use which method. To answer these questions it is first necessary to understand exactly at what level of the quantum computing stack these methods apply. Roughly speaking, the realization of any quantum computation—from defining the computational task to its actual implementation on the physical hardware—can be described in terms of four levels of abstraction or 'stack'. The topmost layer of the stack comprises a very high level description of the problem and its corresponding quantum algorithm in terms of a quantum circuit that maps the qubits and gates. In the next layer, the quantum circuit is expressed in terms of native one- and two-qubit gates, taking into account the experimental platform, available hardware and qubit connectivity. At this stage the expressed gate network is often compressed and compiled and error correction schemes are also introduced. The third layer involves control engineering, wherein the native gates are translated into EM field pulses that implement the corresponding gates on respective qubits, while the final stage is related to the interfacing and implementation of these control fields on the actual physical platform.

Quite evidently, the GRAPE algorithm accomplishes the task at the third level of the stack. On the other hand, it is less obvious on what exact level of the stack the spin-echo based methods fit in. These methods describe how to implement a given compiled network of native two-qubit gates (here, CZ gates) using spin-echoes in an efficient manner; however, they do not provide any information on how to implement the echo pulses (NOT gates). As a result, the spin-echo based method can be thought of as being at the interface of the second and third level of the stack.

Since the two approaches do not belong to the same level of the stack, it is not possible to perform a one-to-one comparison between them. Moreover, GRAPE is applicable for designing *any* given quantum gate, whereas the spin-echo methods are relevant only for producing two-qubit gates, and only in specific systems that have zz-couplings between qubits. In fact, when using GRAPE, one can directly implement any quantum circuit with arbitrary gates without the need for decomposing the circuit into native gates; essentially omitting the operations on the second level of the stack. However, the computational complexity of the GRAPE algorithm restricts its utility to only a handful of qubits. Beyond a few tens of qubits, the linear-programming based method is convenient for implementing native two-qubit gates, which when combined with single-qubit rotations gives a universal set. Such NISQ type systems usually have limited qubit connectivity, in which case, the sub-system GRAPE methods described in Chapter 2 are useful for designing the echo pulses as well as other single-qubit rotations. In systems which are engineered to have only local qubit connectivity, the obvious choice is to use the graph colouring approach for implementing two-qubit gates. Also in such engineered systems, each qubit can be addressed individually which makes it easier to implement single-qubit gates. In summary, the methods described in this thesis have specific regimes of operation and thus need to be used accordingly, and in conjunction with each other.

Bibliography

[1] J. P. Dowling and G. J. Milburn, *Quantum technology: the second quantum revolution*, Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences **361**, 1655 (2003).

[2] G. Milburn, *Schrödinger's Machines: The Quantum Technology Reshaping Everyday Life* (W.H. Freeman Company, 1997).

[3] D. Deutsch, *Quantum theory, the Church-Turing principle and the universal quantum computer*, Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences **400**, 97 (1985).

[4] C. H. Bennett and D. P. DiVincenzo, *Quantum information and computation*, Nature **404**, 247 (2000).

[5] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition* (Cambridge University Press, 2010).

[6] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, *Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels*, Phys. Rev. Lett. **70**, 1895 (1993).

[7] N. Gisin and R. Thew, *Quantum communication*, Nature Photonics **1**, 165 (2007).

[8] S. Wiesner, *Conjugate Coding*, SIGACT News **15**, 7888 (1983).

[9] C. H. Bennett and G. Brassard, *Quantum cryptography: Public key distribution and coin tossing*, in *Proceedings of IEEE International Conference on Computers, Systems, and Signal Processing* (Bangalore, India, 1984) pp. 175–179.

[10] A. K. Ekert, *Quantum cryptography based on Bell's theorem*, Phys. Rev. Lett. **67**, 661 (1991).

[11] C. H. Bennett, F. Bessette, G. Brassard, L. Salvail, and J. Smolin, *Experimental quantum cryptography*, Journal of Cryptology **5**, 3 (1992).

[12] N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden, *Quantum cryptography*, Rev. Mod. Phys. **74**, 145 (2002).

[13] S. Lloyd, *Universal Quantum Simulators*, Science **273**, 1073 (1996).

[14] I. Buluta and F. Nori, *Quantum Simulators*, Science **326**, 108 (2009).

[15] I. M. Georgescu, S. Ashhab, and F. Nori, *Quantum simulation*, Rev. Mod. Phys. **86**, 153 (2014).

[16] T. B. Pittman, Y. H. Shih, D. V. Strekalov, and A. V. Sergienko, *Optical imaging by means of two-photon quantum entanglement*, Phys. Rev. A **52**, R3429 (1995).

[17] L. A. Lugiato, A. Gatti, and E. Brambilla, *Quantum imaging*, Journal of Optics B: Quantum and Semiclassical Optics **4**, S176 (2002).

[18] P.-A. Moreau, E. Toninelli, T. Gregory, and M. J. Padgett, *Imaging with quantum states of light*, Nature Reviews Physics **1**, 367 (2019).

[19] C. L. Degen, F. Reinhard, and P. Cappellaro, *Quantum sensing*, Rev. Mod. Phys. **89**, 035002 (2017).

[20] S. Pirandola, B. R. Bardhan, T. Gehring, C. Weedbrook, and S. Lloyd, *Advances in photonic quantum sensing*, Nature Photonics **12**, 724 (2018).

[21] V. Giovannetti, S. Lloyd, and L. Maccone, *Quantum-Enhanced Measurements: Beating the Standard Quantum Limit*, Science **306**, 1330 (2004).

[22] V. Giovannetti, S. Lloyd, and L. Maccone, *Quantum Metrology*, Phys. Rev. Lett. **96**, 010401 (2006).

[23] V. Giovannetti, S. Lloyd, and L. Maccone, *Advances in quantum metrology*, Nature Photonics **5**, 222 (2011).

[24] H. M. Wiseman and G. J. Milburn, *Quantum Measurement and Control* (Cambridge University Press, 2009).

[25] D. Dong and I. R. Petersen, *Quantum control theory and applications: a survey*, IET Control Theory Applications **4**, 2651 (2010).

[26] S. J. Glaser, U. Boscain, T. Calarco, C. P. Koch, W. Köckenberger, R. Kosloff, I. Kuprov, B. Luy, S. Schirmer, T. Schulte-Herbrüggen, D. Sugny, and F. K. Wilhelm, *Training Schrödinger's cat: quantum optimal control*, The European Physical Journal D **69**, 279 (2015).

[27] A. Acín, I. Bloch, H. Buhrman, T. Calarco, C. Eichler, J. Eisert, D. Esteve, N. Gisin, S. J. Glaser, F. Jelezko, S. Kuhr, M. Lewenstein, M. F. Riedel, P. O. Schmidt, R. Thew, A. Wallraff, I. Walmsley, and F. K. Wilhelm, *The quantum technologies roadmap: a European community view*, New Journal of Physics **20**, 080201 (2018).

[28] J. Preskill, *Quantum Computing in the NISQ era and beyond*, Quantum **2**, 79 (2018).

[29] P. Benioff, *The computer as a physical system: A microscopic quantum mechanical Hamiltonian model of computers as represented by Turing machines*, Journal of Statistical Physics **22**, 563 (1980).

[30] R. P. Feynman, *Simulating physics with computers*, International Journal of Theoretical Physics **21**, 467 (1982).

[31] B. Schumacher, *Quantum coding*, Phys. Rev. A **51**, 2738 (1995).

[32] S. Lloyd, *A Potentially Realizable Quantum Computer*, Science **261**, 1569 (1993).

[33] D. E. Deutsch, *Quantum computational networks*, Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences **425**, 73 (1989).

[34] D. E. Deutsch, A. Barenco, and A. Ekert, *Universality in quantum computation*, Proceedings of the Royal Society of London. Series A: Mathematical and Physical Sciences **449**, 669 (1995).

[35] 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).

[36] D. P. DiVincenzo, *Quantum gates and circuits*, Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences **454**, 261 (1998).

[37] A. Y. Kitaev, *Quantum computations: algorithms and error correction*, Russian Mathematical Surveys **52**, 1191 (1997).

[38] C. M. Dawson and M. A. Nielsen, *The Solovay-Kitaev Algorithm*, Quantum Info. Comput. **6**, 8195 (2006).

[39] D. P. DiVincenzo, *The Physical Implementation of Quantum Computation*, Fortschritte der Physik **48**, 771 (2000).

[40] S. A. Rice, M. Zhao, *et al.*, *Optical Control of Molecular Dynamics* (Wiley, New York, 2000).

[41] M. Shapiro and P. Brumer, *Principles of the Quantum Control of Molecular Processes* (2003).

[42] C. Brif, R. Chakrabarti, and H. Rabitz, *Control of quantum phenomena: past, present and future*, New Journal of Physics **12**, 075008 (2010).

[43] J. P. Palao and R. Kosloff, *Quantum Computing by an Optimal Control Algorithm for Unitary Transformations*, Phys. Rev. Lett. **89**, 188301 (2002).

[44] J. P. Palao and R. Kosloff, *Optimal control theory for unitary transformations*, Phys. Rev. A **68**, 062308 (2003).

[45] T. D. Ladd, F. Jelezko, R. Laflamme, Y. Nakamura, C. Monroe, and J. L. O'Brien, *Quantum computers*, Nature **464**, 45 (2010).

[46] I. Buluta, S. Ashhab, and F. Nori, *Natural and artificial atoms for quantum computation*, Reports on Progress in Physics **74**, 104401 (2011).

[47] J. I. Cirac and P. Zoller, *Quantum Computations with Cold Trapped Ions*, Phys. Rev. Lett. **74**, 4091 (1995).

[48] A. Steane, *The ion trap quantum information processor*, Applied Physics B **64**, 623 (1997).

[49] A. Sørensen and K. Mølmer, *Quantum Computation with Ions in Thermal Motion*, Phys. Rev. Lett. **82**, 1971 (1999).

[50] D. Kielpinski, C. Monroe, and D. J. Wineland, *Architecture for a large-scale ion-trap quantum computer*, Nature **417**, 709 (2002).

[51] H. Häffner, C. Roos, and R. Blatt, *Quantum computing with trapped ions*, Physics Reports **469**, 155 (2008).

[52] D. G. Cory, A. F. Fahmy, and T. F. Havel, *Ensemble quantum computing by NMR spectroscopy*, Proceedings of the National Academy of Sciences **94**, 1634 (1997).

[53] W. S. Warren, *The Usefulness of NMR Quantum Computing*, Science **277**, 1688 (1997).

[54] N. A. Gershenfeld and I. L. Chuang, *Bulk Spin-Resonance Quantum Computation*, Science **275**, 350 (1997).

[55] D. G. Cory, M. D. Price, and T. F. Havel, *Nuclear magnetic resonance spectroscopy: An experimentally accessible paradigm for quantum computing*, Physica D: Nonlinear Phenomena **120**, 82 (1998).

[56] I. L. Chuang, N. Gershenfeld, M. G. Kubinec, and D. W. Leung, *Bulk quantum computation with nuclear magnetic resonance: theory and experiment*, Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences **454**, 447 (1998).

[57] Y. Nakamura, Y. A. Pashkin, and J. S. Tsai, *Coherent control of macroscopic quantum states in a single-Cooper-pair box*, Nature **398**, 786 (1999).

[58] Y. Makhlin, G. Scöhn, and A. Shnirman, *Josephson-junction qubits with controlled couplings*, Nature **398**, 305 (1999).

[59] J. M. Raimond, M. Brune, and S. Haroche, *Manipulating quantum entanglement with atoms and photons in a cavity*, Rev. Mod. Phys. **73**, 565 (2001).

[60] 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).

[61] I. Chiorescu, Y. Nakamura, C. J. P. M. Harmans, and J. E. Mooij, *Coherent Quantum Dynamics of a Superconducting Flux Qubit*, Science **299**, 1869 (2003).

[62] A. Blais, R.-S. Huang, A. Wallraff, S. M. Girvin, and R. J. Schoelkopf, *Cavity quantum electrodynamics for superconducting electrical circuits: An architecture for quantum computation*, Phys. Rev. A **69**, 062320 (2004).

[63] A. Wallraff, D. I. Schuster, A. Blais, L. Frunzio, R.-. S. Huang, J. Majer, S. Kumar, S. M. Girvin, and R. J. Schoelkopf, *Strong coupling of a single photon to a superconducting qubit using circuit quantum electrodynamics*, Nature **431**, 162 (2004).

[64] J. Clarke and F. K. Wilhelm, *Superconducting quantum bits*, Nature **453**, 1031 (2008).

[65] E. Knill, R. Laflamme, and G. J. Milburn, *A scheme for efficient quantum computation with linear optics*, Nature **409**, 46 (2001).

[66] P. Kok, W. J. Munro, K. Nemoto, T. C. Ralph, J. P. Dowling, and G. J. Milburn, *Linear optical quantum computing with photonic qubits*, Rev. Mod. Phys. **79**, 135 (2007).

[67] J. L. O'Brien, *Optical Quantum Computing*, Science **318**, 1567 (2007).

[68] D. Loss and D. P. DiVincenzo, *Quantum computation with quantum dots*, Phys. Rev. A **57**, 120 (1998).

[69] B. E. Kane, *A silicon-based nuclear spin quantum computer*, Nature **393**, 133 (1998).

[70] A. Imamoglu, D. D. Awschalom, G. Burkard, D. P. DiVincenzo, D. Loss, M. Sherwin, and A. Small, *Quantum Information Processing Using Quantum Dot Spins and Cavity QED*, Phys. Rev. Lett. **83**, 4204 (1999).

[71] R. Vrijen, E. Yablonovitch, K. Wang, H. W. Jiang, A. Balandin, V. Roychowdhury, T. Mor, and D. DiVincenzo, *Electron-spin-resonance transistors for quantum computing in silicon-germanium heterostructures*, Phys. Rev. A **62**, 012306 (2000).

[72] J. R. Petta, A. C. Johnson, J. M. Taylor, E. A. Laird, A. Yacoby, M. D. Lukin, C. M. Marcus, M. P. Hanson, and A. C. Gossard, *Coherent Manipulation of Coupled Electron Spins in Semiconductor Quantum Dots*, Science **309**, 2180 (2005).

[73] A. P. Nizovtsev, S. Y. Kilin, F. Jelezko, T. Gaebal, I. Popa, A. Gruber, and J. Wrachtrup, *A quantum computer based on NV centers in diamond: Optically detected nutations of single electron and nuclear spins*, Optics and Spectroscopy **99**, 233 (2005).

[74] J. Wrachtrup and F. Jelezko, *Processing quantum information in diamond*, Journal of Physics: Condensed Matter **18**, S807 (2006).

[75] R. Hanson, L. P. Kouwenhoven, J. R. Petta, S. Tarucha, and L. M. K. Vandersypen, *Spins in few-electron quantum dots*, Rev. Mod. Phys. **79**, 1217 (2007).

[76] M. V. G. Dutt, L. Childress, L. Jiang, E. Togan, J. Maze, F. Jelezko, A. S. Zibrov, P. R. Hemmer, and M. D. Lukin, *Quantum Register Based on Individual Electronic and Nuclear Spin Qubits in Diamond*, Science **316**, 1312 (2007).

[77] R. Hanson and D. D. Awschalom, *Coherent manipulation of single spins in semiconductors*, Nature **453**, 1043 (2008).

[78] J. R. Weber, W. F. Koehl, J. B. Varley, A. Janotti, B. B. Buckley, C. G. Van de Walle, and D. D. Awschalom, *Quantum computing with defects*, Proceedings of the National Academy of Sciences **107**, 8513 (2010).

[79] L. Childress and R. Hanson, *Diamond NV centers for quantum computing and quantum networks*, MRS Bulletin **38**, 134138 (2013).

[80] C. Monroe, D. M. Meekhof, B. E. King, W. M. Itano, and D. J. Wineland, *Demonstration of a Fundamental Quantum Logic Gate*, Phys. Rev. Lett. **75**, 4714 (1995).

[81] D. Deutsch and R. Jozsa, *Rapid solution of problems by quantum computation*, Proceedings of the Royal Society of London. Series A: Mathematical and Physical Sciences **439**, 553 (1992).

[82] S. Gulde, M. Riebe, G. P. T. Lancaster, C. Becher, J. Eschner, H. Häffner, F. Schmidt-Kaler, I. L. Chuang, and R. Blatt, *Implementation of the Deutsch–Jozsa algorithm on an ion-trap quantum computer*, Nature **421**, 48 (2003).

[83] J. Jones, R. Hansen, and M. Mosca, *Quantum Logic Gates and Nuclear Magnetic Resonance Pulse Sequences*, Journal of Magnetic Resonance **135**, 353 (1998).

[84] I. L. Chuang, L. M. K. Vandersypen, X. Zhou, D. W. Leung, and S. Lloyd, *Experimental realization of a quantum algorithm*, Nature **393**, 143 (1998).

[85] N. Linden, H. Barjat, and R. Freeman, *An implementation of the DeutschJozsa algorithm on a three-qubit NMR quantum computer*, Chemical Physics Letters **296**, 61 (1998).

[86] R. Marx, A. F. Fahmy, J. M. Myers, W. Bermel, and S. J. Glaser, *Approaching five-bit NMR quantum computing*, Phys. Rev. A **62**, 012310 (2000).

[87] J. A. Jones, M. Mosca, and R. H. Hansen, *Implementation of a quantum search algorithm on a quantum computer*, Nature **393**, 344 (1998).

[88] I. L. Chuang, N. Gershenfeld, and M. Kubinec, *Experimental Implementation of Fast Quantum Searching*, Phys. Rev. Lett. **80**, 3408 (1998).

[89] L. M. K. Vandersypen, M. Steffen, M. H. Sherwood, C. S. Yannoni, G. Breyta, and I. L. Chuang, *Implementation of a three-quantum-bit search algorithm*, Applied Physics Letters **76**, 646 (2000).

[90] L. M. K. Vandersypen, M. Steffen, G. Breyta, C. S. Yannoni, M. H. Sherwood, and I. L. Chuang, *Experimental realization of Shor's quantum factoring algorithm using nuclear magnetic resonance*, Nature **414**, 883 (2001).

[91] D. G. Cory, M. D. Price, W. Maas, E. Knill, R. Laflamme, W. H. Zurek, T. F. Havel, and S. S. Somaroo, *Experimental Quantum Error Correction*, Phys. Rev. Lett. **81**, 2152 (1998).

[92] E. Knill, R. Laflamme, R. Martinez, and C. Negrevergne, *Benchmarking Quantum Computers: The Five-Qubit Error Correcting Code*, Phys. Rev. Lett. **86**, 5811 (2001).

[93] A. Ekert, R. Jozsa, R. Penrose, R. Laflamme, E. Knill, W. H. Zurek, P. Catasti, and S. V. S. Mariappan, *NMR Greenberger Horne Zeilinger states*, Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences **356**, 1941 (1998).

[94] M. A. Nielsen, E. Knill, and R. Laflamme, *Complete quantum teleportation using nuclear magnetic resonance*, Nature **396**, 52 (1998).

[95] D. Leung, L. Vandersypen, X. Zhou, M. Sherwood, C. Yannoni, M. Kubinec, and I. Chuang, *Experimental realization of a two-bit phase damping quantum code*, Phys. Rev. A **60**, 1924 (1999).

[96] S. Somaroo, C. H. Tseng, T. F. Havel, R. Laflamme, and D. G. Cory, *Quantum Simulations on a Quantum Computer*, Phys. Rev. Lett. **82**, 5381 (1999).

[97] J. A. Jones and M. Mosca, *Approximate Quantum Counting on an NMR Ensemble Quantum Computer*, Phys. Rev. Lett. **83**, 1050 (1999).

[98] J. A. Jones, V. Vedral, A. Ekert, and G. Castagnoli, *Geometric quantum computation using nuclear magnetic resonance*, Nature **403**, 869 (2000).

[99] E. Knill, R. Laflamme, R. Martinez, and C.-H. Tseng, *An algorithmic benchmark for quantum information processing*, Nature **404**, 368 (2000).

[100] L. M. K. Vandersypen, M. Steffen, G. Breyta, C. S. Yannoni, R. Cleve, and I. L. Chuang, *Experimental Realization of an Order-Finding Algorithm with an NMR Quantum Computer*, Phys. Rev. Lett. **85**, 5452 (2000).

[101] A. M. Childs, I. L. Chuang, and D. W. Leung, *Realization of quantum process tomography in NMR*, Phys. Rev. A **64**, 012314 (2001).

[102] L. Viola, E. M. Fortunato, M. A. Pravia, E. Knill, R. Laflamme, and D. G. Cory, *Experimental Realization of Noiseless Subsystems for Quantum Information Processing*, Science **293**, 2059 (2001).

[103] F. Bloch, W. W. Hansen, and M. Packard, *Nuclear Induction*, Phys. Rev. **69**, 127 (1946).

[104] E. M. Purcell, H. C. Torrey, and R. V. Pound, *Resonance Absorption by Nuclear Magnetic Moments in a Solid*, Phys. Rev. **69**, 37 (1946).

[105] W. P. Aue, E. Bartholdi, and R. R. Ernst, *Twodimensional spectroscopy. Application to nuclear magnetic resonance*, The Journal of Chemical Physics **64**, 2229 (1976).

[106] L. E. Kay, M. Ikura, R. Tschudin, and A. Bax, *Three-dimensional triple-resonance NMR spectroscopy of isotopically enriched proteins*, Journal of Magnetic Resonance (1969) **89**, 496 (1990).

[107] M. Ikura, L. E. Kay, and A. Bax, *A novel approach for sequential assignment of proton, carbon-13, and nitrogen-15 spectra of larger proteins: heteronuclear triple-resonance three-dimensional NMR spectroscopy. Application to calmodulin*, Biochemistry **29**, 4659 (1990).

[108] J. Cavanagh, W. J. Fairbrother, A. G. Palmer III, and N. J. Skelton, *Protein NMR Spectroscopy: Principles and Practice* (Academic Press: San Diego, 1996).

[109] L. Emsley and G. Bodenhausen, *Gaussian pulse cascades: New analytical functions for rectangular selective inversion and in-phase excitation in NMR*, Chemical Physics Letters **165**, 469 (1990).

[110] S. McDonald and W. S. Warren, *Uses of shaped pulses in NMR: A primer*, Concepts in Magnetic Resonance **3**, 55 (1991).

[111] H. Geen and R. Freeman, *Band-selective radiofrequency pulses*, Journal of Magnetic Resonance (1969) **93**, 93 (1991).

[112] R. Freeman, *Shaped radiofrequency pulses in high resolution NMR*, Progress in Nuclear Magnetic Resonance Spectroscopy **32**, 59 (1998).

[113] M. H. Levitt, *Symmetrical composite pulse sequences for NMR population inversion. I. Compensation of radiofrequency field inhomogeneity*, Journal of Magnetic Resonance (1969) **48**, 234 (1982).

[114] M. H. Levitt, *Composite pulses*, Progress in Nuclear Magnetic Resonance Spectroscopy **18**, 61 (1986).

[115] H. K. Cummins and J. A. Jones, *Use of composite rotations to correct systematic errors in NMR quantum computation*, New Journal of Physics **2**, 6 (2000).

[116] E. L. Hahn, *Spin Echoes*, Phys. Rev. **80**, 580 (1950).

[117] J. Jones and E. Knill, *Efficient Refocusing of One-Spin and Two-Spin Interactions for NMR Quantum Computation*, Journal of Magnetic Resonance **141**, 322 (1999).

[118] D. W. Leung, I. L. Chuang, F. Yamaguchi, and Y. Yamamoto, *Efficient implementation of coupled logic gates for quantum computation*, Phys. Rev. A **61**, 042310 (2000).

[119] U. Haeberlen and J. S. Waugh, *Coherent Averaging Effects in Magnetic Resonance*, Phys. Rev. **175**, 453 (1968).

[120] L. M. Vandersypen, *Experimental Quantum Computation with Nuclear Spins in Liquid Solution*, Ph.D. thesis, Stanford University (2001).

[121] L. M. K. Vandersypen and I. L. Chuang, *NMR techniques for quantum control and computation*, Rev. Mod. Phys. **76**, 1037 (2005).

[122] K. Pachler and P. Wessels, *Selective Population Inversion (SPI). A pulsed double resonance method in FT NMR spectroscopy equivalent to INDOR*, Journal of Magnetic Resonance (1969) **12**, 337 (1973).

[123] S. Sørensen, R. Hansen, and H. Jakobsen, *Assignments and relative signs of 13C-X coupling constants in 13C FT NMR from selective population transfer (SPT)*, Journal of Magnetic Resonance (1969) **14**, 243 (1974).

[124] H. Jakobsen, S. Linde, and S. Sørensen, *Sensitivity enhancement in 13C FT NMR from selective population transfer (spt) in molecules with degenerate proton transitions*, Journal of Magnetic Resonance (1969) **15**, 385 (1974).

[125] K. Pachler and P. Wessels, *Sensitivity gain in a progressive-saturation selective population inversion NMR experiment*, Journal of Magnetic Resonance (1969) **28**, 53 (1977).

[126] A. Barenco, D. Deutsch, A. Ekert, and R. Jozsa, *Conditional Quantum Dynamics and Logic Gates*, Phys. Rev. Lett. **74**, 4083 (1995).

[127] G. A. Morris and R. Freeman, *Enhancement of nuclear magnetic resonance signals by polarization transfer*, Journal of the American Chemical Society **101**, 760 (1979).

[128] M. Steffen, *A prototype quantum computer using nuclear spins in liquid solution*, Ph.D. thesis, Stanford University (2003).

[129] C. A. Ryan, C. Negrevergne, M. Laforest, E. Knill, and R. Laflamme, *Liquid-state nuclear magnetic resonance as a testbed for developing quantum control methods*, Phys. Rev. A **78**, 012328 (2008).

[130] A. Abragam, *Principles of Nuclear Magnetism* (Clarendon Press, Oxford, UK, 1961).

[131] R. R. Ernst, G. Bodenhausen, and A. Wokaun, *Principles of Nuclear Magnetic Resonance in One and Two Dimensions* (Oxford University Press, 1987).

[132] M. H. Levitt, *Spin Dynamics: Basics of Nuclear Magnetic Resonance* (Wiley, 2008).

[133] R. Freeman, *Spin Choreography: Basic Steps in High Resolution NMR* (Spektrum, 1997).

[134] P. J. Hore, J. A. Jones, and S. Wimperis, *NMR: The Toolkit. How pulse sequences work* (Oxford University Press, 2015).

[135] C. P. Slichter, *Principles of Magnetic Resonance* (Springer, Berlin, 1996).

[136] M. Mehring, *Concepts of spin quantum computing*, Applied Magnetic Resonance **17**, 141 (1999).

[137] J. Jones, *NMR Quantum Computation: A Critical Evaluation*, Fortschritte der Physik **48**, 909 (2000).

[138] D. Cory, R. Laflamme, E. Knill, L. Viola, T. Havel, N. Boulant, G. Boutis, E. Fortunato, S. Lloyd, R. Martinez, C. Negrevergne, M. Pravia, Y. Sharf, G. Teklemariam, Y. Weinstein, and W. Zurek, *NMR Based Quantum Information Processing: Achievements and Prospects*, Fortschritte der Physik **48**, 875 (2000).

[139] J. A. Jones, *Quantum computing and nuclear magnetic resonance*, PhysChemComm **4**, 49 (2001).

[140] J. Jones, *NMR quantum computation*, Progress in Nuclear Magnetic Resonance Spectroscopy **38**, 325 (2001).

[141] T. F. Havel, D. G. Cory, S. Lloyd, N. Boulant, E. M. Fortunato, M. A. Pravia, G. Teklemariam, Y. S. Weinstein, A. Bhattacharyya, and J. Hou, *Quantum information processing by nuclear magnetic resonance spectroscopy*, American Journal of Physics **70**, 345 (2002).

[142] C. Ramanathan, N. Boulant, Z. Chen, D. G. Cory, I. Chuang, and M. Steffen, *NMR Quantum Information Processing*, Quantum Information Processing **3**, 15 (2004).

[143] I. Oliveira, R. Sarthour Jr, T. Bonagamba, E. Azevedo, and J. C. Freitas, *NMR quantum information processing* (Amsterdam, The Netherlands: Elsevier, 2007).

[144] D. Suter and T. S. Mahesh, *Spins as qubits: Quantum information processing by nuclear magnetic resonance*, The Journal of Chemical Physics **128**, 052206 (2008).

[145] J. A. Jones, *Quantum computing with NMR*, Progress in Nuclear Magnetic Resonance Spectroscopy **59**, 91 (2011).

[146] B. Criger, G. Passante, D. Park, and R. Laflamme, *Recent advances in nuclear magnetic resonance quantum information processing*, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences **370**, 4620 (2012).

[147] R. M. Serra and I. S. Oliveira, *Nuclear magnetic resonance quantum information processing*, Philosophical Transactions. Series A, Mathematical, Physical, and Engineering Sciences **370**, 4615 (2012).

[148] T. S. Mahesh, *Quantum information processing by NMR*, Resonance **20**, 1053 (2015).

[149] G. Bhole and J. A. Jones, *Practical pulse engineering: Gradient ascent without matrix exponentiation*, Frontiers of Physics **13**, 130312 (2018).

[150] G. Bhole, T. Tsunoda, P. J. Leek, and J. A. Jones, *Rescaling Interactions for Quantum Control*, Phys. Rev. Applied **13**, 034002 (2020).

[151] T. Tsunoda, G. Bhole, S. A. Jones, J. A. Jones, and P. J. Leek, *Efficient Hamiltonian programming in qubit arrays with nearest-neighbor couplings*, Phys. Rev. A **102**, 032405 (2020).

[152] G. Bhole, J. A. Jones, C. Marletto, and V. Vedral, *Witnesses of non-classicality for simulated hybrid quantum systems*, Journal of Physics Communications **4**, 025013 (2020).

[153] M. Violaris, G. Bhole, J. A. Jones, V. Vedral, and C. Marletto, *Transforming pure and mixed states using an NMR quantum homogenizer*, Phys. Rev. A **103**, 022414 (2021).

[154] M. Hughes, M. D. Frye, R. Sawant, G. Bhole, J. A. Jones, S. L. Cornish, M. R. Tarbutt, J. M. Hutson, D. Jaksch, and J. Mur-Petit, *Robust entangling gate for polar molecules using magnetic and microwave fields*, Phys. Rev. A **101**, 062308 (2020).

[155] N. Khaneja, T. Reiss, C. Kehlet, T. Schulte-Herbrüggen, and S. J. Glaser, *Optimal control of coupled spin dynamics: design of NMR pulse sequences by gradient ascent algorithms*, Journal of Magnetic Resonance **172**, 296 (2005).

[156] M. K. Henry, C. Ramanathan, J. S. Hodges, C. A. Ryan, M. J. Ditty, R. Laflamme, and D. G. Cory, *Fidelity Enhancement by Logical Qubit Encoding*, Phys. Rev. Lett. **99**, 220501 (2007).

[157] C. A. Ryan, O. Moussa, J. Baugh, and R. Laflamme, *Spin Based Heat Engine: Demonstration of Multiple Rounds of Algorithmic Cooling*, Phys. Rev. Lett. **100**, 140501 (2008).

[158] C. A. Ryan, M. Laforest, and R. Laflamme, *Randomized benchmarking of single- and multi-qubit control in liquid-state NMR quantum information processing*, New Journal of Physics **11**, 013034 (2009).

[159] G. Passante, O. Moussa, C. A. Ryan, and R. Laflamme, *Experimental Approximation of the Jones Polynomial with One Quantum Bit*, Phys. Rev. Lett. **103**, 250501 (2009).

[160] O. Moussa, C. A. Ryan, D. G. Cory, and R. Laflamme, *Testing Contextuality on Quantum Ensembles with One Clean Qubit*, Phys. Rev. Lett. **104**, 160501 (2010).

[161] A. M. Souza, J. Zhang, C. A. Ryan, and R. Laflamme, *Experimental magic state distillation for fault-tolerant quantum computing*, Nature Communications **2**, 169 (2011).

[162] D. Lu, N. Xu, R. Xu, H. Chen, J. Gong, X. Peng, and J. Du, *Simulation of Chemical Isomerization Reaction Dynamics on a NMR Quantum Simulator*, Phys. Rev. Lett. **107**, 020501 (2011).

[163] J. Zhang, T.-C. Wei, and R. Laflamme, *Experimental Quantum Simulation of Entanglement in Many-Body Systems*, Phys. Rev. Lett. **107**, 010501 (2011).

[164] O. Moussa, M. P. da Silva, C. A. Ryan, and R. Laflamme, *Practical Experimental Certification of Computational Quantum Gates Using a Twirling Procedure*, Phys. Rev. Lett. **109**, 070504 (2012).

[165] N. Xu, J. Zhu, D. Lu, X. Zhou, X. Peng, and J. Du, *Quantum Factorization of 143 on a Dipolar-Coupling Nuclear Magnetic Resonance System*, Phys. Rev. Lett. **108**, 130501 (2012).

[166] J. Zhang, R. Laflamme, and D. Suter, *Experimental Implementation of Encoded Logical Qubit Operations in a Perfect Quantum Error Correcting Code*, Phys. Rev. Lett. **109**, 100503 (2012).

[167] J. Zhang, M.-H. Yung, R. Laflamme, A. Aspuru-Guzik, and J. Baugh, *Digital quantum simulation of the statistical mechanics of a frustrated magnet*, Nature Communications **3**, 880 (2012).

[168] X. Peng, Z. Luo, W. Zheng, S. Kou, D. Suter, and J. Du, *Experimental Implementation of Adiabatic Passage between Different Topological Orders*, Phys. Rev. Lett. **113**, 080404 (2014).

[169] Z. Li, H. Zhou, C. Ju, H. Chen, W. Zheng, D. Lu, X. Rong, C. Duan, X. Peng, and J. Du, *Experimental Realization of a Compressed Quantum Simulation of a 32-Spin Ising Chain*, Phys. Rev. Lett. **112**, 220501 (2014).

[170] D. Lu, H. Li, D.-A. Trottier, J. Li, A. Brodutch, A. P. Krismanich, A. Ghavami, G. I. Dmitrienko, G. Long, J. Baugh, and R. Laflamme, *Experimental Estimation of Average Fidelity of a Clifford Gate on a 7-Qubit Quantum Processor*, Phys. Rev. Lett. **114**, 140505 (2015).

[171] Z. Li, X. Liu, N. Xu, and J. Du, *Experimental Realization of a Quantum Support Vector Machine*, Phys. Rev. Lett. **114**, 140504 (2015).

[172] D. Lu, T. Xin, N. Yu, Z. Ji, J. Chen, G. Long, J. Baugh, X. Peng, B. Zeng, and R. Laflamme, *Tomography is Necessary for Universal Entanglement Detection with Single-Copy Observables*, Phys. Rev. Lett. **116**, 230501 (2016).

[173] W. Ma, Z. Ma, H. Wang, Z. Chen, Y. Liu, F. Kong, Z. Li, X. Peng, M. Shi, F. Shi, S.-M. Fei, and J. Du, *Experimental Test of Heisenberg's Measurement Uncertainty Relation Based on Statistical Distances*, Phys. Rev. Lett. **116**, 160405 (2016).

[174] T. Xin, D. Lu, J. Klassen, N. Yu, Z. Ji, J. Chen, X. Ma, G. Long, B. Zeng, and R. Laflamme, *Quantum State Tomography via Reduced Density Matrices*, Phys. Rev. Lett. **118**, 020401 (2017).

[175] J. Li, R. Fan, H. Wang, B. Ye, B. Zeng, H. Zhai, X. Peng, and J. Du, *Measuring Out-of-Time-Order Correlators on a Nuclear Magnetic Resonance Quantum Simulator*, Phys. Rev. X **7**, 031011 (2017).

[176] K. Li, Y. Wan, L.-Y. Hung, T. Lan, G. Long, D. Lu, B. Zeng, and R. Laflamme, *Experimental Identification of Non-Abelian Topological Orders on a Quantum Simulator*, Phys. Rev. Lett. **118**, 080502 (2017).

[177] W. Zheng, Z. Ma, H. Wang, S.-M. Fei, and X. Peng, *Experimental Demonstration of Observability and Operability of Robustness of Coherence*, Phys. Rev. Lett. **120**, 230504 (2018).

[178] B.-X. Wang, M.-J. Tao, Q. Ai, T. Xin, N. Lambert, D. Ruan, Y.-C. Cheng, F. Nori, F.-G. Deng, and G.-L. Long, *Efficient quantum simulation of photosynthetic light harvesting*, npj Quantum Information **4**, 52 (2018).

[179] Z. Li, X. Liu, H. Wang, S. Ashhab, J. Cui, H. Chen, X. Peng, and J. Du, *Quantum Simulation of Resonant Transitions for Solving the Eigenproblem of an Effective Water Hamiltonian*, Phys. Rev. Lett. **122**, 090504 (2019).

[180] A. Spörl, T. Schulte-Herbrüggen, S. J. Glaser, V. Bergholm, M. J. Storcz, J. Ferber, and F. K. Wilhelm, *Optimal control of coupled Josephson qubits*, Phys. Rev. A **75**, 012302 (2007).

[181] P. Rebentrost and F. K. Wilhelm, *Optimal control of a leaking qubit*, Phys. Rev. B **79**, 060507 (2009).

[182] R. Fisher, F. Helmer, S. J. Glaser, F. Marquardt, and T. Schulte-Herbrüggen, *Optimal control of circuit quantum electrodynamics in one and two dimensions*, Phys. Rev. B **81**, 085328 (2010).

[183] D. J. Egger and F. K. Wilhelm, *Optimized controlled-Z gates for two superconducting qubits coupled through a resonator*, Superconductor Science and Technology **27**, 014001 (2013).

[184] V. Nebendahl, H. Häffner, and C. F. Roos, *Optimal control of entangling operations for trapped-ion quantum computing*, Phys. Rev. A **79**, 012312 (2009).

[185] P. Schindler, J. T. Barreiro, T. Monz, V. Nebendahl, D. Nigg, M. Chwalla, M. Hennrich, and R. Blatt, *Experimental Repetitive Quantum Error Correction*, Science **332**, 1059 (2011).

[186] T. Monz, D. Nigg, E. A. Martinez, M. F. Brandl, P. Schindler, R. Rines, S. X. Wang, I. L. Chuang, and R. Blatt, *Realization of a scalable Shor algorithm*, Science **351**, 1068 (2016).

[187] R. S. Said and J. Twamley, *Robust control of entanglement in a nitrogen-vacancy center coupled to a ^{13}C nuclear spin in diamond*, Phys. Rev. A **80**, 032303 (2009).

[188] G. Waldherr, Y. Wang, S. Zaiser, M. Jamali, T. Schulte-Herbrüggen, H. Abe, T. Ohshima, J. Isoya, J. F. Du, P. Neumann, and J. Wrachtrup, *Quantum error correction in a solid-state hybrid spin register*, Nature **506**, 204 (2014).

[189] F. Dolde, V. Bergholm, Y. Wang, I. Jakobi, B. Naydenov, S. Pezzagna, J. Meijer, F. Jelezko, P. Neumann, T. Schulte-Herbrüggen, J. Biamonte, and J. Wrachtrup, *High-fidelity spin entanglement using optimal control*, Nature Communications **5**, 3371 (2014).

[190] X. Rong, J. Geng, F. Shi, Y. Liu, K. Xu, W. Ma, F. Kong, Z. Jiang, Y. Wu, and J. Du, *Experimental fault-tolerant universal quantum gates with solid-state spins under ambient conditions*, Nature Communications **6**, 8748 (2015).

[191] P. Rembold, N. Oshnik, M. M. Müller, S. Montangero, T. Calarco, and E. Neu, *Introduction to quantum optimal control for quantum sensing with nitrogen-vacancy centers in diamond*, AVS Quantum Science **2**, 024701 (2020).

[192] O. Sørensen, G. Eich, M. Levitt, G. Bodenhausen, and R. Ernst, *Product operator formalism for the description of NMR pulse experiments*, Progress in Nuclear Magnetic Resonance Spectroscopy **16**, 163 (1984).

[193] I. I. Rabi, *On the Process of Space Quantization*, Phys. Rev. **49**, 324 (1936).

[194] I. I. Rabi, *Space Quantization in a Gyrating Magnetic Field*, Phys. Rev. **51**, 652 (1937).

[195] W. G. Proctor and F. C. Yu, *The Dependence of a Nuclear Magnetic Resonance Frequency upon Chemical Compound*, Phys. Rev. **77**, 717 (1950).

[196] H. S. Gutowsky, D. W. McCall, and C. P. Slichter, *Coupling among Nuclear Magnetic Dipoles in Molecules*, Phys. Rev. **84**, 589 (1951).

[197] E. L. Hahn and D. E. Maxwell, *Spin Echo Measurements of Nuclear Spin Coupling in Molecules*, Phys. Rev. **88**, 1070 (1952).

[198] N. F. Ramsey and E. M. Purcell, *Interactions between Nuclear Spins in Molecules*, Phys. Rev. **85**, 143 (1952).

[199] F. Bloch and A. Siegert, *Magnetic Resonance for Nonrotating Fields*, Phys. Rev. **57**, 522 (1940).

[200] R. Bellman, *On the Theory of Dynamic Programming*, Proceedings of the National Academy of Sciences of the United States of America **38**, 716 (1952).

[201] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mishchenko, *The mathematical theory of optimal processes* (Wiley, New York, NY, 1962).

[202] E. B. Lee and L. Markus, *Foundations of optimal control theory* (Wiley, New York, NY, 1967).

[203] D. E. Kirk, *Optimal Control Theory; An Introduction* (Englewood Cliffs, N.J., Prentice-Hall., 1970).

[204] A. E. Bryson and Y. Ho, *Applied Optimal Control* (Hemisphere, Washington, D.C., 1975).

[205] I. M. Gelfand and S. V. Fomin, *Calculus of variations*, Dover Books on Mathematics (Dover, Mineola, NY, 2000).

[206] D. Liberzon, *Calculus of variations and optimal control theory: a concise introduction* (Princeton University Press, Princeton, NJ, 2011).

[207] D. D'Alessandro, *Introduction to quantum control and dynamics* (Taylor & Francis Ltd, Hoboken, NJ, 2007).

[208] V. Jurdjevic, *Geometric Control Theory*, Cambridge Studies in Advanced Mathematics (Cambridge University Press, Cambridge, 1996).

[209] B. Bonnard, S. J. Glaser, and D. Sugny, *A Review of Geometric Optimal Control for Quantum Systems in Nuclear Magnetic Resonance*, Advances in Mathematical Physics **2012**, 857493 (2012).

[210] J. Zhang, *Geometric method in quantum control*, Chinese Science Bulletin **57**, 2223 (2012).

[211] N. Khaneja, R. Brockett, and S. J. Glaser, *Time optimal control in spin systems*, Phys. Rev. A **63**, 032308 (2001).

[212] N. Khaneja, S. J. Glaser, and R. Brockett, *Sub-Riemannian geometry and time optimal control of three spin systems: Quantum gates and coherence transfer*, Phys. Rev. A **65**, 032301 (2002).

[213] H. Yuan and N. Khaneja, *Time optimal control of coupled qubits under nonstationary interactions*, Phys. Rev. A **72**, 040301 (2005).

[214] D. Sugny, C. Kontz, and H. R. Jauslin, *Time-optimal control of a two-level dissipative quantum system*, Phys. Rev. A **76**, 023419 (2007).

[215] M. Lapert, Y. Zhang, M. Braun, S. J. Glaser, and D. Sugny, *Singular Extremals for the Time-Optimal Control of Dissipative Spin* $\frac{1}{2}$ *Particles*, Phys. Rev. Lett. **104**, 083001 (2010).

[216] E. Assémat, M. Lapert, Y. Zhang, M. Braun, S. J. Glaser, and D. Sugny, *Simultaneous time-optimal control of the inversion of two spin-*$\frac{1}{2}$ *particles*, Phys. Rev. A **82**, 013415 (2010).

[217] A. Garon, S. J. Glaser, and D. Sugny, *Time-optimal control of SU(2) quantum operations*, Phys. Rev. A **88**, 043422 (2013).

[218] G. C. Hegerfeldt, *Driving at the Quantum Speed Limit: Optimal Control of a Two-Level System*, Phys. Rev. Lett. **111**, 260501 (2013).

[219] J. Nocedal and S. J. Wright, *Numerical Optimization*, 2nd ed. (Springer, New York, USA, 2006).

[220] A. P. Peirce, M. A. Dahleh, and H. Rabitz, *Optimal control of quantum-mechanical systems: Existence, numerical approximation, and applications*, Phys. Rev. A **37**, 4950 (1988).

[221] R. Kosloff, S. Rice, P. Gaspard, S. Tersigni, and D. Tannor, *Wavepacket dancing: Achieving chemical selectivity by shaping light pulses*, Chemical Physics **139**, 201 (1989).

[222] J. Somlói, V. A. Kazakov, and D. J. Tannor, *Controlled dissociation of I2 via optical transitions between the X and B electronic states*, Chemical Physics **172**, 85 (1993).

[223] W. Zhu, J. Botina, and H. Rabitz, *Rapidly convergent iteration methods for quantum optimal control of population*, The Journal of Chemical Physics **108**, 1953 (1998).

[224] W. Zhu and H. Rabitz, *Noniterative algorithms for finding quantum optimal controls*, The Journal of Chemical Physics **110**, 7142 (1999).

[225] E. M. Fortunato, M. A. Pravia, N. Boulant, G. Teklemariam, T. F. Havel, and D. G. Cory, *Design of strongly modulating pulses to implement precise effective Hamiltonians for quantum information processing*, The Journal of Chemical Physics **116**, 7599 (2002).

[226] Y. Maday and G. Turinici, *New formulations of monotonically convergent quantum control algorithms*, The Journal of Chemical Physics **118**, 8191 (2003).

[227] Y. Ohtsuki, G. Turinici, and H. Rabitz, *Generalized monotonically convergent algorithms for solving quantum optimal control problems*, The Journal of Chemical Physics **120**, 5509 (2004).

[228] Y. Ohtsuki, Y. Teranishi, P. Saalfrank, G. Turinici, and H. Rabitz, *Monotonically convergent algorithms for solving quantum optimal control problems described by an integrodifferential equation of motion*, Phys. Rev. A **75**, 033407 (2007).

[229] I. I. Maximov, Z. Toner, and N. C. Nielsen, *Optimal control design of NMR and dynamic nuclear polarization experiments using monotonically convergent algorithms*, The Journal of Chemical Physics **128**, 184505 (2008).

[230] P. Doria, T. Calarco, and S. Montangero, *Optimal Control Technique for Many-Body Quantum Dynamics*, Phys. Rev. Lett. **106**, 190501 (2011).

[231] T. Caneva, T. Calarco, and S. Montangero, *Chopped random-basis quantum optimization*, Phys. Rev. A **84**, 022326 (2011).

[232] D. M. Reich, M. Ndong, and C. P. Koch, *Monotonically convergent optimization in quantum control using Krotov's method*, The Journal of Chemical Physics **136**, 104103 (2012).

[233] S. Hou, L. Wang, and X. Yi, *Realization of quantum gates by Lyapunov control*, Physics Letters A **378**, 699 (2014).

[234] G. Bhole, V. S. Anjusha, and T. S. Mahesh, *Steering quantum dynamics via bang-bang control: Implementing optimal fixed-point quantum search algorithm*, Phys. Rev. A **93**, 042339 (2016).

[235] D. Lucarelli, *Quantum optimal control via gradient ascent in function space and the time-bandwidth quantum speed limit*, Phys. Rev. A **97**, 062346 (2018).

[236] S. Machnes, E. Assémat, D. Tannor, and F. K. Wilhelm, *Tunable, Flexible, and Efficient Optimization of Control Pulses for Practical Qubits*, Phys. Rev. Lett. **120**, 150401 (2018).

[237] H. J. Sussmann and V. Jurdjevic, *Controllability of nonlinear systems*, Journal of Differential Equations **12**, 95 (1972).

[238] V. Jurdjevic and H. J. Sussmann, *Control systems on Lie groups*, Journal of Differential Equations **12**, 313 (1972).

[239] V. Ramakrishna, M. V. Salapaka, M. Dahleh, H. Rabitz, and A. Peirce, *Controllability of molecular systems*, Phys. Rev. A **51**, 960 (1995).

[240] F. Albertini and D. D'Alessandro, *Notions of controllability for bilinear multilevel quantum systems*, IEEE Transactions on Automatic Control **48**, 1399 (2003).

[241] S. G. Schirmer, H. Fu, and A. I. Solomon, *Complete controllability of quantum systems*, Phys. Rev. A **63**, 063410 (2001).

[242] V. Ramakrishna and H. Rabitz, *Relation between quantum computing and quantum controllability*, Phys. Rev. A **54**, 1715 (1996).

[243] J. L. Dodd, M. A. Nielsen, M. J. Bremner, and R. T. Thew, *Universal quantum computation and simulation using any entangling Hamiltonian and local unitaries*, Phys. Rev. A **65**, 040301 (2002).

[244] M. D. Bowdrey, D. K. Oi, A. Short, K. Banaszek, and J. Jones, *Fidelity of single qubit maps*, Physics Letters A **294**, 258 (2002).

[245] M. A. Nielsen, *A simple formula for the average gate fidelity of a quantum dynamical operation*, Physics Letters A **303**, 249 (2002).

[246] V. Bergholm, W. Wieczorek, T. Schulte-Herbrüggen, and M. Keyl, *Optimal control of hybrid optomechanical systems for generating non-classical states of mechanical motion*, Quantum Science and Technology **4**, 034001 (2019).

[247] P. W. Shor, *Fault-tolerant quantum computation*, in *Proceedings of 37th Conference on Foundations of Computer Science* (1996) pp. 56–65.

[248] J. Preskill, *Reliable quantum computers*, Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences **454**, 385 (1998).

[249] E. Knill, R. Laflamme, and W. H. Zurek, *Resilient Quantum Computation*, Science **279**, 342 (1998).

[250] E. Knill, *Quantum computing with realistically noisy devices*, Nature **434**, 39 (2005).

[251] D. Aharonov and M. Ben-Or, *Fault-Tolerant Quantum Computation with Constant Error Rate*, SIAM Journal on Computing **38**, 1207 (2008).

[252] P. W. Shor, *Scheme for reducing decoherence in quantum computer memory*, Phys. Rev. A **52**, R2493 (1995).

[253] A. M. Steane, *Error Correcting Codes in Quantum Theory*, Phys. Rev. Lett. **77**, 793 (1996).

[254] A. R. Calderbank and P. W. Shor, *Good quantum error-correcting codes exist*, Phys. Rev. A **54**, 1098 (1996).

[255] A. Steane, *Multiple-particle interference and quantum error correction*, Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences **452**, 2551 (1996).

[256] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, *Mixed-state entanglement and quantum error correction*, Phys. Rev. A **54**, 3824 (1996).

[257] R. Laflamme, C. Miquel, J. P. Paz, and W. H. Zurek, *Perfect Quantum Error Correcting Code*, Phys. Rev. Lett. **77**, 198 (1996).

[258] D. Gottesman, *Class of quantum error-correcting codes saturating the quantum Hamming bound*, Phys. Rev. A **54**, 1862 (1996).

[259] A. R. Calderbank, E. M. Rains, P. W. Shor, and N. J. A. Sloane, *Quantum Error Correction and Orthogonal Geometry*, Phys. Rev. Lett. **78**, 405 (1997).

[260] E. Knill and R. Laflamme, *Theory of quantum error-correcting codes*, Phys. Rev. A **55**, 900 (1997).

[261] C. Gidney and M. Ekerå, *How to factor 2048 bit RSA integers in 8 hours using 20 million noisy qubits*, arXiv:1905.09749 (2019).

[262] H. K. Cummins, G. Llewellyn, and J. A. Jones, *Tackling systematic errors in quantum logic gates with composite rotations*, Phys. Rev. A **67**, 042308 (2003).

[263] B. Rowland and J. A. Jones, *Implementing quantum logic gates with gradient ascent pulse engineering: principles and practicalities*, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences **370**, 4636 (2012).

[264] C. Ryan, *Characterization and Control in Large Hilbert spaces.*, Ph.D. thesis, University of Waterloo (2008).

[265] R. Fisher, *Optimal control of multi-level quantum systems*, Ph.D. thesis, Technische Universität München (2010).

[266] D. L. Goodwin, *Advanced optimal control methods for spin systems*, Ph.D. thesis, University of Southampton (2017).

[267] J. A. Nelder and R. Mead, *A Simplex Method for Function Minimization*, The Computer Journal **7**, 308 (1965).

[268] J. H. Holland, *Adaptation in Natural and Artificial Systems* (University of Michigan Press, Ann Arbor, MI, 1975) second Edition, 1992.

[269] D. E. Goldberg and J. H. Holland, *Genetic Algorithms and Machine Learning*, Machine Learning **3**, 95 (1988).

[270] J. H. Holland, *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control and Artificial Intelligence* (MIT Press, Cambridge, MA, USA, 1992).

[271] S. Machnes, U. Sander, S. J. Glaser, P. de Fouquières, A. Gruslys, S. Schirmer, and T. Schulte-Herbrüggen, *Comparing, optimizing, and benchmarking quantum-control algorithms in a unifying programming framework*, Phys. Rev. A **84**, 022305 (2011).

[272] H. Hogben, M. Krzystyniak, G. Charnock, P. Hore, and I. Kuprov, *Spinach A software library for simulation of spin dynamics in large spin systems*, Journal of Magnetic Resonance **208**, 179 (2011).

[273] J. Johansson, P. Nation, and F. Nori, *QuTiP: An open-source Python framework for the dynamics of open quantum systems*, Computer Physics Communications **183**, 1760 (2012).

[274] J. Johansson, P. Nation, and F. Nori, *QuTiP 2: A Python framework for the dynamics of open quantum systems*, Computer Physics Communications **184**, 1234 (2013).

[275] Z. Toner, T. Vosegaard, C. Kehlet, N. Khaneja, S. J. Glaser, and N. C. Nielsen, *Optimal control in NMR spectroscopy: Numerical implementation in SIMPSON*, Journal of Magnetic Resonance **197**, 120 (2009).

[276] S. G. Schirmer and P. de Fouquieres, *Efficient algorithms for optimal control of quantum dynamics: the Krotov method unencumbered*, New Journal of Physics **13**, 073029 (2011).

[277] M. R. Hestenes and E. Stiefel, *Methods of conjugate gradients for solving linear systems*, Journal of Research of the National Bureau of Standards **49**, 409 (1952).

[278] L. Lasdon, S. Mitter, and A. Waren, *The conjugate gradient method for optimal control problems*, IEEE Transactions on Automatic Control **12**, 132 (1967).

[279] J. R. Shewchuk, *An Introduction to the Conjugate Gradient Method Without the Agonizing Pain*, Technical Report, Carnegie-Mellon University, USA (1994).

[280] R. Fletcher and C. M. Reeves, *Function minimization by conjugate gradients*, The Computer Journal **7**, 149 (1964).

[281] Y. H. Dai and Y. Yuan, *A Nonlinear Conjugate Gradient Method with a Strong Global Convergence Property*, SIAM Journal on Optimization **10**, 177 (1999).

[282] C. G. Broyden, *The Convergence of a Class of Double-rank Minimization Algorithms 1. General Considerations*, IMA Journal of Applied Mathematics **6**, 76 (1970).

[283] R. Fletcher, *A new approach to variable metric algorithms*, The Computer Journal **13**, 317 (1970).

[284] D. Goldfarb, *A Family of Variable-Metric Methods Derived by Variational Means*, Mathematics of Computation **24**, 23 (1970).

[285] D. F. Shanno, *Conditioning of Quasi-Newton Methods for Function Minimization*, Mathematics of Computation **24**, 647 (1970).

[286] P. de Fouquières, S. Schirmer, S. Glaser, and I. Kuprov, *Second order gradient ascent pulse engineering*, Journal of Magnetic Resonance **212**, 412 (2011).

[287] D. C. Liu and J. Nocedal, *On the limited memory BFGS method for large scale optimization*, Mathematical Programming **45**, 503 (1989).

[288] R. H. Byrd, J. Nocedal, and R. B. Schnabel, *Representations of quasi-Newton matrices and their use in limited memory methods*, Mathematical Programming **63**, 129 (1994).

[289] R. M. Wilcox, *Exponential Operators and Parameter Differentiation in Quantum Physics*, Journal of Mathematical Physics **8**, 962 (1967).

[290] T. Levante, T. Bremi, and R. Ernst, *Pulse-Sequence Optimization with Analytical Derivatives. Application to Deuterium Decoupling in Oriented Phases*, Journal of Magnetic Resonance, Series A **121**, 167 (1996).

[291] K. Aizu, *Parameter Differentiation of QuantumMechanical Linear Operators*, Journal of Mathematical Physics **4**, 762 (1963).

[292] C. Moler and C. Van Loan, *Nineteen Dubious Ways to Compute the Exponential of a Matrix*, SIAM Review **20**, 801 (1978).

[293] C. Moler and C. Van Loan, *Nineteen Dubious Ways to Compute the Exponential of a Matrix, Twenty-Five Years Later*, SIAM Review **45**, 3 (2003).

[294] N. J. Higham, *The Scaling and Squaring Method for the Matrix Exponential Revisited*, SIAM Journal on Matrix Analysis and Applications **26**, 1179 (2005).

[295] A. H. Al-Mohy and N. J. Higham, *A New Scaling and Squaring Algorithm for the Matrix Exponential*, SIAM Journal on Matrix Analysis and Applications **31**, 970 (2010).

[296] T. Gradl, A. Spörl, T. Huckle, S. J. Glaser, and T. Schulte-Herbrüggen, *Parallelising Matrix Operations on Clusters for an Optimal Control-Based Quantum Compiler*, in *Euro-Par 2006 Parallel Processing*, edited by W. E. Nagel, W. V. Walter, and W. Lehner (Springer Berlin Heidelberg, Berlin, Heidelberg, 2006) pp. 751–762.

[297] T. Schulte-Herbrüggen, A. Spörl, K. Waldherr, T. Gradl, S. J. Glaser, and T. Huckle, *The HLRB Cluster as Quantum CISC Compiler*, in *High Performance Computing in Science and Engineering, Garching/Munich 2007*, edited by S. Wagner, M. Steinmetz, A. Bode, and M. Brehm (Springer Berlin Heidelberg, Berlin, Heidelberg, 2009) pp. 517–533.

[298] K. Waldherr, T. Huckle, T. Auckenthaler, U. Sander, and T. Schulte-Herbrüggen, *Fast 3D Block Parallelisation for the Matrix Multiplication Prefix Problem*, in *High Performance Computing in Science and Engineering, Garching/Munich 2009* (Springer Berlin Heidelberg, Berlin, Heidelberg, 2010) pp. 39–50.

[299] T. Auckenthaler, M. Bader, T. Huckle, A. Spörl, and K. Waldherr, *Matrix exponentials and parallel prefix computation in a quantum control problem*, Parallel Computing **36**, 359 (2010), parallel Matrix Algorithms and Applications.

[300] M. D. Bowdrey, J. A. Jones, E. Knill, and R. Laflamme, *Compiling gate networks on an Ising quantum computer*, Phys. Rev. A **72**, 032315 (2005).

[301] Y. S. Weinstein, T. F. Havel, J. Emerson, N. Boulant, M. Saraceno, S. Lloyd, and D. G. Cory, *Quantum process tomography of the quantum Fourier transform*, The Journal of Chemical Physics **121**, 6117 (2004).

[302] N. Margolus and L. B. Levitin, *The maximum speed of dynamical evolution*, Physica D: Nonlinear Phenomena **120**, 188195 (1998).

[303] S. Deffner and S. Campbell, *Quantum speed limits: from Heisenberg's uncertainty principle to optimal quantum control*, Journal of Physics A: Mathematical and Theoretical **50**, 453001 (2017).

[304] S. Ashhab, P. C. de Groot, and F. Nori, *Speed limits for quantum gates in multiqubit systems*, Phys. Rev. A **85**, 052327 (2012).

[305] T. E. Skinner, K. Kobzar, B. Luy, M. R. Bendall, W. Bermel, N. Khaneja, and S. J. Glaser, *Optimal control design of constant amplitude phase-modulated pulses: Application to calibration-free broadband excitation*, Journal of Magnetic Resonance **179**, 241 (2006).

[306] P. Coote, C. Anklin, W. Massefski, G. Wagner, and H. Arthanari, *Rapid convergence of optimal control in NMR using numerically-constructed toggling frames*, Journal of Magnetic Resonance **281**, 94 (2017).

[307] P. W. Coote, S. A. Robson, A. Dubey, A. Boeszoermenyi, M. Zhao, G. Wagner, and H. Arthanari, *Optimal control theory enables homonuclear decoupling without Bloch–Siegert shifts in NMR spectroscopy*, Nature Communications **9**, 3014 (2018).

[308] J. H. M. Jensen, F. S. Møller, J. J. Sørensen, and J. F. Sherson, *Approximate Dynamics Lead to More Optimal Control: Efficient Exact Derivatives*, arXiv:2005.09943 (2020).

[309] J. P. S. Peterson, H. Katiyar, and R. Laflamme, *Fast Simulation of Magnetic Field Gradients for Optimization of Pulse Sequences*, arXiv:2006.10133 (2020).

[310] G. Bhole and T. S. Mahesh, *Rapid Exponentiation using Discrete Operators: Applications in Optimizing Quantum Controls and Simulating Quantum Dynamics*, arXiv:1707.02162 (2017).

[311] V. Strassen, *Gaussian elimination is not optimal*, Numerische Mathematik **13**, 354 (1969).

[312] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms, Third Edition* (The MIT Press, 2009).

[313] D. Coppersmith and S. Winograd, *Matrix multiplication via arithmetic progressions*, Journal of Symbolic Computation **9**, 251 (1990), computational algebraic complexity editorial.

[314] J. Alman and V. V. Williams, *A Refined Laser Method and Faster Matrix Multiplication*, arXiv:2010.05846 (2020).

[315] T. Xin, S. Huang, S. Lu, K. Li, Z. Luo, Z. Yin, J. Li, D. Lu, G. Long, and B. Zeng, *NMRCloudQ: a quantum cloud experience on a nuclear magnetic resonance quantum computer*, Science Bulletin **63**, 17 (2018).

[316] L. Viola and S. Lloyd, *Dynamical suppression of decoherence in two-state quantum systems*, Phys. Rev. A **58**, 2733 (1998).

[317] J. J. L. Morton, A. M. Tyryshkin, A. Ardavan, S. C. Benjamin, K. Porfyrakis, S. A. Lyon, and G. A. D. Briggs, *Bang–bang control of fullerene qubits using ultrafast phase gates*, Nature Physics **2**, 40 (2006).

[318] K. Mølmer and A. Sørensen, *Multiparticle Entanglement of Hot Trapped Ions*, Phys. Rev. Lett. **82**, 1835 (1999).

[319] G. S. Paraoanu, *Microwave-induced coupling of superconducting qubits*, Phys. Rev. B **74**, 140504 (2006).

[320] A. Parra-Rodriguez, P. Lougovski, L. Lamata, E. Solano, and M. Sanz, *Digital-analog quantum computation*, Phys. Rev. A **101**, 022305 (2020).

[321] Y. Wu, Y. Wang, X. Qin, X. Rong, and J. Du, *A programmable two-qubit solid-state quantum processor under ambient conditions*, npj Quantum Information **5**, 9 (2019).

[322] A. Bermudez, F. Jelezko, M. B. Plenio, and A. Retzker, *Electron-Mediated Nuclear-Spin Interactions between Distant Nitrogen-Vacancy Centers*, Phys. Rev. Lett. **107**, 150503 (2011).

[323] H. Haas, D. Puzzuoli, F. Zhang, and D. G. Cory, *Engineering effective Hamiltonians*, New Journal of Physics **21**, 103011 (2019).

[324] D. Hayes, S. T. Flammia, and M. J. Biercuk, *Programmable quantum simulation by dynamic Hamiltonian engineering*, New Journal of Physics **16**, 083027 (2014).

[325] J. Welch, D. Greenbaum, S. Mostame, and A. Aspuru-Guzik, *Efficient quantum circuits for diagonal unitaries without ancillas*, New Journal of Physics **16**, 033040 (2014).

[326] P. Krantz, M. Kjaergaard, F. Yan, T. P. Orlando, S. Gustavsson, and W. D. Oliver, *A quantum engineer's guide to superconducting qubits*, Applied Physics Reviews **6**, 021318 (2019).

[327] J. M. Gambetta, J. M. Chow, and M. Steffen, *Building logical qubits in a superconducting quantum computing system*, npj Quantum Information **3**, 2 (2017).

[328] 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).

[329] D. Leung, *Simulation and reversal of n -qubit Hamiltonians using Hadamard matrices*, Journal of Modern Optics **49**, 1199 (2002).

[330] K. G. Beauchamp, *Applications of Walsh and related functions* (Academic Press, 1984).

[331] M. S. Lynn, *On the Schur product of H-matrices and non-negative matrices, and related inequalities*, Mathematical Proceedings of the Cambridge Philosophical Society **60**, 425431 (1964).

[332] R. G. Bland, *The Allocation of Resources by Linear Programming*, Scientific American **244**, 126 (1981).

[333] G. B. Dantzig, *Reminiscences about the origins of linear programming*, Operations Research Letters **1**, 43 (1982).

[334] I. Adler, M. G. C. Resende, G. Veiga, and N. Karmarkar, *An implementation of Karmarkar's algorithm for linear programming*, Mathematical Programming **44**, 297 (1989).

[335] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C (2nd Ed.): The Art of Scientific Computing* (Cambridge University Press, 1992).

[336] J. Du, J. Zhu, M. Shi, X. Peng, and D. Suter, *Experimental observation of a topological phase in the maximally entangled state of a pair of qubits*, Phys. Rev. A **76**, 042121 (2007).

[337] R. Freeman, T. A. Frenkiel, and M. H. Levitt, *Composite Z pulses*, Journal of Magnetic Resonance (1969) **44**, 409 (1981).

[338] J. A. Jones, *Designing short robust not gates for quantum computation*, Phys. Rev. A **87**, 052317 (2013).

[339] N. Boulant, E. M. Fortunato, M. A. Pravia, G. Teklemariam, D. G. Cory, and T. F. Havel, *Entanglement transfer experiment in NMR quantum information processing*, Phys. Rev. A **65**, 024302 (2002).

[340] D. Collins, K. W. Kim, W. C. Holton, H. Sierzputowska-Gracz, and E. O. Stejskal, *NMR quantum computation with indirectly coupled gates*, Phys. Rev. A **62**, 022304 (2000).

[341] N. Khaneja, B. Heitmann, A. Spörl, H. Yuan, T. Schulte-Herbrüggen, and S. J. Glaser, *Shortest paths for efficient control of indirectly coupled qubits*, Phys. Rev. A **75**, 012322 (2007).

[342] J. A. J. Hall, *Towards a practical parallelisation of the simplex method*, Computational Management Science **7**, 139 (2010).

[343] J. A. Jones, *Robust Ising gates for practical quantum computation*, Phys. Rev. A **67**, 012317 (2003).

[344] I. L. Chuang and M. A. Nielsen, *Prescription for experimental determination of the dynamics of a quantum black box*, Journal of Modern Optics **44**, 2455 (1997).

[345] H. F. Hofmann, *Complementary Classical Fidelities as an Efficient Criterion for the Evaluation of Experimentally Realized Quantum Operations*, Phys. Rev. Lett. **94**, 160504 (2005).

[346] L. H. Pedersen, N. M. Møller, and K. Mølmer, *Fidelity of quantum operations*, Physics Letters A **367**, 47 (2007).

[347] K. Mayer and E. Knill, *Quantum process fidelity bounds from sets of input states*, Phys. Rev. A **98**, 052326 (2018).

[348] J. K. Kruschke, *Doing Bayesian Data Analysis* (Academic Press, Oxford, 2011).

[349] F. Arute *et al.*, *Quantum supremacy using a programmable superconducting processor*, Nature **574**, 505 (2019).

[350] P. Spring, T. Tsunoda, B. Vlastakis, and P. Leek, *Modeling Enclosures for Large-Scale Superconducting Quantum Circuits*, Phys. Rev. Applied **14**, 024061 (2020).

[351] P. Sengupta and L. P. Pryadko, *Scalable Design of Tailored Soft Pulses for Coherent Control*, Phys. Rev. Lett. **95**, 037202 (2005).

[352] A. De and L. P. Pryadko, *Universal Set of Scalable Dynamically Corrected Gates for Quantum Error Correction with Always-on Qubit Couplings*, Phys. Rev. Lett. **110**, 070503 (2013).

[353] A. De and L. P. Pryadko, *Dynamically corrected gates for qubits with always-on Ising couplings: Error model and fault tolerance with the toric code*, Phys. Rev. A **89**, 032332 (2014).

[354] K. Appel and W. Haken, *The Solution of the Four-Color-Map Problem*, Scientific American **237**, 108 (1977).

[355] A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M. Chow, and J. M. Gambetta, *Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets*, Nature **549**, 242 (2017).

[356] E. Farhi, J. Goldstone, and S. Gutmann, *A quantum approximate optimization algorithm*, arXiv:1411.4028 (2014).

[357] K. Temme, S. Bravyi, and J. M. Gambetta, *Error Mitigation for Short-Depth Quantum Circuits*, Phys. Rev. Lett. **119**, 180509 (2017).

[358] Y. Li and S. C. Benjamin, *Efficient Variational Quantum Simulator Incorporating Active Error Minimization*, Phys. Rev. X **7**, 021050 (2017).

[359] A. Kandala, K. Temme, A. D. Córcoles, A. Mezzacapo, J. M. Chow, and J. M. Gambetta, *Error mitigation extends the computational reach of a noisy quantum processor*, Nature **567**, 491 (2019).

[360] A. Shaka, J. Keeler, T. Frenkiel, and R. Freeman, *An improved sequence for broadband decoupling: WALTZ-16*, Journal of Magnetic Resonance (1969) **52**, 335 (1983).

[361] M. Anwar, D. Blazina, H. Carteret, S. Duckett, and J. Jones, *Implementing Grovers quantum search on a para-hydrogen based pure state NMR quantum computer*, Chemical Physics Letters **400**, 94 (2004).

[362] M. S. Anwar, J. A. Jones, D. Blazina, S. B. Duckett, and H. A. Carteret, *Implementation of NMR quantum computation with parahydrogen-derived high-purity quantum states*, Phys. Rev. A **70**, 032324 (2004).

[363] M. Anwar, D. Blazina, H. Carteret, S. Duckett, and J. Jones, *Implementing Grovers quantum search on a para-hydrogen based pure state NMR quantum computer*, Chemical Physics Letters **400**, 94 (2004).

[364] J. Baugh, O. Moussa, C. A. Ryan, A. Nayak, and R. Laflamme, *Experimental implementation of heat-bath algorithmic cooling using solid-state nuclear magnetic resonance*, Nature **438**, 470 (2005).

[365] J. M. Fernandez, S. Lloyd, T. Mor, and V. Roychowdhury, *Algorithmic Cooling of Spins: A practicable method for increasing polarization*, International Journal of Quantum Information **02**, 461 (2004).

[366] L. J. Schulman, T. Mor, and Y. Weinstein, *Physical Limits of Heat-Bath Algorithmic Cooling*, Phys. Rev. Lett. **94**, 120501 (2005).

[367] E. Knill, I. Chuang, and R. Laflamme, *Effective pure states for bulk quantum computation*, Phys. Rev. A **57**, 3348 (1998).

[368] S. L. Braunstein, C. M. Caves, R. Jozsa, N. Linden, S. Popescu, and R. Schack, *Separability of Very Noisy Mixed States and Implications for NMR Quantum Computing*, Phys. Rev. Lett. **83**, 1054 (1999).

[369] R. Schack and C. M. Caves, *Classical model for bulk-ensemble NMR quantum computation*, Phys. Rev. A **60**, 4354 (1999).

[370] Y. Mori, R. Sawae, M. Kawamura, T. Sakata, and K. Takarabe, *Quantum circuits for an effective pure state in NMR quantum computer*, International Journal of Quantum Chemistry **105**, 758 (2005).

[371] M. Kawamura, B. Rowland, and J. A. Jones, *Preparing pseudopure states with controlled-transfer gates*, Phys. Rev. A **82**, 032315 (2010).

[372] X. Kong, T. Xin, S. Wei, B. Wang, Y. Wang, K. Li, and G. Long, *Implementation of multiparty quantum clock synchronization*, arXiv:1708.06050 (2017).

[373] A. W. Overhauser, *Polarization of Nuclei in Metals*, Phys. Rev. **92**, 411 (1953).

[374] C. Marletto and V. Vedral, *Gravitationally Induced Entanglement between Two Massive Particles is Sufficient Evidence of Quantum Effects in Gravity*, Phys. Rev. Lett. **119**, 240402 (2017).

[375] D. C. Murphy and K. R. Brown, *Controlling error orientation to improve quantum algorithm success rates*, Phys. Rev. A **99**, 032318 (2019).

[376] M. Goldman, *Quantum description of high-resolution NMR in liquids* (Clarendon Press Oxford, 1988).

[377] J. G. Filgueiras, T. O. Maciel, R. E. Auccaise, R. O. Vianna, R. S. Sarthour, and I. S. Oliveira, *Experimental implementation of a NMR entanglement witness*, Quantum Information Processing **11**, 1883 (2012).

[378] L. Xiao and J. A. Jones, *NMR analogues of the quantum Zeno effect*, Physics Letters A **359**, 424 (2006).

[379] M. Ziman, P. Stelmachovic, V. Buzek, M. Hillery, V. Scarani, and N. Gisin, *Quantum homogenization*, arXiv:quant-ph/0110164 (2001).

[380] V. Scarani, M. Ziman, P. Štelmachovič, N. Gisin, and V. Bužek, *Thermalizing Quantum Machines: Dissipation and Entanglement*, Phys. Rev. Lett. **88**, 097905 (2002).

[381] H. K. Cummins, C. Jones, A. Furze, N. F. Soffe, M. Mosca, J. M. Peach, and J. A. Jones, *Approximate Quantum Cloning with Nuclear Magnetic Resonance*, Phys. Rev. Lett. **88**, 187901 (2002).

[382] J. C. Hoch and A. S. Stern, *NMR Data Processing* (Wiley, New York, 1996).

[383] D. DeMille, *Quantum Computation with Trapped Polar Molecules*, Phys. Rev. Lett. **88**, 067901 (2002).

[384] S. F. Yelin, K. Kirby, and R. Côté, *Schemes for robust quantum computation with polar molecules*, Phys. Rev. A **74**, 050301 (2006).

[385] G. Quéméner and P. S. Julienne, *Ultracold Molecules under Control!*, Chemical Reviews **112**, 4949 (2012).

[386] S. Moses, J. Covey, M. Miecnikowski, D. Jin, and J. Ye, *New frontiers for quantum gases of polar molecules*, Nature Physics **13**, 13 (2017).

[387] G. Wendin, *Quantum information processing with superconducting circuits: a review*, Reports on Progress in Physics **80**, 106001 (2017).

[388] M. Kjaergaard, M. E. Schwartz, J. Braumüller, P. Krantz, J. I.-J. Wang, S. Gustavsson, and W. D. Oliver, *Superconducting Qubits: Current State of Play*, Annual Review of Condensed Matter Physics **11**, 369 (2020).

[389] J. Mur-Petit, J. Pérez-Ríos, J. Campos-Martínez, M. I. Hernández, S. Willitsch, and J. J. García-Ripoll, *Toward a Molecular Ion Qubit*, in *Architecture and Design of Molecule Logic Gates and Atom Circuits*, edited by N. Lorente and C. Joachim (Springer Berlin Heidelberg, Berlin, Heidelberg, 2013) pp. 267–277.

[390] R. Sawant, J. A. Blackmore, P. D. Gregory, J. Mur-Petit, D. Jaksch, J. Aldegunde, J. M. Hutson, M. R. Tarbutt, and S. L. Cornish, *Ultracold polar molecules as qudits*, New J. Phys. **22**, 013027 (2020).

[391] A. Micheli, G. K. Brennen, and P. Zoller, *A toolbox for lattice-spin models with polar molecules*, Nature Physics **2**, 341 (2006).

[392] A. V. Gorshkov, S. R. Manmana, G. Chen, J. Ye, E. Demler, M. D. Lukin, and A. M. Rey, *Tunable superfluidity and quantum magnetism with ultracold polar molecules*, Phys. Rev. Lett. **107**, 115301 (2011).

[393] A. V. Gorshkov, K. R. Hazzard, and A. M. Rey, *Kitaev honeycomb and other exotic spin models with polar molecules*, Molecular Physics **111**, 1908 (2013).

[394] J. A. Blackmore, L. Caldwell, P. D. Gregory, E. M. Bridge, R. Sawant, J. Aldegunde, J. Mur-Petit, D. Jaksch, J. M. Hutson, B. E. Sauer, M. R. Tarbutt, and S. L. Cornish, *Ultracold molecules for quantum simulation: rotational coherences in CaF and RbCs*, Quantum Science and Technology **4**, 014010 (2019).

[395] P. Rosson, M. Kiffner, J. Mur-Petit, and D. Jaksch, *Characterizing the phase diagram of finite-size dipolar Bose-Hubbard systems*, Phys. Rev. A **101**, 013616 (2020).

[396] S. V. Alyabyshev, M. Lemeshko, and R. V. Krems, *Sensitive imaging of electromagnetic fields with paramagnetic polar molecules*, Phys. Rev. A **86**, 013409 (2012).

[397] J. Mur-Petit and J. J. García-Ripoll, *Measuring molecular electric dipoles using trapped atomic ions and ultrafast laser pulses*, Phys. Rev. A **91**, 012504 (2015).

[398] S. Truppe, H. J. Williams, M. Hambach, L. Caldwell, N. J. Fitch, E. A. Hinds, B. E. Sauer, and M. R. Tarbutt, *Molecules cooled below the Doppler limit*, Nature Physics **13**, 1173 (2017).

[399] L. Anderegg, B. L. Augenbraun, Y. Bao, S. Burchesky, L. W. Cheuk, W. Ketterle, and J. M. Doyle, *Laser cooling of optically trapped molecules*, Nature Physics **14**, 890 (2018).

[400] H. J. Williams, L. Caldwell, N. J. Fitch, S. Truppe, J. Rodewald, E. A. Hinds, B. E. Sauer, and M. R. Tarbutt, *Magnetic Trapping and Coherent Control of Laser-Cooled Molecules*, Phys. Rev. Lett. **120**, 163201 (2018).

[401] L. Anderegg, L. W. Cheuk, Y. Bao, S. Burchesky, W. Ketterle, K.-K. Ni, and J. M. Doyle, *An optical tweezer array of ultracold molecules*, Science **365**, 1156 (2019).

[402] S. Ospelkaus, K.-K. Ni, G. Quéméner, B. Neyenhuis, D. Wang, M. H. G. de Miranda, J. L. Bohn, J. Ye, and D. S. Jin, *Controlling the Hyperfine State of Rovibronic Ground-State Polar Molecules*, Phys. Rev. Lett. **104**, 030402 (2010).

[403] B. Yan, S. A. Moses, B. Gadway, J. P. Covey, K. R. A. Hazzard, A. M. Rey, D. S. Jin, and J. Ye, *Observation of dipolar spin-exchange interactions with lattice-confined polar molecules*, Nature **501**, 521 (2013).

[404] L. R. Liu, J. D. Hood, Y. Yu, J. T. Zhang, N. R. Hutzler, T. Rosenband, and K. K. Ni, *Building one molecule from a reservoir of two atoms*, Science **360**, 900 (2018).

[405] L. R. Liu, J. D. Hood, Y. Yu, J. T. Zhang, K. Wang, Y.-W. Lin, T. Rosenband, and K.-K. Ni, *Molecular assembly of ground state cooled single atoms*, Phys. Rev. X **9**, 021039 (2019).

[406] F. Motzoi, J. M. Gambetta, P. Rebentrost, and F. K. Wilhelm, *Simple Pulses for Elimination of Leakage in Weakly Nonlinear Qubits*, Phys. Rev. Lett. **103**, 110501 (2009).

[407] J. Koch, T. M. Yu, J. Gambetta, A. A. Houck, D. I. Schuster, J. Majer, A. Blais, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, *Charge-insensitive qubit design derived from the Cooper pair box*, Phys. Rev. A **76**, 042319 (2007).

[408] J. A. Schreier, A. A. Houck, J. Koch, D. I. Schuster, B. R. Johnson, J. M. Chow, J. M. Gambetta, J. Majer, L. Frunzio, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, *Suppressing charge noise decoherence in superconducting charge qubits*, Phys. Rev. B **77**, 180502 (2008).

[409] H. Paik, D. I. Schuster, L. S. Bishop, G. Kirchmair, G. Catelani, A. P. Sears, B. R. Johnson, M. J. Reagor, L. Frunzio, L. I. Glazman, S. M. Girvin, M. H. Devoret, and R. J. Schoelkopf, *Observation of High Coherence in Josephson Junction Qubits Measured in a Three-Dimensional Circuit QED Architecture*, Phys. Rev. Lett. **107**, 240501 (2011).

[410] C. Rigetti, J. M. Gambetta, S. Poletto, B. L. T. Plourde, J. M. Chow, A. D. Córcoles, J. A. Smolin, S. T. Merkel, J. R. Rozen, G. A. Keefe, M. B. Rothwell, M. B. Ketchen, and M. Steffen, *Superconducting qubit in a waveguide cavity with a coherence time approaching 0.1 ms*, Phys. Rev. B **86**, 100506 (2012).

[411] R. Barends, J. Kelly, A. Megrant, D. Sank, E. Jeffrey, Y. Chen, Y. Yin, B. Chiaro, J. Mutus, C. Neill, P. O'Malley, P. Roushan, J. Wenner, T. C. White, A. N. Cleland, and J. M. Martinis, *Coherent Josephson Qubit Suitable for Scalable Quantum Integrated Circuits*, Phys. Rev. Lett. **111**, 080502 (2013).

[412] N. Schuch and J. Siewert, *Natural two-qubit gate for quantum computation using the* XY *interaction*, Phys. Rev. A **67**, 032301 (2003).

[413] J. Majer, J. M. Chow, J. M. Gambetta, J. Koch, B. R. Johnson, J. A. Schreier, L. Frunzio, D. I. Schuster, A. A. Houck, A. Wallraff, A. Blais, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, *Coupling superconducting qubits via a cavity bus*, Nature **449**, 443 (2007).

[414] A. O. Niskanen, K. Harrabi, F. Yoshihara, Y. Nakamura, S. Lloyd, and J. S. Tsai, *Quantum Coherent Tunable Coupling of Superconducting Qubits*, Science **316**, 723 (2007).

[415] F. W. Strauch, P. R. Johnson, A. J. Dragt, C. J. Lobb, J. R. Anderson, and F. C. Wellstood, *Quantum Logic Gates for Coupled Superconducting Phase Qubits*, Phys. Rev. Lett. **91**, 167005 (2003).